

## صاحب امتیاز: انجمن فناوری اطلاعات و ارتباطات ایران

مدیر مسؤول: دکتر مسعود شفیعی

سر دبیر: دکتر محمدشهرام معین

## هیئت تحریریه:

فرخ حجت کاشانی (استاد)، دانشگاه علم و صنعت ایران

سید احمد رضا شرافت (استاد)، دانشگاه تربیت مدرس

مسعود شفیعی (استاد)، دانشگاه صنعتی امیرکبیر

محمد رضا عارف (استاد)، دانشگاه صنعتی شریف

کریم فائز (استاد)، دانشگاه صنعتی امیرکبیر

رضا فرجی دانا (استاد)، دانشگاه تهران

کمال محامدپور (استاد)، دانشگاه خواجه نصیرالدین طوسی

فرخ مروستی (استاد)، دانشگاه صنعتی شریف

سیداحمد معتمدی (استاد)، دانشگاه صنعتی امیرکبیر

روزبه معینی (استاد)، دانشگاه صنعتی امیرکبیر

کاظم اکبری (دانشیار)، دانشگاه صنعتی امیرکبیر

کامبیز بدیع (دانشیار)، مرکز تحقیقات مخابرات ایران

علی اکبر جلالی (دانشیار)، دانشگاه علم و صنعت ایران

حمیدرضا ربیعی (دانشیار)، دانشگاه صنعتی شریف

امیر البدوی (دانشیار)، دانشگاه تربیت مدرس

محمد تشنه لب (دانشیار)، دانشگاه خواجه نصیر طوسی

محمدشهرام معین (دانشیار)، پژوهشگاه ارتباطات و فناوری اطلاعات

## اعضای هیئت مدیره انجمن:

### • اعضای اصلی:

دکتر مسعود شفیعی (رئیس انجمن)

دانشگاه صنعتی امیرکبیر

دکتر عباس محمدی

دانشگاه صنعتی امیرکبیر

مهندس ناصر یوسف پور

شرکت مخابرات همراه اول

دکتر سید علیرضا هاشمی گلپایگانی

دانشگاه صنعتی امیرکبیر

مهندس فرهاد رضایی

شرکت کوشا تجارت نوپدید

دکتر علیرضا اهری لاحق

شرکت داده پردازی ایران

دکتر محمد جوادیان

دانشگاه صنعتی کرمانشاه

مهندس سیدعلی صموتی

دانشگاه فنی و حرفه‌ای شهید شمسی پور

دکتر علیرضا یاری

پژوهشکده فناوری اطلاعات و ارتباطات

دکتر اکبر کاری دولت آبادی

دانشگاه شهید ستاری

دکتر جمشید شنبه زاده

دانشگاه خوارزمی

### • اعضای علی‌البدل:

دکتر مهدی شجری (دانشگاه صنعتی امیرکبیر)

مهندس امیرارسلان داداشی جوردهی (شرکت ارتباطات سیار ایران)

## بازرسان:

### • بازرسان اصلی:

دکتر محمود کمره‌ای (دانشگاه تهران)

### • بازرسان علی‌البدل:

دکتر امیرشهاب شاهمیری (دانشگاه جامع علمی کاربردی)

براساس نامه شماره ۳/۴۸۱۷ مورخ ۱۳۸۶/۶/۱۱ دفتر امور پژوهشی وزارت علوم، تحقیقات و فناوری و ببنایه رأی کمیسیون بررسی نشریات علمی کشور در تاریخ ۱۳۸۶/۴/۲۳، درجه علمی - پژوهشی به این دوفصلنامه اعطا شده است.



# فهرست

دوفصلنامه علمی - پژوهشی فناوری اطلاعات و ارتباطات ایران

سال دوازدهم، شماره‌های ۴۵ و ۴۶، پاییز و زمستان ۱۳۹۹

- 
- ۱ نقش آفرینی عوامل سازمانی بر رفتار خرید مصرف‌کننده در شبکه‌های اجتماعی  
زهره دهدشتی‌شاهرخ، میترا دانش‌پرور، وحید خاشعی، وحید ناصحی‌فر
- ۲۱ ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP  
امیر امینی، علیرضا علی‌نژاد
- ۴۳ بهینه‌سازی توأم اختصاص توان در یک شبکه مخابرات همکارانه با استفاده از دسترسی چندگانه غیرمتعامد  
حمید امیری آرا، محمداقبر مسرور، محمدرضا ذهابی
- ۵۳ یک مدل کنترل دسترسی برای شبکه‌های اجتماعی آنلاین از طریق ارتباطات کاربر به کاربر  
محمدجواد پیران، محمود دی‌پیر
- ۷۱ ارائه مدلی برای بازیابی اطلاعات متنی با استفاده از اعداد فاصله‌ای  
هومان تحیری، فرزاد قهرمانی
- ۹۵ بهبود کارایی و دقت یافتن یال‌های پرتکرار در خلاصه‌سازی gMatrix از جریان‌گراف  
مسعود کاظمی، سیدحسین خواسته، حمیدرضا رخصتی
- ۱۱۵ بیشینه‌سازی طول عمر شبکه‌های حسگر صوتی زیرآبی با جایگذاری بهینه گره‌های رله  
زهره محمدی، محدثه سلیمانپورمقدم، داریوش عباسی‌مقدم، سیامک طالبی
- ۱۳۰ شناسایی عوامل و شاخص‌های مدل سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر  
محمدتقی تقوی‌فرد، محمدعلی عباسپورطهرانی‌فرد، کامران فیضی، رسول نورالسنا
- ۱۴۵ اندازه‌گیری میزان تشابه مسیرهای جهت‌دار بر روی داده‌های هندسی  
زینب سعیدی، محمد فرشی
- ۱۵۹ بررسی رابطه بین مدیریت منابع انسانی الکترونیکی، چابکی و اثربخشی سازمانی در سازمان‌های زیرمجموعه وزارت  
ارتباطات و فناوری اطلاعات  
عبدالرضا بهادری‌فرد، رضا رسولی
- ۱۸۵ ارائه روشی جدید مبتنی بر مدل کوکومو بمنظور افزایش دقت تخمین تلاش در پروژه‌های نرم‌افزاری  
مهدیه سالاری، وحید خطیبی‌بردسیری، عمید خطیبی‌بردسیری
- ۲۰۱ بهبود الگوریتم رقابت استعماری برای حل مسئله جایگذاری نودها در شبکه‌های حسگر بیسیم گرید سه‌بعدی  
همت شیخی، سیددوفا بارخدا، سودابه محمدی
- ۲۱۳ چکیده انگلیسی مقالات
-



# نقش آفرینی عوامل سازمانی بر رفتار خرید مصرف‌کننده در شبکه‌های اجتماعی

زهره دهدشتی شاهرخ\* میترا دانش پرور\*\* وحید ناصحی فر\*\*\* وحید خاشعی\*\*\*\*

\*\*\*،\*\*\*\* عضو هیات علمی دانشگاه علامه طباطبائی

\*\* عضو هیات علمی دانشگاه بصیر

تاریخ دریافت ۱۳۹۹/۰۱/۱۰ تاریخ پذیرش ۱۳۹۹/۰۹/۰۶

نوع مقاله: پژوهشی

## چکیده

صنعت پوشاک در سالهای اخیر با رشد زیادی در شبکه‌های اجتماعی روبه‌رو بوده است. عوامل زیادی بر خرید پوشاک از شبکه‌های اجتماعی مؤثر هستند که یکی از این عوامل مهم و تأثیرگذار، عوامل سازمانی است. از این رو، هدف این پژوهش، بررسی عوامل سازمانی مؤثر بر مصرف‌کننده در خرید از طریق شبکه‌های اجتماعی است. در انجام این تحقیق از روش ترکیبی استفاده شده است. در بخش کیفی با بررسی ادبیات و مصاحبه با فروشندگان فعال در شبکه‌های اجتماعی متونی تهیه و کدگذاری گردید، مؤلفه‌ها در چهار مقوله نهایی دسته‌بندی و در قالب مدل اولیه ارائه شدند. در بخش کمی، بر اساس مدل اولیه، پرسشنامه‌ای تدوین و در اختیار ۳۸۵ نفر از خریداران پوشاک در شبکه‌های قرار گرفت و برای تحلیل داده‌ها از روش معادلات ساختاری و برای بررسی نقش میانجی از آزمون سوبل استفاده شد. این پژوهش تأثیر مستقل و ترکیبی متغیرهای سازمانی مؤثر بر اعتماد افراد در خرید پوشاک از طریق شبکه‌های اجتماعی را مورد بررسی قرار داده است. نتایج بدست آمده نشان داد کیفیت اطلاعات، امنیت معامله، شهرت شرکت و مکان شرکت تأثیر معنی‌داری بر اعتماد افراد، خرید پوشاک از طریق شبکه‌های اجتماعی، تبلیغات توصیه‌ای و وفاداری آنها دارد.

**واژگان کلیدی:** شبکه‌های اجتماعی، رفتار خرید مصرف‌کننده، صنعت پوشاک، عوامل سازمانی.

## ۱. مقدمه

فضای مجازی را برای افراد به گونه‌ای متفاوت با فضای حقیقی فراهم نموده اند [۱].

نتایج تحقیقات نشان می‌دهد درآمد تجارت از طریق شبکه‌های اجتماعی دائماً در حال رشد است و انتظار می‌رود تا سال ۲۰۲۰ به ۸۰ میلیارد برسد. بنابراین شبکه‌های اجتماعی بستر بسیار مهمی برای کسب درآمد کسب و کارها در آینده نزدیک

بکارگیری شبکه‌های اجتماعی در حوزه بازاریابی روز به روز در حال افزایش است و شبکه‌های اجتماعی با ایجاد بستر اجتماعی لازم، باعث جامعه‌پذیری فضای مجازی برای مصرف‌کنندگان شده‌اند و با ایجاد تعامل میان افراد، امکان تجربه خرید در

نویسنده مسئول: میترا دانش پرور

daneshparvar.basir@yahoo.com

است [۲]. شبکه‌های اجتماعی نه تنها ابزار مؤثر بازاریابی هستند بلکه به طور فزاینده‌ای به یک کانال فروش تبدیل شده‌اند. یکی از صنایعی که به شدت تحت تأثیر شبکه‌های اجتماعی قرار گرفته است، خرده فروشی پوشاک می‌باشد؛ به طوری که ۳۴٪ از خریداران پوشاک در انگلیس در سال ۲۰۱۸ بیان کرده‌اند که در خرید پوشاک از شبکه‌های اجتماعی استفاده می‌کنند این نشان از اهمیت این شبکه‌ها در صنعت پوشاک دارد [۳].

بر اساس تحقیقات انجام شده تخمین زده می‌شود که بازار جهانی پوشاک بیش از ۱,۷ تریلیون دلار ارزش داشته باشد و بیشتر مصرف‌کنندگانی که این بازار را ایجاد کرده‌اند، در شبکه‌های اجتماعی حضور فعال دارند. بنابراین شرکت‌های فروش پوشاک با حضور گسترده در شبکه‌های اجتماعی به دنبال استفاده از این سکوی بزرگ جدید هستند [۴].

از سوی دیگر، شبکه‌های اجتماعی با تشویق افراد برای اشتراک تجربیات زندگی، باعث افزایش تقاضا برای پوشاک شده‌اند؛ زیرا افراد تمایل دارند در هر عکس و یا فیلم با ظاهر جدید دیده شوند و این امر خود باعث افزایش تقاضا برای پوشاک شده است. متأسفانه در صنعت پوشاک شرکت‌های داخلی هنوز به طور سنتی فعالیت می‌کنند و نتوانسته‌اند به اندازه کافی مشتری جذب نمایند و بیشتر مشتریان به سمت پوشاک خارجی می‌روند؛ حتی بسیاری از برندهای پوشاک خارجی از طریق همین شبکه‌ها، مشتریان ایرانی را به خود جذب نموده‌اند. بنابراین توجه شرکت‌های ایرانی فعال در صنعت پوشاک به شبکه‌های اجتماعی فرصتی برای آنها فراهم می‌نماید تا علاوه بر معرفی محصولات خود، با برقراری ارتباط با مشتریان، از نظرات آنها آگاهی پیدا کرده و توسط خود مشتریان، محصولات خود را به دیگر مشتریان معرفی کرده و سهم بازار خود را افزایش دهند.

بررسی قرار گرفته است.

مقاله پیش رو در بخش دوم به مروری بر مبانی نظری و پیشینه تحقیق در زمینه متغیرهای سازمانی مؤثر بر خرید افراد در شبکه‌های اجتماعی می‌پردازد؛ سپس در بخش سوم روش تحقیق کاربردی پژوهش بیان می‌گردد. در این پژوهش، روش تحقیق داده بنیاد به عنوان ابزار تحقیقاتی مقاله مورد استفاده واقع شده است، بدین صورت که با فروشندگان پوشاک در شبکه‌های اجتماعی مصاحبه انجام شده و سپس به روش داده بنیاد، متون حاصل از مصاحبه کدگذاری گردیده و نتایج در قالب مقوله‌هایی ارائه شده است که مراحل هر یک به تفصیل در بخش مربوطه توضیح داده می‌شود. بخش چهارم به ارائه نتایج حاصل از تحلیل داده‌های کیفی و همچنین نتایج فرضیه‌های حاصل از مدل پیشنهادی می‌پردازد و در نهایت در بخش پنجم نتایج تحقیق ارائه می‌گردد.

بر اساس تحقیقات انجام شده تخمین زده می‌شود که بازار جهانی پوشاک بیش از ۱,۷ تریلیون دلار ارزش داشته باشد و بیشتر مصرف‌کنندگانی که این بازار را ایجاد کرده‌اند، در شبکه‌های اجتماعی حضور فعال دارند. بنابراین شرکت‌های فروش پوشاک با حضور گسترده در شبکه‌های اجتماعی به دنبال استفاده از این سکوی بزرگ جدید هستند [۴].

از سوی دیگر، شبکه‌های اجتماعی با تشویق افراد برای اشتراک تجربیات زندگی، باعث افزایش تقاضا برای پوشاک شده‌اند؛ زیرا افراد تمایل دارند در هر عکس و یا فیلم با ظاهر جدید دیده شوند و این امر خود باعث افزایش تقاضا برای پوشاک شده است. متأسفانه در صنعت پوشاک شرکت‌های داخلی هنوز به طور سنتی فعالیت می‌کنند و نتوانسته‌اند به اندازه کافی مشتری جذب نمایند و بیشتر مشتریان به سمت پوشاک خارجی می‌روند؛ حتی بسیاری از برندهای پوشاک خارجی از طریق همین شبکه‌ها، مشتریان ایرانی را به خود جذب نموده‌اند. بنابراین توجه شرکت‌های ایرانی فعال در صنعت پوشاک به شبکه‌های اجتماعی فرصتی برای آنها فراهم می‌نماید تا علاوه بر معرفی محصولات خود، با برقراری ارتباط با مشتریان، از نظرات آنها آگاهی پیدا کرده و توسط خود مشتریان، محصولات خود را به دیگر مشتریان معرفی کرده و سهم بازار خود را افزایش دهند.

اما باید توجه داشت محیط مجازی دارای ریسک‌های گوناگونی است. در چنین محیطی، کاربران به دنبال امنیت تضمین شده در معامله و همچنین حفظ حریم خصوصی خود هستند [۵]. محققان معتقدند امنیت معامله یک عامل اصلی محدودکننده خرید در فضای مجازی است و کنترل امنیت معامله و حفظ حریم خصوصی، عامل مهم و تأثیرگذاری بر اعتماد مصرف‌کننده هستند و هرگونه رخنه اطلاعات، تبعات

## ۲. مبانی نظری و مروری بر مطالعات گذشته

امروزه با افزایش محبوبیت شبکه اجتماعی، این شبکه‌ها به یکی از منابع جستجوی اطلاعات برای خرید کالاها و خدمات تبدیل شده‌اند و تأثیر زیادی بر رفتار مصرف‌کننده دارند [۷]. اما باید توجه داشت کاربران فضای مجازی معمولاً در فرایند خرید، نگران امنیت معامله هستند و مصرف‌کنندگانی که سطح بالاتری از امنیت معامله را احساس نمایند، احتمالاً سطح اعتماد در آنها بالاتر خواهد بود و شبکه‌های اجتماعی را به عنوان یک بازار مجازی بالقوه مد نظر قرار خواهند داد [۵]. امنیت معامله شامل ادراک کاربران از میزان امنیت فراهم شده در هنگام خرید از شبکه‌های اجتماعی است [۶]. یکی دیگر از عوامل مؤثر بر خرید مصرف‌کننده در شبکه‌های اجتماعی، شهرت شرکتهای فعال در این حوزه است. شهرت شرکت عبارت است از باور مصرف‌کننده درباره صداقت شرکت و دغدغه و توجه آن به مصرف‌کننده. شرکتی که از شهرت و تصویر ذهنی مطلوبی در بین افراد برخوردار است، از اعتماد بالای افراد نیز برخوردار خواهد بود. به علاوه، شهرت خوب، یک دارایی نامشهود باارزش برای شرکتهای فعال در فضای مجازی محسوب می‌باشد و باعث افزایش اعتماد مصرف‌کنندگان به شرکت می‌گردد [۷]. عامل مهم دیگری که بر خرید افراد از شبکه‌های اجتماعی مؤثر است، کیفیت اطلاعات موجود در سایت شرکت است. کیفیت اطلاعات شامل میزان دقت و کامل بودن اطلاعاتی است که توسط شرکت برای کاربر تهیه می‌شود. خریداران در فضای مجازی، وابستگی زیادی به اطلاعات فراهم شده توسط شرکت دارند. زیرا کاربران در فضای مجازی اطلاعات اندکی درباره کالا و خدمت دارند [۸]. بنابراین مصرف‌کنندگان تمایل به اعتماد به سایتی دارند که اطلاعات دقیق و به موقع به آنها ارائه

می‌نمایند. در این راستا، تحقیقات نشان می‌دهد کیفیت اطلاعات برای سایت‌های تجارت اجتماعی احتمالاً مهمتر از سایر انواع سایت‌های تجارت الکترونیک است. زیرا اطلاعات مربوط به کالاها و خدمات موجود در این سایت‌ها توسط مصرف‌کنندگان تهیه می‌شود [۹]. خلاصه ای از تحقیقات انجام شده در جدول ۱ ارائه شده است.

از این رو، امروزه اهمیت شبکه‌های اجتماعی بر تصمیم‌گیری مصرف‌کنندگان به سرعت در حال افزایش است، به طوری که محققان تخمین می‌زنند ۹۷٪ از مصرف‌کنندگان نظرهای سایرین را بررسی می‌کنند و تحت تأثیر آنها قرار می‌گیرند [۱۰] [۱۱]. از این رو، شبکه‌های اجتماعی در حال تبدیل شدن به بخش جدایی ناپذیر استراتژی کسب و کار هستند و تعاملات شرکت با مشتریان را به شدت تحت تأثیر قرار می‌دهند [۱۲]. تحقیقات نشان داده است ۷۰٪ از افرادی که از شبکه‌های اجتماعی استفاده می‌کنند به تجربیات یکدیگر گوش می‌دهند، ۶۵٪ درباره کالاها و خدمات شرکت یاد می‌گیرند، ۵۳٪ تجربیات خوب و ۵۰٪ نگرانی‌های خود درباره کالاها و خدمات را بیان می‌کنند [۱۳]. همچنین مصرف‌کنندگان به توصیه دیگر مصرف‌کنندگان، ۹۰٪ اعتماد دارند در حالی که به تبلیغات شرکت ۵۶٪ اعتماد دارند. تحقیق دیگری نیز در این خصوص انجام شده که نشان می‌دهد ۶۲٪ از مصرف‌کنندگان آمریکایی به مرور نظرهای آنلاین دیگر مصرف‌کنندگان در خصوص محصولات مورد نظرشان می‌پردازند و ۹۸٪ از آنها این نظرهای را قابل اعتماد می‌دانند [۱۲]. تحقیقات همچنین نشان دهنده محبوبیت بیشتر شبکه‌های اجتماعی نسبت به موتورهای جستجو در خصوص آگاهی از نظرهای آنلاین مصرف‌کنندگان درباره محصولات مختلف است [۷]. خلاصه- ای از تحقیقات انجام شده در جدول ۱ ارائه شده است.

جدول ۱. خلاصه تحقیقات پیشین

شماره مرجع	متغیرها	عنوان تحقیق	یافته‌ها
[۱۴]	امنیت مبادله و تصمیم خرید	تجارت اجتماعی در بستر موبایل	امنیت مبادله تأثیر معنی داری بر تصمیم خرید دارد.
[۵]	کیفیت اطلاعات، تمایل به خرید	بررسی عوامل مؤثر تمایل به خرید در تجارت اجتماعی	کیفیت اطلاعات سایتهای تجارت اجتماعی تأثیر مثبت بر تمایل به خرید افراد دارد.
[۸]	اعتبار اطلاعات، مفید بودن اطلاعات، تمایل به خرید	ارائه الگویی برای سازگاری تجارت اجتماعی	اعتبار اطلاعات و مفید بودن اطلاعات سایت‌های شبکه‌های اجتماعی تأثیر زیادی بر تمایل به خرید آنها دارد.
[۱۵]	ریسک اطلاعات، جستجوی اطلاعات و نیاز به اطلاعات، تمایل به خرید	نقش اطلاعات در شبکه‌های اجتماعی	ریسک اطلاعات، جستجوی اطلاعات و نیاز به اطلاعات، تأثیر معنی داری بر تمایل به خرید دارند.
[۱۶]	کیفیت اطلاعات و تمایل به خرید	بررسی عوامل مؤثر بر تمایل افراد به شبکه‌های اجتماعی	حمایت اطلاعاتی و کیفیت اطلاعات تأثیر معنی داری بر تمایل افراد به وب سایت‌های شبکه اجتماعی دارد.
[۱۷]	کیفیت اطلاعات، تمایل به خرید	مدل تصمیم‌گیری خرید مبتنی بر تعاملات اجتماعی در شبکه‌های اجتماعی	کیفیت اطلاعات تأثیر معنی داری بر تمایل افراد به خرید از شبکه اجتماعی دارد.
[۱۸]	اعتماد، امنیت معامله، تمایل به خرید	نگرانی از امنیت و تمایل به خرید در شبکه‌های اجتماعی	امنیت معامله تأثیر زیادی بر اعتماد دارد و اعتماد نیز تأثیر معنی داری بر تمایل به خرید در شبکه‌های اجتماعی دارد.
[۳]	اعتماد، شهرت شرکت و کیفیت اطلاعات	بررسی تأثیر اعتماد و عملکرد اعتماد بر تجارت اجتماعی	کیفیت اطلاعات و شهرت شرکت تأثیر معنی داری بر اعتماد افراد دارد و اعتماد تأثیر مثبت بر تمایل به خرید در از طریق شبکه‌های اجتماعی دارد.



## ۱-۲ توسعه فرضیه‌ها و مدل مفهومی

توسعه سریع شبکه‌های اجتماعی، پتانسیل عظیمی را برای انتقال تجارت الکترونیک از یک محیط محصول محور به یک محیط اجتماعی و متمرکز بر مشتری فراهم کرده است به همین دلیل، مدیران به دنبال توسعه ساختارهای مناسب برای فروش از طریق شبکه‌های اجتماعی هستند [۱۴] [۱۹]. از آنجایی که شبکه‌های اجتماعی باعث ارتباط فروشنده و خریدار می‌گردد، فروشندگان باید با تاکید بر موضوعاتی نظیر اعتماد و همچنین حریم خصوصی و امنیت، بر ادراک و واکنش‌های مصرف‌کنندگان تأثیر بگذارند؛ اگرچه مدیریت امنیت در فضای مجازی سخت‌تر است و کاربران هیچ کنترلی بر مسائل امنیتی در فضای مجازی ندارند [۲۰]. همچنین شهرت نیک شرکت، کیفیت وب سایت و تعامل با سایر خریداران از طریق شبکه‌های اجتماعی ممکن است باعث اعتماد مشتری شوند و نگرانی‌های پس از خرید را کاهش دهند [۲۱].

### اعتماد

اعتماد یکی از عوامل مهم در خرید از طریق شبکه‌های اجتماعی است [۱۶]. اعتماد به معنای باور ذهنی مصرف‌کننده در مورد نحوه انجام تعهدات فروشنده نسبت به مصرف‌کننده است که در خرید از فضای مجازی اهمیت بسیار زیادی دارد و بدون وجود اعتماد، روابط میان شرکتها و مصرف‌کنندگان در فضای مجازی امکان پذیر نمی‌باشد [۳] [۱۶] [۲۲]. اعتماد باعث کاهش تردیدهای رفتاری در رفتار خرید مصرف‌کننده از یک سو و افزایش قدرت کنترل مصرف‌کننده در معاملات در شبکه‌های اجتماعی از سوی دیگر می‌گردد که این قدرت باعث افزایش تصمیم خرید مصرف‌کننده می‌شود [۲۳]. محققان زیادی بر اهمیت اعتماد در خرید از شبکه‌های اجتماعی تاکید کرده‌اند که از جمله این آنها می‌توان به [۳] - [۲۰] - [۱۷] - [۲۱] - [۳] - [۲۴] - [۲۵] - [۲۶] - [۱۶] - [۲۳] - [۲۷] - [۲۸] - [۲۹] اشاره نمود. با توجه به اهمیت تأثیر اعتماد بر موفقیت تجارت اجتماعی و بر اساس مصاحبه‌های انجام گرفته می‌توان فرضیه زیر را مطرح نمود:

فرضیه ۱: اعتماد تأثیر معنی‌داری بر تصمیمی مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد.

### عوامل سازمانی

محققان معتقدند ادراک مصرف‌کننده از شهرت یک سایت شبکه اجتماعی، نقش بسیار مهمی در ایجاد اعتماد در مصرف‌کنندگان نسبت به آن سایت دارد. بنابراین شهرت خوب یک شرکت، معمولاً بین افراد منتشر می‌شود و به افزایش اعتماد در بین مصرف‌کنندگان منجر می‌گردد [۱]. شهرت شرکت به عنوان یک دارایی نامشهود کمیاب و ارزشمند و پایدار است که به سختی توسط رقبا قابل تقلید است. شهرت شرکت منجر به مزیت رقابتی می‌گردد و چنین شرکت‌هایی، سود بالاتری را تجربه می‌کنند که این شهرت برای کارکنان و مشتریانشان جذاب است [۳۰]. پژوهشگران زیادی بر اهمیت شهرت شرکت در تمایل افراد به خرید از شبکه‌های اجتماعی تاکید داشته‌اند که از جمله آنها می‌توان به [۳۰] - [۳] - [۱] - [۳۱] اشاره نمود. با توجه به اهمیت تأثیر عوامل سازمانی و بویژه شهرت شرکت در موفقیت تجارت اجتماعی و بر اساس مصاحبه‌های انجام گرفته می‌توان فرضیه زیر را مطرح نمود:

فرضیه ۲: عوامل سازمانی تأثیر معنی‌داری بر تصمیم مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد.

فرضیه ۱-۲: شهرت شرکت تأثیر معنی‌داری بر اعتماد مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد. برخی از محققان کیفیت اطلاعات را شامل ادراک عمومی مصرف‌کننده درباره کامل بودن و صحت اطلاعات ارائه شده درباره خدمات پیشنهادی و فرایند انجام معامله خرید می‌دانند و معتقدند زمانی که شبکه‌های اجتماعی مختلف دارای اطلاعات مشابه درباره محصول هستند، عاملی که باعث انتخاب یک شبکه اجتماعی خاص توسط مصرف‌کننده می‌شود، ادراک وی از کیفیت اطلاعاتی است که ارائه شده است. زیرا ادراک مصرف‌کننده درباره کیفیت اطلاعات، باعث اعتماد وی به آن شبکه می‌گردد [۲۲]. محققان زیادی بر اهمیت کیفیت اطلاعات در تصمیم خرید افراد از شبکه‌های اجتماعی تاکید داشته‌اند که از جمله آنها می‌توان به [۳۲] -

[۱۶]- [۳۰]- [۱]- [۳۱]- [۵]- [۲۲]- [۳۳]- [۲۵]- [۳۴]- [۳۵] اشاره نمود. با توجه به اهمیت تأثیر کیفیت اطلاعات بر موفقیت تجارت اجتماعی و بر اساس مصاحبه‌های انجام گرفته می‌توان فرضیه زیر را مطرح نمود:

فرضیه ۲-۲: کیفیت اطلاعات شرکت تأثیر معناداری بر اعتماد مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد.

ایجاد امنیت معامله در فضای مجازی مشکل‌تر است؛ زیرا از یک سو مدیریت امنیت در فضای مجازی سخت‌تر است و از سوی دیگر کاربران هیچ کنترلی بر مسائل امنیتی در فضای مجازی ندارند. تعدادی از محققان معتقدند که کاربران فضای مجازی معمولاً در فرایند خرید، نگران امنیت معامله هستند. آنها معتقدند امنیت معامله یک عامل اصلی محدودکننده خرید در فضای مجازی است و کنترل امنیت معامله و حفظ حریم خصوصی، عامل مهم و تأثیرگذاری بر اعتماد مصرف‌کننده هستند. بنابراین مصرف‌کنندگانی که سطح بالاتری از امنیت معامله را احساس نمایند، احتمالاً سطح اعتماد در آنها بالاتر خواهد بود و شبکه‌های اجتماعی را به عنوان یک بازار مجازی بالقوه مد نظر قرار خواهند داد [الف]. تحقیقات زیادی بر اهمیت معامله در تصمیم خرید افراد از شبکه‌های اجتماعی تاکید داشته‌اند که از جمله این آنها می‌توان به تحقیقات ( [۲۱]- [۳۰]- [۱]- [۳۱]- [۳۶]- [۳۵]) اشاره نمود. با توجه به اهمیت تأثیر امنیت معامله بر موفقیت تجارت اجتماعی و بر اساس مصاحبه‌های انجام گرفته می‌توان فرضیه زیر را مطرح نمود: فرضیه ۲-۳: امنیت معامله تأثیر معناداری بر اعتماد مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد.

بسیاری از افراد ترجیح می‌دهند از شبکه اجتماعی مربوط به شرکتی خرید کنند که دارای محل فروش آفلاین می‌باشد؛ زیرا این امر باعث اعتماد فرد به آن شبکه شده و نگرانی فرد در هنگام خرید را کاهش می‌دهد. از سوی دیگر نتایج حاصل از مصاحبه با فروشندگان پوشاک در شبکه‌های اجتماعی نشان می‌دهد که برخورداری شرکت‌های فعال در شبکه‌های اجتماعی از محل فروش آفلاین و همچنین موقعیت جغرافیایی آن مکان، تأثیر زیادی بر اعتماد افراد به شبکه‌های اجتماعی و تمایل آنها به خرید پوشاک از این شبکه‌ها دارد. بنابراین بر اساس مصاحبه‌های انجام گرفته می‌توان فرضیه زیر را مطرح نمود:

فرضیه ۲-۴: محل فروش آفلاین شرکت تأثیر معنی‌داری بر اعتماد مصرف‌کننده به خرید پوشاک در شبکه‌های اجتماعی دارد.

اعتماد به شبکه‌های اجتماعی نه تنها بر تصمیم مصرف‌کننده برای خرید پوشاک تأثیر می‌گذارد بلکه باعث خرید مجدد، وفاداری و توصیه به دیگران برای خرید از سایت شبکه اجتماعی نیز می‌گردد. بنابراین بر اساس مصاحبه‌های انجام شده می‌توان فرضیه‌های زیر را مطرح نمود:

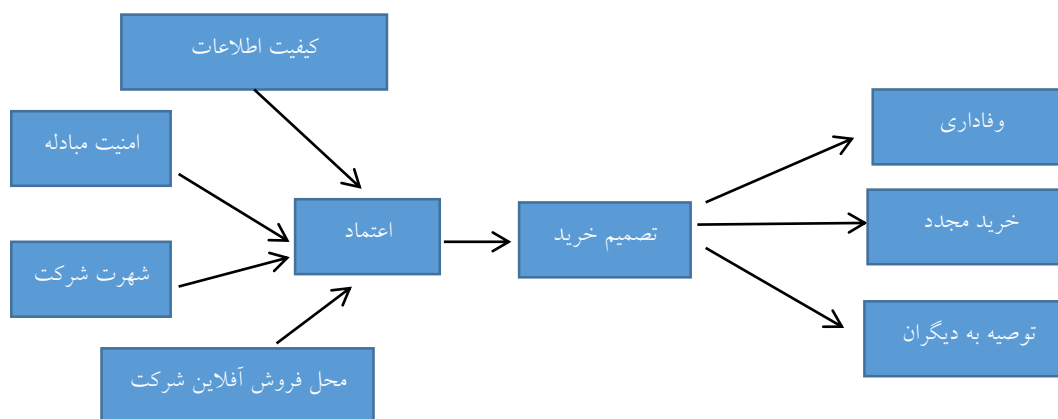
فرضیه ۳: تصمیم خرید مشتریان تأثیر معناداری بر وفاداری آنها به خرید پوشاک از سایت‌های اجتماعی دارد.

فرضیه ۴: تصمیم خرید مشتریان تأثیر معناداری بر خرید مجدد آنها از سایت‌های اجتماعی دارد.

فرضیه ۵: تصمیم خرید مشتریان تأثیر معناداری بر توصیه آنها به دیگران برای خرید پوشاک از سایت‌های اجتماعی دارد.

جدول ۲. متغیرهای سازمانی مؤثر بر رفتار خرید مصرف کننده در شبکه های اجتماعی (استخراج شده از ادبیات موجود)

شماره مرجع	[۱۱]	[۱۰]	[۹]	[۸]	[۷]	[۶]	[۵]	[۴]	[۳]	[۲]	[۱]	[۶۸]	[۶۷]	[۶۶]	[۶۵]	[۶۴]
سال	۲۰۱۵	۲۰۱۷	۲۰۱۹	۲۰۱۸	۲۰۱۵	۲۰۱۸	۲۰۱۵	۲۰۱۶	۲۰۱۵	۲۰۱۵	۲۰۱۸	۲۰۱۵	۲۰۱۸	۲۰۱۶	۲۰۱۶	۲۰۱۶
اعتماد	•			•	•	•	•	•	•	•	•		•	•	•	•
کیفیت اطلاعات	•	•		•	•		•	•	•	•						
امنیت مبادله	•	•		•	•								•	•		
شهرت شرکت	•	•		•	•						•		•	•		



شکل ۱. مدل اولیه عوامل سازمانی مؤثر بر رفتار خرید مصرف کننده در شبکه های اجتماعی

### ۳. روش شناسی پژوهش

با توجه به اینکه در این پژوهش از داده‌های کمی و کیفی استفاده شده است، این پژوهش از حیث نوع داده و نحوه

گردآوری و تحلیل داده‌ها از نوع تحقیقات ترکیبی است. امروزه "شیوه‌های تحقیق ترکیبی" بیشترین کاربرد را داشته و شامل گردآوری و تحلیل داده‌های کمی و کیفی

است [۴۰]. این شیوه پژوهش، بر جمع‌آوری، تحلیل و تلفیق داده‌های کمی و کیفی در یک پژوهش مجزا یا مجموعه‌ای از پژوهش‌ها تمرکز می‌کند و مهم‌ترین فرض آن این است که با بکارگیری ترکیب رویکردهای کمی و کیفی، درک بهتری نسبت به زمانی که هر رویکرد به صورت جداگانه به کار برده می‌شود، از موضوع تحقیق به دست می‌دهد. روش تحقیق ترکیبی، انواع مختلفی دارد که طرح سه بعدی، طرح ترکیبی، طرح توضیحی و طرح اکتشافی چهار نوع رایج آن می‌باشد. با توجه به ماهیت و کارکرد هر یک از انواع چهارگانه روش تحقیق ترکیبی، به نظر می‌رسد طرح اکتشافی مناسب‌ترین نوع از روش تحقیق ترکیبی است که در این پژوهش می‌تواند مورد استفاده قرار گیرد. هدف از این روش تحقیق دو مرحله‌ای، تحت تأثیر قرار دادن نتایج فاز اول (کیفی) بر فاز دوم (کمی) می‌باشد [۴۰]. عوامل

بسیاری بر خرید از طریق شبکه‌های اجتماعی مؤثر هستند اما این پژوهش به دنبال شناسایی عوامل و متغیرهای تاثیرگذاری است که در اختیار و تحت کنترل شرکت هستند که در این پژوهش "عوامل سازمانی" نامیده شده‌اند. بنابراین روش تحقیق این پژوهش در فاز کیفی از حیث هدف، اکتشافی و از نظر نتیجه توسعه‌ای محسوب می‌شود. روش تحقیق این پژوهش در فاز کمی از حیث هدف، تبیینی و از نظر نتیجه، کاربردی محسوب می‌شود و استراتژی پژوهش در این فاز از نوع پیمایش می‌باشد. جامعه آماری این پژوهش در فاز کیفی شامل خبرگان و مدیران شرکتهای فعال صنعت پوشاک در شبکه‌های اجتماعی و در فاز کمی شامل کاربران شبکه اجتماعی که تجربه خرید از این شبکه‌ها داشته‌اند، می‌باشد. خلاصه‌ای از روش کلی پژوهش، در جدول ۳ بیان شده است:

جدول ۳. خلاصه‌ای از روش کلی پژوهش

فاز دوم	فاز اول	مبنای طبقه بندی
تبیینی	اکتشافی	روش تحقیق بر اساس هدف
کاربردی	توسعه‌ای	روش تحقیق بر اساس نتیجه پژوهش
کمی	کیفی	روش تحقیق بر اساس نوع داده‌ها
غیراحتمالی در دسترس	قضاوتی و هدفمند	روش نمونه‌گیری
پرسشنامه	مصاحبه عمیق	روش جمع‌آوری داده‌ها
معادلات ساختاری	داده بنیاد	روش تجزیه و تحلیل داده‌ها

فضای مجازی محصولات خود را ارائه می‌نمودند، مصاحبه انجام دادند که این مصاحبه تا اشباع نظری ادامه یافت. در مرحله بعد، داده‌های حاصل از مصاحبه‌ها، کدگذاری شدند. کدگذاری، روشی است که در خلال آن، داده‌ها تجزیه، مفهوم‌سازی و به شکل نظریه در می‌آیند [۴۱]. برای تحلیل داده‌ها در روش داده بنیاد، از سه روش کدگذاری استفاده می‌شود: کدگذاری باز، محوری و انتخابی. روند خرد کردن، مقایسه کردن، مفهوم‌پردازی و مقوله-پردازی داده‌ها را کدگذاری باز می‌گویند [۴۲]. خرد کردن

در این پژوهش ابتدا با بررسی گسترده ادبیات تحقیق، متغیرهای مرتبط با موضوع تحقیق استخراج شد (جدول ۲) که در بخش توسعه فرضیه‌ها ارائه شده است. سپس در بخش مطالعه کیفی، بر اساس مصاحبه‌های عمیق انجام شده با فروشندگان پوشاک در شبکه‌های اجتماعی، مؤلفه‌های سازمانی مؤثر بر خرید پوشاک از طریق شبکه‌های اجتماعی شناسایی شدند. برای این منظور، محققان با ۱۴ نفر از فروشندگان پوشاک که بیش از ۱۰ سال در صنعت پوشاک حضور فعال داشته و به مدت حداقل ۵ سال در

بدین معناست که جمله، عبارت، مضمون و یا تصویری به چند عنصر محتوایی تقسیم شود؛ به نوعی که هر عنصر نام اختصاصی خود را بگیرد. این مرحله اساس روش کدگذاری باز است.

مقایسه کردن فرایندی است که طی آن، دائمی داده‌ها با مفاهیم و مقولات ایجادشده با هدف ارتباط میان مفاهیم و مقولات و شاخص‌های آن‌ها و داده‌های واقعی مورد مقایسه قرار می‌گیرد. این عمل مداوم با طرح سلسله سولاتی نظیر چه چیز، چگونه، چه زمانی، کجا، با چه هدفی میسر می‌شود. مفهوم پردازی، ایجاد واژه‌هایی برای نامیدن جداگانه وقایع، حوادث، رخدادها و پدیده‌هاست. مقوله‌پردازی مرحله طبقه‌بندی مفاهیم است. مقوله‌ها در پی مقایسه مفاهیم و تشابه پدیده‌ها کشف شده و مفاهیم در مرحله‌ای بالاتر بانظمی دقیق‌تر دسته بندی می‌شوند و مقوله را می‌سازند. مقوله مفهومی است که از سایر مفاهیم انتزاعی‌تر است.

در کدگذاری باز، متون مورد مطالعه که می‌تواند یک جمله، یک بند و حتی صحبت‌های مصاحبه شونده‌ها باشد، به اجزای کوچک‌تری تقسیم شده و طی فرایند دائمی مقایسه، مفهوم پردازی و مقوله‌بندی می‌شوند. در نتیجه ایجاد نام برای پدیده‌ها در قالب مفاهیم، گام نخست تجزیه و تحلیل در کدگذاری باز محسوب می‌شود. این مفاهیم باید در عین نزدیک بودن به متن دارای سطحی از انتزاع نیز باشند تا بتوانند چندین پدیده مشابه را شامل شوند. اهمیت این مرتبه از کدگذاری به قدری زیاد است که بازبینی چندباره و گاهی تغییر برخی مفاهیم را پیش از مرحله مقوله‌پردازی سبب می‌شود. این امر سبب تولید مفاهیم بسیاری می‌گردد [۴۲]؛ در این مرحله محدودیتی در تعداد مقوله‌ها و کدها موجود نیست.

کدگذاری محوری، فرایند ربط‌دهی مقوله‌ها به زیر مقوله‌ها و پیوند دادن مقوله‌ها در سطح ویژگی و ابعاد است. این کدگذاری به این دلیل "محوری" نامیده شده است که کدگذاری حول محور یک مقوله تحقق می‌یابد [۴۲]. در این مرحله، مقوله‌ها، ویژگی‌ها و ابعاد حاصل از کدگذاری

باز، تدوین شده و سر جای خود قرار می‌گیرد تا دانش فزاینده‌ای در مورد روابط ایجاد گردد. در این مرحله، پژوهشگر، یک مقوله مرحله کدگذاری باز را انتخاب و آن را در مرکز فرایندی که در حال بررسی آن است (به عنوان پدیده مرکزی، مقوله هست) قرار می‌دهد و سپس، دیگر مقوله‌ها را به آن ربط می‌دهد. این مقوله‌ها در نمودار الگوی کدگذاری قابل مشاهده است [۴۲].

در کدگذاری باز، تحلیل‌گر به پدید آوردن مقوله‌ها و ویژگی‌های آن‌ها می‌پردازد و سپس می‌کوشد تا مشخص کند که چگونه مقوله‌ها در طول بدهای تعیین شده تغییر می‌کنند. در کدگذاری محوری، مقوله‌ها به طور نظام‌مند بهبودیافته و با زیر مقوله‌ها پیوند داده می‌شوند. با این حال، این‌ها هنوز مقوله‌های اصلی نیستند که در نهایت برای تشکیل یک آرایش نظری بزرگ تر یکپارچه شوند، به طوری که نتایج تحقیق، شکل نظریه پیدا کنند. "کدگذاری انتخابی" فرایند یکپارچه سازی و بهبود مقوله‌هاست [۴۲].

با ورود داده‌های حاصل از بخش کیفی به نرم افزار مکس-کیودی‌ای و کدگذاری آنها به روش‌های فوق، مدل اولیه عوامل سازمانی مؤثر بر خرید پوشاک از طریق شبکه‌های اجتماعی تدوین گردید. سپس بر اساس نتایج بخش کیفی و مدل اولیه، پرسشنامه‌ای در قالب طیف لیکرت ۵ گزینه‌ای طراحی شد و توسط ۳۸۵ نفر که از طریق شبکه‌های اجتماعی اقدام به خرید پوشاک می‌نمودند و به روش نمونه-گیری غیراحتمالی در دسترس انتخاب شده بودند تکمیل گردید. در فرایند تکمیل پرسشنامه، از روش‌های الکترونیکی و سنتی بهره گرفته شد. در روش الکترونیکی، نسخه الکترونیک پرسشنامه در سایت‌های فروش پوشاک در شبکه‌های اجتماعی و همچنین سایت‌های عمومی شبکه‌های اجتماعی قرار داده شد که ۲۴۰ پرسشنامه تکمیل گردید. همچنین ۳۵۰ پرسشنامه نیز به صورت دستی توزیع شد که از بین آنها ۱۴۵ پرسشنامه از قابلیت تجزیه و تحلیل برخوردار بود.

#### ۴. یافته‌های پژوهش

##### ۴-۱ بخش کیفی

شده توسط پژوهشگر در این نظریه، برای تشریح فرایندها، شامل انواع بسیاری از داده‌های کیفی است، از جمله مشاهدات، گفت و شنودها، مصاحبه‌ها، اسناد دولتی، خاطرات پاسخ دهندگان و مجلات و تأملات شخصی خود پژوهشگر [۴۰].

در این پژوهش، معیار جست‌وجو در ابتدا بررسی مقالات موجود در زمینه عوامل سازمانی مؤثر بر خرید از طریق شبکه‌های اجتماعی بود که نتیجه آن در بخش مبانی نظری ارائه شده است. سپس با ۱۴ نفر از فروشندگان پوشاک که از طریق شبکه‌های اجتماعی اقدام به فروش محصولات خود می‌نمایند، مصاحبه‌ای انجام شد. پس از مصاحبه‌های صورت گرفته با اعضای نمونه آماری، فایل‌های صوتی ایشان به فایل متنی تبدیل شد و به روش پدیدار شناسی، کدگذاری باز انجام پذیرفت. پس از جمع بندی و دسته‌بندی هر یک از مفاهیم مذکور و با توجه به ادبیات مورد مطالعه، کدگذاری محوری ابعاد عوامل سازمانی به صورت زیر صورت گرفت:

این پژوهش به دنبال یافتن عوامل سازمانی مؤثر بر خرید پوشاک از طریق شبکه‌های اجتماعی است که برای این منظور از روش کیفی تئوری داده بنیاد استفاده شده است. رویکرد نظریه داده بنیاد یک نوع روش پژوهش کیفی است که به طور استقرایی یک سلسله رویه‌های سامان‌مند را به کار می‌گیرد تا نظریه‌ای درباره پدیده مورد مطالعه ایجاد کند [۴۲]. این روش به هیچ نوع داده خاص، راهبرد پژوهشی، یا علائق نظری مشخص تعهدی ندارد و تنها در پی توسعه مفاهیم است. در واقع، ترکیب داده‌ها در قالب چهارچوب‌های مفهومی است؛ که در رفت و برگشت دائم میان داده‌ها و تحویل‌ها ساخته می‌شود یعنی از جمع‌آوری نظام‌مند داده‌ها به دست می‌آید [۴۳]. این روش اولین بار در دهه ۱۹۶۰ توسط بارنی گلیسروآنسلم استراوس ارائه شد [۴۴]. پژوهش بر اساس نظریه داده بنیاد، به دنبال تولید فهمی از یک فرایند به کمک زنجیره‌ای از کنش‌ها و برهم کنش‌ها بین افراد و وقایع مربوط به یک موضوع است [۴۰]. داده‌های جمع‌آوری

جدول ۴. کدگذاری محوری عوامل سازمانی مؤثر بر خرید پوشاک در شبکه‌های اجتماعی در بخش کیفی

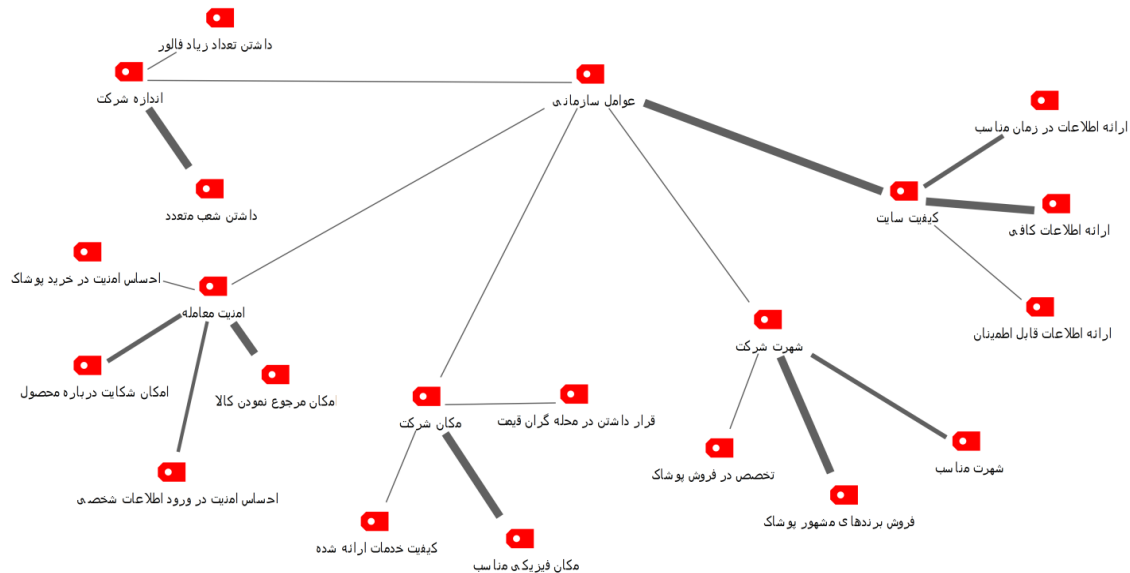
مقوله	مفاهیم متناظر	تعداد تکرار
کیفیت اطلاعات	ارائه اطلاعات قابل اطمینان	۱۲
	ارائه اطلاعات کافی	۱۴
	ارائه اطلاعات در زمان مناسب درباره محصول	۱۰
شهرت شرکت	برخورداری شرکت از شهرت مناسب	۱۱
	فروش برندهای مشهور پوشاک*	۱۳
	تخصص شرکت در فروش پوشاک*	۸
محل فروش آفلاین شرکت*	برخورداری شرکت از مکان فیزیکی*	۱۴
	قرار گرفتن مکان فیزیکی در محله گران قیمت*	۹
	کیفیت خدمات ارائه شده توسط شبکه‌های اجتماعی*	۷
امنیت معامله	احساس امنیت هنگام ورود اطلاعات شخصی	۱۴
	امکان تعویض و مرجوع نمودن کالا*	۱۲
	امکان شکایت درباره محصول*	۸
	احساس امنیت در فرایند خرید پوشاک	۱۲
اندازه شرکت در	داشتن شعب متعدد	۱۰

شبکه های اجتماعی	داشتن فالور زیاد	۸
اعتماد		۱۴
توصیه به دیگران		۱۱
خرید مجدد		۱۰
وفاداری		۱۳

نکته: موارد ستاره دار، متغیرهایی هستند که از تحقیق کیفی بدست آمده‌اند و در ادبیات مورد بررسی قرار نگرفته‌اند.

سازمانی مؤثر بر خرید پوشاک از طریق شبکه های اجتماعی  
به صورت شکل ۲ بدست آمد:

در مرحله بعد، ابعاد، مقوله ها و مفاهیم متناظر آنها وارد نرم افزار مکس کیودی‌ای<sup>۱</sup> شده و کدگذاری شدند. بر اساس کدگذاری‌های انجام شده در نرم افزار، مدل اولیه عوامل

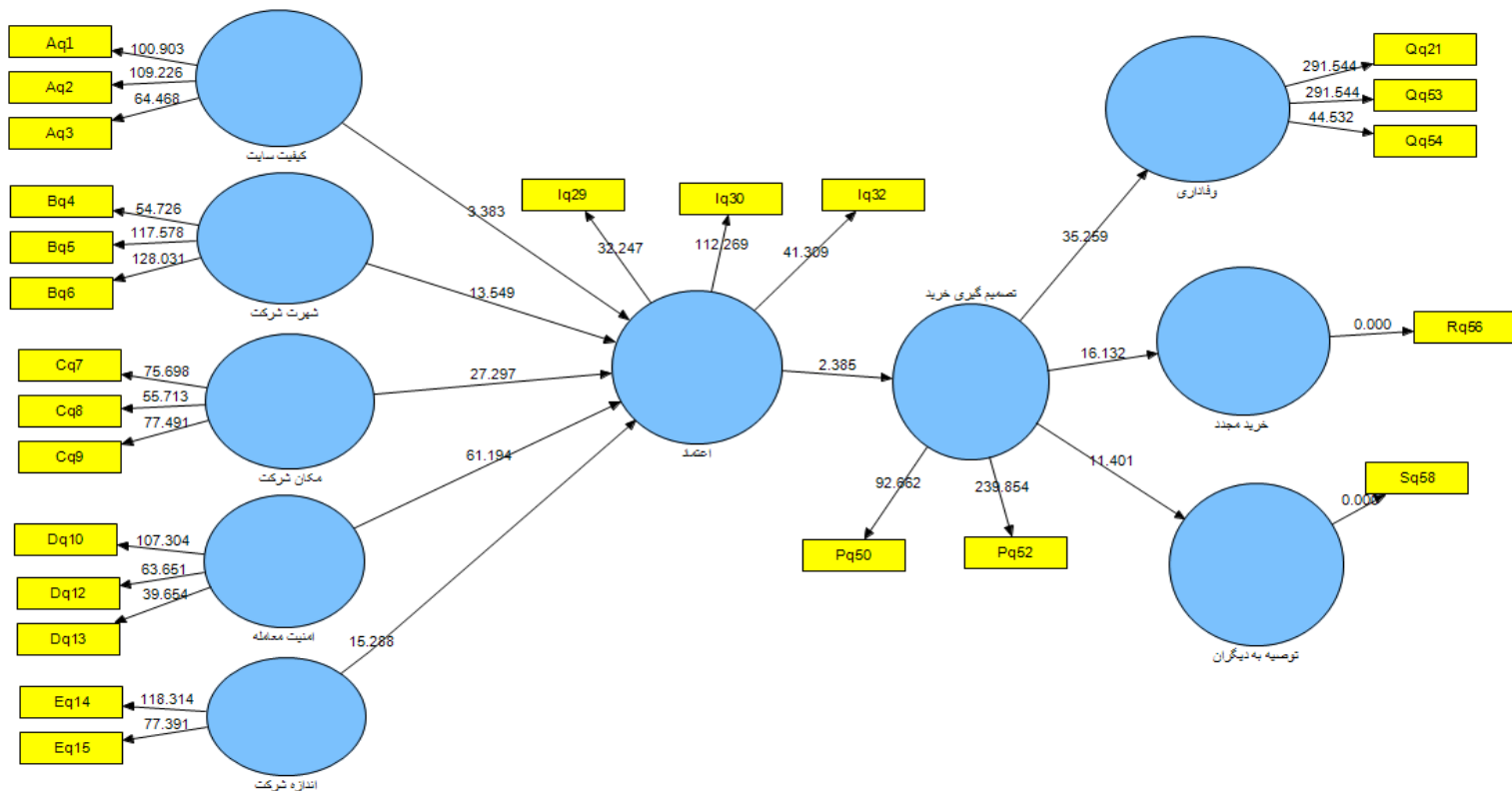


شکل ۲: مدل اولیه عوامل سازمانی مؤثر بر خرید پوشاک از طریق شبکه های اجتماعی بر اساس خروجی نرم افزار مکس کیودی‌ای

#### ۲-۴ بخش کمی

بر اساس نتایج بخش کیفی، فرضیه‌های تحقیق تدوین شد و از طریق پرسشنامه مورد بررسی قرار گرفت و داده‌های حاصل از پرسشنامه، تحلیل گردید. برای بررسی اعتبار سازه‌ها و همچنین آزمون فرضیه‌های پژوهش از مدل‌یابی معادلات ساختاری استفاده شد. قبل از ورود به معادلات ساختاری و نحوه استفاده از آزمون‌های آن در قالب نرم‌افزارهای موجود، ابتدا از آزمون کولموگروف – اسمیرنوف برای بررسی توزیع داده‌ها استفاده گردید که نتایج نشان داد سطح معنی داری برای تمامی متغیرها

کمتر از ۰/۰۵ می باشد که این نشان از عدم نرمال بودن توزیع داده‌ها است. لذا تصمیم بر آن شد که از روش‌های ناپارامتریک در راستای تحلیل داده‌ها استفاده گردد. برای اطمینان از برازش مدل اندازه‌گیری، نیاز به انجام آزمون معناداری روابط بین شاخص‌ها و مؤلفه‌ها است. از این رو مدل در حالت معناداری اجرا شده و ضرایب معناداری کنترل شده است.



شکل ۳. ضرایب معناداری برای مدل اندازه‌گیری کلی تأثیر عوامل بر ابعاد

برای اندازه‌گیری برخوردار هستند و لذا روایی سازه نشان می‌دهد که نشانگرها، ساختارهای عاملی مناسبی را جهت اندازه‌گیری ابعاد مورد مطالعه در مدل تحقیق فراهم می‌آوردند.

علاوه بر روایی سازه که برای بررسی اهمیت نشانگرهای انتخاب شده برای اندازه‌گیری سازه‌ها به کار می‌رود، روایی تشخیصی<sup>۳</sup> نیز مورد بررسی قرار گرفت. روایی سازه بدین معنا است که هر نشانگر فقط سازه خود را اندازه‌گیری کند و ترکیب آنها به گونه‌ای باشد که تمام سازه‌ها به خوبی از یکدیگر تفکیک شوند. این فرایند با کمک شاخص میانگین واریانس استخراج شده صورت گرفت که مشخص شد تمام سازه‌های مورد مطالعه دارای میانگین واریانس استخراج شده<sup>۴</sup> بالاتر از ۰/۴ هستند که این ضرایب در جدول شماره ۵

با توجه به ضرایب معناداری درج شده بر روی خطوط متصل به سؤالات مشخص است، که همه سؤالات دارای معناداری مناسب هستند. لذا با اطمینان می‌توان گفت که شاخص‌های موجود در مدل اندازه‌گیری تحقیق به نحو مناسبی قادر به اندازه‌گیری و پیش‌بینی تغییرات مؤلفه‌های موجود در مدل می‌باشند.

به منظور آزمون فرضیات تحقیق از روش معادلات ساختاری با استفاده از روش حداقل مربعات جزئی<sup>۱</sup> استفاده شد و برای بررسی دقت و اهمیت نشانگرهای انتخاب شده برای اندازه‌گیری سازه‌ها، از روایی سازه<sup>۲</sup> استفاده گردید که نتایج آن در جدول ۵ گزارش شده است. بر اساس نتایج این جدول، نشانگرهای باقی مانده در هر سازه به دلیل داشتن مقدار بیشتر از ۰/۵ از اهمیت لازم

۳. Discriminant Validity  
۴. AVE

۱. PLS  
۲. Construct Validity



معیار دوم بررسی برازش مدل‌های اندازه‌گیری، روایی همگرا است که به بررسی میزان همبستگی هر سازه با سؤالات (شاخص‌ها) خود می‌پردازد. همانطور که نتایج جدول ۵ نشان می‌دهد همه متغیرهای پنهان دارای شاخص قابل قبولی هستند که نشان از برازش خوب مدل دارد.

نشان داده شده است. در نهایت جهت بررسی پایایی نیز از شاخص پایایی ترکیبی، اشتراکی و آلفای کرونباخ استفاده شد که نتایج در جدول ۵ آمده است و ضرایب پایایی متغیرهای پنهان در این پژوهش را به درستی نشان می‌دهد. از آنجایی که پایایی ترکیبی و اشتراکی متغیرها دارای ضرایب مناسب بالای ۰,۷ بوده لذا برازش مناسب مدل اندازه‌گیری تأیید می‌شود.

جدول ۵. نتایج ارزیابی مدل اندازه‌گیری

متغیر	آلفای کرونباخ	بار عاملی	پایایی ترکیبی	پایایی اشتراکی	متوسط واریانس احراز شده	روایی همگرا
کیفیت اطلاعات	۰,۶۳۵	معنادار	۰,۸۴۵۳۷۴	۰,۷۳۲۲۴۹	۰,۷۳۲۲۴۹	تأیید
شهرت شرکت	۰,۶۴۸	معنادار	۰,۸۱۰۳۳	۰,۵۸۸۰۲۲	۰,۵۸۸۰۲۲	تأیید
محل فروش آفلاین شرکت	۰,۶۰۳	معنادار	۰,۸۳۰۷۶	۰,۷۱۱۳۵۶	۰,۷۱۱۳۵۶	تأیید
امنیت معامله	۰,۶۷۲	معنادار	۰,۸۵۷۱۴۷	۰,۷۵۰۳۵	۰,۷۵۰۳۵	تأیید
اندازه شرکت	۰,۷۴۸	معنادار	۰,۸۵۵۴۰۴	۰,۶۶۳۶۴۸	۰,۶۶۳۶۴۸	تأیید
اعتماد	۰,۷۰۵	معنادار	۰,۸۳۲۵۳۷	۰,۶۲۴۲۶۶	۰,۶۲۴۲۶۶	تأیید
تصمیم خرید	۰,۵۸۱	معنادار	۰,۸۲۳۲۰۴	۰,۷۰۰۵۳۵	۰,۷۰۰۵۳۵	تأیید
وفاداری	۰,۸۳۷	معنادار	۰,۹۰۱۶۰۶	۰,۷۵۳۵۶۷	۰,۷۵۳۵۶۷	تأیید
خرید مجدد	۱	معنادار	۱	۱	۱	تأیید
توصیه به دیگران	۱	معنادار	۱	۱	۱	تأیید

تصمیم خرید	۰,۲۸	متوسط
وفاداری	۰,۴۴	قوی
خرید مجدد	۰,۵۱	قوی
توصیه به دیگران	۰,۳۹	قوی

### - معیار $R^2$ یا ضریب تعیین

معیار  $R^2$  یا ضریب تعیین، معیاری است که نشان از قدرت پیش‌بینی تأثیر یک متغیر برونزا از یک متغیر درونزا دارد، به عبارت بهتر بیانگر این است که چه مقدار از واریانس متغیرهای درون‌زا توسط متغیرهای برون‌زا تبیین شده است. سه مقدار ۰,۱۹، ۰,۳۳ و ۰,۶۷ به عنوان مقدار ملاک برای مقادیر ضعیف، متوسط و قوی  $R^2$  در نظر گرفته می‌شود. همانطور که در جدول ۶ مشاهده می‌شود متغیرهای درونزای این پژوهش دارای ضریب تعیین مناسبی هستند:

جدول ۶. ضرایب  $R^2$  مربوط به متغیرهای درونزا

نام مؤلفه	ضریب	نتیجه
اعتماد	۰,۳۳	متوسط

### - معیار افزونگی

این معیار از حاصل ضرب مقادیر اشتراکی سازه‌ها در مقادیر مربوط به ضریب تعیین آنها به دست می‌آید و نشانگر مقدار تغییرپذیری شاخصهای یک سازه درون‌زا است که از یک یا چند سازه برون‌زا تأثیر می‌پذیرد.

$$GoF = \sqrt{\text{Communalities} \times R^2}$$

$$= \sqrt{0,702 \times 0,39} = 0,541$$

با توجه به سه مقدار ۰,۰۱، ۰,۲۵، و ۰,۳۶ به عنوان مقادیر ضعیف، متوسط و قوی برای GOF، حاصل شدن ۰,۵۴۱ برای GOF، نشان از برازش کلی قوی مدل دارد. جدول ۷ نشان دهنده خلاصه نتایج مدل ساختاری پژوهش برای آزمون فرضیه‌های مربوط به متغیرها در حالت اثرات

مستقیم می‌باشد. با توجه به مقادیر ارائه شده در این جدول، زمانی که عدد معناداری بیش از مقدار (۱/۹۶) و یا کمتر از مقدار (۱/۹۶-) حاصل شود، و همچنین سطح معنی داری کمتر یا مساوی ۰/۰۵ باشد، در این صورت فرضیه فوق تأیید می‌گردد.

جدول ۷. نتایج بررسی فرضیه‌ها

روابط	مسیر	ضریب مسیر	آماره t	نتیجه
۱	کیفیت اطلاعات	اعتماد	۳,۳۸۳	تأیید
۲	شهرت شرکت	اعتماد	۱۳,۵۴۹	تأیید
۳	محل فروش آفلاین شرکت	اعتماد	۲۷,۲۹۷	تأیید
۴	امنیت معامله	اعتماد	۶۱,۱۹۴	تأیید
	اندازه شرکت	اعتماد	۱۵,۲۸۸	تأیید
۵	عوامل سازمانی	اعتماد	۹,۰۴۸	تأیید
۶	اعتماد	تصمیم خرید	۳,۱۲۳	تأیید
۷	تصمیم خرید	وفاداری	۳۲,۸۹	تأیید
۸	تصمیم خرید	توصیه به دیگران	۸,۸۰۱	تأیید
۹	تصمیم خرید	خرید مجدد	۱۳,۲۵۱	تأیید

بر اساس نتایج جدول ۷، کیفیت اطلاعات تأثیر معنی‌داری بر اعتماد مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد. همچنین شهرت شرکت، محل فروش آفلاین شرکت و امنیت مبادله تأثیر معنی‌داری بر اعتماد مصرف‌کننده به خرید پوشاک از طریق شبکه‌های اجتماعی دارد. و در نهایت سازه عوامل سازمانی بر اعتماد تأثیر مثبت و معنی‌داری دارد. همچنین تصمیم خرید تأثیر معناداری بر خرید مجدد، وفاداری به سایت‌های شبکه اجتماعی پوشاک و توصیه به دیگران برای خرید از این سایت‌ها دارد.

#### ۴-۳ بررسی نقش متغیر اعتماد در ارتباط بین عوامل

##### سازمانی و تصمیم خرید

برای بررسی نقش متغیر اعتماد در ارتباط بین عوامل سازمانی و تصمیم خرید از آزمون سو بل استفاده شده است که نتایج حاصل از بررسی نقش میانجی اعتماد، در جدول ۷ نمایه شده است.

با توجه به اینکه مقدار بدست آمده برای تمامی متغیرهای میانجی از ۱/۹۶ بیشتر است لذا در سطح اطمینان ۹۵٪ معنادار بودن تأثیر میانجی اعتماد تایید می‌شود. به عبارت دیگر تأثیر میانجی اعتماد در رابطه بین عوامل سازمانی و

تصمیم خرید، تأثیر میانجی اعتماد در رابطه بین کیفیت وب سایت و تصمیم خرید، تأثیر میانجی اعتماد در رابطه بین شهرت شرکت و تصمیم خرید، تأثیر میانجی اعتماد در

رابطه بین محل فروش آفلاین شرکت و تصمیم خرید و تأثیر میانجی اعتماد در رابطه بین امنیت مبادله و تصمیم خرید تایید می‌گردد.

جدول ۸. نتایج آزمون سو بل

مقدار آماره t	رابطه
۴/۲۲	میانجیگری اعتماد در رابطه بین عوامل سازمانی و تصمیم خرید
۳/۸	میانجیگری اعتماد در رابطه بین کیفیت وب سایت و تصمیم خرید
۱۸	میانجیگری اعتماد در رابطه بین شهرت شرکت و تصمیم خرید
۳/۶۶	میانجیگری اعتماد در رابطه بین محل فروش آفلاین شرکت و تصمیم خرید
۶	میانجیگری اعتماد در رابطه بین امنیت مبادله و تصمیم خرید

## ۵. نتیجه گیری و پیشنهادات

در سالهای اخیر شبکه‌های اجتماعی در حال تبدیل شدن به بخش جدایی ناپذیر استراتژی کسب و کار هستند و تعاملات شرکت با مشتریان را به شدت تحت تأثیر قرار داده‌اند. عوامل گوناگونی بر خرید افراد از شبکه‌های اجتماعی مؤثر هستند و تحقیقات بسیاری به بررسی این عوامل پرداخته‌اند. اما تاکنون عوامل مربوط به شرکت که در این مقاله به عنوان "عوامل سازمانی" نامیده شده است، مورد بررسی دقیق و عمیق قرار نگرفته است. از این رو، این پژوهش به دنبال شناسایی عوامل سازمانی مؤثر بر رفتار مصرف‌کننده در خرید پوشاک از طریق شبکه‌های اجتماعی است.

در ابتدا با ۱۴ فروشنده پوشاک فعال در شبکه‌های اجتماعی مصاحبه شد و با تحلیل داده‌ها توسط نرم افزار مکس-کیودی‌ای، مدل اولیه عوامل سازمانی مؤثر بر شبکه‌های اجتماعی ارائه گردید. بر اساس نتایج بدست آمده در بخش کیفی، مؤلفه‌های مؤثر بر خرید پوشاک از طریق شبکه‌های اجتماعی عبارتند از: کیفیت اطلاعات، شهرت شرکت، محل فروش آفلاین شرکت و امنیت مبادله. سپس بر اساس نتایج

بخش کیفی، فرضیه‌های تحقیق تدوین شده و پس از آزمون کمی مدل، نتایج فرضیه‌ها مشخص گردید. از دیدگاه عملی، نتایج بدست آمده نشانگر عوامل مؤثر بر افزایش اعتماد است که باعث موفقیت تجارت اجتماعی در تعاملات میان شرکت و مصرف‌کننده می‌گردد. به طور ویژه در این پژوهش عوامل تعیین کننده بالقوه و مهم اعتماد مصرف‌کنندگان در شبکه‌های اجتماعی و نهایتاً تأثیر آنها بر احتمال خرید افراد و شواهد تجربی درباره تأثیر نسبی هریک از این عوامل بر اعتماد و تصمیم خرید مصرف‌کنندگان ارائه شده است. بنابراین نتایج بدست آمده به شرکت‌های فروش پوشاک در شبکه‌های اجتماعی توصیه می‌کند تا مکانیزم‌های اعتماد سازی را با تمرکز بر عوامل تعیین شده در این پژوهش، بکار گیرند.

بر اساس نتایج بدست آمده شهرت شرکت تأثیر به سزایی بر اعتماد افراد به شبکه‌های اجتماعی آن شرکت و به دنبال آن تمایل آنها به خرید پوشاک از طریق این شبکه‌ها دارد که این یافته با نتایج تحقیقات [۴۵]، [۲۳]، [۱]، [۳]، [۴۶] همراستا است. بنابراین به شرکت‌های فعال در شبکه‌های اجتماعی پیشنهاد می‌شود، یک حساب کاربری تجاری بسازند و در بخش پروفایل سایت اجتماعی خود، به خوبی

هشتگ‌هایی که مخاطبان، برای دیدن مطالب محبوبشان بیشتر از آنها استفاده می‌کنند هشتگ مناسب خود را انتخاب نمایند.

نتایج تحقیق نشان می‌دهد یکی از عوامل بازدارنده برای خرید پوشاک از شبکه‌های اجتماعی، احساس نگرانی فرد از امنیت معامله است که این نتایج با تحقیقات [۳۱]، [۲۱]، [۶]، [۳۴] و [۴۷] هم راستا است. در این راستا پیشنهاد می‌شود با برقراری امکان شکایت، امکان مرجوع نمودن و تعویض محصول، گارانتی محصول، امکان پرداخت وجه در هنگام تحویل محصول و همچنین استفاده از درگاه مطمئن، احساس نگرانی را در افراد کاهش دهند. همچنین با ارائه کد رهگیری (صادر شده توسط اداره پست) به مشتریان، این امکان را برای مشتریان فراهم نمایند تا از وضعیت سفارش خود مطلع گردند. همچنین پیشنهاد می‌شود در قسمت پروفایل شبکه اجتماعی خود، اطلاعات واقعی درباره کسب و کار و ... بنویسند تا زمینه اعتماد مشتری فراهم گردد. از سوی دیگر پیشنهاد می‌شود با انتشار شماره تلفن و یا اکانتی در پروفایل سایت شبکه اجتماعی خود همیشه پاسخگوی مشتریان باشند تا اعتماد آنها را بدست آورند.

نتایج تحقیق نشان می‌دهد محل فروش آفلاین شرکت تأثیر زیادی بر تصمیم خرید پوشاک توسط افراد از طریق شبکه‌های اجتماعی دارند، که در این راستا پیشنهاد می‌شود در سایت اجتماعی شرکت، آدرس و شماره تلفن درج گردد؛ زیرا برخورداری شرکت از محل فروش آفلاین، آدرس و شماره تلفن، باعث اعتماد افراد می‌شود. حتی در صورتی که شرکت قادر به برخورداری از محل فروش آفلاین به طور دائمی نمی‌باشد، پیشنهاد می‌شود در دوره‌های زمانی مختلف اقدام به فروش آفلاین نماید.

با توجه به اینکه در این پژوهش تأثیر عوامل سازمانی بر تصمیم خرید افراد بررسی شده است، از این رو، پیشنهاد می‌شود در پژوهش‌های آتی تأثیر این عوامل بر سایر ابعاد رفتار خرید مصرف‌کننده نظیر فرایند تصمیم‌گیری خرید نظیر تشخیص مساله، جستجوی اطلاعات، ارزیابی گزینه‌ها و رفتار پس از خرید مورد بررسی قرار گیرد. همچنین با توجه به اینکه این پژوهش در حوزه B2C انجام شده است، پیشنهاد می‌شود مدل عوامل سازمانی در بازار B2B نیز

کسب و کار خود را توصیف نمایند. با این کار نه تنها تأثیرگذاری بیشتری بر کاربران خواهند داشت بلکه می‌توانند از ابزارهای تحلیلی شبکه‌های اجتماعی نیز برای کسب اطلاعات دقیق و بیشتر درباره مخاطبان خود استفاده نمایند. همچنین پیشنهاد می‌شود شرکت‌ها با انتشار محتوای مناسب و تمرکز بر فروش پوشاک، خود را متخصص فروش این محصول معرفی نمایند و با این کار صلاحیت خود را در زمینه پوشاک به مشتریان نشان دهند. نتایج نشان داد کیفیت اطلاعات شامل جذابیت اطلاعات، ارائه اطلاعات قابل اطمینان، مورد نیاز و به روز، تأثیر به سزایی بر تمایل افراد به خرید از طریق شبکه‌های اجتماعی دارد که این نتایج با تحقیقات [۱۶]، [۳۱]، [۵]، [۲۲] و [۲۵] هم راستا است. با توجه به اینکه در سایت‌های شبکه‌های اجتماعی، اطلاعات اولین چیزی است که مشتری با آن برخورد می‌نماید، قطعاً جذابیت این اطلاعات و کیفیت آنها اهمیت بسیار زیادی در جذب کاربران شبکه‌های اجتماعی دارد. بنابراین پیشنهاد می‌شود با ارائه اطلاعات جذاب، مرتبط، به روز و قابل اطمینان درباره محصولات، همچنین با تولید محتوا درباره موضوعات جذاب مانند ویژگی‌های خوب شرکت، نوآوری‌ها و موفقیت‌هایش و پشت صحنه کسب و کار و کارکنان، بر اعتماد و رفتار خرید افراد تأثیر گذارند.

در راستای تولید محتوا، پیشنهاد می‌شود کسب و کارهای فروش پوشاک، برای تولید محتوای سایت شبکه اجتماعی-شان، استراتژی و برنامه داشته باشند؛ بدین معنا که باید زمان‌بندی ارسال محتوا و همچنین ماهیت محتوایی که قرار است در زمان‌های تعیین شده منتشر گردد، مشخص شود. همچنین پیشنهاد می‌شود از هشتگ‌های مناسب برای دسته‌بندی پیام‌ها و محتواها استفاده گردد. هشتگ‌ها می‌توانند به فراگیر شدن محتوای کسب و کار و افزایش فروش در شبکه‌های اجتماعی کمک بسیار زیادی کنند. در این راستا شرکت‌ها می‌توانند از هشتگ‌های پرکاربرد استفاده نمایند. همچنین پیشنهاد می‌شود شرکت‌ها برای خود، یک هشتگ مخصوص داشته باشند و در بخش پروفایل شبکه‌های اجتماعی خود نیز درج نمایند. برای این کار پیشنهاد می‌شود با بررسی هشتگ‌های رقبا و همچنین بررسی

*Commerce Research and Applications*,  
۱۲، ۲۰۱۳، ۳۳۷-۳۴۶

[۸] Tien, D. H., Rivas, A. A. & Liao, Y. Examining the influence of customer-to-customer electronic word-of-mouth on purchase intention in social networking sites, *Asia Pacific Management Review*, ۲۰۱۸، ۱-۱۲.

[۹] Doha, A., Ghasemaghaei, M. Social bundling: A novel method to enhance consumers' intention to purchase online bundles, *Journal of Retailing and Consumer Services*, ۳۵، ۲۰۱۷، ۱۰۶-۱۱۷.

[۱۰] Wang Tien. 'Social identity dimensions and consumer behavior in social media', *Asia Pacific Management Review*, ۲۲، ۲۰۱۷، ۴۵-۵۱

[۱۱] Ahmad, Shimi Naurin; Laroche, Michel. 'Analyzing electronic word of mouth: A social commerce construct', *International Journal of Information Management*, ۳۷، ۳، ۲۰۱۷، ۲۰۲-۲۱۳

[۱۲] Swani, K.; Milne, G.; Brown, B.; Assaf, G. & Donthud, N. What messages to post? Evaluating the popularity of social media communications in business versus consumer markets, *Industrial Marketing Management* ۶۲، ۲۰۱۷، ۷۷-۸۷.

[۱۳] Chun Jung Won, Lee Moon J. 'Increasing individuals' involvement and WOM intention on Social Networking Sites: Content matters', *Computers in Human Behavior*, ۶۰، ۲۰۱۶، ۲۲۳-۲۳۲

[۱۴] Williams, M. Social commerce and the mobile platform: Payment and security perceptions of potential users, *Computers in Human Behavior*, ۲۰۱۸

[۱۵] Hilverda, F., Kuttschreuter, M., & Giebels, E. Social media mediated interaction with peers, experts and anonymous authors: Conversation partner

مورد بررسی قرار گیرد. مهمترین محدودیتی که محققان با آن مواجه بوند، نگرانی فروشندگان از افشای روش‌ها و راهبردهای بازاریابی آنها و مقاومت در برابر ارائه اطلاعات بود.

## مراجع

[۱] Kim, Sanghyun; Park, Hyunsun. 'Effects of various characteristics of social commerce (s-commerce) on consumers' trust and trust performance', *International Journal of Information Management*, ۳۳، ۲۰۱۳، ۳۱۸-۳۳۲

[۲] Chen, Jun; Shen, Xiao-Liang. 'Consumers' decisions in social commerce context: An empirical investigation', *Decision Support Systems*, ۷۹، ۲۰۱۵، ۵۵-۶۴

[۳] Kim, S. & Noh, M. 'Determinants Influencing Consumers Trust and Trust Performance of Social Commerce and Moderating Effect of Experience'. *Information Technology Journal*, ۱، ۱۰، ۲۰۱۲، ۱۳۹۶ - ۱۳۸۰.

[۴] Fibre ۲ Fashion. Social Shopping - where social media meets clothing retail, ۲۰۱۸، <http://Fibre۲Fashion.com>

[۵] Chen, Ch., & Chang, Y. What drives purchase intention on Airbnb? Perspectives of consumer reviews, information quality, and media richness, Telematics and Informatics, ۳۵، ۲۰۱۸، ۱۵۱۲-۱۵۲۳

[۶] Doha, A., Elnahla, N., Mcshane, L. Social commerce as social networking, *Journal of Retailing and Consumer Services*, ۴۷، ۲۰۱۹، ۳۰۷-۳۲۱.

[۷] Wang Jyun-Cheng, Chang Ching-Hui. 'How online social ties and product-related risks influence purchase intentions: A Facebook experiment', *Electronic*

- purchase intentions, *Journal of Business Research*, ۷۱, ۲۰۱۷, ۱۳۳-۱۴۱.
- [۲۴] Fu ,S., Yan,Q., Feng,G. Who will attract you? Similarity effect among users on online purchase intention of movie tickets in the social shopping context, *International Journal of Information Management*, ۴۰, ۲۰۱۸, ۸۸-۱۰۲.
- [۲۵] Erkan, I.& Evans, Ch. ,The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption, *Computers in Human Behavior*, ۶۱, ۲۰۱۶, ۴۷-۵۵
- [۲۶] Chakraborty, L.; Bagchi-Sen, R. 'Online shopping intention in the context of data breach in online retail stores:An examination of older and younger adults', *Decision Support Systems*, ۸۳, ۲۰۱۶, ۴۷ - ۵۶
- [۲۷] Xiang, L.; Zheng, X.; Lee, M.; Zhao, D. 'Exploring consumers' impulse buying behavior on social commerce platform: The role of parasocial interaction'. *International Journal of Information Management*, ۳۶, ۲۰۱۶, ۳۳۳-۳۴۷
- [۲۸] Lee, Jieun; Hong, Ilyoo B. 'Predicting positive user responses to social media advertising: The roles of emotional appeal, informativeness, and creativity', *International Journal of Information Management*, ۳۶, ۲۰۱۶, ۳۶۰-۳۷۳
- [۲۹] Amaro, S. ; Duarte, P. 'An integrative model of consumers' intentions to purchase travel Online', *Tourism Management* , ۴۶, ۲۰۱۵, ۶۴-۷۹
- [۳۰] Hajli Nick. 'Social commerce constructs and consumer's intention to buy'. *International Journal of Information Management*, ۳۵ , ۲۰۱۵, ۱۸۳-۱۹۱
- and message framing effects on risk, *Food Quality and Preference*, ۵۶, ۲۰۱۸, ۱۰۷-۱۱۸
- [۱۶] Lal Prerna. 'Analyzing determinants influencing an individual's intention to use social commerce website', *Future Business Journal*, ۳, ۲۰۱۷, ۷۰-۸۵
- [۱۷] Wang, Yichuan; Yu; Chiahui. 'Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning', *International Journal of Information Management*, ۳۷(۳), ۲۰۱۷, ۱۷۹-۱۸۹
- [۱۸] Ahmad, Shimi Naurin; Laroche, Michel. 'Analyzing electronic word of mouth: A social commerce construct', *International Journal of Information Management*, ۳۷( ۳), ۲۰۱۷, ۲۰۲-۲۱۳
- [۱۹] Farivar, S., Turel, O., & Yuan, Y. Skewing users' rational risk considerations in social commerce: An empirical examination of the role of social identification, *Information & Management*, ۵۵, ۲۰۱۸, ۱۰۳۸-۱۰۴۸.
- [۲۰] Hou, R.; Wu, J.; Du, S. Customer social network affects marketing strategy: A simulation analysis based on competitive diffusion model, *Physica* , ۴۶۹ ,۲۰۱۸, ۶۴۴-۶۵۳.
- [۲۱] Kempt, S. Social media use jumps in Q۱ despite privacy fears. Retrieved from: <http://www.Wearsocial.com/۲۰۱۸/۲۴/۰۴>
- [۲۲] Ponte, E.; Carvajal, E.; Escobar-Rodríguez , T. Influence of trust and perceived value on the intention to purchase travel online: Integrating the effects of assurance on trust antecedents , *Tourism Management*, ۴۷ , ۲۰۱۵, ۲۸۶-۳۰۲
- [۲۳] Hajli, N, Sims, j. H. Zadeh, A. & Odile, M. A social commerce investigation of the role of trust in a social networking site on

- [۳۸] Agag, G. & El-Masry, A. 'Understanding consumer intention to participate in online travel community and effects on consumer intention to purchase travel online and WOM: An integration of innovation diffusion theory and TAM with trust'. *Computers in Human Behavior*, ۶۰, ۲۰۱۶, ۲۰۱۶, ۹۷-۱۱۱
- [۳۹] Yahia, I.B., Al-Neama, N., & Kerbache, L. Investigating the drivers for social commerce in social media platforms: Importance of trust, social support and the platform perceived usage, *Retailing and Consumer Services*, ۴۱, ۲۰۱۸, ۱۱-۱۹
- [۴۰] کرسول. جان، کلارک، ویکی. *روش‌های پژوهش ترکیبی*، علیرضا کیامنش، چاپ دوم، تهران: آبیژ، ۱۳۹۴.
- [۴۱] اووه. فلیک. *درآمدی بر تحقیق کیفی*، جلیلی. هادی، نشر نی، ۱۳۸۸.
- [۴۲] استراس. آنسلم، کوربین. جولیت. *اصول روش تحقیق کیفی (نظریه‌مبنایی؛ رویه‌ها و شیوه‌ها)*، رحمت‌پور. رحمت‌الله، پژوهشگاه علوم انسانی و مطالعات فرهنگی، ۱۳۸۵.
- [۴۳] ذکائی. م، نظریه و پژوهش در روش‌های کیفی، فصلنامه علوم اجتماعی، دانشگاه علامه طباطبائی، شماره ۱۷، ۱۳۸۱.
- [۴۴] سفیری. خدیجه. *روش تحقیق کیفی*، نشر پیام نور، ۱۳۸۷.
- [۴۵] Tan, O. How Social Media Contributed to the Rise of Fast Fashion, ۲۰۱۷, <https://www.adweek.com>
- [۴۶] Kim, E., Hong, T. 'Segmentating Customers in Online Stores from Factors that Affect the Customer's Intention to Purchase'. *International Conference on Information Society*, ۲۰۱۰, ۳۸۳-۳۸۸.
- [۴۷] Nguyen, D. Social media marketing strategy, Bachelor Thesis Degree Program in International Business, HAAGA-HELIA university of applied sciences, ۲۰۱۷
- [۳۱] Esmaeili, Leila; Mardani, Shahla; Mutallebi, Muhamad; Hashemi Golpayegani, Seyyed Alireza. 'Studying the Affecting Factors on Trust in Social Commerce', *International Journal of advanced studies in Computer Science and Engineering IJASCSE*, ۴, ۶, ۲۰۱۵
- [۳۲] Lin, Y., Hsu, C., Chen, M., & Fang, Ch. New gratifications for social word-of-mouth spread via mobile SNSs: Uses and gratifications approach with a perspective of media technology, *Telematics and Informatics*, ۳۴, ۴, ۲۰۱۸, ۳۸۲-۳۹۷
- [۳۳] Huang, Zhao; Benyoucef Morad. 'From e-commerce to social commerce: A close look at design features'. *Electronic Commerce Research and Applications*, ۱۲, ۲۰۱۳, ۲۴۶-۲۵۹.
- [۳۴] Fang, Hui; Zhang, Jie; Şensoy, Murat; Magnenat-Thalmann, Nadia. 'Reputation mechanism for e-commerce in virtual reality environments', *Electronic Commerce Research and Applications*, ۱۳, ۶, ۲۰۱۴, ۴۰۹-۴۲۲
- [۳۵] Asshidin Nor Hazlin Nor, Abidin Nurazariah, Borhan Hafizzah Bashira. 'Consumer Attitude and Uniqueness towards International Products'. *Procedia Economics and Finance*, ۳۵, ۲۰۱۶, Pages ۶۳۲-۶۳۸
- [۳۶] Park, J., Ahn, J., Thavisay, T., & Ren, T. Examining the role of anxiety and social influence in multi-benefits of mobile payment service, *Journal of Retailing and Consumer Services*, ۴۷, ۲۰۱۹, ۱۴۰-۱۴۹.
- [۳۷] Hwang, In Jeong; Lee, Bong Gyou; Kim, Ki Youn. 'Information Asymmetry, Social Networking Site Word of Mouth, and Mobility Effects on Social Commerce in Korea'. *Cyberpsychology, Behavior, and Social Networking*, ۱۷, ۲, ۲۰۱۳

نقش آفرینی عوامل سازمانی بر رفتار خرید مصرف‌کننده در شبکه‌های اجتماعی





# ارائه یک روش ترکیبی از ANP و DEMATEL جهت ارزیابی سیستم‌های ERP

امیر امینی\* علیرضا علی‌نژاد\*\*

\*دانش‌آموخته کارشناسی ارشد مهندسی صنایع، دانشکده مهندسی صنایع، دانشگاه صنعتی ارومیه

\*\*دانشیار، دانشکده مهندسی صنایع و مکانیک، دانشگاه آزاد اسلامی، واحد قزوین

تاریخ دریافت: ۱۳۹۹/۰۳/۳۱ تاریخ پذیرش: ۱۳۹۹/۰۹/۱۲

نوع مقاله: پژوهشی

## چکیده

این پژوهش با استفاده از روش تحلیل پوششی داده‌ها به ارزیابی عملکرد سیستم‌های برنامه‌ریزی منابع سازمان ۱۸ نمونه شرکت تولیدی پرداخته است تا مشخص شود آیا اهداف تعریف شده برای سیستم‌های ERP توانسته‌اند عملکرد بعد از پیاده‌سازی این سیستم را تحت تاثیر قرار دهند. با توجه به شناسایی مؤلفه‌های مؤثر در پیاده‌سازی برنامه‌ریزی منابع سازمان با استفاده از مرور تحقیقات پیشین، مصاحبه و نظرات خبرگان شاخص‌های ارزیابی عملکرد این سیستم شناسایی شدند. سپس مهمترین شاخص‌های ورودی با استفاده از نظر خبرگان به روش دیمتل فازی و شاخص‌های خروجی به روش تحلیل شبکه‌ای رتبه‌بندی گردیدند. در این رتبه‌بندی شاخص‌های زمان صرف شده جهت اجرا، زیر ساخت‌های اجرا، آموزش و اقدامات حمایتی کاربر شاخص‌های برتر ورودی، و سه شاخص افزایش بهره‌وری، مدیریت صحیح منابع و میزان رضایت‌مندی کاربر نیز به عنوان شاخص‌های برتر خروجی انتخاب گردیدند. با استفاده از شاخص‌های منتخب به ارزیابی عملکرد سیستم برنامه‌ریزی منابع سازمان شرکت‌های منتخب پرداخته شده است. نتایج این تحقیق در شناسایی نقاط قوت و ضعف شرکت‌ها در مقایسه با بهترین شرکت شناخته شده و هر چه بهتر ساختن سیستم ERP خود مفید واقع خواهد شد.

**واژگان کلیدی:** برنامه‌ریزی منابع سازمان، دیمتل، فرایند تحلیل شبکه‌ای، تحلیل پوششی داده‌ها، ارزیابی کارایی.

## ۱. مقدمه

اختیار استفاده‌کنندگان آن قرار دهند، از ابزارهای مفیدی هستند که سازمان‌ها برای افزایش قابلیت‌های خود، بهبود عملکرد، تصمیم‌گیری بهتر و دستیابی به مزیت رقابتی از آن استفاده می‌کنند [۱]. شاید یکی از جامع‌ترین سیستم‌های اطلاعاتی موجود که در سال‌های اخیر به طور گسترده‌ای مورد استفاده قرار گرفته، سیستم برنامه‌ریزی منابع سازمان باشد. ظرفیت<sup>۱</sup> ERP در یکپارچه سازی فرآیندهای سازمان و اطلاعات حوزه‌های عملکردهای مختلف از طریق یک پایگاه داده متمرکز باعث گردیده است که

قرن بیست و یکم، صده اطلاعات و انفجار آن نام‌گذاری شده و این موضوع اهمیت، تعیین‌کنندگی و مدیریت مناسب اطلاعات را در همه عرصه‌های زندگی بشر روشن می‌سازد. یکی از حوزه‌هایی که ارزیابی عملکرد برای آن لازم و ضروری به نظر می‌رسد، مدیریت فناوری و نوآوری در بنگاه‌هاست. سیستم‌های اطلاعاتی کارا و یکپارچه که بتوانند همه فعالیت‌ها و وظایف موجود در یک سازمان را تحت پوشش قرار داده و اطلاعات لازم و ضروری را به‌موقع در

<sup>۱</sup> Enterprise Resource Planning (ERP)

نویسنده مسئول: امیر امینی، پست الکترونیکی: A.Amini@ine.uut.ac.ir

استفاده از روش تحلیل پوششی داده‌ها، کارایی این سیستم مورد ارزیابی قرار گرفته است. نتایج تحقیق حاضر می‌تواند در شناسایی نقاط ضعف و قوت شرکت‌ها و بهبود سیستم ERP آن‌ها مفید واقع شود.

ساختار پژوهش حاضر به شرح زیر است. پیشینه نظری پژوهش در بخش دوم مورد بررسی قرار می‌گیرد. بخش سوم روش و فرایند اجرای پژوهش را تبیین می‌نماید. یافته‌های پژوهش در بخش چهارم مورد تجزیه و تحلیل قرار می‌گیرند. در انتها نتایج، محدودیت‌ها و پیشنهاد برای پژوهش‌های آتی در بخش پایانی مطرح خواهند شد.

## ۲. ادبیات پژوهش

براساس تعریف عمومی سیستم برنامه‌ریزی منابع سازمان یک بسته نرم‌افزاری قابل تنظیم و از پیش استاندارد شده تجاری است که هدف آن یکپارچگی اطلاعات و جریان اطلاعات بین کلیه بخش‌های سازمان از جمله مالی، حسابداری، منابع انسانی، زنجیره عرضه و مدیریت مشتریان با رویکرد مشتری‌گرایی و پاسخ به بازار می‌باشد. برنامه‌ریزی منابع سازمان یک پایگاه داده، یک برنامه کاربردی و یک واسط یکپارچه در تمامی سازمان است [۷].

در ادبیات فناوری اطلاعات و ارتباطات، این اصطلاح به مجموعه‌های نرم افزاری مستقل ولی مرتبط با یکدیگر اطلاق می‌شود که برای دستیابی به مقاصد خاصی طراحی شده‌اند. امکان تبادل سریع اطلاعات بین این نرم افزارها و همسانی محیط کاری در آنها از مهمترین ویژگی‌های این مجموعه‌هاست. از سوی دیگر رابط کاربری در کلیه آنها مشابه یکدیگر است [۸]. یکپارچگی سیستم ERP به معنای یکپارچگی از لحاظ امکان افزودن ماژول‌های جدید، یکپارچگی از لحاظ پوشش دادن کلیه فرآیندهای مورد نیاز در یک بخش سازمان، یکپارچگی از لحاظ عدم تکرار برخی از فرآیندها و یا حتی بخش‌هایی از یک فرآیند و یکپارچگی از لحاظ اطلاعاتی و عدم تکرار داده‌ها در کل سیستم می‌باشد. به طور کلی یکپارچگی به مفهوم پوشش کامل نرم‌افزار بر روی فرآیندهایی است که نرم افزار برای آنها پیاده‌سازی می‌شود. به عنوان مثال با پیاده‌سازی ماژول حسابداری و مالی، تمامی فرآیندهایی که در واحد حسابداری و مالی سازمان تعریف شده‌اند، تحت پوشش قرار می‌گیرند [۹].

سیر تکاملی سیستم برنامه‌ریزی منابع سازمان از یک سو متأثر از تکامل سخت‌افزارها و نرم‌افزارهای رایانه‌ای و از سوی دیگر متأثر از تحولات صورت گرفته در زمینه مدیریت سازمان‌ها و محیط رقابتی حاکم بر آن‌هاست. برنامه‌ریزی منابع سازمان به منظور غلبه بر مشکلات سیستم‌های عملیاتی موجود در سازمان‌ها که از اوایل دهه ۱۹۶۰ میلادی توسعه پیدا نموده بودند، بوجود آمد [۱۰، ۱۱ و ۱۲].

قرن صاحب‌نظران این سیستم‌ها را به عنوان نیاز موفقیت در قرن بیست و یک معرفی کنند [۲].

ERP مجموعه‌ای منسجم از برنامه‌هایی است که برای پشتیبانی از فعالیت‌های اصلی سازمان نظیر تولید، مالی و حسابداری، فروش و بازاریابی و منابع انسانی به کار گرفته می‌شود [۳]. پیاده‌سازی سیستم‌های ERP زمان و منابع قابل توجهی را از سازمان می‌گیرد. نمونه‌های زیادی از شکست پروژه‌های برنامه‌ریزی منابع سازمان با وجود سرمایه‌گذاری بالا گزارش شده است [۴].

ارائه‌دهندگان سیستم‌های برنامه‌ریزی منابع سازمان ادعا می‌کنند که محصول‌شان بارها آزمایش شده و تحت تجربیات فراوان ایجاد گردیده است و این مهم آن‌ها را قادر می‌سازد که راه‌حل‌های فوق‌العاده‌ای را برای بخش‌های مختلف صنعت و خدمات ارائه کنند. این واقعیت در بسیاری از سازمان‌ها ملموس است، اما تجارب نشان می‌دهد این محصولات در برخی از سازمان‌ها نتوانسته‌اند آن‌گونه که باید مفید و کارگشا باشند، بنابراین پرداختن به عوامل موفقیت سیستم‌های ERP در سازمان‌های مختلف از چالش‌های مهم محققان از زمان پیدایش ERP بوده است [۵].

## ۱-۱ نوآوری پژوهش

صنایع تولیدی و خدماتی در ایران در صدر توجه برای بکارگیری فناوری‌های نوین بوده‌اند. در سال‌های اخیر با توجه به پیاده‌سازی نمونه‌های موفق و ناموفق از پروژه‌های برنامه‌ریزی منابع سازمان در شرکت‌های ایرانی مسأله شناخت عوامل موفقیت این پروژه‌ها در کانون توجه قرار گرفته است [۶]. در پژوهش‌های انجام شده اغلب به بررسی چگونگی استقرار مدیریت منابع سازمانی و دلایل شکست پروژه‌های اجرای آن پرداخته شده و برخی پژوهش‌ها نیز مباحث نرم‌افزاری ERP را مطرح نموده‌اند و بررسی مدیریت چنین سیستمی و ارزیابی الگوهای مناسب جهت عملکرد بهینه آن کمتر مورد توجه قرار گرفته است. بنابراین به منظور پوشش شکاف تحقیقاتی مشاهده شده نیاز است این سیستم‌ها در واحدهای صنعتی و تولیدی بصورت دقیق بررسی گردند و سیاست‌های بهبود اجرای آن‌ها معرفی و ارائه شود. از این رو پژوهش حاضر به دنبال بررسی کارایی استفاده از برنامه‌ریزی منابع سازمان در شرکت‌های تولیدی بر اساس مدل‌های تصمیم‌گیری و تحلیل پوششی داده‌ها است تا بتواند علل و چالش‌های عدم موفقیت اجرای برنامه‌ریزی سازمان را شناسایی نماید. این مهم با بکارگیری روش‌های تصمیم‌گیری چند معیاره جهت شناسایی معیارهای ارزیابی این سیستم در سازمان انجام خواهد شد. بدین منظور مهمترین معیارهای موجود در منابع کتابخانه‌ای و مطالعه موردی شناسایی و وزن‌دهی شده سپس با

بنابراین، تضمین موفقیت بعد از پیاده‌سازی به موضوع تحت تمرکز این تحقیق ERP تبدیل شده است. ساموئلنیکو و بریسون در پژوهش خود تحت عنوان استفاده از تحلیل پوششی داده‌ها (DEA) برای نظارت بر عملکرد مبتنی بر بهره‌وری سازمان بازده محور بیان داشتند که ماهیت رقابتی محیط کسب و کار نیازمند سازماندهی مبتنی بر بهره‌وری است که از میزان نسبی آن در کارایی و کارایی نسبت به رقبا برخوردار باشد [۱۹]. چانگ<sup>۴</sup> و همکاران بیان داشتند که پس از وقوع رسوایی‌های مالی متعدد در سراسر جهان، اهمیت مسائل مرتبط از جمله کنترل داخلی و امنیت اطلاعات تا حد زیادی افزایش یافته است. این مطالعه یک چارچوب کنترل داخلی ایجاد می‌کند که قابل اعمال در سیستم برنامه‌ریزی منابع سازمانی است [۲۰]. سودهامان و تنگاول<sup>۵</sup> در پژوهش خود تحت عنوان تجزیه و تحلیل بهره‌وری از پروژه‌های ERP بیان داشتند که برنامه‌ریزی منابع سازمانی حساس به تغییرات در محیط کسب‌وکار هستند و سرعت افزایش تغییر در تجارت جهانی، مدیران پروژه ERP را به چالش می‌کشد. سیستم استقرار ERP نشان می‌دهد که موفقیت سیستم ERP به شدت به سختی فرآیندهای کیفیت نرم‌افزار بستگی دارد [۲۱]. لال<sup>۶</sup> و تیاراچاکول تجزیه و تحلیل پوششی داده‌ها (DEA) را در فرایند انتخاب یک سیستم ERP بررسی کرده و عوامل مربوط به شکست و موفقیت پیاده‌سازی را بیان نمودند [۱۶].

به منظور ارائه چهارچوبی جهت انتخاب سیستم ERP، آلاسکاری و همکاران<sup>۷</sup> به این نتیجه رسیدند هم سازمان‌های تولیدی و هم خدماتی جهت بکارگیری سیستم‌های ERP مناسب لازم است معیارهای اجرای موفقیت‌آمیز و کاهش ریسک را منحصر در نظر داشته باشند [۲۲]. رویوو و همکاران<sup>۸</sup>، تأثیر برخی از قابلیت‌های قابل تعمیم ERP از جمله تعاملات، تجزیه و تحلیل‌ها، پورتال وب و افزودنی‌ها را بر بکارگیری و ارزش‌گذاری ERP بررسی کردند. نتایج پژوهش ایشان مفاهیم جدیدی را در تأثیرپذیری مثبت ارزش سرمایه‌گذاری و کسب‌وکار از قابلیت‌های ERP معذب نمود [۲۳]. از تحقیقات داخلی می‌توان به مطالعه فیضی و همکاران با موضوع ارائه رویکردی تلفیقی از QFD، FAHP و VIKOR را به منظور انتخاب مناسب‌ترین سامانه ERP ارائه کردند [۲۴]. در پژوهشی با عنوان "انتخاب سیستم برنامه‌ریزی منابع سازمان با بهره‌گیری از

ERP در واقع نسل جدیدی از سیستم‌های اطلاعاتی است که (MIS<sup>۲</sup>) نسل قدیمی‌تر آن به حساب می‌آید. به طور کلی روند رشد و تغییر سیستم‌های کامپیوتری را می‌توان به سه نسل تقسیم کرد [۱۳].

- سیستم‌های جزیره‌ای [۱۴].
- سیستم‌های استخراج اطلاعات [۱۵].
- سیستم‌های تسهیل‌کننده [۱۶ و ۱۷].

## ۱-۲ چالش‌های پیاده‌سازی سیستم برنامه‌ریزی منابع سازمان

عواملی که باعث بحران در سیستم‌های اطلاعاتی موجود سازمان‌ها شده و در زمره چالش‌های پیاده‌سازی ERP به حساب می‌آیند به شرح ذیل می‌باشند: ۱- فناوری با کمترین احساس نیاز وارد کشور شده است و مورد استفاده ناصحیح قرار گرفته است. ۲- در سازمان‌ها سیستم‌های موجود دستی، مکانیزه شدن فرآیندهای صحیح کاری. اکثر سازمان‌ها و شرکت‌های بزرگ کشور در دهه‌های اخیر فقط توانسته‌اند برای سیستم‌های قدیمی و دستی خود بطور جزیره‌ای نرم‌افزار تهیه و آن‌ها را مکانیزه کنند و در قالب رایانه‌های منفرد و یا تحت شبکه اجرا نمایند. ۳- هزینه‌های اجرای این سیستم و مدت زمان بازگشت سرمایه زیاد است. ۴- ساختار سازمانی اکثر سازمان‌های کشور غیرمنعطف است. ERP یک سیستم فرآیندنگر است و این دیدگاه در تناقض با روند انجام فعالیت‌های اکثر سازمان‌ها و شرکت‌های کشور که بصورت وظیفه‌ای انجام می‌شود، است. ۵- فرآیندهای موجود سازمان با فرآیندهای بهینه ERP تطابق ندارد. ۶- تبدیل داده‌ها یکی از معضلات بزرگ در پیاده‌سازی ERP است. ۷- هزینه‌های جانبی تحمیلی بر سازمان هنگام اجرای ERP سنگین است. ۸- کارکنان، عمده‌ترین عوامل ناکامی اجرای پروژه ERP هستند.

## ۲-۲ پیشینه پژوهش

ژو<sup>۳</sup> و همکاران بیان داشتند که سیستم‌های برنامه‌ریزی منابع سازمانی (ERP) در سطح جهان پیاده‌سازی شده‌اند و در طول دهه گذشته پیاده‌سازی آن‌ها به شکل گسترده تحت مطالعه قرار گرفته است [۱۸]. باوجود این، بسیاری از سازمان‌ها هنوز در حال تقلا زدن برای کسب مزایایی از سیستم‌های ERP پیاده‌سازی شده هستند.

<sup>۱</sup> Lall ana Teyarachakul

<sup>۲</sup> Alaskari et al.

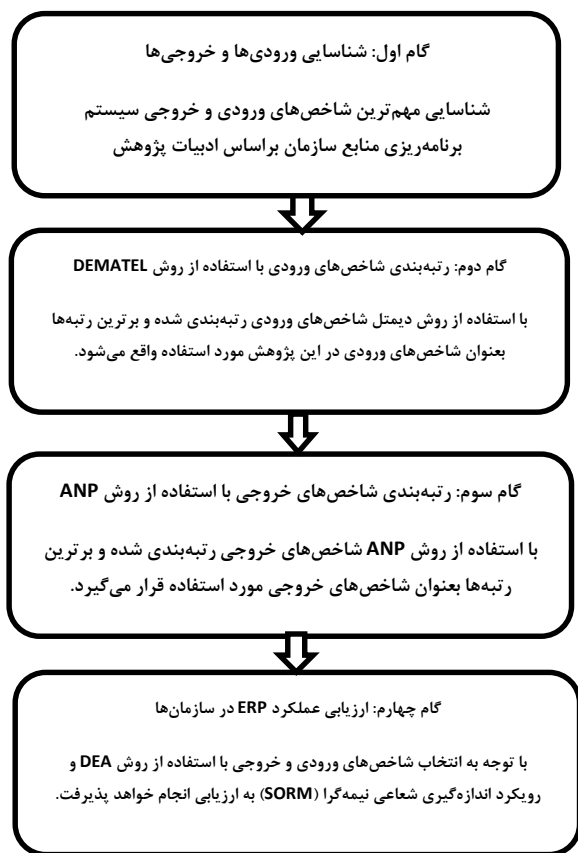
<sup>۳</sup> Ruivoet al.

<sup>۴</sup> Management Information Systems

<sup>۵</sup> Zhu et al.

<sup>۶</sup> Chang

<sup>۷</sup> Thangavel and Sudhaman



شکل ۱. روش انجام تجزیه و تحلیل پژوهش

### ۱-۳ تکنیک دیمتل فازی

تکنیک دیمتل برای اولین بار بین سال‌های ۱۹۷۲ تا ۱۹۷۶ به عنوان روشی برای حل مسائل پیچیده مطرح شد. این تکنیک بر اساس تئوری گراف ساخته شده و مبتنی بر گراف‌های جهت‌داری است که به کمک آن‌ها می‌تواند با تفکیک عوامل دخیل در مسئله به دو گروه علت و معلول مسائل را به روشی ساده حل کند [۲۷]. از آنجا که تکنیک دیمتل قادر به تصمیم‌گیری در شرایط عدم اطمینان نبود، تکنیک دیمتل فازی ارائه شد. روش دیمتل فازی می‌تواند با استفاده از متغیرهای زبانی فازی، تصمیم‌گیری در شرایط عدم اطمینان محیطی را آسان کند [۲۸]. این تکنیک در زمینه‌های تولید، مدیریت سازمان، سیستم‌های اطلاعات و علوم اجتماعی کاربرد دارد و می‌تواند تمام مشکلات پیشروی سازمان را با به کارگیری تصمیم‌گیری گروهی حل نماید [۲۹]. گام‌های اجرای روش دیمتل فازی به شرح زیر است:

گام اول: طراحی پرسشنامه و ایجاد ماتریس مقایسات زوجی هر خبره. در گام نخست جهت تعیین معیارهایی برای تصمیم‌گیری به تنظیم پرسشنامه پرداخته و در اختیار خبرگان قرار می‌دهیم، بعد از

روش تلفیقی سوآرا و آراس خاکستری"، حسینی دهشیری و عرب، معیارهای قابلیت کارکردی نرم‌افزار، پیاده‌سازی مدیریت پروژه و قابلیت کیفی نرم‌افزار را به‌عنوان بااهمیت‌ترین معیارهای انتخاب سیستم برنامه‌ریزی منابع سازمان معرفی کردند [۲۵]. محقر و همکاران (۱۳۹۵) بهبود روش‌ها و فناوری‌های کلیدی با استفاده از گسترش عملکرد کیفیت، تحلیل پوششی داده‌ها و منطق فازی را مورد بررسی قرار دادند و به این نتیجه رسیدند که رویکرد ارائه شده در مقایسه با روش‌های پیشین دارای مزیت‌های بهره‌گیری از شاخص‌های متعدد کمی و کیفی، رتبه‌بندی ساده‌تر و امکان استفاده از قضاوت‌های انسانی را داراست [۲۶].

پژوهش حاضر درصدد است با شناسایی مؤلفه‌های مؤثر در پیاده‌سازی ERP، مهمترین شاخص‌های ورودی و خروجی تأثیرگذار و تعیین‌کننده در ارزیابی این سیستم‌ها را استخراج و در نمونه آماری مورد مطالعه پیاده‌سازی نماید. این مهم با استفاده از رویکرد پیشنهادی پژوهش که در بخش بعدی تشریح خواهد شد، انجام پذیرفته است.

### ۳. روش‌شناسی پژوهش

این پژوهش به ارزیابی عملکرد سیستم ERP، در ۱۸ شرکت تولیدی که سیستم ERP را پیاده‌سازی کرده‌اند، پرداخته‌است. ۹ مورد از این شرکت‌ها که با معرفی‌شان در این پژوهش موافقت کرده‌اند به شرح ذیل می‌باشند و سایر شرکت‌ها بصورت ناشناس مورد مطالعه واقع شده‌اند: ۱- شرکت فولاد هیربد زرنديه، ۲- شرکت تولیدی و بازرگانی نیلپر، ۳- شرکت تولیدی پارس حیات، ۴- شرکت سیم و کابل ابهر، ۵- شرکت ساینا، ۶- شرکت فولاد پاسارگاد، ۷- شرکت تولیدی مازینور، ۸- شرکت ایران ترانسفو، و ۹- شرکت سپهر منور. روش اجرایی این پژوهش با هدف ارزیابی کارایی پیاده‌سازی سیستم برنامه‌ریزی منابع سازمان طراحی شده است. الگوریتم اجرایی پژوهش حاضر براساس طی مراحل نظام‌مند و مبتنی بر روش تحقیق علمی طراحی شده است. گزینش شاخص‌های ورودی و خروجی توسط خبرگان و از طریق پرسشنامه انجام شده است. علت انتخاب این افراد برخوردار از آن‌ها از تخصص کافی، دانش و تجربیات مناسب در این حوزه است. به منظور جمع‌آوری داده‌ها از مطالعه کتابخانه‌ای و مصاحبه استفاده شده است که یکی از ابزارهای رایج تحقیق و روشی مستقیم برای جمع‌آوری داده‌های پژوهش می‌باشد. همچنین بررسی دقیق و مصاحبه با کارشناسان و خبرگان آشنا با مفاهیم یکپارچه‌سازی منابع سازمان جهت شناسایی شاخص‌ها برای ارزیابی عملکرد سیستم ERP در شرکت‌های تولیدی انجام شده است. هریک از مراحل انجام شده در راستای نیل به هدف تحقیق در شکل

$$[u_{ij}^t] = H_u \times (I - H_u)^{-1} \quad (7)$$

گام پنجم: فازی‌زدایی ماتریس روابط کل فازی. جهت انجام فازی‌زدایی از رابطه (۸) استفاده می‌شود:

$$B = \frac{(a_1 + a_r + 2 \times a_r)}{4} \quad (8)$$

گام ششم: محاسبه جمع سطرها و ستون‌ها. پس از تشکیل ماتریس رابطه کل که قبلاً به روش بدست آوردن آن اشاره شد، جمع سطر-های این ماتریس نشان‌دهنده تاثیر مستقیم و غیرمستقیمی که معیار  $\bar{I}$  بر روی معیارهای دیگر می‌گذارد، می‌باشد و جمع ستون‌های این ماتریس نشان‌دهنده تاثیر مستقیم و غیرمستقیمی است که معیار  $\bar{J}$  از معیارهای دیگر دریافت کرده است.

$$\bar{D} = (\bar{D}_i)_{n \times 1} = \left[ \sum_{j=1}^n \bar{T}_{ij} \right]_{n \times 1} \quad (9)$$

$$\bar{R} = (\bar{R}_i)_{1 \times n} = \left[ \sum_{j=1}^n \bar{T}_{ij} \right]_{1 \times n} \quad (10)$$

گام هفتم: محاسبه  $\bar{D} - \bar{R}$  و  $\bar{D} + \bar{R}$

$(\bar{D}_i + \bar{R}_i)$ : نشان‌دهنده اهمیت شاخص‌ها

$(\bar{D}_i - \bar{R}_i)$ : نشان‌دهنده رابطه بین معیارها

$\bar{D}_i - \bar{R}_i > 0$ : معیار اثرگذار می‌باشد

$\bar{D}_i - \bar{R}_i < 0$ : معیار اثر پذیر می‌باشد

گام هشتم: رسم نمودار روابط علی و معلولی. نمودار رابطه علی و معلولی بیانگر رابطه علت و معلولی بین معیارها است. در این نمودار محور افقی نشان‌دهنده  $\bar{D} + \bar{R}$  و محور عمودی نشان‌دهنده  $\bar{D} - \bar{R}$  می‌باشد. معیارهایی که بالای خط افق قرار گیرند علت‌ها و معیارهایی که پایین خط افق قرار گیرند معلول‌ها را نمایش می‌دهند. گام نهم: محاسبه ارزش آستانه روابط<sup>۹</sup> و رسم نقشه روابط شبکه<sup>۱۰</sup>. در این گام ارزش آستانه که برابر میانگین عناصر ماتریس ارتباط کامل می‌باشد محاسبه گردیده، با توجه به این مقدار تمامی مقادیر ماتریس ارتباط کامل که کوچک‌تر از این عدد باشند را مساوی صفر قرار داده و آن رابطه را علی در نظر نمی‌گیریم. با این روش می‌توان از روابط جزئی صرف‌نظر کرده و شبکه روابط قابل اعتنا را ترسیم کرد. تنها روابطی که مقادیر آنها در ماتریس ارتباط کامل از مقدار آستانه بزرگتر باشد در NRM نمایش داده خواهد شد.

### ۳-۲ روش فرآیند تحلیل شبکه‌ای

از تلفیق فرآیند تحلیل سلسله مراتبی<sup>۱۱</sup> و ماتریس تأثیرات وابستگی متقابل بین معیارها فرآیند تحلیل شبکه‌ای حاصل می‌شود. به کمک

جمع‌آوری پرسشنامه‌ها داده‌ها که به زبان فازی می‌باشد باید به اعداد فازی تبدیل شوند، جهت انجام این عمل از جدول زیر استفاده شده است.

جدول ۱. تبدیل عبارات کلامی به اعداد فازی [۲۹]

میزان تأثیر گذاری عبارات کلامی	اعداد فازی مثلثی
بدون تأثیر	(۰/۰، ۰/۱، ۰/۳)
تأثیر کم	(۰/۰، ۱/۰، ۳/۵)
تأثیر متوسط	(۰/۰، ۳/۰، ۵/۷)
تأثیر زیاد	(۰/۰، ۵/۰، ۷/۹)
تأثیر خیلی زیاد	(۰/۰، ۷/۱، ۹)

گام دوم: تشکیل ماتریس ارتباط مستقیم فازی. روابط بین معیارها بصورت یک ماتریس نمایش داده شده است. چون از دیدگاه چند کارشناس استفاده شده از میانگین ساده نظرات استفاده می‌شود و ماتریس ارتباط مستقیم را تشکیل می‌دهیم. به منظور محاسبه میانگین حسابی نظرات خبرگان از رابطه (۱) استفاده می‌شود:

$$\bar{z} = \frac{\bar{x}^1 \oplus \bar{x}^2 \oplus \bar{x}^3 \oplus \dots \oplus \bar{x}^p}{p} \quad (1)$$

$p$  تعداد خبره‌ها،  $\bar{x}^1$  ماتریس مقایسه زوجی خبره اول،  $\bar{x}^2$  ماتریس مقایسه زوجی خبره دوم،  $\bar{x}^p$  ماتریس مقایسه زوجی خبره  $p$ ام،  $\bar{z}$  عدد فازی مثلثی  $(\bar{l}_{ij}, \bar{m}_{ij}, \bar{u}_{ij})$  است.

گام سوم: نرمالایز کردن ماتریس ارتباط مستقیم فازی. جهت نرمال‌سازی ماتریس ارتباط مستقیم فازی که در گام قبل بدست آمد، از رابطه (۲) استفاده می‌شود:

$$\bar{H}_{ij} = \frac{\bar{z}_{ij}}{r} = \left( \frac{\bar{l}_{ij}}{r}, \frac{\bar{m}_{ij}}{r}, \frac{\bar{u}_{ij}}{r} \right) = (\bar{l}^r, \bar{m}^r, \bar{u}^r) \quad (2)$$

$$r = \max_{1 \leq i \leq n} (\sum_{j=1}^n u_{ij})$$

گام چهارم: تشکیل ماتریس روابط کل فازی. با استفاده از روابط ماتریس روابط کلی فازی محاسبه می‌شود:

$$T = \lim_{k \rightarrow \infty} (\bar{H}^1 \oplus \bar{H}^2 \oplus \dots \oplus \bar{H}^k) = X(I - X)^{-1} \quad (3)$$

$T$ : ماتریس ارتباط کامل

$I$ : ماتریس همانی

هر درایه  $T$  یک عدد فازی به صورت زیر است:

$$\bar{t}_{ij} = (\bar{l}_{ij}^t, \bar{m}_{ij}^t, \bar{u}_{ij}^t) \quad (4)$$

$$[\bar{l}_{ij}^t] = H_l \times (I - H_l)^{-1} \quad (5)$$

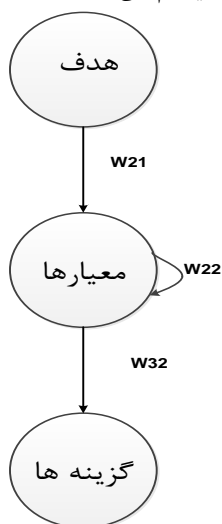
$$[\bar{m}_{ij}^t] = H_m \times (I - H_m)^{-1} \quad (6)$$

<sup>۱۱</sup> Analytic Hierarchy Process (AHP)

<sup>۹</sup> Threshold Value

<sup>۱۰</sup> Network Relations Map (NRM)

ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP



شکل ۲. نمونه‌ای از ساختار شبکه‌ای

گام ۲: تشکیل ماتریس مقایسات زوجی و تعیین بردارهای اولویت. مشابه مقایسات زوجی که در روش AHP انجام می‌شود عناصر تصمیم در هر یک از خوشه‌ها، بر اساس میزان اهمیت آن‌ها در ارتباط با معیارهای کنترلی دو به دو تقسیم می‌شوند. تاثیر هر عنصر بر روی عنصر دیگر از طریق بردار ویژه قابل ارائه است. اهمیت نسبی عناصر بر اساس مقیاس ۹ کمیته ساعتی سنجیده می‌شود. در این قسمت بردار اهمیت داخلی محاسبه می‌شود که نشانگر اهمیت نسبی (ضریب اهمیت) عناصر یا خوشه‌هاست، که از طریق رابطه (۱۱) بدست می‌آید:

$$Aw = \lambda_{max}W \quad (11)$$

که در آن:

$A$ : ماتریس مقایسه زوجی معیارها

$W$ : بردار ویژه (ضریب اهمیت)

$\lambda_{max}$ : بزرگترین مقدار ویژه عددی

جهت محاسبه بردار ویژه  $W$  ساعتی (۱۹۹۹) چندین روش ارائه کرده است و در صورتی که محاسبات قرار است بدون استفاده از نرم‌افزار خاصی انجام شوند، بهتر است از روش تقریب میانگین هندسی استفاده شود.

سلسله مراتب یک‌طرفه در سطوح مختلف تصمیم، AHP یک چهارچوب تصمیم‌گیری ارائه می‌کند که کاستی‌های ذاتی مراحل ارزیابی و اندازه‌گیری را برطرف کند، اما ناتوان از ارزیابی وابستگی‌های احتمالی داخلی در بین عوامل است. لذا روش فرآیند تحلیل شبکه‌ای<sup>۱۲</sup> ابداع گردید که هر موضوع و مسئله‌ای را به مثابه شبکه‌ای از معیارها، زیر معیارها و گزینه‌ها که با یکدیگر در خوشه‌هایی جمع شده‌اند، در نظر می‌گیرد، تا بازخورها و تعاملات موجود بین عوامل و خوشه‌ها را نیز در نظر بگیرد. بنابراین روش مذکور تبعات و آثار پیچیده تأثیرات متقابل را بویژه هنگامی که ریسک و عدم اطمینان وجود دارد، بهتر لحاظ می‌کند. هدف روش تصمیم‌گیری تحلیل شبکه‌ای ساختارمند کردن فرآیند تصمیم‌گیری با توجه به یک سناریو متأثر از فاکتورهای چندگانه مستقل از هم می‌باشد. این تکنیک، تصمیم‌گیری چندمعیاره را به وسیله جایگزینی شبکه به جای سلسله مراتب بهبود می‌بخشد. در این رویکرد، هدف مسئله بالاترین نقطه به شمار آمده و سطوح میانی شامل فاکتورهای نمایانگر سطوح بالاتر می‌باشند. آخرین سطح شامل گزینه‌ها یا فعالیت‌هایی است که برای رسیدن به هدف باید به آن‌ها توجه کرد [۳۰].

تمامی عناصر در یک شبکه می‌توانند، به هر شکل دارای ارتباط با یکدیگر باشند. به عبارت دیگر در یک شبکه، بازخورد و ارتباط متقابل بین و میان خوشه‌ها امکان پذیر است [۳۱]، بنابراین ANP را می‌توان متشکل از دو قسمت دانست: سلسله مراتب کنترلی و ارتباط شبکه‌ای. سلسله مراتب کنترلی ارتباط بین هدف، معیارها و زیر معیارها را شامل شده و بر ارتباط درونی سیستم تاثیرگذار است و ارتباط شبکه‌ای وابستگی بین عناصر و خوشه‌ها را شامل می‌شود [۳۲]. فرآیند تحلیل شبکه‌ای شامل چهار مرحله می‌باشد که عبارتست از:

گام ۱: ساخت مدل و تبدیل مسئله/ موضوع به یک ساختار شبکه‌ای موضوع/ مسئله باید به طور آشکار و روشن به یک سیستم منطقی، مثل یک شبکه تبدیل شود. این ساختار شبکه‌ای را می‌توان از طریق طوفان مغزی و یا هر روش مناسب دیگری چون دلفی، یا روش گروه اسمی بدست آورد. شکل (۲) نمونه‌ای از ساختار شبکه‌ای را به نمایش می‌گذارد.

<sup>۱۲</sup> Analytic Network Process (ANP)

ANP سنتی سخت و حتی غیرممکن می‌سازد. برای رفع این مشکل و لحاظ نمودن عدم قطعیت در قضاوت تصمیم‌گیرندگان از رویکردهای زبانی و منطق فازی استفاده می‌شود.

### ۳\_۳ تحلیل پوششی داده‌ها

تحلیل پوششی داده‌ها (DEA<sup>۱۳</sup>) یک مدل برنامه‌ریزی ریاضی برای ارزیابی کارایی واحدهای تصمیم‌گیرنده‌ای (DMU<sup>۱۴</sup>) است که چندین ورودی و چندین خروجی دارند. اندازه‌گیری کارایی به دلیل اهمیت آن در ارزیابی عملکرد یک شرکت یا سازمان همواره مورد توجه محققین قرار گرفته است. فارل، با استفاده از روشی همانند اندازه‌گیری کارایی در مباحث مهندسی، به اندازه‌گیری کارایی برای واحد تولیدی اقدام کرد. موردی که فارل برای اندازه‌گیری کارایی مد نظر قرار داد شامل یک ورودی و یک خروجی بود [۳۳]. چارنز، کوپر و رودز دیدگاه فارل را توسعه دادند و الگویی را ارائه کردند که توانایی اندازه‌گیری کارایی با چندین ورودی و خروجی را داشت. این الگو، تحت عنوان تحلیل پوششی داده‌ها نام گرفت و نخستین بار در رساله دکترای ادوارد رودز<sup>۱۵</sup> و به راهنمایی کوپر<sup>۱۶</sup> تحت عنوان ارزیابی پیشرفت تحصیلی دانش‌آموزان مدارس ملی آمریکا در سال ۱۹۷۶، در دانشگاه کارنگی مورد استفاده قرار گرفت از آن جا که این الگو توسط چارنز<sup>۱۷</sup>، کوپر و رودز ارائه گردید، به الگوی (CCR) که از حروف اول نام سه فرد یاد شده تشکیل شده است، معروف گردید [۳۴] و در سال ۱۹۷۸ در مقاله‌ای با عنوان اندازه‌گیری کارایی واحدهای تصمیم‌گیرنده ارائه شد [۳۵].

در واقع تحلیل پوششی داده‌ها مبتنی بر بهینه‌سازی با استفاده از برنامه‌ریزی خطی است که به آن روش ناپارامتریک نیز گفته می‌شود. در این روش منحنی مرزی کارا از نقاطی که بوسیله برنامه‌ریزی خطی تعیین می‌شود، ترسیم می‌گردد. برای تعیین این نقاط می‌توان از دو فرض بازدهی ثابت و متغیر نسبت به مقیاس استفاده کرد. روش برنامه‌ریزی خطی پس از یک سری بهینه‌سازی مشخص می‌کند که آیا واحد تصمیم‌گیرنده مورد نظر روی مرز کارایی قرار گرفته است و یا خارج آن قرار دارد؟ بدین وسیله واحدهای کارا و ناکارا از یکدیگر تفکیک می‌شوند. تکنیک DEA تمام داده‌ها را تحت پوشش قرار داده و به همین دلیل تحلیل پوششی داده‌ها نامیده شده است [۳۶]. بعدها بنکر<sup>۱۸</sup>، چارنز و کوپر با تغییر در مدل CCR مدل جدیدی را عرضه کردند که بر اساس حروف اول نام خانوادگی آنان به مدل

گام ۳: تشکیل سوپرماتریس و تبدیل آن به سوپرماتریس حد. برای دستیابی به اولویت‌های کلی در یک سیستم با تأثیرات متقابل، بردارهای اولویت داخلی (یعنی Wهای محاسبه شده) در ستون‌های مناسب یک ماتریس وارد می‌شوند. در نتیجه یک سوپرماتریس (درواقع یک ماتریس تقسیم‌بندی شده) که هر بخش از این ماتریس ارتباط بین دو خوشه در یک سیستم را نشان می‌دهد، بدست می‌آید. بعنوان مثال سوپرماتریس شبکه نمایش داده شده بدین صورت خواهد بود:

$$(12) \quad W_n = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 \\ w_{21} & w_{22} & 0 \\ 0 & w_{32} & 1 \end{bmatrix} \end{matrix}$$

این نوع سوپرماتریس را سوپرماتریس اولیه می‌نامند. با جایگزینی بردار اولویت‌های داخلی (ضرایب اهمیت) عناصر و خوشه‌ها در سوپرماتریس اولیه سوپرماتریس ناموزون بدست می‌آید. در مرحله بعد سوپرماتریس موزون از طریق ضرب مقادیر سوپرماتریس ناموزون در ماتریس خوشه‌ای محاسبه می‌شود. سپس از طریق نرمالیزه کردن سوپرماتریس موزون، سوپرماتریس از نظر ستونی به حالت تصادفی تبدیل می‌شود [۳۲].

در مرحله سوم سوپرماتریس حد با به توان رساندن تمامی عناصر سوپرماتریس موزون تا زمانی که واگرایی حاصل شود (از طریق تکرار)، یا به عبارت دیگر تمامی عناصر سوپرماتریس همانند هم شوند، محاسبه می‌شود.

گام ۴: انتخاب گزینه برتر. اگر سوپرماتریس تشکیل شده در مرحله سوم کل شبکه را در نظر گرفته باشد، یعنی گزینه‌ها نیز در سوپرماتریس لحاظ شده باشند، اولویت کلی گزینه‌ها از ستون مربوط به گزینه‌ها در سوپرماتریس حد نرمالیزه شده قابل حصول است. اگر سوپرماتریس، فقط بخشی از شبکه که وابستگی متقابل دارند را شامل شود و گزینه‌ها در سوپرماتریس در نظر گرفته نشوند، محاسبات بعدی لازم است صورت بگیرد تا اولویت کلی گزینه‌ها بدست آید. گزینه‌ای که بیشترین اولویت کلی را داشته باشد به عنوان برترین گزینه برای موضوع مورد نظر انتخاب می‌شود.

پیچیدگی و عدم اطمینان موجود در مسائل تصمیم‌گیری دنیای واقعی، بعضاً انجام مقایسه‌های دقیق و واقع‌بینانه را با استفاده از

<sup>۱۳</sup> William W. Cooper

<sup>۱۴</sup> Abraham Charnes

<sup>۱۵</sup> Rajiv D. Banker

<sup>۱۳</sup> Data Envelopment Analysis

<sup>۱۴</sup> Decision Making Unit

<sup>۱۵</sup> Edward Rhodes



ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP

$$\sum_{j=1}^n \lambda_j x_{ij} \geq y_{ro}, \quad r = 1, \dots, s$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$\lambda_j \geq 0, \quad j = 1, \dots, n$$

در مدل ثانویه BCC محدودیت متناظر با اضافه شدن متغیر آزاد در علامت  $w$  در مسئله اولیه، محدودیت  $\sum_{j=1}^n \lambda_j = 1$  است. در این مدل  $\theta$  نسبت کاهش ورودی‌های واحد تحت بررسی را برای بهبود کارایی نشان می‌دهد. یک واحد در این مدل کاراست اگر و تنها اگر دو شرط زیر برای آن برقرار باشد:

الف)  $\theta^* = 1$  شود.

ب) تمامی متغیرهای کمکی مقدار صفر داشته باشند.

از آنجا که مدل‌های DEA استاندارد قادر به اداره مجموعه داده غیرمنفی نیستند. هدف کلی استفاده از روش SORM به کار بردن تغییرات ماهرانه ذکر شده برای مدل BCC استاندارد است. فرض کنید  $I$  و  $R$  به ترتیب مجموعه‌ای از شاخص‌های ورودی‌ها و خروجی‌های غیرمنفی باشند. از نظر ریاضی  $I = \{i : x_i \geq 0\}$  و  $R = \{r : y_r \geq 0, r = 0, \dots, m\}$   $\{1, \dots, s\}$  به علاوه، فرض کنید  $I$  و  $K$  مجموعه‌ای از شاخص‌های یک ورودی و یک خروجی با کوچکترین داده منفی یا معادل  $I \in 1, R$   $k \in R^C \cap I^C$  باشند. علی‌نژاد و همکاران (۲۰۱۵)،  $x_i$  و  $y_k$  را با دو مقدار غیرمنفی به عنوان  $x_i^1, x_i^2, y_k^1, y_k^2$  به صورت زیر جایگزین کردند [۳۸]:

$$x_{ij}^1 = \begin{cases} x_{ij}, & x_{ij} > 0 \\ 0, & x_{ij} \leq 0 \end{cases} \quad x_{ij}^2$$

$$= \begin{cases} 0, & x_{ij} > 0 \\ -x_{ij}, & x_{ij} \leq 0 \end{cases}$$

$$y_{kj}^1 = \begin{cases} y_{kj}, & y_{kj} > 0 \\ 0, & y_{kj} \leq 0 \end{cases} \quad y_{kj}^2$$

$$= \begin{cases} 0, & y_{kj} > 0 \\ -y_{kj}, & y_{kj} \leq 0 \end{cases}$$

در ادامه مدلی برای اندازه‌گیری رتبه کارایی  $DMU_0$  پیشنهاد شده است:

$$\min Z = \theta \quad (18)$$

$$\text{s.t.} \quad \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{i0} \quad i \in I$$

$$\sum_{j=1}^n \lambda_j x_{ij}^1 \leq \theta x_{i0}^1 \quad i \in I^C$$

BCC شهرت یافت. مدل BCC مدلی از انواع مدل‌های تحلیل پوششی داده‌هاست که ارزیابی کارایی نسبی واحدهایی با بازده متغیر نسبت به مقیاس می‌پردازد. مدل نسبت BCC برای ارزیابی کارایی واحد تحت بررسی صفر بصورت زیر است [۳۷]:

$$\text{Max } Z. \quad \text{مدل نسبت BCC ورودی محور}$$

$$= \frac{\sum_{r=1}^s u_r y_{rj} + w}{\sum_{i=1}^m v_i x_{ij}}$$

$$\text{St:} \quad \begin{matrix} u_r, v_i & \geq 0 \\ \sum_{r=1}^s u_r y_{rj} + w & \leq 1 \\ \sum_{i=1}^m v_i x_{ij} & \leq 1 \end{matrix} \quad \begin{matrix} \text{W آزاد} \\ \text{در علامت} \end{matrix}$$

$$(j = 1, 2, \dots, n)$$

- مدل مضربی BCC ورودی محور:

مدل نسبت BCC یک مدل برنامه‌ریزی غیرخطی است که برای خطی کردن آن به دو شیوه می‌توان عمل کرد: حداکثر کردن صورت و ثابت نگه داشتن مخرج، که مدل مضربی BCC ورودی محور را ایجاد می‌کند، یا حداقل کردن مخرج و ثابت نگه داشتن صورت کسر که ایجاد کننده مدل مضربی BCC خروجی محور است. مدل مضربی اولیه (ورودی محور) بصورت زیر خواهد بود:

$$\text{Max } Z. = \sum_{r=1}^s u_r y_{rj} + w \quad (15)$$

$$\text{s.t.}:$$

$$\sum_{i=1}^m v_i x_{ij} = 1$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} + w \leq 0$$

$$(j = 1, 2, \dots, n)$$

$$u_r, v_i \geq 0$$

W آزاد در علامت

تفاوت این مدل با مدل CCR در وجود متغیر آزاد در علامت  $w$  است. علامت متغیر  $w$  در این مدل نوع بازده به مقیاس را بصورت زیر تعیین می‌کند:

الف: هرگاه  $w < 0$  باشد نوع بازده به مقیاس، کاهش‌ی است.

ب: هرگاه  $w = 0$  باشد نوع بازده به مقیاس، ثابت است.

پ: هرگاه  $w > 0$  باشد نوع بازده به مقیاس، افزایش است.

ثانویه مدل BCC به صورت زیر است. متغیر متناظر با محدودیت اول مسئله با  $\theta$  و متغیرهای متناظر با سایر محدودیت‌ها با  $\lambda_j$  به نمایش گذاشته شده است.

$$\min Z = \theta \quad (16)$$

$$\text{s.t.} \quad \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{i0}, \quad i = 1, \dots, m$$

جدول ۳. مقیاس ارزیابی فازی

میزان تأثیرگذاری عبارات کلامی	اعداد فازی مثلثی / اعداد قطعی مورد استفاده در پرسشنامه
بدون تأثیر	۰/(۰,۰/۰,۱/۳)
تأثیر کم	۱/(۰/۰,۱/۰,۳/۵)
تأثیر متوسط	۲/(۰/۰,۳/۰,۵/۷)
تأثیر زیاد	۳/(۰/۰,۵/۰,۷/۹)
تأثیر خیلی زیاد	۴/(۰/۰,۷/۱,۰/۹)

با استفاده از نتیجه نهائی این مقایسات ماتریس روابط مستقیم فازی برای معیارها تشکیل شده که در جدول ۳ نشان داده شده است. در این ماتریس‌ها  $\tilde{x}_{ij} = (l_{ij} + m_{ij} + u_{ij})$  اعداد فازی مثلثی می‌باشند. به صورت نمونه دو جدول مقایسات زوجی کارشناسان اول و دوم نمایش داده شده است:

$$\sum_{j=1}^n \lambda_j x_{ij}^v \geq \theta x_{io}^v \quad i \in I^c$$

$$\sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad r \in R$$

$$\sum_{j=1}^n \lambda_j y_{rj}^1 \geq y_{ro}^1 \quad r \in R^c$$

$$\sum_{j=1}^n \lambda_j y_{rj}^v \leq y_{ro}^v \quad r \in R^c$$

$$\sum_{j=1}^n \lambda_j = 1 \quad \lambda_j \geq 0, \forall j$$

در نتیجه در مدل فوق:

$$\frac{|I| + |R| + 2(|I^c| + |R^c|) + 1}{n + 1}$$

(۱۹) متغیرها

#### ۴. تجزیه و تحلیل داده‌ها

#### ۴\_۱ انتخاب شاخص‌های ورودی با استفاده از روش

#### دیمتل فازی

با توجه به بررسی متون علمی تدوین شده در حوزه ERP شاخص‌های ورودی جهت بررسی به شرح زیر می‌باشند، که سه شاخص برتر به روش دیمتل فازی انتخاب گردیده و جهت ارزیابی کارائی سیستم‌های یکپارچه‌سازی منابع سازمان مورد استفاده قرار می‌گیرد.

جدول ۲. شاخص‌های ورودی

نماد	شرح عامل	نماد	شرح عامل
C1	پیچیدگی اجرا	C2	زمان صرف شده برای اجرا
C3	آموزش و اقدامات حمایتی کاربر	C4	زیر ساخت‌های اجرای
C5	هزینه‌های بسترسازی	C6	آموزش پرسنل
C7	پشتیبانی مناسب سازمان		

گام‌های بررسی روابط درونی و چگونگی تأثیر معیارها با استفاده از روش دیمتل فازی به شرح زیر می‌باشد:

گام اول: تهیه پرسشنامه و ایجاد ماتریس مقایسات زوجی جهت بررسی روابط درونی معیارها از ۷ کارشناس متخصص در این حوزه درخواست شده تا به مقایسات زوجی بین شاخص‌های ورودی از نقطه نظر میزان تأثیر معیار I بر معیار J در ستون پرداخته (بدین منظور پرسشنامه طراحی گردیده و از گزینه‌های زبانی و اعداد فازی که در جدول ۳ آمده است جهت مقایسه استفاده شده است).

جدول ۴. ماتریس تصمیم معیارهای خبره اول و دوم به عنوان نمونه

C <sub>۷</sub>			C <sub>۶</sub>			C <sub>۵</sub>			C <sub>۴</sub>			C <sub>۳</sub>			C <sub>۲</sub>			C <sub>۱</sub>			DM <sub>۱</sub>	
۱	۰/۹	۰/۷	۱	۰/۹	۰/۷	۰/۷	۰/۵	۰/۳	۰/۵	۰/۳	۰/۱	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۳	۰/۱	۰	۰	C <sub>۱</sub>
۱	۰/۹	۰/۷	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۳	۰/۱	۰	۰/۷	۰/۵	۰/۳	۰/۳	C <sub>۲</sub>
۰/۵	۰/۳	۰/۱	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۳	۰/۱	۰	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۳	C <sub>۳</sub>
۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۳	۰/۱	۰	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۳	C <sub>۴</sub>
۰/۹	۰/۷	۰/۵	۰/۹	۰/۷	۰/۵	۰/۳	۰/۱	۰	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۵	C <sub>۵</sub>
۰/۹	۰/۷	۰/۵	۰/۳	۰/۱	۰	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۳	۰/۱	۰	۰	C <sub>۶</sub>
۰/۳	۰/۱	۰	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۳	۰/۱	۰	۰/۳	۰/۱	۰	۰	C <sub>۷</sub>
C <sub>۷</sub>			C <sub>۶</sub>			C <sub>۵</sub>			C <sub>۴</sub>			C <sub>۳</sub>			C <sub>۲</sub>			C <sub>۱</sub>			DM <sub>۲</sub>	
۰/۵	۰/۳	۰/۱	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۵	۰/۳	۰/۱	۰/۷	۰/۵	۰/۳	۰/۳	۰/۱	۰	۰	C <sub>۱</sub>
۰/۵	۰/۳	۰/۱	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۹	۰/۷	۰/۵	۱	۰/۹	۰/۷	۰/۳	۰/۱	۰	۰/۷	۰/۵	۰/۳	۰/۳	C <sub>۲</sub>
۰/۹	۰/۷	۰/۵	۱	۰/۹	۰/۷	۰/۷	۰/۵	۰/۳	۱	۰/۹	۰/۷	۰/۳	۰/۱	۰	۱	۰/۹	۰/۷	۱	۰/۹	۰/۷	۰/۹	C <sub>۳</sub>
۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۳	۰/۱	۰	۰/۳	۰/۱	۰	۱	۰/۹	۰/۷	۱	۰/۹	۰/۷	۰/۹	۰/۷	۰/۵	۰/۳	C <sub>۴</sub>
۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۳	۰/۱	۰	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۷	۰/۵	۰/۳	۰/۵	C <sub>۵</sub>
۰/۹	۰/۷	۰/۵	۰/۳	۰/۱	۰	۰/۵	۰/۳	۰/۱	۰/۵	۰/۳	۰/۱	۰/۹	۰/۷	۰/۵	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۳	C <sub>۶</sub>
۰/۳	۰/۱	۰	۰/۵	۰/۳	۰/۱	۰/۷	۰/۵	۰/۳	۰/۵	۰/۳	۰/۱	۰/۷	۰/۵	۰/۳	۰/۹	۰/۷	۰/۵	۰/۵	۰/۳	۰/۱	۰	C <sub>۷</sub>

میانگین حسابی ساده نظرات استفاده کرده و ماتریس ارتباط مستقیم را بدست می‌آوریم، که در جدول ۶ نمایش داده شده است.

گام دوم: محاسبه ماتریس ارتباط مستقیم فازی در تکنیک دیمتل گروهی زمانی که از دیدگاه چند خبره استفاده می‌شود جهت تشکیل ماتریس ارتباط مستقیم از

جدول ۵. ماتریس تصمیم متوسط

C <sub>۷</sub>			C <sub>۶</sub>			C <sub>۵</sub>			C <sub>۴</sub>			C <sub>۳</sub>			C <sub>۲</sub>			C <sub>۱</sub>			ردیف
۰/۷۰	۰/۵۳	۰/۳۳	۰/۸۶	۰/۷۰	۰/۵۰	۰/۶۴	۰/۴۴	۰/۲۴	۰/۸۰	۰/۶۱	۰/۴۱	۰/۷۶	۰/۵۶	۰/۳۶	۰/۸۳	۰/۶۴	۰/۴۴	۰/۳۰	۰/۱۰	۰/۰۰	C <sub>۱</sub>
۰/۷۶	۰/۵۹	۰/۳۹	۰/۷۶	۰/۵۶	۰/۳۶	۰/۷۷	۰/۵۹	۰/۳۹	۰/۷۹	۰/۵۹	۰/۳۹	۰/۸۰	۰/۶۱	۰/۴۱	۰/۳۰	۰/۱۰	۰/۰۰	۰/۷۶	۰/۵۶	۰/۳۶	C <sub>۲</sub>
۰/۷۹	۰/۵۹	۰/۳۹	۰/۸۳	۰/۶۴	۰/۴۴	۰/۷۳	۰/۵۳	۰/۳۳	۰/۷۷	۰/۵۹	۰/۳۹	۰/۳۰	۰/۱۰	۰/۰۰	۰/۷۱	۰/۵۳	۰/۳۳	۰/۷۴	۰/۵۶	۰/۳۶	C <sub>۳</sub>
۰/۷۰	۰/۵۰	۰/۳۰	۰/۸۰	۰/۶۱	۰/۴۱	۰/۷۰	۰/۵۰	۰/۳۱	۰/۳۰	۰/۱۰	۰/۰۰	۰/۷۷	۰/۵۹	۰/۳۹	۰/۸۳	۰/۶۴	۰/۴۴	۰/۸۴	۰/۶۴	۰/۴۴	C <sub>۴</sub>
۰/۸۱	۰/۶۱	۰/۴۱	۰/۷۶	۰/۵۶	۰/۳۶	۰/۳۰	۰/۱۰	۰/۰۰	۰/۷۳	۰/۵۳	۰/۳۳	۰/۷۴	۰/۵۶	۰/۳۶	۰/۷۰	۰/۵۰	۰/۳۰	۰/۷۳	۰/۵۳	۰/۳۳	C <sub>۵</sub>
۰/۸۱	۰/۶۱	۰/۴۱	۰/۳۰	۰/۱۰	۰/۰۰	۰/۶۷	۰/۴۷	۰/۲۷	۰/۷۰	۰/۵۰	۰/۳۰	۰/۷۶	۰/۵۶	۰/۳۶	۰/۷۱	۰/۵۳	۰/۳۳	۰/۶۱	۰/۴۱	۰/۲۴	C <sub>۶</sub>
۰/۳۰	۰/۱۰	۰/۰۰	۰/۶۴	۰/۴۴	۰/۲۴	۰/۶۹	۰/۵۰	۰/۳۰	۰/۷۰	۰/۵۰	۰/۳۰	۰/۶۴	۰/۴۴	۰/۲۴	۰/۶۷	۰/۴۷	۰/۲۷	۰/۵۰	۰/۳۰	۰/۱۳	C <sub>۷</sub>

گام سوم: نرمالایز کردن ماتریس ارتباط مستقیم فازی جهت نرمالایز کردن ماتریس ارتباط مستقیم از رابطه ۲ استفاده کرده و ماتریس زیر را بدست آورده‌ایم.

جدول ۶. ماتریس ارتباط مستقیم نرمال شده فازی

	C <sub>۷</sub>			C <sub>۶</sub>			C <sub>۵</sub>			C <sub>۴</sub>			C <sub>۳</sub>			C <sub>۲</sub>			C <sub>۱</sub>			M
C <sub>۱</sub>	-/۱۴	-/۱۱	-/۰۷	-/۱۷	-/۱۴	-/۱۰	-/۱۳	-/۰۹	-/۰۵	-/۱۶	-/۱۲	-/۰۸	-/۱۵	-/۱۱	-/۰۷	-/۱۷	-/۱۳	-/۰۹	-/۰۶	-/۰۲	-/۰۰	
C <sub>۲</sub>	-/۱۵	-/۱۲	-/۰۸	-/۱۵	-/۱۱	-/۰۷	-/۱۶	-/۱۲	-/۰۸	-/۱۶	-/۱۲	-/۰۸	-/۱۶	-/۱۲	-/۰۸	-/۰۶	-/۰۲	-/۰۰	-/۱۵	-/۱۱	-/۰۷	
C <sub>۳</sub>	-/۱۶	-/۱۲	-/۰۸	-/۱۷	-/۱۳	-/۰۹	-/۱۵	-/۱۱	-/۰۷	-/۱۶	-/۱۲	-/۰۸	-/۱۶	-/۰۲	-/۰۰	-/۱۴	-/۱۱	-/۰۷	-/۱۵	-/۱۱	-/۰۷	
C <sub>۴</sub>	-/۱۴	-/۱۰	-/۰۶	-/۱۶	-/۱۲	-/۰۸	-/۱۴	-/۱۰	-/۰۶	-/۰۶	-/۰۲	-/۰۰	-/۱۶	-/۱۲	-/۰۸	-/۱۷	-/۱۳	-/۰۹	-/۱۷	-/۱۳	-/۰۹	
C <sub>۵</sub>	-/۱۶	-/۱۲	-/۰۸	-/۱۵	-/۱۱	-/۰۷	-/۰۶	-/۰۲	-/۰۰	-/۱۵	-/۱۱	-/۰۷	-/۱۵	-/۱۱	-/۰۷	-/۱۴	-/۱۰	-/۰۶	-/۱۵	-/۱۱	-/۰۷	
C <sub>۶</sub>	-/۱۶	-/۱۲	-/۰۸	-/۱۶	-/۰۲	-/۰۰	-/۱۴	-/۱۰	-/۰۵	-/۱۴	-/۱۰	-/۰۶	-/۱۵	-/۱۱	-/۰۷	-/۱۴	-/۱۱	-/۰۷	-/۱۲	-/۰۸	-/۰۵	
C <sub>۷</sub>	-/۰۶	-/۰۲	-/۰۰	-/۱۳	-/۰۹	-/۰۵	-/۱۴	-/۱۰	-/۰۶	-/۱۴	-/۱۰	-/۰۶	-/۱۳	-/۰۹	-/۰۵	-/۱۴	-/۱۰	-/۰۶	-/۱۰	-/۰۶	-/۰۳	

گام چهارم: تشکیل ماتریس روابط کل فازی. با استفاده از روابط ۳، ۵، ۶ و ۷ ماتریس ارتباط کامل را بدست آورده‌ایم که در جدول ۷ ماتریس تمامی روابط صرف نظر شده شده است.

جدول ۷. ماتریس روابط کل فازی

	C <sub>۷</sub>			C <sub>۶</sub>			C <sub>۵</sub>			C <sub>۴</sub>			C <sub>۳</sub>			C <sub>۲</sub>			C <sub>۱</sub>			T
C <sub>۱</sub>	۳/۳۱	-/۳۴	-/۱۲	۳/۳۸	-/۳۷	-/۱۵	۳/۰۸	-/۳۰	-/۰۹	۳/۳۸	-/۳۴	-/۱۲	۳/۲۶	-/۳۴	-/۱۲	۳/۲۶	-/۳۵	-/۱۴	۳/۰۰	-/۲۳	-/۰۵	
C <sub>۲</sub>	۳/۳۵	-/۳۵	-/۱۳	۳/۳۸	-/۳۵	-/۱۲	۳/۱۲	-/۳۲	-/۱۲	۳/۳۰	-/۳۴	-/۱۳	۳/۲۹	-/۳۴	-/۱۳	۳/۱۹	-/۳۵	-/۰۵	۳/۱۱	-/۳۱	-/۱۱	
C <sub>۳</sub>	۳/۳۱	-/۳۴	-/۱۳	۳/۳۶	-/۳۶	-/۱۴	۳/۰۸	-/۳۱	-/۱۱	۳/۲۵	-/۳۴	-/۱۲	۳/۱۶	-/۳۵	-/۰۵	۳/۲۳	-/۳۳	-/۱۱	۳/۰۷	-/۳۱	-/۱۱	
C <sub>۴</sub>	۳/۳۵	-/۳۳	-/۱۱	۳/۴۱	-/۳۶	-/۱۴	۳/۱۲	-/۳۱	-/۱۱	۳/۳۲	-/۳۵	-/۰۵	۳/۳۰	-/۳۴	-/۱۳	۳/۳۰	-/۳۵	-/۱۴	۳/۱۳	-/۳۳	-/۱۳	
C <sub>۵</sub>	۳/۲۶	-/۳۴	-/۱۳	۳/۳۸	-/۳۳	-/۱۲	۳/۹۴	-/۳۲	-/۰۴	۳/۱۹	-/۳۲	-/۱۱	۳/۱۸	-/۳۲	-/۱۲	۳/۱۶	-/۳۱	-/۱۱	۳/۰۱	-/۳۰	-/۱۰	
C <sub>۶</sub>	۳/۱۳	-/۳۳	-/۱۲	۳/۰۷	-/۳۴	-/۰۵	۳/۰۰	-/۳۸	-/۰۹	۳/۰۶	-/۳۰	-/۱۰	۳/۰۶	-/۳۱	-/۱۱	۳/۰۵	-/۳۰	-/۱۱	۳/۸۷	-/۲۷	-/۰۹	
C <sub>۷</sub>	۳/۳۸	-/۳۱	-/۰۴	۳/۳۸	-/۳۷	-/۰۹	۳/۶۶	-/۳۶	-/۰۹	۳/۳۱	-/۳۷	-/۰۹	۳/۳۰	-/۳۷	-/۰۸	۳/۲۹	-/۳۷	-/۰۹	۳/۶۲	-/۳۲	-/۰۶	

گام پنجم: فازی زدایی. در این گام جهت بدست آوردن جواب قطعی عمل فازی زدایی را با استفاده از رابطه ۸ انجام داده که نتیجه در جدول ۸ به نمایش در آمده است.

جدول ۸. ماتریس روابط کل قطعی

	C <sub>۷</sub>	C <sub>۶</sub>	C <sub>۵</sub>	C <sub>۴</sub>	C <sub>۳</sub>	C <sub>۲</sub>	C <sub>۱</sub>
C <sub>۱</sub>	۱/۰۳	۱/۰۷	۰/۹۴	۱/۰۲	۱/۰۱	۱/۰۳	۰/۸۸
C <sub>۲</sub>	۱/۰۴	۱/۰۵	۰/۹۷	۱/۰۳	۱/۰۳	۰/۹۴	۰/۹۶
C <sub>۳</sub>	۱/۰۳	۱/۰۵	۰/۹۵	۱/۰۱	۰/۹۳	۱/۰۰	۰/۹۵

ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP

۱/۰۳	۱/۰۶	۰/۹۶	۰/۹۴	۱/۰۳	۱/۰۳	۰/۹۸	C۴
۱/۰۲	۱/۰۲	۰/۸۶	۰/۹۸	۰/۹۹	۰/۹۷	۰/۹۳	C۵
۰/۹۸	۰/۹۰	۰/۸۹	۰/۹۴	۰/۹۵	۰/۹۴	۰/۸۷	C۶
۰/۸۱	۰/۸۸	۰/۸۲	۰/۸۶	۰/۸۵	۰/۸۶	۰/۷۸	C۷

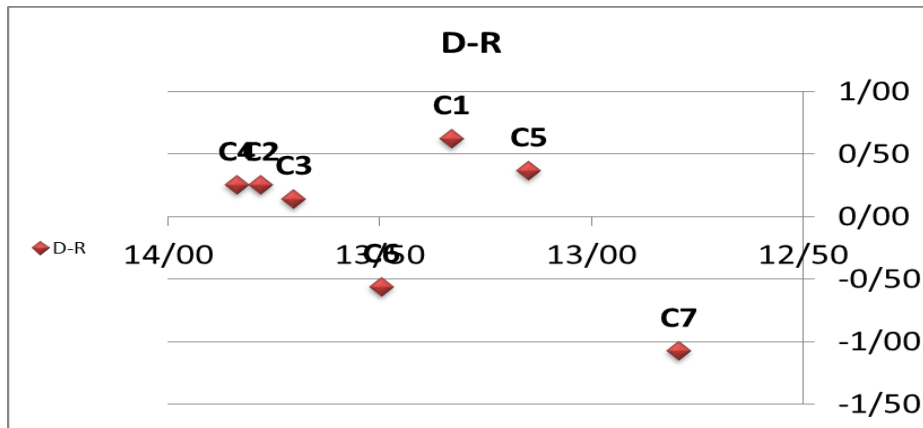
گام هشتم: محاسبه جمع سطرها و ستون‌های ماتریس ارتباط کامل.

جدول ۹. محاسبه جمع سطرها و ستون‌ها

C۷	C۶	C۵	C۴	C۳	C۲	C۱	
۵/۸۶	۶/۴۷	۶/۷۶	۷/۰۵	۶/۹	۷/۰۲	۶/۹۸	D
۶/۹۴	۷/۰۳	۶/۳۹	۶/۷۹	۶/۷۸	۶/۷۶	۶/۳۵	R

نشان دهنده D-R است. معیارهایی که بالای خط افق قرار دارند علت‌ها و معیارهایی که پایین خط افق قرار دارند معلول-ها را نمایش می‌دهند.

گام هشتم: رسم نمودار روابط علی و معلولی. نمودار روابط علی و معلولی نشان‌دهنده رابطه علت و معلولی میان معیارها می‌باشد. محور افقی نشان دهنده D+R و محور عمودی

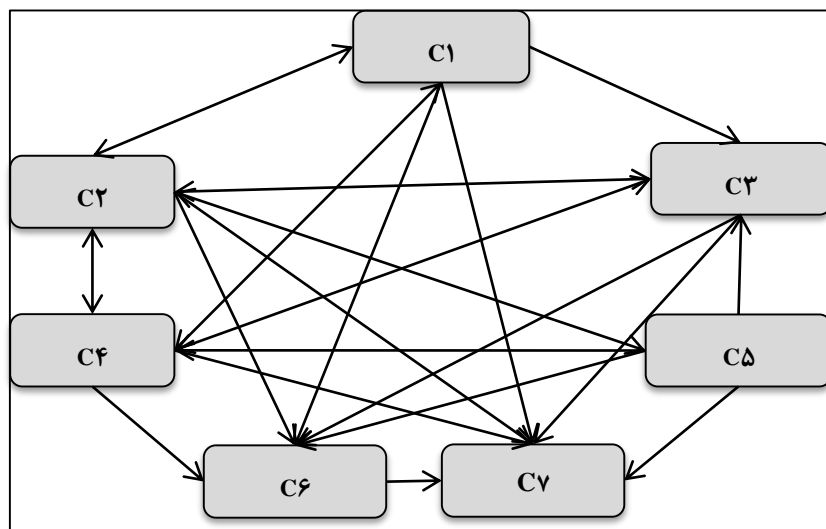


شکل ۳. دیاگرام روابط علی و معلولی

جدول ۱۰. الگوی روابط معنی دار معیارها

C۷	C۶	C۵	C۴	C۳	C۲	C۱	
X	X		X	X	X		C۱
X	X	X	X	X		X	C۲
X	X		X		X		C۳
X	X			X	X	X	C۴
X	X		X	X	X		C۵
X							C۶
							C۷

گام نهم: محاسبه ارزش آستانه و رسم نقشه روابط شبکه. جهت رسم نقشه روابط شبکه مقدار آستانه که برابر با میانگین عناصر ماتریس ارتباط کامل می‌باشد را بدست می‌آوریم. در این پژوهش مقدار آستانه برابر است با ۰,۹۶، با توجه به این مقدار تمامی مقادیر ماتریس رابطه کل که کوچک‌تر از این عدد باشند را مساوی با صفر قرار می‌دهیم و آن رابطه را علی در نظر نمی‌گیریم. در نقشه روابط شبکه تنها عواملی که مقدارشان در ماتریس رابطه کل بزرگ‌تر از مقدار آستانه باشد نمایش داده می‌شوند.



شکل ۴. نقشه روابط شبکه

#### ۲\_۴ رتبه‌بندی شاخص‌های ورودی با استفاده از روش DEMATEL

بررسی نتایج نهایی بدست آمده از اجرای دیمتل فازی:

جدول ۱۱. نتایج بدست آمده از تکنیک دیمتل

D-R	D+R	R	D	T
-۰/۶۲	۱۳/۳۳	۶/۳۵	۶/۹۸	پیچیدگی اجرا
-۰/۲۵	۱۳/۷۸	۶/۷۶	۷/۰۲	زمان صرف شده برای اجرا
-۰/۱۴	۱۳/۷۰	۶/۷۸	۶/۹۲	آموزش و اقدامات حمایتی کاربر
-۰/۲۵	۱۳/۸۴	۶/۷۹	۷/۰۵	زیر ساخت‌های اجرای
-۰/۳۷	۱۳/۱۵	۶/۳۹	۶/۷۶	هزینه‌های بسترسازی
-۰/۵۶	۱۳/۵۰	۷/۰۳	۶/۴۷	آموزش پرسنل
-۱/۰۸	۱۲/۷۹	۶/۹۴	۵/۸۶	پشتیبانی مناسب سازمان

کمترین تأثیرپذیری را از سایر معیارها دارد. بردار افقی (D + R)، میزان تأثیر و تأثر عامل مورد نظر در سیستم است. به عبارت دیگر هرچه مقدار D + R عاملی بیشتر باشد، آن عامل تعامل بیشتری با سایر عوامل سیستم دارد و از اهمیت بیشتری برخوردار است. براین اساس شاخص «زیر ساخت‌های اجرای»، «زمان صرف شده برای اجرا» و «آموزش و اقدامات حمایتی کاربر» بیشترین تعامل را با سایر معیارهای مورد مطالعه دارند، و جهت ارزیابی عملکرد سیستم‌های ERP مورد استفاده قرار داده شده است. بردار عمودی (D-R)،

در جدول فوق جمع عناصر هر سطر (D) نشانگر میزان تأثیرگذاری آن عامل بر سایر عامل‌های سیستم است. براین اساس «زیر ساخت‌های اجرای» از بیشترین تأثیرگذاری برخوردار است. شاخص «زمان صرف شده برای اجرا» و «پیچیدگی اجرا» با تأثیرگذاری تقریباً مشابه در جایگاه دوم و سوم قرار دارد. جمع عناصر ستون (R) برای هر عامل نشانگر میزان تأثیرپذیری آن عامل از سایر عامل‌های سیستم است. براین اساس شاخص «آموزش پرسنل» از میزان تأثیرپذیری بسیار زیادی برخوردار است. معیار «پیچیدگی اجرا» نیز

ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP

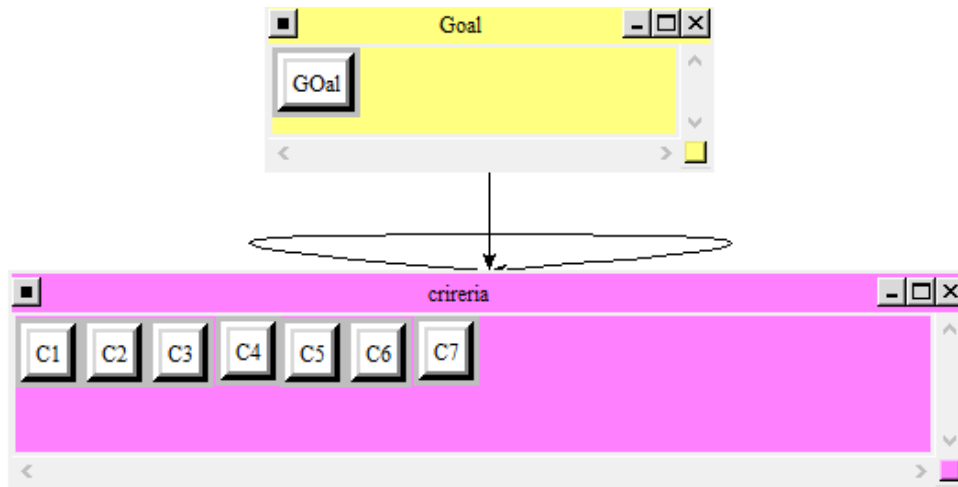
سیستم‌های ERP از میان ۷ خروجی که از مطالعه ادبیات پیشین در این زمینه بدست آورده شده است، با استفاده از تکنیک ANP فازی پرداخته خواهد شد. شاخص‌های خروجی بدست آمده از مطالعات انجام شده به شرح جدول ۱۲ است:

سیستم‌های ERP در شکل (۵) ارائه شده است. این مهم با توجه به ساختار شبکه‌ای سوپرماتریس اولیه که به شرح زیر است، انجام می‌شود.

(۱)

$$W = \begin{matrix} \text{هدف} & \text{معیارها} \\ \text{معیارها} & \begin{pmatrix} w_{21} & w_{22} \end{pmatrix} \end{matrix}$$

نمایش شبکه‌ای مدل در نرم افزار سوپر دسیژن مطابق شکل زیر می‌باشد:



شکل ۵. ساختار شبکه‌ای در نرم افزار Super Decisions

شده است. از آنجایی که در این پژوهش از قضاوت گروهی که توسط ۷ خبره صورت گرفته استفاده می‌شود، ماتریس پیش‌رو از میانگین هندسی نظرات گروهی بدست آمده است.

قدرت تأثیرگذاری هر عامل را نشان می‌دهد. بطور کلی اگر  $D - R$  مثبت باشد، متغیر یک متغیر علی محسوب می‌شود و اگر منفی باشد، معلول محسوب می‌شود.

### ۳\_۴ انتخاب شاخص‌های خروجی با تکنیک ANP فازی

در این بخش به انتخاب سه خروجی برتر جهت ارزیابی عملکرد

جدول ۱۲. شاخص‌های خروجی

نماد	شرح عامل	نماد	شرح عامل
C۱	شاخص افزایش بهره‌وری سازمان	C۲	مدیریت صحیح منابع
C۳	میزان رضایتمندی کاربر	C۴	انطباق عملیاتی نرم‌افزار
C۵	پویایی واحدها	C۶	کاهش هزینه‌های تکنولوژی اطلاعات
C۷	توانمندسازی سازمان در توسعه		

گام‌های تکنیک ANP به شرح زیر می‌باشد:

جهت پیاده‌سازی تکنیک ANP در این پژوهش از نرم افزار سوپر دسیژن<sup>۱۹</sup> استفاده شده است. گام اول ایجاد ساختار شبکه‌ای.

این پژوهش بدنبال اولویت‌بندی معیارهای اصلی که پیشتر ذکر گردید، می‌باشد. مدل شبکه‌ای جهت اولویت‌بندی ۷ معیار ورودی

گام دوم تشکیل ماتریس مقایسات زوجی و کنترل سازگاری آن‌ها \*مقایسه زوجی معیارها ( $W_{21}$ ):

مقایسه زوجی معیارهای ۷ گانه بر اساس مقیاس ۹ کمیته ساعتی (۱۹۹۹) انجام می‌گیرد [۳۲]. نتیجه مقایسه زوجی معیارها و همچنین بردار موزون حاصل از آن یعنی  $W_{21}$  در جدول (۱۳) ارائه

<sup>۱۹</sup> Super Decision

جدول ۱۳. جدول مقایسات زوجی معیارها

معیارها	C <sub>۱</sub>	C <sub>۲</sub>	C <sub>۳</sub>	C <sub>۴</sub>	C <sub>۵</sub>	C <sub>۶</sub>	C <sub>۷</sub>
C <sub>۱</sub>	۱	۱	۱	۱	۲	۳	۳
C <sub>۲</sub>	۱	۱	۳	۲	۲	۴	۵
C <sub>۳</sub>	۱	۰,۳	۱	۱	۱	۲	۳
C <sub>۴</sub>	۱	۰,۵	۱	۱	۲	۳	۴
C <sub>۵</sub>	۰,۵	۰,۵	۱	۰,۵	۱	۲	۵
C <sub>۶</sub>	۰,۳	۰,۲۵	۰,۵	۰,۳	۰,۵	۱	۲
C <sub>۷</sub>	۰,۳	۰,۲	۰,۳	۰,۲۵	۰,۲	۰,۵	۱

$\text{ضریب ناسازگاری (CR)} = 0,024$

\*مقایسه زوجی وابستگی‌های دورنی معیارها (ماتریس  $W_{22}$ )  
 برای درک وابستگی‌های متقابل بین معیارها، مقایسه زوجی بین معیارها به منظور دستیابی به عناصر ماتریس  $W_{22}$  و بر اساس مقیاس ۹ کمیتی ساعتی (۱۹۹۹) انجام می‌شود [۳۲]. برای نحوه

محاسبه ضریب اهمیت هر یک از معیارهای اصلی (با توجه به وابستگی متقابل بین آن‌ها) مقایسه زوجی ۶ گانه دیگر (با کنترل کردن معیار اول یعنی C<sub>۱</sub>) در جدول ۱۴ ارائه شده است:

جدول ۱۴. مقایسات زوجی معیارها با توجه به وابستگی درونی آن‌ها با کنترل C<sub>۱</sub>

معیارها	C <sub>۲</sub>	C <sub>۳</sub>	C <sub>۴</sub>	C <sub>۵</sub>	C <sub>۶</sub>	C <sub>۷</sub>
C <sub>۲</sub>	۱	۱	۲	۳	۲	۱
C <sub>۳</sub>	۱	۱	۱	۲	۳	۲
C <sub>۴</sub>	۰,۵	۱	۱	۴	۳	۲
C <sub>۵</sub>	۰,۳	۰,۵	۰,۲۵	۱	۱	۱
C <sub>۶</sub>	۰,۵	۰,۳	۰,۳	۱	۱	۱
C <sub>۷</sub>	۱	۰,۵	۰,۵	۱	۱	۱

$\text{ضریب ناسازگاری (CR)} = 0,040$

\* تشکیل سوپرماتریس ناموزون:  
 با توجه به اینکه کلیه ماتریس‌های مقایسه‌ای موجود در ساختار سوپر ماتریس ناموزون ( $W_{21}, W_{22}$ ) محاسبه شده و سازگاری آن‌ها نیز کنترل شده است، می‌توان با جایگزین کردن این ماتریس‌ها در سوپر ماتریس اولیه سوپرماتریس ناموزون را به شرح جدول زیر بدست آورد.

شش ماتریس مقایسه زوجی دیگر شبیه ماتریس ارائه شده در جدول (۱۴) تشکیل شده و ضریب ناسازگاری هر یک از آن‌ها کنترل شده است تا بتوان ماتریس مربوط به وابستگی‌های متقابل معیارها ( $W_{22}$ ) را محاسبه کرد. مانند مرحله قبل بدلیل استفاده از قضاوت گروهی از میانگین هندسی استفاده شده است. بدلیل حجم بالای جداول از نمایش سایر جداول صرف نظر شده است.

گام سوم: محاسبه سوپر ماتریس حد

جهت محاسبه سوپر ماتریس حد باید مراحل زیر را طی کرد:



جدول ۱۵. سوپر ماتریس ناموزون

Cluster Node Labels		criteria							Goal
		C1	C2	C3	C4	C5	C6	C7	GOal
criteria	C1	0.000000	0.348332	0.366894	0.345932	0.336037	0.302102	0.280130	0.186665
	C2	0.240130	0.000000	0.264386	0.232700	0.254705	0.296744	0.195852	0.269171
	C3	0.223278	0.165149	0.000000	0.130083	0.157011	0.185765	0.178813	0.134894
	C4	0.232408	0.184050	0.127105	0.000000	0.095905	0.109615	0.188658	0.173851
	C5	0.088582	0.131762	0.089640	0.118300	0.000000	0.055259	0.095334	0.127861
	C6	0.092129	0.073947	0.088266	0.089383	0.066820	0.000000	0.061212	0.065160
	C7	0.123472	0.096760	0.063709	0.083602	0.089522	0.050514	0.000000	0.042398
Goal	GOal	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

\* تشکیل سوپر ماتریس موزون: موزون را در ماتریس خوشه‌ای ضرب کرد. ماتریس خوشه‌ای میزان تاثیر گذاری هر یک از خوشه‌ها برای دستیابی به اهداف مطالعه را منعکس می‌کند. سوپر ماتریس موزون این پژوهش در جدول ۱۶ ارائه شده است.

جدول ۱۶. سوپر ماتریس موزون

Cluster Node Labels		criteria							Goal
		C1	C2	C3	C4	C5	C6	C7	GOal
criteria	C1	0.000000	0.348332	0.366894	0.345932	0.336037	0.302102	0.280130	0.186665
	C2	0.240130	0.000000	0.264386	0.232700	0.254705	0.296744	0.195852	0.269171
	C3	0.223278	0.165149	0.000000	0.130083	0.157011	0.185765	0.178813	0.134894
	C4	0.232408	0.184050	0.127105	0.000000	0.095905	0.109615	0.188658	0.173851
	C5	0.088582	0.131762	0.089640	0.118300	0.000000	0.055259	0.095334	0.127861
	C6	0.092129	0.073947	0.088266	0.089383	0.066820	0.000000	0.061212	0.065160
	C7	0.123472	0.096760	0.063709	0.083602	0.089522	0.050514	0.000000	0.042398
Goal	GOal	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

\* محاسبه سوپر ماتریس حد: هدف از به حد رساندن سوپر ماتریس موزون این است که تأثیر نسبی دراز مدت هر یک از عناصر آن در یکدیگر حاصل شود. برای واگرایی ضریب اهمیت هر یک از عناصر ماتریس موزون، آن را به توان  $k$  که

جدول ۱۷. سوپر ماتریس حد نرمال شده

Cluster Node Labels	criteria							Goal
	C1	C2	C3	C4	C5	C6	C7	GOal
criteria	C1	0.252477	0.252477	0.252477	0.252477	0.252477	0.252477	0.252477
	C2	0.197238	0.197238	0.197238	0.197238	0.197238	0.197238	0.197238
	C3	0.151593	0.151593	0.151593	0.151593	0.151593	0.151593	0.151593
	C4	0.147217	0.147217	0.147217	0.147217	0.147217	0.147217	0.147217
	C5	0.091570	0.091570	0.091570	0.091570	0.091570	0.091570	0.091570
	C6	0.075660	0.075660	0.075660	0.075660	0.075660	0.075660	0.075660
	C7	0.084244	0.084244	0.084244	0.084244	0.084244	0.084244	0.084244
Goal	GOal	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

نتایج نهایی رتبه‌بندی معیارها در این پژوهش به شرح جدول ۱۸ می‌باشد:

جدول ۱۸. رتبه‌بندی معیارها

رتبه نهایی	وزن نرمال	شرح شاخص	نماد
۱	۰,۲۵۲۴۸	شاخص افزایش بهره‌وری سازمان	C۱
۲	۰,۱۹۷۲۴	مدیریت صحیح منابع	C۲
۳	۰,۱۵۱۵۹	میزان رضایتمندی کاربر	C۳
۴	۰,۱۴۷۲۲	انطباق عملیاتی نرم‌افزار	C۴
۵	۰,۰۹۱۵۷	پویایی واحد ها	C۵
۷	۰,۰۷۵۶۶	کاهش هزینه‌های تکنولوژی اطلاعات	C۶
۶	۰,۰۸۴۲۴	توانمندسازی سازمان در توسعه	C۷

با توجه به ارزیابی صورت پذیرفته از شاخص‌های ورودی و خروجی جهت ارزیابی عملکرد برنامه‌ریزی منابع سازمان در شرکت‌های مورد مطالعه، به جمع‌آوری اطلاعات سازمان‌ها پرداخته شده و اطلاعات بدست آمده در هر سازمان براساس شاخص‌های ورودی و خروجی به شرح زیر است:

همان‌گونه که در جدول (۱۸) مشاهده می‌گردد، شاخص‌های افزایش بهره‌وری، مدیریت صحیح منابع و میزان رضایتمندی کاربر در رتبه‌های اول تا سوم قرار گرفته و بعنوان شاخص‌های خروجی در بررسی عملکرد ERP در سازمان مورد استفاده واقع شده‌اند.

#### ۴\_۴ رتبه‌بندی عملکرد ERP در شرکت‌های مورد مطالعه

ارائه یک روش ترکیبی از DEA، DEMATEL و ANP جهت ارزیابی سیستم‌های ERP

جدول ۱۹. اطلاعات جمع‌آوری شده ورودی‌ها و خروجی‌ها

ردیف	DMU	شاخص ورودی رتبه اول(ماه)	شاخص ورودی رتبه دوم(هزینه)	شاخص ورودی رتبه سوم نمره±۱۰	شاخص خروجی رتبه اول(درصد)	شاخص خروجی رتبه دوم نمره ±۱۰	شاخص خروجی رتبه سوم نمره ±۱۰
۱	هیرید زرنديه	۱۳	۱۵۰۰	۵	۲۰	-۱	۳
۲	بازرگانی نیلپر	۱۲	۱۳۰۰	۹	۵۰	۵	۵
۳	مازینور	۱۸	۱۹۰۰	۹	۷۰	۷	۸
۴	کابل ابهر	۲۴	۱۷۰۰	۶	۳۰	-۳	۶
۵	ساینا	۱۵	۱۴۰۰	۷	۵۰	۶	۷
۶	پاسارگاد	۱۳	۱۰۰۰	۵	۲۰	-۴	۷
۷	پارس حیات	۱۲	۱۵۰۰	۴	۱۰	-۲	۶
۸	منور	۱۹	۱۸۰۰	۸	۶۰	-۲	۸
۹	ترانسفو	۱۲	۱۳۰۰	۵	۳۰	۳	۳
۱۰	A	۱۵	۱۰۰۰	۲	۱۰	۳	۳
۱۱	B	۱۱	۱۲۰۰	۵	۱۵	۴	۱
۱۲	C	۱۶	۱۵۰۰	۵	۴۵	-۱	۵
۱۳	D	۲۰	۱۶۰۰	۴	۳۰	۴	۴
۱۴	E	۲۵	۱۶۰۰	۹	۴۰	۵	۵
۱۵	F	۲۷	۱۷۰۰	۳	۲۵	۶	۲
۱۶	G	۱۴	۱۲۰۰	۷	۲۰	۴	۱
۱۷	H	۱۷	۱۵۰۰	۶	۳۵	۵	۵
۱۸	I	۲۰	۱۷۰۰	۸	۳۰	۶	۴

پس از جمع‌آوری اطلاعات از شرکت‌های مورد مطالعه با مصاحبه با خبرگان، بوسیله نرم‌افزار GAMS به ارزیابی عملکرد سیستم‌های ERP که نتایج بررسی به شرح جدول ۲۰ می‌باشد:

با استفاده از روش تحلیل پوششی داده‌ها با رویکرد SORM پرداخته شده

جدول ۲۰. محاسبه کارایی سیستم ERP شرکت‌های تولیدی

اولویت	DMU	Efficiency
۱	پارس حیات	۱
	فولاد پاسارگاد	۱
	A	۱
	B	۱
۲	ایران ترانسفو	۰/۹۴
۳	نیلپر	۰/۹۲
۴	هیرید زرنديه	۰/۸۹
۵	G	۰/۸۸
۶	ساینا	۰/۷۹
۷	C	۰/۷۷
۸	H	۰/۷۳
۹	D	۰/۶۹
۱۰	F	۰/۶۶
۱۱	سپهر منور	۰/۶۳
۱۲	مازینور- I-E	۰/۶۲
۱۳	کابل ابهر	۰/۵۸

با توجه به ارزیابی انجام شده، سیستم ERP شرکت‌های پارس حیات، فولاد پاسارگاد، A، B، دارای بیشترین کارایی (یک) هستند و شرکت کابل ابهر ناکارترین شرکت در این پژوهش می‌باشد. با توجه به اینکه مدل تحلیل پوششی داده‌های مورد استفاده در این پژوهش ورودی محور می‌باشد، نتایج حاصله نشانگر این است که شرکت‌های ناکارا می‌توانند با تغییراتی در میزان ورودی‌های سیستم ERP خود، کارایی این سیستم را بهبود بخشند.

## ۵. جمع‌بندی و نتیجه‌گیری

از مهم‌ترین عوامل موفقیت و کسب مزیت رقابتی برای سازمان‌ها در شرایط پیچیده و در حال تحول امروزی استفاده از سیستم‌های اطلاعاتی است. یکی از بروزترین دستاوردهای فناوری اطلاعات، سیستم‌های برنامه‌ریزی منابع سازمان است که توجه سازمان‌های پیشرو در استفاده از فناوری اطلاعات را به خود معطوف نموده و باعث بوجود آمدن بازار بزرگی در ارائه این محصول شده است. ERP بسته‌ای نرم‌افزاری است که به سازمان‌ها کمک می‌کند تا کارایی و اثربخشی استفاده از منابعش را ارتقا دهد. این سیستم‌ها ارتباطات فرآیندهای تجاری و بکارگیری نرم‌افزای آن‌ها را تسهیل می‌نمایند و یکپارچگی و امنیت را در کلیه بخش‌های کسب‌وکار موجب می‌شوند. مزیت دیگر آن‌ها قابلیت ارائه خدمات به کلیه سازمان‌ها با هر حجم فعالیت، اندازه و گستردگی جغرافیایی است. به تاسی از کشورها و سازمان‌های پیشرو، صنایع تولیدی و خدماتی کشورمان در سال‌های اخیر در صدد بکارگیری فناوری‌های نوین بوده‌اند و نمونه‌های متنوعی از پروژه‌های برنامه‌ریزی منابع سازمان ارائه گردیده است. لذا آسیب‌شناسی عوامل موفقیت این پروژه‌ها در کانون توجه کارشناسان و برنامه‌ریزان و مدیران قرار گرفته است. پژوهش‌های پیشین بیشتر به چگونگی استقرار، دلایل شکست و مباحث نرم‌افزاری ERP پرداخته‌اند و ارزیابی الگوهای مناسب جهت عملکرد بهینه و نحوه مدیریت آن کمتر مورد توجه قرار گرفته است. به منظور پوشش شکاف تحقیقاتی مذکور، پژوهش حاضر با هدف بررسی دقیق سیستم‌های ERP در ۱۸ نمونه واحدهای صنعتی و تولیدی انجام شده تا مشخص شود آیا اهداف تعریف شده برای سیستم‌های ERP توانسته‌اند عملکرد بعد از پیاده‌سازی آن را تحت تاثیر قرار دهند و در پایان سیاست‌های بهبود اجرای آن‌ها معرفی شود. به منظور حفظ تقریبی ماهیت

استدلال‌های انسانی و بهبود بهره‌وری تصمیم‌گیری از منطق فازی بهره گرفته شده است. بدین منظور، مهم‌ترین شاخص‌های مؤثر در پیاده‌سازی برنامه‌ریزی منابع سازمان شناسایی و با استفاده از نظر خبرگان به روش دیمتل فازی وزن‌دهی شده‌اند. همچنین مهم‌ترین شاخص‌های خروجی نیز به روش فرآیند تحلیل شبکه‌ای رتبه‌بندی گردیدند. سه شاخص زمان صرف شده جهت اجرا، زیر ساخت‌های اجرا، آموزش و اقدامات حمایتی کاربر، شاخص‌های منتخب ورودی، و شاخص‌های افزایش بهره‌وری، مدیریت صحیح منابع و میزان رضایت‌مندی کاربر نیز به عنوان شاخص‌های برتر خروجی انتخاب گردیدند. در پایان، با استفاده از روش تحلیل پوششی داده‌ها و براساس شاخص‌های منتخب، کارایی این سیستم‌ها مورد ارزیابی قرار گرفته است. رتبه‌بندی این شاخص‌ها می‌تواند کمک شایانی به سازمان‌های تولیدی که تصمیم به پیاده‌سازی سیستم ERP دارند، در جهت اجرای موفقیت آمیز این سیستم کرده و در سازمان‌هایی که این سیستم را پیاده سازی کرده‌اند در راستای ارزیابی و بهینه‌سازی سیستم بکار گرفته شود. پس از تعیین شاخص‌های ورودی و خروجی سیستم ERP، ارزیابی عملکرد این سیستم با استفاده از روش تحلیل پوششی داده‌ها با رویکرد SORM، که رویکرد نوینی در جهت ارزیابی عملکرد سیستم‌ها با داده‌های منفی باشد، انجام گرفته و شرکت‌های کارا در این زمینه مشخص گردید که شامل چهار شرکت پارس حیات، فولاد پاسارگاد، A و B می‌باشد. و سایر شرکت‌های مورد بررسی فاقد سیستم ERP کارا بودند.

با مقایسه معیارهای شناسایی شده توسط صارمی و عابدینی (۱۳۸۶) با معیارهای تحقیق حاضر، اقدامات حمایتی و رضایت‌مندی کاربر، زیرساخت فناوری، و بهره‌وری و مدیریت صحیح منابع (در قالب توانمندی توانمندی سازمان) مورد تأیید واقع شد [۳۹]. همچنین پژوهش محمودی و احمدی (۱۳۹۰)، نیز اهمیت معیارهایی چون زیرساخت مناسب و آموزش را تأیید می‌نماید [۴۰].

علاوه بر این یافته‌های این تحقیق با تحقیقات مشابه خارجی از جمله مالدونادو (۲۰۰۹) و گانش و مهتا (۲۰۱۰) که عوامل موفقیت پروژه‌های ERP را در آمریکای جنوبی و هند مورد بررسی قرار داده‌اند، اشتراکات زیادی دارد [۴۱ و ۴۲].

می‌توان ارائه داد، به شرح زیر می‌باشد:

- \* پیشنهاد می‌شود دسته‌بندی و رتبه‌بندی شاخص‌های موفقیت پیاده‌سازی ERP در چارچوب حوزه‌های نه‌گانه مدیریت پروژه طبق استاندارد دانش مدیریت پروژه<sup>۲۰</sup> انجام شود.
- \* پیشنهاد می‌شود شناسایی و دسته‌بندی عوامل مؤثر بر پیاده‌سازی موفقیت‌آمیز ERP با ترکیب روش‌های ارزیابی مانند کارت امتیازی متوازن انجام شود.
- \* پیشنهاد می‌شود جهت افزایش دقت بررسی کارایی از مدل‌های تحلیل پوششی داده‌های فازی استفاده شود.
- \* پیشنهاد می‌شود به ارزیابی عملکرد سیستم‌های ERP استاندارد و سیستم‌های ERP بومی‌سازی شده بطور جداگانه پرداخته و کارایی این دو سیستم با یکدیگر مقایسه شود.

## ۱.۵ محدودیت‌های پژوهش

- انجام پژوهش با محدودیت‌ها و مسایل مختلفی روبه‌رو بوده است که شامل دو محدودیت محقق و محدودیت‌های خارج از کنترل محقق می‌باشد که در ادامه به آن‌ها اشاره می‌شود:
- ۱- از محدودیت‌های تحقیق می‌توان به محدودیت‌های جمع‌آوری برخی داده‌های پژوهش با استفاده از مصاحبه اشاره نمود. محدودیت‌هایی مانند: محدودیت ذاتی مصاحبه، عدم امکان بررسی کامل میزان دقت پاسخ‌دهندگان و عدم تمایل بعضی از پاسخ‌دهندگان به همکاری.
  - ۲- عدم دسترسی به اطلاعات و سوابق گذشته سازمان‌های مورد مطالعه.

## ۲.۵ پیشنهاد برای پژوهش‌های آتی

پیشنهاداتی که با توجه به پژوهش انجام شده و مطالعات صورت گرفته

## مراجع

- [۷]. Bhumgara A., Sayyed I., Enterprise Resource Planning Systems, International Journal of Advances in Engineering & Technology, ۱۰(۲), ۲۰۱۷, ۲۸۳.
- [۸]. رضایی هانیه، سعید جهانیان، بررسی تاثیر اجرای مازل خرید ERP بر مدیریت زنجیره تامین مورد مطالعه: سیستم خرید شرکت ذوب آهن اصفهان، دومین کنفرانس بین‌المللی مدیریت و حسابداری، موسسه آموزش عالی صالحان، تهران، ایران، ۱۳۹۶.
- [۹]. امیری فرزاد، فیضی کمره حدیث، خامه‌ای منا، نظری الهام، نقش مدیریت اطلاعات (IM) و سیستم مدیریت منابع (ERP) در مدیریت دانش استراتژیک و تصمیم‌گیری، دومین کنفرانس ملی مدیریت مهندسی، موسسه آموزش عالی مهرآستان گیلان، آستانه اشرفیه، ایران، ۱۳۹۶.
- [۱۰]. Saade R. G., Nijher H., Critical success factors in enterprise resource planning implementation: A review of case studies, Journal of Enterprise Information Management, ۲۹(۱), ۷۲-۹۶, ۲۰۱۶.
- [۱۱]. Costa C. J., Ferreira E., Bento F., Aparicio, M., Enterprise resource planning adoption and satisfaction determinants, Computers in Human Behavior, ۶۳, ۶۵۹-۶۷۱, ۲۰۱۶.
- [۱۲]. رسولی فرزاد، طبقه‌بندی عوامل مهم و حیاتی برنامه ریزی منابع سازمانی، همایش پژوهش‌های مدیریت و علوم انسانی، موسسه
- [۱]. امینی، علی، بنیادی نایینی، علی؛ محمدی، مهدی احذزاده نمین، مهناز، ارزیابی عملکرد مدیریت فناوری و نوآوری در دو مرحله توانمندسازها و نتایج با استفاده از تحلیل پوششی داده‌ها، مدیریت توسعه فناوری، ۴، ۳-۳۳، ۶۸-۱۳۹۵.
- [۲]. Clayton M.J., Delphi: a technique to harness expert opinion for critical decision-making tasks in education, Educational Psychology, ۱۷(۴), ۳۷۳-۳۸۶, ۲۰۰۶.
- [۳]. محمودی جعفر، احمدی فریدون، تعیین عوامل کلیدی موفقیت ERP در سازمان‌های دولتی ایران، فصلنامه تجارت و اقتصاد نوین، شماره ۱۳، صفحه ۶۷-۷۸، ۱۳۹۳.
- [۴]. Tsai W.H., Lin S.J., Lin W.R., Liu J.Y., The relationship between planning & control risk and ERP project success, Proceedings of the IEEE IEEM, ۱۸۳۵-۱۸۳۹, ۲۰۰۹.
- [۵]. Kahraman C., Çebı S., A new multi-attribute decision making method: Hierarchical fuzzy axiomatic design, Expert Systems with Applications, ۳۶(۳), ۴۸۴۸-۴۸۶۱, ۲۰۰۹.
- [۶]. زارع رواسان، احد، دیلمی، هما، مهارت‌های موردنیاز مدیران پروژه‌های پیاده‌سازی سیستم‌های برنامه‌ریزی منابع سازمان، مطالعات مدیریت کسب و کار هوشمند، دوره ۲، شماره ۸، ۶۷-۸۸، ۱۳۹۳.

- ERP system, *Information & Management*, ۵۱(۲), ۱۸۷-۲۰۵, ۲۰۱۴.
- [۲۱]. Sudhaman P., Thangavel C., Efficiency analysis of ERP projects-software quality perspective, *International Journal of Project Management*, ۳۳(۴), ۹۶۱-۹۷۰, ۲۰۱۵.
- [۲۲]. Alaskari, O., Pinedo-Cuenca, R., & Ahmad, M. M., Framework for Selection of ERP System: Case Study. *Procedia Manufacturing*, ۳۸, ۶۹-۷۵, ۲۰۱۹.
- [۲۳]. Ruivo, P., Johansson, B., Sarker, S., & Oliveira, T., The relationship between ERP capabilities, use, and value. *Computers in Industry*, ۱۱۷, ۱۰۳۲-۹, ۲۰۲۰.
- [۲۴]. فیضی، کامران، ثابت مطلق، محمد، عابدینی نادینی، مهدی، به‌کارگیری رویکرد تلفیقی QFD، FAHP و VIKOR به‌منظور انتخاب مناسب‌ترین سامانه ERP، مطالعات مدیریت کسب و کار هوشمند، دوره ۳، شماره ۱۰، ۱-۲۰، ۱۳۹۳.
- [۲۵]. حسینی دهشیری، جلال‌الدین، عرب، علیرضا، انتخاب سیستم برنامه‌ریزی منابع سازمان با بهره‌گیری از روش تلفیقی سوارا و آراس خاکستری، مطالعات مدیریت کسب و کار هوشمند، دوره ۵، شماره ۱۸، ۷۳-۱۰۳، ۱۳۹۵.
- [۲۶]. محقر، علی؛ خلوصی، محمدهادی، خلوصی، سمیه، بهبود روش فناوری‌های کلیدی با استفاده از گسترش عملکرد کیفیت، تحلیل پوششی داده‌ها و منطق فازی، مدیریت توسعه فناوری، ۴، ۴، ۱۰۷-۱۳۴، ۱۳۹۵.
- [۲۷]. Wu, W. W. & Lee, Y. T., Developing global managers' competencies using the fuzzy DEMATEL method, *Expert Systems with Applications*, ۳۲(۲), ۴۹۹-۵۰۷, ۲۰۰۷.
- [۲۸]. Ozer U., Hassan K., Unal A., An integrated DWMATEL and Fuzzy ANP techniques for evaluation and selection of outsourcing provider for a telecommunication company, *Computers & Industrial Engineering*, ۸۶, ۱۳۷-۱۴۶, ۲۰۱۵.
- [۲۹]. Chang B., Chang, C.W., & Wu, CH., Fuzzy DEMATEL method for developing supplier selection criteria, *Expert system with Applications*, ۳۸(۳), ۱۸۵۰-۱۸۵۸, ۲۰۱۱.
- پژوهشی مدیریت مدبر، دانشگاه تهران، تهران، ایران، ۱۳۹۵.
- [۱۳]. Rashid, M. A., Hossain, L., & Patrick, J. D., The evolution of ERP systems: A historical perspective, In *Enterprise Resource Planning: Solutions and Management* (pp. ۳۵-۵۰). IGI Global, ۲۰۱۲.
- [۱۴]. Muscatello J. R., Small, M. H., and Chen, I. J., Implementing Enterprise Resource Planning (ERP) Systems in Small and Midsize Manufacturing Firms, *International Journal of Operations and Management*, ۲۳(۸), ۸۵۰-۸۷۱, ۲۰۰۳.
- [۱۵]. Rosenthal P. H., and Park, L. J., Managing Information Systems Textbooks: Assessing their Orientation Toward Potential General Managers, California State University, Losangeles, California, ۲۰۰۹.
- [۱۶]. Lall V., Teyarachakul S., Enterprise resource planning (ERP) system selection: a data envelopment analysis (DEA) approach, *Journal of Computer Information Systems*, ۴۷(۱), ۱۲۳-۱۲۷, ۲۰۱۶.
- [۱۷]. Kofahi I., Alryalat H., Enterprise Resource Planning (ERP) Implementation Approaches and the Performance of Procure-to-Pay Business Processes: (Field Study in Companies that Implement Oracle ERP in Jordan), *International Journal of Information Technology Project Management (IJITPM)*, ۸(۱), ۵۵-۷۱, ۲۰۱۷.
- [۱۸]. Zhu, Y., Li, Y., Wang, W., & Chen, J., What leads to post-implementation success of ERP? An empirical study of the Chinese retail industry, *International Journal of Information Management*, ۳۰(۳), ۲۶۵-۲۷۶, ۲۰۱۰.
- [۱۹]. Samoilenko, S., & Osei-Bryson, K. M., Using Data Envelopment Analysis (DEA) for monitoring efficiency-based performance of productivity-driven organizations: Design and implementation of a decision support system. *Omega*, ۴۱(۱), ۱۳۱-۱۴۲, ۲۰۱۳.
- [۲۰]. Chang S. I., Yen, D. C., Chang, I. C., & Jan D., Internal control framework for a compliant

- analysis, *Management science*, ۳۰(۹), ۱۰۷۸-۱۰۹۲, ۱۹۸۴.
- [۳۸]. Alinezhad, A., Kakavand, H., Kazemi, A., & Javad, M. O. M., Developing a multi-objective BCC model in grey environment. *International Journal of Modelling in Operations Management*, ۵(۱), ۷۲-۸۲, ۲۰۱۵.
- [۳۹]. صارمی، محمود، عابدینی، مهدی، ارزیابی و استخراج شاخص‌های مرتبط با آمادگی صنعت خودروسازی جهت پیاده‌سازی ERP، دانش مدیریت، ۷۷ (۲۰)، ۱۵-۳۱، ۱۳۸۶.
- [۴۰]. محمودی، جعفر، احمدی، فریدون، تعیین عوامل کلیدی موفقیت تعیین عوامل کلیدی موفقیت ERP در سازمان‌های دولتی ایران، فصلنامه اقتصاد و تجارت نوین، ۳ (۱۳)، ۶۷-۸۸، ۱۳۹۰.
- [۴۱]. Maldonado, M., Factors impacting the success of ERP implementations in small and medium enterprises: an empirical assessment from Latin America. In Proceedings of the Fourteenth Americas Conference on Information Systems, July, ۲۰۰۹.
- [۴۲]. Ganesh, L., & Mehta, A., Critical success factors for successful enterprise resource planning implementation at Indian SMEs. *International Journal of Business, Management and Social Sciences*, ۱(۱), ۶۵-۷۸, ۲۰۱۰.
- [۳۰]. Gulcin B., Gizem C., A novel hybrid MCDM approach based on fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS to evaluate green suppliers, *Expert Systems with Applications* ۳۹.۳۰۰۰.۳۰۱۱, ۲۰۱۲.
- [۳۱]. Ervural, B. C., Zaim, S., Demirel, O. F., Aydin, Z., & Delen, D., An ANP and fuzzy TOPSIS- based SWOT analysis for Turkey's energy planning, *Renewable and Sustainable energy reviews*, ۱۸۲, ۱۵۳۸-۱۵۵۰, ۲۰۱۸.
- [۳۲]. Saaty, T. L., Fundamentals of the analytic network process, in *proceedings of the ۵<sup>th</sup> international symposium on the analytic hierarchy process*, ۱۲-۱۴, ۱۹۹۹.
- [۳۳]. Farrell, M. J., The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Vol. ۱۲۰, pp.۲۵۳-۲۹۰, ۱۹۵۷.
- [۳۴]. Charnes, A., Cooper, W. W., & Rhodes, E., Measuring the efficiency of decision making units, *European journal of operational research*, ۲(۶), ۴۲۹-۴۴۴, ۱۹۷۸.
- [۳۵]. مهرگان، محمدرضا. تحلیل پوششی داده‌ها: مدل‌های کمی در ارزیابی عملکرد سازمان‌ها، نشر کتاب دانشگاهی، چاپ سوم، تهران، ۱۳۹۵.
- [۳۶]. معین‌الدینی سعید، عادل آذر، (۱۳۹۳)، ارزیابی عملکرد سازمان بر اساس شاخص‌های جوایز کیفیت - رویکرد DEA، دو ماهنامه علمی - پژوهشی دانشگاه شاهد، سال یازدهم، شماره ۸، ۱۳۹۳.
- [۳۷]. Banker, R. D., Charnes, A., & Cooper, W. W., Some models for estimating technical and scale inefficiencies in data envelopment

## بهینه‌سازی توأم اختصاص توان در یک شبکه مخابرات همکارانه با استفاده از دسترسی چندگانه غیرمتعامد

حمید امیری آرا\* محمدباقر مسرور\*\* محمدرضا ذهابی\*\*\*

\* دکتري، دانشکده مهندسی برق و کامپیوتر - دانشگاه صنعتی نوشیروانی بابل

\*\* کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر - دانشگاه صنعتی نوشیروانی بابل

\*\*\* استادیار، دانشکده مهندسی برق و کامپیوتر - دانشگاه صنعتی نوشیروانی بابل

تاریخ دریافت: ۱۳۹۹/۰۲/۱۳ تاریخ پذیرش: ۱۳۹۹/۰۹/۱۷

نوع مقاله: پژوهشی

### چکیده

در این مقاله یک شبکه فراسو با دو کاربر به ترتیب در نقش فرستنده و رله، و یک ایستگاه مرکزی به عنوان گیرنده در نظر گرفته شده است. هدف این است که ضریب بهینه سمبل‌های سیگنال غیرمتعامد و اختصاص توان بهینه در منبع-رله به منظور بیشینه کردن متوسط مجموع نرخ در یک سیستم مخابرات همکارانه با استفاده از تکنیک دسترسی چندگانه غیرمتعامد، تعیین شود. برای دستیابی به این اهداف، ابتدا متوسط مجموع نرخ سیستم مخابرات همکارانه با استفاده از دسترسی چندگانه غیرمتعامد با رله کدگشایی و ارسال، در کانال مستقل رایلی محاسبه گردید. سپس در گام اول، مسئله بهینه‌سازی ضریب سمبل‌های سیگنال غیرمتعامد در این سیستم به ازای هر اختصاص توانی به صورت ریاضی بیان شده و رابطه بسته برای حل تقریبی آن پیشنهاد شده است. در گام دوم، مسئله بهینه‌سازی اختصاص توان بهینه ارسالی از منبع-رله، به ازای ضرایب سیگنال مشخص معرفی و حل گردید. در انتها، مسئله بهینه‌سازی توأم ضریب سمبل‌های سیگنال غیرمتعامد و اختصاص توان بررسی می‌شود و الگوریتمی برای بهینه‌سازی توأم این دو پارامتر پیشنهاد شده است. الگوریتم پیشنهادی این مقاله نشان می‌دهد که بهینه‌سازی توأم ضریب سمبل‌های سیگنال غیرمتعامد و اختصاص توان منجر به دستیابی به متوسط نرخ بالاتری نسبت به بهینه‌سازی جداگانه هر یک از این پارامترها می‌شود. همچنین، شبیه‌سازی و نتایج عددی برای تأیید روابط تئوری ارائه گردیده است، که شبیه‌سازی‌ها بهره  $3 \text{ dB}$  را برای سیستم بهینه شده توأم با استفاده از الگوریتم پیشنهادی در مقایسه با سیستم غیر بهینه نشان می‌دهند.

**واژگان کلیدی:** مخابرات همکارانه، دسترسی چندگانه غیرمتعامد، بهینه‌سازی، اختصاص توان.

ظرفیت سیستم را افزایش دهد. راه حل استفاده از گوناگونی بهره‌گیری از سیستم مخابرات همکارانه (CRS) است.

در مراجع [۱]، [۲] نشان داده شده است که سیستم مخابرات همکارانه با یک رله نیمه‌دوسویه که در آن مقصد دو کپی از یک سیگنال یکسان را در دو فاز مختلف از منبع و رله، دریافت و ترکیب می‌کند، باعث بهبود در کیفیت سیگنال دریافتی می‌شود. با

### ۱. مقدمه

گوناگونی همکارانه فناوری شاخصی در شبکه‌های بیسیم است که می‌تواند با محوشدگی مقابله کند، ناحیه پوششش را گسترش و

نویسنده مسئول: محمد رضا ذهابی zahabi@nit.ac.ir



ترکیب NOMA و کدهای فضا-زمان الموتی به نتایج قابل توجهی دست پیدا کردند. در این طرح برخلاف طرح‌های قبلی از چند آنتن در فرستنده و رله استفاده شده است، و نویسندگان توانستند رابطه بسته برای مجموع ظرفیت ارگودیک و احتمال قطع بدست آورند. شبیه‌سازی‌ها نشان می‌دهد که این طرح در سیگنال به نویز متوسط و بالا، به نتایج بهتری نسبت به آنچه که در مراجع [۴] و [۱۱] انجام شده است کسب کرده است.

با توجه به دانش نویسندگان، هیچ مرجع قبلی تاکنون به بهینه‌سازی توأم ضرایب سمبل‌های سیگنال متعامد و اختصاص توان به منبع-رله در سیستم‌های همکارانه با استفاده از NOMA با هدف بیشینه‌سازی مجموع نرخ قابل دستیابی پرداخته است. در مقاله حاضر، علاوه بر بهینه‌سازی ضرایب سمبل‌های سیگنال متعامد در تکنیک NOMA، بهینه‌سازی اختصاص توان برای منبع-رله در سیستم همکارانه نیز انجام گرفته است. با کمک روابط بسته بدست آمده در این مقاله، این امکان فراهم شده تا بهینه‌سازی توأم در سیستم همکارانه با استفاده از NOMA صورت گیرد که این باعث بهبود متوسط مجموع نرخ<sup>۱۰</sup> (ASR) سیستم نسبت به کارهای گذشته شده است.

در ادامه مقاله به صورت زیر سازماندهی شده است:

- بخش دوم مدل سیستم را معرفی می‌کند.
- در بخش سوم، متوسط مجموع نرخ برای شبکه مخابرات همکارانه با استفاده از NOMA در حالت فرا سو، در کانال رایلی محاسبه گردید.
- در بخش چهارم، به بهینه‌سازی توأم ضرایب سمبل‌های سیگنال متعامد و اختصاص توان منبع-رله برای رسیدن به حداکثر مجموع نرخ پرداخته شده است.
- نتایج و بررسی آن‌ها در بخش پنجم قرار داده شده است.

## ۲. مدل سیستم

یک سیستم رله همکارانه فرا سو<sup>۱۱</sup> (U-CRS) مطابق شکل (۱) شامل یک کاربر به عنوان منبع (S)، کاربری دیگر در نقش رله نیمه دوسویه (R)، و یک ایستگاه پایه به عنوان مقصد (D) را در نظر بگیرید.

در این مقاله فرض شده تمام مسیرها (منبع-رله، منبع-مقصد و رله-مقصد) قابل دسترسی باشند. ضرایب کانال در مسیرهای S-D، S-R و R-D به ترتیب با  $h_{SR}$ ،  $h_{SD}$  و  $h_{RD}$  نام‌گذاری و این ضرایب،

این‌حال، این نوع ارسال همکارانه رله‌ای، راندمان طیفی سیستم را به خاطر انتقال در دو فاز جداگانه محدود می‌کند. مرجع [۳] یک سیستم دسترسی چندگانه غیرمتعامد<sup>۱</sup> (NOMA) فراسو<sup>۲</sup> را پیشنهاد می‌کند، که در آن کاربران بسته‌های خود را با استفاده از کد فضا-زمان ارسال می‌کنند تا در گیرنده بهره‌برداری کامل را از گوناگونی فضایی<sup>۳</sup> ببرند. در مرجع [۴] یک سیستم همکارانه با NOMA ارائه شده است که با ارسال چندگانه سیگنال‌ها، بهره‌وری طیفی را بهبود بخشد. تکنیک NOMA به فرستنده اجازه می‌دهد که چندین سیگنال اطلاعات را به طور هم‌زمان به چند گیرنده که در همکاری با هم هستند برساند، بنابراین در مخابرات همکارانه به جهت افزایش کارایی طیف می‌توان توأم دو سیگنال داده را از منبع ارسال کرد [۵]. در این طرح، اختصاص توان در منبع برای دو سیگنال انجام می‌شود و با کدینگ برهم‌نهی<sup>۴</sup> برای مقصد و رله ارسال می‌شوند. در فاز اول سیگنال با توان بالاتر با توجه به سیگنال به نویز بالایی که دارد در مقصد آشکار سازی می‌شود و در رله ابتدا سیگنال با توان بالا آشکار سازی می‌شود و سپس سیگنال با توان پایین با استفاده از حذف تداخل پیاپی<sup>۵</sup> (SIC) آشکار سازی می‌شود. در فاز دوم سیگنال با توان پایین تر آشکار سازی شده در رله برای مقصد ارسال می‌شود. بنابراین در دو فاز دو سیگنال از منبع به مقصد ارسال می‌شود.

در مرجع [۶]، یک رابطه بسته برای محاسبه احتمال قطع شبکه NOMA همکارانه تقویت و ارسال<sup>۶</sup> (AF) پیشنهاد شده است. در مرجع [۷]، احتمال قطع و ظرفیت مجموع ارگودیک در یک سیستم NOMA همکارانه با رله رمزگشایی و ارسال<sup>۷</sup> (DF) اندازه‌گیری شد. در مراجع [۸] و [۹] یک استراتژی انتخاب رله برای بررسی عملکرد احتمال قطع و مرتبه چندگانگی سیستم همکارانه NOMA پیشنهاد شده است. در این روش کاربر دومی را با شرایط بهتر کانال به عنوان یک رله همکار DF برای انتقال اطلاعات با روش NOMA انتخاب می‌کنند. در مرجع [۱۰]، مفهوم NOMA همکارانه در مخابرات سلولی با کمک کاربر برای تقویت قابلیت اطمینان ارتباطات در لبه سلول ارائه شده است. در مرجع [۱۱] نویسندگان، گیرنده استفاده شده در مرجع [۴] را بهبود بخشیدند. در این مدل سیستم گیرنده می‌تواند به طور توأم هر دو سمبل ارسالی از مبدأ را به وسیله ترکیب‌کننده بیشینه نرخ<sup>۸</sup> (MRC) و کدگشایی کند. این طرح نسبت به طرح مطرح شده در مرجع [۴] دارای ظرفیت ارگودیک بهتر و احتمال قطع<sup>۹</sup> پایین‌تری است. در مرجع [۱۲] نویسندگان با

<sup>۷</sup> Decode and forward

<sup>۸</sup> Maximum Ratio Combining

<sup>۹</sup> Outage Probability

<sup>۱۰</sup> Average sum rate

<sup>۱۱</sup> Uplink-Cooperative Relay System

<sup>۱</sup> Non Orthogonal Multiple Access

<sup>۲</sup> Uplink

<sup>۳</sup> Space diversity

<sup>۴</sup> Super Position Coding

<sup>۵</sup> Signal Interference Cancellation

<sup>۶</sup> Amplify and forward

$$\gamma_{SD} = \frac{|h_{SD}|^2 a_1 b P_T}{|h_{SD}|^2 a_2 b P_T + \sigma^2} \quad (5)$$

بیان کرد. در فاز زمانی دوم، رله سمبل کدگشایی شده  $S_2$  را با توان  $(1-b)P_T$  برای مقصد ارسال می‌کند. فرض بر این است که رله به درستی سمبل  $S_2$  را کدگشایی می‌کند، سیگنال دریافتی در مقصد در فاز زمانی دوم به صورت

$$r_{RD} = h_{RD} \sqrt{(1-b)P_T} S_2 + n_{RD} \quad (6)$$

بیان می‌شود، که در آن نویز سفید گوسی اضافه‌شونده با واریانس  $\sigma^2$  است. سیگنال به نویز دریافتی برای سمبل  $S_2$  در رابطه (۶) برابر با

$$\gamma_{RD} = \frac{|h_{RD}|^2 (1-b)P_T}{\sigma^2} \quad (7)$$

می‌شود. در نتیجه در مخابرات همکارانه با NOMA، مقصد می‌تواند دو سمبل  $S_1$  و  $S_2$  را در دو بازه زمانی دریافت کند، بر خلاف مخابرات همکارانه معمول که در آن گیرنده در دو بازه زمانی تنها یک سیگنال داده را دریافت می‌کند.

### ۳. محاسبه متوسط مجموع نرخ

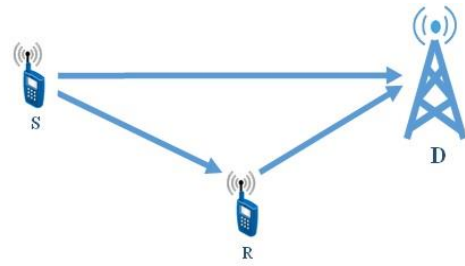
عبارت  $\lambda \triangleq |h_{RD}|^2$ ،  $\lambda_{SR} \triangleq |h_{SR}|^2$ ،  $\lambda_{SD} \triangleq |h_{SD}|^2$  و  $\rho \triangleq \frac{P_T}{\sigma^2}$  را در نظر بگیرید که  $\rho$  سیگنال به نویز ارسالی می‌باشد. توجه شود که  $\frac{bP_T}{\sigma^2}$  سیگنال به نویز ارسالی در منبع و  $\frac{(1-b)P_T}{\sigma^2}$  سیگنال به نویز ارسالی در رله می‌باشد. همچنین واریانس ضرایب کانال نیز بصورت  $\beta_{SD} = E|h_{SD}|^2$ ،  $\beta_{SR} = E|h_{SR}|^2$  و  $\beta_{RD} = E|h_{RD}|^2$  بیان می‌شود. از آنجایی که سمبل  $S_1$  باید هم در مقصد و هم با روش SIC در رله کدگشایی شود، نرخ قابل دستیابی برای سمبل  $S_1$  از طریق روابط (۳) و (۵) به صورت

$$\begin{aligned} C_{S_1} &= \frac{1}{\gamma} \min\{\log_2(1 + \gamma_{SD}), \log_2(1 + \gamma_{SR}^1)\} \\ &= \frac{1}{\gamma} \log_2 \left\{ 1 + \frac{\min\{\lambda_{SD}, \lambda_{SR}\} (1 - a_2) b \rho}{\min\{\lambda_{SD}, \lambda_{SR}\} b \rho a_2 + 1} \right\} \\ &= \frac{1}{\gamma} \log_2 \{1 + \min\{\lambda_{SD}, \lambda_{SR}\} b \rho\} - \\ &\quad \frac{1}{\gamma} \log_2 \{1 + \min\{\lambda_{SD}, \lambda_{SR}\} a_2 \rho\} \end{aligned} \quad (8)$$

قابل محاسبه است. باید به این نکته توجه شود که خط دوم رابطه (۸) بر اساس افزایشی بودن  $\gamma_{SD}$  و  $\gamma_{SR}$  نوشته شده است. همچنین در مرحله آخر از  $a_2 = 1 - a_1$  استفاده شده است. بر این اساس که نرخ نقطه به نقطه رله کدگشایی و ارسال بر مبنای ضعیف‌ترین مسیر استوار است و با استفاده از روابط (۴) و (۷) نرخ قابل دستیابی برای سمبل  $S_2$  برابر

$$\begin{aligned} C_{S_2} &= \frac{1}{\gamma} \min\{\log_2(1 + \gamma_{SR}), \log_2(1 + \gamma_{RD})\} \\ &= \frac{1}{\gamma} \log_2(1 + \min\{\lambda_{SR} a_2 b \rho, \lambda_{RD} (1-b) \rho\}) \end{aligned} \quad (9)$$

می‌شود. از روابط (۸) و (۹)، مجموع نرخ قابل دستیابی CRS با استفاده از NOMA بدست می‌آید:



شکل ۱. مدل سیستم مخابرات همکارانه پایه در حالت فراسو شامل یک کاربر منبع (S)، یک کاربر به عنوان رله (R) و ایستگاه پایه به عنوان مقصد (D)

متغیرهای تصادفی رایلی با واریانس  $\beta_{SD}$ ،  $\beta_{SR}$  و  $\beta_{RD}$  در نظر گرفته می‌شوند. همچنین به خاطر تأثیرات تلفات مسیر و سایه‌افکنی که مسیر D-S را ضعیف‌تر از مسیر R-S می‌کند، شرایط  $\beta_{SD} < \beta_{SR}$  برقرار است.

توان کل ارسال را با  $P_T = P_S + P_R$  نمایش می‌دهیم و توان منبع و رله به ترتیب با  $P_S = bP_T$  و  $P_R = (1-b)P_T$  قابل بیان است که در آن  $b$  ضریب اختصاص توان منبع و رله است. در CRS با استفاده از NOMA،

نرخ  $\sqrt{a_1 b P_T} S_1 + \sqrt{a_2 b P_T} S_2$  را با توجه به استاندارد NOMA در حین بازه زمانی اول، برای رله و مقصد ارسال می‌کند. که در آن  $S_i$   $i$  امین سمبل اطلاعات،  $E[|S_i|^2] = 1$  و ضریب اختصاص توان برای سمبل‌ها می‌باشد. توجه داشته باشید که مطابق فرضیات NOMA شرایط  $a_1 > a_2$  و  $a_1 + a_2 = 1$  برقرار است.

سیگنال دریافتی در رله و مقصد در بازه زمانی اول عبارت است از:

$$r_{SR} = h_{SR} (\sqrt{a_1 b P_T} S_1 + \sqrt{a_2 b P_T} S_2) + n_{SR} \quad (1)$$

$$r_{SD} = h_{SD} (\sqrt{a_1 b P_T} S_1 + \sqrt{a_2 b P_T} S_2) + n_{SD} \quad (2)$$

$n_{SD}$  و  $n_{SR}$  به عنوان نویز سفید اضافه‌شونده گوسی با واریانس  $\sigma^2$  و مستقل فرض می‌شود. رله سمبل  $S_1$  را با در نظر گرفتن  $S_2$  به عنوان نویز، کدگشایی، سپس آن را کنار می‌گذارد تا سمبل  $S_2$  را از رابطه (۱) کدگشایی کند، این عمل SIC نام دارد. بنابراین سیگنال به نویز دریافتی برای سمبل  $S_1$  و  $S_2$  در رله به ترتیب برابر با

$$\gamma_{SR}^1 = \frac{|h_{SR}|^2 a_1 b P_T}{|h_{SR}|^2 a_2 b P_T + \sigma^2} \quad (3)$$

$$\gamma_{SR} = \frac{|h_{SR}|^2 a_2 P_T}{\sigma^2} \quad (4)$$

می‌شود، که علامت I در رابطه (۳) بر این نکته اشاره می‌کند که آشکارسازی  $S_1$  در رله با وجود تداخل ناشی از  $S_2$  می‌باشد. همچنین با فرض اینکه عمل SIC سمبل  $S_1$  را به طور کامل و بدون خطا حذف کرده، رابطه (۴) محقق می‌شود. در سمت دیگر، مقصد  $S_2$  را به عنوان نویز شناخته تا سمبل  $S_1$  را از رابطه (۲) بدست آورد. پس سیگنال به نویز دریافتی برای سمبل  $S_1$  در سمت مقصد را می‌توان به صورت

$$\ln\left(\frac{\sigma^2}{abP_T\beta_{SR}} + \frac{\sigma^2}{a(1-b)P_T\beta_{RD}}\right)$$

همانطور که مشاهده می‌شود بر خلاف رابطه دقیق متوسط مجموع نرخ (رابطه ۱۴)) که بصورت انتگرالی است، حد بالای بدست آمده در رابطه (۱۵) تابع ساده‌ای از ضریب سمبل‌های سیگنال نامتعامد ( $a$ ) و توان اختصاص یافته به منبع-رله ( $b$ ) است. همچنین از کران بالا برای متوسط مجموع نرخ بدست آمده ملاحظه می‌شود که بر خلاف سیستم‌های متعامد حتی در سیگنال به نویز بی‌نهایت، متوسط مجموع نرخ سیستم NOMA مقدار محدودی به دست می‌آید که این ناشی از وجود تداخل سمبل‌های سیگنال نامتعامد می‌باشد. در ادامه مقاله از روابط (۱۴) و (۱۵) برای ارزیابی عملکرد متوسط مجموع نرخ سیستم به صورت دقیق و مجانبی استفاده می‌شود.

#### ۴. محاسبه ضرایب اختصاص توان بهینه

در این بخش، تعیین ضریب سمبل‌های سیگنال نامتعامد و اختصاص توان بهینه منبع-رله سیستم همکارانه با هدف بیشینه نمودن متوسط مجموع نرخ بیت برای ساختار شبکه همکارانه با استفاده از NOMA مد نظر می‌باشد. ابتدا، در زیر بخش اول، رابطه بسته برای بهینه‌سازی اختصاص توان منبع-رله به ازای هر ضریب سمبل‌های سیگنال متعامد، به جهت بهینه‌سازی سیستم همکارانه شروع می‌شود. در زیر بخش دوم، ضریب سمبل‌های سیگنال متعامد به ازای هر اختصاص توانی از منبع-رله بهینه می‌گردد، تا عملکرد تکنیک NOMA بیشینه گردد. در انتها بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعامد و توان منبع-رله به صورت ریاضی بیان شده و یک الگوریتم برای تعیین جواب بهینه مسئله بهینه‌سازی توأم پیشنهاد می‌شود.

#### ۴-۱ بهینه‌سازی اختصاص توان منبع-رله

در این بخش هدف بهینه‌سازی اختصاص توان منبع-رله در شبکه‌های مخابرات همکارانه که بصورت

$$P_s^*, P_r^* = \operatorname{argmin} \bar{C}^{NOMA}$$

$$\text{subject to: } P_s + P_r = P_T$$

قابل تعریف است، آن‌که در آن  $P_T$  نشان دهنده توان کل انتقال یا به عبارت دیگر حداکثر توانی که یک بسته مجاز است در سرا سر انتقال از منبع به مقصد مصرف کند، می‌باشد. بدین منظور با کمک نسبت توان ( $b$ ) و جایگذاری توان منبع  $P_s = rP$  و توان رله  $P_r = (1-r)P$  در رابطه (۱۵)، تابع هدف مسئله به صورت

$$\bar{C}_{\rho \rightarrow \infty}^{NOMA} = \{E_C + \ln\left(\frac{\sigma^2}{abP_T\beta_{SD}\beta_{SR}} + \frac{\sigma^2}{a(1-b)P_T\beta_{RD}}\right)\} \times \quad (16)$$

$$C^{NOMA} = C_{S_1} + C_{S_2} \quad (10)$$

در ادامه برای سهولت نگارش  $a_r$  را با  $a$  نشان می‌دهیم. با تعریف  $Y \triangleq \min\{\lambda_{SR}b\rho a, \lambda_{RD}(1-b)\rho\}$  و  $X \triangleq \min\{\lambda_{SD}, \lambda_{SR}\}$  و استفاده از تابع چگالی احتمال  $f_{\lambda_\delta}(x) = \frac{1}{\beta_\delta} e^{-x/\beta_\delta}$  برای  $\delta \in \{SD, SR, RD\}$ ، تابع توزیع تجمعی  $X$  و  $Y$  به ترتیب  $F_X(x) = 1 - e^{-x(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})}$  و  $F_Y(y) = 1 - e^{-y(\frac{1}{\beta_{SR}b\rho a} + \frac{1}{\beta_{RD}(1-b)\rho})}$  می‌شود. با مشتق‌گیری از  $F_X(x)$  و  $F_Y(y)$ ، تابع چگالی احتمال  $X$  و  $Y$  به ترتیب می‌شود:

$$f_X(x) = \left(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}}\right) e^{-x(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})} \quad (11)$$

$$f_Y(y) = \left(\frac{1}{\beta_{SR}b\rho a} + \frac{1}{\beta_{RD}(1-b)\rho}\right) \times e^{-y(\frac{1}{\beta_{SR}b\rho a} + \frac{1}{\beta_{RD}(1-b)\rho})} \quad (12)$$

با استفاده از روابط (۱۰)، (۱۱) و (۱۲) متوسط مجموع نرخ به صورت

$$\bar{C}^{NOMA} = \int_0^\infty \left\{ \frac{1}{\gamma} \log_\gamma(1 + xb\rho) - \log_\gamma(1 + xb\rho a) \right\} f_X(x) dx + \int_0^\infty \left\{ \log_\gamma(1 + y) \right\} f_Y(y) dy \quad (13)$$

قابل محاسبه است. با استفاده از رابطه  $\int_0^\infty e^{-\mu x} \ln(1 + \eta x) dx = -\frac{1}{\mu} e^{\mu/\eta} \operatorname{Ei}(-\mu/\eta)$  حل انتگرال رابطه (۱۳) به صورت

$$\begin{aligned} \bar{C}^{NOMA} = & -\frac{\log_\gamma e}{\gamma} \left\{ e^{\frac{1}{b\rho}(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})} \times \right. \\ & \operatorname{Ei}\left(-\frac{1}{b\rho}(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})\right) - \\ & e^{\frac{1}{b\rho a}(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})} \operatorname{Ei}\left(-\frac{1}{b\rho a}(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}})\right) \\ & \left. + e^{\frac{1}{\beta_{SR}b\rho a} + \frac{1}{\beta_{RD}(1-b)\rho}} \times \right. \\ & \left. \operatorname{Ei}\left(-\frac{1}{\beta_{SR}b\rho a} - \frac{1}{\beta_{RD}(1-b)\rho}\right) \right\} \end{aligned} \quad (14)$$

قابل بیان است، که در آن  $\operatorname{Ei}(\cdot)$  نشان‌دهنده تابع انتگرال نمایی است. با استفاده از تقریب‌های  $1 + x \approx e^x$  برای  $x$  کوچک و  $\operatorname{Ei}(-x) \approx E_C + \ln(x)$  [مرجع ۱۳ معادله (۸,۲۱۲,۱)]، که  $E_C$  نشان دهنده ثابت اوپلر است، و با در نظر گرفتن  $\rho_t = \frac{bP_T}{\sigma^2}$  و  $\rho_r = \frac{(1-b)P_T}{\sigma^2}$  یک عبارت مجانبی برای رابطه (۱۴) بدین گونه بدست می‌آید:

$$\bar{C}_{\rho \rightarrow \infty}^{NOMA} = -\frac{\log_\gamma e}{\gamma} \times \left\{ E_C + \ln\left(\frac{\sigma^2}{bP_T} \left(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}}\right)\right) - \ln\left(\frac{\sigma^2}{abP_T} \left(\frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}}\right)\right) + \right. \quad (15)$$

## ۴-۲ بهینه‌سازی ضریب سمبل‌های سیگنال متعامد

ضریب اختصاص توان سمبل‌های سیگنال متعامد تأثیر بسیاری بر عملکرد تکنیک دسترسی چندگانه غیرمتعامد دارد و با بهینه‌سازی آن می‌توان متوسط نرخ را بیشینه نمود. این مسئله بهینه‌سازی به صورت

$$S_1^*, S_2^* = \operatorname{argmin} \bar{C}^{NOMA}$$

$$\text{subject to: } S_1 + S_2 = P_S$$

قابل بیان است، مشابه روند بخش قبلی، بدین منظور با تعریف  $g \triangleq \gamma \left( \frac{1}{\beta_{SR} b \rho a} + \frac{1}{\beta_{RD} (1-b) \rho} \right)$  و با مشتق‌گیری از رابطه (۱۳) نسبت به  $a$  رابطه زیر بدست می‌آید:

$$\frac{\partial \bar{C}^{NOMA}}{\partial a} = -\frac{\log_2 e}{2a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \times \int_0^\infty \frac{x}{\left( \frac{1}{b \rho a} + x \right)} e^{-x \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right)} dx + \frac{\log_2 e}{2} \left( \frac{1}{\beta_{SR} b \rho a^2} \right) \times \left( \frac{1}{\beta_{SR} b \rho a} + \frac{1}{\beta_{RD} (1-b) \rho} \right)^{-1} \times \int_0^\infty \left( \frac{g}{\frac{1}{\beta_{SR} b \rho a} + \frac{1}{\beta_{RD} (1-b) \rho} + g} \right) e^{-g} dg \quad (20)$$

با استفاده از [مرجع ۱۳ معادله (۵، ۳۵۳، ۳)]، رابطه (۲۰) را می‌توان به صورت

$$\frac{\partial \bar{C}^{NOMA}}{\partial a} = \frac{\log_2 e}{2} \left\{ - \left( \frac{1}{b \rho a} \right) \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \times e^{\frac{1}{b \rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right)} \operatorname{Ei} \left( - \frac{1}{b \rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \right) - \frac{1}{a} + \left( \frac{1}{\beta_{SR} b \rho a} \right) e^{\left( \frac{1}{\beta_{SR} b \rho a} + \frac{1}{\beta_{RD} (1-b) \rho} \right)} \times \operatorname{Ei} \left( - \frac{1}{\beta_{SR} b \rho a} - \frac{1}{\beta_{RD} (1-b) \rho} \right) + \left( \frac{1}{\beta_{RD} a (1-b)} + \frac{1}{\beta_{SR} b \rho a} \right) \right\} \quad (21)$$

بازنویسی کرد. سپس با اعمال تقریب‌های  $e^x \approx 1 + x$  و  $\operatorname{Ei}(-x) \approx E_c + \ln(x)$  که تقریب برای سیگنال به نویز بالا هستند، و  $\beta_{SR} a + \beta_{RD} \approx \frac{1}{\beta_{SR} b \rho a} + \frac{1}{\beta_{RD} (1-b) \rho} \approx \frac{1}{\beta_{SR} b \rho a}$  و  $\beta_{RD}$  برای  $a$  و  $b$  های کوچک، رابطه (۲۱) به صورت

$$\frac{\partial \bar{C}^{NOMA}}{\partial a} = -\frac{\log_2 e}{2 b \rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \times \left\{ E_c + \ln \left( \frac{1}{b \rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \right) \right\} + \frac{\log_2 e}{2 \beta_{SR} b \rho a} \times \left\{ E_c + \ln \left( \frac{1}{\beta_{SR} b \rho a} \right) \right\} - \frac{(\log_2 e) \beta_{SR} b}{2 \beta_{RD} (1-b)} \quad (22)$$

قابل بیان است. با مشتق گرفتن مجدد از رابطه (۲۲) نسبت به  $a$  داریم:

$$\left. \left( \frac{\sigma^2}{ab P_T \beta_{SR}} + \frac{\sigma^2}{a(1-b) P_T \beta_{RD}} \right) \right\} = \frac{-\log_2 e}{2} \left\{ E_c + \ln \left( \frac{\sigma^2}{P_T} \right) + \ln \left( \frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}} \right) \right\} = \frac{1}{2} \log_2 \frac{P_T}{\sigma^2} - \frac{1}{2} \log_2 e E_c - \frac{1}{2} \log_2 \left( \frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}} \right)$$

قابل بازنویسی است. برای حل مسئله بهینه‌سازی اختصاص توان می‌توان از کلیه عبارات ثابت نسبت به متغیرهای  $b$  در رابطه (۱۶) صرف‌نظر نمود و حل مسئله بهینه‌سازی معادل با برابر صفر قرار دادن مشتق رابطه (۱۶) می‌گردد:

$$\frac{\partial \bar{C}_{\rho \rightarrow \infty}^{NOMA}}{\partial b} = -\frac{1}{2 \ln 2} \frac{1}{\frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}}} = 0 \quad (17)$$

برای اثبات مقعر بودن رابطه (۱۷) از آن یک بار دیگر نسبت به  $b$  مشتق گرفته و مشاهده می‌شود:

$$\frac{\partial^2 \bar{C}_{\rho \rightarrow \infty}^{NOMA}}{\partial b^2} = -\frac{1}{2 \ln 2} \times \left( \frac{a^2}{(1-b)^4 \beta_{RD}^2} + \frac{3}{b^4 \beta_{SR}^2} + \frac{2a((1-b)^2 + b^2)}{b^3(1-b)^3 \beta_{SR} \beta_{RD}} \right) \left( \frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}} \right)^2 \quad (18)$$

رابطه (۱۸) باید همواره کوچک‌تر از صفر باشد، از آنجا که  $\beta_{RD}$ ،  $\beta_{SR}$  و  $a$  و  $b$  اعدادی مثبت و همچنین  $b$  و  $a$  اعدادی بین صفر و یک هستند  $< 0$  همواره برقرار است. در نتیجه مقدار بهینه ضریب اختصاص توان منبع-رله را می‌توان با حل معادله رابطه (۱۷) به صورت

$$b^* = \frac{1}{1 + \sqrt{a \left( \frac{\beta_{SD}}{\beta_{RD}} \right)}} \quad (19)$$

بدست آورد. رابطه (۱۹) مقدار بهینه ضریب اختصاص توان منبع-رله  $b^*$  برحسب ضریب اختصاص توان سمبل‌های سیگنال متعامد  $a$  را ارائه می‌دهد. با استفاده از این نسبت بهینه توان  $a^*$ ، توان بهینه منبع و رله به ترتیب از روابط

$$P_S^* = b^* P$$

$$P_R^* = (1 - b^*) P$$

قابل محاسبه است. حال که ضریب اختصاص توان منبع-رله به جهت بهبود عملکرد سیستم همکارانه محاسبه گردید، در ادامه به دنبال بهینه‌سازی ضریب سمبل‌های سیگنال متعامد به جهت بهبود عملکرد تکنیک NOMA برای دستیابی به متوسط نرخ حداکثری هستیم.

جایگذاری آن در رابطه (۲۹) مقدار ضریب اختصاص توان بهینه سیگنال،  $a^*$  را محاسبه نمود.

### ۳-۴ بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعامد و اختصاص توان منبع-رله

در بخش آخر هدف بهینه‌سازی ضریب سمبل‌های سیگنال متعامد به همراه بهینه‌سازی توان برای منبع-رله است، که به کارایی بالاتر برای شبکه در عوض پیچیدگی بیشتر می‌رسد. مسئله بهینه‌سازی توأم به صورت زیر قابل بیان است:

$$a^*, b^* = \operatorname{argmax} C^{NOMA} \quad (31)$$

$$0 \leq a \leq 1$$

$$0 \leq b \leq 1$$

الگوریتم بهینه‌سازی توأم ضرایب اختصاص توان منبع-رله و سمبل‌های سیگنال متعامد
۱- مقدار اولی به برای $a$ در نظر گرفته می‌شود (مثلاً $a = 0.5$ )
۲- محاسبه $b$ جدی با استفاده از رابطه $b_{new}^* = \frac{1}{1 + \sqrt{a \times \left(\frac{\beta_{SD}}{\beta_{RD}}\right)}}$
۳- با $b$ بدست آمده در بند ۲ و رابطه $a_{new}^* = e^{(-\frac{1}{\alpha}W(\omega) + E_c + \psi)}$ بروزرسانی می‌شود.
۴- تکرار بند دوم و سوم تا به مقدار $a^*$ و $b^*$ بهی نه همگرا شود.

مقرر بودن این مسئله بهینه‌سازی در پیوست اثبات شده است. حال مسئله (۳۱) را می‌توان بدون از دست دادن کلیت مسئله به صورت دو گام جداگانه نمایش داد، یعنی:

$$a^*, b^* = \operatorname{argmax}_{a \in [0,1]} \left( \operatorname{argmax}_{b \in [0,1]} C^{NOMA} \right) \quad (32)$$

با استفاده از رابطه ضریب اختصاص توان بهینه در منبع-رله (رابطه (۱۹)) و جایگذاری آن در رابطه ضریب بهینه سمبل‌های سیگنال متعامد (رابطه (۲۹))، بهینه‌سازی توأم اختصاص توان در منبع-رله و ضریب بهینه سمبل‌های سیگنال متعامد با استفاده از الگوریتم زیر قابل استخراج است.

همچنین در شکل ۲ روندنمای الگوریتم بهینه‌سازی توأم پیشنهادی این مقاله نمایش داده شده است.

$$\frac{\partial^2 \bar{C}^{NOMA}}{\partial a^2} = \frac{\log_2 e}{2\rho b a^2} \left[ \frac{2}{\beta_{SD}} \times \left\{ E_c + \ln \left( \frac{1}{b\rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \right) \right\} + \frac{2}{\beta_{SR}} \ln \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) + \frac{1}{\beta_{SD}} - \frac{2}{\beta_{SR}} \left( \ln \frac{1}{\beta_{SR}} \right) \right] < 0 \quad (23)$$

اگر نامساوی رابطه (۲۳) برقرار باشد، اثبات می‌شود رابطه (۲۲) نسبت به  $a$  مقعر است. رابطه (۲۳) در صورتی برقرار است که  $a > \frac{1}{b\rho} e^{E_c + \ln \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) + \frac{\beta_{SD}}{\beta_{SR}} \ln \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) + \frac{1}{\beta_{SD}} - \frac{\beta_{SD}}{\beta_{SR}} \ln \left( \frac{1}{\beta_{SR}} \right)}$  باشد، از آنجا  $\rho$  عددی بسیار بزرگ می‌باشد، این نامساوی برقرار است و مسئله بهینه‌سازی یک جواب بهینه کلی دارد. با معادل صفر قراردادن رابطه (۲۲)، مقدار  $a^2$  به صورت زیر بدست می‌آید:

$$a^2 = \frac{(1-b)\beta_{RD} \ln a - (1-b)\beta_{RD} E_c - \beta_{SD}\beta_{SR} b^2 \rho}{b^2 \rho \beta_{SD} \beta_{SR}} - \frac{(1-b)\beta_{RD} (\ln b \rho \beta_{SR})}{\beta_{SR} b^2 \rho} + \frac{(1-b)\beta_{RD}}{\beta_{SR} b^2 \rho} \times \quad (24)$$

$$\left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \ln \left( \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right)^{-1} b \rho \right)$$

با تغییر متغیر  $\zeta \triangleq \ln a$ ، رابطه (۲۴) را می‌توان به صورت:

$$e^{2\zeta} = \frac{(1-b)\beta_{RD}}{\beta_{SD}\beta_{SR} b^2 \rho} \zeta - \frac{(1-b)\beta_{RD} E_c}{b^2 \rho \beta_{SD} \beta_{SR}} + \frac{(1-b)\beta_{RD}}{b^2 \beta_{SR} \rho} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \times \ln \left( \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right)^{-1} b \rho \right) - \frac{(1-b)\beta_{RD}}{b^2 \rho \beta_{SR}^2} \ln(b\rho\beta_{SR}) \quad (25)$$

بازنویسی کرد. حال با استفاده از تابع لمبرت که با تابع  $W(\cdot)$  [مرجع ۱۴ رابطه (۱)] بیان می‌شود، می‌توان  $\zeta$  را از رابطه (۲۵) بدین‌گونه محاسبه نمود:

$$\zeta = -\frac{1}{\alpha} W(\omega) + E_c + \psi \quad (26)$$

که در آن:

$$\omega = -\frac{2b^2\beta_{SD}\beta_{SR}\rho}{(1-b)\beta_{RD}} e^{\psi E_c + 2\psi} \quad (27)$$

و

$$\psi = -\left( 1 + \frac{\beta_{SD}}{\beta_{SR}} \right) \ln \left( \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right)^{-1} b \rho \right) + \frac{\beta_{SD}}{\beta_{SR}} \ln(\beta_{SR} b \rho) \quad (28)$$

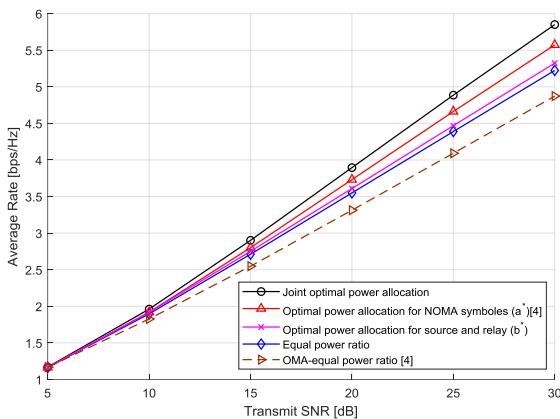
هستند. در نتیجه ضریب بهینه را می‌توان از رابطه

$$a^* = e^{(-\frac{1}{\alpha}W(\omega) + E_c + \psi)} \quad (29)$$

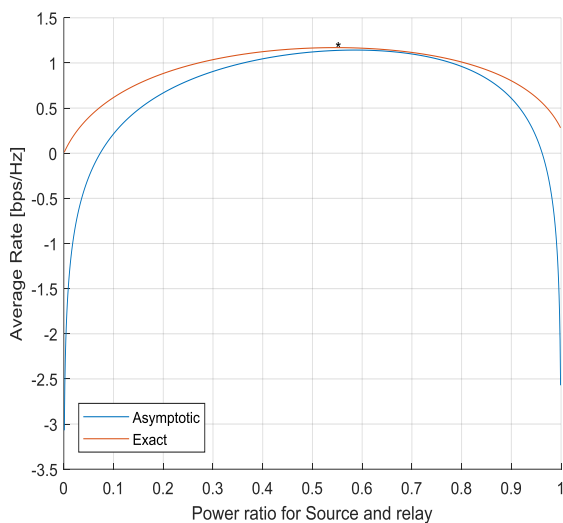
بدست آورد. از آنجا که مقدار  $\omega$  در رابطه (۲۶) منفی است، از رابطه تقریب لمبرت [۱۴] می‌توان برای تقریب  $\zeta$  در رابطه (۲۶) به صورت

$$W(\omega) \approx \ln(\omega) - \frac{\gamma}{\alpha} \times \left[ 1 - \left\{ 1 + \alpha \left( -\frac{1 + \ln(-\omega)}{\gamma} \right)^{\frac{1}{\alpha}} \right\}^{-1} \right] \quad (30)$$

استفاده کرد، که در آن  $\alpha = 0.3205$  است. با کمک رابطه (۳۰) و



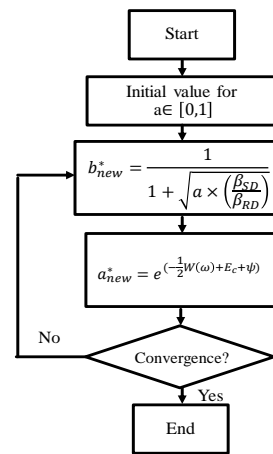
شکل ۳. متوسط نرخ دریافتی نسبت به SNR ارسالی برای چهار طرح: الف) ضریب سمبل‌های سیگنال متعادل برابر ( $a = b$ ) ب) اختصاص توان بهینه در منبع-رله، ج) ضریب سمبل‌های سیگنال متعادل بهینه [۴] و د) بهینه‌سازی توان



شکل ۴. متوسط مجموع نرخ بر حسب ضریب اختصاص توان منبع-رله مختلف در شبکه مخابرات همکارانه NOMA با فرض سیگنال به نویز ۵ dB

بهبود تقریباً ۳ dB در مقایسه حالت الف) یعنی حالت بدون بهینه‌سازی قابل مشاهده است.

برای اینکه اثر رابطه مجانبی (رابطه (۱۴)) و رابطه دقیق (رابطه (۱۵)) متوسط نرخ را در تعیین ضریب اختصاص توان منبع-رله مقایسه کنیم، منحنی شکل (۴) را به نمایش می‌گذاریم. این شکل میزان متوسط نرخ را بر حسب ضریب اختصاص توان منبع-رله مختلف به نشان می‌دهد. همان‌طور که از شکل مشاهده می‌شود،



شکل ۲. روندنمای الگوریتم بهینه‌سازی توأم پیشنهادی

## ۵. نتایج شبیه‌سازی

در این بخش برای تأیید روابط و الگوریتم بدست آمده در بخش قبل شبیه‌سازی‌هایی انجام شده است. با توجه به ساختار سیستم همکارانه این یک فرض معمول است که میزان بهره کانال بین منبع-رله بهتر از واریانس کانال‌های بین منبع-مقصد و رله-مقصد در نظر گرفته شود. در نتیجه در تمامی شبیه‌سازی‌ها بهره کانال بین گرہ‌ها منبع-مقصد، منبع-رله و رله-مقصد به ترتیب به صورت  $\beta_{SD} = 1$ ,  $\beta_{SR} = 10$  و  $\beta_{RD} = 5$  فرض شده است.

هدف از شکل (۳) اثبات این است که در صورت بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعادل و اختصاص توان در رله و مقصد، کارایی شبکه با رله همکار و استفاده از NOMA افزایش می‌یابد. در این شکل، متوسط نرخ بر اساس سیگنال به نویز ارسالی برای چهار طرح مورد بررسی قرار گرفته است: الف) ضریب سمبل‌های سیگنال متعادل برابر ( $a = 1/2$ ) و اختصاص توان منبع-رله ( $b = 1/2$ )، ب) ضریب سمبل‌های سیگنال متعادل برابر ( $a = 1/2$ ) و اختصاص توان منبع-رله بهینه ( $b = b^*$ )، ج) ضریب سمبل‌های سیگنال متعادل بهینه ( $a = a^*$ ) و اختصاص توان منبع-رله بهینه برابر ( $b = 1/2$ ) [۴] و د) اختصاص بهینه توأم ضریب سمبل‌های سیگنال متعادل و توان منبع-رله. همچنین در این شکل متوسط نرخ بر اساس سیگنال به نویز ارسالی برای سیستم در یک شبکه مخابرات همکارانه با استفاده از دسترسی چندگانه متعادل (OMA) نیز از مرجع [۴] برای مقایسه نشان داده شده است.

با توجه به شکل مشخص است که در سیگنال به نویز ۳۰ dB اختصاص توان منبع-رله بهینه و ضریب اختصاص توان سمبل‌های سیگنال متعادل برابر در حالت ب)، متوسط نرخ نزدیک به ۱ dB بهبود نسبت به حالت بدون بهینه‌سازی الف) دارد. بعلاوه، با استفاده از الگوریتم پیشنهادی بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعادل و توان منبع-رله در حالت د)، در شکل میزان

<sup>۱</sup> Orthogonal Multiple Access

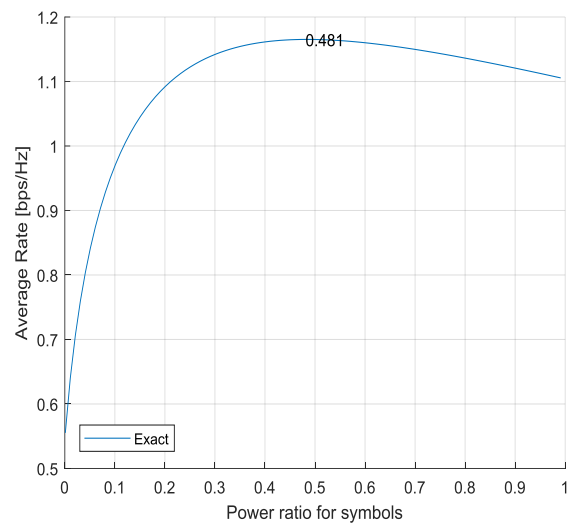
مقدار اختصاص توان بهینه محاسبه شده از رابطه تقریبی (رابطه (۱۹)) کاملاً بر مقدار بهینه رابطه دقیق (که از جستجوی سراسری بدست آمده است) انطباق دارد. همچنین از این شکل مشخص است که اختصاص توان منبع-رله به شدت بر کارایی متوسط مجموع نرخ شبکه تأثیرگذار است و هرچه قدر از مقدار اختصاص توان بهینه فاصله گرفته شود میزان متوسط نرخ کمتر می‌شود و کارایی کاهش پیدا می‌کند.

شکل (۵) میزان متوسط نرخ را به ازای ضریب سمبل‌های سیگنال متعامد مختلف نشان می‌دهد. مشابه شکل (۴)، در این شکل می‌توان تأثیر بهینه‌سازی ضریب سمبل‌های سیگنال متعامد در افزایش متوسط مجموع نرخ در شبکه مخابرات همکارانه NOMA را مشاهده نمود. در این نمودار بر خلاف شکل (۴) تنها رابطه دقیق متوسط مجموع نرخ رسم شده است، زیرا از این رابطه برای محاسبه ضریب سمبل‌های سیگنال متعامد بهینه ( $a^*$ ) استفاده شده است. همان‌طور که از رابطه (۱۴) مورد انتظار بود ضریب سمبل‌های سیگنال متعامد بر کارایی شبکه NOMA به شدت تأثیرگذار است، به ویژه زمانی که مقدار ضریب بسیار کوچک  $a \rightarrow 0$  در نظر گرفته شود.

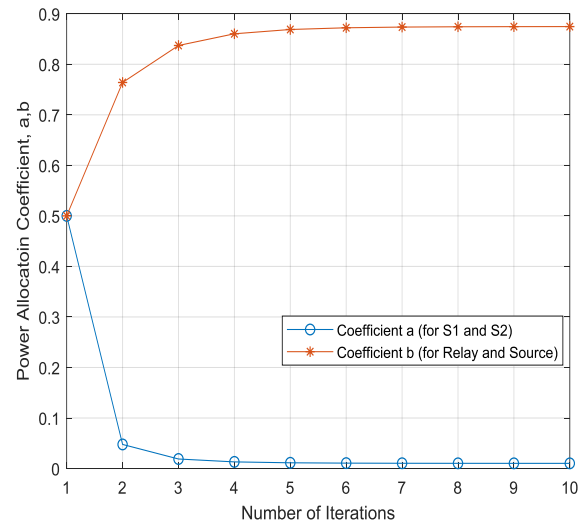
در انتها، به بررسی همگرایی الگوریتم پیشنهادی بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعامد و توان منبع-رله می‌پردازد. به عنوان مثالی از رفتار همگرایی الگوریتم پیشنهادی در این مقاله، مقدار ضریب سمبل‌های سیگنال متعامد ( $a$ ) و ضرایب اختصاص توان منبع-رله ( $b$ ) را در هر مرحله از الگوریتم در شکل (۶) نشان دادیم. این شکل مقدار این ضرایب اختصاص توان را بر حسب تکرارهای مختلف الگوریتم نشان می‌دهد که در نهایت به مقادیر  $a^* = 0.0103$  و  $b^* = 0.874$  همگرا می‌شود. از شکل مشخص است که الگوریتم با تعداد تکرار معقول (تقریباً ۵ تکرار) همگرا می‌شود و ضرایب بهینه محاسبه می‌شوند، بنابراین الگوریتم پیشنهادی در این مقاله از سرعت همگرایی خوبی برخوردار است. البته می‌توان نشان داد که این سرعت همگرایی برای سایر شرایط کانال نیز برقرار است.

## ۶. نتیجه‌گیری

در این مقاله، یک سیستم مخابرات همکارانه با استفاده از دسترسی چندگانه غیرمتعامد (NOMA) با رله DF در کانال‌های با محوشوندگی رایلی ارائه شده است. در بخش اول عملکرد متوسط مجموع نرخ (ASR) این سیستم محاسبه گردید. در بخش بعد، از آنجایی که اختصاص توان سمبل‌های سیگنال غیرمتعامد تأثیر قابل ملاحظه‌ای بر عملکرد سیستم NOMA می‌گذارد، طراحی بهینه ضریب سمبل‌های سیگنال متعامد بیان شد. همچنین از آنجا که کنترل توان در شبکه‌های همکارانه با رله



شکل ۵. متوسط مجموع نرخ نسبت به ضرایب سمبل‌های سیگنال متعامد مختلف در شبکه مخابرات همکارانه NOMA. با فرض سیگنال به نویز ۵ dB.



شکل ۶. همگرایی الگوریتم بهینه‌سازی توأم ضریب سمبل‌های سیگنال متعامد ( $a$ ) و اختصاص توان منبع-رله ( $b$ ). مقادیر ضرایب اختصاص توان بهینه بر حسب تعداد تکرار الگوریتم

Computing Electronics and Control, ۱۶(۵), ۱۹۶۶-۱۹۷۳.

- [۷] Liu, H., Ding, Z., Kim, K. J., Kwak, K. S., & Poor, H. V. (۲۰۱۸). Decode-and-forward relaying for cooperative NOMA systems with direct links. *IEEE Transactions on Wireless Communications*, ۱۷(۱۲), ۸۰۷۷-۸۰۹۳.
- [۸] Wang, Z., & Peng, Z. (۲۰۱۹). Secrecy performance analysis of relay selection in cooperative NOMA systems. *IEEE Access*, ۷, ۸۶۲۷۴-۸۶۲۸۷.
- [۹] Zou, D., Deng, D., Rao, Y., Li, X., & Yu, K. (۲۰۱۹). Relay selection for cooperative NOMA system over correlated fading channel. *Physical Communication*, ۳۵, ۱۰۰۷۰۲.
- [۱۰] Ghous, M., Abbas, Z. H., Abbas, G., Hassan, A. K., & Moinuddin, M. (۲۰۲۰). Transmit beamformer based performance analysis and diversity gains of cell edge user in cooperative MISO-NOMA system. *Physical Communication*, ۱۰۱۱۰۲.
- [۱۱] Xu, M., Ji, F., Wen, M., & Duan, W. (۲۰۱۶). Novel receiver design for the cooperative relaying system with non-orthogonal multiple access. *IEEE Communications Letters*, ۲۰(۸), ۱۶۷۹-۱۶۸۲.
- [۱۲] Kader, M. F., & Shin, S. Y. (۲۰۱۶). Cooperative relaying using space-time block coded non-orthogonal multiple access. *IEEE Transactions on Vehicular Technology*, ۶۶(۷), ۵۸۹۴-۵۹۰۳.
- [۱۳] Jeffrey, A., & Zwillinger, D. (Eds.). (۲۰۰۷). *Table of integrals, series, and products*. Elsevier.
- [۱۴] Chatzigeorgiou, I. (۲۰۱۳). Bounds on the Lambert function and their application to the outage analysis of user cooperation. *IEEE Communications Letters*, ۱۷(۸), ۱۵۰۵-۱۵۰۸.
- [۱۵] Gilbert, G. T. (۱۹۹۱). Positive definite matrices and Sylvester's criterion. *The American Mathematical Monthly*, ۹۸(۱), ۴۴-۴۶.

دارای اهمیت است، تخصصی بهی‌توان بی‌منبع و رله نی‌محقق شده است. در انتها، الگوریتمی برای حل مسئله بهی‌توان سازی توأم ضربی سمبل‌های سی‌گنال متعامد و اختصاص توان منبع-رله در شبکه‌های همکارانه NOMA، به منظور تضمین عملکرد بهی‌توان و مطمئن این سی‌ستم‌ها در عمل پی‌شنهاد شده است. نتایج نشان می‌دهند که اگر اختصاص توان به سمبل‌های NOMA بهی‌توان صورت‌نگی‌رد و با اختصاص توان بی‌منبع و رله بصورت بهی‌توان در نظر گرفته نشود، عملکرد متوسط مجموع نرخ سی‌ستم کاهش می‌ابد. همچنین نتایج شبیه‌سازی بهی‌توان سازی توأم اختصاص توان با الگوریتم پی‌شنهادی، بهره ۱dB در مقایسه با حالتی که در آن تنها اختصاص توان برای سی‌گنال صورت گرفته باشد و بهره ۲dB در مقایسه با حالتی که فقط اختصاص توان برای منبع-رله انجام شده باشد، را نشان می‌دهد.

## مراجع

- [۱] Amiriara, H., Zahabi, M. R., & Meghdadi, V. (۲۰۱۸, December). Joint Power and Location Optimization of Relay for Amplify-and-Forward Cooperative Relaying. In ۲۰۱۸ International Conference on Internet of Things, Embedded Systems and Communications (IINTEC) (pp. ۹۷-۱۰۲). IEEE.
- [۲] Amiriara, H., Zahabi, M. R., & Meghdadi, V. (۲۰۲۰). Joint power-location optimization in AF cooperative relay systems with Nakagami-m channel. *Physical Communication*, ۱۰۱۰۶۷.
- [۳] Seo, J. B., Jin, H., Joung, J., & Jung, B. C. (۲۰۲۰). Uplink NOMA Random Access Systems With Space-Time Line Code. *IEEE Transactions on Vehicular Technology*, ۶۹(۴), ۴۵۲۲-۴۵۲۶.
- [۴] Kim, J. B., & Lee, I. H. (۲۰۱۵). Capacity analysis of cooperative relaying systems using non-orthogonal multiple access. *IEEE Communications Letters*, ۱۹(۱۱), ۱۹۴۹-۱۹۵۲.
- [۵] Zeng, M., Hao, W., Dobre, O. A., & Ding, Z. (۲۰۲۰). Cooperative NOMA: State of the Art, Key Techniques, and Open Challenges. *IEEE Network*, ۳۴(۵), ۲۰۵-۲۱۱.
- [۶] Do, D. T., & Nguyen, T. T. T. (۲۰۱۸). Exact Outage Performance Analysis of Amplify-and-Forward-Aware Cooperative NOMA. *TELKOMNIKA Telecommunication*



$$\frac{\partial^2 C(a, b)}{\partial a \partial b} = -\frac{1}{\gamma \ln \gamma} \times \left( \frac{1}{b^\gamma (1-b)^\gamma \beta_{RD} \beta_{SD} \left[ \frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}} \right]^\gamma} \right) < 0.$$

$$\frac{\partial^2 C(a, b)}{\partial b^2} = -\frac{1}{\gamma \ln \gamma} \times \left( \frac{a^\gamma \left( \frac{1}{(1-b)^\gamma \beta_{RD}} + \frac{\gamma}{b^\gamma \beta_{SR}} + \frac{\gamma a ((1-b)^\gamma + b^\gamma)}{b^\gamma (1-b)^\gamma \beta_{SR} \beta_{RD}} \right)}{\left( \frac{1}{b \beta_{SR}} + \frac{a}{(1-b) \beta_{RD}} \right)^\gamma} \right) < 0.$$

### پیوست‌ها

. اثبات مقعر بودن رابطه (۳۱):

برای بررسی مقعر بودن مسئله بهینه‌سازی، ابتدا ماتریس هسین  $C(a, b)$  را بصورت زیر تعریف می‌کنیم:

$$H = \begin{bmatrix} \frac{\partial^2 C(a, b)}{\partial a^2} & \frac{\partial^2 C(a, b)}{\partial a \partial b} \\ \frac{\partial^2 C(a, b)}{\partial a \partial b} & \frac{\partial^2 C(a, b)}{\partial b^2} \end{bmatrix}$$

درایه‌های این ماتریس عبارتند از:

$$\frac{\partial^2 C(a, b)}{\partial a^2} = \frac{\log_\gamma e}{\gamma \rho b a^\gamma} \left[ \frac{\gamma}{\beta_{SD}} \left\{ E_c + \ln \left( \frac{1}{b \rho a} \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) \right) \right\} + \frac{\gamma}{\beta_{SR}} \ln \left( \frac{1}{\beta_{SD}} + \frac{1}{\beta_{SR}} \right) + \frac{1}{\beta_{SD}} - \frac{\gamma}{\beta_{SR}} \left( \ln \frac{1}{\beta_{SR}} \right) \right] < 0.$$

هر چهار درایه با توجه به مقادیر  $a, b \in \{0, 1\}$  کاملاً منفی می‌باشند، همچنین دترمینان ماتریس  $H$  نیز که به صورت  $\det[H] = \left( \frac{\partial^2 C(a, b)}{\partial a^2} \times \frac{\partial^2 C(a, b)}{\partial b^2} - \left( \frac{\partial^2 C(a, b)}{\partial a \partial b} \right)^2 \right)$  نوشته می‌شود، مقداری کاملاً منفی دارد. از آنجایی که ماتریس  $H$  یک ماتریس هرمیتی است و با توجه به معیار سیلوستر<sup>۱</sup> در مرجع [۱۵]، چون تمام درایه‌ها مقادیر کاملاً منفی دارند و مقدار دترمینان آن نیز منفی است، بنابراین مقادیر ویژه  $H$  نیز کوچک‌تر از صفر است و تابع نسبت به  $a$  و  $b$  مقعر است.

<sup>۱</sup> Sylvester's criterion

## یک مدل کنترل دسترسی برای شبکه‌های اجتماعی آنلاین از طریق ارتباطات کاربر به کاربر

\* محمدجواد پیران \* محمود دی‌پیر

\* دانشجوی دکتری، دانشگاه آزاد اسلامی واحد علوم و تحقیقات

\*\* محمود دی‌پیر، دانشیار، دانشگاه هوایی شهید ستاری

تاریخ دریافت ۱۳۹۹/۰۱/۱۵ تاریخ پذیرش ۱۳۹۹/۰۹/۲۹

نوع مقاله: پژوهشی

### چکیده

با فراگیر شدن شبکه‌های اجتماعی و رشد افزون اطلاعات به اشتراک‌گذاری شده در آن‌ها، کاربران این شبکه‌ها در معرض تهدیدهای بالقوه‌ی امنیت و حریم خصوصی داده‌ها قرار دارند. تنظیمات حریم خصوصی گنجانده‌شده در این شبکه‌ها به کاربران کنترل کاملی در جهت مدیریت و خصوصی‌سازی دسترسی به اطلاعات اشتراکی‌شان توسط کاربران دیگر نمی‌دهد. در این مقاله به کمک مفهوم گراف اجتماعی، یک مدل جدید کنترل دسترسی کاربر به کاربر پیشنهاد شد که امکان بیان سیاست‌های حریم شخصی و کنترل دسترسی دقیق‌تر و حرفه‌ای‌تری را برحسب الگو و عمق روابط میان کاربران در شبکه‌های اجتماعی فراهم می‌کند. در این مقاله با به‌کارگیری روش شاخصه‌ای منظم، روابط غیرمستقیم در میان کاربران موردبررسی و تحلیل قرار گرفته و سیاست‌های دقیق‌تری نسبت به مدل‌های قبلی ارائه‌شده است. ارزیابی نتایج نشان داد، در مورد ۱۰ همسایه برای هر کاربر، تجمع احتمال یافته‌های یک مسیر واجد شرایط، به ترتیب برای سه حلقه شماره‌ده اول برابر ۱، ۱۰/۵ و ۶۷/۳ درصد است، و نهایتاً برای شماره‌ده چهارم به ۱۰۰ درصد می‌رسد که با افزایش مشخصه شمارشی تعریف‌شده، زمان اجرای متوسط الگوریتم پیشنهادی و سایر الگوریتم‌های ارائه‌شده در روش‌های قبلی افزایش می‌یابد اما برای حدود بالاتر مشخصه شمارشی، الگوریتم پیشنهادی بهتر از الگوریتم‌های قبلی عمل می‌کند.

**واژگان کلیدی:** شبکه‌های اجتماعی، امنیت شبکه، حریم خصوصی، حلقه شماره‌ده، امنیت داده‌ها.

## ۱- مقدمه

شبکه های اجتماعی در زندگی روزانه نقش قابل توجهی داشته و شکل ارتباطات تعاملی و به اشتراک گذاری اطلاعات را در میان مردم تغییر داده اند. کاربران حجم گسترده ای از اطلاعات خود را در شبکه های اجتماعی آنلاین با اهداف مختلفی به اشتراک می گذارند. به اشتراک گذاری اطلاعات و ارتباط میان کاربران که مبتنی بر ارتباطات اجتماعی است تحت عنوان روابط<sup>۱</sup> تعریف می شود [۱]. با توجه به این که بیشتر کاربران باهدف ارتباط با افراد و مخاطبان شناخته شده خود به شبکه های اجتماعی آنلاین می پیوندند، آن ها اغلب اطلاعات حساس یا خصوصی زیادی را در مورد خود منتشر کرده و به اشتراک می گذارند. با توجه به شهرت شبکه های اجتماعی آنلاین و رشد انفجاری اطلاعات به اشتراک گذاری شده در آن ها، کاربران شبکه های اجتماعی آنلاین در معرض تهدیدات بالقوه ای امنیت و حریم خصوصی داده های خود قرار دارند. بنابراین جرائم مرتبط با حریم خصوصی و امنیت در شبکه های اجتماعی آنلاین توجه تعداد زیادی از رسانه ها و جوامع تحقیقاتی را به خود جلب کرده است [۲ و ۱]. این رخدادها نشان دهنده ضرورت کنترل دسترسی مؤثر است که موجب محافظت از داده ها نسبت به دسترسی نامشخص و تأیید نشده در شبکه های اجتماعی آنلاین می گردد [۳].

کنترل دسترسی در شبکه های اجتماعی آنلاین نشان دهنده مشخصات و خصوصیات منحصر به فرد متعددی نسبت به کنترل دسترسی معمول می باشد. در کنترل دسترسی اجباری و نقش محور، یک سیاست کنترل دسترسی سیستم گسترده توسط مدیر امنیت لحاظ می شود. در کنترل دسترسی اختیاری، مالک و ناشر منبع سیاست کنترل دسترسی را تعریف می کند. در سیستم های شبکه های اجتماعی آنلاین کاربران انتظار دارند تا دسترسی را نسبت به منابع و فعالیت های مرتبط خود تنظیم کنند؛ بنابراین دسترسی به اطلاعات در شبکه های اجتماعی آنلاین در معرض سیاست های تعیین شده از سوی کاربر قرار دارد [۴]. به جز مالک اطلاعات منبع، برخی از کاربران مرتبط نیز ممکن است کنترل نمایش اطلاعات منبع یا کاربر را در اختیار داشته باشند. جهت

ممانعت کاربران از دسترسی به محتوای ناخواسته یا نامناسب، سیاست های تعیین شده از سوی کاربر که موجب ارائه نحوه ی در نظر گرفتن اطلاعات دسترسی مورد نیاز کاربران است به کار گرفته می شوند؛ بنابراین بایستی سیستم این سیاست های نسبی را از سوی کاربران در دسترس و کاربران هدف همراه با سیاست های تعیین شده ی سیستم جمع آوری کرده و آن ها را برای تصمیم گیری در مورد کنترل جمعی آگاه کند. در شبکه های اجتماعی آنلاین دسترسی به منابع به طور معمول بر اساس روابط بین دسترسی کاربر و کنترل کاربر در گراف اجتماعی کنترل می شود. این نوع از کنترل دسترسی رابطه محور، وجود یک رابطه خاص یا مجموعه ای از روابط را بین کاربران در نظر گرفته و سیاست های کنترل دسترسی را برحسب این روابط کاربر با کاربر<sup>۲</sup> بیان می کنند [۵].

در این مقاله یک مدل کنترل دسترسی رابطه محور کاربر به کاربر جدید پیشنهاد می شود که امکان بیان سیاست های کنترل دسترسی دقیق تر و حرفه ای تری را برحسب الگو و عمق روابط میان کاربران در شبکه فراهم می کند. نوع الگو، بیان کننده الگوی انواع روابط موجود در مسیر رابطه از کاربر در دسترس به سمت کاربر هدف است. در این تحقیق یک روش شاخصه ای منظم از کاراکترها را به کار گرفته و علائم کمی نیز جهت بیان مسیر رابطه مورد استفاده قرار می گیرند که نشان دهنده روابط غیرمستقیم در میان کاربران است. استفاده از رابطه مبتنی بر عبارات منظم و روابط چندگانه به ما توانایی لازم را در جهت ارائه سیاست های مشخص می دهد که این قابلیت در مدل های قبلی وجود نداشت و با توجه به اطلاعات ما، این مدل به عنوان اولین مدل کنترل دسترسی رابطه ای برای شبکه های اجتماعی آنلاین با این قابلیت شناخته می شود. در ادامه در قسمت ۱، مروری بر مهم ترین پژوهش های صورت گرفته در خصوص کنترل دسترسی برای شبکه های اجتماعی آنلاین انجام شده است؛ سپس در قسمت ۲، روش پیشنهادی بیان می شود. ارزیابی نتایج در قسمت ۳ انجام شده و در نهایت قسمت ۴ به نتیجه گیری و قسمت ۵ به پیشنهادهایی جهت پژوهش محققین آتی پرداخته است.

## ۲- مروری بر پیشینه ی پژوهش

رنجبر، مومن زاده و ارفعی نیا، در پژوهشی با عنوان مدل کنترل دسترسی در شبکه های اجتماعی، با یک تکنیک کنترل دسترسی

<sup>۲</sup> User-to-User

<sup>۱</sup> Relationships

کاربردهای جذابی که این اطلاعات را مورد تحلیل قرار می‌دهند را معرفی نمودند [۸]. این امر منجر به معرفی تعدادی بخش‌های پژوهشی جدید برای محققین گردید. با مرور تحقیق‌سورانی و همکاران، طبقه‌بندی عناوین پژوهشی در مورد شبکه‌های اجتماعی آنلاین مطرح گردید. این طبقه‌بندی شامل ۱۷ زیر عنوان یا زیر بخش تحقیقاتی بود که به همراه برخی تحقیقات انجام‌شده در این زیر بخش‌ها معرفی شده است. بر طبق پیامدهای (جزئی قابل توجه و بعضاً فجیع) ارتباط اطلاعات شخصی و خصوصی یکی از بخش‌های پژوهشی که محققین روی آن بسیار پژوهش کرده‌اند حریم خصوصی در شبکه‌های اجتماعی آنلاین می‌باشد. بعد از ارائه کلیاتی پیرامون زیر بخش‌های پژوهشی مختلف سایت‌های شبکه اجتماعی بر زیر بخش حفاظت از حریم خصوصی در شبکه‌های اجتماعی تأکید بیشتری کرده و جنبه‌های مختلفی از آن به همراه طبقه‌بندی این جنبه‌ها را معرفی نمودند. یانگ و همکاران تلاشی بر تدوین سامانه‌های توصیه‌گر کاملاً پیش‌برنده رتبه‌بندی / بازخورد چرخه محور انجام دادند که بر استنتاج چرخه‌های صداقت اجتماعی با مقوله خاص تمرکز می‌کند و از داده رتبه‌بندی فعلی همراه با داده شبکه اجتماعی گرفته‌شده است. چندین واریانس مختلف مرتب‌شده دوستان با چرخه‌های مبتنی بر سطوح تخصصی استنباط شده آن‌ها را طرح‌ریزی می‌کند. از طریق آزمایش‌های صورت گرفته بر داده‌های عمومی موجود شرح می‌دهد که الگوهای پیشنهادی چرخه محور می‌تواند از اطلاعات صداقت اجتماعی کاربر که به‌دقت بیشتر منتج می‌شود را مورد استفاده بهتر قرار دهد [۹]. ساردو و همکاران یک چارچوب تجربی را پیشنهاد کرده‌اند که با این چارچوب پروفایل جعلی در لیست دوستان امکان‌پذیر است. باین حال این چارچوب به سایت شبکه‌ی اجتماعی خاصی به نام فیس‌بوک محدود شده است. این چارچوب داده‌ها را از لیست دوستان استخراج می‌کند و آن را برای طبقه‌بندی آن‌ها به‌عنوان داده‌های واقعی یا جعلی مورد استفاده قرار می‌دهد و از روند یادگیری ماشین نظارت شده و تحت نظارت استفاده می‌کند [۱۰]. یانگ و همکاران در مقاله‌ای، پروفایل ده میلیون کاربر از بزرگ‌ترین میکرو بلاگ کشور چین و تعداد ۴۱ میلیون پروفایل تویتر به‌طور تجربی مطالعه شدند. با مطالعه توپولوژی‌ها و ضا به این ویژگی پایدار پی بردند که کاربران تابع انتشار اطلاعات مطلوب برحسب حداقل زمان تکثیر دست هستند. سپس این دلیل را از دیدگاه تئوریک بررسی کرده و پی بردند که ساختار شبکه به حداقل انتشار

مؤثر پیشنهاد گردید تا مجاز یا غیر مجاز بودن درخواست دسترسی به منابع مشترک را با توجه به شرایط طرفین چندگانه تعیین نمایند [۶]. روش کنترل دسترسی پیشنهادی برای تجزیه و تحلیل حفظ حریم خصوصی کاربر به کاربر می‌رود. یکی از مشکلات شبکه‌های اجتماعی که کاربران با آن روزهان رو به رو می‌شوند افراد ناشناسی که برای سوء استفاده از اطلاعات شخصی کاربران منبع به آن‌ها درخواست دوستی می‌دهند و اکثر کاربران منبع بدون هیچ اطلاعاتی راجع به درخواست دهنده ندانسته درخواست را قبول می‌کنند که این امر باعث آسیب‌پذیری می‌شود. مکانیسم‌های امنیتی موجود در شبکه‌های اجتماعی نمی‌تواند به طور مؤثر از اطلاعات خصوصی کاربران محافظت کند. در پژوهش مذکور مدلی ارائه شد که در صد آسیب‌پذیری کاربران شبکه‌های اجتماعی را کمتر می‌کند. مدل ارائه شده با گذاشتن یک قانون به نام قانون دوست مشترک که باعث زیاد شدن اعتماد بین کاربر منبع و کاربر درخواست‌دهنده در شبکه اجتماعی می‌شود. به طوری که اگر تعداد دوستان مشترک بین کاربر منبع و کاربر درخواست‌دهنده زیاد باشد آن درخواست در قسمت درخواست‌های قابل اعتماد قرار می‌گیرد و اگر تعداد دوستان مشترک کم باشد آن درخواست در قسمت درخواست‌های غیرقابل اعتماد قرار می‌گیرند. بنابراین کاربر منبع متوجه می‌شود که چه کاربرانی قابل اعتماد هستند و چه کاربرانی غیرقابل اعتماد می‌باشند. اما در آخر خود کاربر منبع تصمیم می‌گیرد که چه درخواست‌هایی را قبول کند و یا چه درخواست‌هایی را رد کند.

چنگ و همکاران در تحقیقی به بررسی تصمیم برای استفاده از یک شبکه اجتماعی آنلاین به‌صورت کنش اجتماعی عمدی مفهوم سازی و تأثیر نسبی سه مد فرآیندهای نفوذ و تأثیر اجتماعی (رعایت و تطابق درونی سازی و شناسایی) بر کاربرد کنش اجتماعی عمدی (قصد جمعی) مورد بررسی قرار دادند. نتایج بررسی بر روی تحقیق تجربی کاربران فیس بوک (تعداد افراد = ۳۸۹ نفر) نشان داد، قصد جمعی برای استفاده از سایت شبکه اجتماعی بر طبق هندجار ذهنی و هویت اجتماعی تعیین می‌گردد. به علاوه هویت اجتماعی ساختار نهفته درجه دومی است که از مؤلفه‌های درجه اول شناختی ارزیابی و عاطفی تشکیل شده است [۷]. سورانی و همکاران در مقاله‌ای به برخی از پژوهش‌ها و تحقیقات انجام‌شده پیرامون اطلاعات سایت‌های شبکه اجتماعی با مرور برخی از

### ۳-۱ اصول مقدماتی

مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر<sup>۳</sup> که توسط چنگ و همکاران معرفی شده است [۷]، روابط کاربر به کاربر<sup>۴</sup> در شبکه های اجتماعی آنلاین<sup>۵</sup> را به منظور صدور مجوز<sup>۶</sup> دریافت کرده و یک ز بان تعیین خط مشی<sup>۷</sup> مبتنی بر عبارات منظم<sup>۸</sup> تعریف می کند. ابتدا مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر را که مدل پیشنهادی ما بر اساس آن است به طور خلاصه بیان می کنیم.

### ۳-۱-۱ علامت های پایه

علامت  $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n, \sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\}$  در مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر، مجموعه ای از انواع روابط را مشخص می کند. با در نظر گرفتن یک رابطه از نوع  $\sigma_i \in \Sigma$ ، معکوس رابطه بصورت  $\sigma_i^{-1} \in \Sigma$  خواهد بود. یک اقدام<sup>۹</sup> دارای یک فرم فعال و یک فرم منفعل است، که به ترتیب با اقدام و اقدام<sup>-۱</sup> نشان داده می شوند. اگر "آلیس" بتواند "باب" را پوک (یک اصطلاح در شبکه های اجتماعی)<sup>۱۰</sup> کند، از دیدگاه "آلیس" عملکرد پوک ولی از دیدگاه "باب" پوک<sup>-۱</sup> خواهد بود.

### ۳-۱-۲ اجزای مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر

اجزای این مدل در شکل (۱)، نشان داده شده است. کاربر در حالی دسترسی<sup>۱۱</sup> نشان دهنده یک کاربر فعال است که درخواست دسترسی به یک هدف را دارد و حامل خط مشی های کاربر در حال دسترسی<sup>۱۲</sup> است. هر اقدام نشان دهنده عملیاتی است که توسط یک کاربر در حال دسترسی در برابر یک هدف آغاز شده است. یک عملکرد برای کاربر در حال دسترسی به صورت اقدام و برای هدف به صورت اقدام<sup>-۱</sup> نشان داده می شود. اهداف می توانند کاربران هدف<sup>۱۳</sup> باشند (به عنوان مثال، یک کاربر، کاربر دیگر را پوک کند و منبع هدف<sup>۱۴</sup> باشد). همچنین یک کاربر کنترل کننده<sup>۱۵</sup> وجود دارد که دسترسی به یک منبع را کنترل می کند و مانند مالک منبع

اطلاعات برحسب حداقل زمان تکثیر وابسته است. این مطالعه کارایی انتشار اطلاعات در میکرو بلاگ شبکه های اجتماعی آنلاین را نشان می دهد. این کارایی با انجام آزمایش های تجربی تحلیل ثغوریکی و شبیه سازی عددی تائید می شود [۱۱]. لایت و همکاران در مقاله ای به یکی از روش های نظریه پردازی درباره تجربه افراد از سایت های شبکه اجتماعی به مفهوم قطع ارتباط پرداختند. با معرفی و توسعه یک دیدگاه مبتنی بر قطع ارتباط نشان دادند که استراتژی های قطع ارتباط توقف و جلوگیری جزء الزامات عملیاتی کاربران و صاحبان سایت هایی نظیر فیس بوک است. در واقع پژوهش آنها اثبات کرد که قطع ارتباط در این محیط ها را نباید فقط به عنوان روشی برای مقاومت در برابر این سایت ها و ترک آنها در نظر گرفت بلکه این قطع ارتباط می تواند به عنوان یک عامل تسهیل گر اجتماعی اقتصادی نیز عمل کند [۱۲].

با بررسی پیشینه ی پژوهش، یک مدل کنترل دسترسی رابطه محور کاربر به کاربر جدید برای شبکه های اجتماعی آنلاین پیشنهاد شده است که امکان بیان سیاست های کنترل دسترسی دقیق تر و حرفه ای تری را برحسب الگو و عمق روابط میان کاربران در شبکه را فراهم می کند. سیاست های کنترل دسترسی در مورد کاربران و منابع برحسب اقدام ثبت شده انواع روابط، نقطه شروع ارزیابی و تعداد جهش ها در مسیر تشکیل شده اند. روش پیشنهادی، شاخصه ای منظم از کاراکترها را به کار گرفته و از علائم کمی جهت بیان مسیر رابطه استفاده می کند که نشان دهنده روابط غیرمستقیم در میان کاربران می باشد. استفاده از رابطه مبتنی بر عبارات منظم و روابط چندگانه موجب توانایی ارائه سیاست های مشخص نسبت به مدل های انجام شده در کارهای پیشین می گردد. این مدل به عنوان اولین مدل کنترل دسترسی رابطه ای برای شبکه های اجتماعی آنلاین با این قابلیت شناخته می شود.

### ۳-۲ روش پیشنهادی

در این قسمت به بیان مفاهیم پایه و سپس روش پیشنهادی پرداخته شده است.

<sup>10</sup> Poke

<sup>11</sup> Accessing user (u<sub>a</sub>)

<sup>12</sup> Accessing User Policies (AUP)

<sup>13</sup> Target User (TU)

<sup>14</sup> Target Resource (TR)

<sup>15</sup> Controlling User (CU)

<sup>3</sup> User-to-User Relationship-based Access Control (UURAC)

<sup>4</sup> User-to-User

<sup>5</sup> Online Social Networks (OSNs)

<sup>6</sup> Authorization

<sup>7</sup> Policy

<sup>8</sup> Expression-based

<sup>9</sup> Action

است. نمودار اجتماعی یک شبکه اجتماعی آنلاین به صورت سه‌گانه  $G = \langle U, E, \Sigma \rangle$  مدل‌سازی می‌شود؛ که در آن،  $U$  یک مجموعه محدود از کاربران ثبت‌شده در سیستم است و به صورت گره‌هایی در نمودار نشان داده شده است.

در این تعریف  $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n, \sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\}$  یک مجموعه محدود از انواع روابط را نشان می‌دهد، که هر شاخص<sup>۲۴</sup> از نوع  $\sigma$  نشان‌دهنده یک نوع ارتباط است که در سیستم پشتیبانی می‌شود و  $E \subseteq U \times U \times \Sigma$  نشان‌دهنده پال‌های گراف اجتماعی و مجموعه‌ای از روابط موجود کاربر است. برای هر  $\sigma_i \in \Sigma$  یک  $\sigma_i^{-1} \in \Sigma$  وجود دارد که نشان‌دهنده معکوس رابطه نوع  $\sigma_i$  است. برای جلوگیری از بروز خطا و بالا بردن دقت، می‌بایست روابط اصلی و معکوس آن‌ها همواره بر روی نمودار اجتماعی به طور هم‌زمان وجود داشته باشند. با در نظر گرفتن یک کاربر  $u \in U$ ، یک کاربر  $v \in U$  و یک رابطه نوع  $\sigma \in \Sigma$ ، رابطه  $(u, v, \sigma)$  بیان می‌کند که رابطه نوع  $\sigma$  از کاربر  $u$  شروع و در  $v$  پایان می‌پذیرد. این رابطه همیشه دارای فرم معادل  $(v, u, \sigma^{-1})$  است.

### ۳-۲ مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر

در این بخش، مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر را توسعه می‌دهیم تا معیارها و اجرای خط‌مشی ReBAC مشخصه آگاه<sup>۲۵</sup> ساده‌تر شود.<sup>۲۶</sup>

### ۳-۲-۱ مشخصه‌ها<sup>۲۷</sup> در شبکه‌های اجتماعی آنلاین

شبکه‌های اجتماعی آنلاین اطلاعات زیادی مربوط به مشخصه‌های کاربران و منابع را نگهداری می‌کنند. کاربران اطلاعات پروفایل موردنیاز در شبکه‌های اجتماعی مانند نام، سن، جنسیت و غیره را نگه می‌دارند. هنگامی که یک بخش از منبع در شبکه اجتماعی آنلاین آپلود می‌شود، ارائه‌دهنده منبع<sup>۲۸</sup> نیز می‌تواند بعضی از

دارای رابطه کاربر-به-منبع<sup>۱۶</sup> با خودش است. دسترسی به یک کاربر هدف بر اساس خط‌مشی‌های کاربر هدف<sup>۱۷</sup> کنترل می‌شود که با استفاده از روابط بین کاربران هدف و کاربران درخواست‌کننده دسترسی پیکربندی می‌شود. این درحالی است که دسترسی به یک منبع هدف بر اساس خط‌مشی‌های منبع هدف<sup>۱۸</sup> کنترل می‌شود که با استفاده از روابط بین کاربر کنترل‌کننده و درخواست‌کننده دسترسی پیکربندی شده است.

یک درخواست دسترسی  $\langle u_a, \text{act}, \text{target} \rangle$  نشان‌دهنده شروع یک دسترسی است که  $\text{act} \in \text{Act}$  نوع دسترسی درخواست شده [توسط کاربر در حال دسترسی به هدف] را تعیین می‌کند. اگر  $u_a$  درخواست برقراری ارتباط متقابل با یک کاربر دیگر داشته باشد، هدف  $u_t \in U$  خواهد بود. اگر  $u_a$  تلاش کند به منبعی که متعلق به یک کاربر دیگر است یعنی  $u_c$  دسترسی پیدا کند، هدف همان منبع  $r_t \in R$  خواهد بود که در آن  $R$  یک مجموعه محدود از منابع در شبکه اجتماعی آنلاین است. یک خط‌مشی، شامل قوانینی است که تنظیمات لازم جهت صدور مجوز را تعیین می‌کنند. خط‌مشی‌ها می‌توانند به صورت مشخص شده توسط سیستم<sup>۱۹</sup> یا مشخص شده توسط کاربر<sup>۲۰</sup> باشند. خط‌مشی‌های مشخص شده توسط سیستم<sup>۲۱</sup>، قوانین دستگاهی گسترده‌ای هستند که توسط سیستم شبکه اجتماعی آنلاین اجرا می‌شوند درحالی که سیاست‌های مشخص شده توسط کاربر برای کاربران یا منابع خاص اعمال می‌شود. سیاست‌های مشخص شده توسط کاربر شامل TUP, AUP و TRP است که به ترتیب همان خط‌مشی‌های کاربر در حال دسترسی، کاربر هدف و منبع هدف هستند.

### ۳-۱-۳- نمودار اجتماعی<sup>۲۲</sup>

همان‌طور که در شکل ۲ نشان داده شده است، شبکه اجتماعی آنلاین به‌عنوان یک گراف ساده علامت‌دار جهت‌دار خلاصه شده است که در آن هر گره یک کاربر و هر یال<sup>۲۳</sup> مربوط به یک رابطه کاربر با کاربر

<sup>۲۵</sup> Attribute-Aware

<sup>۲۶</sup> تأکید می‌کنیم که در مدل کنترل دسترسی مبتنی بر روابط کاربر به کاربر فقط روابط کاربر به کاربر در نظر گرفته می‌شود بنابراین منابع تنها می‌توانند به‌عنوان هدف مسیر ارتباط واقع شوند و مسیر ارتباط فقط می‌تواند شامل کاربران باشد.

<sup>۲۷</sup> Attributes

<sup>۲۸</sup> Resource Provider

<sup>۱۶</sup> User-to-Resource (U<sup>2</sup>R)

<sup>۱۷</sup> Target User Policies (TUP)

<sup>۱۸</sup> Target Resource Policies (TRP)

<sup>۱۹</sup> System-Specified (SS)

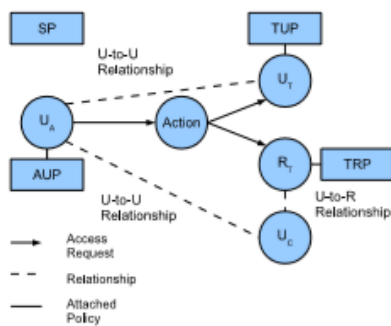
<sup>۲۰</sup> User-Specified (US)

<sup>۲۱</sup> System-Specified Policies (SSP)

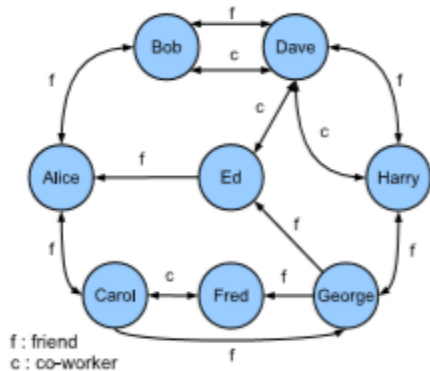
<sup>۲۲</sup> Social Graph

<sup>۲۳</sup> Edge

<sup>۲۴</sup> Specifier



شکل ۱. اجزای مدل [۷]



شکل ۲. یک گراف اجتماعی ساده [۷]

همچنین گره ها می توانند برای یک یا چند شیء در نظر گرفته شوند. یک نمونه برای مشخصه های چند یال، اعتماد انتقال پذیری<sup>۳۲</sup> بین دو گره است که به طور مستقیم متصل نیستند<sup>۳۳</sup>. به عنوان مثال، ارزش اعتماد دو یا چند یال باید در نظر گرفته شود تا اعتماد کلی بین کاربر در حال دسترسی و کاربر هدف/کنترل کننده محاسبه شود. مشخصه هایی که گره های چندگانه را توصیف می کنند بیشتر در شبکه های اجتماعی آنلاین دیده می شوند، مانند سن متوسط، موقعیت متداول یا عادت مشترک بین افراد. مشخصه های گره و یال مربوطه نیز می توانند برای ایجاد ترکیبات خط مشی ای مونتاژ شوند. به عنوان مثال، "آلیس" ممکن است یک خط مشی ای را مشخص کند که "فقط کاربرانی که دارای بیش از ۰/۵ اعتماد با "باب" هستند بتوانند به اطلاعات او دسترسی داشته باشند."

**مشخصه تعداد:** مشخصه ی تعداد مواردی را توصیف می کند که با هیچ گره و یالی مرتبط نباشند. این مشخصه، الزام رخداد برای

اطلاعات جانبی<sup>۲۹</sup> را در مورد منبع پیوست کند. می توانیم خط مشی های مبتنی بر این اطلاعات مشخصه مرتبط با کاربران و منابع را تعریف کنیم. باین حال، اکثر سیستم های ReBAC بر روی برخی از جنبه های خاص روابط مانند نوع، عمق و قدرت تمرکز کرده اند. این امر باعث می شود که ReBAC نسبتاً ساده و کارآمد باشد اما استفاده از ReBAC را از نظر قابلیت کنترل محدود می کند. در سال های اخیر، مطالعات بر روی کنترل دسترسی مبتنی بر مشخصه<sup>۳۰</sup> ABAC نشان داده است که اطلاعات متنی مختلف کاربر، منبع و محاسبات محیط می تواند برای کنترل های بسیار انعطاف پذیر و دقیق تر استفاده شود [۷]. باین حال، راه حل های فعلی ABAC به راحتی قابل استفاده در ReBAC در شبکه های اجتماعی آنلاین نیستند. در حالی که مدل های معمول ABAC فقط مشخصه های کاربر در حال دسترسی، منبع هدف و گاهی اوقات محیط محاسبات را در نظر می گیرند، ReBAC مشخصه آگاه باید مشخص کند که کدام صفات و صفات چه کسانی (یعنی مشخصه های کاربر، مشخصه های ارتباط) در مسیر ارتباط بین کاربران در حال دسترسی و کاربران هدف/کنترل کننده باید بررسی شود. برای ReBAC مشخصه آگاه، سه نوع مشخصه را به صورت زیر تعیین می کنیم: مشخصه گره (کاربر، منبع)، مشخصه یال (رابطه) و مشخصه تعداد.<sup>۳۱</sup>

**مشخصه گره:** کاربران و منابع به صورت گره هایی بر روی نمودار اجتماعی نمایش داده می شوند. کاربران مشخصه هایی دارند که هویت و ویژگی های خود مانند نام، سن، جنسیت و غیره را تعریف می کنند. ویژگی های منبع ممکن است شامل عنوان، مالک، تاریخ و غیره باشد.

**مشخصه یال:** هر یال با مشخصه هایی مرتبط است که ویژگی های یال را توصیف می کنند. چنین مشخصه هایی ممکن است شامل رابطه وزن ها، نوع ها و غیره باشد.

<sup>۳۲</sup> Transitive Trust

<sup>۳۳</sup> منظور از trust ایجاد اعتماد برای اشتراک گذاری اطلاعات می باشد.

<sup>۲۹</sup> Metadata

<sup>۳۰</sup> Attribute-Based Access Control (ABAC)

<sup>۳۱</sup> Count Attribute

## جدول ۱. کمیت سنج‌های مشخصه

$\forall [+m, -n]$	تمام اشخاص از $m^{\text{th}}$ به $n^{\text{th}}$ ام آخر به طوری که، $m + n \leq h$ و $n$ عدد صحیح غیر منفی هستند و $h$ یک حد حلقه شمارنده است.
$\forall [+m, +n]$	تمام اشخاص از $m^{\text{th}}$ به $n^{\text{th}}$ ام آخر به طوری که $m \leq n \leq h$ باشد.
$\forall [-m, -n]$	تمام اشخاص از $m^{\text{th}}$ ام آخر به $n^{\text{th}}$ ام آخر به طوری که $h \geq m \geq n$ باشد.
$\exists [+m, -n]$	یک نهاد از $m^{\text{th}}$ به $n^{\text{th}}$ ام آخر به طوری که $m + n \leq h$ باشد.
$\exists [+m, +n]$	یک نهاد از $m^{\text{th}}$ به $n^{\text{th}}$ ام به طوری که $m \leq n \leq h$ باشد.
$\exists [-m, -n]$	یک نهاد از $m^{\text{th}}$ ام آخر به $n^{\text{th}}$ ام آخر به طوری که $h \geq m \geq n$ باشد.
$\forall \{ \mathcal{N}^{\pm} \}$	تمام نهادهایی که در این مجموعه وجود دارند.
$\exists \{ \mathcal{N}^{\pm} \}$	یک نهادی که در این مجموعه وجود دارد.

برای بیان یک محدوده یا موقعیت دقیق در مسیر، از علائم + و - برای نشان دادن جهت‌های مستقیم (از ابتدا) و عقب (از انتها) استفاده می‌کنیم، بعد از این علامت یک عدد که موقعیت را از جلو یا عقب نشان می‌دهد قرار می‌گیرد. توجه داشته باشید که شاخص برای کاربران از ۰ شروع می‌شود، در حالی که شاخص برای روابط از ۱ آغاز می‌شود. به عنوان مثال، برای کاربران، +۰ به معنای کاربر شروع‌کننده و ۱- نشان‌دهنده آخرین کاربر با رابطه از نوع دوم<sup>۳۵</sup> موجود در مسیر است؛ در حالی که برای روابط، ۱+ نشان‌دهنده اولین رابطه در مسیر و ۲- آخرین رابطه از نوع دوم است. علامت‌های مثبت و منفی در دو سطر گذشته، به جای معنا و مفهوم طبیعی ریاضی، جهت مستقیم یا عقب را نشان می‌دهند.

یک قانون خط‌مشی مبتنی بر مشخصه متشکل از یک کمیت سنج تعیین‌کننده مقدار گره /یال خاص، یک عملکرد بولی از مشخصه‌های (گره/یال) و یک پیش‌بین<sup>۳۶</sup> مشخصه تعداد  $\text{count} \geq i$  به شرح زیر است:

$\langle\langle \text{quantifier}, f(\text{ATTR}(N), \text{ATTR}(E)), \text{count} \geq i \rangle\rangle$

ویژگی‌های مسیر مبتنی بر مشخصه را نشان می‌دهد و پایین‌ترین حد وقوع چنین مسیری را مشخص می‌کند.

## ۲-۲-۳ فرمول‌بندی خط‌مشی مبتنی بر مشخصه

خط‌مشی مبتنی بر مشخصه، نیازهای کنترل دسترسی را مشخص می‌کند که مربوط به مشخصه‌های کاربران و روابط آن‌هاست. در اینجا، به‌طور قراردادی زبان اساسی مبتنی بر خط‌مشی مشخصه‌ها را تعریف می‌کنیم.

- $E$  و  $N$  به ترتیب گره‌ها و یال‌ها هستند.
- $N A_k (1 \leq k \leq K)$  و  $E A_l (1 \leq l \leq L)$  به ترتیب مشخصه‌های از پیش تعریف‌شده برای گره‌ها و یال‌ها هستند که  $K$  تعداد مشخصه‌های گره و  $L$  تعداد مشخصه‌های یال است.
- $\text{ATTR}(n)$  و  $\text{ATTR}(e)$  به ترتیب تخصیص مشخصه برای گره  $n$  و یال  $e$  است، به طوری که  $\text{ATTR}(n) \in N A_1 \times N A_2 \times \dots \times N A_K$  و  $\text{ATTR}(e) \in E A_1 \times E A_2 \times \dots \times E A_L$  هر مشخصه تنها دارای یک مقدار در دامنه تعریف شده برای آن است.

در مسیر ارتباط بین دو کاربر در شبکه‌های اجتماعی آنلاین، ممکن است کاربران بسیار دیگری وجود داشته باشند که با روابط متفاوت متصل شده‌اند. هر کاربر یا رابطه دارای مشخصه‌هایی است که می‌تواند برای تعیین قوانین کنترل دسترسی استفاده شود. در برخی موارد، مشخصه‌های تمام کاربران یا روابط در مسیر باید مورد توجه قرار گیرد اما گاهی اوقات تنها مشخصه‌های برخی از کاربران یا روابط استفاده می‌شود. همان‌طور که در جدول (۱) نشان داده شده است، از کمیت سنج<sup>۳۴</sup> جامع  $\forall$  و کمیت سنج همگانی  $\exists$  به ترتیب برای نشان دادن "همه" و "حداقل یک" کاربر(ها) یا رابطه (ها) استفاده می‌کنیم. علامت  $[]$  برای نشان دادن محدوده‌ها در مسیر رابطه استفاده می‌شود، در حالی که  $\{\}$  مجموعه‌ای از کاربران/روابط واقع شده در یک فاصله خاص در مسیر بین کاربر در حال دسترسی و کاربر هدف/کنترل‌کننده را نشان می‌دهد.

<sup>۳۵</sup> Predicate

<sup>۳۴</sup> Quantifier

<sup>۳۶</sup> Second Last User



غیرفعال یک عمل به خط‌مشی‌های مشخص شده سیستم اعمال نمی‌شود، زیرا این خط‌مشی‌ها با هیچ موجودیت خاص در عمل مرتبط نیست. با این حال، هنگام تعیین خط‌مشی یک سیستم برای یک منبع، می‌توانیم به صورت اختیاری منبع را از لحاظ نوع منبع (r.typevalue, r.typevalue) اصلاح کنیم. جدول ۴ قالب<sup>۳۷</sup> قواعد گراف را با استفاده از فرم Backus-Naur<sup>۳۸</sup> تعریف می‌کند. هر قانون گراف یک گره شروع شونده<sup>۳۹</sup> و یک قانون مسیر<sup>۴۰</sup> را مشخص می‌کند. گره شروع شونده، کاربری را نشان می‌دهد که در آن ارزیابی خط‌مشی در آن شروع شده است. یک قانون مسیر نشان‌دهنده یک مجموعه از خصوصیات مسیر است. هر مشخصه مسیر شامل یک جفت (مسیر<sup>۴۱</sup>، حلقه شمارنده) است که الگوی مسیر ارتباط بین دو کاربر و حداکثر تعداد یال‌های مسیر را تعیین می‌کند که برای دسترسی پیدا کردن باید برآورده شود. مشخصه‌های مسیر چندگانه را می‌توان با ربط دهنده "∧" و مجزا ساز "∨" متصل کرد. "-" بر روی مشخصه‌های مسیر نشان‌دهنده عدم وجود جفت مشخص شده الگوی ارتباطی و حد حلقه شمارنده است. الگوی مسیر ارتباط مسیر نشان‌دهنده یک دنباله از تصریح کننده‌های<sup>۴۲</sup> نوع از گره شروع شونده به گره ارزیابی است. برخلاف مدل کنترل دسترسی رابطه محور کاربر به کاربر، به منظور تسهیل خط‌مشی‌های مبتنی بر مشخصه عبارت جدید AttPolicy را به دستورالعمل اضافه می‌کنیم. AttPolicy می‌تواند پس از مشخصه تمام مسیر (مسیر، حلقه شمارنده) یا یک بخش از الگوی مسیر یافت شود. چیزی که به (مسیر، حلقه شمارنده) اعمال می‌شود خط‌مشی مبتنی بر مشخصه فراگیر گفته می‌شود. هنگامی که آن از یک بخش مسیر پیروی می‌کند، یک خط‌مشی مبتنی بر مشخصه‌های محلی است که تنها برای این بخش قابل استفاده است. برای سادگی، مثال‌ها از این پس تنها از مشخصه‌های مبتنی بر ویژگی‌های فراگیر استفاده می‌کنند. اکنون نشان می‌دهیم که چگونه قوانین مبتنی بر مشخصه‌ها می‌توانند به برخی از نمونه‌های UURACA اعمال شوند.

مثال ۱: خط‌مشی مشخصه گره و مشخصه تعداد. "آلیس" می‌خواهد مشخصات خود را برای کاربرانی که حداقل پنج دوست از

توجه داشته باشید که یک کمیت سنج به یک تابع گره/یال اعمال می‌شود، اما به پیش‌بین مشخصه تعداد اعمال نمی‌شود. به عنوان مثال، R<sup>۳</sup> قانونی را بیان می‌کند که "حداقل باید پنج ارتباط مشترک بین درخواست کنندگان و مالک باشد، که شغل آنان دانش‌آموز است". در R<sup>۲</sup> و R<sup>۳</sup>، از پیش‌بین مشخصه تعداد استفاده نمی‌شود و این به عنوان "-" نشان داده شده است که به طور پیش فرض  $count \leq 1$  را نشان می‌دهد. در اینجا، این خط‌مشی حاوی یک قانون است که نشان می‌دهد "یک کاربر که از طریق بزرگسالانی که آدرسشان "تگزاس" است می‌تواند دسترسی داشته باشد". R<sup>۳</sup> نیازمند آن است که در مسیر بین کاربر در حال دسترسی و کاربر هدف/در حال کنترل، کاربران باید در سه فاصله خاص با بزرگسالان باشند.

- R<sup>۱</sup> :  $(\exists [+1, -1], occupation(u) = \text{"student"}, count \geq 5)$
- R<sup>۲</sup> :  $(\forall [+1, -1], (age(u) \geq 18) \wedge (address(u) = \text{"Texas"}), -)$
- R<sup>۳</sup> :  $(\forall \{+1, +2, -1\}, (age(u) \geq 18), -)$

### ۳-۲-۳ مشخصه‌های خط‌مشی

خط‌مشی‌های مبتنی بر مشخصه در مسیرهای ارتباطی خاص بین کاربر در حال دسترسی و کاربر هدف/در حال کنترل اعمال می‌شود. برای این منظور، زبان توصیف خط‌مشی مبتنی بر عبارات منظم پیشنهاد شده در [۷] را گسترش می‌دهیم. جدول ۲، فهرستی از نمادهای مورد استفاده در زبان توصیف خط‌مشی را تعریف می‌کند. خط‌مشی‌های مدل کنترل دسترسی رابطه محور کاربر به کاربر مشخصه آگاه شامل دو بخش است: یک عملیات درخواست شده و یک قانون گراف که دسترسی را بر اساس نمودار اجتماعی تعیین می‌کند. همان طور که در جدول ۳ نشان داده شده است، انواع مختلفی از خط‌مشی‌ها را شناسایی کرده‌ایم. عملیات به شکل غیرفعال اقدام<sup>۴۱</sup> در خط‌مشی کاربر هدف و خط‌مشی منبع هدف نشان داده شده‌اند، چرا که منبع/کاربر هدف همیشه گیرنده عمل است. خط‌مشی منبع هدف دارای پارامتر اضافی u<sub>c</sub> است که نشان‌دهنده کاربر در حال کنترل منابع است. تمایز فرم فعال و

<sup>41</sup> Path

<sup>42</sup> Specifiers

<sup>37</sup> Syntax

<sup>38</sup> Backus-Naur Form (BNF)

<sup>39</sup> Startingnode

<sup>40</sup> Pathrule

سیاست سیستم برای کاربر	<اقدام، نمودار>
سیاست سیستم برای منابع	<r.typevalue>, <r.typevalue>

جدول ۴. دستور زبان برای قوانین گراف

<p>“StartingNode”, ”PathRule”) → قوانین گرافیکی          Connective  AttP athSpecExp  AttP athSpecExp → قوانین مسیر          PathRule          AttP athSpecExp → PathSpecExp  PathSpecExp : ”AttP olicy          ∨  ∧ → ارتباطات          PathSpec  ”-”PathSpec → مسیره‌های ویژه          ”HopCount“, ”AttP ath”) → مسیره‌های اختصاصی          ”HopCount“, ”EmptySet”) → مسیره‌های اختصاصی          اعداد → شمارنده حلقه          مسیر   مسیر → پیوست مسیر          مسیر   مسیر   انواع توالی → مسیر          ∅ → فضای خالی          انواع توالی ”   انواع پیوست توالی   پیوست توالی → انواع توالی          ”AttP olicy : ”TypeExp  TypeExp → پیوست انواع توالی          کلمات انواع مشخص کننده  انواع مشخص کننده → انواع خروجی ها          برای پردازی از تحلیل کننده اختصاصی استفاده کنید → خط و مشی          ua ut uc → گره شروع          σ<sub>1</sub> σ<sub>2</sub>   ...  σ<sub>n</sub> σ<sub>1</sub><sup>-1</sup> σ<sub>2</sub><sup>-1</sup>   ...  σ<sub>n</sub><sup>-1</sup>  Σ → انواع ویژه کلمات          ” + ”   ”* ” → کلمات          [۰ - ۹]+ → اعداد</p>
---

مثال ۲: خط‌مشی مشخصه یال. آلیس به کاربران اجازه دسترسی به تصویر ۱ را می‌دهد در صورتی که کاربر در فاصله ۳ حلقه (جهش) قرار داشته باشد و بتواند در مسیری با حداقل اعتماد (۰/۵) روابط دوستانه در هر حلقه به او دست یابند. چنین خط‌مشی‌ای به شرح زیر است:

- P۳: <read, Photo, (ua, ((f\*, ۳) : ∇[+۱, -۱], trust(r) ≥ ۰, ۵, ))>

سیستم قبل از اعطای دسترسی، هر یال موجود در مسیر را بررسی می‌کند تا اطمینان حاصل شود که ارزش اعتماد آن مطابق نیاز است.

مثال ۳: اتخاذ یک خط‌مشی مدل کنترل دسترسی رابطه محور کاربر به کاربر. خط‌مشی زیر فقط شامل الزامات مبتنی بر رابطه (f\*, ۳) می‌باشد، درحالی که مشخصه‌های گره/یال و مشخصه تعداد هر دو خالی هستند:

- P۴: <poke, (ua, (f\*, ۳) : ∃[+۰, -۰], , ))>

دانش‌آموزان مشترک دارند به اشتراک بگذارد. او می‌تواند خط‌مشی زیر را برای دوستان - دوستان خود مشخص کند:

- P۱: <profile access, (ua, ((ff, ۲) : ∃[+۱, -۱], occupation(u) = “student”, count ≥ ۵))>

اگر او بخواهد به کسی که یک دوست مشترک "باب" را با او به اشتراک گذاشته است، اجازه بدهد پروفایلش را ببیند، این خط‌مشی را می‌توان به شرح زیر بیان کرد:

- P۲: <profile access, (ua, ((ff, ۲) : ∃[+۱, -۱], name(u) = “Bob”, ))>

برای P۱، سیستم نیاز به پیدا کردن مسیره‌هایی دارد که با (f f, ۲) مطابقت داشته و مشخصه شغلی کاربران در مسیره‌ها را بررسی کند. اگر حداقل پنج مسیر با این مشخصات وجود داشته باشد، به U<sub>a</sub> اجازه داده می‌شود تا اطلاعات پروفایل هدف را ببیند. برای P۲، یک بار مسیر (f f, ۲) پیدا شده است و نام کاربر در مسیر برابر با "باب" است، و لذا سیستم اجازه دسترسی به آن را خواهد داد.

جدول ۲. علامت‌های مشخصات خط‌مشی

اتصال بده دو کارکتر را با $\sigma \in \Sigma$ یا بال مجموعه خودش پایان - به - پایان که نشانه ی یکسری وقایع از انواع رابطه‌هاست.	الحاق (۰)
علامت اتحاد وهمبستگی $\sigma$ که بزرگتر/مساوی صفر است	ستاره (*)
متصل کردن $\sigma$ را یک یا چند بار نشان می‌دهد.	جمع (+)
وقایع $\sigma$ را صفر یا یک بار را نشان می‌دهد.	علامت سؤال (?)
تقسیم مشخصات چندین مسیر را نشان می‌دهد.	اتصال ناپیوسته (V)
نشانه‌گر پیوند مشخصات چندین مسیر است.	رابطه متصل (∧)
عدم وجود جفت مشخص رابطه را نشان می‌دهد.	علامت منفی (-)
الگوی رابطه و سیاست های مبتنی بر ویژگی را از هم جدا می‌کند.	دو نقطه (:)

جدول ۳. نمایش خط‌مشی کنترل دسترسی

سیاست کاربر	<اقدام، نمودار>
سیاست کاربر هدف	<act <sup>-1</sup> , نمودار>
سیاست منبع هدف	<act <sup>-1</sup> , u <sub>c</sub> , نمودار>

که بر اساس یک استراتژی منسجم، هماهنگ و یا اولویت بندی شده عمل می کند.

**الگوریتم بررسی مسیر مشخصه-آگاه:** همان طور که در شکل ۲ نشان داده شده است الگوریتم بررسی مسیر برای عبور از نمودار گرافیکی  $G$  از یک استراتژی جستجوی اول عمق (DFS)<sup>۴۳</sup> از گره شروع  $S$  استفاده می کند. هدف پیدا کردن مسیرهای ارتباط بین گره شروع  $S$  و گره ارزیابی  $t$  است که این مسیرها الزامات خط مشی را برآورده سازند. الگوی مسیر  $path$  و محدودیت حلقه شمارنده الزامات مبتنی بر رابطه را مشخص می کنند درحالی که  $globalattpol$  نشان دهنده قوانین مبتنی بر مشخصه است.

الگوریتم ۱: ارزیابی دسترسی
۱. (Policy Collecting Phase) ۲. if target = $u_t$ then ۳. $AUP \leftarrow u_a$ 's policy for act, $TUP \leftarrow u_t$ 's policy for act <sup>-1</sup> , $SP \leftarrow$ system's policy for act ۴. else ۵. $u_c \leftarrow$ owner( $r_t$ ), $AUP \leftarrow u_a$ 's policy for act, $TRP \leftarrow u_c$ 's policy for act <sup>-1</sup> on $r_t$ , $SP \leftarrow$ system's policy for act, $r$ .type ۶. (Policy Evaluation Phase) ۷. for all policy in AUP, TUP/TRP and SP do ۸.     Extract graph rules (start, path rule) from policy ۹. for all graph rule extracted do ۱۰.     Determine the starting node, specified by start, where the path evaluation starts ۱۱.     Determine the evaluating node which is the other user involved in access ۱۲.     Extract path rules path rule from graph rule ۱۳.     Extract each path spec path, hopcount and/or attribute rule attpolicy from path rules ۱۴.     Simultaneously path-check each path spec and evaluate the corresponding attribute rule using Algorithm ۳-۲ ۱۵.     Evaluate a combined result based on conjunctive or disjunctive connectives between path specs ۱۶.     Compose the final result from the result of each policy

الگوریتم ۲: DFSPATHCHECKER( $G$ , path, hopcount, $s$ , $t$ , $globalattpol$ )
۱. $DFA \leftarrow REtoDFA(path)$ ; $currentPath \leftarrow NIL$ ; $d \leftarrow 0$ ۲. $stateHistory \leftarrow$ DFA starts at the initial state

مدل UURACA به طور یکپارچه با مدل کنترل دسترسی رابطه محور کاربر به کاربر، سازگار است. مثال ۳ نشان می دهد که چگونه خط مشی مدل کنترل دسترسی رابطه محور کاربر به کاربر، سیاست را می توان در UURACA اتخاذ کرد.

#### ۴- الگوریتم

این بخش به بررسی دسترسی UURACA می پردازد. مدل کنترل دسترسی رابطه محور کاربر به کاربر، یک الگوریتم بررسی مسیر را فراهم می کند تا مسیر صحیح بین درخواست دهنده دسترسی و هدف (و یا مالک منبع) را که مطابق با نیازهای ReBAC است، پیدا کند. برای اجرای خط مشی های مبتنی بر مشخصه، ارزیابی دسترسی باید در زمان بررسی مسیر، سیاست های مبتنی بر مشخصه را در نظر بگیرد. ممکن است فردی بررسی کردن مشخصه را بر روی مسیرهای حاصل که توسط الگوریتم پژوهش حاضر پیدا می شود، اجرا کند. باین حال، این احتمال وجود دارد که این کار ناکارآمد باشد. در این تحقیق، یک الگوریتم مسیریابی اصلاح شده را برای ترکیب کردن ارزیابی یک خط مشی مبتنی بر مشخصه در طی روند یافتن مسیر، ارائه می کنیم.

**روش ارزیابی دسترسی:** درخواست های دسترسی می توانند به روش توضیح داده شده در الگوریتم ۱ ارزیابی شوند. برای یک درخواست دسترسی اقدام ( $Ua$ )، سیستم خط مشی  $Ua$  را در مورد اقدام بررسی کرده و بر اساس آن خط مشی اقدام<sup>-1</sup> برای هدف و همچنین سیاست مشخص شده در سیستم برای اقدام را به دست می آورد. ماژول تصمیم گیری، مشخصات مسیر، حلقه شمارنده و قوانین مبتنی بر مشخصه attpolicy را از این خط مشی ها استخراج می کند. این روش، الگوریتم بررسی مسیر را برای تعیین نتیجه هر خط مشی اجرا می کند. در طی بررسی مسیر، ماژول تصمیم گیری نیز باید تمام مشخصه های مربوطه را پیگیری کند و مطمئن شود که آن ها بر اساس خط مشی های مبتنی بر مشخصه ها برآورده می شوند. در نهایت، نتایج تمام خط مشی های انتخاب شده در ارزیابی، یک نتیجه واحد را تشکیل می دهند. وجود خط مشی های چند کاربره ممکن است منازعات خط مشی ها را افزایش دهد. برای حل این مسئله، می توانیم سیاست حل مناقشات مرجع [۷] را اتخاذ کنیم

<sup>۴۳</sup> Depth-First Search (DFS)

```

transition  $\sigma$  is not at accepting state
۲۷. break
۲۸. case  $\xi v \in \text{currentPath}$  and  $v = t$  and
transition  $\sigma$  is invalid for DFA
۲۹. break
۳۰. case  $\circ v \in \text{currentPath}$  and  $v \neq t$  and
transition  $\sigma$  is invalid for DFA
۳۱. break
۳۲. case  $\imath v \in \text{currentPath}$  and  $v \neq t$  and
transition  $\sigma$  is valid for DFA
۳۳.  $d \leftarrow d + 1$ ;  $\text{currentPath} \leftarrow \text{currentPath} \cdot (u, v, \sigma)$ 
۳۴.  $\text{currentState} \leftarrow \text{DFA takes transition } \sigma$ 
۳۵.  $\text{stateHistory} \leftarrow \text{stateHistory} \cdot (\text{currentState})$ 
۳۶. if (DFST(v)) then
۳۷. return TRUE
۳۸. else
۳۹.  $d \leftarrow d - 1$ ;  $\text{currentPath} \leftarrow \text{currentPath} \setminus (u, v, \sigma)$ 
۴۰.  $\text{attrList} \leftarrow \text{attrList} \setminus (\text{ATTR}(v), \text{ATTR}(u, v, \sigma))$ 
۴۱.  $\text{previousState} \leftarrow \text{last element in stateHistory}$ 
۴۲. DFA backs off the last taken transition  $\sigma$  to  $\text{previousState}$ 
۴۳.  $\text{stateHistory} \leftarrow \text{stateHistory} \setminus (\text{previousState})$ 
۴۴. return FALSE

```

مثال خط‌مشی P1 را در بخش قبلی در نظر بگیریم:

$\langle \text{profile access}, (u_a, ((f, 2): \exists[+1, -1], \text{occupation}(u) = \text{"student"}, \text{count} \geq 5)) \rangle$

گرامر قواعد زبانی<sup>۴۴</sup>، ابتدا گره شروع شونده  $u_a$  را استخراج می‌کند و قوانین مبتنی بر رابطه  $(f, 2)$  و قوانین مبتنی بر مشخصه  $\text{"student"}, \text{count} \geq 5$  را تجزیه می‌کند.

سپس الگوریتم<sup>۲</sup> یک DFA (ماشین‌های خودکار محدود قطعی<sup>۴۵</sup>) را از عبارات منظم  $f, f$  ایجاد می‌کند. این کار با استفاده از تابع (REtoDFA) انجام می‌شود. متغیرهای  $\text{currentPath}$  و  $\text{stateHistory}$  به ترتیب به  $NIL$  و حالت اولیه DFA مقداردهی می‌شوند. قواعد مبتنی بر مشخصه به سه بخش تقسیم می‌شود:

$\exists[+1, -1]$ ،  $\text{"occupation}(u) = \text{'student'}$  and  $\text{"count} \geq 5$ ".  $\exists$ ، کل مسیر را بین درخواست‌کننده دسترسی و

```

۳. Extract the quantifier symbol and interval/set
information from globalattpol
۴. Get the required rules for attributes of edges and
nodes f(ATTR(E),ATTR(N))
۵. Fetch the requirements of count attribute "count
 $\geq i$ ". If it is omitted, "count  $\geq 1$ ".
۶. Assign temporary space for attributes according
to the size of the interval/set and the hopcount
limit
۷. Initialize counter  $\text{count} \leftarrow \cdot$ 
۸. if  $\text{hopcount} \neq \cdot$  then
۹. return DFST(s)

```

الگوریتم<sup>۳</sup>. DFST(u)

```

۱. if  $d + 1 > \text{hopcount}$  then
۲. return FALSE
۳. else
۴. for all  $(v, \sigma)$  where  $(u, v, \sigma)$  in G do
۵. switch
۶. case  $\imath v \in \text{currentPath}$ 
۷. break
۸. case  $\imath v \in \text{currentPath}$  and  $v = t$  and DFA with
transition  $\sigma$  is at accepting state
۹. if  $v$  and  $(u, v, \sigma)$  is within the range specified
by quantifier then
۱۰.  $\text{attrList} \leftarrow \text{attrList} \cdot (\text{ATTR}(v), \text{ATTR}(u, v, \sigma))$ 
۱۱. if  $f(\text{ATTR}(v), \text{ATTR}(u, v, \sigma)) = \text{TRUE}$  then
۱۲.  $\text{count} \leftarrow \text{count} + 1$ 
۱۳. if  $\text{count} \geq i$  then
۱۴.  $d \leftarrow d + 1$ ;  $\text{currentPath} \leftarrow \text{currentPath} \cdot (u, v, \sigma)$ 
۱۵.  $\text{currentState} \leftarrow \text{DFA takes transition } \sigma$ 
۱۶.  $\text{stateHistory} \leftarrow \text{stateHistory} \cdot (\text{currentState})$ 
۱۷. return TRUE
۱۸. else
۱۹.  $\text{attrList} \leftarrow \text{attrList} \setminus (\text{ATTR}(v), \text{ATTR}(u, v, \sigma))$ 
۲۰. else
۲۱.  $d \leftarrow d + 1$ ;  $\text{currentPath} \leftarrow \text{currentPath} \cdot (u, v, \sigma)$ 
۲۲.  $\text{currentState} \leftarrow \text{DFA takes transition } \sigma$ 
۲۳.  $\text{stateHistory} \leftarrow \text{stateHistory} \cdot (\text{currentState})$ 
۲۴. return TRUE
۲۵. break
۲۶. case  $\imath v \in \text{currentPath}$  and  $v = t$  and
transition  $\sigma$  is valid for DFA but DFA with

```

<sup>45</sup> Deterministic Finite Automata (DFA)

<sup>44</sup> Grammar

هدف (یا صاحب منبع) که عملکرد تابع گره زیر را اعمال می کند، تعیین می کند.

occupation(u) = 'student' " یک تابع مشخصه های گره است که شغل کاربران را در مسیر بررسی می کند. پیش بین مشخصه تعداد "count attribute"  $\leq 5$  تعداد مشخصی از مسیرهای ارتباط و اجد شرایط را مشخص می کند. در این مثال برای ذخیره مقادیر مشخصه گره ها و یال ها در طول گذران، برای مشخصه های ۱ گره و ۲ یال نیاز به فضا داریم. به طور کلی، اگر فاصله  $[a, -b]$  باشد و محدودیت حلقه شمارنده برابر C باشد، باید فضایی را برای مشخصه های  $(1 + c - a - b)$  گره و  $(c - a - b)$  یال اختصاص دهیم.

پس از مقداردهی d (شاخص حلقه شمارنده) به ۰، الگوریتم ۲ تابع تراوایی<sup>۴۶</sup> (تابع گذار، تابع عبور) (DFS DFST) را که در الگوریتم ۳ نشان داده شده است از گره شروع شونده اجرا می کند. با توجه به گره u، الگوریتم ابتدا مطمئن می شود که یک گام روبه جلو، محدودیت حلقه شمارنده را نقض نمی کند. در غیر این صورت، باید خارج و به گره قبلی برگردد. اگر گذار بیشتری مجاز باشد، الگوریتم شروع به انتخاب یک یال  $(\sigma, v, u)$  از مجموعه ای از تمام یال های پیشامد می کند که از u یک به یک خارج می شوند. با توجه به الگوی مسیر  $ff$  در مثال، در مرحله اول، الگوریتم به طور خاص به دنبال یک یال مشاهده نشده نوع f می گردد که در گره دیگری به جز گره ارزیابی متوقف می شود (مورد ۶). اگر چنین لبه ای پیدا شود، بیابید بگوییم  $(u_a, u_1, f)$ ، الگوریتم مقدار d را یک واحد افزایش داده است، یال را به currentPath اضافه می کند، DFA را از حالت اولیه با حرکت گذار f حرکت می دهد و تاریخچه وضعیت DFA را بروز رسانی می کند. همچنین مشخصه های متناظر یال  $(f, u_a), u_1$  و گره  $u_1$  را به لیست مشخصه attrList برای ارزیابی بعدی اضافه می کند، زیرا  $u_1$  یک hop از  $u_a$  دور است و بنابراین در محدوده  $[+1, -1]$  قرار دارد. الگوریتم سپس به اجرای (DFST) در گره جدید  $u_1$  ادامه می دهد. از گره  $u_1$ ، فرآیند قبلی را با چک کردن حد حلقه شمارنده و انتخاب یال های پیشامد جدید مجدداً تکرار می کند. از آنجا که حد حلقه شمارنده برابر ۲ است، الگوریتم باید یک یال مشاهده نشده نوع f را پیدا کند که در (case t ۲) پایان می یابد. هنگامی که یال  $(u_1, t, f)$  کشف می شود، الگوریتم برای پیدا کردن

مشخصه متناظر برای ارزیابی ادامه می یابد.  $[1, -1]$  نشان می دهد که ما همچنین نیاز به بررسی مشخصه آخرین گره دوم در مسیر که  $u_1$  است، داریم. از آنجایی که ما قبلاً ویژگی های  $u_1$  را به لیست اضافه کردیم، الگوریتم به سادگی تابع ویژگی  $(u_1, ATTR)$  f را اجرا می کند تا ببیند آیا آن الزامات را برآورده می کند یا خیر. اگر الزامات برآورده شوند، بعد از آن مقدار مشخصه تعداد را که در این مورد count است، بررسی می کنیم. این خط مشی بیان می کند که به پنج مسیر مجاز نیاز دارد، بنابراین الگوریتم باید شمارنده را افزایش داده و به گره قبلی بازگردد تا ۴ مسیر دیگر جستجو کند. اگر  $(u_a, u_1, f)$  مسیر پنجم پیدا شده توسط ما باشد، DFST (u) الگوریتم ۱، باید مقدار درست و همچنین تمام فراخوانی های (DFST) قبلی خود را نیز بازگرداند. در نهایت، این کار سبب می شود الگوریتم ۲ مقدار درست را برگرداند (خروجی الگوریتم یک مقدار صحیح و درست باشد)، که نشان می دهد تعداد مسیرهای ضروری مناسب برای خط مشی را پیدا کرده ایم. اگر مشخصه گره/یال با الزامات مطابقت نداشته باشد، الگوریتم مشخصه ها را از لیست حذف می کند (خط ۱۸-۱۹) و یال بعدی را امتحان می کند.

پس از اتمام جستجوی یال در این سطح و بازگشت به فراخوانی قبلی (DFST) (خط ۳۸-۴۳)، باید لبه را رها کرده و تمام متغیرها را به مقادیر قبلی بازنشانی کند. الگوریتم ۲ پس از اینکه تمام یال های پیشامد سبب شوند  $u_a$  جستجوی ناموفقی داشته باشد مقدار نادرست<sup>۴۷</sup> را برمی گرداند.

اثبات صحت این الگوریتم اساساً همانند الگوریتم حلقه شمارنده [۷] است. الگوریتم جدید نه لبه های بیشتری را به حساب می آورد و هنوز بین  $O(dmin^{Hopcount})$  و  $O(dmax^{Hopcount})$  محدود می شود، جایی که  $dmin$  و  $dmax$  نشان دهنده حداقل و حداکثر درجه خروجی<sup>۴۸</sup> گره و حلقه شمارنده نشان دهنده محدودیت حلقه شمارنده است. بررسی ویژگی مشخصه هنگامی که الگوریتم یک مسیر واجد شرایط ممکن را پیدا می کند سربار<sup>۴۹</sup> اضافی را نشان می دهد. هزینه های سربار متناسب با مقدار مشخصه ها و نوع توابع مشخصه در نظر گرفته شده است که به ساختار نمودار اجتماعی مرتبط نیست.

<sup>۴۸</sup> Out-Degree

<sup>۴۹</sup> Overhead

<sup>۴۶</sup> traversal

<sup>۴۷</sup> False

## ۵- ارزیابی نتایج

به آن اضافه کنند، که در آن انواع رابطه به‌طور یکنواخت توزیع شده بود. باین‌حال، افزودن دستی اطلاعات نوع به مجموعه‌های داده‌های واقعی، ممکن است رفتارهای واقعی کاربر را نشان ندهد و در نتیجه تمامیت مجموعه‌های داده‌ها را خراب می‌کند و مقدار داده‌های واقعی را کاهش می‌دهد. علاوه بر این، مجموعه‌های داده‌های مختلف، خواص متفاوتی دارند که آن‌ها را غیرقابل مقایسه می‌سازد. از این‌رو، داده‌های ترکیبی به جایگزینی برای ما تبدیل می‌شود، که در آن می‌توانیم گراف‌های اجتماعی متفاوت را تحت کنترل خود بیآوریم و چند خاصیت ویژه‌ی این گراف‌ها را آنالیز می‌کنیم. برای تولید گراف‌های اجتماعی ترکیبی، چون آزمایش ما روی مقایسه‌ی شدت متمرکز است، درجه‌ی پیشروی هر گره را به تعدادی ثابت در هر گراف تنظیم می‌کنیم. انتخاب مقصد هر لبه به صورت تصادفی انجام می‌شود.

در مجموعه‌ی اول آزمایش‌ها، عملکرد الگوریتم‌های پیشنهادی و الگوریتم مقاله [۷] را با توجه به قواعد حد تعداد حلقه‌ی شمارنده متفاوت بررسی می‌کنیم. به‌طور خاص، پارامترها را به ۱۰۰۰ کاربر و یک نوع رابطه برای مجموعه‌ی آزمایش‌ها تنظیم می‌کنیم. هر کاربر دارای تعداد همسایه‌های یکسانی است، که به‌طور تصادفی در میان ۹۹۹ کاربر دیگر انتخاب می‌شوند. انواع مختلف الگوهای مسیر، شامل سرشماری (شمارش یک‌به‌یک) و \* -الگو، در قواعد به کار گرفته‌شده‌اند تا تأثیر حد تعداد حلقه‌ی شمارنده را روی عملکرد الگوریتم‌ها تحقیق کنند.

در مجموعه‌ی دوم آزمایش‌ها، قصد داریم تا عملکرد الگوریتم‌ها را در تعداد لبه‌های متفاوتی مطالعه کنیم که باید پیموده شود (یعنی درجه‌ی متوسط گره‌ها در گراف)، تا مقیاس‌پذیری رویکرد خود را نسبت به گراف تراکم نشان دهیم. مانند حالت قبل، ۱۰۰۰ کاربر را انتخاب می‌کنیم، ولی دو نوع از روابط را ایجاد می‌کنیم، به نام‌های «(friend)» و «(coworker)»، و به‌طور تصادفی هر رابطه را بین کاربران با یکی از این دو نوع اختصاص می‌دهیم. تعداد همسایه‌ها برای هر کاربر در مقادیر ۱۰۰، ۲۰۰، ۵۰۰ و ۱۰۰۰ در نظر گرفته‌شده است. تنها دو نوع از رابطه وجود دارد و گراف اجتماعی در واقعیت یک گراف پراکنده است، ۱۰۰۰ همسایه برای هر ۱۰۰۰ کاربر، ارزیابی گراف اجتماعی را مترکم می‌کند. سپس قواعد مختلف را روی این ۴ گراف اجرا می‌کنیم تا تفاوت‌های آن‌ها را با هم مقایسه کنیم.

در این بخش، نتایجی را نشان می‌دهیم که از مطالعات عملکردی ما روی الگوریتم پیشنهادی بررسی مسیر در مقایسه با الگوریتم مقاله [۷] به‌دست‌آمده است. الگوریتم‌ها را در متلب ۲۰۱۴ پیاده‌سازی کردیم و دو مجموعه از آزمایش‌ها را طراحی کردیم تا اجرای ارزیابی درخواست دسترسی را با استفاده از دو الگوریتم آزمایش کنیم. یک تصمیم‌گیرنده‌ی کنترل دسترسی را با چک‌کننده‌های مسیر الگوریتم پیشنهادی و الگوریتم مقاله [۷] روی یک ماشین مجازی با ۴ گیگابایت حافظه و پردازنده‌ی مرکزی ۴ هسته‌ای ۲.۵۳ GHz تعبیه کردیم. گراف‌های اجتماعی که باید آزمایش شوند، در پایگاه داده‌ی MySQL روی ماشین آزمایش به همراه قواعد کنترلی نمونه ذخیره شدند تا برای آزمایش آماده شوند. قواعد نمونه و درخواست‌های کنترلی را طراحی کردیم تا اطلاعات لازم گراف را برای تصمیم‌گیری‌های دسترسی گردآوری کنیم. سپس زمانی را ارزیابی کردیم که طول می‌کشد تا الگوریتم یک بررسی مسیر را روی گراف کامل کند و نتیجه‌ی آن را به تصمیم‌گیرنده بازگرداند.

## ۵-۱ مجموعه داده‌های تحقیق

به هنگام طراحی آزمایش‌ها، دو پارامتر گراف را در نظر می‌گیریم: تعداد حلقه‌ی شمارنده (عمق) و درجه (عرض). اگرچه تعداد کل گره‌ها در سیستم می‌تواند بر عملکرد و مقیاس‌پذیری بسیاری از سیستم‌های گراف تأثیر بگذارد، در سیستم ما، الگوریتم‌ها کل گراف را بررسی نمی‌کنند بلکه تنها مسیریابی با حلقه‌ی شمارنده‌ی محدود ناشی از یک گره را بررسی می‌کنند. بنابراین، تعداد کل گره‌ها با توجه به عملکرد چندان قابل توجه نیست. در واقع، حد تعداد حلقه‌ی شمارنده و تعداد لبه‌ها است که باید در هر حلقه‌ی شمارنده بررسی شود که به‌اندازه‌ی مسأله کمک می‌کند و از این‌رو عملکرد سیستم ما را تحت تأثیر قرار می‌دهد.

یک مسأله‌ی مهم در این ارزیابی شامل انتخاب مجموعه‌های داده‌های نمایشی است. چندین مجموعه‌ی داده‌ی عمومی وجود دارد که از سیستم‌های شبکه‌های اجتماعی آنلاین دنیای واقعی با داده‌های واقعی زیاد گردآوری شده‌اند. باین‌حال، اغلب آن‌ها تنها یک نوع رابطه را در نظر می‌گیرند یا از اصلاً از هیچ رابطه‌ای پشتیبانی نمی‌کنند. قبلاً در یک آنالیز مرتبط [۱۳]، نویسندگان مجموعه‌های داده‌های اصلی را اصلاح کرده بودند تا اطلاعات نوع را

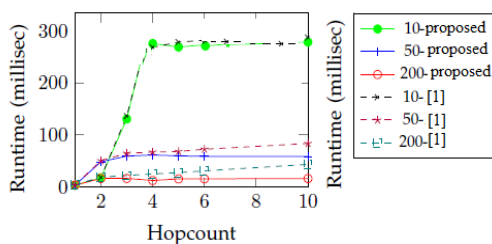
یافتن یک مسیر واجد شرایط انتخاب می کند، درحالی که الگوریتم پیشنهادی از سرریز در گراف های متراکم آسیب نمی بیند.

با توجه به ایده ی کلاسیک «۶ درجه جدایی» و یافته های «آزمایش های جهان کوچک» [۳]، هیچ جفت از افراد در فاصله ی بیشتر از ۶ اتصال متوسط قرار نگرفته اند. بکستروم و همکاران در [۱۵] نشان دادند که فاصله ی متوسط روی گراف اجتماعی کنونی فیسبوک، کمتر از ۶ درجه است و با رشد فیسبوک به ۴,۷۴ کاهش یافته است. بر اساس این یافته ها، برای موارد درست مسیره های شمارشی، حد حلقه شمارنده را برابر ۴ در نظر گرفتیم، زیرا مجموعه ی داده های ما نسبت به فیسبوک کمتر است. همان طور که در شکل ۳ (ب) نشان داده شده است، وقتی حد حلقه شمارنده افزایش یابد، هزینه ی زمانی با الگوریتم مقاله [۷] به طور قابل توجهی افزایش می یابد، زیرا بدون اینکه جستجو روی تمامی لبه ها را در سطح فعلی انجام ندهد، به حلقه شمارنده بعدی نمی رود؛ درحالی که انتخاب حلقه شمارنده بالاتر، عملکرد الگوریتم پیشنهادی را بدتر نمی کند. شکل ۳ (ج) مقایسه ی بین دو الگوریتم را در حالت های مورد غلط نشان می دهد. حالت های مورد غلط، در واقع بدترین حالت را برای بررسی مسیر نشان می دهد که در آن، هر دو الگوریتم باید به طور جامع، تمامی مسیره های محتمل را در حد حلقه شمارنده از گره آغازین جستجو کرد. بنابراین، دو الگوریتم به طور هم زمان در دو تنظیم شمارشی و \* -الگوی اجرا می شوند. با افزایش حلقه شمارنده، هزینه های زمانی الگوریتم افزایش می یابد.

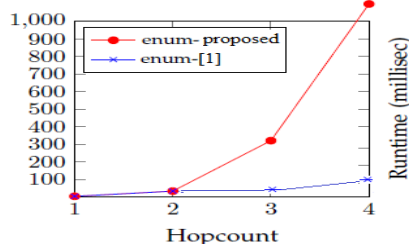
در یک قاعده ی کنترل دسترسی معین، به طور تصادفی ۱۰۰۰ جفت متفاوت از گره های هدف و درخواست دهنده را از گراف انتخاب می کنیم و هر الگوریتم را ۵ بار روی این ۱۰۰۰ جفت گره اجرا می کنیم. هر اندازه گیری، نتیجه ی متوسط این ۵۰۰۰ اجرا است. برای انجام مقایسه ی عادلانه بین موارد درست و غلط، قواعد مختلفی را طراحی کردیم تا به ۵۰۰۰ مورد درست و ۵۰۰۰ مورد غلط برسیم. برای مقایسه ی مساوی بین موارد درست تنظیمات مختلف، تعداد کاربران انتخاب شده را طوری مقیاس زدیم که بتوانیم از تعداد مشابه موارد درست به نتایج مورد نظر دست یابیم.

## ۵-۲ نتایج تجربی

شکل ۳ نتایج مجموعه ی اول آزمایش ها را نشان می دهد. الگوریتم های پیشنهادی و الگوریتم مقاله [۷] را با استفاده از قواعد با حدود حلقه شمارنده مختلف در هر دو حالت مورد غلط و مورد درست، با هم مقایسه کردیم. برای موارد درست مسیره های \* -الگو، شکل ۳ (الف) نشان می دهد که زمان اجرای متوسط چگونه با توجه به افزایش در حد حلقه شمارنده تغییر می کند. برای انجام یک قیاس گسترده تر، در این تست خاص، مقادیر ۱۰، ۵۰ و ۲۰۰ را که نزدیک به ۱۹۰ برابر تعداد دوست های متوسط بنا به ادعای فیسبوک است [۱۴] برای تعداد همسایه های هر کاربر تنظیم کردیم. مسیره های \* -الگو انعطاف پذیری بیشتری نسبت به مسیره های شماری در بررسی مسیر دارند. در واقع، نتایج \* -الگو، گذشت زمانی کوتاه ترین مسیر واجد شرایط را ثبت می کند. همان طور که انتظار داریم، با افزایش حلقه شمارنده، زمان اجرای مورد نیاز برای هر دو الگوریتم افزایش می یابد، ولی وقتی تعداد آن به ۴ برسد، روند آن هموار می شود. این نشان می دهد که یک مسیر واجد شرایط را می توان همواره بین دو کاربر در ۴ حلقه شمارنده در این تنظیمات یافت. یک محاسبه ی احتمال هم این یافته ها را تأیید می کند. در مورد ۱۰ همسایه برای هر کاربر، تجمع احتمال یافته های یک مسیر واجد شرایط، به ترتیب برای سه حلقه شمارنده اول برابر ۱، ۱۰/۵ و ۶۷/۳ درصد است، و نهایتاً برای حلقه شمارنده چهارم به ۱۰۰ درصد می رسد. احتمال در دو گراف متراکم دیگر، در سه حلقه شمارنده اول به ۱۰۰ درصد می رسد. همچنین یافتیم که الگوریتم پیشنهادی کمی بهتر از الگوریتم BFS برای حد بالای حلقه شمارنده در گراف های پراکنده عمل می کند، زیرا الگوریتم مقاله [۷] کاوش های طولی را پیش از



الف

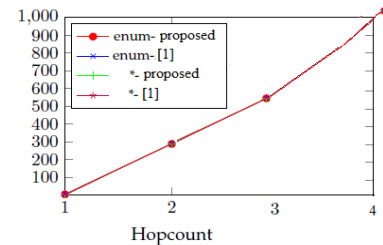


در شکل ۴ (الف و ب) نشان داده شده است، مشخص است که افزایش حد حلقه شمارنده، زمان یافتن یک مسیر واجد شرایط را برای هر دو الگوریتم افزایش می‌دهد، زیرا فضای جستجو افزایش می‌یابد. همچنین مشاهده می‌کنیم که وقتی با قواعد ۲ حلقه شمارنده سروکار داشته باشیم، هزینه‌های زمانی با افزایش در درجات گره کاهش می‌یابد. این اساساً به این دلیل است که احتمال بیشتری وجود دارد تا سریع‌تر یک مسیر واجد شرایط را بین دو گره در گراف متراکم نسبت به گراف پراکنده پیدا کرد، هرچند که بدترین زمان محتمل برای گراف متراکم بیشتر است. برای قواعد ۳ حلقه شمارنده، الگوریتم مقاله [۷] باید تمامی مسیرهای احتمالی را در ۲ حلقه شمارنده اول جستجو کند و بعد به جستجوی حلقه شمارنده سوم بپردازد، در نتیجه با افزایش درجه‌ی گره، زمان بیشتری طول می‌کشد تا یک انطباق را بیابد. در مقابل، الگوریتم پیشنهادی نسبت به هم‌متای الگوریتم مقاله [۷] خود در مواجهه با قواعد ۳ حلقه شمارنده و بیشتر، بهتر عمل می‌کند. مانند مجموعه آزمایش‌های اول، به نتایج مشابهی برای دو الگوریتم در حالت‌های مورد غلط دست یافتیم. در انطباق با آنالیز قبلی ما روی پیچیدگی، نتایج مشاهده شده از ۴ گراف اجتماعی متفاوت، افزایش زمان نسبت به درجه‌ی گره را نشان می‌دهد.

## ۶- نتیجه‌گیری و پیشنهادها

در این مقاله، یک مدل کنترل دسترسی رابطه محور کاربر به کاربر توسعه‌یافته برای شبکه‌های اجتماعی آنلاین ارائه داد که از خط‌مشی‌های مبتنی بر رابطه و مبتنی بر مشخصه برای تعیین دسترسی استفاده کرد. اطلاعات مشخصه کاربران و روابط آن‌ها به‌عنوان نمودار اجتماعی در شبکه‌های اجتماعی آنلاین با توجه به کنترل دسترسی مهم است. در این تحقیق، خط‌مشی‌های مبتنی بر مشخصه را فرمول‌بندی شد و دستور زبان را برای مشخصه‌های خط‌مشی توسعه دادیم. زبان خط‌مشی از بیان الزامات در خصوص برخی یا همه کاربران و روابط در مسیر پشتیبانی می‌کند. نتایج ما نشان می‌دهد که هر دو پارامتر درجه‌ی گره و حد حلقه شمارنده، به طور قابل توجهی بر عملکرد دو الگوریتم تأثیر می‌گذارد. در چند مورد، جستجوی یک مسیر واجد شرایط با طول ۳ حلقه شمارنده، مدت‌زمان زیادی طول می‌کشد که سیستم و کاربر نمی‌توانند آن را تحمل کنند. با این حال، گراف‌های اجتماعی در واقعیت بزرگ و پراکنده هستند. علاوه بر این، افراد اغلب تمایل دارند تا با سایر

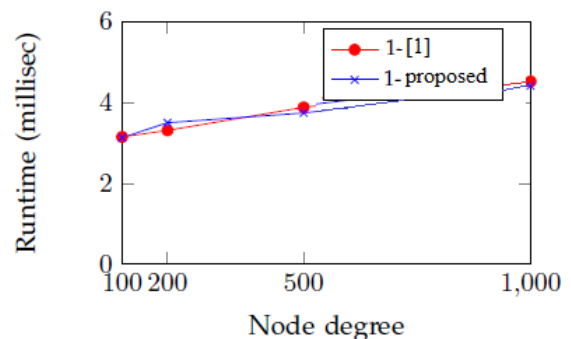
ب



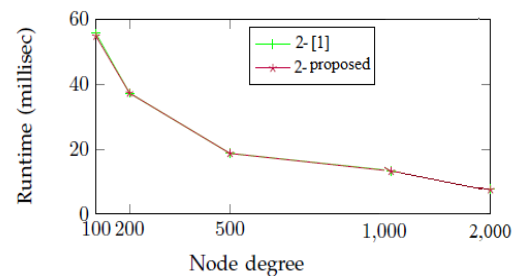
ج

شکل ۳. مقایسه حلقه شمارنده در الگوریتم پیشنهادی و

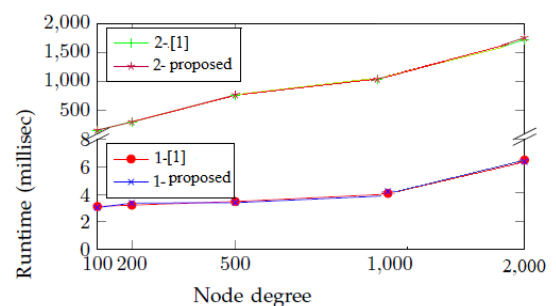
الگوریتم مقاله [۷]



الف) حلقه شمارنده ۱- سناریوی درست



ب) ۲ حلقه شمارنده- سناریوی درست



ج) سناریوی غلط

شکل ۴. مقایسه الگوریتم پیشنهادی و الگوریتم مقاله [۷] روی

درجه گره

شکل ۴ مقایسه‌ی عملکردی دو الگوریتم را روی گراف‌ها با درجات متفاوت گره نشان می‌دهد. در حالت‌های مورد درست، همان‌طور که



- [۴] B. Carminati, E. Ferrari, R. Heatherly, M. Kantarcioglu, and B. Thuraisingham. "A semantic web based framework for social network access control," In Proceedings of the 14th ACM SACMAT, pp. ۱۷۷-۱۸۶, ۲۰۰۹.
- [۵] B. Carminati, E. Ferrari, R. Heatherly, M. Kantarcioglu, and B. Thuraisingham. "Semantic web-based social network access control," *Computers and Security*, p. ۳۰, ۲۰۱۱.
- [۶] فاطمه رنجبر کاهریزی، حسین مومن زاده و حسن ارفعی - [۶] نیا، "مدل کنترل دسترسی در شبکه های اجتماعی"، نخستین کنفرانس تدبیر علوم کامپیوتر، مهندسی برق، ارتباطات و فناوری اطلاعات ایران در جهان اسلام، اردیبهشت ۱۳۹۸.
- [۷] Y. Cheng, J. Park, and R. Sandhu, "An Access Control Model for Online Social Networks Using User-to-User Relationships," *IEEE Transactions on Dependable and Secure Computing*, Vol. ۱۳, no. ۴, ۲۰۱۶.
- [۸] Mohammad Soryani, Behrooz Minaei, "Social Networks Research Aspects: A Vast and Fast Survey Focused on the Issue of Privacy in Social Network Sites," arXiv, ۲۰۱۲.
- [۹] X. Yang, H. Steck, Y. Liu, "Circle-based recommendation in online social networks," *KDD '12 Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. ۱۲۶۷-۱۲۷۵, ۲۰۱۲.
- [۱۰] A.J. Sarode, A. Mishra, "Audit and Analysis of Impostors: An experimental approach to detect fake profile in online

کاربران در یک فاصله ی نزدیک رابطه داشته باشند، بنابراین حلقه شمارنده بزرگ در عمل چندان مرسوم نیست. اگر کاربران قواعد را با محدودیت های کم و حد حلقه شمارنده کوچک تعیین کنند، الگوریتم ها قادر هستند تا یک نتیجه را در یک زمان کوتاه و منطقی بازگردانند. همچنین به سیستم پیشنهاد می دهیم تا یک تایم اضافی را برای هر جستار دسترسی اضافه کند تا از منتظر ماندن برای این حالت های شدید اجتناب شود. راه محتمل دیگر برای کاهش طول ها این است که به کاربران اجازه داده شود تا یک دید سفارشی و شخصی از گراف اجتماعی داشته باشند و مسیرهای کوتاه را ایجاد کنند تا به طور مکرر برای الگوهای رابطه به کار روند. مشاهده ی مهم دیگر از آزمایش های ما این است که اگرچه آن ها تقریباً دارای عملکرد یکسان برای قواعد ۱ و ۲ حلقه شمارنده هستند، الگوریتم پیشنهادی برای قواعد با مقادیر حلقه شمارنده متوسط (مانند ۳، ۴، ۵ و ...) نسبت به الگوریتم مقاله [۷] بهتر است. بنابراین، الگوریتم های بررسی مسیر را بر اساس روش پیشنهادی و روش مقاله [۷] ارائه کردیم و پیچیدگی الگوریتم ها را آنالیز نمودیم. امکان پذیری رویکرد خود را با بررسی اثبات اصل پیاده سازی دو الگوریتم نشان دادیم. در حالی که می توان مدل پیشنهادی را برای کنترل های مبتنی بر مشخصه های حتی دقیق تر توسعه داد، مدل پیشنهادی یک مکانیسم پایه ای جامع برای ReBAC فراهم می کند که اجازه کنترل دسترسی به مشخصه های متمایز را می دهد.

## مراجع

- [۱] D.M. Boyd, and N.B. Ellison, "Social network sites: Definition, history, and scholarship," *J. Computer-Mediated Communication*, Vol. ۱۳, no. ۱, pp. ۲۱۰-۲۳۰, ۲۰۰۷.
- [۲] H. Gao, J. Hu, T. Huang, J. Wang, and Y. Chen, "Security issues in online social networks," *Internet Computing*, IEEE, Vol. ۱۵, no. ۴, pp. ۵۶-۶۳, ۲۰۱۱.
- [۳] G. Bruns, P.W. Fong, I. Siahaan, and M. Huth, "Relationship based access control: its expression and enforcement through hybrid logic," In Proceedings of the second CODASPY, pp. ۱۱۷-۱۲۴, ۲۰۱۲.

- [۱۸] P. W. Fong. "Relationship-based access control: protection model and policy language," *In Proceedings of the first CODASPY*, pp. ۱۹۱-۲۰۲, ۲۰۱۱.
- [۱۹] P. W. Fong, M. Anwar, and Z. Zhao. "A privacy preservation model for facebook-style social network systems," *In Computer Security-ESORICS ۲۰۰۹*, pp. ۳۰۳-۳۲۰, ۲۰۰۹.
- [۲۰] P. W. Fong and I. Siahaan. "Relationship-based access control policies and their policy languages," *In Proceedings of the ۱۷<sup>th</sup> SACMAT*, pp. ۵۱-۶۰, ۲۰۱۱.
- [۲۱] S. R. Kruk, S. Grzonkowski, A. Gzella, T. Woroniecki, and H.-C. Choi, "D-FOAF: Distributed identity management with access rights delegation," *In The Semantic Web-ASWC*, pp. ۱۴۰-۱۵۴, ۲۰۰۶.
- [۲۲] M.K. Christy. Cheung, K.O. Matthew Lee, "A theoretical model of intentional social action in online social networks, ELSEVIER," *Decision Support Systems*, Vol. ۴۹, no. ۲, pp. ۲۴-۳۰, ۲۰۱۰.
- [۲۳] B. Carminati, E. Ferrari, and J. Girardi. "Performance analysis of relationship-based access control in osns," *In IEEE IRI*, pp. ۴۴۹-۴۵۶, ۲۰۱۲.
- [۲۴] S. Milgram, "The small world problem," *Psychology today*, Vol. ۲, no. ۱, pp. ۶۰-۶۷, ۱۹۶۷.
- social network," *ICCCT ۱۵*, pp. ۲۵-۲۷, ۲۰۱۵.
- [۱۱] Yong Li, Mengjiong Qian, Depeng Jin, Pan Hui, Athanasios V. Vasilakos, "Revealing the efficiency of information diffusion in online social networks of microblog," *Information Sciences*, no. ۲۹۳ (۲۰۱۵) ۳۸۳-۳۸۹.
- [۱۲] Ben Light and Elija Cassidy, "Strategies for the suspension and prevention of connection: Rendering disconnection as socioeconomic lubricant with Facebook," *New Media & Society*, Vol. ۱۶, no. ۷, pp. ۱۱۶۹-۱۱۸۴, ۲۰۱۴.
- [۱۳] Y. Cheng, J. Park, and R. Sandhu, "Relationship-Based Access Control for Online Social Networks: Beyond User-to-User Relationships, Paper presented at the SocialCom/PASSAT, ۲۰۱۲.
- [۱۴] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. "The anatomy of the facebook social graph," *CoRR*, ۲۰۱۱.
- [۱۵] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, "Four degrees of separation," *CoRR*, ۲۰۰۷.
- [۱۶] B. Carminati, E. Ferrari, and A. Perego. "A decentralized security framework for web-based social networks," *Int.J.o Info. Security and Privacy*, Vol. ۲, no. ۴, ۲۰۰۸.
- [۱۷] B. Carminati, E. Ferrari, and A. Perego. "Enforcing access control in web-based social networks," *ACM Trans. Inf. Syst. Secur*, Vol. ۱۳, no. ۱, ۲۰۰۹.

یک مدل کنترل دسترسی برای شبکه های اجتماعی آنلاین از طریق ارتباطات کاربر به کاربر

## ارائه مدلی برای بازیابی اطلاعات متنی با استفاده از اعداد فاصله‌ای

فرزاد قهرمانی\* هومان تحیری\*\*

\* دانشجوی دکتری بخش مهندسی و علوم کامپیوتر و فناوری اطلاعات، دانشکده مهندسی برق و کامپیوتر، دانشگاه شیراز

\*\* استادیار بخش مهندسی و علوم کامپیوتر و فناوری اطلاعات، دانشکده برق مهندسی و کامپیوتر، دانشگاه شیراز

تاریخ پذیرش: ۱۳۹۹/۱۰/۰۲

تاریخ دریافت: ۱۳۹۹/۰۶/۰۷

نوع مقاله: پژوهشی

### چکیده

با گسترش و توسعه وب و افزایش محتوای آنلاین، اهمیت سیستم‌های بازیابی اطلاعات که بتوانند با دقت بالاتری به نیازهای اطلاعاتی کاربران پاسخ دهند، بیشتر از پیش مشخص است. یک بخش مهم در طراحی هر سیستم بازیابی اطلاعات، انتخاب روشی مناسب برای مدل کردن آن سیستم است که در این راستا تعیین روش وزن‌دهی به لغات جهت بیان میزان اهمیت آنها در اسناد و پرس‌وجوها، نقش به‌سزایی دارد. روش‌های مختلفی در خصوص چگونگی وزن‌دهی به لغات ارائه شده که غالباً یک وزن عددی را تخصیص می‌دهند اما نمی‌توان با قطعیت گفت که بهترین روش وزن‌دهی کدام است. با توجه به ابهام و عدم قطعیتی که در این زمینه وجود دارد، در این مقاله مدلی ارائه شده که به جای استفاده از یک مقدار وزنی، با استفاده از وزن‌های بدست آمده از تعدادی روش وزن‌دهی پایه که به دقت انتخاب شده‌اند، برای هر لغت بازه‌ای از وزن‌ها را به عنوان یک وزن فاصله‌ای محاسبه می‌کند. در این مدل با انجام تجمیع مناسب، میزان ارتباط هر سند با پرس‌وجوی ورودی نیز به صورت یک وزن فاصله‌ای تعیین شده و برحسب آنها می‌توان با استفاده از یکی از سه روش پیشنهادی، اسناد را رتبه‌بندی کرد. در آزمایش‌های انجام شده بر روی مجموعه داده‌های معتبر Cranfield و Medline، اثرات نرمال‌سازی طول بردار وزن‌های پایه، استفاده از مؤلفه‌های مختلف در فاکتور فرکانس لغت و فاکتور فرکانس مجموعه مورد مطالعه و بحث قرار گرفته است و مشخص شد که انتخاب مجموعه‌ای مناسب از روش‌های وزن‌دهی پایه برای اعمال روش پیشنهادی، به همراه استفاده از روش رتبه‌بندی مناسب، تأثیر به‌سزایی در بهبود بازدهی سیستم خواهد داشت. با انتخاب‌های مناسب، برای دو مجموعه داده مذکور به ترتیب MAP با مقادیر ۰,۴۳۳۲۳ و ۰,۵۴۵۸۰ بدست آمد. این نتایج نشان داد که روش پیشنهادی نه تنها باعث بهبود نسبت به هر یک از روش‌های وزن‌دهی پایه می‌شود، بلکه در مقایسه با چند روش وزن‌دهی پیچیده اخیر نیز بهتر عمل می‌کند.

**واژگان کلیدی:** بازیابی اطلاعات متنی، رتبه‌بندی اسناد، وزن‌دهی لغات، اعداد فاصله‌ای، وزن فاصله‌ای

### ۱. مقدمه

گیرد. خصوصاً هدف یک موتور جستجوی وب، بازیابی صفحاتی از وب است که به نیازهای اطلاعاتی کاربر که برحسب تعدادی کلمه

توسعه روزافزون وب به همراه قابلیت دسترسی به حجم زیادی از محتوای آنلاین، باعث شده که تحقیقات زیادی در زمینه‌ی موتورهای جستجو<sup>۱</sup> یا سیستم‌های بازیابی اطلاعات (IRS<sup>۲</sup>) صورت

<sup>۱</sup> Search engines

<sup>۲</sup> Information Retrieval Systems

نویسنده مسئول: هومان تحیری tahayori@shirazu.ac.ir

با ناظر و بدون ناظر تقسیم می‌شود. روش‌های وزن‌دهی لغت با ناظر که از اطلاعات ضویت کلاس اسناد آموزشی در گروه‌های از قبل تعریف شده استفاده می‌کنند، عمدتاً از تکنیک‌های یادگیری ماشین استفاده می‌کنند [۱۱]. در این تحقیق روش‌های رتبه‌بندی و روش‌های وزن‌دهی آماری بدون یادگیری مدنظر می‌باشند.

در برخی از روش‌های بازیابی اطلاعات متنی، کاربرد به طریقی میزان اهمیت لغات مورد جستجو را وارد می‌کند که نهایتاً به عنوان یک وزن عددی محاسبه و برای لغات منظور می‌شود؛ در برخی دیگر کاربرد فقط پرس‌وجوی خود را در قالب یک متن یا چند لغت وارد کرده و سیستم IR به طریقی مشابه با وزنی که برای لغات اسناد محاسبه می‌کند، وزنی عددی نیز برای لغات پرس‌وجو محاسبه می‌نماید. اگر به وزن‌های تخصیص داده شده به لغات پرس‌وجوها و اسناد طبق روش‌های مختلف وزن‌دهی توجه شود، مشاهده می‌شود که این وزن‌ها برای یک لغت مقادیر مختلفی را در بر گرفته که ممکن است نزدیک به هم بوده یا از هم دور باشند، و سؤالی که پیش می‌آید این است که واقعاً وزن در نظر گرفته شده توسط کدام روش مناسب‌تر بوده و میزان اهمیت آن لغت را بهتر نشان می‌دهد؟ همچنین آیا کاربرد از اطلاعاتی/مقادیری که به عنوان وزن و میزان اهمیت وارد کرده مطمئن بوده است؟ با توجه به نداشتن پاسخی قطعی در خصوص سؤالات مذکور، در این مقاله به جای استفاده از یک مقدار وزنی، به طریقی ساده و به صورت آفلاین، برای هر لغت بازه‌ای از وزن‌ها به عنوان یک وزن فاصله‌ای<sup>۷</sup> در نظر گرفته می‌شود که بدون افزایش پیچیدگی می‌تواند منجر به افزایش دقت شود. در واقع استفاده از یک بازه یا فاصله به عنوان وزن لغات پرس‌وجو و اسناد، عدم قطعیت و عدم اطمینان موجود در انتساب وزن عددی به لغات را تعدیل ساخته و می‌تواند به محاسبه میزان ارتباط دقیق‌تر بین اسناد و پرس‌وجوها کمک کند.

روش پیشنهادی در این مقاله برای تخصیص دادن وزن فاصله‌ای به لغات، استفاده از وزن‌های محاسبه شده توسط تعدادی روش وزن‌دهی مختلف می‌باشد. بدین منظور ابتدا تعدادی از روش‌های وزن‌دهی موجود را مورد مطالعه قرار داده و با در نظر گرفتن ویژگی‌های بکار رفته در آنها، زیرمجموعه‌ای از آنها به عنوان روش‌های وزن‌دهی پایه<sup>۸</sup> انتخاب می‌شوند؛ سپس بر اساس وزن‌های تعیین شده برای هر لغت توسط این روش‌های پایه و انجام پردازش‌های لازم، یک وزن فاصله‌ای برای آن لغت تعیین می‌شود. با توجه به لغات درون هر پرس‌وجو و با انجام تجمیع مناسب بر روی وزن‌های فاصله‌ای این لغات، میزان ارتباط هر سند با پرس‌وجوی

کلیدی در یک پرس‌وجو<sup>۱</sup> بیان شده، مرتبط باشند [۱]. در واقع هدف یک سیستم IR این است که از بین اطلاعات جمع‌آوری شده، اقلام مرتبط<sup>۲</sup> با پرس‌وجوی کاربر را استخراج کرده و اقلام ناخواسته و غیرمرتبط را فیلتر کرده و کنار بگذارد. در این راستا، سیستم IR با نمایش، ذخیره‌سازی و دسترسی به اسناد<sup>۳</sup> سر و کار دارد که می‌توانند شامل متون، تصاویر، انیمیشن، صوت، اقلام چندرسانه‌ای، صفحات وب، توثیت، وبلاگ و دیگر قطعات اطلاعاتی باشند [۲]. هرچند در این مقاله اسنادی که به صورت متنی و نوشتاری هستند، مورد توجه می‌باشند.

مساله اصلی در تمام سیستم‌های IR، تطبیق دادن اسناد با پرس‌وجو و پیش‌بینی این موضوع است که کاربرد کدام اسناد را مرتبط با خواست خود خواهد دانست [۳]. در نمایش اسناد و پرس‌وجوها غالباً با نگاشتی از لغات به آنها کار داریم، که طبق [۲] لغات یا کلمات کلیدی مجموعه را به ترتیب به "اسناد" و "پرس‌وجوها" نگاشت می‌کنند و این غالباً با نسبت دادن وزن به این لغات صورت می‌گیرد که بیانگر میزان اهمیت آنها در سند یا پرس‌وجوی مربوطه است.

در طول عمری که از بازیابی اطلاعات می‌گذرد، روش‌های مختلفی در خصوص چگونگی وزن‌دهی به لغات در اسناد و پرس‌وجوها ارائه شده است، از جمله [۴] تا [۹]. علاوه بر IR، وزن‌دهی به لغات در حوزه‌های دیگری مانند طبقه‌بندی متون<sup>۴</sup> [۱۰] تا [۱۳] و تحلیل احساسات<sup>۵</sup> [۱۴]، نیز کاربرد داشته و نقش مهمی بازی می‌کند. اهمیت وزن‌دهی مناسب، و تأثیر زیاد آن در بهبود نتایج نهایی در تمامی این مطالعات مشهود بوده است [۶]، [۷]، و [۱۲].

یادگیری رتبه‌بندی<sup>۶</sup> زمینه دیگری از تحقیقات است که یادگیری ماشین، بازیابی اطلاعات، و پردازش زبان طبیعی را با یکدیگر درمی‌آمیزد. یادگیری رتبه‌بندی در معنای عام و گسترده، اشاره به استفاده از هر روش یادگیری ماشین برای رتبه‌بندی دارد. در معنای خاص و محدود، یادگیری رتبه‌بندی به تکنیک‌های یادگیری ماشین برای ساخت مدل رتبه‌بندی، در ایجاد رتبه‌بندی و تجمیع رتبه‌بندی اشاره دارد [۱۵]. استفاده از یادگیری ماشین در IR و طبقه‌بندی متون و غیره، فقط محدود به رتبه‌بندی نبوده و در مرحله وزن‌دهی به لغات نیز کاربرد دارد. بطور کلی دو نوع اصلی وزن‌دهی به لغات وجود دارد. یکی وزن‌دهی لغت معنایی، که از معانی دسته‌ها و لغات موجود در پایگاه‌های دانشی مانند WordNet بهره می‌برد؛ دوم وزن‌دهی لغت آماری، که به چگونگی ظهور یک لغت در یک سند یا گروهی از اسناد از نقطه‌نظر آماری مرتبط است و خود به دو دسته

<sup>۵</sup> Sentiment Analysis

<sup>۶</sup> Learning to rank

<sup>۷</sup> Interval Weight

<sup>۸</sup> Basic Weighting Methods

<sup>۱</sup> Relevant

<sup>۲</sup> Documents

<sup>۳</sup> Text Classification

معرفی مدل‌های مر سوم IR، خصوصیات آنها و توسعه‌هایی که از جنبه مفاهیم مختلف صورت گرفته‌است، پرداخته می‌شود.

## ۱.۲ مدل بولی<sup>۶</sup>

مدل بولی از نخستین مدل‌های بازیابی اطلاعات است که حتی می‌توان روش کار نخستین دستگاه مکانیکی [۱۸]، و استفاده‌های اولیه از کامپیوتر برای بازیابی اطلاعات [۱۹] تا [۲۱]، را جزء آن دانست. بطور کلی در مدل بولی اسناد و پرس‌وجوها به صورت مجموعه‌ای از لغات شاخص دیده می‌شوند که به آن "کیسه لغات"<sup>۷</sup> گفته می‌شود. در جستجوی بولی وزن لغات در هر سند یا پرس‌وجو، ۰ یا ۱ است؛ یعنی در صورتی که در سند یا پرس‌وجو هیچ تطابقی از لغت وجود نداشته باشد، درجه تطابق ۰، و در صورتی که بطور کامل مطابقت داشته باشد، درجه تطابق ۱ در نظر گرفته می‌شود. مدل بولی در بازیابی اطلاعات مبتنی بر محاسبات منطقی بولی است. با یک OR بولی، یک سند برای برآورده کردن یک پرس‌وجو باید یکی از لغات پرس‌وجو را داشته باشد. در صورتیکه برای پرس‌وجوی AND بولی، سند باید شامل تمامی لغات پرس‌وجو باشد تا پرس‌وجو برآورده شده و ۱ را برگرداند. این مدل به دلیل سادگی و درک مستقیم آن، به طور گسترده‌ای در سیستم‌های اولیه پذیرفته و استفاده شد. مشکل این مدل آن است که تطابق‌ها رتبه‌بندی نمی‌شوند و مشخص نمی‌کند کدام سند بیشترین رخداد از لغت را دارد. برای OR بولی، هیچ ارجحیتی بین لغات پرس‌وجو وجود ندارد و همگی وزن یکسانی دارند. بنابراین اسناد با تطابق برابر در لیست برگشتی ممکن است حاوی لغات متفاوتی باشند [۲].

## ۲.۲ مدل احتمالی<sup>۸</sup>

نخستین بار لان<sup>۹</sup> در سال ۱۹۵۷ پیشنهاد نسبت دادن امتیازی به اسناد بر اساس روش احتمالاتی و اهمیت کلمات کلیدی در اسناد را داد بطوریکه این امتیاز نشان‌دهنده میزان ارتباط اسناد با پرس‌وجو باشد [۲۲]. او همچنین در سال ۱۹۵۸ پیشنهاد استفاده از فرکانس رویداد کلمه را به عنوان اندازه‌گیری مفیدی از ارزش کلمه، مطرح کرد که بعداً به عنوان وزن فرکانس لغت (TF<sup>۱۰</sup>) شناخته شد [۴]. مؤثر بودن این روش نسبت به جستجوی بولی در آزمایش‌های زیادی نشان داده شده است [۲۳]. با توسعه ایده‌های اولیه، در [۲۴] اصولی برای رتبه‌بندی احتمالاتی تعریف شد که مشخص می‌کرد چگونه اسناد بطور بهینه بر اساس معیارهای احتمالاتی و با توجه به

ورودی نیز به صورت یک وزن فاصله‌ای تعیین شده و برحسب آنها می‌توان با استفاده از یکی از سه روش پیشنهاد شده در مقاله، اسناد را رتبه‌بندی<sup>۱</sup> کرد. در اینجا ضمن اجرای روش پیشنهادی بر روی دو مجموعه داده معتبر، در آزمایش‌ها نشان داده شده که انتخاب زیرمجموعه‌ای مناسب از روش‌های وزندهی پایه برای اعمال روش پیشنهادی، به همراه استفاده از روش رتبه‌بندی مناسب، تأثیر به سزایی در بهبود بازدهی سیستم خواهد داشت؛ لذا در آزمایش‌های صورت گرفته بر روی ۲۴ روش وزندهی پایه، تأثیر سه دسته از معیارها در انتخاب زیرمجموعه‌های وزنی مورد مطالعه قرار گرفته و نهایتاً مناسب‌ترین ترکیب معیارها در انتخاب زیرمجموعه‌های وزنی، در کنار مناسب‌ترین روش رتبه‌بندی تعیین شده است.

در ادامه، در بخش ۲ مروری بر کارهای مرتبط و مبانی نظری خواهد شد. در بخش ۳ مقدمات لازم در خصوص اعداد فاصله‌ای، و در بخش ۴ روش وزندهی پیشنهادی تشریح شده است. آزمایش‌های انجام شده و نتایج بدست آمده در بخش ۵ مورد بحث قرار گرفته، و نهایتاً بخش ۶ مربوط به نتیجه‌گیری است.

## ۲. کارهای مرتبط و مبانی نظری

در وب، رتبه‌بندی اسناد به سه دسته کلی مبتنی بر محتوا<sup>۲</sup>، مبتنی بر پیوند<sup>۳</sup>، و مبتنی بر پیوند-محتوا<sup>۴</sup> تقسیم می‌شود [۱۱]. در رتبه‌بندی مبتنی بر محتوا، برای بازیابی صفحات وب مرتبط با یک پرس‌وجو از محتوای صفحات به عنوان ویژگی‌هایی برای رتبه‌بندی استفاده می‌شود. رتبه‌بندی مبتنی بر پیوند یا اتصال، بر روی اطلاعات ساختاری مانند تعداد پیوندهای اشاره شده به یک صفحه وب یا تعداد لینک‌های خروجی تمرکز دارد که نشان از محبوبیت صفحات دارند. از جمله روش‌های متعارف در این زمینه می‌توان به روش‌های HITS<sup>۵</sup> و Page Rank<sup>۱</sup> اشاره کرد؛ در [۱۶] الگوریتم-های مختلف در این زمینه مرور شده و مورد مقایسه قرار گرفته‌اند. روش رتبه‌بندی مبتنی بر پیوند-محتوا، برای یافتن تعادل مناسب بین دو مورد قبل مطرح شده است که مقاله [۱۷] نمونه‌ای از این روش است که با استفاده از ویژگی‌های نسبتاً ساده اسناد وب، رتبه‌بندی مناسبی را با استفاده از معماری برنامه‌نویسی ژنتیکی ارائه می‌دهد. تحقیق پیش رو جزء دسته مبتنی بر محتوا بوده و بر روی روش‌های وزندهی لغات در بازیابی اطلاعات متنی تمرکز دارد؛ به همین دلیل، در این بخش با تکیه بر روش‌های وزندهی، به

<sup>۶</sup> Boolean model

<sup>۷</sup> Bag of terms

<sup>۸</sup> Probability model

<sup>۹</sup> Luhn

<sup>۱۰</sup> Term Frequency

<sup>۱</sup> Ranked

<sup>۲</sup> Content-based ranking

<sup>۳</sup> Hyperlink-based ranking

<sup>۴</sup> Hyperlink-content-based ranking

<sup>۵</sup> Hyperlink Induced Topic Search

زاویه بین بردارهای سند و پرس‌وجو را به عنوان معیار شباهت میان آنها مطرح کرد که یکی از معیارهای مهم در این زمینه است [۲۵]. اسپارک جونز در سال ۱۹۷۲، ایده استفاده از فرکانس سند معکوس (IDF<sup>۹</sup>) را مطرح کرد که به لغاتی که در تعداد اسناد کمتری مشاهده می‌شدند وزن بیشتری می‌داد [۳۵]. به دنبال آن ایده ترکیب دو وزن TF و IDF (یعنی TF-IDF<sup>۱۰</sup>) خیلی سریع مطرح شد که از جمله می‌توان به [۲۶] اشاره کرد که حاصل ضرب آن دو را به عنوان وزن لغات در نظر گرفت. سالتون و همکاران در سال ۱۹۷۵ یک مدل فضای برداری را برای شاخص‌بندی خودکار ارائه دادند [۳۷].

با توجه به اهمیت انتخاب وزن مناسب برای لغات در بازیابی اطلاعات از اسناد متنی با روش فضای برداری، سالتون و باکلی در [۶] به اهمیت سه فاکتور در وزن‌دهی لغات اشاره کردند که از ضرب آنها در یکدیگر می‌توان وزن هر لغت را تعیین کرد، این فاکتورها عبارتند از: فرکانس لغت (TF)، فرکانس سند معکوس (IDF) (یا فرکانس مجموعه معکوس<sup>۱۱</sup>)، و فاکتور نرمال‌سازی طول بردار<sup>۱۲</sup>. آنها برای هر یک از فاکتورها، مؤلفه‌های وزن‌دهی نشان‌داده شده در جدول ۱ را در نظر گرفتند تا با انتخاب یکی از مؤلفه‌ها برای هر فاکتور، و ضرب آنها در یکدیگر بتوان به یک روش وزن‌دهی لغت رسید. بدین ترتیب امکان انتخاب ترکیبات متنوعی برای وزن‌دهی لغات اسناد یا پرس‌وجوها وجود داشت که برحسب مؤلفه‌های بکار رفته برای فاکتورها در هر روش، آن روش با یک نام سه‌جزئی<sup>۱۳</sup> معرفی می‌گردد. با توجه به اینکه روش‌های وزن‌دهی در نظر گرفته شده برای لغات اسناد و لغات پرس‌وجو می‌تواند متفاوت باشد، آنها در آزمایش‌های خود، ترکیبات وزنی مختلف را توسط دو سه‌جزئی نشان دادند که این دو به ترتیب روش‌های در نظر گرفته شده برای وزن‌دهی لغات اسناد و وزن‌دهی لغات پرس‌وجو را نشان می‌دادند (مثلاً ترکیب وزنی  $tfc.nfx$  به معنی در نظر گرفتن وزن  $\frac{tf \cdot \log \frac{N}{n}}{\sqrt{\sum_{vector} (tf \cdot \log \frac{N}{n})^2}}$  برای لغات اسناد و وزن  $\log \frac{N}{n} \cdot \left(0.5 + 0.5 \frac{tf}{\max tf}\right)$  برای لغات پرس‌وجو می‌باشد). آنها در آزمایش‌های زیادی که بر روی چند مجموعه داده انجام دادند، به مقایسه نتایج استفاده از ترکیبات مختلف پرداخته و ترکیباتی را به عنوان روش‌های برتر وزن‌دهی لغات معرفی کردند.

معیارهای ارزیابی تعریف شده، رتبه‌بندی می‌شوند. از آن زمان تحقیقات زیادی انجام شده و مانند مدل‌های دیگر، نسخه‌های متعددی از این مدل نیز ارائه شده است که بخشی از آنها در [۲۵] آمده است.

نمایش‌های دیگری از مدل احتمالاتی وجود دارد که شامل مدل‌های  $BM_1, BM_{11}, BM_{15}, BM_{25}$  هستند. این مدل‌ها شبیه هم و طرفدار بهترین تطابق<sup>۱</sup> هستند که  $BM_{25}$  معمولترین آنها بوده و به دلیل نتایج خوبی که داشته، در تحقیقات زیادی به عنوان مبنای مقایسه<sup>۲</sup> استفاده شده است [۷] و [۲۶] تا [۲۸]. از این روش که با عنوان Okapi-BM<sub>25</sub> نیز نام برده می‌شود، انواع مختلفی ارائه شده است که در [۲۹] با انجام یک مطالعه تکرارپذیری<sup>۳</sup> بر روی هشت نوع مختلف آن نشان داده است که تفاوت معنی‌داری بین آنها وجود ندارد.

مدل‌سازی زبانی سند<sup>۴</sup> یک رویکرد مهم دیگر در روش‌های آماری برای بازیابی اطلاعات است [۳۰] و [۳۱]. این نوع روش‌ها ذاتاً مولد<sup>۵</sup> هستند، تا آنجا که هدف آنها تخمین مدلی از فرایند تولید سند است. در بازیابی اطلاعات، زمانی که مدل‌های سند تخمین زده شدند، اسناد می‌توانند بر اساس احتمال مدل سندشان در تولید پرس‌وجو، رتبه‌بندی شوند [۳۲]. مدل‌های زبانی سندی که از طریق یک فرآیند Pólya چندمتغیره، یک ویژگی خودتقویت‌کننده<sup>۶</sup> را ارائه می‌دهند، به طور قابل توجهی اثربخشی بازیابی را افزایش می‌دهند [۳۳]. کامینز [۳۲] این روش را بیشتر بکار گرفته و فرآیند کلی‌تری برای مدل‌سازی سند آماری ارائه داده است. او با ارائه روش‌های جدیدی در تخمین پارامترهای فرآیند Pólya، مدل‌های  $GSPUD_{mc}$  و  $GSPUD_{bst}$  را معرفی کرد که در آنها کارایی بازیابی در مقایسه با مدل‌های قبلی افزایش یافت.

## ۳.۲ مدل فضای برداری<sup>۷</sup>

مدل معمول دیگر در IR مدل فضای برداری است. نخستین بار سوئیتزر<sup>۸</sup> [۵] بود که به اسناد و پرس‌وجوها به صورت بردارهایی  $n$  بعدی از لغات منحصربفرد مجموعه نگاه کرد. در اینجا به جای سیستم تطابق/عدم تطابق مدل بولی، از وزن لغات برای نشان دادن میزان اهمیت آنها در سند یا پرس‌وجوی مربوطه استفاده می‌شود. تابع تطابق، معیار فاصله بردار پرس‌وجو از بردار سند است، و بر اساس این میزان فاصله می‌توان یک لیست رتبه‌بندی شده در خروجی ارائه داد [۲]. سالتون در [۳۴] ایده استفاده از کسینوس

<sup>۸</sup> Switzer

<sup>۹</sup> Inverse Document Frequency

<sup>۱۰</sup> Term Frequency-Inverse Document Frequency

<sup>۱۱</sup> Inverse Collection Frequency

<sup>۱۲</sup> Vector Length Normalization

<sup>۱۳</sup> Triple

<sup>۱</sup> Best match

<sup>۲</sup> Baseline

<sup>۳</sup> Reproducibility

<sup>۴</sup> Document Language Modelling

<sup>۵</sup> Generative

<sup>۶</sup> Self-reinforcing

<sup>۷</sup> Vector space model

جدول ۱. مؤلفه‌های وزن‌دهی لغت [۶]

مؤلفه فرکانس لغت		
b	۱,۰	وزن دودویی که در صورت حضور لغت در یک بردار، برابر ۱ می‌باشد (فرکانس لغت نادیده گرفته می‌شود)
t	tf	فرکانس لغت خام (تعداد دفعاتی که یک لغت در یک سند یا پرس‌وجو رخ می‌دهد)
n	$0.5 + 0.5 \frac{tf}{\max tf}$	فرکانس لغت نرمال شده افزایش یافته <sup>۱</sup> (فاکتور tf، توسط بیشترین مقدار tf در بردار نرمال شده، و مجدداً برای قرار گرفتن بین ۰,۵ و ۱ نرمال شده است)
مؤلفه فرکانس مجموعه		
x	۱,۰	بدون تغییر در وزن؛ استفاده از مؤلفه فرکانس لغت اصلی (b, t, یا n)
f	$\log \frac{N}{n}$	ضرب مؤلفه tf اصلی، در یک مؤلفه فرکانس مجموعه معکوس (N تعداد کل اسناد در مجموعه، و n تعداد اسنادی است که یک لغت در آنها مشاهده شده است)
p	$\log \frac{N - n}{n}$	ضرب مؤلفه tf، در یک مؤلفه فرکانس مجموعه معکوس احتمالی
مؤلفه نرمال‌سازی طول بردار		
x	۱,۰	بدون تغییر؛ استفاده از فاکتورهای بدست آمده از فرکانس لغت و فرکانس مجموعه به تنهایی (بدون نرمال‌سازی)
c	$\frac{1}{\sqrt{\sum_{vector} w_i^2}}$	استفاده از نرمال‌سازی کسینوسی؛ تقسیم هر وزن لغت w بر فاکتوری که نشان‌دهنده طول بردار اقلیدسی است

## ۴.۲ مدل‌های بازیابی فازی

مدل‌های بازیابی فازی یا مدل‌های بولی توسعه یافته که بر پایه مجموعه‌های فازی تعریف شده‌اند، به منظور رسیدگی به ابهام و عدم دقت کاربران (که در پرس‌وجوهای آنها انعکاس می‌یابد) و ارتباط جزئی بین اسناد و پرس‌وجوها پیشنهاد شدند [۳]. این مدل‌های توسعه یافته، وزن‌های لغت و توابع عضویت یک مدل مجموعه فازی را ترکیب کرده و تطابق‌هایی که به دقت یک مدل بولی هستند را برمی‌گردانند. از جمله این مدل‌ها می‌توان به Paice، MMM<sup>۵</sup>، و P-Norm اشاره کرد [۴۳]. توسعه‌های فازی سیستم‌های IR و تعمیم‌های فازی مدل بازیابی بولی در [۴۴] و [۴۵] نیز مرور شده است.

## ۳. مقدمات

در این بخش برخی از مفاهیم اولیه که در مقاله مورد استفاده قرار گرفته است تعریف خواهد شد.

ابراهیم و لاند-سیلوا [۸] برای بهبود وزن‌دهی لغات، از روش وزن‌دهی فرکانس لغت همراه با میانگین رویداد لغت (TF-ATO<sup>۲</sup>) استفاده کردند که در آن وزن هر لغت در هر سند، از تقسیم TF آن لغت در سند بر میانگین رویداد لغات موجود در آن سند بدست می‌آید. آنها همچنین از یک روش متمایزکننده (DA<sup>۳</sup>) بر پایه بردار مرکز ثقل اسناد، به منظور حذف وزن‌های کم ارزش از اسناد، استفاده کردند و در آزمایش‌های خود تأثیر مثبت هر یک از این دو ایده جدید وزن‌دهی را در دو حالت با حذف کلمات توقف<sup>۴</sup> و بدون حذف کلمات توقف نشان دادند.

تحقیقات زیاد دیگری نیز در راستای بهبود روش TF-IDF انجام شده است که با در نظر گرفتن فاکتورها و مؤلفه‌هایی دیگر سعی بر این مهم داشته‌اند، از جمله این تحقیقات می‌توان به [۱۰]، [۱۳]، و [۳۸] تا [۴۲] اشاره کرد.

<sup>۱</sup> Augmented

<sup>۲</sup> Term Frequency with Average Term Occurrence

<sup>۳</sup> Discriminative Approach

<sup>۴</sup> Stop-Words

<sup>۵</sup> Min, Max, and Mixed



### ۱.۳ اعداد فاصله‌ای<sup>۱</sup>

عدد فاصله‌ای  $A$  زیرمجموعه‌ای از اعداد حقیقی است که یک محدوده پیوسته و متصل به هم را در بر می‌گیرد.

$$A = \{x \in \mathbf{R}, l_a \leq x \leq r_a\} = [l_a, r_a] \quad (1)$$

$r_a$  و  $l_a$  به ترتیب حدود چپ (پایین) و راست (بالا) این عدد هستند و  $l_a \leq r_a$ ؛ همانطور که مشاهده می‌شود تمامی مقادیر بین این دو کرانه متعلق به  $A$  می‌باشند.

بطور کلی اگر  $A$  و  $B$  دو عدد فاصله‌ای با مقادیر  $A = [l_a, r_a]$  و  $B = [l_b, r_b]$  باشند، برخی از مفاهیم لازم به صورت زیر تعریف می‌شوند [۴۶]:

- پهنا یا وسعت یک عدد فاصله‌ای:

$$width(A) = r_a - l_a \quad (2)$$

- نقطه میانی یا مرکز یک عدد فاصله‌ای:

$$midpoint(A) = \frac{l_a + r_a}{2} \quad (3)$$

- جمع دو عدد فاصله‌ای:

$$A + B = [l_a + l_b, r_a + r_b] \quad (4)$$

- تفریق دو عدد فاصله‌ای:

$$A - B = [l_a - r_b, r_a - l_b] \quad (5)$$

- ضرب دو عدد فاصله‌ای:

$$A \times B = [\min(l_a \times l_b, l_a \times r_b, r_a \times l_b, r_a \times r_b), \max(l_a \times l_b, l_a \times r_b, r_a \times l_b, r_a \times r_b)] \quad (6)$$

- تقسیم دو عدد فاصله‌ای:

$$\frac{A}{B} = A \times \frac{1}{B}, \quad \cdot \notin B \quad \text{و} \quad \frac{1}{B} = \left[ \frac{1}{r_b}, \frac{1}{l_b} \right] \quad (7)$$

توجه شود که بجز پهنا و نقطه میانی یک عدد فاصله‌ای که حاصل آنها یک عدد حقیقی است، در بقیه موارد حاصل هر یک از عملیات نیز یک عدد فاصله‌ای خواهد بود.

### ۲.۳ میانگین وزنی فاصله‌ای (IWA)<sup>۲</sup>

این نوع میانگین‌گیری شبیه به میانگین‌گیری وزنی معمولی است که در آن برای داده‌ها و وزن آنها، به جای اعداد حقیقی از اعداد فاصله‌ای استفاده می‌شود و نتیجه میانگین وزنی نهایی آن نیز نه یک عدد، بلکه یک بازه عددی (عدد فاصله‌ای) خواهد بود. فرض کنید مجموعه‌های  $X$  و  $W$  هر یک شامل  $N$  عدد فاصله‌ای باشند

که به ترتیب داده‌ها و وزن‌های متناظر هر یک را در بر می‌گیرند [۴۷]:

$$x_i = [l_{x_i}, r_{x_i}], \quad x_i \in X \quad i = 1, 2, \dots, N \quad (8)$$

$$w_i = [l_{w_i}, r_{w_i}], \quad w_i \in W \quad i = 1, 2, \dots, N$$

در اینصورت میانگین وزنی فاصله‌ای  $X_i$  های مختلف، عدد فاصله‌ای  $Y$  بوده و به صورت زیر محاسبه می‌شود:

$$Y = \frac{\sum_{i=1}^N x_i \times w_i}{\sum_{i=1}^N w_i} = [l_y, r_y] \quad (9)$$

در رابطه (۹) نیاز به انجام عملیات ضرب، جمع، و تقسیم بین اعداد فاصله‌ای می‌باشد که جهت انجام آنها از روابط (۳) تا (۷) استفاده می‌شود.

### ۴. روش پیشنهادی وزن‌دهی

در این مقاله که بر پایه مدل فضای برداری عمل می‌شود، روش پیشنهادی کلی این است که بر اساس وزن‌های مختلفی که از چند روش مختلف وزن‌دهی برای یک لغت بدست آمده‌اند، یک وزن فاصله‌ای (بازه‌ای) برای آن لغت محاسبه شود تا بدین ترتیب هر یک از وزن‌های اولیه بتوانند به نحوی استفاده شده و اثرگذار باشند. در اینجا ماژول تطابق نیز در محاسبات خود بر اساس اصول اعداد فاصله‌ای عمل کرده و نهایتاً میزان ارتباط اسناد مختلف با هر پرس‌وجو را در قالب یک سری اعداد فاصله‌ای ارائه می‌دهد که بر مبنای آنها اسناد رتبه‌بندی می‌شوند. مراحل روش پیشنهادی و جزئیات بیشتر در ادامه شرح داده شده است.

### ۱.۴ پیش‌پردازش داده‌ها

جهت پالایش اولیه مستندات، از معمول‌ترین عملیات در مرحله پیش‌پردازش استفاده می‌گردد: حذف کلمات توقف<sup>۳</sup> (کلماتی که بی‌معنی بوده و تاثیری در برآورد میزان ارتباط اسناد و پرس‌وجوها ندارند)، انجام ریشه‌یابی<sup>۴</sup> (یکسان گرفتن کلماتی که ریشه مشترکی دارند)، و نادیده گرفتن بزرگ و کوچک بودن حروف<sup>۵</sup>.

### ۲.۴ انتخاب روش‌های وزن‌دهی پایه<sup>۶</sup>

همانطور که اشاره شد روش‌های وزن‌دهی مختلفی در تحقیقات پیشین معرفی شده‌اند. در این مرحله بایستی تعدادی از این روش‌ها را به عنوان روش‌های وزن‌دهی پایه انتخاب کرد تا از آنها جهت ساخت وزن‌های فاصله‌ای لغات فرهنگ بدست‌آمده، استفاده نمود. با

<sup>۴</sup> Stemming

<sup>۵</sup> Case sensitivity

<sup>۶</sup> Basic weighting methods

<sup>۱</sup> Interval Numbers

<sup>۲</sup> Interval Weighted Average

<sup>۳</sup> Stop words

که در آن  $wq_{m,t}$  وزن مربوط به لغت  $t$  در آن پرس‌وجو، طبق روش وزندهی پایه  $m$  می‌باشد.

#### ۴.۴ نرمال‌سازی وزن‌های پایه

وزن‌های مختلفی که در مرحله قبل برای هر لغت بدست آمد، ممکن است بسته به روش‌های وزندهی انتخابی، مقادیر مختلفی را در بر بگیرند. مثلاً ممکن است وزن‌های بدست آمده برای لغات طی یک روش، تماماً در بازه  $[0,1]$ ، یا بازه  $[0,5,1]$  قرار گرفته، و یا مقادیری بیش از ۱، یا حتی منفی را نیز در بر گیرند. بنابراین در این مرحله باید تمامی وزن‌های پایه مربوط به لغات مختلف سند  $d$  را که در روش وزندهی  $m$  بدست آمده‌اند، در بازه  $[0,1]$  نرمال‌سازی کرد تا بدین ترتیب اثر پراکندگی وزن‌های روش‌های مختلف برای هر لغت تعدیل گردد. نرمال‌سازی می‌تواند بر اساس رابطه زیر انجام شود [۴۸]:

$$nwd_{d,m,t} = \frac{wd_{d,m,t} - wd_{d,m}^{min}}{wd_{d,m}^{max} - wd_{d,m}^{min}} \quad t \in \{1, 2, \dots, n_t\} \quad (12)$$

جائیکه  $wd_{d,m}^{max}$  و  $wd_{d,m}^{min}$  به ترتیب کمترین و بیشترین مقادیر وزنی پایه مجموعه لغات در سند  $d$ ، طی روش وزندهی  $m$ ، و  $nwd_{d,m,t}$  وزن پایه نرمال‌شده مربوط به لغت  $t$  در سند  $d$  طبق روش وزندهی پایه  $m$  می‌باشند. به طریق مشابه بایستی وزن‌های پایه بدست آمده برای لغات پرس‌وجو را نیز نرمال‌سازی کرد تا  $nwd_{d,m,t}$ ‌های مختلف برای بردار پرس‌وجوی  $Q_m$  بدست آید. توجه شود که انجام نرمال‌سازی وزن‌ها، به معنی ۱ شدن طول بردار وزن‌ها نمی‌باشد.

#### ۵.۴ محاسبه وزن فاصله‌ای

در این مرحله بایستی به طریقی مناسب با استفاده از وزن‌های پایه نرمال‌شده مختلفی که از روش‌های مختلف وزندهی برای هر لغت در هر سند یا پرس‌وجو بدست آمده، یک وزن فاصله‌ای برای آن لغت بدست آورد. در این تحقیق این وزن‌های فاصله‌ای بر اساس انحراف استاندارد وزن‌های مختلف در پیرامون میانگین آنها تشکیل می‌شود و برای لغات اسناد مختلف طبق رابطه زیر بدست می‌آیند:

$$iwd_{d,t} = [lwd_{d,t}, rwd_{d,t}] \quad (13)$$

$$= [avg_{d,t} - std_{d,t}, avg_{d,t} + std_{d,t}]$$

جائیکه  $lwd_{d,t}$  و  $rwd_{d,t}$  به ترتیب حد چپ و راست بازه در نظر گرفته شده برای وزن فاصله‌ای، و  $avg_{d,t}$  و  $std_{d,t}$  به ترتیب میانگین و انحراف استاندارد وزن‌های پایه نرمال‌شده مختلفی هستند که برای لغت  $t$  در سند  $d$  بدست آمده بودند و به صورت زیر محاسبه می‌شوند:

توجه به اینکه روش‌های وزندهی مختلف مطرح شده در [۶] که در بخش قبل شرح داده شدند از تنوع زیادی برخوردار بوده و پایه خیلی از روش‌های وزندهی در تحقیقات بعدی بوده‌اند، در این تحقیق نیز از آنها به عنوان روش‌های وزندهی پایه استفاده شده است. البته در اینجا به مؤلفه‌های فاکتور فرکانس لغت در جدول ۱، مؤلفه دیگری تحت عنوان مؤلفه فرکانس لغت نرمال‌شده (معمولی، نه افزایش‌یافته) با فرمول  $\frac{tf}{\max tf}$  (و با نماد  $Z$ ) اضافه شده تا توسط آن بتوان مقدار فاکتور TF لغات هر سند را با توجه به ماکزیمم TF لغات آن سند بین ۰ و ۱ محدود کرد. در این صورت امکان ایجاد ۲۴ ترکیب متنوع به عنوان روش‌های وزندهی پایه برای لغات اسناد یا پرس‌وجوها بوجود خواهد آمد که در جدول ۳ قابل مشاهده هستند.

واضح است که انتخاب روش‌های وزندهی اولیه مناسب، از اهمیت بالایی در بهبود نتایج نهایی برخوردار بوده و ممکن است در نظر گرفتن تمامی این ۲۴ روش وزندهی، جهت انجام بقیه مراحل روی آنها لازم نباشد و انتخاب زیرمجموعه‌ای مناسب از آنها کفایت کند. در بخش ۵ آزمایش‌های مختلفی در این خصوص انجام شده و انتخاب زیرمجموعه‌های مختلفی مورد بحث و بررسی قرار گرفته است.

#### ۳.۴ محاسبه وزن‌های پایه

در این مرحله با استفاده از روش‌های مختلف وزندهی پایه که در مرحله قبل انتخاب شده‌اند، برای هر لغت تعدادی وزن پایه و اولیه تعیین می‌شود. بدین ترتیب اگر پس از انجام پیش‌پردازش‌ها، تعداد  $n_t$  لغت منحصر بفرد در فرهنگ لغات وجود داشته باشد، می‌توان هر سند  $d$  را در روش وزندهی  $m$  به صورت برداری از وزن لغات به فرم زیر نمایش داد:

$$D_{d,m} = \langle wd_{d,m,1}, wd_{d,m,2}, \dots, wd_{d,m,n_t} \rangle \quad (10)$$

$$d \in \{1, 2, \dots, n_d\} \quad m \in \{1, 2, \dots, n_m\}$$

که در آن  $n_d$  تعداد اسناد در مجموعه،  $n_m$  تعداد روش‌های مختلف در نظر گرفته شده جهت بدست آوردن وزن‌های پایه، و  $wd_{d,m,t}$  وزن پایه مربوط به لغت  $t$  در سند  $d$  طبق روش وزندهی پایه  $m$  می‌باشد. همچنین به طریق مشابه می‌توان هر پرس‌وجو را در روش وزندهی  $m$  به صورت برداری از وزن لغات به فرم زیر نمایش داد:

$$Q_m = \langle wq_{m,1}, wq_{m,2}, \dots, wq_{m,n_t} \rangle \quad (11)$$

$$m \in \{1, 2, \dots, n_m\}$$

$$Relevancy(D_d, Q) = IWA(D_d, Q) = \quad (16)$$

$$\frac{\sum_{t=1}^{n_t} (iwd_{d,t} \times iwq_t)}{\sum_{t=1}^{n_t} iwq_t} = [l_d, r_d] \quad d: 1, 2, \dots, n_d$$

$l_d$  و  $r_d$  به ترتیب حد چپ (پایین) و راست (بالا) عدد فاصله‌ای نتیجه شده می‌باشند. با اینکه ظاهراً سری عملیات صورت و مخرج رابطه (۱۶) باید بر روی کل لغات ( $n_t$ ) انجام شوند اما در عمل تنها کافی است که به تعداد لغات پرس‌وجو (بعد از مرحله پیش‌پردازش) و بر روی آنها انجام شوند، زیرا که وزن فاصله‌ای لغات غیرموجود در پرس‌وجو صفر بوده و در محاسبات بی‌تأثیر هستند. این موضوع سرعت محاسبات را بالا می‌برد.

در رابطه (۱۶) در واقع مقدار وضعیت بازیابی (RSV) سند  $D_d$  به صورت یک عدد فاصله‌ای محاسبه می‌شود؛ پس از محاسبه RSV اسناد مختلف، برای رتبه‌بندی آنها می‌توان به یکی از سه روش زیر اقدام کرد:

- رتبه‌بندی اسناد بر اساس مقدار چپ ( $Lrank^1$ ): در اینجا بایستی اسناد را بر اساس مقدار چپ (حد پایین) ارتباط‌های بدست آمده با پرس‌وجو ( $l_d$ ها) به صورت نزولی مرتب کرد تا رتبه آنها بدست آید.
- رتبه‌بندی اسناد بر اساس مقدار راست ( $Rrank^2$ ): در اینجا بایستی اسناد را بر اساس مقدار راست (حد بالای) ارتباط‌های بدست آمده با پرس‌وجو ( $r_d$ ها) به صورت نزولی مرتب کرد تا رتبه آنها بدست آید.
- رتبه‌بندی اسناد بر اساس مقدار نقطه میانی ( $Mrank^3$ ): در اینجا برای بدست آوردن رتبه اسناد، ابتدا نقطه میانی (مرکز) اعداد فاصله‌ای بدست آمده به عنوان ارتباط اسناد با پرس‌وجو را طبق رابطه (۲) بدست آورده و سپس این مقادیر را به صورت نزولی مرتب کرده تا رتبه اسناد بدست آید.

## ۵. نتایج و بحث

به دلیل محدودیت سخت‌افزاری در کار بر روی مجموعه داده‌های بزرگ، آزمایش‌ها بر روی دو مجموعه داده محک<sup>۴</sup> Cranfield و Medline انجام شد که موضوع آنها به ترتیب در زمینه علوم مربوط به گاز و هوا<sup>۵</sup> و پزشکی می‌باشد [۲۶]. اطلاعات آماری این مجموعه‌ها در جدول ۲ مشاهده می‌شود [۶].

$$avg_{d,t} = \frac{\sum_{m=1}^{n'_m} nwd_{d,m,t}}{n'_m}, \quad n'_m \leq n_m \quad (14)$$

$$std_{d,t} = \sqrt{\frac{\sum_{m=1}^{n'_m} (nwd_{d,m,t} - avg_{d,t})^2}{n'_m}} \quad (15)$$

همانطور که قبلاً اشاره شد ممکن است در برخی موارد از تمامی روش‌های وزن‌دهی پایه در نظر گرفته شده استفاده نشود و تنها زیرمجموعه‌ای از آنها در نظر گرفته شود؛ در همین خصوص،  $n'_m$  تعداد روش‌های وزن‌دهی پایه مورد استفاده می‌باشد. وزن‌های فاصله‌ای لغات پرس‌وجو ( $iWq_t$ ها) نیز به طریق مشابه محاسبه می‌شود.

## ۶.۴ رتبه‌بندی اسناد

پس از اینکه در مرحله قبل برای هر یک از لغات اسناد و پرس‌وجوها بازه‌ای از وزن‌ها به عنوان وزن فاصله‌ای تعیین شد، حال سوال این است که جهت رتبه‌بندی اسناد، میزان ارتباط بین پرس‌وجوها و اسناد چگونه باید مشخص شود؟ می‌بایست توجه شود که کاربر نیازهای اطلاعاتی خود را در قالب پرس‌وجوی ورودی از سیستم می‌خواهد و وزن لغات در آن پرس‌وجو میزان اهمیت هر لغت را تعیین می‌کند؛ لذا هرچه سندی حاوی لغات خواسته شده در پرس‌وجو بوده و خصوصاً لغات مهمتر پرس‌وجو، در آن سند وزن بیشتری داشته باشند، می‌توان گفت که آن سند به پرس‌وجوی خواسته شده مرتبط‌تر است. برای نشان دادن این میزان ارتباط می‌توان از رابطه (۱۶) استفاده کرد. در عمل رابطه (۱۶) مبین محاسبه میانگین وزنی بر روی لغات سند است که هدف آن برآورد یک وزن کلی برای سند در ارتباط با پرس‌وجوی ورودی است. به منظور تأثیر دادن اهمیت لغات پرس‌وجو، وزن این لغات نیز به عنوان ضریب وزنی در محاسبه دخالت داده شده است. در اینجا چون وزن‌های محاسبه شده برای لغات به صورت فاصله‌ای هستند، از روش محاسبه میانگین وزنی فاصله‌ای استفاده می‌گردد که فرمول کلی آن در رابطه (۹) آورده شد؛ کافی است برای محاسبه میزان ارتباط بین سند  $D_d$  و پرس‌وجوی  $Q$ ، در رابطه (۹) وزن‌های فاصله‌ای لغات سند ( $iwd_{d,t}$ ها) را به عنوان داده‌های اصلی، و وزن‌های فاصله‌ای لغات پرس‌وجو ( $iWq_t$ ها) را به عنوان وزن‌های داده‌های اصلی در نظر گرفت. بدین ترتیب داریم:

<sup>۴</sup> Benchmark  
<sup>۵</sup> Aerodynamics

<sup>۱</sup> Left rank  
<sup>۲</sup> Right rank  
<sup>۳</sup> Midpoint rank

جدول ۲. اطلاعات آماری مجموعه داده‌ها (شامل میانگین طول بردار و انحراف استاندارد طول بردار) [۶]

مجموعه داده	تعداد بردارها (اسناد یا پرس‌وجوها)	میانگین طول بردارها (تعداد لغات)	انحراف استاندارد طول بردارها	میانگین فرکانس لغات در بردارها
Cranfield documents	۱۳۹۸	۵۳،۱۳	۲۲،۵۳	۱،۵۸
queries	۲۲۵	۹،۱۷	۳،۱۹	۱،۰۴
Medline documents	۱۰۳۳	۵۱،۶۰	۲۲،۷۸	۱،۵۴
queries	۳۰	۱۰،۱۰	۶،۰۳	۱،۱۲

جائیکه  $n_q$  تعداد پرس‌وجو‌ها، و  $AP(q)$  میانگین دقت مربوط به پرس‌وجوی  $q$  است.

- معیار دقت در نقطه  $n$  ( $P@n$ ): این معیار، نسبت تعداد اسناد مرتبط را به کل اسناد در برترین  $n$  سند ارائه شده، مشخص می‌کند؛ یعنی [۱۶]:

$$P@n = \frac{\# \text{ of relevant documents in top } n \text{ results}}{n} \quad (20)$$

در اینجا معیارهای  $P@5$ ،  $P@10$ ، و  $P@20$  مورد ارزیابی قرار می‌گیرند.

در ادامه آزمایش‌های انجام‌شده و تحلیل‌های لازم شرح داده خواهد شد.

## ۱.۵ روش‌های وزن‌دهی پایه

همانطور که در بخش ۴-۲ اشاره شد، در این تحقیق ۲۴ روش وزن‌دهی پایه در نظر گرفته شد که نهایتاً بایستی زیرمجموعه‌هایی مناسب از آنها جهت اعمال روش‌های پیشنهادی بر روی آنها، انتخاب شوند. در اینجا ابتدا هر یک از این روش‌ها به تنهایی به عنوان روش وزن‌دهی لغات اسناد و پرس‌وجوها انتخاب، و وزن‌های بدست آمده از هر روش طبق رابطه (۱۲) در بازه  $[0,1]$  نرمال‌سازی شدند. سپس با استفاده از فرمول مرسوم ضرب بردارهای سند و پرس‌وجو در یکدیگر، شباهت و ارتباط اسناد با پرس‌وجوها محاسبه و رتبه‌بندی شدند. نتایج  $MAP$  هر دو مجموعه داده برای تمامی ۲۴ روش وزن‌دهی، در جداول ۳ و ۴ مشاهده می‌شود.

اگر به نتایج جدول ۳ دقت شود، نکاتی دیده می‌شود که توجه به آنها می‌تواند به انتخاب زیرمجموعه‌هایی مناسب جهت وزن‌های پایه

پس از انجام پیش‌پردازش‌های اشاره شده در بخش ۴-۱ بر روی هر یک از مجموعه داده‌ها، آزمایش‌های مختلفی در خصوص چگونگی انتخاب زیرمجموعه‌های وزنی پایه برای اسناد و پرس‌وجوها و انجام رتبه‌بندی بر اساس روش‌های پیشنهادی صورت گرفت. کارآیی روش وزن‌دهی پیشنهادی بر اساس ساس دآوری‌های اسناد مرتبط به هر پرس‌وجو (که به صورت دودویی انجام شده برای هر مجموعه داده موجود است)، و برحسب معیار میانگین متوسط دقت ( $MAP^1$ ) و دقت در نقطه  $n$  ( $P@n$ ) ارزیابی می‌شود.

- معیار میانگین متوسط دقت ( $MAP$ ): این معیار در قالب یک عدد، مؤثر بودن سیستم<sup>۲</sup> را نشان داده و به طور گسترده‌ای مورد استفاده قرار می‌گیرد [۴۹]. متوسط دقت ( $AP$ ) ترکیبی از دقت و بازخوانی<sup>۳</sup> است که، برای یک پرس‌وجو، بر روی لیست رتبه‌بندی شده اسناد، اعمال شده و به صورت زیر محاسبه می‌شود [۷]:

$$AP = \frac{1}{R} \sum_{r=1}^{n_d} P(d_r) \cdot Rel(d_r) \quad (17)$$

$$R = \sum_{r=1}^{n_d} Rel(d_r) \quad (18)$$

جائیکه  $d_r$  سند در رتبه  $r$ ،  $P(d_r)$  دقت در رتبه  $r$ ،  $R$  تعداد اسناد مرتبط به پرس‌وجو در مجموعه، و  $Rel(d_r)$  دآوری دودویی (۰ و ۱) انجام شده برای سند رتبه  $r$  از نظر داشتن ارتباط با پرس‌وجو است. پس از اینکه  $AP$  برای تمامی پرس‌وجوهای یک مجموعه داده بدست آمد، میانگین آنها به عنوان  $MAP$  در نظر گرفته خواهد شد، یعنی:

$$MAP = \frac{\sum_{q=1}^{n_q} AP(q)}{n_q} \quad (19)$$

<sup>۲</sup> Recall

<sup>۱</sup> Mean Average Precision

<sup>۳</sup> System effectiveness

جدول ۳. روش‌های وزن‌دهی پایه، به همراه مقدار MAP هر روش بر روی مجموعه داده‌های Cranfield و Medline

شماره روش	روش‌های ۱-۱۲ (بدون نرمال‌سازی طول بردار)												روش‌های ۱۳-۲۴ (همراه با نرمال‌سازی طول بردار)											
	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲	۲۳	۲۴
نام سه‌جزئی روش	bxx	bfx	bpx	lxx	lfx	lpx	nx	nfx	mpx	zxx	zfx	zdx	bxc	bfc	bpc	lxc	lfc	lpc	nxc	nfc	mpc	zxc	zfc	zdc
Cranfield	۰.۳۷۱۱۹	۰.۳۲۱۲۴	۰.۳۱۴۴۱	۰.۳۴۰۰۴	۰.۳۸۷۱	۰.۳۷۴۱۳	۰.۳۶۵۹۶	۰.۳۷۷۹۹	۰.۳۶۶۰۷	۰.۳۴۰۰۴	۰.۳۸۷۱	۰.۳۷۴۱۳	۰.۳۱۷۷۵	۰.۳۳۸۲۵	۰.۳۳۰۱۴	۰.۳۷۷۳۳	۰.۴۰۳۴۱	۰.۳۹۴۸۵	۰.۳۷۵۲۹	۰.۳۸۸۵۵	۰.۳۷۹۵۶	۰.۳۷۲۳۱	۰.۴۰۳۴۱	۰.۳۹۴۸۵
Medline	۰.۴۲۶۳۳	۰.۴۷۲۹۵	۰.۴۷۲۹۷	۰.۴۶۳۱۹	۰.۵۰۱۸	۰.۵۰۰۸۴	۰.۴۸۱۵۷	۰.۵۱۵۹۸	۰.۵۱۵۲۱	۰.۴۴۳۱۹	۰.۵۰۱۸	۰.۵۰۰۸۸	۰.۴۵۸۳۶	۰.۴۸۲۷۱	۰.۴۸۲۱۴	۰.۴۶۷۳۹	۰.۵۱۵۲۵	۰.۵۱۴۸۶	۰.۴۹۴۱۹	۰.۵۲۳۶۵	۰.۵۲۳۳	۰.۴۶۷۳۹	۰.۵۱۵۲۵	۰.۵۱۴۸۶

وزن‌های پایه طبق رابطه (۱۲) نیز این وضعیت باقی مانده و وزن این لغات از لغاتی که اصلاً در سند وجود نداشته‌اند کمتر می‌شود؛ بنابراین طبیعی است که در چنین مواردی، روش‌های مذکور ضعیف‌تر عمل کنند.

### ۲.۵ اعمال روش پیشنهادی بر روی زیرمجموعه‌هایی از وزن‌های پایه

با توجه به اینکه ایده اصلی این تحقیق، تعیین وزن‌های فاصله‌ای برای لغات با اعمال روش پیشنهادی بر روی زیرمجموعه‌ای از وزن‌های پایه است که با روش‌های مختلف وزن‌دهی بدست آمده‌اند، بر اساس مشاهدات جدول ۳ و نکات برگرفته از آن که در بخش ۵-۱ اشاره شد، در انتخاب زیرمجموعه‌های مختلف از ۲۴ روش وزن‌دهی پایه، می‌توان موارد زیر را در نظر گرفت که در جدول ۵ نیز با جزئیات بیشتری آورده شده‌اند:

- أ) در نظر گرفتن زیرمجموعه‌هایی فقط از روش‌های ۱-۱۲ (A<sub>۱</sub>)، فقط از روش‌های ۱۳-۲۴ (A<sub>۲</sub>)، یا از روش‌های ۱-۲۴ (A<sub>۳</sub>)
- ب) انتخاب روش‌های فاقد مؤلفه Z (B<sub>۱</sub>)، روش‌های فاقد مؤلفه n (B<sub>۲</sub>)، یا روش‌های شامل هر دو مؤلفه Z و n (B<sub>۳</sub>)
- ج) انتخاب زیرمجموعه‌هایی شامل روش‌های ۱ یا ۱۳ (C<sub>۱</sub>)، فاقد روش‌های ۱ و ۱۳ (C<sub>۲</sub>)، و علاوه بر کنار گذاشتن روش‌های ۱ و ۱۳ فاقد مؤلفه p در فاکتور فرکانس مجموعه (C<sub>۳</sub>)

در واقع برای هر یک از سه دسته از عواملی که ممکن است در انتخاب زیرمجموعه‌های وزنی تأثیرگذار باشند، سه گزینه برای

جدول ۴. بیشترین، کمترین، و میانگین مقادیر MAP بدست آمده از ۲۴ روش وزن‌دهی پایه

	Cranfield	Medline
بیشترین	۰.۴۰۳۴۱	۰.۵۲۳۶۵
کمترین	۰.۲۷۱۱۹	۰.۴۲۶۳۳
میانگین	۰.۳۶۲۰۹	۰.۴۸۹۰۴

کمک کند؛ از جمله:

- استفاده از روش‌های ۱۳-۲۴، به دلیل نرمال بودن طول بردار وزن‌ها، در تمامی موارد نسبت به استفاده از روش‌های ۱-۱۲ متناظر بهتر است.
- روش‌های وزن‌دهی ۱ و ۱۳ (bxx و bxc)، به دلیل دادن مقادیر وزنی بولی (۰ و ۱) به لغات، نسبت به سایر روش‌ها نتایج ضعیف‌تری دارند.
- استفاده از مؤلفه پیشنهادی Z به جای مؤلفه n در فاکتور اول (یعنی نرمال‌سازی معمولی TF در وزنهای ۱-۱۲ و ۲۲-۲۴ به جای نرمال‌سازی افزایش‌یافته آن در وزنهای ۷-۹ و ۱۹-۲۱)، در مجموعه داده Cranfield باعث نتایج بهتری شده است.
- استفاده از مؤلفه p به جای مؤلفه f در فاکتور دوم (یعنی فرکانس مجموعه معکوس احتمالی به جای فرکانس مجموعه معکوس)، در برخی موارد باعث نتایج ضعیف‌تری می‌شود. علت این موضوع به خاطر این است که در محاسبه مؤلفه p (یعنی  $\log \frac{N-n}{n}$ )، در مواردی که برای یک لغت داریم:  $n > \frac{N}{4}$ ، حاصل لگاریتم و در نتیجه وزن لغت منفی می‌شود؛ یعنی وزن چنین لغاتی حتی از لغاتی که اصلاً در سند وجود ندارند و وزن صفر گرفته‌اند هم کمتر می‌شود. با این حساب حتی پس از نرمال‌سازی

(۱۲) در بازه [۰،۱] نرمال سازی شد، بر اساس روش پیشنهادی در بخش ۴-۵ (رابطه (۱۳)) وزن فاصله‌ای لغات مشخص گردید و در ادامه طبق بخش ۴-۶ نیز بر اساس میانگین وزنی فاصله‌ای (IWA)، میزان ارتباط اسناد با پرس‌وجوی ورودی تعیین شد و نهایتاً بر اساس یکی از روش‌های سه‌گانه، رتبه‌بندی اسناد صورت گرفته و طبق روابط (۱۷ تا ۲۰) با بدست آوردن مقدار MAP و  $P@n$ ، نتیجه هر روش مورد ارزیابی قرار گرفت.

جدول ۶، نتایج MAP مربوط به این آزمایش‌ها را بر روی تمامی ۲۷ زیرمجموعه ممکن از وزن‌های پایه، در مجموعه داده‌های Cranfield و Medline، و برحسب روش‌های رتبه‌بندی مختلف (Rrank, Lrank, Mrank) نشان می‌دهد. در این جدول هر ردیف مربوط به یک زیرمجموعه از روش‌های وزن‌دهی پایه است که در سه ستون سمت چپ جدول، شرایط در نظر گرفته شده جهت انتخاب زیرمجموعه‌های وزنی پایه (طبق جدول ۵) برای آن ردیف معرفی شده است. بر اساس این شرایط، شماره روش‌های وزنی پایه انتخابی مورد استفاده در زیرمجموعه مربوط به هر ردیف از جدول دقیقاً مشخص می‌شود که در ستون چهارم جدول آورده شده است. در آزمایش‌های اولیه، زیرمجموعه‌های وزنی یکسانی برای اسناد پرس‌وجوها استفاده شد؛ هرچند در آزمایش‌ها مشاهده شد که کنار گذاشتن روش‌های وزن‌دهی شامل مؤلفه  $p$ ، از زیرمجموعه‌های وزنی، ممکن است برای اسناد مفید باشد اما برای پرس‌وجوها کارآیی سیستم را پایین می‌آورد و بهتر است این روش‌های وزن‌دهی در زیرمجموعه‌های وزنی پرس‌وجوها حتماً مورد استفاده قرار گیرند؛ به همین دلیل در جدول ۶ بجز در ردیف‌هایی که در ستون چهارم صراحتاً برای پرس‌وجوها زیرمجموعه‌ی دیگری ذکر شده، در بقیه موارد از زیرمجموعه‌های وزنی یکسانی برای اسناد و پرس‌وجوها استفاده شده است. به این ترتیب به عنوان مثال، سطر اول جدول ۶ بیان‌کننده این است که اگر در تشکیل زیرمجموعه وزنی، از وزن‌های پایه ۱ تا ۱۲ ( $A_1$ ) که نرمال سازی طول بردار بر روی آنها انجام نشده استفاده گردد، بطوریکه فاقد روش‌های وزن‌دهی برگرفته از مؤلفه  $p$  بوده ( $B_1$ ) و شامل روش‌های وزن‌دهی بولی ۱ یا ۱۳ باشند ( $C_1$ )، در اینصورت زیرمجموعه انتخابی مورد آزمایش شامل روش‌های وزن‌دهی پایه ۱ تا ۹ (یا به عبارتی روش‌های  $bxx$  تا  $np_x$  از جدول ۳) خواهد بود که در شش ستون بعد نتایج MAP برای دو مجموعه داده و با استفاده از هر یک از سه روش مختلف رتبه‌بندی آورده شده است. برای هر زیرمجموعه در هر مجموعه داده، بهترین روش رتبه‌بندی با زمینه طوسی مشخص شده است. همچنین مقادیر ضخیم شده در هر ستون، معرف زیرمجموعه‌ای است که بهترین نتیجه را برای آن روش رتبه‌بندی، در آن مجموعه

انتخاب در نظر گرفته و هر کدام با نمادهایی معرفی شده‌اند. با در نظر گرفتن این موارد و انتخاب گزینه‌های مختلف، بطور کلی امکان جدول ۵. نمادهای مورد استفاده جهت معرفی زیرمجموعه‌های وزنی پایه انتخابی

نمادهای مربوط به نرمال سازی طول	
$A_1$	در نظر گرفتن زیرمجموعه‌هایی فقط شامل روش‌های وزن‌دهی پایه بدون نرمال سازی طول (روش‌های ۱: $bxx$ تا ۱۲: $zpx$ )
$A_2$	در نظر گرفتن زیرمجموعه‌هایی فقط شامل روش‌های وزن‌دهی پایه همراه با نرمال سازی طول (روش‌های ۱۳: $bxc$ تا ۲۴: $zpc$ )
$A_3$	در نظر گرفتن زیرمجموعه‌هایی بدون محدودیت در این خصوص، یعنی زیرمجموعه‌هایی شامل هم روش‌های وزن‌دهی پایه بدون نرمال سازی طول و هم روش‌های وزن‌دهی پایه همراه با نرمال سازی طول (روش‌های ۱: $bxx$ تا ۲۴: $zpc$ )
نمادهای مربوط به مؤلفه فرکانس لغت انتخابی	
$B_1$	در نظر گرفتن زیرمجموعه‌هایی فقط شامل روش‌های وزن‌دهی پایه که در آنها از مؤلفه $Z$ استفاده نشده (عدم استفاده از روش‌های ۱۰: $zxx$ تا ۱۲: $zpx$ و ۲۲: $zxc$ تا ۲۴: $zpc$ )
$B_2$	در نظر گرفتن زیرمجموعه‌هایی فقط شامل روش‌های وزن‌دهی پایه که در آنها از مؤلفه $n$ استفاده نشده (عدم استفاده از روش‌های ۷: $nxx$ تا ۹: $np_x$ و ۱۹: $nxc$ تا ۲۱: $npc$ )
$B_3$	در نظر گرفتن زیرمجموعه‌هایی از روش‌های وزن‌دهی پایه بدون محدودیت در مؤلفه فرکانس لغت مورد استفاده (یعنی در زیرمجموعه‌های وزنی هم از روش‌های شامل مؤلفه $Z$ می‌توان استفاده کرد و هم از روش‌های شامل مؤلفه $n$ )
نمادهای مربوط به استفاده از روش بولی و روش‌های وزن‌دهی شامل مؤلفه $p$ فرکانس مجموعه	
$C_1$	در نظر گرفتن زیرمجموعه‌هایی از روش‌های وزن‌دهی پایه بدون محدودیت در این خصوص (یعنی در زیرمجموعه‌های وزنی هم از روش‌های وزن‌دهی بولی استفاده می‌شود و هم از روش‌های وزن‌دهی شامل مؤلفه $p$ )
$C_2$	در نظر گرفتن زیرمجموعه‌هایی فاقد روش‌های وزن‌دهی پایه بولی (عدم استفاده از روش‌های ۱: $bxx$ و ۱۳: $bxc$ )
$C_3$	در نظر گرفتن زیرمجموعه‌هایی فاقد روش‌های وزن‌دهی پایه بولی (روش‌های ۱: $bxx$ و ۱۳: $bxc$ ) و همچنین فاقد روش‌های وزن‌دهی شامل مؤلفه $p$ فرکانس مجموعه (یعنی روش‌های ۳: $bpx$ ، ۶: $tp_x$ ، ...، ۲۴: $zpc$ )

انتخاب ۲۷ زیرمجموعه مختلف از ۲۴ روش پایه وجود دارد که در آزمایش‌ها مورد بررسی و ارزیابی قرار گرفته‌اند. در آزمایش‌ها برای هر یک از این زیرمجموعه‌ها، پس از اینکه بر اساس هر یک از روش‌های وزن‌دهی در زیرمجموعه، وزن‌های پایه‌ی لغات مختلف اسناد و پرس‌وجوها طبق روابط (۱۰) و (۱۱) تعیین و طبق رابطه

داده، داشته است. نتایج این جدول از جنبه‌های مختلفی قابل بحث می‌باشد که در ادامه به آنها اشاره خواهد شد.

جدول ۶. مقدار MAP زیرمجموعه‌های وزنی مختلف، در روش وزن‌دهی فاصله‌ای، برحسب روش‌های رتبه‌بندی مختلف، بر روی مجموعه داده‌های Cranfield و Medline (مقادیر با زمینه طوسی، بیشترین مقدار بین سه روش رتبه‌بندی در هر مجموعه داده، و مقادیر ضخیم شده، بیشترین مقدار در هر ستون هستند)

			Cranfield			Medline			
			Lrank	Rrank	Mrank	Lrank	Rrank	Mrank	
			روش‌های وزن‌دهی پایه مورد استفاده در زیرمجموعه‌ها						
روش‌های ۱-۱۳: (بدون نرمال‌سازی طول)	B <sub>r</sub> :Lacking z component	C <sub>r</sub> : including ۱ or ۱۳	۱-۹	۰,۳۹۰۱۱	۰,۳۷۴۶۱	۰,۳۹۸۲۶	۰,۵۲۷۴۶	۰,۴۹۸۵۳	۰,۵۰۹۹۶
		C <sub>r</sub> : excluding ۱ & ۱۳	۲-۹	۰,۳۹۰۰۶	۰,۴۰۳۲۳	۰,۴۱۸۲۴	۰,۵۳۲۱۳	۰,۵۲۲۲۶	۰,۵۲۹۵۴
		C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۷-۸ (weighting subset for queries: ۲-۹)	۰,۳۹۹۹۳	۰,۴۰۲۰۵	۰,۴۱۷۰۸	۰,۵۳۴۸۹	۰,۵۱۷۱۲	۰,۵۲۸۶۴
	B <sub>r</sub> :Lacking n component	C <sub>r</sub> : includ. ۱ or ۱۳	۱-۶, ۱۰-۱۲	۰,۳۷۲۰۳	۰,۳۸۵۰۶	۰,۴۰۷۷۵	۰,۴۸۲۶۴	۰,۵۰۴۷۰	۰,۵۲۳۴۲
		C <sub>r</sub> : exclud. ۱ & ۱۳	۲-۶, ۱۰-۱۲	۰,۳۸۹۸۷	۰,۳۹۶۹۱	۰,۴۰۶۲۱	۰,۵۰۴۳۸	۰,۵۲۴۸۹	۰,۵۲۸۶۷
		C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۱۰-۱۱ (queries: ۲-۶, ۱۰-۱۲)	۰,۳۹۳۴۱	۰,۳۹۹۳۷	۰,۴۰۸۱۷	۰,۴۹۹۹۷	۰,۵۱۹۲۹	۰,۵۲۱۵۲
	B <sub>r</sub> :including either z or n component	C <sub>r</sub> : includ. ۱ or ۱۳	۱-۱۲	۰,۳۸۶۶۴	۰,۳۸۶۰۷	۰,۴۰۸۴۲	۰,۵۱۳۰۲	۰,۵۰۳۲۱	۰,۵۱۹۷۲
		C <sub>r</sub> : exclud. ۱ & ۱۳	۲-۱۲	۰,۳۹۱۲۴	۰,۴۰۲۹۷	۰,۴۱۳۰۴	۰,۵۱۹۱۷	۰,۵۲۳۱۵	۰,۵۳۲۱۹
			C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۷-۸, ۱۰-۱۱ (queries: ۲-۱۲)	۰,۳۹۹۰۳	۰,۴۰۱۳۳	۰,۴۱۵۴۳	۰,۵۱۶۳۹	۰,۵۱۹۱۱
روش‌های ۱۳-۲۴: (همراه با نرمال‌سازی طول)	B <sub>r</sub> :Lacking z component	C <sub>r</sub> : includ. ۱ or ۱۳	۱۳-۲۱	۰,۳۷۵۸۵	۰,۴۲۷۹	۰,۴۳۲۲۳	۰,۵۳۲۹۸	۰,۵۲۵۶۷	۰,۵۳۷۹۸
		C <sub>r</sub> : exclud. ۱ & ۱۳	۱۴-۲۱	۰,۳۸۱۶۴	۰,۴۲۵۴۴	۰,۴۳۲۳۵	۰,۵۳۴۴۱	۰,۵۲۳۶۹	۰,۵۴۱۳۴
		C <sub>r</sub> : excluding ۱, ۱۳ & p	۱۴, ۱۶-۱۷, ۱۹-۲۰ (queries: ۱۴-۲۱)	۰,۳۹۷۹۹	۰,۴۲۶۹۸	۰,۴۳۲۴۲	۰,۵۴۳۱۳	۰,۵۲۱۰۰	۰,۵۳۹۰۵
	B <sub>r</sub> :Lacking n component	C <sub>r</sub> : includ. ۱ or ۱۳	۱۳-۱۸, ۲۲-۲۴	۰,۴۱۴۳۲	۰,۴۱۹۲۷	۰,۴۲۹۶	۰,۵۴۳۹۴	۰,۵۱۴۴۰	۰,۵۲۸۸۶
		C <sub>r</sub> : exclud. ۱ & ۱۳	۱۴-۱۸, ۲۲-۲۴	۰,۴۱۱۱۱	۰,۴۱۶۲۳	۰,۴۲۴۷۶	۰,۵۴۲۰۰	۰,۵۱۲۵۰	۰,۵۲۵۹۴
		C <sub>r</sub> : excluding ۱, ۱۳ & p	۱۴, ۱۶-۱۷, ۲۲-۲۳ (queries: ۱۴-۱۸, ۲۲-۲۴)	۰,۴۱۹۳۷	۰,۴۱۱۹۹	۰,۴۲۲۳۶	۰,۵۴۳۷۰	۰,۵۰۶۶۸	۰,۵۲۱۱۵
	B <sub>r</sub> :including either z or n component	C <sub>r</sub> : includ. ۱ or ۱۳	۱۳-۲۴	۰,۴۰۳۵۵	۰,۴۲۴۳۵	۰,۴۳۱۸۹	۰,۵۴۵۸۰	۰,۵۱۹۲۷	۰,۵۳۵۹۷
		C <sub>r</sub> : exclud. ۱ & ۱۳	۱۴-۲۴	۰,۴۰۵۸۷	۰,۴۲۳۱۲	۰,۴۲۹۵۱	۰,۵۴۳۳۰	۰,۵۱۹۱۶	۰,۵۳۶۱۹
			C <sub>r</sub> : excluding ۱, ۱۳ & p	۱۴, ۱۶-۱۷, ۱۹-۲۰, ۲۲-۲۳ (queries: ۱۴-۲۴)	۰,۴۱۵۳۸	۰,۴۲۱۶۹	۰,۴۳۱۳۲	۰,۵۴۵۴۴	۰,۵۱۴۷۱
روش‌های ۲۴-۱ (شامل هر دو دسته)	B <sub>r</sub> :Lacking z component	C <sub>r</sub> : includ. ۱ or ۱۳	۱-۹, ۱۳-۲۱	۰,۲۷۶۲۷	۰,۴۰۴۸۷	۰,۴۱۵۷۲	۰,۴۴۱۷۰	۰,۵۱۳۴۹	۰,۵۲۱۰۷
		C <sub>r</sub> : exclud. ۱ & ۱۳	۲-۹, ۱۴-۲۱	۰,۳۴۵۳	۰,۴۲۱۹	۰,۴۲۵۰۲	۰,۵۱۱۶۲	۰,۵۲۸۶۹	۰,۵۳۱۱۳
		C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۷-۸, ۱۴, ۱۶-۱۷, ۱۹-۲۰ (queries: ۲-۹, ۱۴-۲۱)	۰,۳۶۳۰۳	۰,۴۲۰۶۶	۰,۴۲۵۸۲	۰,۵۱۷۵۹	۰,۵۳۰۶۳	۰,۵۳۳۹۴
		C <sub>r</sub> : includ. ۱ or ۱۳	۱-۶, ۱۰-۱۲, ۱۳-۱۸, ۲۲-	۰,۲۷۱۱۸	۰,۴۱۲۳۳	۰,۴۱۸۸۶	۰,۳۷۹۷۱	۰,۵۲۲۶۲	۰,۵۳۱۳۰
	C <sub>r</sub> : exclud. ۱ & ۱۳	۲-۶, ۱۰-۱۲, ۱۴-۱۸, ۲۲-	۰,۳۶۹۷	۰,۴۱۱۱۳	۰,۴۱۷۲۱	۰,۴۹۳۴۹	۰,۵۲۶۴۶	۰,۵۲۹۶۱	

B <sub>r</sub> :Lacking n component	C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۱۰-۱۱, ۱۴, ۱۶- ۱۷, ۲۲-۲۳ (queries: ۲-۶, ۱۰-۱۲, ۱۴-۱۸, ۲۲-۲۴)	۰,۳۷۹۳۳	۰,۴۱۵۷۷	۰,۴۱۷۸۹	۰,۴۹۱۹۸	۰,۵۲۳۰۰	۰,۵۲۴۳۵
B <sub>r</sub> :including either z or n component	C <sub>r</sub> : includ. ۱ or ۱۳	۱-۲۴	۰,۲۹۲۰۴	۰,۴۱۴۵۴	۰,۴۲۲۲۷	۰,۴۳۰۰۶	۰,۵۲۲۶۹	۰,۵۳۰۰۴
	C <sub>r</sub> : exclud. ۱ & ۱۳	۲-۱۲, ۱۴-۲۴	۰,۳۵۷۱۳	۰,۴۱۷۷۲	۰,۴۲۱۲۹	۰,۴۹۷۹۰	۰,۵۲۹۸۲	۰,۵۳۴۰۰
	C <sub>r</sub> : excluding ۱, ۱۳ & p	۲, ۴-۵, ۷-۸, ۱۰-۱۱, ۱۴, ۱۶-۱۷, ۱۹-۲۰, ۲۲-۲۳ (queries: ۲-۱۲, ۱۴-۲۴)	۰,۳۶۸۳۱	۰,۴۲۱۱	۰,۴۲۳۱۸	۰,۴۹۵۵۲	۰,۵۲۹۱۵	۰,۵۳۰۲۹

## ۵\_۲\_۱ ارزیابی کلی نتایج

آمده از زیرمجموعه‌های وزنی مختلف، برای روش‌های رتبه‌بندی Lrank, Rrank و Mrank در مجموعه داده Cranfield، به ترتیب ۰,۰۳۹، ۰,۰۱۴ و ۰,۰۰۹، و در مجموعه داده Medline، به ترتیب ۰,۰۰۴، ۰,۰۰۸ و ۰,۰۰۷ بدست آمد. پراکندگی کمتر بدین معنی است که وابستگی نتایج به زیرمجموعه‌های انتخاب شده از وزن‌های پایه، کمتر بوده و صرف‌نظر از زیرمجموعه انتخابی، بیشتر می‌توان به نتایج بدست آمده اطمینان کرد.

با محاسبه میانگین نتایج روش‌های مختلف در هر یک از مجموعه داده‌ها مشاهده می‌شود که میانگین نتایج روش وزن‌دهی فاصله‌ای، در مجموعه داده‌های Cranfield و Medline، به ترتیب ۰,۴۰۲۴۳ و ۰,۵۱۹۳۴ می‌باشد که نه تنها از میانگین نتایج روش‌های وزن‌دهی پایه متناظر هر مجموعه داده در جدول ۵ بهتر هستند بلکه نزدیک به بیشترین مقدار این مجموعه داده‌ها در جدول ۵ هستند که همگی نشان‌دهنده تأثیر مثبت در نظر گرفتن وزن‌های فاصله‌ای است.

## ۵\_۲\_۲ اثر نرمال‌سازی طول وزن‌های پایه

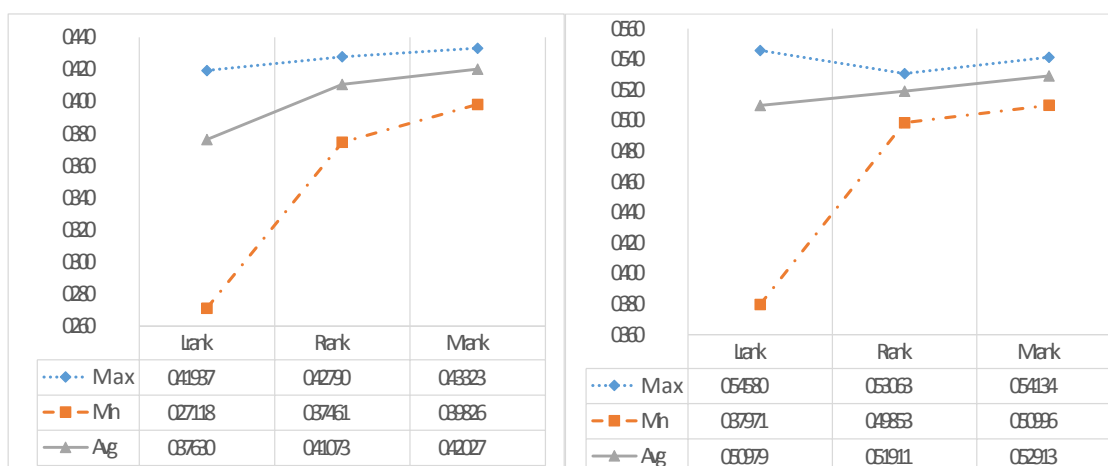
جهت بررسی اینکه نرمال‌سازی طول وزن‌های پایه مورد استفاده در زیرمجموعه‌های وزنی، چه تأثیری در نتایج دارد، میانگین نتایج MAP روش‌های وزن‌دهی در هر یک از نواحی A<sub>۱</sub>، A<sub>۲</sub> و A<sub>۳</sub>، برای هر یک از سه روش رتبه‌بندی Lrank, Rrank و Mrank، در هر یک از مجموعه‌های Cranfield و Medline، در جدول ۶ بدست آمد که در شکل ۲ در قالب سه خط مختلف قابل مشاهده هستند.

با نگاه به جدول ۶، مشاهده می‌شود که در مجموعه داده Creanfield، برای کلیه زیرمجموعه‌ها بهترین نتایج زمانی بدست آمده است که رتبه‌بندی بر اساس "مرکز" اعداد فاصله‌ای مربوط به شباهت‌ها (Mrank) صورت گرفته است. در مجموعه داده Medline نیز غالباً در استفاده از زیرمجموعه‌هایی از A<sub>۱</sub> و A<sub>۳</sub> رتبه‌بندی Mrank، و در استفاده از زیرمجموعه‌هایی از A<sub>۲</sub> (که اتفاقاً شامل بهترین نتیجه هم هست)، رتبه‌بندی Lrank بهتر بوده است.

شکل ۱ که بر اساس جدول ۶ بدست آمده، مقادیر بیشترین، کمترین، و میانگین MAP تمامی ۲۷ زیرمجموعه وزنی مختلف از روش‌های وزن‌دهی پایه را به ازاء هر یک از سه روش رتبه‌بندی پیشنهادی (Lrank, Rrank و Mrank)، در هر یک از مجموعه داده‌های Cranfield و Medline، نشان می‌دهد. در اینجا نیز از روی مقادیر میانگین نتایج در سه روش رتبه‌بندی مختلف مشاهده می‌شود که در هر دو مجموعه داده، بطور متوسط نتایج روش رتبه‌بندی Rrank بهتر از Lrank و روش رتبه‌بندی Mrank بهتر از آن دو بوده است.

با مقایسه فواصل بین مقادیر بیشترین و کمترین، در روش‌های رتبه‌بندی مختلف مشاهده می‌شود که در هر دو مجموعه داده، پراکندگی نتایج بدست آمده از زیرمجموعه‌های وزنی مختلف، در روش رتبه‌بندی Rrank کمتر از Lrank، و در روش رتبه‌بندی Mrank کمتر از آن دو بوده است. این موضوع با محاسبه انحراف استاندارد نتایج نیز تایید شد، جاییکه انحراف استاندارد نتایج بدست

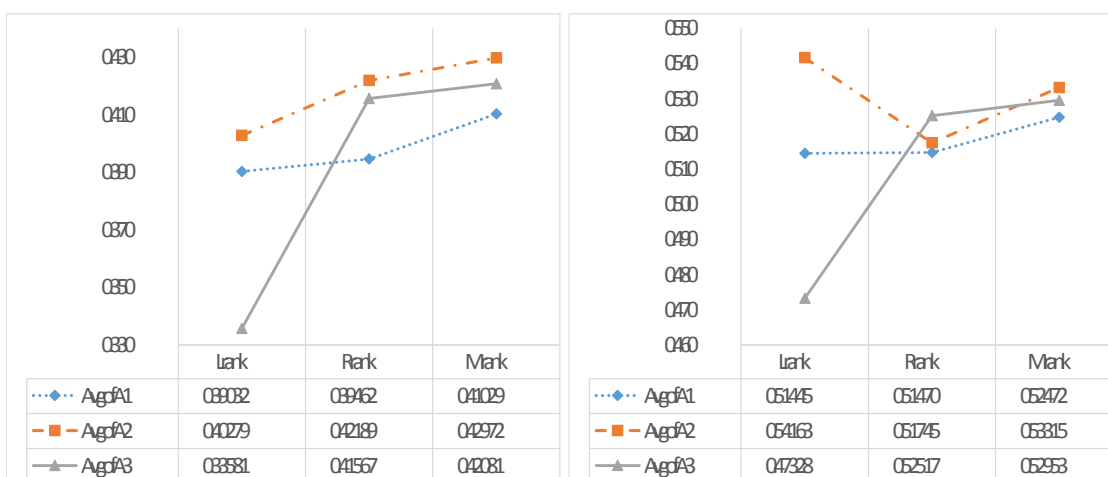




(a) Cranfield

(b) Medline

شکل ۱. مقایسه بیشترین، کمترین، و میانگین MAP تمامی ۲۷ زیرمجموعه وزنی، در روش‌های رتبه‌بندی مختلف جدول ۶، در مجموعه داده‌های Medline و Cranfield



(a) Cranfield

(b) Medline

شکل ۲. تاثیر نرمال‌سازی طول: میانگین MAP زیرمجموعه‌های وزنی نواحی  $A_1$ ،  $A_2$  و  $A_3$ ، در روش‌های رتبه‌بندی مختلف جدول ۶

### ۵\_۲\_۳ اثر استفاده از مؤلفه‌های نرمال‌سازی TF مختلف، در وزن‌های پایه

شکل ۳ از میانگین‌گیری بر روی نتایج MAP زیرمجموعه‌های وزنی در هر یک از نواحی  $B_1$ ،  $B_2$  و  $B_3$ ، برای هر یک از سه روش رتبه‌بندی Lrank، Rrank و Mrank، در مجموعه داده‌های Cranfield و Medline، در جدول ۶ بدست آمده تا تأثیر استفاده از هر یک مشخص شود. همانطور که مشاهده می‌شود، بطور کلی در هر دو مجموعه داده، استفاده از زیرمجموعه‌های وزنی  $B_1$  و  $B_3$  منجر به نتایج بهتری نسبت به  $B_2$  شده است؛ یعنی بهتر است در زیرمجموعه‌های انتخابی حتماً از روش‌های وزندهی پایه شامل مؤلفه  $n$ ، استفاده شود. در شکل ۳ مشاهده می‌شود که در هر دو مجموعه داده، در روش رتبه‌بندی Mrank نتایج بهتری بدست آمده است اما اگر نتایج بدست آمده در بخش ۵-۲-۲ نیز مدنظر

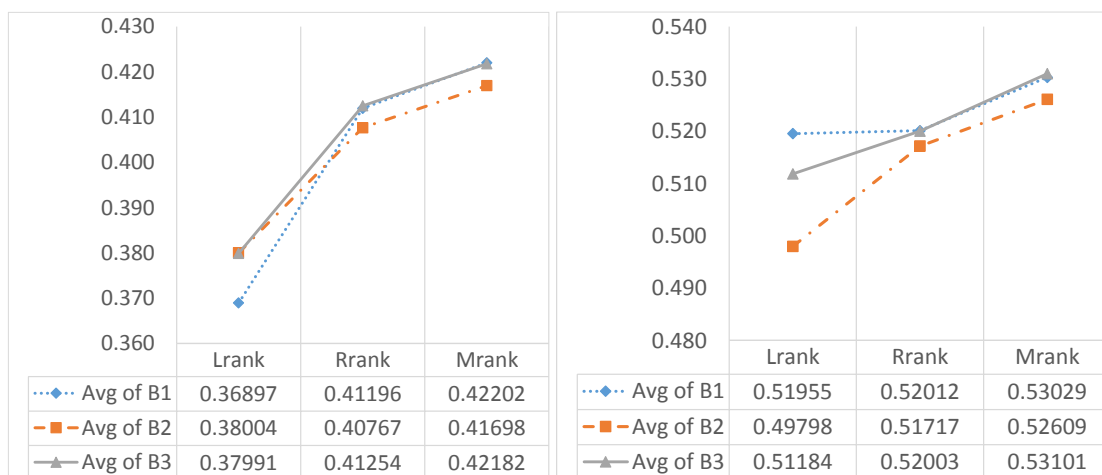
همانطور که مشاهده می‌شود، در کل، نتایج بدست آمده از زیرمجموعه‌های وزنی ناحیه  $A_2$  مناسب‌تر از  $A_3$ ، و این نیز غالباً بهتر از  $A_1$  بوده است که نشان‌دهنده تأثیر مثبت نرمال‌سازی طول بردار وزن‌های پایه مورد استفاده است. البته در روش رتبه‌بندی Rrank در مجموعه داده Medline، نتایج  $A_3$  بهتر از  $A_2$  شده، اما همانطور که مشاهده می‌شود و در بخش ۵-۲-۱ نیز اشاره شد، در این مجموعه داده نتایج روش‌های رتبه‌بندی Lrank و Mrank مناسب‌تر بوده و بیشتر مدنظر هستند. همچنین مشاهده می‌شود که در استفاده از زیرمجموعه‌های وزنی ناحیه  $A_2$ ، در مجموعه داده Cranfield استفاده از روش رتبه‌بندی Mrank، و در مجموعه داده Medline استفاده از روش رتبه‌بندی Lrank، مناسب‌تر بوده است.

رتبه‌بندی  $Lrank$ ،  $Rrank$  و  $Mrank$ ، برای هر دو مجموعه داده در جدول ۶ بدست آمده است. همانطور که مشاهده می‌شود، استفاده از روش‌های وزن‌دهی بولی در زیرمجموعه وزن‌های پایه ( $C_1$ )، در هر دو مجموعه داده باعث کاهش در نتایج MAP شده است. بعد از کنار گذاشتن روش‌های وزن‌دهی بولی، بطور متوسط، در مجموعه داده‌های Cranfield عدم استفاده از وزن‌های وزن‌دهی شامل مؤلفه  $p$  فرکانس مجموعه در زیرمجموعه‌های وزن‌دهی مناسب‌تر است (یعنی  $C_3$  بهتر از  $C_2$ )، در حالیکه در مجموعه داده Medline، بجز در روش  $Lrank$  در روش‌های  $Rrank$  و  $Mrank$  برعکس است.

قرار داده و تأثیر این مؤلفه‌ها فقط در ناحیه برتر  $A_2$  مورد بررسی قرار گیرد، مشاهده می‌شود که در  $A_2$ ، در مجموعه داده Cranfield، زیرمجموعه‌های وزنی  $B_1$  در روش رتبه‌بندی  $Mrank$  با میانگین  $0.43267$ ، بهترین نتایج را داده‌اند. در مجموعه داده Medline، نیز با اینکه نتایج این ناحیه خوب هستند اما بهترین نتایج به ازاء زیرمجموعه‌های وزنی  $B_3$  در روش رتبه‌بندی  $Lrank$  با میانگین  $0.54485$ ، بدست آمده است.

### ۴\_۲\_۵ اثر استفاده از روش‌های بولی، و روش‌های وزن‌دهی شامل مؤلفه $P$

شکل ۴ از میانگین‌گیری بر روی نتایج MAP زیرمجموعه‌های وزنی در هر یک از نواحی  $C_1$ ،  $C_2$  و  $C_3$ ، در هر یک از سه روش



(a) Cranfield

(b) Medline

شکل ۳. تأثیر مؤلفه‌های نرمال‌سازی TF مختلف: میانگین MAP زیرمجموعه‌های وزنی نواحی  $B_1$ ،  $B_2$  و  $B_3$ ، در روش‌های رتبه‌بندی مختلف جدول ۶

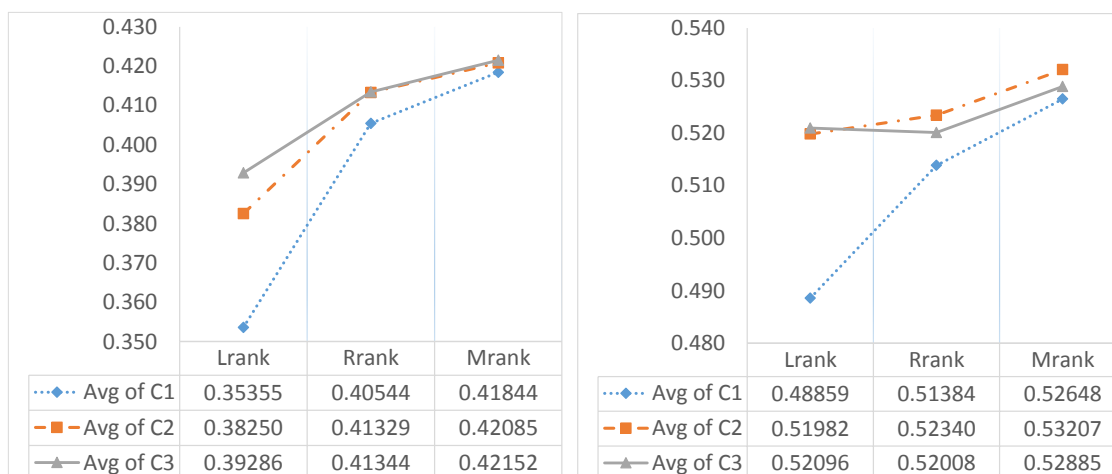
اگر به پنج روش برتر در هر یک از مجموعه داده‌ها که در جدول ۶ زیر آنها خط کشیده شده توجه شود نیز به همان نتایجی که در بحث‌های قبلی اشاره شد، رسیده و مشاهده می‌گردد که:

- در مجموعه Cranfield بهترین نتایج مربوط به استفاده از زیرمجموعه‌هایی از روش‌های وزن‌دهی  $13-24$  بوده ( $A_2$ ) که در آنها حتماً از وزن‌های پایه شامل مؤلفه  $n$  استفاده شده ( $B_1$  یا  $B_2$ ) و جهت رتبه‌بندی نهایی از  $Mrank$  استفاده شده است. در این شرایط، استفاده از زیرمجموعه‌های وزنی  $C_1$  و سپس  $C_3$  مناسب‌تر بوده‌اند، یعنی استفاده از وزن‌های بولی و وزن‌های شامل مؤلفه  $p$  در زیرمجموعه‌های وزنی ( $C_1$ )، یا عدم استفاده از هیچیک از آنها ( $C_3$ )

در شکل ۴ نیز شاید در کل، در هر دو مجموعه داده، روش رتبه‌بندی  $Mrank$  نتایج بهتری داده است اما اگر با توجه به نتایج بخش ۵-۲-۲ فقط ناحیه برتر  $A_2$  مورد بررسی قرار گیرد، مشاهده می‌شود که در  $A_2$ ، در مجموعه داده Cranfield، زیرمجموعه‌های وزنی  $C_1$  در روش رتبه‌بندی  $Mrank$  با میانگین  $0.43157$ ، و در مجموعه داده Medline، زیرمجموعه‌های وزنی  $C_3$  در روش رتبه‌بندی  $Lrank$  با میانگین  $0.54409$ ، بهترین نتایج را داده‌اند.

### ۵\_۲\_۵ بررسی بهترین روش‌ها

در جدول ۶، نتایج مربوط به ۸۱ آزمایش مختلف بر روی دو مجموعه داده آورده شد و در بخش‌های قبل اثر روش‌ها، پارامترها، و مؤلفه‌های مختلف در این نتایج مورد بررسی و بحث قرار گرفت.



(a) Cranfield

(b) Medline

شکل ۴. تاثیر استفاده از روش‌های بولی و مؤلفه  $p$ : میانگین MAP زیرمجموعه‌های وزنی نواحی  $C_1, C_2, C_3$ . در روش‌های رتبه‌بندی مختلف جدول ۶

- در ۲۰ درصد، از هر دو سری وزن‌های پایه‌ی با نرمال‌سازی طول و بدون نرمال‌سازی طول ( $A_2$ ) استفاده شده است، اما زیرمجموعه‌هایی که در آنها فقط وزن‌های پایه‌ی بدون نرمال‌سازی طول ( $A_1$ ) بکار رفته باشد، مشاهده نمی‌شود.
- در ۵۰ درصد از آنها، از وزن‌های پایه فاقد مؤلفه  $Z$  در فاکتور TF (زیرمجموعه‌هایی از  $B_1$ )، در ۲۰ درصد از روش‌ها از وزن‌های فاقد مؤلفه  $n$  (زیرمجموعه‌هایی از  $B_2$ )، و در ۳۰ درصد دیگر از هر دو مؤلفه (زیرمجموعه‌هایی از  $B_3$ ) استفاده شده است.
- ۳۰ درصد از این روش‌ها، از وزن‌های پایه بولی و وزن‌های شامل مؤلفه  $p$  در فرکانس مجموعه (زیرمجموعه‌هایی از  $C_1$ ) استفاده کرده‌اند، در ۳۰ درصد، بدون استفاده از وزن‌های بولی، از وزن‌های پایه شامل مؤلفه  $p$  در فرکانس مجموعه (زیرمجموعه‌هایی از  $C_2$ ) استفاده شده و در ۴۰ درصد باقیمانده علاوه بر عدم استفاده از وزن‌های بولی، از روش‌های فاقد مؤلفه  $p$  (زیرمجموعه‌هایی از  $C_3$ ) استفاده شده است.

- در مجموعه Medline بهترین نتایج مربوط به استفاده از زیرمجموعه‌هایی از روش‌های وزن‌دهی ۱۳-۲۴ بوده ( $A_2$ ) که در آنها حتماً از وزن‌های پایه شامل مؤلفه  $Z$  استفاده شده ( $B_2$  یا  $B_3$ ) و جهت رتبه‌بندی نهایی از Lrank استفاده شده است. در این شرایط نیز مانند Cranfield، استفاده از زیرمجموعه‌های وزنی  $C_1$  و سپس  $C_3$  مناسب‌تر بوده‌اند.

مشاهده می‌گردد که در رسیدن به بهترین نتایج، خیلی از شرایط بین دو مجموعه داده یکسان است اما به هر حال در مواردی که با هم اختلاف دارند، چه مؤلفه‌ها و روش‌هایی را در نظر گرفت که صرف‌نظر از خصوصیات مجموعه داده، نتایج بهتری بدهد؟ برای رسیدن به پاسخ این سوال، ابتدا رتبه‌ی هر یک از ۸۱ روش مذکور در مجموعه داده مربوطه بدست آمد و سپس میانگین رتبه این روش‌ها در دو مجموعه داده محاسبه و به صورت صعودی مرتب شدند تا برترین روش‌ها بدست آیند. در جدول ۷ میانگین رتبه‌ی ده روش برتر در دو مجموعه داده مشاهده می‌شود، ضمن اینکه مؤلفه‌ها و روش‌های بکار رفته در هر روش به همراه رتبه و MAP آنها در هر یک از مجموعه داده‌ها آورده شده است. طبق جدول ۷ مشاهده می‌شود که:

- در ۸۰ درصد از ده روش برتر، از وزن‌های پایه‌ی با نرمال‌سازی طول در زیرمجموعه‌ها ( $A_2$ ) استفاده شده و

جدول ۷. روش‌های برتر وزن‌دهی و بازیابی، به همراه رتبه و مقدار MAP هر یک در هر مجموعه داده و متوسط رتبه آنها در دو مجموعه

	Methods				Rank / MAP		Mean rank for ۲ datasets
	Length Normalization	Using n & z components	Using p components	Ranking Method	Cranfield	Medline	
۱	A <sub>r</sub>	B <sub>۱</sub>	C <sub>۱</sub>	Mrank	۱ / .۴۳۳۲۳	۱۰ / .۸۳۷۹۱	۵,۵
۱	A <sub>r</sub>	B <sub>۱</sub>	C <sub>r</sub>	Mrank	۳ / .۴۳۲۳۸	۸ / .۸۴۱۳۴	۵,۵
۱	A <sub>r</sub>	B <sub>۱</sub>	C <sub>r</sub>	Mrank	۲ / .۴۳۲۴۲	۹ / .۸۳۹۰۸	۵,۵
۴	A <sub>r</sub>	B <sub>r</sub>	C <sub>۱</sub>	Mrank	۴ / .۴۳۱۸۹	۱۲ / .۸۳۸۹۷	۸
۵	A <sub>r</sub>	B <sub>r</sub>	C <sub>r</sub>	Mrank	۷ / .۴۲۹۸۱	۱۱ / .۸۳۶۱۹	۹
۶	A <sub>r</sub>	B <sub>r</sub>	C <sub>r</sub>	Mrank	۵ / .۴۳۱۳۲	۲۰ / .۸۳۱۸۶	۱۲,۵
۷	A <sub>r</sub>	B <sub>۱</sub>	C <sub>r</sub>	Mrank	۱۰ / .۴۲۸۸۲	۱۶ / .۸۳۳۹۴	۱۳
۸	A <sub>r</sub>	B <sub>r</sub>	C <sub>r</sub>	Lrank	۲۴ / .۴۱۹۳۷	۴ / .۸۴۳۷۰	۱۴
۹	A <sub>r</sub>	B <sub>۱</sub>	C <sub>r</sub>	Mrank	۱۲ / .۴۲۸۰۲	۲۲ / .۸۳۱۱۳	۱۷
۱۰	A <sub>r</sub>	B <sub>r</sub>	C <sub>۱</sub>	Mrank	۶ / .۴۲۹۶	۳۰ / .۸۲۸۸۶	۱۸

روش رتبه‌بندی Mrank، شاید عدم استفاده از وزن‌های بولی (C<sub>r</sub>) نتایج بهتری را داشته باشد.

#### ۵\_۲\_۶ بررسی معیار P@N

باید توجه داشت که اکثر کاربران موتورهای جستجو تنها نتایج ابتدای لیست اسناد برگشتی را بررسی کرده و برای ۸۰ درصد آنها تنها سه نتیجه برگشتی ابتدایی جالب می‌باشد [۵۰]. برای بررسی نتایج از این نقطه نظر، در جدول ۸ مقادیر معیارهای P@۵، P@۱۰ و P@۲۰، برای وزن‌های پایه ۱۳ تا ۲۴ (که زیرمجموعه‌های انتخابی از آنها در بخش قبل مناسب‌تر تشخیص داده شدند)، و در جدول ۹ مقادیر این معیارها برای وزن‌های فاصله‌ای بدست آمده از زیرمجموعه‌های مختلف وزن‌های ۱۳ تا ۲۴، به ازاء روش‌های رتبه‌بندی مختلف، در دو مجموعه داده Cranfield و Medline آورده شده است.

- در ۹۰ درصد از موارد از روش رتبه‌بندی Mrank، و در ۱۰ درصد از روش Lrank استفاده شده است، اما در هیچیک از ده روش برتر از روش رتبه‌بندی Rrank استفاده نشده است.

با توجه به جدول ۷ و بررسی‌های فوق، می‌توان گفت که بطورکلی در روش پیشنهادی با ۲۴ روش وزن‌دهی پایه در نظر گرفته شده، استفاده از زیرمجموعه‌هایی از وزن‌های پایه که نرمال‌سازی طول روی آنها انجام شده (A<sub>r</sub>)، فاقد مؤلفه Z در فاکتور TF هستند (B<sub>۱</sub>)، در کنار استفاده از روش رتبه‌بندی Mrank مناسب‌تر است. با نگاه به شرایط سه سطر اول جدول ۷، مشاهده می‌شود که استفاده یا عدم استفاده از وزن‌های بولی و وزن‌های پایه شامل مؤلفه p در فرکانس مجموعه (C<sub>i</sub>های مختلف)، تأثیری در میانگین رتبه‌های آنها نداشته است؛ با این وجود با توجه به شکل ۴، در

جدول ۸. مقادیر دقت در برترین ۵، ۱۰، و ۲۰ سند ارائه شده، به ازاء روش‌های وزن‌دهی پایه ۱۳ تا ۲۴، در هر مجموعه داده

شماره و نام سه جزئی روش	Cranfield			Medline			
	P@۵	P@۱۰	P@۲۰	P@۵	P@۱۰	P@۲۰	
	۱۳	bxc	۰,۳۳۹۵۶	۰,۲۴۴۴۴	۰,۱۶۳۷۸	۰,۷۲۶۶۷	۰,۵۸۰۰۰
۱۴	bfc	۰,۳۴۹۳۳	۰,۲۵۶۰۰	۰,۱۶۹۱۱	۰,۶۳۳۳۳	۰,۵۸۳۳۳	۰,۵۱۱۶۷
۱۵	bpc	۰,۳۴۳۱۱	۰,۲۴۹۷۸	۰,۱۶۶۶۷	۰,۶۴۰۰۰	۰,۵۸۶۶۷	۰,۵۰۶۶۷
۱۶	txc	۰,۳۹۲۰۰	۰,۲۷۷۳۳	۰,۱۸۳۷۸	۰,۶۷۳۳۳	۰,۵۷۰۰۰	۰,۴۶۸۳۳
۱۷	tfc	۰,۴۲۷۵۶	۰,۳۰۳۱۱	۰,۲۰۱۱۱	۰,۶۶۶۶۷	۰,۶۰۳۳۳	۰,۵۳۳۳۳
۱۸	tpc	۰,۴۱۲۴۴	۰,۲۹۷۷۸	۰,۱۹۶۲۲	۰,۶۶۰۰۰	۰,۶۰۶۶۷	۰,۵۳۳۳۳
۱۹	nxc	۰,۴۰۹۷۸	۰,۲۸۲۲۲	۰,۱۸۲۶۷	۰,۷۲۶۶۷	۰,۶۱۶۶۷	۰,۵۰۰۰۰
۲۰	nfc	۰,۴۱۰۶۷	۰,۲۸۷۵۶	۰,۱۸۷۱۱	۰,۶۷۳۳۳	۰,۶۳۶۶۷	۰,۵۳۵۰۰
۲۱	npc	۰,۴۰۲۶۷	۰,۲۸۵۳۳	۰,۱۸۴۰۰	۰,۶۷۳۳۳	۰,۶۳۳۳۳	۰,۵۳۸۳۳
۲۲	zxc	۰,۳۹۲۰۰	۰,۲۷۷۳۳	۰,۱۸۳۷۸	۰,۶۷۳۳۳	۰,۵۷۰۰۰	۰,۴۶۸۳۳
۲۳	zfc	۰,۴۲۷۵۶	۰,۳۰۳۱۱	۰,۲۰۱۱۱	۰,۶۶۶۶۷	۰,۶۰۳۳۳	۰,۵۳۳۳۳
۲۴	zpc	۰,۴۱۲۴۴	۰,۲۹۷۷۸	۰,۱۹۶۲۲	۰,۶۶۰۰۰	۰,۶۰۶۶۷	۰,۵۳۳۳۳
بیشترین مقدار		۰,۴۲۷۵۶	۰,۳۰۳۱۱	۰,۲۰۱۱۱	۰,۷۲۶۶۷	۰,۶۳۶۶۷	۰,۵۳۸۳۳
میانگین مقادیر		۰,۳۹۳۲۶	۰,۲۸۰۱۵	۰,۱۸۴۶۳	۰,۶۷۲۷۸	۰,۵۹۹۷۲	۰,۵۱۰۴۲

روش پیشنهادی از اکثر نتایج روش‌های وزن‌دهی پایه در جدول ۴ نیز بهتر شده است. بنابراین بر اساس هر دو معیار، اثربخشی ایجاد وزن‌های فاصله‌ای بر اساس وزن‌های پایه، نسبت به استفاده از هر یک از وزن‌های پایه به تنهایی، کاملاً مشهود می‌باشد. همچنین باید توجه داشت که محاسبه وزن‌های پایه اولیه مورد نظر و محاسبه وزن‌های فاصله‌ای برای لغات مختلف اسناد، به صورت آفلاین صورت گرفته و در زمان ورود یک پرس‌وجوی جدید، تنها محاسبه میزان ارتباط آن با اسناد مختلف (آن هم فقط بر اساس لغات درون پرس‌وجو) انجام می‌گیرد؛ بنابراین روش پیشنهادی ضمن افزایش دقت، از روشی ساده استفاده کرده و افزایش پیچیدگی و زمان اجرا را به همراه ندارد.

همانگونه که در جداول ۸ و ۹ مشاهده می‌شود، در وزن‌های فاصله‌ای نیز مانند وزن‌های پایه، مقادیر دقت در  $n$ های کوچکتر که بیشتر مورد توجه کاربران است، بالاتر می‌باشد. همچنین در مقایسه جداول ۶ و ۹، مشاهده می‌گردد که در معیار  $P@n$  نیز مانند معیار MAP در مجموعه داده Cranfield نتایج رتبه‌بندی بر اساس Mrank، و در مجموعه داده Medline نتایج رتبه‌بندی بر اساس Lrank مناسب‌تر بوده است. با مقایسه جداول ۸ و ۹ می‌توان متوجه شد که میانگین مقادیر بدست آمده از زیرمجموعه‌های مختلف در روش وزن‌دهی فاصله‌ای، به ازاء هر یک از روش‌های رتبه‌بندی، از میانگین روش‌های وزن‌دهی پایه‌ی مربوطه بهتر شده است. در روش وزن‌دهی فاصله‌ای، در مجموعه داده Cranfield، مقادیر میانگین Mrank (و در مجموعه داده Medline، مقادیر میانگین Lrank) حتی از بیشترین مقادیر معادل در جدول ۸ هم بهتر شده‌اند. با مقایسه نتایج MAP جدول ۶ با جداول ۴ و ۵ نیز مشاهده می‌گردد که نه تنها میانگین نتایج روش‌های پیشنهادی مختلف در این تحقیق از میانگین نتایج استفاده از هر یک از وزن‌های پایه به تنهایی، با اختلاف بهتر است، بلکه نتایج MAP

جدول ۹. مقادیر دقت در برترین ۵، ۱۰ و ۲۰ سند ارائه شده، برای زیرمجموعه‌های مختلف از روش‌های وزن‌دهی پایه ۱۳ تا ۲۴، در روش وزن‌دهی فاصله‌ای، برحسب روش‌های رتبه‌بندی مختلف، در هر مجموعه داده (مقادیر با زمینه طوسی، بیشترین مقدار بین سه روش رتبه‌بندی، و مقادیر ضخیم شده، بیشترین مقدار در هر ستون هستند)

	روش رتبه‌بندی	روش‌های وزن‌دهی پایه مورد استفاده در زیرمجموعه‌ها	Cranfield			Medline			
			P@۵	P@۱۰	P@۲۰	P@۵	P@۱۰	P@۲۰	
روش‌های ۱۳-۲۴ (همراه با نرمال‌سازی طول): A <sub>r</sub>	Lrank	(B <sub>۱</sub> .C <sub>۱</sub> )	۱۳-۲۱	۰,۴۰۳۵۶	۰,۲۸۴۰۰	۰,۱۸۶۲۲	۰,۷۲۶۶۷	۰,۶۴۰۰۰	۰,۵۳۸۲۳
			Rrank	۰,۴۵۲۴۴	۰,۳۱۲۰۰	۰,۲۰۵۳۳	۰,۷۱۳۳۳	۰,۶۴۶۶۷	۰,۵۴۰۰۰
			Mrank	۰,۴۵۶۰۰	۰,۳۱۴۶۷	۰,۲۰۶۶۷	۰,۷۲۶۶۷	۰,۶۵۰۰۰	۰,۵۴۰۰۰
	Lrank	(B <sub>۱</sub> .C <sub>۲</sub> )	۱۴-۲۱	۰,۴۱۰۶۷	۰,۲۸۴۸۹	۰,۱۸۶۰۰	۰,۷۱۳۳۳	۰,۶۵۰۰۰	۰,۵۴۱۶۷
			Rrank	۰,۴۴۷۱۱	۰,۳۱۰۶۷	۰,۲۰۶۶۷	۰,۶۹۳۳۳	۰,۶۴۳۳۳	۰,۵۴۳۳۳
			Mrank	۰,۴۶۰۴۴	۰,۳۱۲۰۰	۰,۲۰۷۵۶	۰,۷۲۰۰۰	۰,۶۶۰۰۰	۰,۵۳۶۶۷
	Lrank	(B <sub>۱</sub> .C <sub>۳</sub> )	۱۴, ۱۶-۱۷, ۱۹-۲۰ (queries: ۱۴-۲۱)	۰,۴۱۴۲۲	۰,۲۹۳۳۳	۰,۱۹۵۵۶	۰,۷۲۶۶۷	۰,۶۵۳۳۳	۰,۵۴۶۶۷
			Rrank	۰,۴۴۸۸۹	۰,۳۱۱۱۱	۰,۲۰۵۳۳	۰,۷۰۶۶۷	۰,۶۴۰۰۰	۰,۵۳۱۶۷
			Mrank	۰,۴۵۷۷۸	۰,۳۱۵۱۱	۰,۲۰۸۴۴	۰,۷۴۰۰۰	۰,۶۴۳۳۳	۰,۵۴۰۰۰
	Lrank	(B <sub>۲</sub> .C <sub>۱</sub> )	۱۳-۱۸, ۲۲-۲۴	۰,۴۴۵۳۳	۰,۳۰۴۸۹	۰,۱۹۸۶۷	۰,۷۲۶۶۷	۰,۶۵۶۶۷	۰,۵۴۸۲۳
			Rrank	۰,۴۴۱۷۸	۰,۳۱۱۵۶	۰,۲۰۴۴۴	۰,۷۰۶۶۷	۰,۶۲۳۳۳	۰,۵۳۰۰۰
			Mrank	۰,۴۵۷۷۸	۰,۳۱۲۴۴	۰,۲۰۹۳۳	۰,۷۱۳۳۳	۰,۶۳۳۳۳	۰,۵۳۸۲۳
	Lrank	(B <sub>۲</sub> .C <sub>۲</sub> )	۱۴-۱۸, ۲۲-۲۴	۰,۴۴۰۸۹	۰,۳۰۳۱۱	۰,۱۹۶۸۹	۰,۷۳۳۳۳	۰,۶۵۶۶۷	۰,۵۵۶۶۷
			Rrank	۰,۴۳۸۲۲	۰,۳۰۹۷۸	۰,۲۰۴۰۰	۰,۷۰۰۰۰	۰,۶۳۰۰۰	۰,۵۲۶۶۷
			Mrank	۰,۴۵۰۶۷	۰,۳۱۰۶۷	۰,۲۰۶۴۴	۰,۷۰۰۰۰	۰,۶۳۶۶۷	۰,۵۳۵۰۰
	Lrank	(B <sub>۲</sub> .C <sub>۳</sub> )	۱۴, ۱۶-۱۷, ۲۲-۲۳ (queries: ۱۴-۱۸, ۲۲-۲۴)	۰,۴۴۲۶۷	۰,۳۰۹۳۳	۰,۲۰۴۲۲	۰,۷۴۶۶۷	۰,۶۵۳۳۳	۰,۵۵۶۶۷
			Rrank	۰,۴۳۷۳۳	۰,۳۰۳۱۱	۰,۲۰۳۱۱	۰,۷۱۳۳۳	۰,۶۱۶۶۷	۰,۵۲۱۶۷
			Mrank	۰,۴۴۷۱۱	۰,۳۱۰۶۷	۰,۲۰۶۲۲	۰,۷۱۳۳۳	۰,۶۴۳۳۳	۰,۵۳۰۰۰
	Lrank	(B <sub>۳</sub> .C <sub>۱</sub> )	۱۳-۲۴	۰,۴۲۵۷۸	۰,۲۹۸۲۲	۰,۱۹۳۵۶	۰,۷۴۶۶۷	۰,۶۵۶۶۷	۰,۵۵۳۳۳
			Rrank	۰,۴۴۶۲۲	۰,۳۱۴۲۲	۰,۲۰۳۷۸	۰,۷۰۶۶۷	۰,۶۲۶۶۷	۰,۵۳۱۶۷
			Mrank	۰,۴۵۶۰۰	۰,۳۱۶۰۰	۰,۲۰۹۳۳	۰,۷۲۶۶۷	۰,۶۴۶۶۷	۰,۵۳۶۶۷
	Lrank	(B <sub>۳</sub> .C <sub>۲</sub> )	۱۴-۲۴	۰,۴۳۴۶۷	۰,۲۹۶۰۰	۰,۱۹۵۵۶	۰,۷۲۶۶۷	۰,۶۵۶۶۷	۰,۵۵۶۶۷
			Rrank	۰,۴۴۷۱۱	۰,۳۱۳۷۸	۰,۲۰۴۸۹	۰,۷۰۰۰۰	۰,۶۱۶۶۷	۰,۵۳۵۰۰
			Mrank	۰,۴۵۶۸۹	۰,۳۱۲۴۴	۰,۲۰۸۲۲	۰,۷۱۳۳۳	۰,۶۵۰۰۰	۰,۵۳۸۲۳
Lrank	(B <sub>۳</sub> .C <sub>۳</sub> )	۱۴, ۱۶-۱۷, ۱۹-۲۰, ۲۲-۲۳ (queries: ۱۴-۲۴)	۰,۴۳۷۳۳	۰,۳۰۰۸۹	۰,۲۰۳۳۳	۰,۷۳۳۳۳	۰,۶۶۰۰۰	۰,۵۵۱۶۷	
		Rrank	۰,۴۴۱۷۸	۰,۳۰۸۸۹	۰,۲۰۵۳۳	۰,۷۰۰۰۰	۰,۶۳۳۳۳	۰,۵۳۳۳۳	
		Mrank	۰,۴۵۴۲۲	۰,۳۱۰۲۲	۰,۲۰۹۱۱	۰,۷۲۶۶۷	۰,۶۴۳۳۳	۰,۵۳۳۳۳	
<b>میانگین مقادیر Lrank</b>			۰,۴۲۸۳۵	۰,۲۹۷۱۹	۰,۱۹۵۵۶	۰,۷۳۱۱۱	۰,۶۵۳۷۰	۰,۵۵۰۰۰	
<b>میانگین مقادیر Rrank</b>			۰,۴۴۴۵۴	۰,۳۱۰۵۷	۰,۲۰۴۷۷	۰,۷۰۴۴۴	۰,۶۳۰۷۴	۰,۵۳۲۵۹	
<b>میانگین مقادیر Mrank</b>			۰,۴۵۵۲۱	۰,۳۱۲۶۹	۰,۲۰۷۹۳	۰,۷۲۰۰۰	۰,۶۴۵۱۹	۰,۵۳۶۴۸	

پیشنهادی، بلکه نتایج ترکیب بهینه بدست آمده نیز از نتایج تمامی دیگر روش‌ها بهتر شده است که این نشان‌دهنده تأثیر مثبت استفاده از وزن‌دهی فاصله‌ای در روش پیشنهادی است.

### ۶. نتیجه‌گیری

در این تحقیق نشان داده شد که می‌توان به جای در نظر گرفتن یک مقدار وزنی برای لغات اسناد و پرس‌وجوها، از تمامی وزن‌هایی که توسط تعدادی روش وزن‌دهی مختلف تعیین شده به عنوان وزن‌های پایه بهره برد و به روشی ساده و به صورت آفلاین، یک فاصله (بازه) وزنی برای لغات بدست آورد که بدون افزایش پیچیدگی و زمان اجرا، باعث دقت می‌شود. در این مقاله ۲۴ روش وزن‌دهی پایه در نظر گرفته شد که از نظر مؤلفه‌های بکار رفته در فاکتور فرکانس لغت، فاکتور مجموعه، و فاکتور نرمال‌سازی طول بردار وزنی، تفاوت‌هایی با هم داشتند؛ بر این اساس، آزمایش‌های مختلفی صورت گرفت تا در کنار بررسی سه روش رتبه‌بندی مختلف، تأثیر بکارگیری مؤلفه‌های مختلف فاکتورهای فوق را در وزن‌های پایه بررسی کرده، تا بتوان بهترین زیرمجموعه‌های وزنی از ۲۴ روش وزن‌دهی پایه را جهت ایجاد وزن‌های فاصله‌ای تعیین کرد. در این بررسی‌ها نتایج زیر بدست آمد:

- در مجموعه داده Cranfield، بهترین نتایج مربوط به استفاده از زیرمجموعه‌هایی از روش‌های وزن‌دهی پایه است که نرمال‌سازی طول بردار روی آنها انجام شده و در آنها حتماً از وزن‌های پایه شامل مؤلفه  $n$  استفاده شده، و جهت رتبه‌بندی نهایی از روش Mrank استفاده شده است. در این شرایط، استفاده از وزن‌های بولی و وزن‌های شامل مؤلفه  $p$  در زیرمجموعه‌های وزنی، و سپس عدم استفاده از هیچیک از آنها در زیرمجموعه‌های وزنی، مناسب‌تر بوده‌اند.
- در مجموعه Medline نیز بهترین نتایج مربوط به استفاده از زیرمجموعه‌هایی از روش‌های وزن‌دهی پایه است که نرمال‌سازی طول بردار روی آنها انجام شده، اما در آنها باید حتماً از وزن‌های پایه شامل مؤلفه  $Z$  استفاده شده و جهت رتبه‌بندی نهایی از Lrank استفاده شده باشد. در این شرایط نیز مانند Cranfield، استفاده از وزن‌های بولی و وزن‌های شامل مؤلفه  $p$  در زیرمجموعه‌های وزنی، و سپس عدم استفاده از هیچیک از آنها در زیرمجموعه‌های وزنی، مناسب‌تر بوده‌اند.

در آزمایش‌های انجام شده مشخص شد که استفاده از وزن‌های فاصله‌ای می‌تواند باعث بهبود اثربخشی بازیابی اطلاعات متنی

### ۳.۵ مقایسه با دیگر روش‌ها

جهت مقایسه با دیگر روش‌های وزن‌دهی، مقالات زیادی مورد بررسی قرار گرفتند؛ در برخی از مقالات معیار MAP مدنظر نبوده و در برخی نیز نتایج بر اساس انجام آزمایش بر روی گزیده‌ای از پرس‌وجوها (نه تمامی آنها) ارائه شده است و امکان مقایسه وجود نداشت. در جدول ۱۰، بهترین نتایج روش پیشنهادی در بین ترکیبات مختلف پارامترهای بررسی شده برای تعیین وزن‌های فاصله‌ای لغات، که در واقع مربوط به بهترین ترکیب در هر مجموعه داده هستند، به همراه نتایج ترکیب بهینه بر روی هر دو مجموعه داده  $(A_2, B_1, C_2, Mrank)$  که در بخش ۵-۲-۵ مشخص شد، آورده شده است. این نتایج با بهترین نتایج روش بهترین تطابق  $BM_{25}$  [۷] و [۲۶]، روش وزن‌دهی فرکانس لغت همراه با میانگین رویداد لغت (TF-ATO) در دو حالت عدم استفاده و استفاده از روش متمایزکننده (DA) به منظور حذف وزن‌های کم ارزش از اسناد [۸]، و مدل‌های زبانی سند  $GSPUD_{mc}$  و  $GSPUD_{bs_t}$  در فرآیند Pólya [۳۲]، مقایسه شده است. به دلیل موجود نبودن مقادیر  $P@n$  در این مقالات، امکان مقایسه از منظر این معیار میسر نبود.

جدول ۱۰. مقایسه مقدار MAP روش پیشنهادی با دیگر روش‌ها

		مجموعه داده	
		Cranfield	Medline
روش/مدل			
Wistchel [۷]	$BM_{25}$	۰,۴۱۶	۰,۵۰۹
Cummins [۲۶]	$BM_{25}$	۰,۴۲۲۳	۰,۵۳۴۳
Ibrahim and Landa-Silva [۸]	TF-ATO without DA TF-ATO with DA	۰,۳۵۴۷ ۰,۳۹۹۸	- -
Cummins [۳۲]	$GSPUD_{bs_t}$ $GSPUD_{mc}$	۰,۴۲۷ ۰,۴۳۲	۰,۵۲۳ ۰,۵۳۳
روش پیشنهادی	بهترین نتایج بین روش‌های مختلف ترکیب بهینه پارامترها $(A_2, B_1, C_2, Mrank)$	۰,۴۳۳۲۳	۰,۵۴۵۸۰
		۰,۴۳۳۳۵	۰,۵۴۱۳۴

همانگونه که مشاهده می‌شود، در هر دو مجموعه داده Cranfield و Medline، نه فقط بهترین نتایج بدست آمده در روش

- [۵] P. Switzer, "Vector Images in Document Retrieval," in *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*, Washington, ۱۹۶۴.
- [۶] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information processing & management*, vol. ۲۴, no. ۵, pp. ۵۱۳-۵۲۳, ۱۹۸۸.
- [۷] R. Cummins, *The evolution and analysis of term-weighting schemes in information retrieval*, Galway: National University of Ireland, ۲۰۰۸.
- [۸] O. A. S. Ibrahim and D. Landa-Silva, "Term frequency with average term occurrences for textual information retrieval," *Soft Computing*, vol. ۲۰, no. ۸, pp. ۳۰۴۵-۳۰۶۱, ۲۰۱۶.
- [۹] K. Goslin and M. Hofmann, "A Wikipedia powered state-based approach to automatic search query enhancement," *Information Processing & Management*, vol. ۵۴, no. ۴, pp. ۷۲۶-۷۳۹, ۲۰۱۸.
- [۱۰] K. Chen, Z. Zhang, J. Long and H. Zhang, "Turning from TF-IDF to TF-IGM for term weighting in text classification," *Expert Systems with Applications*, vol. ۶۶, pp. ۲۴۵-۲۶۰, ۲۰۱۶.
- [۱۱] S. Plansangkiet, *New weighting schemes for document ranking and ranked query suggestion*, PhD diss., University of Essex, ۲۰۱۷.
- [۱۲] D. Kandé, R. M. Marone, S. Ndiaye and F. Camara, "A Novel Term Weighting Scheme Model," in *Proceedings of the ۴th International Conference on Frontiers of Educational Technologies (ICFET ۱۸)*, Moscow, ۲۰۱۸.
- [۱۳] T. Dogan and A. K. Uysal, "Improved inverse gravity moment term weighting for
- نسبت به هریک از وزن‌های پایه شود. با بررسی که بر روی ده ترکیب برتر در بین ترکیب‌های مورد آزمایش صورت گرفت، مشخص شد که صرف‌نظر از خصوصیات مجموعه داده‌ها، جهت ایجاد زیرمجموعه‌های وزنی و بدست آوردن بهترین نتایج، استفاده از شرایط زیر مناسب‌تر است:
- استفاده از زیرمجموعه‌هایی از وزن‌های پایه که نرمال‌سازی طول روی بردار وزنی آنها صورت گرفته است.
  - استفاده از زیرمجموعه‌هایی از وزن‌های پایه که فاقد مؤلفه Z در فاکتور TF هستند.
  - استفاده از روش رتبه‌بندی Mrank جهت رتبه‌بندی میزان ارتباط اسناد مختلف با یک پرس‌وجو.
  - در شرایط فوق، هرچند استفاده یا عدم استفاده از وزن‌های بولی و وزن‌های پایه شامل مؤلفه p فرکانس مجموعه، تأثیر زیادی در نتایج ندارد، اما شاید عدم استفاده از وزن‌های بولی مناسب‌تر باشد.
- در تحقیقات آتی می‌توان بر روی دیگر روش‌های وزندهی پایه و دیگر مجموعه داده‌ها، آزمایش‌ها را انجام داده و در کنار آن بر روی دیگر روش‌ها یا فاکتورها در تعیین زیرمجموعه‌های وزنی، تعیین وزن‌های فاصله‌ای و غیره کار کرد.

## مراجع

- [۱] S. Marrara, G. Pasi and M. Viviani, "Aggregation operators in information retrieval," *Fuzzy Sets and Systems*, vol. ۳۲۴, pp. ۳-۱۹, ۲۰۱۷.
- [۲] D. H. Kraft and E. Colvin, *Fuzzy Information Retrieval*, North Carolina: Morgan and Claypool, ۲۰۱۷.
- [۳] D. H. Kraft, E. Colvin, G. Bordogna and G. Pasi, "Fuzzy information retrieval systems: A historical perspective," in *Fifty Years of Fuzzy Logic and its Applications*, Springer, Cham, ۲۰۱۵, pp. ۲۶۷-۲۹۶.
- [۴] H. P. Luhn, "The automatic creation of literature abstracts," *IBM Journal of research and development*, vol. ۲, no. ۲, pp. ۱۵۹-۱۶۵, ۱۹۵۸.



- research and development*, vol. ۱, no. ۴, pp. ۳۰۹-۳۱۷, ۱۹۵۷.
- [۲۳] K. S. Jones, Information retrieval experiment, Newton, MA: Butterworth-Heinemann, ۱۹۸۱.
- [۲۴] S. E. Robertson, "The probability ranking principle in IR," *Journal of documentation*, vol. ۳۳, no. ۴, pp. ۲۹۴-۳۰۴, ۱۹۷۷.
- [۲۵] M. Sanderson and W. B. Croft, "The history of information retrieval research," *Proceedings of the IEEE*, vol. ۱۰۰, no. Special Centennial Issue, pp. ۱۴۴۴-۱۴۵۱, ۲۰۱۲.
- [۲۶] H. F. Witschel, "Global term weights in distributed environments," *Information Processing & Management*, vol. ۴۴, no. ۳, pp. ۱۰۴۹-۱۰۶۱, ۲۰۰۸.
- [۲۷] Y. Gupta, A. Saini and A. K. Saxena, "A new fuzzy logic based ranking function for efficient information retrieval system," *Expert Systems with Applications*, vol. ۴۲, no. ۳, pp. ۱۲۲۳-۱۲۳۴, ۲۰۱۵.
- [۲۸] A. I. Kadhim, "Term Weighting for Feature Extraction on Twitter: A Comparison Between BM<sub>25</sub> and TF-IDF," in *International Conference on Advanced Science and Engineering (ICOASE)*, ۲۰۱۹.
- [۲۹] C. Kamphuis, A. P. de Vries, L. Boytsov and J. Lin, "Which BM<sub>25</sub> Do You Mean? A Large-Scale Reproducibility Study of Scoring Variants," in *European Conference on Information Retrieval*, Cham, ۲۰۲۰.
- [۳۰] J. M. Ponte and W. B. Croft, "A language modeling approach to information retrieval," in *In Proceedings of the ۲۱st annual international ACM SIGIR conference on Research and development in information retrieval*, ۱۹۹۸.
- [۳۱] C. Zhai and J. Lafferty, "A study of smoothing methods for language models applied to information retrieval," *ACM text classification*, "Expert Systems with Applications", vol. ۱۳۰, pp. ۴۵-۵۹, ۲۰۱۹.
- [۱۴] S. Balbi, M. Misuraca and G. Scepti, "Combining different evaluation systems on social media for measuring user satisfaction," *Information Processing & Management*, vol. ۵۴, no. ۴, pp. ۶۷۴-۶۸۵, ۲۰۱۸.
- [۱۵] H. Li, "Learning to rank for information retrieval and natural language processing," *Synthesis Lectures on Human Language Technologies*, vol. ۴, no. ۱, pp. ۱-۱۱۳, ۲۰۱۱.
- [۱۶] S. Gugnani, T. Bihany and R. K. Roul, "A complete survey on web document ranking," *International Journal of Computer Applications (۹۷۵ ۸۸۸۷)*, vol. ICACEA, no. ۲, pp. ۱-۷, ۲۰۱۴.
- [۱۷] A. H. Keyhanipour, M. Piroozmand and K. Badie, "A GP-adaptive web ranking discovery framework based on combinative content and context features," *Journal of Informetrics*, vol. ۳, no. ۱, pp. ۷۸-۸۹, ۲۰۰۹.
- [۱۸] E. Goldberg, "Statistical machine". U.S. Patent ۱۸۳ ۸۳۸ ۹۲۹-۱۹۳۱, ۱۹۳۱.
- [۱۹] J. E. Holmstrom, "Section III. Opening plenary session," in *The Royal Society Scientific Information Conference*, London, U.K., ۱۹۴۸.
- [۲۰] H. F. Mitchell Jr, "The use of the univ AC FAC-tronic system in the library reference field," *American Documentation*, vol. ۴, no. ۱, pp. ۱۶-۱۷, ۱۹۵۳.
- [۲۱] M. Taube, C. D. Gull and I. S. Wachtel, "Unit terms in coordinate indexing," *American Documentation*, vol. ۳, no. ۴, pp. ۲۱۳-۲۱۸, ۱۹۵۲.
- [۲۲] H. P. Luhn, "A statistical approach to mechanized encoding and searching of literary information," *IBM Journal of*

- distribution," *Information Processing & Management*, vol. ۵۴, no. ۳, pp. ۳۷۰-۳۷۹, ۲۰۱۸.
- [۴۱] R. Lakshmi and S. Baskar, "Novel Term Weighting Schemes for Document Representation based on Ranking of Terms and Fuzzy Logic with Semantic Relationship of Terms," *Expert Systems with Applications*, vol. ۱۳۷, pp. ۴۹۳-۵۰۳, ۲۰۱۹.
- [۴۲] F. Carvalho and G. P. Guedes, *TF-IDFC-RF: A Novel Supervised Term Weighting Scheme*, arXiv preprint arXiv:۲۰۰۳.۰۷۱۹۳, ۲۰۲۰.
- [۴۳] W. B. Frakes and R. Baeza-Yates, Eds., *Information retrieval: Data structures & algorithms*, vol. ۳۳۱, Englewood Cliffs, NJ: prentice Hall, ۱۹۹۲.
- [۴۴] G. Bordogna and G. Pasi, "Controlling retrieval through a user-adaptive representation of documents," *International Journal of Approximate Reasoning*, vol. ۱۲, no. ۳-۴, pp. ۳۱۷-۳۳۹, ۱۹۹۵.
- [۴۵] D. H. Kraft, G. Bordogna and G. Pasi, "An extended fuzzy linguistic approach to generalize Boolean information retrieval," *Information Sciences-Applications*, vol. ۲, no. ۳, pp. ۱۱۹-۱۳۴, ۱۹۹۵.
- [۴۶] Y. Y. Yao, "Interval-set algebra for qualitative knowledge representation," in *Proceedings of ICCI'۹۳: ۹th International Conference on Computing and Information*, ۱۹۹۳.
- [۴۷] J. M. Mendel and D. Wu, *Perceptual computing: Aiding people in making subjective judgments*, vol. ۱۳, John Wiley & Sons, ۲۰۱۰.
- [۴۸] J. Han, J. Pei and M. Kamber, *Data mining: concepts and techniques*, Elsevier, ۲۰۱۱.
- [۴۹] A. Turpin and F. Scholer, "User performance versus precision measures for *Transactions on Information Systems (TOIS)*, vol. ۲۲, no. ۲, pp. ۱۷۹-۲۱۴, ۲۰۰۴.
- [۳۲] R. Cummins, *Modelling Word Burstiness in Natural Language: A Generalised Polya Process for Document Language Models in Information Retrieval*, arXiv preprint arXiv:۱۷۰۸.۰۶۰۱۱, ۲۰۱۷.
- [۳۳] R. Cummins, J. H. Paik and Y. Lv, "A Pólya urn document language model for improved information retrieval," *ACM Transactions on Information Systems (TOIS)*, vol. ۳۳, no. ۴, p. ۲۱, ۲۰۱۵.
- [۳۴] G. Salton, *Automatic Information Organization and Retrieval*, New York: McGraw-Hill, ۱۹۶۸.
- [۳۵] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of documentation*, vol. ۲۸, no. ۱, pp. ۱۱-۲۱., ۱۹۷۲.
- [۳۶] G. Salton and C.-S. Yang, "On the specification of term values in automatic indexing," *Journal of documentation*, vol. ۲۹, no. ۴, pp. ۳۵۱-۳۷۲, ۱۹۷۳.
- [۳۷] G. Salton, A. Wong and C.-S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. ۱۸, no. ۱۱, pp. ۶۱۳-۶۲۰, ۱۹۷۵.
- [۳۸] F. S. Al-Anzi, D. AbuZeina and S. Hasan, "Utilizing standard deviation in text classification weighting schemes," *Int J Innov Comput Inf Control*, vol. ۱۳, no. ۴, pp. ۱۳۸۵-۱۳۹۸, ۲۰۱۷.
- [۳۹] J. Beel, S. Langer and B. Gipp, "Tf-iduf: A novel term-weighting scheme for user modeling based on users' personal document collections," in *iConference ۲۰۱۷*, Wuhan, China, ۲۰۱۷.
- [۴۰] L. Bernauer, E. J. Han and S. Y. Sohn, "Term discrimination for text search tasks derived from negative binomial

and technology behind search, ۲ ed.,  
Harlow: England: Pearson Education Ltd.,  
۲۰۱۱.

simple search tasks," in *In Proceedings of the ۲۹th annual international ACM SIGIR conference on Research and development in information retrieval*, ۲۰۰۶.

[۵۰] R. Baeza-Yates and B. Ribeiro-Neto,  
Modern Information Retrieval: the concepts



## بهبود کارائی و دقت یافتن یال‌های پرتکرار در خلاصه سازی gMatrix از جریان گراف

مسعود کاظمی\* سیدحسین خواسته\*\* حمیدرضا رخصتی\*\*\*

\*دانشجوی کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی

\*\*استادیار، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی

\*\*\*کارشناسی ارشد هوش مصنوعی و رباتیک، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی

تاریخ پذیرش: ۱۳۹۹/۰۹/۲۶

تاریخ دریافت: ۱۳۹۹/۰۳/۰۵

نوع مقاله: پژوهشی

### چکیده

در سیستم‌های کاربردی، گراف‌ها با دامنه وسیعی از راس‌ها وجود دارند و یال‌ها به سرعت زیادی در قالب جریان گراف تولید می‌شوند. یکی از مسائل موجود در جریان‌های گراف سنگین که به صورت لحظه‌ای وارد می‌شوند پیدا کردن زیرگراف‌های پرتکرار است. خلاصه‌های جریان مبتنی بر طرح، مانند count-min، اطلاعات گره‌های پرتکرار را با دقت قابل قبولی نگهداری می‌کنند ولی ساختار گراف اصلی را از دست می‌دهند. از بین این روش‌ها، gMatrix ساختاری می‌باشد که مشخصات گراف اصلی را نیز حفظ می‌کند. این روش از توابع درهم‌ساز مختلف، برای ذخیره‌ی خلاصه‌ی جریان گراف استفاده کرده و به کمک این توابع و معکوس آنها، زیرگراف‌های پرتکرار را به دست می‌آورد. به دلیل داشتن حجم کمتر از جریان اصلی، gMatrix معمولاً به پرس و جوها با دقت بالایی پاسخ نمی‌دهد. همچنین این روش از مشکل مرتبه‌ی زمانی بالا در پاسخ به پرس و جوها هم رنج می‌برد. در این مقاله روش جدیدی ارائه شده است که به ازای هزینه‌ی کم حافظه‌ی مصرفی، زمان پاسخگویی به پرس و جو زیرگراف پرتکرار را به صورت چشم‌گیری کاهش می‌دهد. همچنین الگوریتم ارائه شده با افزایش استقلال بین توابع در هم سازی با استفاده از روش شباهت برداری گساین، احتمال برخورد عناصر در هم سازی شده را کاهش می‌دهد. نتایج آزمایشات تجربی که به زبان C++ پیاده‌سازی شده است و بر روی داده‌های شبکه اجتماعی فرندستر اجرا شده است، نشان می‌دهد که روش پیشنهادی برای یافتن زیرگراف‌های پرتکرار پیچیدگی زمانی و دقت یافتن این زیر گراف‌ها را بهبود می‌بخشد.

**واژگان کلیدی:** جریان گراف، خلاصه‌سازی مبتنی بر طرح، gMatrix، توابع درهم‌ساز، شباهت برداری گساین

## ۱. مقدمه

محدودیت‌هایی که در حفظ و نگهداری جریان گراف و همچنین محدودیت‌هایی که در پردازش تمام ساختار یک گراف سنگین (که به صورت جریان وارد می‌شود) داریم، عملاً بسیاری از پرس و جوهای موجود به شکل کنونی خود قابل استفاده نیستند. برای پردازش سریع جریان‌های گراف روش‌های زیادی در انواع مختلف از جمله پنجره لغزان<sup>۷</sup> و طرح گراف<sup>۸</sup> معرفی شده‌اند. روش‌های مبتنی بر طرح یک خلاصه از جریان گراف را نگهداری می‌کنند و به محض ورود یال‌های جدید در گذر زمان آن خلاصه را بروزرسانی می‌کنند. از جمله خلاصه‌های جریان مبتنی بر طرح می‌توان به count-min اشاره کرد که می‌تواند اطلاعات تکرار گره‌ها با تکرار بالا را با دقت قابل قبولی نگهداری کند. [۱] اما این خلاصه مبتنی بر طرح ساختار گراف اصلی را از دست می‌دهد. مگر اینکه ما اطلاعات راس شروع و پایان هر یال را ذخیره کنیم که این کار بسیار هزینه بر و نشدنی می‌باشد. به عنوان مثال اکثر روش‌های موجود می‌توانند راس‌ها و یال‌های تکرار بالا را تشخیص دهند، اما آنها قادر به جواب پرس و جوهای پیچیده‌تر مانند پیدا کردن مسیریایی که از یال‌های با تکرار بالا تشکیل شده است نخواهند بود. در طرف دیگر، خلاصه مبتنی بر طرح gMatrix که یک طرح سه بعدی می‌باشد، جریان‌های گراف سنگین را در حالی که اطلاعاتی از رفتارهای ساختاری گراف مذکور نگهداری می‌کند، به صورت بلادرنگ خلاصه می‌کند. در نتیجه gMatrix قادر به پاسخگویی پرس و جوهای ساختاری مهم از قبیل دسترسی پذیری از طریق یال‌های پرتکرار نیز خواهد بود. [۲]

## ۲. کارهای پیشین

### ۲\_۱ روش طرح count-min

در الگوریتم count-min با استفاده از یک روش درهم سازی، فرکانس حدودی تعداد زیادی از آیتم‌های مجزا در یک جریان داده محاسبه می‌شوند. در این الگوریتم به تعداد  $W =$

در علوم کامپیوتر، جریان داده<sup>۱</sup> به یک دنباله از داده‌ها که با گذشت زمان در دسترس خواهند بود گفته می‌شود. می‌توان یک جریان را آیتم‌های موجود روی یک تسمه نقاله در نظر گرفت که به صورت تک تک پردازش می‌شوند و نه به صورت جمعی. پردازش جریان‌ها به صورت کاملاً متفاوتی از پردازش داده‌ها به صورت دسته‌ای صورت می‌گیرد. از آنجایی که ممکن است جریان‌ها نامحدود باشند، توابع معمولی نمی‌توانند روی جریان‌ها به صورت کلی استفاده شوند. به صورت رسمی جریان‌ها را داده<sup>۲</sup> نمی‌دانند چرا که داده محدود می‌باشد، حال آنکه جریان‌ها بالقوه نامحدود هستند و شبه‌داده<sup>۳</sup> نامیده می‌شوند. علاوه بر آن معمولاً جریان‌ها حجیم هستند. به صورتی که نگهداری و ذخیره تمام اجزای آن امکان‌پذیر نمی‌باشد. حتی در صورت امکان ذخیره این حجم داده، به دلیل پویایی این داده‌ها و تغییر آن در هر ورود داده جدید در گذر زمان، عملاً اجرای دوباره و دوباره یک عملیات بر روی کل داده‌هایی که تاکنون دریافت شده است امکان‌پذیر نمی‌باشد و بسیار کار پرهزینه‌ای می‌باشد. یکی از انواع جریان‌ها، جریان گراف<sup>۴</sup> می‌باشد. جریان گراف در قالب جریانی از رخدادهای صورت گرفته روی یک گراف تعریف می‌شود. در واقع جریان گراف به جای مجموعه‌ای ثابتی از رئوس و یال‌ها و مقادیر آنها، رخدادهایی در گذر زمان می‌دهد که اطلاعاتی در مورد راس‌ها و یال‌ها ارائه می‌دهد. در نتیجه جریان گراف می‌تواند شامل یال‌های یک گراف که به صورت تدریجی وارد می‌شوند باشد. دامنه گسترده‌ای از داده‌ها مانند شبکه‌های حمل و نقل، شبکه‌های آی‌پی<sup>۵</sup> و شبکه‌های اجتماعی به این تقسیم بندی تعلق دارند.

بر روی گراف‌ها پرس و جوهای کاربردی مختلفی تعریف شده است که در حوزه‌های مختلف اطلاعاتی از ساختار کلی گراف و همچنین اطلاعات جزئی مربوط به گره یا یال مشخصی را می‌دهند. حال آنکه در جریان گراف به خاطر وجود

<sup>۵</sup> IP

<sup>۶</sup> Query

<sup>۷</sup> Sliding Window

<sup>۸</sup> Graph Sketch

<sup>۱</sup> Data Stream

<sup>۲</sup> Data

<sup>۳</sup> Codata

<sup>۴</sup> Graph Stream

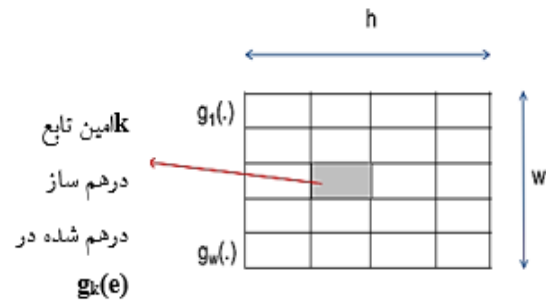
مجموع فرکانس‌های تمام آیتم‌ها تا این لحظه در نظر گرفته می‌شود.

## ۲-۲ روش gMatrix

در این مقاله فرض شده است که داده‌های گراف  $G = (V, E)$  به صورت جریان داده‌ای یالی دریافت می‌شود. هر راس  $i \in V$  از مجموعه  $N = \{1, 2, \dots, n\}$  برداشته می‌شود و هر یال  $(i, j) \in E$  یک یال جهت دار است. جریان داده ورودی گراف مورد نظر شامل عضوهایی به صورت  $(i_1, j_1, f_{i_1 j_1}), (i_2, j_2, f_{i_2 j_2}), \dots, (i_t, j_t, f_{i_t j_t}), \dots$  می‌باشد. اینجا  $(i_t, j_t, f_{i_t j_t})$  نشان دهنده دریافت یال  $t$ -ام جریان داده با تکرار  $f_{i_t j_t}$  می‌باشد. در بسیاری از کاربردها، مقدار  $f_{i_t j_t}$  به صورت طبیعی برابر یک است، ولی ما یک تکرار عمومی غیر از یک را برای کلیت بخشیدن به نتایج در نظر می‌گیریم. برای مثال در کاربردهای مخابراتی، عدد تکرار  $f_{i_t j_t}$  می‌تواند نشان دهنده تعداد ثانیه‌هایی باشد که شخص  $i_t$  با شخص  $j_t$  با شروع در زمان  $t$  در حال مکالمه بوده است. یک یال می‌تواند چندین بار در یک جریان داده ظاهر شود. زمانی که یک پرس و جو را مطرح می‌کنیم، مثلاً پیدا کردن تعداد تکرار یال  $(i, j)$ ، به دنبال جمع تکرارهای یال  $(i, j)$  تا کنون هستیم. درحالی که روش‌های طرح محور می‌توانند با کمی تغییر برای پرس و جوهای زمانی نیز استفاده شوند، آنها را در این قسمت بررسی نمی‌کنیم. همچنین توجه کنید که نتایج ما با فرض بر اینکه یال  $(i, j)$  همیشه به صورت مرتب شده است به سادگی می‌توانند به گراف‌های غیر جهت دار تعمیم داده شوند.

پرس و جوهای تعریف شده در روش gMatrix، تکرار یال (محاسبه تکرار یک یال  $(i, j)$ )، یال‌های پرتکرار (پیدا کردن همه یال‌هایی که تکرار آنها بیشتر از یک حد آستانه داده شده  $F$  می‌باشد)، مجموع تکرار راس (محاسبه مجموع تکرار تمام یال‌های ورودی و خروجی برای یک راس  $i$ )، راس‌های پرتکرار (پیدا کردن همه راس‌های که مجموع تکرار آنها (روی همه یال‌های ورودی و خروجی از آنها) از یک حد آستانه داده

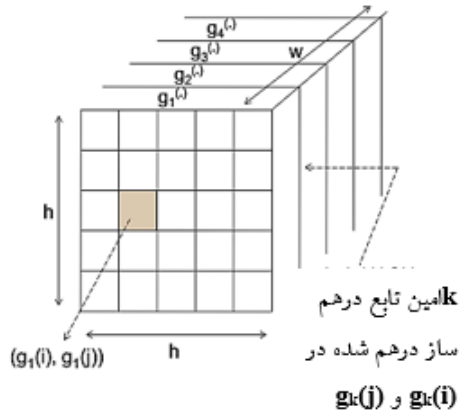
$\ln(\frac{1}{\delta})$  عدد تابع درهم سازی که دو به دو مستقل می‌باشند استفاده می‌شود و هر کدام به یک عدد در بازه  $h = [0, \frac{e}{\delta}]$  به صورت یکنواخت نگاشت می‌شود. که پایه لگاریتم طبیعی است. همچنین داده ساختار مورد نظر از یک آرایه دو بعدی با ابعاد  $h * w$  تشکیل شده است. همان‌طور که در شکل ۱) مشخص است، هر تابع درهم سازی با یکی از  $w$  آرایه یک بعدی  $h$  خانه‌ای مطابقت دارد.



شکل ۱. طرح count-min برای جریان داده [1]

در یک جریان داده یک بعدی که عناصر آن دامنه‌های وسیعی دارند وقتی المنت جدیدی وارد می‌شود تمام  $w$  تابع درهم سازی روی آن اجرا می‌شود تا به یک عدد در بازه  $[0, h - 1]$  نگاشت شوند. سپس مقادیر این  $w$  خانه یکی اضافه می‌شوند. به جهت محاسبه تعداد یک آیتم، مجموعه  $w$  خانه مطابق با آن که هر یک را  $w$  تابع درهم سازی می‌دهند بازیابی شده و مقدار کمینه میان این خانه‌ها پیدا می‌شوند. اگر فرض کنیم  $Ct$  تعداد واقعی این آیتم باشد، مقدار تخمین زده شده در کمترین حالت با  $Ct$  برابر است. از آنجایی که شمارش غیر منفی مد نظر است و به دلیل امکان وجود تداخل در خانه‌های درهم سازی ممکن است تخمین زیاد انجام پذیرد. همچنین در یک جریان داده با  $L$  داده ورودی، تخمین برآورده شده برای خانه  $Ct$  در بیشترین مقدار با احتمال  $(1 - \delta)$  برابر با  $(Ct + \epsilon * L)$  خواهد بود. به همین شکل در یک رخداد که آیتم‌ها فرکانس‌های مربوط به خود را دارند، باید تعداد هر آیتم را با فرکانس مربوطه آن افزایش داد. در این حالت  $L$  را

شده  $F$  بیشتر می‌باشد)، مجموع تکرار زیرگراف (محاسبه مجموع تکرار همه یال‌های یک زیرگراف که روی زیرمجموعه‌ای از راس‌های داده شده  $S$  می‌باشد) و قابل دستیابی بودن (یافتن اینکه آیا یک راس مبدأ<sup>۹</sup> به راس دیگری به عنوان مقصد<sup>۱۰</sup> از طریق یال‌های با تکرار بیشتر از آستانه داده شده  $F$  قابل دسترسی می‌باشد و یا خیر).



شکل ۲. طرح gMatrix برای جریان‌های گراف [۲]

## ۲\_۳\_۱ به روزرسانی خلاصه ی gMatrix

در دو پرس و جوی اول فقط تکرار یال‌ها مطرح می‌باشد که با طرح count-min هم قابل پاسخ گویی می‌باشند. به همین صورت پرس و جوی سوم و چهارم فقط به تکرار راس‌ها نیازمند می‌باشد که مجدداً با یک طرح بر روی راس‌ها قابل پاسخ گویی می‌باشد. اما دو پرس و جوی آخر به دنبال یافتن رفتار ساختاری جریان گراف مورد نظر بر مبنای تکرار یال‌های آن می‌باشند. که این پرس و جوها با استفاده از طرح‌های کنونی موجود قابل پاسخ گویی نیستند و باید اطلاعات شروع و پایان همه یال‌ها نیز ذخیره شود که بسیار سنگین می‌باشد. ساختار gMatrix طراحی شده یک ماتریس سه بعدی می‌باشد. به جای آرایه یک راه حل کاملاً متفاوت، gMatrix بر پایه اصول ساختار count-min طراحی شده است.

## ۲\_۳\_۲ ساختار gMatrix

برای بروز رسانی gMatrix، کافی است در شروع تمام خانه‌ها را با ۰ مقدار دهی اولیه کنیم و برای هر یال ورودی  $(i, j)$  با فرکانس  $f(i, j)$  یال درهم سازی شده متناظر  $(g_k(i), g_k(j))$  را برای تمام مقادیر  $k$  از ۱ تا  $w$  محاسبه میکنیم و سپس فرکانس هر یک از این خانه را با مقدار  $f(i, j)$  افزایش دهیم. میزان فضای مورد استفاده در gMatrix از مرتبه  $O(h * h * w)$  می‌باشد و همچنین پیچیدگی زمانی برای بروز رسانی یک یال ورودی از جریان برابر با  $O(2 * w)$  می‌باشد [۳, ۴].

## ۲\_۳\_۲ تابع درهم ساز و عکس درهم سازی

به جهت کارا بودن gMatrix، باید توابع درهم ساز دو به دو مستقل باشند. به علاوه، به جهت پردازش پرس و جوی

<sup>۱۰</sup> Destination

<sup>۹</sup> Source



۱. در گام اول، gMatrix اسکن می‌شود و تمام یال‌های درهم سازی شده که تکرار آنها از حد آستانه  $F$  بیشتر است برای توابع درهم ساز مختلف پیدا می‌شوند.

۲. در گام دوم، برای هر یال  $(p, q)$  پر تکرار تحت تابع درهم ساز  $g_k$ ، مجموعه تمام یال‌های گراف اصلی پر تکرار ممکن با استفاده از تکنیک عکس نگاشت درهم سازی پیدا شده که با  $( )$  نمایش داده می‌شود.

(۲)

$$E_k = \{(i, j) : i \in g_k^{-1}(p), j \in g_k^{-1}(q), V(g_k(p), g_k(q), k) \geq F\}$$

در نهایت، تصادم<sup>۱۳</sup> مجموعه یال‌های مختلف  $E_k$  برای همه  $k \in (1, w)$  یال‌های پر تکرار را به ما می‌دهد. توجه داشته باشید که گام دوم پرهزینه ترین گام محاسباتی این روش می‌باشد. به همین منظور، تکنیک‌های بهینه سازی برای پرس و جو یال‌های پر تکرار نیز ارائه می‌شوند [۸، ۹، ۱۰].

### ۳. روش پیشنهادی

در این مقاله مشکلات روش gMatrix از دو جهت بهبود یافته است. در رویکرد اول سعی شده است تا زمان اجرای یافتن یال‌های پر تکرار کاهش یابد و در رویکرد دوم به استفاده از توابع درهم‌سازی مستقل سعی شده است تا نرخ مثبت کاذب در یافتن این یال‌ها بهبود یابد.

### ۳\_۱ بهینه‌سازی زمان اجرا به کمک داده ساختارها

الگوریتم موجود در روش gMatrix برای محاسبه‌ی اشتراک مجموعه‌ها از روش مناسبی استفاده نمی‌کند. برای سریع‌تر کردن مرحله یافتن اشتراک دو مجموعه از داده ساختار درخت قرمز-سیاه استفاده می‌کنیم که با توجه به شناخته بودن این درخت از توضیح نحوه کارکرد آن خودداری می‌کنیم. در این قسمت مراحل استفاده از داده‌ساختار قرمز-سیاه و نحوه بهینه‌سازی پرس و جو یال‌های پر تکرار به صورت زیر

یال‌های پر تکرار و همچنین پرس و جوهای مبتنی بر ساختار که در قسمت‌های بالا ذکر شده اند، به محاسبه عکس نگاشت تابع درهم ساز نیاز است. اگر فرض کنیم تابع  $g$  یک تابع در هم ساز می‌باشد، ما معکوس این تابع را با  $g^{-1}$  به صورت (۱) تعریف می‌کنیم.

$$g^{-1}(p) = \{i : g(i) = p\} \quad (۱)$$

طول عکس نگاشت درهم سازی فقط  $[P/h]$  باشد و از طریق تعمیم الگوریتم اقلیدس به صورت کارآیی با پیچیدگی زمانی  $O(\frac{P}{h} * \log(P))$  محاسبه می‌شود. [۵]

### ۲\_۴ محاسبه پرس و جوی تکرار یال

برای محاسبه تکرار یک یال  $(i, j)$  با استفاده از gMatrix، مقادیر  $V(g_k(i), g_k(j), k)$  برای  $w$  مقدار مختلف  $k$  محاسبه می‌شوند. کمینه این مقادیر به عنوان تخمین تکرار  $Q(i, j)$  از یال  $(i, j)$  در نظر گرفته می‌شود. این تخمین را با نماد  $Q'(i, j)$  نشان می‌دهیم. پیچیدگی زمانی مورد نیاز برای تخمین تکرار یک یال  $O(2w)$  است که  $w$  تعداد توابع درهم ساز می‌باشد [۶، ۷].

### ۲\_۵ محاسبه پرس و جوی یال‌های پر تکرار

یال‌های پر تکرار تمام یال‌های با تکرار بیشتر از یکی آستانه مشخص شده توسط کاربر می‌باشند. از آنجایی که gMatrix یک طرح احتمالی است، نمی‌توان تمام یال‌های پر تکرار را به صورت قطعی پیدا کرد. در نتیجه، یک روش ارائه شده است که هیچ منفی کاذب<sup>۱۱</sup>ی را بازیابی نمی‌کند، اما یک یال ممکن است نمونه مثبت کاذب<sup>۱۲</sup> باشد. متعاقباً احتمال مثبت کاذب بودن یک یال نیز محاسبه می‌شود. این روش دو گام دارد:

<sup>۱۳</sup> Collision<sup>۱۱</sup> False Negative<sup>۱۲</sup> False Positive

۴. همه یال‌های ممکن از مجموعه گره‌های  $IntersectedQ1$  به مجموعه گره‌های  $IntersectedQ2$  را می‌سازیم.

۵. حذف کردن منفی کاذب‌ها از یال‌های تولید شده و ارائه جواب نهایی

### ۳\_۲ تحلیل پیچیدگی زمانی

پیچیدگی پیمایش کردن خلاصه gMatrix برابر  $O(h^2 w)$  می‌باشد. فرض می‌کنیم که  $k$ -مین تابع درهم ساز از gMatrix که  $1 \leq k \leq w$  در مجموع تعداد  $C_k$  خانه با تعداد تکرار بیشتر یا مساوی آستانه  $F$  دارد [۱۱]. آنگاه بر اساس طراحی ما برای مدولار بودن تابع‌های درهم ساز، هر کدام از این خانه‌ها به  $[P/h]$  راس‌های مبدا و  $[P/h]$  راس‌های مقصد مربوط می‌شوند که نیاز به زمان  $O(2C_k [P/h] \log P)$  برای یافتن تمامی این مجموعه راس‌ها داریم. نهایتاً، همانگونه که بحث شد اشتراک را روی مجموعه راس‌ها می‌گیریم. بنابراین در مجموع پیچیدگی زمانی پرس و جوی یال‌های پرتکرار برابر است با رابطه (۳).

$$O\left(h^2 w + 2 \left[\frac{P}{h}\right] \log P \sum_{k=1}^w C_k\right) \quad (3)$$

با توجه به مرحله‌ای که در بخش قبل توضیح داده شد، برای هر تابع درهم‌ساز شماره  $k$ ، تعداد  $h^2$  خانه بررسی می‌شود که اگر فرض کنیم به تعداد  $C_k$  خانه مقدار بیشتر از حد آستانه  $F$  داشته باشند و از آنجایی که برای هر گره تعداد  $\left[\frac{P}{h}\right]$  مقدار عکس نگاشت بدست می‌آید، در نتیجه در بیشترین حالت (همه مقادیر عکس نگاشت‌ها یکتا باشند) به تعداد  $C_k \left[\frac{P}{h}\right]$  مقدار در هر درخت قرار می‌گیرد که مرتبه زمانی قرار دادن هر مقدار در درخت  $O(\log(C_k \left[\frac{P}{h}\right]))$  خواهد بود. پس برای هر تابع درهم‌ساز مرتبه زمانی کلی  $O\left(h^2 + 2C_k \left[\frac{P}{h}\right] \log(C_k \left[\frac{P}{h}\right])\right)$  می‌باشد.

توضیح داده می‌شود. توجه داشته باشید که درخت قرمز-سیاه فقط مقادیر یکتا را نگهداری می‌کند.

۱. دو درخت قرمز-سیاه  $IntersectedQ1$  و  $IntersectedQ2$  را تعریف کنید.

۲. برای تابع درهم ساز اول، برای هر خانه  $(p, q)$  از gMatrix که مقدار آن از آستانه  $F$  بیشتر است مراحل الف و ب را انجام دهید:

أ. عکس نگاشت  $p$  را بدست آورید و همه مقادیر را در  $IntersectedQ1$  قرار دهید.

ب. عکس نگاشت  $q$  را بدست آورید و همه مقادیر را در  $IntersectedQ2$  قرار دهید.

۳. برای تابع درهم ساز ۲ تا  $w$  مراحل الف تا ث را انجام دهید:

أ. دو درخت قرمز-سیاه  $Q1$  و  $Q2$  را تعریف کنید.

ب. برای هر خانه  $(p, q)$  از gMatrix که مقدار آن از آستانه  $F$  بیشتر است مراحل ۱ و ۲ را انجام دهید:

۱. عکس نگاشت  $p$  را بدست

آور و همه مقادیر را در  $Q1$  قرار دهید.

۲. عکس نگاشت  $q$  را بدست

آور و همه مقادیر را در  $Q2$  قرار دهید.

ت. عناصری از  $IntersectedQ1$  که در  $Q1$  نیستند را حذف کنید.

ث. عناصری از  $IntersectedQ2$  که در  $Q2$  نیستند را حذف کنید.

کلی الگوریتم ارائه شده برای یافتن توابع درهم‌ساز مستقل پرداخته می‌شود و نحوه‌ی تبدیل مسئله به یک مسئله در فضای برداری ارائه می‌شود.

### ۳\_۳\_۱ شباهت توابع درهم ساز با معیار فاصله

تابع درهم سازی مانند  $h_i(x)$  را به صورت زیر تعریف می‌کنیم:

(۵)

$$x_h^i = h_i(x)$$

در رابطه (۵)  $x$  را ورودی تابع درهم سازی و  $h_i(x)$  را تابع در هم سازی  $i$ -ام و  $x_h^i$  را خروجی تابع در هم سازی  $i$ -ام می‌نامیم. برای ارزیابی دو تابع درهم سازی  $h_i(x)$  و  $h_j(x)$  در اولین گام، برداری به طول دلخواه  $K \geq 1000$  از اعداد طبیعی در فاصله‌ی  $[1, P-1]$  در نظر می‌گیریم و خروجی تابع‌های درهم‌ساز را برای این بردار محاسبه کرده و بردارهای  $x_{h_i}$  و  $x_{h_j}$  می‌نامیم. با یافتن فاصله برای دو بردار بدست آمده، میزان شباهت این دو تابع درهم سازی به صورت زیر محاسبه می‌شود:

(۶)

$$Sim(h_i, h_j) = d_7(x_{h_i}, x_{h_j})$$

در رابطه شماره‌ی (۶) تابع محاسبه‌ی فاصله یعنی  $d_7(x_{h_i}, x_{h_j})$  یکی از سه روش فاصله‌ی اقلیدسی، فاصله‌ی کساین و یا اطلاعات متقابل است. رابطه‌های ریاضی هر یک از این معیارهای فاصله در بخش قبلی توضیح داده شد. در ادامه به بررسی نحوه‌ی استفاده از هر یک از معیارهای فاصله‌ی برای محاسبه‌ی شباهت دو تابع درهم‌ساز یعنی مقدار عددی تابع  $Sim(h_i, h_j)$  پرداخته می‌شود.

### ۳\_۳\_۲ شباهت با فاصله‌ی اقلیدسی

فاصله‌ی اقلیدسی به عنوان ساده‌ترین روش محاسبه‌ی فاصله یا شباهت در مسائل فضاهای برداری در نظر گرفته می‌شود.

همچنین برای توابع درهم‌ساز شماره ۲ تا  $w$ ، علاوه بر این، یک بار روی درخت عکس نگاشت‌های تابع درهم‌ساز اول حرکت می‌کنیم و تمام مقادیر را در درخت تابع درهم‌ساز کنونی جستجو می‌کنیم که در مجموع مرتبه زمانی  $O(wC_k \left[ \frac{P}{h} \right] \log(C_k \left[ \frac{P}{h} \right]))$  خواهد بود [۱۲]. در نتیجه به صورت کلی مرتبه زمانی الگوریتم پیشنهادی در رابطه (۴) آمده است. همانطور که در رابطه‌های (۳) و (۴) مشخص است پیچیدگی زمانی الگوریتم از ضرب تعداد ضریب  $C_k$  به جمع آنها تقلیل یافته است.

(۴)

$$O(wh^2 + \sum_{k=1}^w C_k \left[ \frac{P}{h} \right] \log(C_k \left[ \frac{P}{h} \right]) + \sum_{k=2}^w C_k \left[ \frac{P}{h} \right] \log(C_k \left[ \frac{P}{h} \right]))$$

### ۳\_۳\_۳ بهبود دقت پرس و جو به کمک توابع درهم‌ساز مستقل

وجود وابستگی بین توابع درهم‌ساز باعث می‌شود که در پاسخ به پرس و جوهای مختلف، و در مرحله‌ی ابهام‌زدایی برای یافتن پاسخ پرس و جو، تصادم بسیار زیادی به وجود آید. این مشکل باعث می‌شود که دقت پاسخگویی به پرس و جوها در روش gMatirx کاهش یابد. برای مرتفع کردن این مشکل باید از توابع درهم‌سازی استفاده کرد که وابستگی کمتری به یکدیگر داشته باشند تا بتوانند روی جدول‌های درهم‌سازی مختلف تصادم کمتری ایجاد کنند. برای این منظور از روش بردارهای شباهت یا معیارهای فاصله بین توابع درهم‌ساز استفاده می‌شود. هرچقدر که میزان فاصله‌ی برداری بین توابع درهم‌ساز بیشتر باشد، باعث می‌شود که تصادم کمتری در ابهام‌زدایی توابع درهم‌ساز ایجاد شود [۱۳، ۱۴]. در این پژوهش از روش‌های بردار فاصله‌ی کساین، اطلاعات متقابل و فاصله‌ی اقلیدسی استفاده شده است. در بخش بعدی به روش

که  $n(x)$  تعداد رخدادهای مقدار  $x$  در  $x_{h_i}$  است. این احتمال به همین ترتیب برای هر  $y \in \mathcal{Y}$ ،  $y_{h_j}$ ، تعریف می‌شود:

(۱۰)

$$p(y \in x_{h_j}) = \frac{n(y)}{K}$$

که در آن  $n(y)$  تعداد رخدادهای مقدار  $y$  در  $x_{h_j}$  است.

احتمال توام  $p(x, y)$  نیز به این ترتیب تعریف می‌شود که در این رابطه  $n(x, y)$  تعداد رخدادهایی به صورت  $(x, y)$  است که  $x$  عضوی از بردار  $x_{h_i}$  و  $y$  عضوی از بردار  $x_{h_j}$  باشد [۱۶].

(۱۱)

$$p(x, y) = \frac{n(x, y)}{K * K}$$

### ۳-۳-۴ شباهت با فاصله‌ی کساین

با یافتن فاصله‌ی کساین برای دو بردار بدست آمده، میزان شباهت این دو تابع در هم سازی به صورت زیر محاسبه می‌شود:

(۱۲)

$$\begin{aligned} Sim(h_i, h_j) &= d_C(x_{h_i}, x_{h_j}) = \cos(\theta) \\ &= \frac{x_{h_i} \cdot x_{h_j}}{\|x_{h_i}\| * \|x_{h_j}\|} \end{aligned}$$

که در رابطه‌ی بالا  $\cdot$  نماد ضرب داخلی بین دو بردار  $x_{h_i}$  و  $x_{h_j}$  است که در بخش فاصله‌ی کساین تعریف شد. همچنین  $\theta$  هم زاویه‌ی بین این دو بردار است. در نهایت معیار شباهت دو بردار  $x_{h_i}$  و  $x_{h_j}$  به کمک فاصله‌ی کساین یعنی به صورت زیر تعریف می‌شود [۱۷, ۱۸]:

در این مسئله به سادگی می‌توان میزان شباهت دو تابع درهم‌ساز را به کمک فاصله‌ی دو بردار درهم‌سازی شده به کمک این دو تابع یعنی  $x_{h_i}$  و  $x_{h_j}$  به صورت رابطه (۷) در نظر گرفت:

(۷)

$$\begin{aligned} Sim(h_i, h_j) &= d_E(x_{h_i}, x_{h_j}) \\ &= \sqrt{\sum_{k=1}^K (x_{h_i}^k - x_{h_j}^k)^2} \end{aligned}$$

که در این رابطه،  $K$  همان طول بردار در نظر گرفته شده است و  $x_{h_i}^k$  درایه‌ی  $k$ -ام از نمایش برداری خروجی درهم‌سازی شده‌ی درایه‌های بردار ابتدایی به کمک تابع درهم‌سازی  $i$ -ام یعنی  $h_i$  است [۱۵].

### ۳-۳-۳ شباهت با اطلاعات متقابل

با توجه به تعریف اطلاعات متقابل در بخش قبلی، می‌توان شباهت دو بردار  $x_{h_i}$  و  $x_{h_j}$  را از طریق اطلاعات متقابل به دست آورد.

(۸)

$$\begin{aligned} Sim(h_i, h_j) &= d_I(x_{h_i}, x_{h_j}) \\ &= \sum_{y \in x_{h_j}} \sum_{x \in x_{h_i}} p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right) \end{aligned}$$

که در (۸) احتمال رخداد یک  $x \in x_{h_i}$  به صورت زیر تعریف می‌شود:

(۹)

$$p(x \in x_{h_i}) = \frac{n(x)}{K}$$

(۱۶)

$$Dependence(M_i) = \left\{ \frac{\sum_{j=1}^w \sum_{k=j+1}^w sim(h_{ij}, h_{ik})}{\frac{w * (w - 1)}{2}} \right\}$$

که در (۱۶) تابع  $sim(h_{ij}, h_{ik})$  همان تابع محاسبه‌ی فاصله‌ی بین دو تابع درهم‌ساز است که در

(۷)، (۸) و (۹) تعریف شد. همچنین  $w$  تعداد توابع درهم‌ساز مجزای موجود در روش gMatrix است. در انتها با استفاده از مجموعه‌ی توابعی که کمترین وابستگی را نسبت به یکدیگر دارند می‌توانیم بهترین مجموعه‌ی توابع درهم‌ساز را برای gMatrix به صورت رابطه (۱۷) انتخاب کنیم:

(۱۷)

$$BestHashSet(M) = \arg \min_i \{Dependence(M_i)\}$$

با این روش می‌توانیم با اطمینان بیشتری در مورد کاهش برخورد‌های اتفاقی در توابع درهم‌سازی در هنگام ساخت gMatrix پیش رویم. در این راستا، الگوریتم ارائه شده برای یافتن توابع درهم‌ساز مستقل از یکدیگر دارای بخش‌های اصلی زیر است:

۱. مجموعه‌ی  $M$  شامل  $m$  مجموعه‌ی  $w$  تایی از توابع درهم‌سازی مجزا بساز
۲. برای هر مجموعه از توابع درهم‌سازی ساخته شده مثل  $M_i$  که  $i = 1 \dots m$  انجام بده:
  - ا. میزان عدم وابستگی توابع درهم‌ساز عضو را از (۱۶۶) محاسبه کن.
۳. بهترین مجموعه از  $M$  را به کمک (۱۷) به عنوان  $M_b$  انتخاب کن.
۴. از مجموعه‌ی توابع درهم‌ساز  $M_b$  در ساخت نهایی gMatrix استفاده کن.

(۱۳)

$$d_C(x_{h_i}, x_{h_j}) = \cos(\theta) = \frac{\sum_{k=1}^K x_{h_i}^k * x_{h_j}^k}{\sqrt{\sum_{k=1}^K x_{h_i}^k} \sqrt{\sum_{k=1}^K x_{h_j}^k}}$$

### ۳\_۴ روش پیشنهادی برای تعیین توابع درهم‌ساز مستقل

برای صحبت با اطمینان بیشتر در مورد کاهش برخورد‌های اتفاقی، می‌توان تعدادی مجموعه از توابع درهم‌ساز تعریف کرد که هر مجموعه شامل  $w$  تابع درهم‌ساز مجزا است. این مجموعه را  $M$  می‌نامیم که شامل  $m$  مجموعه تابع درهم‌ساز است که هر مجموعه شامل  $w$  تابع درهم‌ساز است. این مجموعه به عنوان ورودی الگوریتم به صورت تصادفی از اعداد اول موجود در رابطه‌های توابع درهم‌ساز تولید می‌شود.

(۱۴)

$M = \{\{h_{i1}, \dots, h_{iw}\}, \dots, \{h_{i1}, \dots, h_{iw}\}, \dots, \{h_{m1}, \dots, h_{mw}\}\}$   
 سپس با استفاده از معیارهای فاصله‌ی تعریف شده، برای هر یک از مجموعه‌های موجود در  $M$  فاصله‌ی بین دونه دوی توابع درهم‌ساز موجود در آن را به کمک روش رابطه (۱۹-۳) محاسبه کرده و میانگین این فاصله‌ها را به عنوان فاصله‌ی نهایی برای آن مجموعه یا میزان عدم وابستگی توابع درهم‌ساز آن مجموعه در نظر می‌گیریم. پس اگر فرض کنیم که مجموعه‌ی توابع درهم‌سازی کاندید  $i$ -ام در مجموعه‌ی اصلی  $M$  به صورت رابطه (۱۵) باشد:

(۱۵)

$$M_i = \{h_{i1}, \dots, h_{iw}\}$$

معیار عدم وابستگی بین توابع درهم‌ساز موجود در مجموعه‌ی  $M_i$  برابر با میانگین شباهت دونه‌دوی توابع درهم‌ساز است و به صورت رابطه (۱۶) تعریف می‌شود:

#### ۴. ارزیابی روش پیشنهادی

کلیده‌ی آزمایش‌های انجام شده در این بخش با کمک یک کامپیوتر شخصی با حافظه‌ی اصلی ۸ گیگابایت، پردازنده‌ی Intel Core i۵ ۲,۷ GHz، روی سیستم عامل مک، انجام شده‌اند. در این پژوهش الگوریتم‌ها توسط زبان برنامه‌نویسی ++C پیاده‌سازی شده‌اند که توضیح کلی روند پیاده‌سازی را در این قسمت خواهیم داشت. توجه داشته باشید که برای هر مرحله برنامه‌ای جداگانه نوشته شده است که شامل برنامه‌ای برای الحاق فرکانس به مجموعه داده فرندستر، برنامه‌ای برای ساخت خلاصه گراف gMatrix از روی جریان داده، برنامه‌ای برای ساخت پایگاه داده‌ی کلی شامل همه اطلاعات مجموعه داده، سه برنامه برای تست و مقایسه سه پرس و جوی مختلف روی داده واقعی و خلاصه gMatrix، می‌باشند.

#### ۴-۱ مجموعه‌های داده‌ی مورد آزمایش

در این پژوهش از مجموعه داده شبکه اجتماعی فرندستر<sup>۱۴</sup> که از سایت<sup>۱۵</sup> دانشگاه استنفورد موجود می‌باشد، به دو شکل متفاوت استفاده شده است و در ادامه به جزئیات این دو شکل متفاوت داده خواهیم پرداخت. فرندستر یک شبکه بازی آنلاین است. قبل از راه اندازی مجدد به عنوان یک وب سایت بازی، فرندستر یک سایت شبکه اجتماعی بود که در آن کاربران می‌توانستند یال دوستی با هم تشکیل دهند. همچنین شبکه اجتماعی فرندستر به کاربران اجازه می‌داد گروه تشکیل دهند که اعضای دیگر نیز می‌توانستند به آن ملحق شوند. ما چنین گروه‌های تعریف شده توسط کاربر را به عنوان جوامع حقیقی می‌شناسیم. برای شبکه اجتماعی، ما زیرگراف القا شده از گره‌هایی که به حداقل یک اجتماع متعلق هستند یا به گره‌های دیگر که آنها متعلق به حداقل یک جامعه

هستند را برمی‌داریم. این داده‌ها توسط پروژه بایگانی وب<sup>۱۶</sup> ارائه شده است، جایی که نمودار کامل موجود است. اطلاعات تکمیلی این مجموعه داده در جدول (۱) آورده شده است [۲۱، ۲۰].

جدول ۱. اطلاعات و خصوصیات مجموعه داده فرندستر

۶۵۶۰۸۳۶۶	تعداد گره‌ها
۱۸۰۶۰۶۷۱۳۵	تعداد یال‌ها
۶۵۶۰۸۳۶۶	تعداد گره‌ها در بزرگترین مولفه ضعیفا همبند <sup>۱۷</sup>
۱۸۰۶۰۶۷۱۳۵	تعداد یال‌ها در بزرگترین مولفه ضعیفا همبند
۶۵۶۰۸۳۶۶	تعداد گره‌ها در بزرگترین مولفه قویا همبند <sup>۱۸</sup>
۱۸۰۶۰۶۷۱۳۵	تعداد یال‌ها در بزرگترین مولفه قویا همبند
۰,۱۶۲۳	ضریب خوشه‌بندی <sup>۱۹</sup> میانگین
۴۱۷۳۷۲۴۱۴۲	تعداد مثلث‌ها
۰,۰۰۵۸۵۹	کسر مثلث‌های بسته
۳۲	قطر (طولانی‌ترین مسیر کوتاه <sup>۲۰</sup> )
۵,۸	قطر موثر <sup>۲۱</sup> ۹۰-درصدی

باید توجه داشت که این مجموعه داده بدون جهت می‌باشد و همچنین فاقد فرکانس برای هر یال می‌باشد. در واقع این مجموعه داده فقط شامل گره شروع و گره پایان هر یال می‌باشد. در نتیجه برای نسبت دادن فرکانس به این مجموعه داده از توزیع زیفان استفاده شده که استفاده از این توزیع در

<sup>۱۹</sup> Clustering

<sup>۲۰</sup> Longest shortest path

<sup>۲۱</sup> ۹۰-percentile

<sup>۱۴</sup> Friendster

<sup>۱۵</sup> Snap.stanford.edu

<sup>۱۶</sup> Web Archive Project

<sup>۱۷</sup> Weakly connected component (WCC)

<sup>۱۸</sup> Strongly connected component (SCC)

بخواند، سپس با استفاده از رابطه کلی توزیع زیفان که در رابطه (۱۸) آورده شده است فرکانس مربوطه را محاسبه کند و به همراه گره شروع و گره پایان یال در یک فایل مجزا به عنوان خروجی چاپ نماید. اگر  $N$  را تعداد عناصر در نظر بگیریم و همچنین  $k$  را مرتبه هر عنصر و  $S$  مقدار ضریب انحراف باشد، رابطه کلی توزیع زیفان به صورت زیر می‌باشد:

$$(18) \quad f(k; s, N) = \frac{1/k^s}{\sum_{n=1}^N (1/n^s)}$$

ما برای هر عنصر مقدار محاسبه شده از توزیع زیفان را در  $10^1$  ضرب می‌کنیم و مقدار سقف نتیجه را به عنوان فرکانس آن عنصر در نظر می‌گیریم. بدین صورت از آنجایی که جمع همه مقادیر تولید شده توسط توزیع زیفان برابر با مقدار ۱ خواهند بود، جمع تمام فرکانس‌های عناصر تولید شده توسط ما نیز  $10^1 \approx$  خواهد بود. توجه داشته باشید که مرتبه عناصر به جهت سادگی، طبق ترتیب ورودی آنها در نظر گرفته شده است. همچنین با توجه به مراحل بالا، دو مجموعه داده مختلف از روی مجموعه داده فرندستر به ازای ضرایب انحراف  $S = 1.0$  و  $S = 1.4$  ساخته شد.

#### ۴\_۳ انتخاب بهترین بردار شباهت

قبل از انجام عملیات خلاصه‌سازی جریان به کمک روش gMatrix باید یک مجموعه تابع درهم‌ساز مناسب انتخاب گردد [۲۴]. در این مرحله با انتخاب یک بردار شباهت از روش‌های مطرح‌شده‌ی فاصله‌ی اقلیدسی، شباهت گساین و اطلاعات متقابل یک روش انتخاب می‌شود و بعد از انجام الگوریتم پیشنهادی مجموعه‌ی توابع درهم‌ساز مناسب مشخص می‌گردد. سپس به کمک این مجموعه مراحل بعدی انجام می‌شود که همان ساخت خلاصه‌ی گراف و پاسخ به پرس و جوها و ارزیابی دقت پاسخ‌گویی است. برای انتخاب از

ادبیات جریان داده [۲۲، ۱] به یک عمل استاندارد تبدیل شده است. همچنین در تولید فرکانس برای یال‌های مجموعه داده از دو ضریب انحراف<sup>۲۲</sup> مختلف  $S = 1.0$  و  $S = 1.4$  استفاده شده است.

شایان ذکر است که دلیل استفاده از مجموعه داده فرندستر با علم به عدم وجود فرکانس در این مجموعه این بوده است که با الحاق مصنوعی فرکانس به این مجموعه داده ما می‌توانیم نتایج خود را برای ضریب انحراف‌های مختلف تست و بررسی کنیم. از طرف دیگر، همانطور که در گزارش شده است، بیشتر مجموعه داده‌های واقعی ضریب انحراف مشترکی در بازه استفاده شده توسط ما را دارند. [۱۶، ۱۷] در جدول (ابعاد مجموعه داده استفاده شده نمایش داده شده است. که حجم جریان<sup>۲۳</sup> نمایانگر حجم جریان گراف با تکرار یال‌ها می‌باشد.

جدول ۲. ابعاد و خصوصیات مجموعه داده فرندستر

حجم جریان	بیشترین فرکانس یال	فرکانس تجمعی جریان	یال‌ها	گره‌ها
۸۰ گیگابایت	۴.۴۳ $\cdot 10^9 (S = 1.0)$ $1.3 \cdot 10^9 (S = 1.4)$	$L = 10^1$	۱,۸ میلیارد	۶۶ میلیون

#### ۴\_۲ افزودن فرکانس به مجموعه داده

در ابتدای امر به جهت آماده‌سازی مجموعه داده فرندستر باید برای هر یال یک فرکانس با استفاده از توزیع زیفان تولید کنیم [۲۳]. به همین منظور برنامه‌ای نوشته شده است تا گره شروع و گره پایان هر یال را از روی فایل مجموعه داده فرندستر

<sup>۲۳</sup> Stream Size

<sup>۲۲</sup> Skew

بین سه معیار فاصله‌ی معرفی شده، بعد از انجام چندین اجرا و مقایسه‌ی نتایج، بهترین روش به عنوان معیار معرفی می‌گردد. برای ساخت بردار شباهت در همه‌ی آزمایش‌ها متغیر  $K = 1000$  در نظر گرفته شده است.

#### ۴-۴ ارزیابی عملکرد روش پیشنهادی

برای ارزیابی یک سیستم خلاصه سازی جریان گراف در مراجع تعدادی پرس و جو تعریف می‌شود. برای ارزیابی عملکرد روش gMatrix، تعداد ۶ پرس و جو و جوی مختلف تعریف شده است. [۲] از این ۶ پرس و جو، دو پرس و جو یافتن فرکانس یک یال و یافتن یال‌های پرتکرار به عنوان پایه‌ای برای سایر پرس و جوها استفاده شده است. بنابراین علاوه بر پرس و جو اصلی مورد تاکید در این پژوهش (یافتن زیرگراف‌های پرتکرار) برای ارزیابی عملکرد روش پیشنهادی این دو پرس و جو هم در نظر گرفته شده است.

برای ارزیابی دو آزمایش مستقل در نظر گرفته شده است. همان‌طور که در بخش ساخت داده گفته شد، در این دو آزمایش متغیر ضریب انحراف توزیع زیفان متفاوت است. همچنین سایر متغیرهای ثابت در این دو آزمایش در

جدول ۳. مقدار متغیرهای مشترک در آزمایش‌ها

نماد	$s$	$K$	$m$	$w$	$h$
آزمایش ۱	۱,۰	H	۱۰۰۰	۱۰	۱۰۰۰
آزمایش ۲	۱,۴	H	۱۰۰۰	۱۰	۱۰۰۰

#### ۴-۴-۱ پرس و جو یافتن فرکانس یک یال

برای این پرس و جو یک برنامه نوشته شده است که ابتدا خلاصه gMatrix، عدد اول  $P$  و توابع درهم ساز را که روی فایل ذخیره شده‌اند می‌خواند و بازسازی می‌کند. سپس ۱۰۰۰۰ یال پرتکرار گراف اصلی را پیدا می‌کند و تعداد ۱۰۰ یال از آنها را به صورت تصادفی انتخاب و فرکانس تخمینی آنها را محاسبه می‌کند. در ادامه با پیدا کردن فرکانس واقعی همین یال‌ها از پایگاه داده ساخته شده، واریانس خطای gMatrix را در برآورد کردن فرکانس این یال‌ها را طبق (۱۹) محاسبه می‌کند. در این رابطه تعداد نمونه‌ها یا  $n$  برابر

جدول) آمده است. به علت اینکه طول جدول درهم‌سازی در gMatrix برابر با  $h=1000$  است، طول بردار سنجش شباهت توابع درهم‌ساز هم برابر با این مقدار قرار می‌گیرد. زیرا خروجی توابع درهم‌ساز در بازه‌ی  $[0, h]$  قرار دارد بنابراین انتخاب  $K = h$  می‌تواند باعث پوشش بیشتر فضای مسئله گردد.



gMatrix و همچنین پرس و جوی آن تغییری ایجاد نشده است، این انتظار را داشتیم که جواب این پرس و جو برای همه روش‌های انتخاب تابع درهم ساز روی یک مجموعه داده یکسان، جوابی مشابه و با دقت یکسانی به دست بدهد.

از نظر ریاضی علت این است که gMatrix برای یافتن فرکانس یک یال، مقدار درون W جدول درهم‌ساز را بررسی می‌کند و از بین آنها کمترین مقدار را به عنوان فرکانس یال درخواست شده در نظر می‌گیرد. با این تعریف انتخاب توابع درهم‌ساز مستقل یا وابسته تغییر در نتایج این پرس و جو نخواهد داشت. در نهایت از بین مقادیر درون W جدول درهم‌ساز در نهایت مقدار یکی از جدول‌ها به عنوان فرکانس یال انتخاب می‌شود. که این مقدار به مقدار خانه‌های جدول‌های دیگر ارتباطی ندارد. به همین علت در این پرس و جو با هر مجموعه‌ی توابع درهم‌ساز ورودی دقت نهایی یافتن فرکانس یک یال تغییر چندانی نخواهد کرد. همچنین از نظر زمان پاسخ‌گویی به این پرس و جو هم به علت عدم تغییر در الگوریتم یافتن خانه‌ی موجود در تابع درهم‌ساز، تغییری بین روش gMatrix و روش پیشنهادی این پژوهش وجود ندارد و زمان پاسخ‌گویی به این پرس و جو در هر دو روش دقیقاً به یک اندازه است.

#### ۴\_۴\_۲ پرس و جوی یافتن یال‌های پر تکرار

برای این پرس و جو نیز یک برنامه نوشته شده است که با توجه به فرکانس تجمعی همه یال‌های جریان گراف که  $L = 10^6$  می‌باشد، پرس و جو را اجرا می‌کند. برای ارزیابی دقت این پرس و جو از کسر یک ده هزارم از فرکانس تجمعی یعنی  $10^6$  می‌باشد، به عنوان  $F$  در نظر گرفته می‌شود. سپس با این مقدار  $F$  تمام یال‌های پر تکرار یال‌ها را با استفاده از خلاصه gMatrix محاسبه می‌شود و در نهایت با بررسی فرکانس هر کدام از پایگاه داده‌ی کلی، مقدار نرخ مثبت کاذب در جواب این پرس و جو را به عنوان پارامتر اصلی ارزیابی دقت در نظر گرفته می‌شود. همچنین زمان پاسخ‌گویی به

تعداد ۱۰۰ یال،  $\mu$  میانگین فرکانس‌ها و  $x_i$  فرکانس بدست آمده توسط هر روش است. مقدار خطای این واریانس برای دو آزمایش مطرح‌شده و به کمک هر یک از معیارهای فاصله‌ی مطرح شده، در جدول (آمده است).

(۱۹)

$$\text{Variance of Error} = \frac{1}{100} \sum_{i=1}^{100} (x_i - \mu)^2$$

جدول ۴. واریانس خطای یافتن فرکانس یک یال

روش	آزمایش ۱	آزمایش ۲
فاصله‌ی اقلیدسی	۱۲۹۵	۸۶۰
اطلاعات مشترک	۱۲۸۶	۸۶۳
شباهت کساین	۱۲۸۷	۸۶۰
gMatrix	۱۲۹۰	۸۵۵

لازم به ذکر است که برای یافتن فرکانس تخمینی هر یال، همه یال‌های درهم ساز مربوطه در هر W تابع درهم ساز را بدست می‌آوریم و بین مقادیر ذخیره شده در این خانه‌ها کوچکترین مقدار را انتخاب می‌کنیم که نزدیک‌ترین مقدار به مقدار واقعی یال در بین آنها می‌باشد. همچنین از آنجایی که در همه روش‌ها یک الگوریتم یکسان استفاده شده است و صرفاً خلاصه ساخته شده متفاوت می‌باشد، در نتیجه از لحاظ مرتبه زمانی همه یکسان هستند و نیازی به مقایسه نمی‌باشد [۲۵].

#### ۴\_۴\_۱ مقایسه و ارزیابی نتایج

همانطور که در

جدول (نمایش داده شده است، واریانس خطا در هر آزمایش، فارغ از نحوه انتخاب توابع درهم‌ساز، مقادیر یکسانی دارد. با توجه به اینکه در نحوه ذخیره‌سازی فرکانس یال‌ها در

کردیم و مجدداً نتایج را یادداشت کردیم. مقایسه این دو آزمایش با روش gMatrix از نظر مقدار نرخ مثبت کاذب (دقت) در

جدول ۲ آورده شده است.

جدول ۲. نرخ مثبت کاذب در یافتن یال‌های پرتکرار برای هر آزمایش با آستانه‌ی  $F = 10^6$

آزمایش / روش	آزمایش ۱: $S = 1.0$	آزمایش ۲: $S = 1.4$
فاصله اقلیدسی	۰,۰۱۳	۰,۰۰۹
اطلاعات مشترک	۰,۰۱۱	۰,۰۰۳
شباهت کساین	۰,۰۰۷	۰,۰۰۱
gMatrix	۰,۰۱۲	۰,۰۰۴

#### ۴\_۲\_۳ مقایسه و تحلیل نرخ مثبت کاذب

در

جدول ۲) که نمایانگر نرخ مثبت کاذب می‌باشد، دیده می‌شود که در هر دو آزمایش (دو مجموعه داده متفاوت) میزان نرخ مثبت کاذب در روش شباهت کساین از روش gMatrix کمتر شده است. بدین ترتیب در آزمایش اول به میزان ۴۱ درصد و در آزمایش دوم به میزان ۷۵ درصد افزایش یافته است. همچنین روش فاصله‌ی اقلیدسی نمی‌تواند بهبود خاصی در نتایج ایجاد کند و حتی نتایج آن از خود gMatrix هم

پرس و جویا به همراه حافظه‌ی مصرفی روش پیشنهادی در مقایسه با روش gMatrix به عنوان پارامترهای دیگری برای ارزیابی کارایی محاسبه می‌شود. اما با توجه به اثبات مرتبه‌ی زمانی برای روش ارائه شده برای ارزیابی کارایی از نظر مرتبه‌ی زمانی، باید از معیارهای مختلف برای آستانه‌ی  $F$  استفاده کرد زیرا مرتبه‌ی زمانی به این معیار وابستگی دارد. برای این پرس و جو دو جدول (۵) و (۶) در ادامه آمده است که به ترتیب نمایانگر دقت الگوریتم (نرخ مثبت کاذب) و دیگری نمایانگر زمان اجرای هر کدام از روش‌ها می‌باشد که به تفکیک هر روش و آزمایش برای مجموعه داده به نمایش گذاشته شده‌اند.

#### ۴\_۲\_۴ نرخ مثبت کاذب

آزمایش‌هایی برای یافتن یال‌های پرتکرار که تکرار بالاتر از یک آستانه مشخصی داشتند انجام داده ایم. مقدار این آستانه تکرار را با درصدی از تکرار جریان بیان می‌کنیم، زیرا خطای تخمین تکرار یک یال نسبت به جمع تکرار جریان حساس می‌باشد. gMatrix تضمین می‌کند که مجموعه یال‌های گزارش شده مجموعه‌ای شامل یال‌های واقعی می‌باشد. نرخ مثبت کاذب را برای پرس و جویا یال‌های پرتکرار به این صورت بیان می‌شود:

نرخ مثبت کاذب = تعداد یال‌هایی که به اشتباه به عنوان یال‌های پرتکرار گزارش شده‌اند تقسیم بر تعداد درست یال‌های پرتکرار، طبق (۰):

(۲۰)

False Positive Rate

$$= \frac{\text{\#Number of edges Incorrectly reported as Heavy Hitters}}{\text{\#Number of reported edges as Heavy Hitters}}$$

گرچه، این نرخ مثبت کاذب کاهش داده شده به قیمت حافظه و زمان اجرای بالاتر به دست آمده است. این پرس و جو را یکبار با استفاده از الگوریتم پیشین یافتن یال‌های پرتکرار و روی gMatrix ساخته شده بدون بهبود توابع درهم‌ساز اجرا و نتایج را یادداشت کردیم. بار دیگر از الگوریتم پیشنهادی بهبود یافته برای یافتن یال‌های پرتکرار و روی gMatrix ساخته شده به همراه بهبود یافته توابع درهم‌ساز استفاده

• ارایه ی نتایج قابل قبول در مقایسه با دیگر روش‌های یافتن فاصله در هنگام کار با بردارهای تنک.

#### ۴\_۳\_۱ زمان پاسخ به پرس و جو

در

جدول ( برای هر یک از روش‌ها میانگین زمان پاسخ به پرس و جوهای یافتن یال‌های پرتکرار نمایش داده می‌شود. در جدول قابل مشاهده است زمان پاسخ‌گویی به پرس و جو ارتباطی به معیار فاصله‌ی شباهت استفاده شده در الگوریتم ندارد اما به علت بهینه‌سازی استفاده شده به کمک داده ساختار درخت قرمز-سیاه این زمان در روش پیشنهادی پژوهش نسبت به روش gMatrix بسیار کمتر است. این نتایج با توجه به بحث مرتبه‌ی زمانی مطرح شده قابل پیش بینی هستند و محاسبات را تایید می‌کند.

جدول ۶. زمان پاسخ به پرس و جوی یال‌های پرتکرار برای هر آزمایش بر حسب ثانیه

مقدار آستانه	$F = 10^6$	$F = 10^7$	$F = 10^8$	$S = 1.0$	$S = 1.4$	$S = 1.4$
آزمایش/رو ش	$S = 1.0$	$S = 1.4$	$S = 1.0$	$S = 1.4$	$S = 1.0$	$S = 1.4$
بهینه‌سازی به کمک داده ساختار درخت قرمز- سیاه	۷۹۴	۳۹۱	۳۱	۳۰	۱	۴
gMatrix	۱۰۵۶ ۲	۷۰۷۱	۶۱۲	۵۴۹	۹	۲۸

#### ۴\_۳\_۲ مقایسه و تحلیل زمان پاسخ به پرس و جو

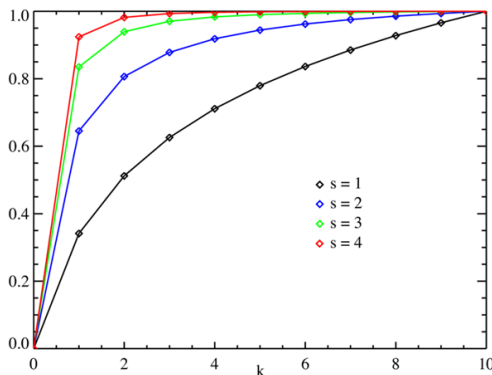
بدتر است. در مورد روش اطلاعات مشترک هم نتایج به دست آمده به رغم داشتن برتری جزئی نسبت به روش gMatrix به این روش بسیار نزدیک هستند. علت این است که روش اطلاعات مشترک برای این مسئله در یافتن احتمال رخداد یک واقعه ضعیف عمل می‌کند. زیرا اطلاعات کافی برای یافتن میزان احتمال یک رخداد مثل  $p(x = x_i)$  را در دست نداریم. در نهایت از بین این سه بردار فاصله بهترین نتایج را روش فاصله‌ی کساین ارائه می‌دهد. دلیل برتری فاصله‌ی کساین نسبت به سایر روش‌ها این است که این فاصله سعی می‌کند بردارها را از نظر زاویه‌ی بین آنها شباهت سنجی کند و به مقادیر هر یک از ابعاد یک بردار وابسته نیست [۲۶].

در این پژوهش، نشان دادیم که استفاده از فاصله‌ی کساین بر روی توابع درهم سازی برای یافتن میزان شباهت و سازگاری آن‌ها از کارایی بالایی برخوردار می‌باشد، زیرا:

- در توابع درهم سازی مورد استفاده در این مطالعه، بزرگی مقدار عدد خروجی تابع درهم سازی در هنگامی که این اعداد بزرگتر از حد بالایی مانند  $F$  باشند، از اهمیت یکسانی برخوردار می‌باشند. به طور مثال،  $x_1$  و  $x_2$  دو خروجی از تابع درهم سازی ۱ و ۲ می‌باشند. با فرض اینکه  $x_1 > F$  و  $x_2 > F$  و  $x_1/x_2 = 10^3$  در گام‌های بعدی الگوریتم ارائه شده بر پایه ی توابع درهم سازی، اهمیت  $x_1$  و  $x_2$  یکسان در نظر گرفته خواهد شد (و نه  $10^3$  برابر). حال می‌توان به اصلی ترین ویژگی فاصله ی کساین اشاره نمود. اهمیت زاویه‌ی ساخته شده در بردارها و نه اندازه ی آن‌ها. از این ویژگی می‌توان به عنوان اثباتی برای کارایی فاصله ی کساین در مساله‌ی ارائه شده یاد کرد.
- پیچیدگی زمانی مناسب فاصله ی کساین با توجه به محدودیت‌های زمانی تعریف شده در مساله

جدول) مربوط به نتایج زمان اجرای پرس و جو ها، زمان پاسخ به پرس و جو برای مقادیر تعیین شده برای  $S$  در توزیع زیفان، با بالا رفتن ضریب انحراف، تعداد یال‌ها با فرکانس بالا کمتر می‌شود اما فرکانس خود این یال‌ها بسیار بیشتر خواهد بود. اما با ضریب انحراف کمتر، تعداد یال‌های پرتکرار بیشتری خواهیم داشت. تعداد این یال‌ها برای مقدار آستانه‌ی  $F$  و برای مقدار ضریب انحراف توزیع زیفان در جدول) آمده است.

توجه اینکه زمان پاسخ به پرس و جو یافتن یال‌های پرتکرار در هر دو روش برای مقدار آستانه‌ی فرکانس  $F = 10^6$  برای مقدار ضریب انحراف  $S = 1.0$  از مقدار ضریب انحراف  $S = 1.4$  بیشتر است همین امر است. زیرا همان‌طور که شکل) نشان می‌دهد، وقتی ضریب انحراف بیشتر باشد یال‌های پرتکرار کم ولی وزن‌های خیلی بالا دارند. اما وقتی ضریب انحراف کمتر باشد تعداد یال‌های پرتکرار بیشتر است اما یال‌ها وزن‌های بسیار بالایی ندارند.



شکل ۳. احتمال فرکانس یال‌ها برای ضریب انحراف‌های مختلف در توزیع زیفان

به عبارت دیگر وقتی ضریب انحراف کمتر باشد توزیع زیفان میزان وزن یال کلی  $L$  را در بین تعداد زیادتری از یال‌ها توزیع می‌کند. به همین علت وقتی فرکانس برابر با  $F = 10^6$  است، برای ضریب انحراف  $S = 1.4$  تعداد یال‌های پرتکرار کمتری داریم. اما وقتی ضریب انحراف  $S = 1.0$  باشد تعداد یال‌های پرتکرار برای فرکانس  $F = 10^6$  بیشتر است. در نهایت هنگامی میزان فرکانس به فرکانس تجمعی جریان گراف در

روش پیشنهادی همان‌طور که در جداول نتایج مشاهده می‌گردد روش پیشنهادی به کمک داده ساختار درخت قرمز-سیاه زمان پاسخ به پرس و جوها نسبت به روش gMatrix بهبود یافته است. به علت مشابه نبودن شرایط اجرای آزمایش با سیستم اجرای آزمایشات در این پژوهش، همه‌ی آزمایشات مقایسه‌ی مرتبه‌ی زمانی و زمان پاسخ‌گویی به پرس و جوها، برای روش gMatrix هم روی سیستم مشابه این پژوهش انجام شده است. [۲] بنابراین زمان اجرای گزارش شده برای روش gMatrix، برای اجرا آن روی محیط شبیه‌سازی مشابه با روش پیشنهادی و با شرایط جریان گراف کاملاً مشابه است. با توجه به اثبات مرتبه‌ی زمانی روش پیشنهادی در بخش ۰ انتظار می‌رود که این بهبود تقریباً یک حالت لگاریتمی باشد. زیرا مرتبه‌ی زمانی اصلی روش gMatrix برابر است با ضرب ضرایب  $C_k$  اما در روش پیشنهادی ما این مرتبه برابر با جمع ضرایب  $C_k$  است. بهبود نمایی در زمان پاسخ به پرس و جو یافتن یال‌های پرتکرار از مقایسه‌ی نتایج گرفته شده در

جدول) قابل مشاهده است.

جدول ۷. تعداد یال‌های پرتکرار با ضریب انحراف برای مقدار آستانه‌ی مختلف

مقدار آستانه	$F = 10^6$	$F = 10^7$	$F = 10^8$	$S = 1.0$	$S = 1.4$
ضریب انحراف	$S = 1.0$	$S = 1.4$	$S = 1.0$	$S = 1.4$	$S = 1.4$
توزیع زیفان					
تعداد یال‌های پرتکرار	۸۶۳۱۶	۴۳۷۵	۴۵	۶۱	۴

در

اشتراک بین نمایه‌های معکوس شده با مرتبه‌ی زمانی لگاریتمی استفاده شد.

۳- نتایج آزمایش‌های تجربی روی داده‌ها مسئله‌ی جریان گراف بهبود زمان پاسخ به پرس و جوهای مطرح شده را از حالت خطی به حالت لگاریتمی نشان می‌دهد. اما در ازای این کاهش مرتبه‌ی زمانی، به علت استفاده از داده ساختار ها، مرتبه‌ی حافظه‌ی مورد استفاده در روش پیشنهادی دو برابر روش gMatrix شده است.

۴- مشکل دوم روش gMatrix دقت پایین در پاسخ به پرس و جو ی یافتن زیرگراف‌های پر تکرار و یال‌های پر تکرار است. با بررسی دقیق‌تر روش gMatrix دریافت می‌شود که در حین یافتن معکوس نمایه‌های توابع درهم‌سازی هرچه توابع درهم‌ساز مستقل‌تر از یکدیگر باشند باعث می‌شود که تفاوت الگوی استفاده از تمام خانه‌های جدول درهم‌ساز بیشتر شود و برعکس هرچه توابع وابسته‌تر باشند احتمال اینکه تابع‌های درهم‌ساز از یک الگوی مشابه برای نگاشت راس‌ها به خانه‌های خود استفاده کنند بیشتر است و این مسئله باعث کاهش دقت روش gMatrix شده است. زیرا این روش برای بهبود روش طرح count-min به جای w جدول درهم‌سازی یک بعدی، w جدول درهم‌سازی دو بعدی را به طور همزمان استفاده می‌کند.

۵- برای حل مشکل توابع درهم‌ساز وابسته در روش gMatrix از روشی استفاده شد تا بتوان شباهت دو تابع درهم‌ساز را به دست آورد. برای این منظور با نگاشت فضای مسئله به یک فضای برداری، برای محاسبه‌ی عدم شباهت دو تابع درهم‌ساز، از معیارهای فاصله‌ی اقلیدسی، اطلاعات مشترک و شباهت کساین استفاده شد. با استفاده از این معیارهای فاصله می‌توان با در نظر گرفتن خروجی هر یک از توابع درهم‌ساز در یک فضای برداری میزان وابستگی الگوی آنها را به دست آورد و توابع درهم‌سازی با وابستگی کمتر را به عنوان ورودی روش gMatrix استفاده کرد.

توزیع زیغان یعنی  $L = 10^{10}$  نزدیک‌تر می‌شود، یعنی  $F = 10^7$  و یا  $F = 10^8$ ، تعداد یال‌های پر تکرار برای ضریب انحراف کمتر یعنی  $S = 10$  کمتر می‌شود. اما برای ضریب انحراف بیشتر یعنی  $S = 10^4$  به علت بالابودن وزن یال‌های ایجاد شده، تعداد یال‌های پر تکرار در این حالت بیشتر خواهد بود.

## ۵. نتیجه‌گیری و راهکارهای آتی

هدف از این پژوهش بهبود کارایی یافتن یال‌های پر تکرار در مسئله‌ی جریان گراف بود. برای حل این مسئله یک روش مبتنی بر توابع درهم‌ساز به نام gMatrix مطرح شده است که دارای مشکلاتی از جمله پیچیدگی زمانی بالا و دقت پایین در پاسخ به پرس و جوهای مطرح در مسئله‌ی جریان گراف است. با بررسی دقیق مراحل و الگوریتم‌های ارائه شده در روش gMatrix، دلایل این دو مشکل و راه حل‌های آنها به شرح زیر می‌باشد:

۱- روش gMatrix برای یافتن یال‌های پر تکرار دارای یک مرحله‌ی ابهام زدایی است که در این مرحله معکوس مقدار توابع درهم‌ساز برای یک نمایه از هر تابع در نظر گرفته می‌شود و در انتها این معکوس‌ها با یکدیگر اشتراک داده می‌شوند تا یال‌های نهایی پر تکرار بدست آیند. برای بخش اشتراک‌گیری بین معکوس نمایه‌های توابع درهم‌ساز، خود روش gMatrix راه‌حلی با مرتبه‌ی زمانی خطی ارائه کرده است. و این مرتبه‌ی زمانی به علت زیاد بودن تعداد نمایه‌های معکوس شده، می‌تواند باعث کندی فرایند پاسخ به پرس و جوها شود.

۲- راه حل ارائه شده برای حل این مشکل استفاده از داده‌ساختارها برای کاهش مرتبه‌ی زمانی مرحله‌ی ابهام زدایی در روش gMatrix برای پاسخ به پرس و جوهای زیر گراف‌های پر تکرار و یال‌های پر تکرار، بود. در این پژوهش از داده ساختار درخت قرمز-سیاه برای یافتن

مثلا روش‌های پوشش بیتی<sup>۲۴</sup>، یا روش‌های مجموعه‌ها مجزا<sup>۲۵</sup>

- استفاده از معیارهای فاصله یا شباهت بردار دیگری مثل جاکارد<sup>۲۶</sup>، همینگ<sup>۲۷</sup> و ...
- استفاده از توابع درهم‌ساز جدول بندی ترکیبی<sup>۲۸</sup> [۲۷]، زیرا در این جدول‌ها گستردگی پخش نقاط در تابع درهم ساز بیشتر است. اما برای یافتن معکوس چنین توابعی باید چاره‌ای اندیشیده شود.

## مراجع

- [۱] G. Cormode and S. Muthukrishnan, "An Improved Data-Stream Summary: The Count-min Sketch and its Applications", *J. of Algorithms*, ۵۵(۱), ۲۰۰۵.
- [۲] Khan, Arijit, and Charu Aggarwal. "Query-friendly compression of graph streams", *Proceedings of the ۲۰۱۶ IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE Press, ۲۰۱۶.
- [۳] M.R.Henzinger, P.Raghavan, and S.Rajagopalan, "Computing on data streams", *External memory algorithms*, pages ۱۰۷-۱۱۸, ۱۹۹۹.
- [۴] J.Feigenbaum, S.Kannan, A.McGregor, S.Suri, and J.Zhang, "On graph problems in a semi-streaming model", *Theoretical Computer Science*, ۳۴۸(۲-۳):۲۰۷-۲۱۶, ۲۰۰۵.
- [۵] A. Gilbert, S. Guha, P. Indyk, Y. Kotidis, S. Muthukrishnan, M. Strauss, "Fast, small-space algorithms for approximate histogram maintenance", in: *Proceedings of the ۳۴th*

۶- نتایج آزمایش‌های تجربی روی داده‌ی مسئله نشان می‌دهند که استفاده از معیارهای فاصله و شباهت مطرح شده می‌تواند باعث افزایش دقت یا کاهش نرخ مثبت کاذب در پاسخ به پرس و جوهای زیر گراف‌های پرتکرار و یال‌های پرتکرار شود. همچنین نتایج نشان می‌دهد که از بین سه روش مطرح شده برای سنجش وابستگی تابع‌های درهم‌ساز روش شباهت کساین بهترین نتایج را در بر دارد. زیرا ماهیت آن طوری است که به جای توجه به خود مولفه‌ی‌ها بردار مثل روش اطلاعات مشترک و یا توجه به اندازه‌ی بردارها مثل روش فاصله‌ی اقلیدسی، به زاویه‌ی بین دو بردار توجه دارد.

۷- از منظر میزان بهبود نتایج بدست آمده، روش پیشنهادی به کمک داده ساختار درخت قرمز-سیاه زمان پاسخ به پرس و جوها نسبت به روش gMatrix بسیار بهبود یافته است، بیش از ۱۰ برابر، البته باید این نکته ذکر شود که پیاده سازی gMatrix توسط خود ما انجام شده است. در مورد نرخ مثبت کاذب نیز، میزان نرخ مثبت کاذب در روش شباهت کساین (بهترین جواب در بین روشهای پیشنهادی) از روش gMatrix کمتر شده است. بدین ترتیب در آزمایش اول به میزان ۴۱ درصد و در آزمایش دوم به میزان ۷۵ درصد بهبود بدست آمده است

## ۵-۱ کارهای آتی

- برای کارهای آتی و مرتفع کردن مشکلات این پژوهش راه‌حلهایی مدنظر است از جمله:
- استفاده از داده‌ساختارهای دیگر که از نظر مرتبه‌ی زمانی مشابه با درخت قرمز-سیاه عمل کنند اما از نظر مرتبه‌ی حافظه‌ی مورد نیاز نسبت به درخت قرمز-سیاه حافظه‌ی کمتری را نیاز داشته باشند.

<sup>۲۷</sup> Hamming Distance

<sup>۲۸</sup> Mixed Tabulation Hash Functions

<sup>۲۴</sup> Bit Mask

<sup>۲۵</sup> Disjoint Set

<sup>۲۶</sup> Jaccard Similarity

- Evaluation of the State-of-the-art", *Data Knowl. Eng.*, ۶۸(۴):۴۱۵-۴۳۰, ۲۰۰۹.
- [۱۷] P. Roy, A. Khan, and G. Alonso, "Augmented Sketch: Faster and More Accurate Stream Processing", In *SIGMOD*, ۲۰۱۶.
- [۱۸] L.S.Buriol, G.Frahling, S.Leonardi, A. Marchetti-Spaccamela, and C. Sohler, "Counting triangles in data streams", In *ACM Symposium on Principles of Database Systems*, pages ۲۵۳-۲۶۲, ۲۰۰۶.
- [۱۹] A.Pavan, K.Tangwongsan, S.Tirthapura, and K.-L.Wu. "Counting and sampling triangles from a graph stream", In *International Conference on Very Large Data Bases*, ۲۰۱۳.
- [۲۰] K. J. Ahn, S. Guha, and A. McGregor, "Graph sketches: sparsification, spanners, and subgraphs", In *ACM Symposium on Principles of Database Systems*, pages ۵-۱۴, ۲۰۱۲.
- [۲۱] H.Jowhari, M.Saglam, and G.Tardos, "Tight bounds for lp samplers, finding duplicates in streams, and related problems", In *ACM Symposium on Principles of Database Systems*, pages ۴۹-۵۸, ۲۰۱۱.
- [۲۲] R.Paghand C.E.Tsourakakis, "Colorful triangle counting and a map-reduce implementation", *Inf. Process. Lett.*, ۱۱۲(۷):۲۷۷-۲۸۱, ۲۰۱۲.
- [۲۳] L.Becchetti, P.Boldi, C.Castillo, and A.Gionis, "Efficient algorithms for large-scale local triangle counting", *TKDD*, ۴(۳), ۲۰۱۰.
- [۲۴] H.Jowhari and M.Ghods, "New streaming algorithms for counting triangles in graphs", In *COCOON*, pages ۷۱۰-۷۱۶, ۲۰۰۵.
- [۲۵] D.M.Kane, K.Mehlhorn, T.Sauerwald, and H.Sun, "Counting arbitrary subgraphs in *ACM Symposium on Theory of Computing*, pp. ۳۸۹-۳۹۸, ۲۰۰۲.
- [۶] B.Bollobás. "Extremal Graph Theory", *Academic Press*, New York, ۱۹۷۸.
- [۷] S.Baswana, "Streaming algorithm for graph spanners single pass and constant processing time per edge", *Inf. Process. Lett.*, ۱۰۶(۳):۱۱۰-۱۱۴, ۲۰۰۸.
- [۸] M. Elkin. "Streaming and fully dynamic centralized algorithms for constructing and maintaining sparse spanners", *ACM Transactions on Algorithms*, ۷(۲):۲۰, ۲۰۱۱.
- [۹] M. Elkin and J. Zhang, "Efficient algorithms for constructing  $(1 + \epsilon, \beta)$ -spanners in the distributed and streaming models", *Distributed Computing*, ۱۸(۵):۳۷۵-۳۸۵, ۲۰۰۶.
- [۱۰] R.Tarjan, "Data Structures and Network Algorithms", *SIAM*, Philadelphia, ۱۹۸۳.
- [۱۱] A.A.Benczur and D.R.Karger, "Approximating s-t minimum cuts in  $O(n^2)$  time", In *ACM Symposium on Theory of Computing*, pages ۴۷-۵۵, ۱۹۹۶.
- [۱۲] D.A.Spielman and S.-H.Teng, "Spectral sparsification of graphs", *SIAM J. Comput.*, ۴۰(۴):۹۸۱-۱۰۲۵, ۲۰۱۱.
- [۱۳] J.D.Batson, D.A.Spielman, and N.Srivastava, "Twice-ramanujan sparsifiers", *SIAM J. Comput.*, ۴۱(۶):۱۷۰۴-۱۷۲۱, ۲۰۱۲.
- [۱۴] K.J.Ahn and S.Guha, "Graph sparsification in the semi-streaming model", In *International Colloquium on Automata, Languages and Programming*, pages ۳۲۸-۳۳۸, ۲۰۰۹.
- [۱۵] J.A.Kelner and A.Levin. "Spectral sparsification in the semi-streaming setting", *Theory Comput. Syst.*, ۵۳(۲):۲۴۳-۲۶۲, ۲۰۱۳.
- [۱۶] N. Manerikar and T. Palpanas, "Frequent Items in Streaming Data: An Experimental

data streams", In *International Colloquium on Automata, Languages and Programming*, pages ۵۹۸–۶۰۹, ۲۰۱۲.

[۲۶] M.Manjunath, K.Mehlhorn, K.Panagiotou, and H.Sun, "Approximate counting of cycles in streams" In *European Symposium on Algorithms*, pages ۶۷۷–۶۸۸, ۲۰۱۱.

[۲۷] Dahlgaard, Søren, Mathias Knudsen, and Mikkel Thorup, "Practical Hash Functions for Similarity Estimation and Dimensionality Reduction", *Advances in Neural Information Processing Systems*, ۲۰۱۷.





## بیشینه‌سازی طول عمر شبکه‌های حسگر صوتی زیرآبی با جایگذاری بهینه گره‌های رله

زهرا محمدی\* محدثه سلیمان پورمقدم\*\* داریوش عباسی مقدم\*\*\* سیامک طالبی\*\*\*\*

\* کارشناس ارشد - گروه مهندسی برق - دانشگاه شهید باهنر کرمان

\*\* استادیار - گروه مهندسی مکانیک - مجتمع آموزش عالی بم

\*\*\* دانشیار - گروه مهندسی برق - دانشگاه شهید باهنر کرمان

\*\*\*\* استاد - گروه مهندسی برق - دانشگاه شهید باهنر کرمان

تاریخ دریافت: ۱۳۹۹/۰۱/۲۵ تاریخ پذیرش: ۱۳۹۹/۰۹/۰۶

نوع مقاله: پژوهشی

### چکیده

شبکه‌های حسگر صوتی زیرآبی به دلیل ویژگی‌های مطلوب خود و کاربردهای عملی گسترده در زمینه‌های ارتباطی مختلف، توجه زیادی را به خود جلب کرده‌اند. از آنجا که گره‌های حسگر زیرآبی دارای هزینه‌ی بالا و پیچیدگی جایگذاری هستند، افزایش طول عمر این شبکه‌ها از اهمیت زیادی برخوردار است. گره‌های رله نقش مهمی در کاهش فاصله مخابراتی و انرژی مصرفی دارند. اما، مسئله مهم قرارگیری بهره‌ور گره‌های رله اطراف گره‌های بحرانی شبکه به منظور جلوگیری از حذف آن‌ها و در نتیجه افزایش طول عمر شبکه است. برای این منظور، در این مقاله روشی جدید به نام جایگذاری بهره‌ور گره رله (ERS) معرفی شده است که شامل فرمول‌بندی کردن مکان هر گره رله به صورت یک مسئله بهینه‌سازی غیرمحدب است. در حقیقت، وجود قیود تفاضل محدب منجر به غیرمحدب شدن مسئله بهینه‌سازی پیشنهادی می‌شود و دستیابی به جواب بهینه را دشوار می‌سازد. از این‌رو، در گام بعد با پیشنهاد یک تبدیل جدید، مسئله مذکور به معادل محدب خود تبدیل می‌شود. مهم‌ترین مزیت مسئله برنامه‌ریزی محدب، قابلیت دستیابی به جواب بهینه مسئله است. نتایج شبیه‌سازی نشان‌دهنده برتری عملکرد روش پیشنهادی در طول عمر و بهره‌وری نسبت به روش ابتکاری پیشین تنظیم گره رله (RA) است.

**واژگان کلیدی:** شبکه‌های حسگر صوتی زیرآبی، گره‌های رله، طول عمر شبکه، مخابرات چندپرسی، برنامه‌ریزی محدب.

شده است [۱]. شبکه‌های حسگر صوتی زیرآبی در مواردی مانند

نظارت بر آلودگی، نظارت بر استخراج نفت و نظارت بر آبی‌پروری

به کار برده می‌شوند، از این‌رو، اخیراً بسیار مورد توجه قرار گرفته‌اند

[۲، ۳]. از جمله کاربردهای دیگر این شبکه‌ها می‌توان به اخطار

سونامی، پیدا کردن لاشه کشتی، کاربردهای نظامی و امنیتی،

### ۱. مقدمه

بخش اعظم سطح کره زمین با آب پوشیده شده است و کنترل آن

حائز اهمیت است. با پیشرفت سریع فناوری شبکه‌های حسگر بی-

سیم، انگیزه‌ای برای گسترش شبکه‌های حسگر صوتی زیرآبی ایجاد

<sup>۱</sup> نویسنده مسئول: سیامک طالبی [siamak.talebi@uk.ac.ir](mailto:siamak.talebi@uk.ac.ir)

آن‌ها در محیط زیرآبی به صورت تنک<sup>۳</sup> انجام می‌شود [۵]. جایگذاری تنک گره‌های حسگر زیرآبی منجر به طولانی شدن فاصله‌ی مخابره<sup>۴</sup> بین آن‌ها و در نتیجه افزایش انرژی مصرفی گره‌های حسگر می‌شود. از آن‌جا که گره‌های حسگر با توجه به الزامات پوشش<sup>۵</sup> و اتصال<sup>۶</sup> شبکه در منطقه هدف پراکنده می‌شوند، کاهش فاصله بین آن‌ها به منظور کاهش انرژی مصرفی عملی نیست. از این‌رو در [۷، ۸]، بکارگیری گره‌های رله بزرگ و پر قدرت<sup>۷</sup> به منظور کاهش فاصله مخابراتی بررسی شده است. این گره‌ها دارای انرژی بیش‌تر و اندازه بزرگ‌تر نسبت به گره‌های حسگر هستند و به عنوان گره‌های مسیریاب<sup>۸</sup> جهت ارسال اطلاعات گره‌های حسگر به سمت ایستگاه پایه عمل می‌کنند. عیب بزرگ این گره‌ها اندازه بزرگ آن‌ها است که در نتیجه به آسانی توسط دشمن قابل شناسایی هستند. بنابراین بکارگیری گره‌های رله کوچک<sup>۹</sup> که دارای اندازه و انرژی یکسان با گره‌های حسگر هستند از لحاظ عملی مناسب‌تر است [۹].

در این مقاله جایگذاری گره‌های رله کوچک در شبکه حسگر صوتی زیرآبی سه‌بعدی در نظر گرفته شده است. در حقیقت، از مشارکت گره‌های رله جهت بهبود حفره انرژی و افزایش طول عمر شبکه حسگر زیرآبی استفاده شده است. علاوه بر آن، استفاده بهینه از انرژی گره‌های رله مسئله‌ای است که در این تحقیق به آن پرداخته شده است. در ادامه، به دسته‌بندی مطالعات صورت گرفته در حوزه افزایش طول عمر شبکه حسگر صوتی زیرآبی پرداخته می‌شود. عده‌ای از محققین روی طراحی پروتکل مسیریابی برای متعادل کردن مصرف انرژی بین گره‌های حسگر متمرکز شده‌اند. برای مثال، در [۱۰، ۱۱] به منظور افزایش طول عمر شبکه حسگر زیرآبی، فاصله بین گره‌های حسگر و انرژی باقی‌مانده<sup>۱۰</sup> هر کدام به عنوان معیار برای انتخاب گره‌های ارسال‌کننده به کار رفته است. دو پروتکل مسیریابی در [۱۲] ارائه شده است، بر مبنای اولین روش پیشنهادی هر گره طبق انرژی باقی‌مانده خود به صورت تک یا چندپرسی با دیگر گره‌ها به مخابره اطلاعات می‌پردازد. هم‌چنین،

اکتشاف منابع زیرآبی و غیره اشاره کرد. گره‌های حسگر موجود در این شبکه‌ها توانایی مشاهده محیط اطراف، جمع‌آوری اطلاعات از محیط، پردازش اندک اطلاعات جمع‌آوری شده و ارسال آن‌ها را دارند. هر کدام از این اعمال باعث صرف انرژی می‌شود، البته اکثر انرژی در انتقال اطلاعات از دست می‌رود [۴].

از آن‌جا که انرژی مصرفی گره‌های حسگر از طریق یک باتری با انرژی محدود تامین می‌شود، این باتری‌ها پس از مدتی تخلیه می‌شوند. با توجه به طبیعت محیط زیرآبی، امکان شارژ دوباره یا عوض کردن باتری گره‌های حسگر زیرآبی وجود ندارد. بنابراین، اتمام انرژی گره‌ها علاوه بر نا کارآمدی آن‌ها در نظارت بر محیط اطراف، منجر به عدم توانایی آن‌ها در ارسال اطلاعات سایر گره‌های حسگر در مخابرات چندپرسی<sup>۱</sup> می‌شود. از سویی دیگر، هزینه‌ی گره‌های حسگر زیرآبی در مقایسه با گره‌های حسگر زمینی بسیار بالا است و در نتیجه افزایش مدت کارکرد این گره‌ها مسئله‌ای مهم است [۵].

ساخت یک شبکه حسگر زیرآبی چندپرسی در یک منطقه بزرگ چندین چالش را به دنبال دارد. برای مثال، گره‌های حسگر نزدیک به ایستگاه پایه علاوه بر ارسال اطلاعات خود به سمت ایستگاه پایه مسئول ارسال اطلاعات بقیه گره‌های حسگر هستند که موجب اتمام سریع‌تر انرژی آن‌ها نسبت به سایر گره‌های حسگر می‌شود. نویسندگان در [۶] ثابت کرده‌اند زمانی که گره‌های نزدیک به ایستگاه پایه تمام انرژی خود را از دست داده‌اند، سایر گره‌ها ۹۳٪ انرژی خود را حفظ کرده‌اند. این مسئله باعث ایجاد حفره انرژی<sup>۲</sup> در شبکه و در نتیجه جزءبندی شدن شبکه می‌شود. در حقیقت، در یک شبکه جزءبندی شده بعضی از گره‌های حسگر با وجود داشتن انرژی توانایی ارسال اطلاعات خود را به سمت ایستگاه پایه ندارند و در نتیجه طول عمر شبکه به پایان می‌رسد.

همان‌گونه که پیش از این بیان شد، هزینه‌ی گره‌های حسگر زیرآبی در مقایسه با هم‌تای زمینی آن‌ها بالاتر است، از این‌رو جایگذاری

<sup>۱</sup> Multi-hop communication

<sup>۲</sup> Energy hole

<sup>۳</sup> Sparse

<sup>۴</sup> Transmission distance

<sup>۵</sup> Coverage

<sup>۶</sup> Connectivity

<sup>۷</sup> Macro relay nodes

<sup>۸</sup> Router nodes

<sup>۹</sup> Micro relay nodes

<sup>۱۰</sup> Residual energy

[۱۹] به منظور کاهش انرژی مصرفی از حرکت گره رله بین دو گره در حال مخابره استفاده شده است. پس از پایان روند مخابره گره‌ها، گره رله به مکان اولیه خود بازمی‌گردد. همچنین به منظور جایگذاری گره‌های رله در شبکه حسگر سه‌بعدی در [۲۰] یک روش ابتکاری به نام تنظیم گره رله<sup>۵</sup> (RA) پیشنهاد شده است. طبق روش پیشنهادی ابتکاری RA، گره‌های رله با توجه به یک تابع چگالی در مکان‌های مختلف شبکه قرار می‌گیرند. بدین صورت که در مناطق با انرژی مصرفی بالاتر، تعداد بیش‌تری گره رله قرار می‌گیرد. اما در این روش مکان گره‌های رله در دو گام به دست می‌آید به گونه‌ای که ابتدا مکان آن‌ها روی سطح آب به صورت تصادفی تعیین می‌شود و سپس عمق آن‌ها تنظیم می‌شود. این نوع روش‌های جایگذاری ابتکاری دست‌یابی به مکان شبه‌بهینه را به دنبال دارد که باعث اتلاف منابع موجود می‌شود.

از سویی دیگر، برخی از نویسندگان با تعریف مسئله بهینه‌سازی ریاضی به جایگذاری گره رله در شبکه پرداخته‌اند [۲۱، ۲۲]. به عنوان مثال، روش پیشنهادی در [۲۲] شامل یک مسئله بهینه‌سازی غیر محدب است. مسائل بهینه‌سازی غیر محدب دارای چندین نقطه بهینه محلی هستند و جواب مسئله به نقطه اولیه روش‌های تکرار بستگی دارد، از این‌رو پیدا کردن نقطه بهینه سراسری دشوار و گاه غیر ممکن است [۲۳]. این عوامل منجر به زمان‌بر شدن و در نتیجه ناکارآمدی روش مطرح شده در [۲۲] می‌گردد. به علاوه، نویسندگان در [۲۱] ابتدا با در نظر گرفتن چیدمان دایروی برای قرارگیری گره‌های حسگر، جایگذاری گره رله را به صورت هم‌زمان با گره‌های حسگر مورد مطالعه قرار داده‌اند. با در نظر گرفتن این فرضیات و مدل سیستم، جایگذاری گره رله به صورت یک مسئله بهینه‌سازی دو هدفه<sup>۶</sup> شامل دو هدف متضاد فرمول‌بندی شده است. به منظور تبدیل مسئله دو هدفه به مسئله تک هدفه از روش مجموع وزنی<sup>۷</sup> استفاده شده است و در نهایت الگوریتم تفاضل محدب<sup>۸</sup> برای حل آن به کار رفته است. در این مقاله، با پیروی از [۲۰] حالت کلی‌تر و عملی‌تری از جایگذاری گره‌های حسگر

نویسندگان در پروتکل پیشنهادی دوم، با تخصیص انرژی بیش‌تر به گره‌هایی که مسئول حمل بار بیش‌تری در شبکه هستند سعی در رفع مسئله حفره انرژی در شبکه داشته‌اند.

به هر حال، از آن‌جا که محیط زیرآبی غیر قابل پیش‌بینی است، طراحی یک پروتکل مسیریابی مناسب نیازمند پیچیدگی محاسباتی بالا است [۱۳]. از این رو برخی رویکردهای کاهش مصرف انرژی حول نحوه جمع‌آوری اطلاعات از گره‌های حسگر توسعه یافته‌اند. به عنوان مثال، در [۱۴، ۱۵] جمع‌آوری داده مبتنی بر خوشه‌بندی<sup>۱</sup> ارائه شده است، در این روش‌ها کل گره‌های حسگر به تعدادی خوشه تقسیم می‌شوند و در هر خوشه یک گره به عنوان سرخوشه<sup>۲</sup> انتخاب می‌شود. پس از آن سرخوشه‌ها مسئول ارسال اطلاعات سایر گره‌های حسگر به سمت ایستگاه پایه می‌شوند. لازم به ذکر است در [۱۴] بعد از انتخاب سرخوشه‌ها، هر خوشه به زیرخوشه‌هایی تقسیم می‌شود. در هر کدام از زیرخوشه‌ها یک گره به عنوان گره نماینده، مسئول جمع‌آوری داده از دیگر گره‌ها می‌شود. همچنین در [۱۵] انرژی باقی‌مانده و مکان گره‌های حسگر به منظور انتخاب بهینه سرخوشه‌ها در نظر گرفته شده است.

در شبکه‌های حسگر صوتی زیر آبی خطای ارسال تحت پیوندهای صوتی طولانی غیر قابل اجتناب است، از این‌رو، در سال‌های اخیر بحث کاهش احتمال خطا مورد توجه برخی از محققین قرار گرفته است. به منظور افزایش قابلیت اطمینان، ابتدایی‌ترین روش درخواست بازفرستی خودکار<sup>۳</sup> است [۱۶]. اما از آن‌جا که این روش بر پایه ارسال دوباره اطلاعات هست، موجب توان مصرفی بالا در شبکه می‌شود. به همین منظور، نویسندگان در [۱۷، ۱۸] به بررسی روش تصحیح خطای رو به جلو<sup>۴</sup> به منظور افزایش طول عمر شبکه حسگر زیر آبی پرداخته‌اند. لازم به ذکر است که در روش تصحیح خطای رو به جلو مسئله بهینگی انرژی با بازفرستی‌های بسیار کم قابل حل هست.

توجه به مسئله جایگذاری گره رله به منظور افزایش طول عمر شبکه حسگر صوتی زیرآبی در [۱۹-۲۲] بررسی شده است. در

<sup>۱</sup> Clustering

<sup>۲</sup> Cluster head

<sup>۳</sup> Automatic Repeat reQuest (ARQ)

<sup>۴</sup> Forward Error Correction

<sup>۵</sup> Relay node Adjustment

<sup>۶</sup> Bi-objective optimization problem

<sup>۷</sup> Weighted sum method (WSM)

<sup>۸</sup> Difference Convex Algorithm

بیشینه‌سازی محدب برابری بهینه محلی<sup>۳</sup> و بهینه سراسری<sup>۴</sup> آن‌ها است، الگوریتم‌هایی که برای حل این مسائل توسعه یافته‌اند از این ویژگی استفاده می‌کنند و بسیار سریع و کارا به جواب بهینه دست می‌یابند.

۳. به منظور ارزیابی عملکرد روش پیشنهادی، به بررسی پیچیدگی زمانی و مقایسه طول عمر (از طریق آزمایشات شبیه‌سازی) با روش RA موجود در [۲۰] پرداخته شده است. طبق بررسی انجام شده، پیچیدگی روش پیشنهادی همانند روش RA و طبق آزمایشات شبیه‌سازی عملکردی بالاتر از آن دارد.

ادامه این مقاله به صورت زیر سازمان‌دهی شده است:

در بخش دوم، مدل سیستم ارائه می‌شود. روش پیشنهادی جای‌گذاری گره رله و مدل معادل محدب آن در بخش سوم به طور مفصل شرح داده می‌شود. در بخش چهارم، به صورت تحلیلی به بررسی بهره‌وری روش پیشنهادی در برابر تغییرات محیط زیر آبی پرداخته می‌شود. بررسی پیچیدگی روش پیشنهادی و پس از آن نتایج ارزیابی روش پیشنهادی در بخش پنجم ارائه و تحلیل خواهند شد. در نهایت در بخش ششم، به جمع‌بندی این تحقیق پرداخته خواهد شد.

## ۲. مدل سیستم

در این بخش به تشریح کامل شبکه حسگر صوتی زیرآبی سه‌بعدی پرداخته می‌شود. برای این منظور، ابتدا مدل سیستم و انرژی مصرفی مشخص می‌شوند، پس از آن تعریف طول عمر شبکه حسگر صوتی ارائه می‌شود. در انتهای بخش، به شرح فرضیات مورد استفاده در تحقیق پرداخته خواهد شد.

شبکه حسگر صوتی زیرآبی، سه‌بعدی و عمق محیط زیرآبی  $H$  در نظر گرفته می‌شود. منطقه هدف به صورت استوانه نشان داده شده در شکل ۱ مدل می‌شود. شبکه دارای سه نوع گره شامل: گره حسگر، گره رله و ایستگاه پایه است. گره‌های حسگر مسئول نظارت بر محیط اطراف و جمع‌آوری اطلاعات از آن هستند و هر یک مجهز به فرستنده/گیرنده صوتی برای مخابرات تحت کانال زیرآبی هستند. اطلاعاتی که توسط هر گره حسگر جمع‌آوری می‌شود، باید به

(جایگذاری تصادفی) در نظر گرفته شده است. هدف اصلی، طراحی یک روش جایگذاری رله بر اساس مسائل بهینه‌سازی محدب است. در حقیقت، به دلیل برابر بودن جواب بهینه محلی با بهینه سراسری در مسائل بهینه‌سازی محدب، دستیابی به جواب بهینه بسیار سریع‌تر است. اما در عمل، بسیاری از مسائل به فرم غیر محدب هستند. در مواجهه با این نوع مسائل، بهترین راه‌حل در صورت امکان تبدیل مسئله غیر محدب به همتای محدب آن است و در غیر این صورت به دست آوردن جواب تقریبی مسئله از طریق تقریب محدب مسئله مورد توجه قرار گرفته است [۲۴].

## ۱-۱ نوآوری و سازمان‌دهی تحقیق

به طور کلی نوآوری مقاله به صورت زیر خلاصه می‌گردد:

۱. با توجه به نقش مهم گره‌های رله در افزایش طول عمر شبکه و به منظور غلبه بر چالش‌های موجود در روش‌های جایگذاری ابتکاری، هدف اصلی این مقاله به دست آوردن مکان بهینه گره‌های رله در شبکه‌های حسگر صوتی زیرآبی است. جهت حصول به این هدف، در ابتدا با پیشنهاد الگوریتم «جایگذاری بهره‌ور گره رله (ERS)» مسئله جایگذاری گره رله به صورت یک مسئله بهینه‌سازی ریاضی فرمول‌بندی می‌شود. نکته حائز اهمیت، اعمال قید عملی به مسئله طرح شده است که قابلیت جایگذاری هم‌زمان و غیرهم‌زمان گره‌های رله با گره‌های حسگر را به دنبال دارد. به بیانی بهتر، طبق روش پیشنهادی نه تنها حداکثر اتصال گره بحرانی با هم‌سایگان عمق پایین‌تر آن تضمین می‌شود بلکه انرژی گره رله تا حد امکان ذخیره می‌گردد. قید مذکور به صورت یک قید خطی (با پیچیدگی پایین) به مسئله اضافه می‌شود.

۲. با توجه به طبیعت غیرمحدب<sup>۲</sup> مسئله طرح شده، دستیابی به مکان بهینه گره‌های رله چالش‌برانگیز است. طبق مطالعات صورت گرفته، در کارهای پیشین به بررسی بهینه جایگذاری گره‌های رله پرداخته نشده است. از این رو، در گام بعد با پیشنهاد یک تبدیل جدید مسئله پیشنهادی به معادل محدب خود تبدیل می‌شود. ویژگی مهم مسائل

<sup>۱</sup> Efficient Relay node Setting

<sup>۲</sup> Non-convex

<sup>۳</sup> Local optimum

<sup>۴</sup> Global optimum

انرژی مصرفی یک گره را مشخص می‌کند. انرژی مصرفی از سالی بین دو گره تابعی از فاصله بین دو گره، نرخ اطلاعات و هم‌چنین ویژگی‌های محیطی است [۱۵]. به طور کلی مقدار انرژی که صرف ارسال یک بیت اطلاعات از گره  $i$  به سمت گره  $j$  می‌شود، با نماد  $p_{ij}$  مشخص می‌شود و از رابطه زیر قابل محاسبه است [۱۵، ۲۰]:

$$p_{ij} = \begin{cases} p_s + e_{fs}d_{ij} & d_{ij} < d_t \\ p_s + e_{mp}d_{ij} & d_{ij} \geq d_t \end{cases} \quad (۲)$$

در رابطه فوق،  $p_s$  انرژی مصرفی برای پردازش یک بیت اطلاعات جهت ارسال است،  $e_{fs}$  و  $e_{mp}$  به ترتیب نشان‌دهنده ضریب تقویت‌کننده ارسال در مدل فضای آزاد و چندمسیری هستند، هم‌چنین،  $d_t$  یک فاصله آستانه است. به طور خاص در رابطه با کانال زیرآبی، ضریب تقویت‌کننده برای هر دو مدل  $a^{d_{ij}}$  است [۱۵، ۲۰]، به طوری که  $a$  ضریب جذب در محیط زیرآبی است. در نتیجه، معادله مصرف انرژی برای ارسال یک بیت اطلاعات از گره  $i$  به سمت گره  $j$  به صورت زیر بازنویسی می‌شود:

$$p_{ij} = \begin{cases} p_s + a^{d_{ij}}d_{ij} & d_{ij} < d_t \\ p_s + a^{d_{ij}}d_{ij} & d_{ij} \geq d_t \end{cases} \quad (۳)$$

ضریب جذب به فرکانس سیگنال صوتی ( $f$ ) ربط دارد و برای فرکانس‌های بالاتر از یک کیلوهرتز بر حسب  $dB/km$  از طریق رابطه زیر محاسبه می‌شود [۲۵]:

$$10 \log \alpha = 0.1 \frac{f^2}{1+f^2} + \frac{4.0f^2}{4100+f^2} + 2.75 \times 10^{-4} f^2 + 0.003 \quad (۴)$$

لازم به ذکر است، این ضریب برای فرکانس‌های پایین‌تر از یک کیلوهرتز به صورت زیر قابل بیان است:

$$10 \log \alpha = 0.02 + 0.1 \frac{f^2}{1+f^2} + 0.11f^2 \quad (۵)$$

در ادامه تعریف طول عمر شبکه حسگر صوتی زیرآبی ارائه می‌شود.

## ۲-۲- طول عمر شبکه

تعاریف متعددی برای طول عمر شبکه‌های حسگر صوتی زیرآبی وجود دارد، از جمله نویسندگان در [۱۳، ۲۰، ۲۶] طول عمر شبکه را مدت زمان اتمام طول عمر اولین گره، تعریف کرده‌اند. در [۲۷] دوره زمانی اتمام طول عمر همه گره‌های شبکه را به عنوان طول عمر شبکه در نظر گرفته‌اند. برخی دیگر آن را مدت زمان پایان طول عمر بخش مشخصی از گره‌ها در نظر گرفته‌اند. با توجه به این که یک مخابره تا زمانی موثر و متعادل است که همه گره‌ها در

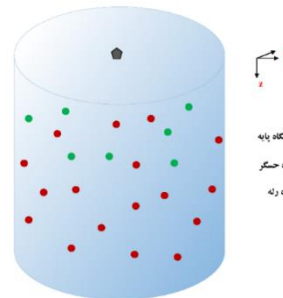
ایستگاه پایه‌ی روی سطح آب برسد. هر گره حسگر علاوه بر جمع‌آوری اطلاعات از محیط اطراف، توانایی رله کردن اطلاعات سایر گره‌های حسگر را دارد، در حقیقت اطلاعات گره حسگری که در فاصله زیاد از ایستگاه پایه قرار دارد با چندین پرش به آن می‌رسد. مکان گره‌های حسگر در منطقه هدف به صورت تصادفی و مجموعه آن‌ها با  $K$  مشخص می‌شود.

گره‌های رله به عنوان گره‌های مسیریاب بین گره‌های حسگر عمل می‌کنند، مجموعه این گره‌ها با  $R$  مشخص می‌شود و هر یک مجهز به فرستنده/گیرنده صوتی برای مخابرات زیرآبی هستند. برخلاف گره‌های حسگر که مکان آن‌ها به صورت تصادفی تعیین می‌شود، مکان گره‌های رله به صورتی دقیق برای بهبود عملکرد شبکه تعیین می‌شود. در این مقاله، گره‌های رله به منظور افزایش طول عمر شبکه به آن اضافه می‌شوند. شایان ذکر است که از دیدگاه عملی، ابتدا مکان گره‌های رله روی سطح آب ثابت می‌شود و سپس با استفاده از سیستم تنظیم عمق<sup>۱</sup>، عمق آن‌ها معین می‌گردد [۲۹].

هر دو نوع گره حسگر و رله دارای انرژی اولیه  $E_p$  و برد مخابراتی  $RT$  هستند،  $d_{ij}$  بیان‌گر فاصله اقلیدسی بین دو گره  $i$  و  $j$  است و به صورت زیر محاسبه می‌شود:

$$d_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \quad (۱)$$

در رابطه فوق،  $(x_i, y_i, z_i)$  و  $(x_j, y_j, z_j)$  به ترتیب مختصات دکارتی گره  $i$  و  $j$  است. در ادامه مدل انرژی مصرفی شبکه حسگر صوتی زیرآبی شرح داده می‌شود.



شکل ۱. قرارگیری گره‌های رله و حسگر در منطقه هدف استوانه‌ای و دستگاه مختصات دکارتی

## ۲-۱- مدل انرژی مصرفی

مقدار انرژی که صرف مخابره اطلاعات و پردازش آن می‌شود، کل

<sup>۱</sup>depth adjustment system

$e_i$	انرژی باقیمانده گره $i$
$E_p$	انرژی اولیه گره‌ها
$(x_i, y_i, z_i)$	مختصات گره $i$
$d_{ij}$	فاصله اقلیدسی دو گره $i$ و $j$
$p_{ij}$	توان مصرفی برای ارسال یک بیت اطلاعات از گره $i$ به گره $j$
$\alpha$	ضریب جذب محیط زیر آبی
$d_t$	فاصله آستانه
$p_s$	توان مصرفی برای پردازش یک بیت اطلاعات جهت ارسال
$p_r$	توان مصرفی برای پردازش یک بیت اطلاعات جهت دریافت
$f$	فرکانس سیگنال صوتی
$\tau_i$	طول عمر گره $i$

شبکه موجود باشند [۱۳]، بنابراین، طول عمر شبکه ( $\tau$ ) مشابه [۱۳، ۲۰، ۲۶] به صورت زیر تعریف می‌شود:

$$\tau = \min(\tau_i), \forall i \in S \cup R, \quad (6)$$

در رابطه فوق،  $\tau_i$  طول عمر گره نوعی  $i$  است و از رابطه زیر قابل محاسبه است [۲۰]:

$$\tau_i \left( \sum_{j \in SURUBS, i \neq j} p_{ij} r_{ij} + p_r \sum_{k \in SURUBS, k \neq i} r_{ki} \right) = e_i \quad (7)$$

لازم به ذکر است،  $p_r$  انرژی مصرفی گره  $i$  جهت دریافت یک بیت اطلاعات و  $r_{ij}$  بیانگر نرخ ارسالی از گره  $i$  به سمت گره  $j$  است. هم‌چنین،  $e_i$  انرژی باقی‌مانده گره  $i$  است. به منظور سهولت خوانندگان، لیست نمادهای مورد استفاده در مقاله در جدول ۱ آورده شده است. فرضیات مورد استفاده در این مقاله در زیر ذکر شده‌اند:

### ۳. روش پیشنهادی جایگذاری گره‌های رله

همان‌گونه که پیش از این بیان شد، در این مقاله جایگذاری گره‌های رله به منظور افزایش طول عمر شبکه مورد بررسی قرار گرفته است. طبق تعریف طول عمر شبکه در رابطه (۶)، گره‌ای که دارای کم‌ترین طول عمر است زمان کارکرد شبکه را مشخص می‌کند و به عنوان گره بحرانی<sup>۱</sup> شناخته می‌شود. بنابراین، با قرار دادن گره رله اطراف گره مذکور طول عمر شبکه افزایش می‌یابد. برای این منظور، روش ERS برای تعیین مکان هر گره رله پیشنهاد می‌شود. هم‌چنین، تعیین مکان دقیق گره رله تحت مسئله برنامه‌ریزی محدب یکی دیگر از مباحثی است که در این مقاله به بررسی آن پرداخته شده است. در ادامه روش پیشنهادی به طور مفصل شرح داده می‌شود.

#### ۳.۱ الگوریتم پیشنهادی ERS

گام اول برای جایگذاری گره رله، شناسایی گره بحرانی (تعیین‌کننده طول عمر شبکه) است، این گره به صورت  $b$  مشخص می‌شود. بدین معنی که:

$$b = \operatorname{argmin}(\tau_j), \forall j \in S \cup R. \quad (9)$$

همان‌گونه که از رابطه (۳) برمی‌آید، مخابره هر گره با دورترین گره هم‌سایه بیش‌ترین انرژی را مصرف می‌کند. در روش پیشنهادی به این مسئله توجه شده است و به منظور افزایش طول عمر شبکه، گره رله در خط واصل گره بحرانی شبکه و دورترین همسایه‌اش قرار

۱. اتصال بین گره‌های حسگر برقرار است و گره‌های رله به منظور افزایش طول عمر شبکه به آن اضافه می‌شوند.

۲. مکان گره‌های حسگر و انرژی باقی‌مانده آن‌ها معلوم است [۲۰].

۳. شبکه حسگر همگن فرض می‌شود و برد مخابراتی همه گره‌ها برابر است.

۴. پهنای باند کانال صوتی زیرآبی محدود فرض می‌شود. این شرط از ظرفیت محدود لینک مشخص شده با نماد  $\beta_{cap}$  منعکس می‌شود. شرط محدود بودن پهنای باند کانال صوتی زیرآبی بیان می‌دارد که مجموع نرخ ارسالی از هر گره بایستی کم‌تر از  $\beta_{cap}$  باشد، بدین معنی که [۲۰]:

$$\sum_{j \in SURUBS, i \neq j} r_{ij} \leq \beta_{cap}, \forall i \in S \cup R. \quad (8)$$

با داشتن اطلاعات کامل از مدل سیستم، در ادامه به شرح روش پیشنهادی برای قرارگیری گره‌های رله در شبکه پرداخته می‌شود.

#### جدول ۱. لیست نمادها و تشریح آن‌ها

نماد	تشریح
$R$	مجموعه گره‌های رله
$S$	مجموعه گره‌های حسگر
$RT$	برد مخابراتی گره‌های حسگر و رله
$r_{ij}$	نرخ ارسالی از گره $i$ به گره $j$

<sup>۱</sup> Critical node

باشد گره رله روی آن منطبق می‌شود. با این نوع جایگذاری، گره رله مسئول مخابره با گره  $f$  تحت فاصله طولانی  $d_{bf}$  می‌شود که منجر به بحرانی شدن وضعیت آن و در نتیجه ناکارآمدی جایگذاری‌اش می‌شود. خصوصاً در جایگذاری هم‌زمان گره‌های حسگر و رله که هر دو نوع گره دارای انرژی باقی‌مانده برابر (همان انرژی اولیه یا  $E_p$ ) هستند انرژی مصرفی گره‌های رله نیز از اهمیت خاصی برخوردار است. از این‌رو، هدف اصلی روش پیشنهادی افزایش طول عمر شبکه حسگر زیرآبی و تضمین ارسال اطلاعات جمع‌آوری شده توسط گره‌های حسگر به سمت ایستگاه پایه است. به بیانی دیگر، چنانچه گره بحرانی  $b$  توانایی ارسال اطلاعات دریافتی از گره‌های همسایه‌اش را نداشته باشد، اتصال او با آن‌ها حفظ می‌شود. حال، مجموعه  $N_b$  به صورت زیر تعریف می‌شود

$$N_b = \{j; r_{jb} > 0\}. \quad (15)$$

که بیان‌گر کل گره‌هایی است که اطلاعات خود را به سمت گره  $b$  ارسال می‌کنند. شایان ذکر است که اتصال همواره دو گره  $b$  و  $j$  در صورتی وجود دارد که

$$\tau_b = \tau_j \quad (16)$$

از جهت دیگر، طبق آن چه پیش از این بیان شد توجه به توان مصرفی گره رله از اهمیت خاصی برخوردار است. بنابراین، به منظور دستیابی به مکان گره رله، یک مسئله بهینه‌سازی مقید در رابطه با طول عمر گره  $b$  به صورت زیر پیشنهاد و طرح می‌شود:

$$(x_r, y_r, z_r, p_{br}, d_{br}) = \operatorname{argmax} \tau_b \quad (17)$$

s.t.

$$\tau_b \leq \max(\tau_k), \forall k \in \{N_b\}, \quad (18)$$

$$x_r = x_b + \frac{x_f - x_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (19)$$

$$y_r = y_b + \frac{y_f - y_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (20)$$

$$z_r = z_b + \frac{z_f - z_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (21)$$

$$p_{br} = p_s + \alpha^{d_{br}} d_{br}^{\gamma}, 0 \leq d_{br} < d_t \quad (22)$$

$$p_{br} = p_s + \alpha^{d_{cr}} d_{cr}^{\xi}, d_t \leq d_{cr} \leq RT \quad (23)$$

$$d_{br}^{\gamma} = (x_b - x_r)^{\gamma} + (y_b - y_r)^{\gamma} + (z_b - z_r)^{\gamma} \quad (24)$$

مسئله فوق جزء مسائل بهینه‌سازی غیرمحدب در نظر گرفته می‌شود. این مسئله دارای تابع هدف محدب (بیان شده در (۲۵))،

می‌گیرد و به عنوان گره مسیریاب بین آن‌ها عمل می‌کند. بنابراین، در گام بعد دورترین همسایه گره  $b$  به صورت زیر شناسایی و  $f$  نام‌گذاری می‌شود:

$$f = \operatorname{argmax}(d_{bj}), \forall j \in \{k; r_{bk} > 0\}. \quad (10)$$

به منظور افزایش طول عمر گره بحرانی  $b$ ، گره رله  $r$  در خط واصل گره‌های  $f$  و  $b$  به صورت نشان داده شده در شکل ۲ قرار می‌گیرد، از مزایای این نوع جایگذاری می‌توان به موارد زیر اشاره کرد:

۱. در این نوع جایگذاری، مطمئناً گره رله در برد مخابراتی دو گره حسگر  $b$  و  $f$  قرار می‌گیرد و توانایی مخابره با آن‌ها را دارد.
۲. نوع جایگذاری رله از نوع کارآمد است، بدین معنی که از مخابره گره رله در فاصله‌ای بیش از طول لینک  $(b, f)$  جلوگیری می‌شود.

در صورتی که بردار مکان گره نوعی  $j$  با  $\vec{j}$  نمایش داده شود (ایستگاه پایه واقع بر مبدا مختصات است)، با توجه به شکل ۲ ارتباط  $\vec{r}$  با  $\vec{f}$  و  $\vec{b}$  به صورت زیر بیان می‌شود:

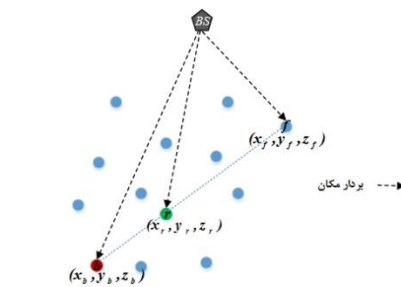
$$\vec{r} = \vec{b} + \frac{\vec{f} - \vec{b}}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (11)$$

در رابطه فوق،  $d_{br}$  فاصله اقلیدسی سه‌بعدی بین دو گره بحرانی و رله است. بنابراین، مختصات گره رله از طریق معادلات زیر به دست می‌آید:

$$x_r = x_b + \frac{x_f - x_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (12)$$

$$y_r = y_b + \frac{y_f - y_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (13)$$

$$z_r = z_b + \frac{z_f - z_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \quad (14)$$



شکل ۲. قرارگیری گره رله در شبکه حسگر زیرآبی طبق روش پیشنهادی

واضح است اگر تنها هدف، بهینه‌سازی طول عمر گره بحرانی  $b$



$$f_i(x) \leq 0, i = 1, 2, \dots, I \quad (28)$$

$$\psi_m(x) - t_m = 0, m = 1, 2, \dots, M \quad (29)$$

$$\Phi_m(x) - t_m = 0, m = 1, 2, \dots, M \quad (30)$$

همان‌گونه که پیش از این بیان شد،  $f_i(x), i = 0, 1, \dots, I$  توابعی محدب هستند. همچنین، دو قید (۳۲) و (۳۳) - که به صورت تفاسیل یک تابع محدب  $(\psi_m(x))$  یا  $(\Phi_m(x))$  و خطی  $(t_m)$  بیان شده‌اند - خاصیت تحدب را دارند (اثبات در پیوست آورده شده است). بنابراین جواب بهینه مسئله (۳۴) با حل مسئله معادل محدب (۲۰) به دست خواهد آمد.

اکنون با تعریف متغیر جدید  $t$  و اعمال تبدیل پیشنهادی فوق، مدل محدب پیشنهادی مسئله به صورت زیر حاصل می‌شود:

$$(x_r, y_r, z_r, p_{br}, d_{br}) = \operatorname{argmax}(t_b)$$

s.t.

$$t_b = \frac{e_b}{\sum_{j \neq f} p_{bj} r_{bj} + p_r \sum_{k \in \text{SUR}}^{k \neq b} r_{kb} + (r_{br} \times p_{br})}$$

$$t_b \leq \max(\tau_k), \forall k \in \{N_b\},$$

$$x_r = x_b + \frac{x_f - x_b}{\|\vec{f} - \vec{b}\|} \times d_{br}$$

$$y_r = y_b + \frac{y_f - y_b}{\|\vec{f} - \vec{b}\|} \times d_{br}$$

$$z_r = z_b + \frac{z_f - z_b}{\|\vec{f} - \vec{b}\|} \times d_{br}$$

$$p_{br} = p_s + \alpha^{d_{br}} d_{br}^\gamma, 0 \leq d_{br} < d_t$$

$$p_{br} = p_s + \alpha^{d_{cr}} d_{cr}^\xi, d_t \leq d_{cr} \leq RT$$

$$t = (x_b - x_r)^\gamma + (y_b - y_r)^\gamma + (z_b - z_r)^\gamma$$

$$d_{br}^\gamma = t \quad (31)$$

بنابراین، مسئله جایگذاری گره رله فوق به دسته مسائل برنامه‌ریزی محدب تعلق دارد، در این مقاله برای حل آن از جعبه ابزار CVX موجود در نرم افزار متلب استفاده شده است.

قیود آفینی<sup>۱</sup> (بیان شده در (۲۶) - (ج.۲۷))، قیود محدب (بیان شده در (۲۸) و (۲۹)) و تفاضل محدب<sup>۲</sup> (قید (۳۰)) است. وجود قید غیر محدب (۳۱)، منجر به غیرمحدب شدن مسئله فوق می‌شود که حصول به مکان بهینه گره رله را دشوار می‌سازد. از این‌رو، جهت غلبه بر این چالش، در ادامه ابتدا فرم کلی این گونه مسائل شناخته و معرفی می‌شوند، سپس با پیشنهاد یک تبدیل جدید مسئله فوق به همتای محدب خود تبدیل می‌شود.

## b. مسئله غیرمحدب جایگذاری گره‌های رله و معادل

### محدب پیشنهادی

در این بخش، ابتدا گروه خاصی از مسائل غیرمحدب شامل تابع هدف محدب، قیود محدب (یا آفینی) و تفاضل محدب معرفی و شرح داده می‌شوند. بعد از آن با پیشنهاد و اعمال یک تبدیل جدید بر این مسائل، به معادل محدب آن‌ها دست خواهیم یافت. فرم کلی مسائل شامل تابع هدف محدب، قیود محدب و تفاضل محدب به صورت زیر فرمول‌بندی می‌شود:

$$x = \operatorname{argmin} f.(x)$$

s.t.

$$f_i(x) \leq 0, i = 1, 2, \dots, I$$

$$h_m(x) = 0, m = 1, 2, \dots, M \quad (25)$$

به گونه‌ای که  $x$  نشان‌دهنده بردار متغیرها و  $f_i(x), i = 0, 1, \dots, I$  توابعی محدب یا آفینی هستند. همچنین،  $h_m(x), m = 1, 2, \dots, M$  توابع تفاضل محدب هستند و به صورت زیر فرمول‌بندی می‌شوند:

$$h_m(x) = \Phi_m(x) - \psi_m(x) \quad (26)$$

همان‌گونه که از نام این نوع توابع مشخص است، دو تابع  $\psi_m(x)$  و  $\Phi_m(x)$  محدب هستند و اصطلاحاً اجزای محدب تابع  $h_m(x)$  نامیده می‌شوند [۲۸].

با پیشنهاد و معرفی متغیر جدید  $t = [t_1, \dots, t_M] \in \mathbb{R}^M$  مسئله (۲۵) به صورت زیر می‌تواند فرمول‌بندی شود:

$$(x, t) = \operatorname{argmin} f.(x) \quad (27)$$

s.t.

<sup>۱</sup> Affine constraints

<sup>۲</sup> Difference convex

$$\begin{aligned} dx_b &\approx dx_r \approx dx_f, \\ dy_b &\approx dy_r \approx dy_f, \\ dz_b &\approx dz_r \approx dz_f \end{aligned} \quad (۳۵)$$

حال می‌توان گفت:

$$d_{br}' \approx d_{br} \quad (۳۶)$$

به طریق مشابه برای تغییرات صورت گرفته در  $\|\vec{f} - \vec{b}\|$  قابل اثبات است. با جایگذاری شرط (۴۵) و (۴۶) در روابط (۴۲)، (۴۳) و (۳۷) داریم:

$$\begin{aligned} x_r &= x_b + \frac{x_f - x_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \\ y_r &= y_b + \frac{y_f - y_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \\ z_r &= z_b + \frac{z_f - z_b}{\|\vec{f} - \vec{b}\|} \times d_{br} \end{aligned}$$

بنابراین، با توجه به هندسه مسئله طرح شده و نکات عملی بیان شده، می‌توان گفت که تغییرات مکانی تاثیر چندانی بر عملکرد روش پیشنهادی نخواهد داشت.

در ادامه به ارزیابی روش پیشنهادی جایگذاری گره رله (ERS) از لحاظ پیچیدگی و عملکرد پرداخته می‌شود و با روش ابتکاری RA موجود در [۲۰] مقایسه می‌شود.

## ۵. ارزیابی عملکرد

در این بخش ابتدا با تحلیل پیچیدگی زمانی روش پیشنهادی و پس از آن با ارائه و تشریح نتایج حاصل از شبیه‌سازی، به ارزیابی عملکرد روش پیشنهادی پرداخته می‌شود.

### a. تحلیل پیچیدگی

مرسوم‌ترین روش برای به دست آوردن مرتبه پیچیدگی یک الگوریتم تعیین تعداد محاسبات آن شامل جمع، ضرب، تقسیم، جستجوی خطی و انتساب است. محاسبات روش پیشنهادی ERS در جدول ۲ آورده شده است. اگر فرض شود که تعداد گره‌های ح سگر و رله برابر  $N$  و  $M$  باشد - با توجه به جدول ۱ - پیچیدگی روش پیشنهادی ERS به مرتبه زیر محدود می‌شود:

$$(N + M)^2 + (N + M) \times iter \quad (۳۸)$$

به‌گونه‌ای که  $iter$  همان تعداد تکرار اجرای مسئله بهینه‌سازی برای رسیدن به جواب بهینه است. با توجه به این که تعداد تکرار نسبت به  $N+M$  کمتر است، بنابراین مرتبه پیچیدگی برابر  $(N + M)^2$  است. طبق محاسبات [۲۰] پیچیدگی روش RA برابر همین مرتبه

## ۴. بررسی بهره‌وری روش پیشنهادی جایگذاری

### گره رله در برابر تغییرات محیط زیرآبی

به منظور حفظ کارایی شبکه‌های حسگر زیرآبی، گره‌های زیرآبی بایستی نسبت به تغییرات شدید زیرآبی و وضعیت پایداری داشته باشند. در حقیقت، تغییرات مکانی زیاد موجب خارج شدن گره‌ها از دید مخابراتی یکدیگر و مختل شدن کل شبکه حسگر می‌شود. در نتیجه این امر، نظارت بر محیط زیرآبی سخت و گاهی غیر ممکن خواهد بود. برای این منظور تمهیداتی چون توسعه سخت افزار گره‌ها و استفاده از گره‌های نیمه ساکن به جای گره‌های شناور آزاد صورت گرفته است. بنابراین، تغییر مکان گره‌های زیرآبی به صورت جزئی و در یک ناحیه محدود فرض می‌شود [۳۰].

از آن جا که مکان گره نوعی  $i$  توسط سه مولفه  $(x_i, y_i, z_i)$  مشخص می‌گردد، جابجایی مکانی آن به صورت  $(dx_i, dy_i, dz_i)$  قابل بیان است. بنابراین،  $(dx_r, dy_r, dz_r)$ ،  $(dx_b, dy_b, dz_b)$  و  $(dx_f, dy_f, dz_f)$  به ترتیب بیان گر تغییر مکان گره رله  $(r)$ ، گره بحرانی  $(b)$ ، دورترین همسایه گره بحرانی  $(f)$  هستند. از طرفی طبق روش پیشنهادی، مکان گره رله توسط سه رابطه (۳۵)، (۳۶) و (۳۷) حاصل می‌شود. با در نظر گرفتن جابجایی هر یک از گره‌ها روابط مذکور به صورت زیر بازنویسی خواهند شد:

$$x_r + dx_r = x_b + dx_b \quad (۳۲)$$

$$+ \frac{x_f + dx_f - x_b - dx_b}{\|\vec{f} - \vec{b}\|}$$

$$\times d_{br}'$$

$$y_r + dy_r = y_b + dy_b \quad (۳۳)$$

$$+ \frac{y_f + dy_f - y_b - dy_b}{\|\vec{f} - \vec{b}\|}$$

$$\times d_{br}'$$

$$z_r + dz_r = z_b + dz_b \quad (۳۴)$$

$$+ \frac{z_f + dz_f - z_b - dz_b}{\|\vec{f} - \vec{b}\|}$$

$$\times d_{br}'$$

بگونه‌ای که:

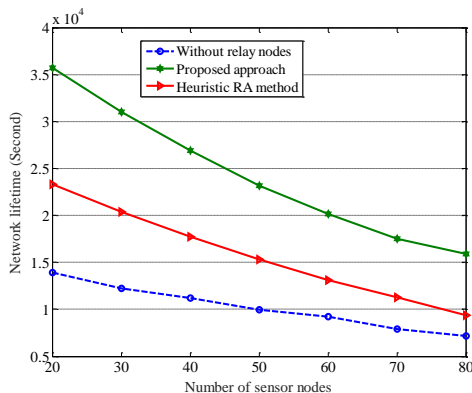
$$d_{br}' = \sqrt{(x_b + dx_b - x_r - dx_r)^2 + (y_b + dy_b - y_r - dy_r)^2 + (z_b + dz_b - z_r - dz_r)^2}$$

با توجه به همگن بودن شبکه ح سگر صوتی زیرآبی (طبق فرض ۳ مقاله) و قرارگیری سه گره  $f$ ،  $b$  و  $r$  در برد مخابراتی یکدیگر (فاصله نسبتاً نزدیک)، می‌توان بیان کرد:

انرژی اولیه گره‌ها	۴۰۰ کیلوژول
نرخ تولیدی هر گره حسگر	۲۰۰-۱۰ بیت بر ثانیه
فاصله آستانه	۸۷ متر

در اولین بخش ارزیابی، عملکرد روش پیشنهادی در سائیزهای (تعداد گره‌های حسگر) مختلف شبکه بررسی شده است. از این رو در شرایطی کاملاً یکسان به جایگذاری تعداد یکسانی از گره‌های رله (۳۰٪ تعداد گره‌های حسگر) به دو روش RA و پیشنهادی (ERS) پرداخته شده است. نتایج این بررسی در شکل ۳ نشان داده شده است. همان‌گونه که از شکل ۳ مشخص است، با افزایش تعداد گره‌های حسگر طول عمر شبکه کاهش می‌یابد. در حقیقت، افزایش تعداد گره‌های حسگر در شبکه منجر به افزایش حجم اطلاعات ارسالی توسط گره بحرانی شبکه و در نتیجه افزایش انرژی مصرفی آن می‌شود؛ این مسئله باعث کاهش طول عمر شبکه می‌شود. به علاوه، همان‌گونه که مشخص است عملکرد روش پیشنهادی ERS به مراتب از روش RA بهتر است. در حقیقت، روش پیشنهادی با بکارگیری مسئله بهینه‌سازی ریاضی و تبدیل آن به یک مسئله محذب توانایی حصول به مکان بهینه گره‌های رله را دارد.

در مقابل، طبق روش ابتکاری RA ابتدا گره‌های رله روی سطح آب جایگذاری می‌شوند و سپس عمق آن‌ها تنظیم می‌شود. جدا سازی مختصات گره رله از سویی و قرارگیری گره‌های رله در مکان‌های پرت از سویی دیگر منجر به عملکرد شبه‌بهینه روش RA شده است.



شکل ۳. بررسی عملکرد روش پیشنهادی برای تعداد مختلف گره حسگر

لازم به ذکر است، در غیاب گره رله در شبکه به دلیل فاصله مخابراتی طولانی بین گره‌های حسگر طول عمر شبکه در پایین‌ترین

است. در ادامه به بررسی عملکرد و طول عمر روش پیشنهادی با حفظ مرتبه پیچیدگی پرداخته می‌شود.

جدول ۲. محاسبات روش پیشنهادی ERS

نوع عملگر	مرتبه اجرایی	دفعات محاسبه	مرتبه پیچیدگی
جستجوی خطی	۱	$2 \times (N + M) - 5$	$N + M$
انتساب	۱	$4 + (6 \times iter)$	$iter$
جمع	۱	$(N + M - 1)^2 + (N + M + 14) \times iter$	$(N + M)^2 + (N + M) \times iter$
تقسیم	۱	$(N + M - 1) + 2 \times iter$	$(N + M) + 2 \times iter$
ضرب	۱	$(N + M - 1)^2 + (N + M + 15) \times iter$	$(N + M)^2 + (N + M) \times iter$

### b. شبیه‌سازی و تحلیل نتایج روش پیشنهادی

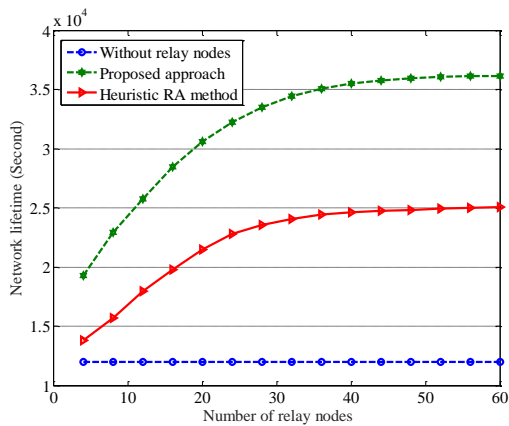
در این بخش، عملکرد روش پیشنهادی با روش ابتکاری RA مقایسه می‌شود. پیش از تشریح کامل بررسی‌های صورت گرفته، پارامترهای مسئله باید تعیین گردند: عمق آب ۲۰۰۰ متر فرض شده است، نرخ تولیدی هر گره حسگر به صورت تصادفی از ۱۰ تا ۲۰۰ بیت بر ثانیه تنظیم شده است، به علاوه انرژی اولیه هر گره ۴۰۰ کیلو ژول انتخاب شده است [۲۰]. به طور کلی تمام پارامترهای معین در جدول ۳ آورده شده است. در ادامه به تشریح بررسی‌های صورت گرفته، پرداخته می‌شود.

جدول ۳. پارامترهای شبیه‌سازی

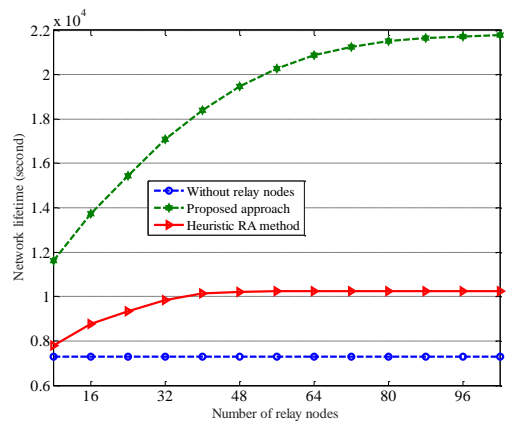
پارامتر	مقدار
عمق آب	۲۰۰۰ متر
برد مخابراتی گره‌ها	۵۰۰ متر
فرکانس سیگنال صوتی	۱ کیلوهرتز
توان مصرفی برای پردازش هر بیت داده ارسالی	۱ میلی‌وات
توان مصرفی برای پردازش هر بیت داده دریافتی	۱ میلی‌وات
ظرفیت پیوند مخابراتی	۱۰ کیلوبیت بر ثانیه

مقدار قرار دارد.

در آزمایش دوم، اثر افزایش تعداد گره‌های رله بر طول عمر شبکه بررسی شده است. برای این منظور، طول عمر شبکه بر حسب تعداد گره رله رسم شده است. این بررسی برای سه سایز شبکه (۲۰، ۴۰ و ۸۰) انجام شده است. در هر سایز، تعداد گره‌های رله از ۰.۱ تا ۱.۸ برابر تعداد گره‌های حسگر تغییر می‌کند. نتایج این ارزیابی در شکل ۴ آورده شده است. همان‌گونه که از شکل ۴ برمی‌آید، در هر دو نوع روش جایگذاری گره رله (روش ابتکاری RA و روش پیشنهادی) با افزایش تعداد گره‌های رله طول عمر شبکه افزایش می‌یابد، اما عملکرد روش پیشنهادی به مراتب از روش ابتکاری RA بیشتر است. همان‌گونه که پیش از این بیان شد، در روش RA ابتدا مکان اولیه گره‌های رله (مختصات XY) روی سطح آب مشخص می‌شود و پس از آن بدون عوض کردن دو مختصه اولیه، عمق آن‌ها تنظیم می‌شود. در حالی که در روش پیشنهادی مکان‌دهی گره رله به صورت یک مسئله بهینه‌سازی ریاضی فرمول‌بندی شده است که با استفاده از یک تبدیل به همتای محدب خود تبدیل شده است. واضح است که پیدا کردن هم‌زمان سه مختصه گره رله تحت یک مسئله بهینه‌سازی محدب باعث عملکرد بالای روش پیشنهادی شده است.



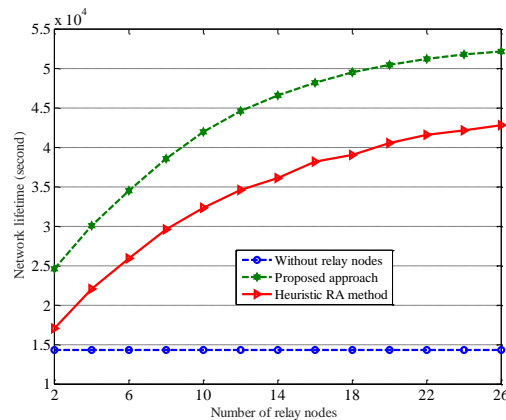
ب. تعداد ۴۰ گره حسگر



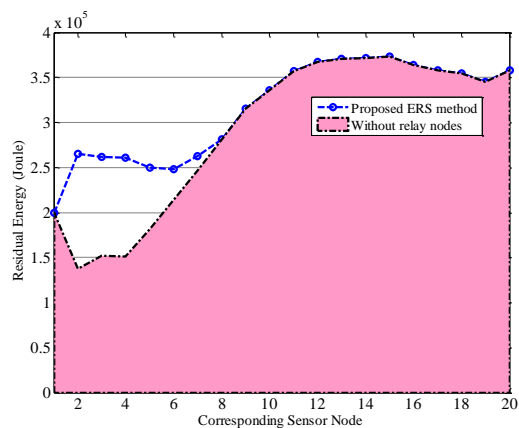
پ. تعداد ۸۰ گره حسگر

شکل ۴: بررسی عملکرد روش پیشنهادی در تعدادهای مختلف گره رله، به ازای سایزهای مختلف شبکه

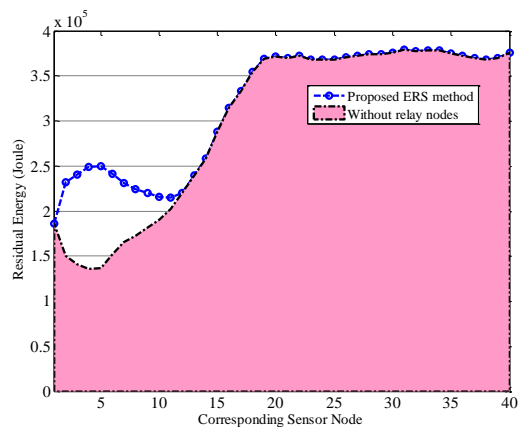
مسئله مهم دیگری که باید مورد توجه قرار داد عملکرد روش‌های جایگذاری در سایزهای مختلف شبکه است. به طور مثال روش ابتکاری RA در سایز بالای شبکه (۸۰) عملکرد بسیار ضعیفی دارد و طول عمر بسیار سریع به بالاترین مقدار خود می‌رسد. در حقیقت، در روش RA به هر میزان که تعداد گره‌های حسگر افزایش می‌یابد احتمال قرارگیری گره رله بین گره‌های حسگر کاهش می‌یابد. بدین معنی که، با افزایش تعداد گره‌های حسگر فاصله بین آن‌ها کاهش می‌یابد و گره‌های رله‌ای که روی سطح آب ثابت شده‌اند کارایی خود را در کاهش فاصله مخابراتی بین گره‌های حسگر از دست می‌دهند. در حالی که در روش پیشنهادی مکان بهینه گره رله تحت یک مسئله برنامه‌ریزی ریاضی به منظور افزایش طول عمر شبکه به



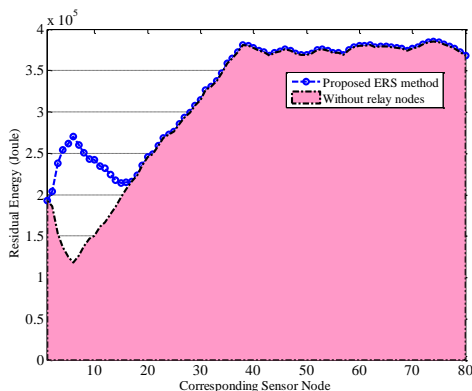
آ. تعداد ۲۰ گره حسگر



آ. تعداد ۲۰ گره حسگر



ب. تعداد ۴۰ گره حسگر



پ. تعداد ۸۰ گره حسگر

شکل ۵. بررسی عملکرد روش پیشنهادی در رفع حفره انرژی

دست می‌آید. لازم به ذکر است که رسیدن به حداکثر عملکرد زمانی اتفاق می‌افتد که طول عمر گره بحرانی به شدت پایین است و گره حسگر روی آن قرار می‌گیرد.

در آخرین ارزیابی، توانایی دو روش RA و ERS پیشنهادی در حل مسئله حفره انرژی در شبکه حسگر صوتی زیرآبی سه‌بعدی بررسی شده است. برای این منظور، برای سایزهای مختلف شبکه (۲۰، ۴۰ و ۸۰ گره حسگر) انرژی باقی‌مانده گره‌های حسگر بررسی شده است. همچنین، عملکرد هر یک از روش‌های RA و روش پیشنهادی در رفع حفره انرژی ناشی از مصرف نامتعادل انرژی مورد ارزیابی قرار گرفته است. نتایج این بررسی در شکل‌های ۵ و ۶ آورده شده است. همان‌گونه که مشخص است، هر دو روش جایگذاری، سعی در متعادل کردن مصرف انرژی بین گره‌های حسگر و رفع حفره انرژی دارند، اما روش پیشنهادی ERS در حل این قضیه مطلوب‌تر عمل می‌کند. در حقیقت، روش پیشنهادی با حفظ اتصال گره بحرانی و همسایگان آن امکان دستیابی به مصرف توان متعادل‌تر بین گره‌های حسگر را دارد.

## ۶. نتیجه‌گیری

در این مقاله مسئله جایگذاری گره رله به منظور افزایش طول عمر شبکه‌های حسگر صوتی زیرآبی مورد مطالعه قرار گرفته است. به منظور قرارگیری بهینه گره‌های رله در شبکه حسگر زیرآبی، روشی به نام ERS پیشنهاد شده است، به گونه‌ای که مکان بهینه گره‌های رله در دو گام مشخص شده است. در گام اول مسئله جایگذاری به صورت یک مسئله برنامه‌ریزی غیرمحدب فرمول‌بندی شده است و در گام بعد با معرفی یک تبدیل جدید مسئله غیرمحدب مذکور به معادل محدب خود تبدیل شده است. نتایج شبیه‌سازی نشان‌دهنده عملکرد مناسب روش پیشنهادی از لحاظ رفع حفره انرژی و طول عمر شبکه است.

مراجع

[۱] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, "A survey of architectures and localization techniques for underwater acoustic sensor networks," *IEEE Communications Surveys & Tutorials*, vol. ۱۳, no. ۳, pp. ۴۸۷-۵۰۲, ۲۰۱۱.

[۲] S. Climent, A. Sanchez, J. V. Capella, N. Meratnia, and J. J. Serrano, "Underwater acoustic wireless sensor networks: advances and future trends in physical, MAC and routing layers," *Sensors*, vol. ۱۴, no. ۱, pp. ۷۹۵-۸۳۳, ۲۰۱۴.

[۳] G. Qiao, Z. Babar, L. Ma, S. Liu, and J. Wu, "MIMO-OFDM underwater acoustic communication systems—A review," *Physical Communication*, vol. ۲۳, pp. ۵۶-۶۴, ۲۰۱۷.

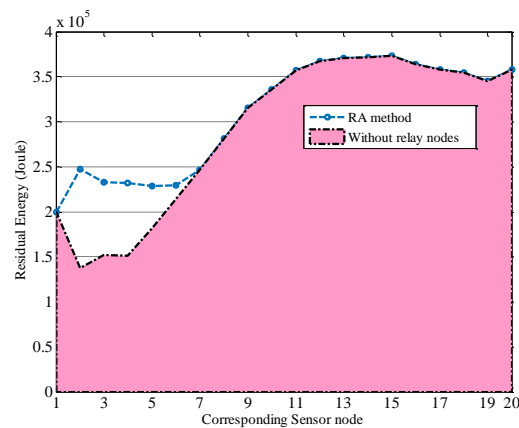
[۴] M. Patil and R. C. Biradar, "A survey on routing protocols in wireless sensor networks," in *Networks (ICON), ۲۰۱۲ ۱۸th IEEE International Conference on*, ۲۰۱۲, pp. ۸۶-۹۱: IEEE.

[۵] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad hoc networks*, vol. ۳, no. ۳, pp. ۲۵۷-۲۷۹, ۲۰۰۵.

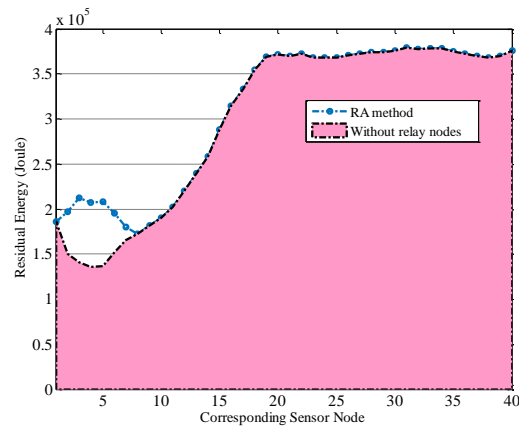
[۶] A. Wadaa, S. Olariu, L. Wilson, M. Eltoweissy, and K. Jones, "Training a wireless sensor network," *Mobile Networks and Applications*, vol. ۱۰, no. ۱-۲, pp. ۱۵۱-۱۶۸, ۲۰۰۵.

[۷] E. L. Lloyd and G. Xue, "Relay node placement in wireless sensor networks," *IEEE Transactions on Computers*, vol. ۵۶, no. ۱, pp. ۱۳۴-۱۳۸, ۲۰۰۷.

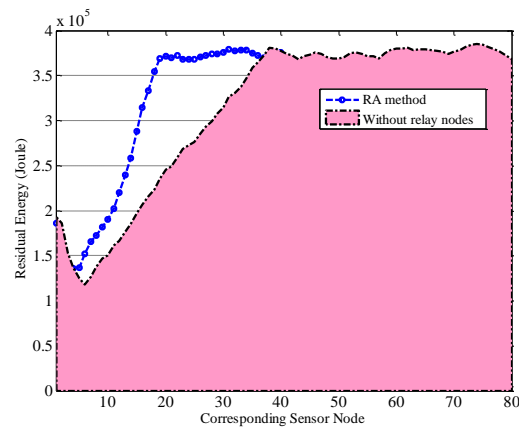
[۸] S. Misra, S. D. Hong, G. Xue, and J. Tang, "Constrained relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Transactions on Networking (TON)*, vol. ۱۸, no. ۲, pp. ۴۳۴-۴۴۷, ۲۰۱۰.



آ. تعداد ۲۰ گره حسگر



ب. تعداد ۴۰ گره حسگر



ت. تعداد ۸۰ گره حسگر

شکل ۶. بررسی عملکرد روش ابتکاری RA در رفع حفره انرژی

- [۱۸] H. U. Yıldız, "Improvement of underwater acoustic sensor networks performance with fountain codes," in *۲۰۱۹ ۲۷th Signal Processing and Communications Applications Conference (SIU)*, ۲۰۱۹, pp. ۱-۴: IEEE.
- [۱۹] B. Das, E. S. Mishra, and S. K. Sethi, "Enhancement of Lifetime of Acoustic Sensor using PEGASIS Algorithm in UASN," *International Journal of Electronics, Electrical and Computational System (Mc Graw Hill Publication)*, vol. ۶, no. ۹, pp. ۵۳۴-۵۴۵, ۲۰۱۷.
- [۲۰] L. Liu, M. Ma, C. Liu, and Y. Shu, "Optimal relay node placement and flow allocation in underwater acoustic sensor networks," *IEEE Transactions on Communications*, vol. ۶۵, no. ۵, pp. ۲۱۴۱-۲۱۵۲, ۲۰۱۷.
- [۲۱] Z. Mohammadi, M. Soleimanpour-Moghadam, M. Askarizadeh, and S. Talebi, "Increasing the Lifetime of Underwater Acoustic Sensor Networks: Difference Convex Approach," *IEEE Systems Journal*, ۲۰۲۰.
- [۲۲] Z. Mohammadi, M. Soleimanpour-Moghadam, S. Talebi, and D. Abbasi-Moghadam, "A new optimization algorithm for relay node setting in underwater acoustic sensor networks," in *۲۰۱۸ ۳rd Conference on Swarm Intelligence and Evolutionary Computation (CSIEC)*, ۲۰۱۸, pp. ۱-۵: IEEE.
- [۲۳] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, ۲۰۰۴.
- [۲۴] C.-Y. Chi, W.-C. Li, and C.-H. Lin, *Convex optimization for signal processing and communications: from fundamentals to applications*. CRC press, ۲۰۱۷.
- [۲۵] W. H. Thorp, "Analytic description of the low-frequency attenuation coefficient," *The Journal of the Acoustical Society of*
- [۹] K. Lu, G. Liu, R. Mao, and Y. Feng, "Relay node placement based on balancing power consumption in wireless sensor networks," *IET wireless sensor systems*, vol. ۱, no. ۱, pp. ۱-۶, ۲۰۱۱.
- [۱۰] A. Wahid, S. Lee, and D. Kim, "An energy-efficient routing protocol for UWSNs using physical distance and residual energy," in *OCEANS ۲۰۱۱ IEEE-Spain*, ۲۰۱۱, pp. ۱-۶: IEEE.
- [۱۱] A. Wahid, S. Lee, and D. Kim, "A reliable and energy-efficient routing protocol for underwater wireless sensor networks," *International Journal of Communication Systems*, vol. ۲۷, no. ۱۰, pp. ۲۰۴۸-۲۰۶۲, ۲۰۱۴.
- [۱۲] N. Javaid, M. Shah, A. Ahmad, M. Imran, M. I. Khan, and A. V. Vasilakos, "An enhanced energy balanced data transmission protocol for underwater acoustic sensor networks," *Sensors*, vol. ۱۶, no. ۴, p. ۴۸۷, ۲۰۱۶.
- [۱۳] M. Akbar, N. Javaid, A. H. Khan, M. Imran, M. Shoaib, and A. Vasilakos, "Efficient data gathering in ۳D linear underwater wireless sensor networks using sink mobility," *Sensors*, vol. ۱۶, no. ۳, p. ۴۰۴, ۲۰۱۶.
- [۱۴] J. U. Khan and H.-S. Cho, "A distributed data-gathering protocol using AUV in underwater sensor networks," *Sensors*, vol. ۱۵, no. ۸, pp. ۱۹۳۳۱-۱۹۳۵۰, ۲۰۱۵.
- [۱۵] K. Wang, H. Gao, X. Xu, J. Jiang, and D. Yue, "An energy-efficient reliable data transmission scheme for complex environmental monitoring in underwater acoustic sensor networks," *IEEE Sensors Journal*, vol. ۱۶, no. ۱۱, pp. ۴۰۵۱-۴۰۶۲, ۲۰۱۶.
- [۱۶] X. Du, K. Li, X. Liu, and Y. Su, "RLT code based handshake-free reliable MAC protocol for underwater sensor networks," *Journal of Sensors*, vol. ۲۰۲۰۱۶, ۱۶.
- [۱۷] H. U. Yıldız, "Maximization of underwater sensor networks lifetime via fountain codes," *IEEE Transactions on Industrial*

- [۲۹] C. Detweiler, M. Doniec, I. Vasilescu, and D. Rus, "Autonomous depth adjustment for underwater sensor networks: Design and applications," *IEEE/ASME Transactions on Mechatronics*, vol. ۱۷, no. ۱, pp. ۱۶-۲۴, ۲۰۱۱.
- [۳۰] N. Nowsheen, G. Karmakar, and J. Kamruzzaman, "PRADD: A path reliability-aware data delivery protocol for underwater acoustic sensor networks," *Journal of Network and Computer Applications*, vol. ۷۵, pp. ۳۸۵-۳۹۷, ۲۰۱۶.
- [۳۱] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, ۲۰۰۴.
- America, vol. ۴۲, no. ۱, pp. ۲۷۰-۲۷۰, ۱۹۶۷.
- [۲۶] J. Cao, J. Dou, and S. Dong, "Balance transmission mechanism in underwater acoustic sensor networks," *International Journal of Distributed Sensor Networks*, vol. ۱۱, no. ۳, p. ۴۶۹۳۴۰, ۲۰۱۵.
- [۲۷] N. Javaid, M. R. Jafri, Z. A. Khan, U. Qasim, T. A. Alghamdi, and M. Ali, "Iamctd :Improved adaptive mobility of courier nodes in threshold-optimized dbr protocol for underwater wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. ۱۰, no. ۱۱, p. ۲۱۳۰۱۲, ۲۰۱۴.
- [۲۸] R. Horst and N. V. Thoai, "DC programming: overview," *Journal of Optimization Theory and Applications*, vol. ۱۰۳, no. ۱, pp. ۱-۴۳, ۱۹۹۹.

همچنین، با توجه به تعریف تابع محدب، عبارت زیر برای تابع محدب  $g(x)$  صدق می‌کند:

$$g(\theta x + (1 - \theta)y) \leq \theta g(x) + (1 - \theta)g(y) \quad (۴۱)$$

با تفریق مساوی (۴۰) از نام مساوی (۴۱) عبارت زیر حاصل می‌شود:

$$g(\theta x + (1 - \theta)y) - f(\theta x + (1 - \theta)y) \leq \theta g(x) + (1 - \theta)g(y) - (\theta f(x) + (1 - \theta)f(y)) \quad (۴۲)$$

بنابراین،

$$k(\theta x + (1 - \theta)y) \leq \theta k(x) + (1 - \theta)k(y) \quad (۴۳)$$

در نتیجه، تابع  $k(x)$  محدب است و مسئله یک مسئله برنامه‌ریزی محدب است.

### پیوست: اثبات تحدب دو قید (۲.۲۰) و

(۲.۲۰) ث)

همان گونه که می‌دانیم، برای بررسی تحدب تابع دلخواه  $k(x)$  شرط باید زیر بررسی شود:

$$k(\theta x + (1 - \theta)y) \leq \theta k(x) + (1 - \theta)k(y) \quad (۳۹)$$

به طوری که  $x$  و  $y$  نقاطی در دامنه تابع  $k(x)$  باشند و  $0 \leq \theta \leq 1$  [۳۱].  $k(x)$  تحت دو قید (۲.۲۰) و (۲.۲۰) ث)

به صورت زیر قابل بیان هست:

$$k(x) = g(x) - f(x)$$

به گونه‌ای که  $f(x)$  و  $g(x)$  به ترتیب بیان‌گر توابع محدب و خطی هستند. واضح است که در تابع خطی  $f(x)$  شرط زیر صدق می‌کند:

$$f(\theta x + (1 - \theta)y) = \theta f(x) + (1 - \theta)f(y) \quad (۴۰)$$



بیشینه‌سازی طول عمر شبکه‌های حسگر صوتی زیرآبی با جایگذاری بهینه گره‌های رله

## شناسایی عوامل و شاخص‌های مدل سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر

محمدتقی تقوی فرد\* محمدعلی عباسپور طهرانی فرد\*\* کامران فیضی\*\*\* رسول نورالسنا\*\*\*\*

\*دانشیار دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی

\*\*دانشجوی دکتری دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی

\*\*\*استاد دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبایی

\*\*\*\*استاد دانشکده مهندسی صنایع، دانشگاه علم و صنعت

تاریخ دریافت: ۱۳۹۹/۰۵/۲۱ تاریخ دریافت: ۱۳۹۹/۰۹/۱۹

نوع مقاله: پژوهشی

### چکیده

یکی از مهم‌ترین بخش‌های رسانه‌ای و سرگرمی در دنیای امروز، صنعت بازی‌های رایانه‌ای است. تعداد کاربران این نوع بازی‌ها به‌صورت مستمر در حال افزایش است و کسب‌وکارهای این صنعت رشد چشمگیری کرده‌اند؛ به‌طوری که این رشد گردش مالی میلیون دلاری را برای شرکت‌های توسعه‌دهنده این نوع بازی‌ها به دنبال داشته است. از این رو، هدف اصلی این پژوهش ارائه مدلی جهت سنجش کیفیت خدمات بر بستر ابر در صنعت بازی‌های رایانه‌ای است. پژوهش حاضر از نظر هدف کاربردی-توسعه‌ای است و جزء تحقیقات کیفی به حساب می‌آید. همچنین این پژوهش از منظر شیوه گردآوری اطلاعات تحقیق «توصیفی-غیرآزمایشی» محسوب می‌شود. در این پژوهش برای سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر، پس از بررسی ادبیات نظری پژوهش و مدل‌های موجود و انجام روش دلفی، ۵۲ شاخص در قالب ۹ بعد اصلی شناسایی شدند. نتایج پژوهش مدلی جامع برای سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر ارائه می‌کند که می‌تواند مورد استفاده صنعت گران و پژوهشگران قرار گیرد.

**واژگان کلیدی:** کیفیت خدمات، مدل‌های سنجش کیفیت خدمات، خدمات مبتنی بر ابر، صنعت بازی‌های رایانه‌ای.

### ۱- مقدمه

که در سال‌های اخیر با پیشرفت‌های تکنولوژیکی به‌سرعت در حال افزایش است. بسیاری از محققان معتقدند بازی‌های ابری نسل آینده صنعت بازی هستند [۳].

بازی‌های ویدئویی مبتنی بر ابر، راه‌حل‌های مقرون‌به‌صرفه، انعطاف‌پذیر و با کارایی بالا را برای کاربران نهایی با منابع محدود محاسباتی ارائه می‌دهند و آن‌ها را قادر می‌سازد تا بازی‌های گرافیکی بالا را با حداقل سیستم موردنیاز تجربه کنند. از آنجایی که بخش مهمی از پردازش محاسباتی، از جمله رندر بازی و فشرده‌سازی ویدئو، در مراکز داده انجام می‌شود، کنترل انتقال اطلاعات در دیتابیس تأثیر مهمی در کیفیت خدمات بازی ابر دارد [۲].

بازی ابری سرویس نوظهوری است که اخیراً در صنعت بازی رشد فراوانی داشته است و در میان محصولات سودآور در بازار نرم‌افزار می‌باشد [۱] و [۲]. امروزه، شرکت‌های بازی‌سازی به استفاده از فرصت‌های میزبانی برنامه‌های کاربردی بازی ابری (بازی ابری) روی آوردند [۱]. بازی ابری، با استفاده از در اختیار داشتن داده‌های بازی ذخیره‌شده در پایگاه داده‌های ابر، محاسبات فزاینده‌ای را به ابر منتقل می‌کند و به بازیکنان امکان دسترسی کامل به محیط بازی شخصی خود را از هر دستگاه (کامپیوتر شخصی، تلفن همراه و غیره) و عملاً در هر جا می‌دهد. این یک روند نسبتاً جدید در صنعت بازی است

کاربران از سوی دیگر، صنعت بازی های رایانه ای را دستخوش تحولات فراوانی کرده است.

با توجه به فضای رقابتی موجود در صنعت بازی های رایانه ای، لزوم دستیابی به سطوح بالای کیفیت برای سازمان های ارائه دهنده بازی های رایانه ای بسیار ضروری به نظر می رسد. از این رو، سازمان ها ناچارند تا کیفیت خدمات ارائه شده به کاربران را مورد بررسی، سنجش و تحلیل قرار دهند.

مدل های موجود برای ارزیابی کیفیت خدمات (مدل های SERVQUAL، SERVPERF و CASERV) و دیگر مدل های کیفیت خدمات سنتی) پاسخگوی جنبه های نوین بازی ابر نیستند. به همین دلیل، این مطالعه به دنبال بررسی ارائه یک رویکرد جایگزین برای اندازه گیری درک مشتری از کیفیت خدمات در صنعت بازی می باشد. این پژوهش با مرور مدل های سنتی و مدرن، زمینه هایی را برای تدوین یک ابزار دقیق تر برای ارزیابی کیفیت خدمات در صنعت بازی فراهم می کند. از این رو، با توجه به مرور ادبیات صورت گرفته توسط محقق، کمبود مدل های منسجم و بنیادی برای حمایت از به کارگیری خدمات رایانش ابری در توسعه محصولات بنگاه های بازی سازی احساس می گردد. در این راستا این پژوهش به دنبال پاسخگویی به این سؤالات است که «مفاهیم بنیادی مدل سنجش کیفیت خدمات صنعت بازی های رایانه ای بر بستر ابر چیست؟» و «ابعاد، شاخص ها و مؤلفه های مدل سنجش کیفیت خدمات صنعت بازی های رایانه ای بر بستر ابر چیست؟».

## ۲- پیشینه پژوهش

### ۲-۱- پیشینه نظری

توسعه بازی یک مفهوم بین رشته ای است که شامل مهندسی نرم افزار، مدیریت و رشته های هنری می باشد. نتایج تحقیقات تجربی نشان می دهد که کیفیت عوامل کسب و کار با عملکرد یک سازمان ارائه دهنده بازی های دیجیتالی ارتباط مثبت دارد [۱۰].

اهمیت کیفیت خدمات ارائه شده به مصرف کنندگان غیرقابل انکار می باشد. مصرف کنندگان به طور جدی به دنبال کیفیت در زندگی و خریدهای خود هستند؛ و در سال های اخیر خواستار دریافت خدمات با کیفیت بالاتر می باشند [۱۱].

امروزه سازمان ها برای فعالیت های روزانه خود به شدت به خدمات فناوری اطلاعات نیازمندند [۱۲]. رایانش ابری مدلی برای دستیابی به شبکه ای همه جایی، راحت و مبتنی بر تقاضا می باشد و بازی به عنوان خدمت مدل ارائه خدمات بازی می باشد که به مشتریان اجازه می دهد تا از بازی های ارائه شده توسط ارائه دهنده بازی بر روی فضای رایانش ابری استفاده نماید [۱۳]. اخیراً، به دلیل دورمان بالا و

معمولاً بازی های کامپیوتری دارای گرافیک بالایی هستند. از این رو، کاربران بازی های کامپیوتری همواره نیاز به سخت افزار به روز و باکیفیت دارند تا بتوانند نیازهای گرافیکی بازی های جدید را برآورده نمایند. این موضوع باعث می شود هر زمان که یک نسخه جدید بازی ناسازگار با سیستمشان منتشر می شود، سخت افزار خود را تغییر دهند. این کار هزینه زیادی را برای کاربر خواهد داشت، زمان تنظیم بازی را افزایش می دهد و دشواری های مربوط به ناسازگاری سخت افزار و نرم افزار را نیز به دنبال خواهد داشت [۴].

در بازی ابر، کاربر می تواند به سخت افزار سیستم های ابر دسترسی پیدا کند و نیاز به تنظیم مجدد سخت افزار خود با انتشار نسخه جدید بازی ندارند. همچنین نگرانی های نصب بازی و یا ناسازگاری نرم افزار با سخت افزار را از بین می برد. یکی دیگر از مزیت های مهم این است که بازی را می توان با اتصال به اینترنت پرسرعت، در هر سیستمی اعم از کامپیوتر، دستگاه های تلفن همراه و حتی در تلویزیون اجرا نمود [۵].

ایده اصلی بازی ابر، این است که همه اجزای بازی از جمله موتور بازی، پردازش هوش مصنوعی و ماژول ها بر روی فضای ابری ارائه می گردد. ورودی های کاربر و فریم های ویدئوی کدگذاری شده بین سرور بازی ابر و بازیکنان بازی منتقل می شود. بازی در سرور اپراتور ذخیره می شود و به طور مستقیم به مشتریانی که به سرور دسترسی دارند، ارائه می گردد. این موضوع به کاربر اجازه می دهد تا با کمترین امکانات به بازی دسترسی داشته باشد و دیگر به ارتقای توانایی کامپیوتر توجه نمی کند. در حال حاضر شرکت هایی نظیر G-Cluster، CiiNOW، OTOY، StreamingMyGame، سونی و مایکروسافت خدمات بازی های ابر را به عموم ارائه می دهند [۶].

تا کنون چندین مدل برای اندازه گیری عملکرد بازی ابر با توجه به پارامترهای سیستم ارائه شده است. زمان تأخیر و کیفیت تصویر سیستم OnLive را مورد سنجش قرار دادند [۸]. روش اندازه گیری دیگری را ارائه شده است که به بررسی تأخیر پاسخ سیستم بازی های رایانه ای می پردازد. در این مقاله، تأخیر پاسخ به سه بخش تقسیم شد: تداخل شبکه، تأخیر پردازش و تأخیر پخش [۹].

امروزه سنت جای خود را به مدرنیته داده است و روند رو به رشد پیشرفت تکنولوژی، جهان امروز همه سازمان ها و به تبع آن زندگی روزمره انسان ها را به چالش کشیده است. در چنین فضایی، سازمان ها ناگزیرند خود را با تکنولوژی های نوین تطبیق دهند. صنعت بازی یکی از صنایع فعال در حوزه تکنولوژی می باشد که همگام با روند رو به رشد تکنولوژی، ناچار است خود را به روز نگه دارد. امروزه بازی های رایانه ای سنتی جای خود را به بازی های ابری داده اند. ایجاد فضای جدید کسب و کار در این صنعت از یک سو و پاسخگویی به نیازهای

در پژوهشی روشی برای ارزیابی ذهنی و عینی و همچنین تجزیه و تحلیل خدمات بازی ابر توسعه دادند. این روش بر اساس یک رویکرد غیر انتزاعی بود که به همین دلیل آن را در انواع مختلف سیستم‌های بازی ابر می‌تواند استفاده کرد. رضایت کاربر از کیفیت تجربه یکی از عوامل مهم در موفقیت خدمات بازی‌های ابر می‌باشد [۲۱].

در پژوهشی مجموعه‌ای از روش‌های اندازه‌گیری ارزیابی کیفیت خدمات در سیستم‌های بازی‌های ابر را ارائه مورد بررسی قرار دادند و اثربخشی طرح‌های را با استفاده از مطالعه موردی دو سیستم شناخته شده بازی ابر یعنی OnLive و StreamMyGame بررسی نمودند. نتایج نشان داد که سیستم OnLive بهتر عمل می‌کند، زیرا که در آن نرخ فریم قابل تنظیم است، کیفیت گرافیکی بهتری دارد و تأخیر پردازش سرور کمتری دارد و پهنای باند کمتری را در شبکه اشغال می‌کند [۲۲].

در پژوهشی نتایجی ذهنی و فیزیولوژیکی ۳۲ بازیکن را در یک بازی تیراندازی مورد بررسی قرار دادند. آن‌ها دریافتند که کیفیت ویدئو بر کیفیت درک شده بازیکن، تجربه بازیکن، ارزش امتیاز و قدرت باند فرکانس آلفا تأثیر می‌گذارد. همچنین نتایج تحقیق نشان داد که (۱) روش‌های ذهنی تنوع کیفی را ارزیابی می‌کنند و (۲) اندازه‌گیری‌های فیزیولوژیکی نفوذ بر بازیکن از لحاظ وضعیت شناختی را کاهش می‌دهد [۲۳].

جنبه‌های مختلف سیستم بازی‌های رایانه‌ای را برای نشان دادن مزایا و قدرت‌های مختلف آن مورد بحث قرار دادند. آن‌ها مفهوم بازی ابر، معماری سیستم ابر و پارامترهای کیفیت سرویس را به طور دقیق مورد بحث قرار دادند. دریافتند که سیستم بازی ابر باید با هدف ارائه بهترین زمان پاسخ و مصرف پهنای باند کم ارائه گردد [۵].

در تحقیقی دو پیشنهاد ارائه نمودند. (۱) طرح‌هایی را برای اعمال پارادایم اخیر شبکه‌های تعریف شده نرم‌افزار برای رایانش ابری پیشنهاد نمودند و (۲) یک کنترل کننده شبکه را پیشنهاد کردند که باعث کاهش تأخیر در بازیابی اطلاعات توسط بازیکنان می‌گردد. نتایج نشان داد که کنترل پیشنهادی آن‌ها، تأخیر را تقریباً ۹ درصد کاهش داده است. این کاهش باعث بهبود در کیفیت تجربه بازی بازیکنان شد [۲].

در پژوهشی به بررسی بهینه‌سازی کیفیت برای سیستم‌های بازی ابری با توجه به چالش‌های موجود، محدودیت‌های سیستم و الزامات خدمات پرداخته شده است. محقق ابتدا به بررسی فن‌آوری‌های فشرده‌سازی بازی‌های ویدئویی با استفاده از سیستم‌های موجود در بازی ابری پرداخت. در گام دوم، محقق یک پارادایم بازی ابری انحصاری جدید را معرفی می‌کند که متشکل از رفت و آمدهای بالای

محبوبیت رایانش ابری، شرکت‌ها بیش از پیش سیستم‌های اطلاعاتی خود را بر روی پایگاه داده رایانش ابری قرار می‌دهند [۱۴].

در سال‌های اخیر، بازی‌های رایانه‌ای بخش عظیمی از سرگرمی‌های مردم را تشکیل می‌دهد. بازی به عنوان خدمت نوید بر طرف شدن محدودیت‌های بازی‌های رایانه‌ای را به کاربران می‌دهد [۱۵]. با مطالعه اسناد و پایگاه‌های استنادی موجود، در این بخش به مرور برخی از مطالعات صورت گرفته در حوزه کیفیت خدمات بازی ابری پرداخته خواهد شد.

در تحقیقی سرویس دهنده‌های دانشگاه دانمارک، لهستان، برزیل و مالزی را مورد بررسی و مقایسه قرار دادند. نتایج نشان داد که بین تأخیر و بازده همبستگی وجود دارد [۱۶].

در پژوهشی به دنبال پاسخگویی به این پرسش بوده‌اند که زمان واقعی تأخیر در سیستم‌های بازی رایانه‌ای ابری چقدر است؟ برای پاسخگویی به این پرسش، محققان زمان تأخیر پاسخگویی را در دو سیستم عامل بازی، یعنی OnLive و StreamMy-Game مورد تحلیل قرار دادند. نتایج نشان داد که زمان تأخیر پاسخگویی سیستم On-Live تقریباً نصف سیستم StreamMyGame می‌باشد [۹]. در پژوهشی تحت عنوان «مدل چندبعدی و سلسله مراتبی از کیفیت خدمات در صنعت بازی» با ایجاد یک چارچوب مفهومی و مقیاس اندازه‌گیری به بررسی کیفیت خدمات در صنعت بازی پرداختند. بر اساس مرور ادبیات گسترده و تحقیق کیفی و تجربی، یک مدل چندبعدی و سلسله مراتبی از کیفیت خدمات برای صنعت بازی ارائه شد. تجزیه و تحلیل داده‌های جمع‌آوری شده از مجموع ۵۱۰ مشتری نشان داد که مدل پیشنهادی به خوبی داده‌ها را در بر می‌گیرد. روایی و قابلیت اطمینان مقیاس اندازه‌گیری با استفاده از پیش‌آزمون بررسی شد [۱۷].

در پژوهشی با استفاده از یک ابزار اندازه‌گیری و مجموعه‌ای از ارائه‌دهندگان بازی‌های ابر، تأخیر در بازی ابری را به صورت کیفی بررسی کردند. آن‌ها دو ارائه‌دهنده بازی و سه بازیکن را مورد بررسی قرار دادند [۱۸].

در پژوهشی به بررسی و تحلیل سیستماتیک سیستم‌عامل‌های پیشرفته بازی ابر پرداختند و برخی از این سیستم‌ها را مقایسه نمودند [۱۹].

در پژوهشی عوامل مؤثر بر کیفیت تجربه بازی‌های تلفن همراه را مورد بررسی قرار دادند. محققان تأثیر نوع بازی، اندازه صفحه‌نمایش و تأخیر شبکه را بر روی ابعاد مختلف ادراک کیفیت تجربه بازیکن بررسی کردند. نتایج نشان داد که اندازه صفحه‌نمایش تأثیر قابل توجهی بر کیفیت کلی دارد [۲۰].

محصولات بنگاه های بازی سازی احساس می گردد که این شکاف در این پژوهش مورد بررسی قرار می گیرد.

## ۲-۲- مدل های کیفیت خدمات

در این بخش، مدل های موجود در زمینه کیفیت خدمات، کیفیت خدمات الکترونیکی، کیفیت خدمات رایانش ابری، کیفیت خدمات بازی به عنوان خدمت از ادبیات موجود تشریح گردید. در ادامه به مرور مدل های کیفیت خدمات بررسی شده پرداخته می شود.

**مدل SERVPERF:** در این تحقیق یک مدل مفهومی و اندازه گیری کیفیت خدمات و ارتباط آن با رضایت مشتری و قصد خرید به دست آمده است. محققان امتیازات مختلف ادراک مشتری را بررسی کرده و نتیجه گیری کرده اند که تنها ادراک مشتری پیش بینی کننده بهتری از کیفیت خدمات است. کرونین و تیلور عقیده دارند عملکرد فعلی سازمان مهم ترین مبنای ارزیابی کیفیت خدمت ارائه شده می باشد و انتظارات مشتری را نباید در مفهوم کیفیت خدمت دخیل دانست [۲۹].

**مدل کیفیت خدمات وب:** در این تحقیق ابعاد کیفیت خدمات در سناریو مبتنی بر وب مورد تجزیه تحلیل قرار گرفت و تعاریف جدید برای ابعاد SERVQUAL شناسایی شد. ابعاد به دست آمده برای کیفیت خدمات در این تحقیق عبارت بودند از پاسخگویی، قابلیت اطمینان، اطمینان، همدلی، کیفیت اطلاعات و ارتباطات یکپارچه [۳۰].

**مدل کیفیت خدمات الکترونیکی:** این مطالعه یک مدل مفهومی برای کیفیت خدمات الکترونیکی ارائه می دهد. ابعاد مدل پیشنهادی عبارتند از طراحی مناسب وبسایت، چگونگی استفاده از تکنولوژی به جهت دسترسی آسان مشتری، جذابیت و فهم وبسایت، پشتیبانی مناسب و نگهداری و مراقبتی که وبسایت می تواند برای مشتریان خود فراهم کند [۳۱].

**مدل E-S-QUAL:** این مدل برای سنجش کیفیت خدماتی که توسط یک وبسایت ارائه می شود تهیه شده است. این مدل شامل ۴ بعد کارایی، در دسترس بودن سیستم، تکمیل چرخه خدمت رسانی و حریم خصوصی [۳۲].

**مدل کیفیت خدمات سیستم های اطلاعاتی:** در این پژوهش مدل سروکوال به جهت استفاده در سیستم های اطلاعاتی مورد بررسی قرار گرفت و ۵ بعد مدل سروکوال به ۴ بعد تغییر کرد. ۴ بعد نهایی به دست آمده بر اساس تحلیل عاملی عبارت بودند از قابلیت اطمینان، پاسخگویی، توافق و عوامل ملموس. همچنین شاخص های متناسب با سنجش هر کدام از این زیر بعدها متناسب با حوزه سیستم های اطلاعاتی تعیین گردید [۳۳].

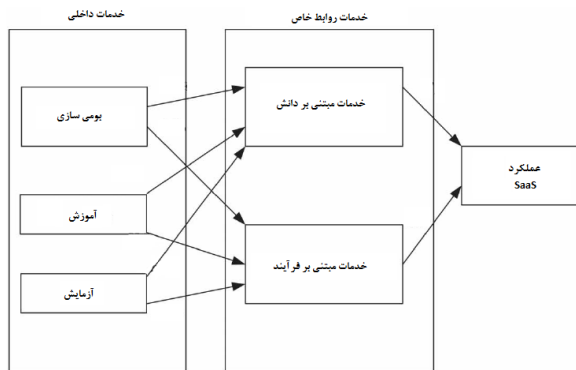
اطلاعات از اجزای بازی به سرور و پایانه های بازیکنان بود. در نهایت، رویکرد عملی برای پارادایم بازی ابری را مورد بررسی قرار داد. آزمایش های انجام شده نشان داد که پارتیشن بندی هوشمند منجر به عملکرد بهتر سیستم می شود [۲۴].

در پژوهشی تحت عنوان به دنبال اندازه گیری کیفیت خدمات مورداستفاده در بازی های ابری موبایلی بودند. عملکرد کیفیت از لحاظ سرعت بیت، از دست دادن بسته و نرخ فریم اندازه گیری شد. نتایج نشان داد که نرخ فریم و نرخ بیت بر تأخیر در شبکه تأثیر دارند [۱۵]. در پژوهشی برای بهبود کیفیت بازی های ویدیویی، یک مدل شبیه سازی برای تأخیر پایان به پایان در شبکه بازی های کامپیوتری و ابر ارائه دادند. تأخیر پایان به پایان عبارت است از میزان تأخیر در اقدام کاربر و نمایش نتایج بر روی صفحه کاربر [۲۵]. در پژوهشی با انجام یک مطالعه آزمایشگاهی ذهنی بر روی ۵۲ بازیکن و دو بازی متفاوت به شناسایی استراتژی های انطباق رمزگذاری ویدئویی مبتنی بر کیفیت تجربه پرداختند. نتایج نشان داد که در شرایط خاص، پهنای باند شناسایی شده و کاهش سرعت فریم منجر به بهبود کیفیت تجربه خواهد شد [۲۶]. در تحقیقی سازمان را قادر ساخت تا به درک بهتر از تأثیر عوامل کسب و کار و نقش آن ها از نظر عملکرد بازی در بازار برسند. سازمان های توسعه دهنده بازی نیاز دارند تا عوامل مختلف کسب و کار را در نظر داشته و تلاش های فعلی خود را برای بهبود عملکرد بازی در بازار را باید افزایش دهند. این مقاله بر روی توسعه یک بازی دیجیتال به بلوغ رسیده برای ارزیابی روند توسعه تمرکز داشت. این مطالعه به بررسی شواهد تجربی و توجیه عوامل کسب و کار در ارزیابی کسب و کار را بعد از بلوغ فرایند توسعه بازی را فراهم نمود [۱۰]. در مقاله ای چارچوب ترکیبی جدیدی را با استفاده از جریان ویدئو و گرافیک آن برای ارائه بازی با کیفیت بالا پیشنهاد نمودند. در چارچوب پیشنهادی، سرویس های ابر نه تنها فریم های ویدئوی کدگذاری شده را انتقال می دهند، بلکه به تدریج اطلاعات گرافیکی را ارسال می کنند. نتایج نشان داد که جریان ترکیبی پیشنهاد شده در مقایسه با جریان ویدئو سنتی می تواند به پایین ترین تحریف در محدودیت های پهنای باند و تأخیر در ارائه پاسخ دست یابد [۲۷].

در پژوهشی به طراحی و پیاده سازی یک سیستم بازی مجهز به قابلیت های شناختی پرداخته شده است. هدف طراحی این سیستم غلبه بر مشکل متغیر بودن کیفیت برای کاربران بوده است. نتایج نشان داد که اجزای نرم افزاری متعادل بین ابر و دستگاه های کاربر منجر به عملکرد بهتر سیستم مانند کاهش در زمان تأخیر می شود [۲۸]. بر اساس مرور پیشینه صورت گرفته، کمبود مدل های منسجم و بنیادی برای حمایت از به کارگیری خدمات رایانش ابری در توسعه

**مدل کیفیت خدمات ابری:** در این تحقیق یک مدل کیفیت خدمات ابری برای تمامی خدمات ارائه شده بر روی ابر تهیه شده است و مطالعه موردی بر روی ۳ شرکت ارائه خدمات ذخیره‌سازی اطلاعات ابری انجام شده است و در پایان نیز شاخص‌هایی به جهت سنجش کیفیت خدمات ابری ارائه شده است. ابعاد کیفیت شناسایی شده در این تحقیق عبارت‌اند از قابلیت استفاده، قابلیت دسترسی، قابلیت اطمینان، پاسخگویی، امنیت و قابلیت تطابق [۳۹].

**مدل عملکرد خدمات SaaS:** مدل مورد استفاده در این پژوهش به قرار شکل زیر می‌باشد [۴۰]:



شکل ۲. مدل عملکرد خدمات SaaS

**مدل کیفیت خدمات برنامه‌ریزی منابع سازمانی بر روی ابر:** در این پژوهش یک مشتری که از خدمات نرم‌افزار برنامه‌ریزی منابع سازمانی بر روی ابر استفاده می‌کرده است مورد مطالعه موردی قرار گرفته و نتایج به دست آمده با مدل سس-کوال مورد مقایسه قرار گرفته است. بر اساس خروجی‌های این تحقیق علاوه بر عوامل مدل سس-کوال یعنی توافق، پاسخگویی، قابلیت اتکا، انعطاف‌پذیری، امکانات و امنیت عوامل دیگری نیز بر کیفیت خدمات مؤثر هستند که این عوامل عبارت‌اند از آموزش، نگهداری، تحقق وعده، یکپارچگی داده، بازخورد، ارتباطات، انتظارات واقعی، ارزیابی هزینه پایین [۱۲].

**مدل کیفیت خدمات نرم‌افزار به‌عنوان خدمت:** در این پژوهش یک مدل کیفیت خدمات نرم‌افزار به‌عنوان خدمت از نگاه مشتریان و با مطالعه موردی مشتریان یک شرکت ارائه‌کننده خدمات نرم‌افزار به‌عنوان خدمت در برزیل توسعه داده شده است. ابعاد کیفیت خدمات به دست آمده در این مدل عبارت‌اند از خدمات مشتری، همپاری مشتری، قابلیت اطمینان، فرایندهای کسب‌وکار، دسترسی، اطلاعات در رابطه با نحوه استفاده از نرم‌افزار به‌عنوان خدمت [۴۱].

**مدل کیفیت خدمات الکترونیکی:** به بررسی رابطه بین کیفیت خدمات الکترونیکی درک شده و ارزش برند پرداختند. آن‌ها با استفاده از داده‌های جمع‌آوری شده از کاربران بانکداری آنلاین در تایوان، یک مدل سلسله‌مراتبی کیفیت خدمات الکترونیکی استخراج

**مدل E-GOVQUAL:** این مطالعه مدلی برای مفهوم‌سازی و ارزیابی خدمات که دولت به‌صورت الکترونیک به شهروندان ارائه می‌کند را توسعه داده است. ابعاد کیفیت خدمات در این مدل کارایی، قابلیت اتکا، پشتیبانی شهروندان و اعتماد به دست آمده است [۳۴].

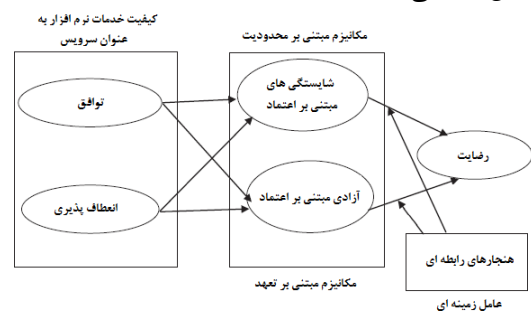
**مدل کیفیت خدمات اینترنت:** در این تحقیقات مدلی برای کیفیت خدماتی که توسط تأمین‌کنندگان اینترنت (ISPs) ارائه می‌شود توسعه یافته است. این تحقیق در صنعت خدمات اینترنتی تایلد انجام شده و ابعاد کیفیت خدمات به دست آمده عبارت‌اند از کیفیت شبکه، خدمات مشتری و پشتیبانی فنی، کیفیت اطلاعات و پشتیبانی وبسایت از اطلاعات و امنیت و حریم شخصی [۳۵].

**مدل SaaS-Qual:** در این تحقیق با تطبیق مدل Servqual یک مدل کیفیت خدمات برای حوزه خدمات نرم‌افزار به‌عنوان خدمت توسعه داده شده است. ابعاد کیفیت خدمات به دست آمده در این تحقیق عبارت‌اند از توافق، پاسخگویی، قابلیت اتکا، انعطاف‌پذیری، امکانات و امنیت [۳۶].

**مدل کیفیت خدمات نرم‌افزار به‌عنوان خدمت:** در ابتدا ۶ ویژگی نرم‌افزار به‌عنوان خدمت به شرح قابلیت استفاده مجدد، دسترسی، مقیاس‌پذیری، پرداخت هزینه استفاده، سفارشی‌سازی، مدیریت داده‌ها توسط تأمین‌کننده استخراج شده و پس از تطابق این ۶ ویژگی با ۵ شاخص کیفیت خدمات به شرح استفاده مجدد، کارایی، قابلیت اطمینان، مقیاس‌پذیری و دسترسی مدلی برای ارزیابی کیفیت خدمات نرم‌افزار به‌عنوان خدمت ارائه شده است [۳۷].

**مدل کیفیت خدمات IT:** در این مقاله یک چارچوب اندازه‌گیری کیفیت خدمات فناوری اطلاعات مورد ارزیابی قرار گرفته است. این چارچوب کیفیت را در ۶ حوزه رضایت مشتری، رفتار خدماتی مشتری، عملکرد فرایندی، کیفیت سیستم‌های اطلاعاتی، کیفیت خدمات فناوری اطلاعات و ارزش خدمات فناوری اطلاعات ارزیابی کرده است.

**مدل کیفیت خدمات SaaS:** مدل مورد استفاده در این پژوهش به قرار شکل (۱) می‌باشد [۳۸]:



شکل ۱. مدل کیفیت خدمات SaaS

دهند. این چارچوب تأثیر قابل توجهی در رقابت سالم بین ارائه دهندگان خدمات ابری برای برآورده کردن کیفیت مورد نیاز کاربران دارد. شاخص های ارائه شده در این چارچوب عبارتند از: پاسخگویی، چابکی در ارائه خدمات، قیمت خدمات، عملکرد خدمات، قابلیت اطمینان، امنیت و حریم خصوصی و قابلیت استفاده [۴۷].

**مدل سلسله مراتبی کیفیت خدمات صنعت بازی: مدلی چندبعدی و سلسله مراتبی جهت سنجش کیفیت خدمات در صنعت بازی ارائه نمودند.** آن ها با ایجاد یک چارچوب مفهومی و مقیاس اندازه گیری به بررسی کیفیت خدمات در صنعت بازی پرداختند. مدل حاضر با مرور ادبیات و نظرسنجی از خبرگان ارائه شده است. تجزیه و تحلیل اطلاعات جمع آوری شده از مجموع ۵۱۰ کاربر نشان داد که مدل پیشنهادی به خوبی داده ها را در بر می گیرد. ابعاد مدل پیشنهادی عبارتند از: کیفیت تعامل، کیفیت محیط فیزیکی، کیفیت خروجی ها (صوتی / تصویری)، رفتار، تجربه و تخصص، حل مسئله، تجهیزات، محدوده فعالیت، طراحی، امنیت، ظرفیت، زمان انتظار [۱۷].

**مدل پذیرش بازی مبتنی بر رایانش ابری: مدلی برای پذیرش بازی مبتنی بر رایانش ابری توسط کاربران ارائه نمودند.** آن ها در این تحقیق عوامل مؤثر بر پذیرش بازی ابر توسط کاربران را مورد بررسی قرار دادند. آن ها معتقد بودند که عدم آگاهی کاربران در مورد این فناوری، منجر به رد کردن فناوری جدید می شود. این مدل با استفاده از مصاحبه و نظرسنجی بر پایه خبرگان استخراج گردیده است. نتایج این تحقیق ویژگی های پلتفرمی که بر فرآیند تصمیم گیری کاربران تأثیر می گذارد بوده است. این ویژگی ها عبارتند از: کیفیت درک شده، در دسترس بودن اینترنت، سودمندی درک شده، درک سهولت استفاده، امنیت درک شده، مالکیت درک شده، علاقه به بازی کردن، هزینه های بازی و در دسترس بودن بازی [۴۸].

همان طور که از مرور مدل های موجود برای ارزیابی کیفیت خدمات (مدل های SERVQUAL، SERVPERF و CASERV و دیگر مدل های کیفیت خدمات سنتی) بر می آید، این مدل ها پاسخگوی جنبه های نوین بازی ابر نیستند. به همین دلیل، وجود مدلی با رویکرد جایگزین برای اندازه گیری درک مشتری از کیفیت خدمات در صنعت بازی با غلبه بر نقاط ضعف مدل های سنتی، مشهود می باشد.

### ۳- روش تحقیق

پژوهش حاضر از نظر هدف «توسعه ای-کاربردی» است و از نظر روش جزء تحقیقات کیفی به حساب می آید. همچنین این پژوهش از منظر شیوه گردآوری اطلاعات تحقیق «توصیفی-غیرآزمایشی» محسوب می شود [۴۸].

نمودند و تأثیر آن را بر اعتماد، رضایت، وفاداری و ارزش برند بررسی نمودند. نتایج نشان داد که کیفیت درک شده که از طریق تعامل با سرویس بانکی آنلاین شکل گرفت، به طور مثبت بر اعتماد و رضایت مشتری تأثیر می گذارد که به نوبه خود بر وفاداری و ارزش برند تأثیر می گذارد [۴۲].

**مدل کیفیت خدمات الکترونیکی دولت: با استفاده از ۳۰ مدل کیفی ارائه خدمات دولت الکترونیک و نظرخواهی از خبرگان مدلی جهت سنجش کیفیت خدمات الکترونیکی ارائه شده توسط دولت ارائه نمودند.** مدل آن ها دارای ۴ بعد اصلی و ۳۰ بعد فرعی بوده است. ابعاد اصلی این مدل عبارتند از: دامنه مدیریت، دامنه اطلاعات، دامنه خدمات و دامنه فنی [۴۳].

**مدل کیفیت خدمات رایانش ابری: مدلی جهت سنجش کیفیت خدمات مبتنی بر اعتماد رایانش ابری ارائه نمود.** تحقیق حاضر یک مدل اعتماد جدید را بر اساس تحقیقات گذشته ارائه داد. مدل ارائه شده با استفاده از چهار پارامتر دسترسی، قابلیت اطمینان، کارایی و یکپارچگی داده ها به سنجش کیفیت خدمات می پردازد. این تحقیق یک توافقنامه سطح خدمات مرکب از کیفیت خدمات مورد نیاز کاربر و قابلیت های ارائه دهندگان خدمات ابر ارائه نمود [۴۴].

**مدل عوامل مؤثر بر کیفیت خدمات بازی های رایانه ای ابری: در تحقیق به ارائه مدلی برای سنجش کیفیت خدمات بازی های رایانه ای پرداختند.** آن ها در این تحقیق کاربران دو بازی کامپیوتری ارائه شده توسط یک ارائه دهنده خدمات بازی های رایانه ای را مورد بررسی قرار دادند. این تحقیق تأثیر پهنای باند، انتقال و تأخیر را بر جنبه های مختلف کیفیت درک شده توسط کاربر مورد بررسی قرار داد. نتایج تحقیق نشان داد که راه اندازی نرم افزار در محیطی محدود تأثیر قابل توجهی بر قضاوت های ذهنی کاربران دارد و نیز محتوای بازی بر معیارهای عملکردی کیفیت درک شده تأثیر می گذارد [۴۵].

**مدل کیفیت خدمات بازی مبتنی بر ابر: در تحقیقی پلتفرمی جهت ارائه خدمات بازی ابری چندکاربره پرداختند.** در انتها به مقایسه پلتفرم پیشنهادی و پلتفرم های معروف OnLive، Ubitus و Gaikai بر پایه شاخص های عملکردی پرداختند. شاخص های عملکردی عبارتند از: بستر نرم افزاری مورد نیاز، قابل انتقال و حمل بودن، وضوح رمزگذاری صوتی و تصویری، پشتیبانی کاربران، مجازی سازی، فرمت رمزگذاری صوتی و تصویری و پشتیبانی گرافیکی [۴۶].

**چارچوب رتبه بندی خدمات رایانش ابری: چارچوبی جهت رتبه بندی خدمات رایانش ابری ارائه نمودند.** مدل ارائه شده در این پژوهش کاربران را قادر می سازد تا خدمات دریافت شده از سوی شرکت های ارائه دهنده خدمات رایانش ابری را مورد سنجش قرار

رویه انجام این تحقیق بر چهار گام استوار بود، در گام اول، مرور ادبیات برای شناسایی مقالات و مدل‌های موجود در زمینه کیفیت خدمات در حوزه‌های فناوری محور و بالأخص صنعت بازی صورت پذیرفت. برای این کار از کلیدواژه‌هایی نظیر مدل‌های موجود در زمینه کیفیت خدمات، کیفیت خدمات الکترونیکی، کیفیت خدمات رایانش ابری، کیفیت خدمات بازی به‌عنوان خدمت استفاده شد. در گام بعد به شناسایی ابعاد، ویژگی‌ها و شاخص‌های مدل‌های موجود جهت سنجش کیفیت خدمات پرداخته شد و این ابعاد، ویژگی‌ها و شاخص‌ها استخراج شدند. در گام سوم به دسته‌بندی ابعاد، ویژگی‌ها و شاخص‌های مدل کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر قالب جدول (۱) پرداخته شد. در گام چهارم پس از طی چهار راند روش دلفی مدل نهایی تحقیق استخراج گردید.

جدول ۱. معیارهای مدل کیفیت خدمات صنعت بازی مبتنی بر ابر

ابعاد اصلی	زیر بعدها	منابع
قابلیت اطمینان پشتیبانی	توانایی کارمندان برای انتقال اعتماد و اطمینان	[۳۴] و [۴۳]
	ارائه به‌موقع خدمات در زمان تعهد داده‌شده	[۲۹] و [۳۲] و [۳۴] و [۴۳] و [۴۴] و [۵۰] و [۵۱] و [۵۲]
	کنترل روند ارائه خدمات	[۴۳]
	وجود اعتماد در روابط کاری	[۱۷] و [۲۹] و [۳۶] و [۳۸] و [۴۳] و [۵۰] و [۵۱] و [۵۴]
	اجرای تعهدات مربوط به قرارداد	[۳۴] و [۳۶] و [۴۳]
	وجود اطلاعات تماس	[۳۴] و [۴۷]
پاسخگویی	اطلاع‌رسانی	[۴۷] و [۵۰] و [۵۵]
	اطلاع دادن زمان دقیق خدمت‌رسانی به کاربران	[۲۹] و [۳۶] و [۴۷] و [۵۳]
	ارائه کانال ارتباطی مناسب	[۳۶] و [۴۷] و [۵۰]
	در دسترس بودن کارکنان	[۲۹] و [۴۳] و [۴۷] و [۴۸]

ابعاد اصلی	زیر بعدها	منابع
شایستگی‌های اخلاقی و فنی کارکنان	ساعت کاری ارائه خدمات پشتیبانی مناسب برای کاربران	[۲۹] و [۳۶] و [۳۸] و [۴۷] و [۵۱]
	ارائه خدمات در کوتاه‌ترین زمان	[۳۲] و [۳۴] و [۳۵] و [۴۴] و [۵۴] و [۵۵] و [۵۶]
	دستورالعمل کاربرپسند	[۳۴] و [۳۵] و [۵۰]
	سؤالات متداول	[۳۴] و [۳۵] و [۵۰]
	امکانات پیگیری درخواست‌های کاربران	[۳۴] و [۳۵] و [۵۰]
	صلاحیت کارکنان	[۲۹] و [۳۴] و [۳۶] و [۳۸] و [۴۰] و [۵۰] و [۵۱] و [۵۳] و [۵۴] و [۵۵]
شایستگی‌های اخلاقی و فنی کارکنان	به‌روز نگه‌داشتن مداوم توانایی‌های فنی کارکنان	[۵۳]
	توجه فردی	[۲۹] و [۳۶] و [۵۱]
	ارائه حداکثر تلاش از جانب کارکنان	[۴۰]
	اشتیاق کارکنان جهت حل مشکل کاربران	[۱۷] و [۲۹] و [۳۴] و [۳۵] و [۳۶] و [۴۰] و [۵۰] و [۵۲] و [۵۳] و [۵۴] و [۵۵]
	نزاکت کارکنان	[۱۷] و [۲۹] و [۳۴] و [۳۶] و [۵۰] و [۵۱] و [۵۴]
	ایجاد رابطه دوستانه با کاربر	[۱۷] و [۳۶] و [۵۰] و [۵۳]
	تناسب فرهنگی بین ارائه‌دهنده خدمات نرم‌افزار به‌عنوان خدمت و دریافت‌کننده خدمت	[۳۶]
	ایجاد همدردی با کاربران	[۲۹] و [۵۱]



ابعاد اصلی	زیر بعدها	منابع
قابلیت اطمینان و دسترسی نرم افزار	ثبات عملکرد خدمات	[۴۲] و [۴۳] و [۴۷] [۵۲]
	اجرای تعهدات مربوط به قرارداد	[۳۴] و [۳۶] و [۴۳] [۴۷]
	ارائه خدمات بدون خطا	[۳۵] و [۳۶] و [۴۳]
	ارائه گواهینامه های کیفیت مناسب	[۳۶]
	در دسترس بودن سیستم برای کاربران	[۳۲] و [۳۶] و [۴۳] [۴۸]
	بازیابی کارآمد فاجعه	[۲۹] و [۳۶] و [۳۸]
	افزونگی سخت افزاری و نرم افزاری	[۲۹] و [۳۶] و [۳۸]
	قابل انتقال و حمل بودن	[۴۶]
	سرعت پردازش اطلاعات و بارگذاری	[۳۲] و [۳۴] و [۳۵] و [۳۸] [۴۳]
	انعطاف پذیری	قابلیت ادغام و همکاری برنامه SaaS با زیرساخت ICT کاربران
مدولار بودن ویژگی های انتخاب شده		[۳۶] و [۳۸]
جذابیت انتخاب خدمات		[۴۲]
امکان اصلاح پارامترهای قراردادی در مراحل بعدی		[۳۶]
امکانات و کاربرپسندی	تناسب امکانات نرم افزار با نیاز کاربران	[۳۲] و [۵۱] و [۵۳]
	داشتن دسترسی به آخرین نسخه های نرم افزاری	[۳۶] و [۳۸] و [۴۸]
	کیفیت فنی نرم افزار	[۴۲]
	ویژگی های استخراج و گزارش دهی اطلاعات	[۳۶]
	تنوع بستر نرم افزاری موردنیاز	[۴۶]
	قیمت خدمات	[۴۷] و [۴۸]
کیفیت اطلاعات	سرعت پردازنده سیستم بازی	[۴۴] و [۴۷]
	پهنای باند ارائه شده	[۴۴]
	پیچیدگی بازی	[۴۵]
	نبودن یا ناچیز بودن تأخیر در بازی	[۴۴]
	ویژگی های پیکربندی بازی	[۳۶]
	نبودن یا ناچیز بودن باگ	[۴۰]
	رابط کاربری بصری جذاب	[۳۶] و [۴۳] و [۴۵]
	کاربرپسند بودن سیستم	[۳۲] و [۳۴] و [۴۲] و [۴۳] و [۴۸]
	قابلیت متناسب سازی رابط کاربری	[۳۲] و [۳۴] و [۳۶] و [۳۸] و [۴۰] [۴۳]
	پشتیبان گیری و بازیابی داده ها	[۳۶] و [۴۶]
عدم سوءاستفاده از اطلاعات	فراهم کردن یک محیط فیزیکی امن برای نگهداری از اطلاعات	[۱۷] و [۳۴] و [۳۶] و [۴۳]
	حفاظت در برابر ویروس	[۳۶] و [۴۳]
	به روز بودن اطلاعات	[۴۲]
	قابل فهم بودن اطلاعات	[۴۲]
تصویر سازمانی	ممیزی های امنیتی و نفوذپذیری منظم	[۱۷] و [۳۶] و [۴۴] [۴۸]
	عدم ارائه اطلاعات شخصی کاربران به دیگران	[۳۲] و [۳۴] و [۳۶] و [۳۸]
	عدم استفاده از اطلاعات شخصی	[۳۴] و [۴۷]
	رمزگذاری صوتی و تصویری	[۱۷] و [۴۶]
	روش دستیابی به نام کاربری و رمز عبور	[۳۴]
	پیام های رمزگذاری شده	[۳۴] و [۳۶] و [۴۴]
تصویر سازمانی	کنترل دسترسی کارمندان به اطلاعات	[۳۴] و [۳۶] و [۴۷]
	مسئولیت پذیری اجتماعی	[۵۲] و [۵۵]
	نگرش قبلی کاربران نسبت به استفاده از خدمات	[۵۲]
	کیفیت گرافیک بازی	[۱۷] و [۴۴] و [۴۵] [۴۶]
مجازی سازی محیط بازی	[۴۴]	
کیفیت صدای بازی	[۴۵]	

شاخص	معیار
سرعت پردازش اطلاعات و بارگذاری	انعطاف‌پذیری (F)
قابلیت ادغام و همکاری بازی با زیرساخت ICT کاربران	
مدولار بودن ویژگی‌های انتخاب‌شده	
جذابیت انتخاب خدمات	
تناسب امکانات بازی با نیاز کاربران	امکانات و کاربرپسندی (U)
داشتن دسترسی به آخرین نسخه‌های بازی	
کیفیت فنی بازی	
ویژگی‌های استخراج و گزارش دهی اطلاعات	
تنوع بستر نرم‌افزاری موردنیاز	
قیمت خدمات	
کیفیت گرافیک بازی	
مجازی‌سازی محیط بازی	
کیفیت صدای بازی	
سرعت پردازنده سیستم بازی	
پهنای باند ارائه شده	
پچیدگی بازی	
افزایش توانمندی‌های افراد	
قابلیت به اشتراک‌گذاری دستاوردهای بازی	
نبود یا ناچیز بودن تأخیر در بازی	
خلاقیت در بازی	
ویژگی‌های پیکربندی بازی	
سرگرمی به‌دست‌آمده از بازی	
استفاده از المان‌های (شخصیت) جذاب عجین شده با فرهنگ ایرانی	
سناریوی جذاب بازی	
مشارکت با دوستان	
نبود یا ناچیز بودن باگ	
رابط کاربری بصری جذاب	
کاربرپسند بودن سیستم	
قابلیت متناسب‌سازی رابط کاربری	
به‌روز بودن اطلاعات	کیفیت اطلاعات

ابعاد اصلی	زیر بعدها	منابع
	تصویر اخلاقی و شهرت شرکت	[۵۲]

#### ۴- یافته‌های پژوهش

هدف از تحقیق حاضر ارائه مدلی جهت سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر بوده است. همان‌طور که از نظر گذشت، در بخش قبل مدل‌های موجود در زمینه کیفیت خدمات، کیفیت خدمات الکترونیکی، کیفیت خدمات رایانش ابری، کیفیت خدمات بازی به‌عنوان خدمت از ادبیات موجود تشریح گردید. پس از بررسی و جمع‌بندی مدل‌های موجود، مدل اولیه تحقیق در قالب جدول (۱) ارائه گردید. سپس با استفاده از روش دلفی و بر پایه نظرخواهی از خبرگان صنعت بازی‌های رایانه‌ای بر بستر ابر در چهار راند، مدل نهایی تحقیق مطابق جدول (۲) و شکل (۳) استخراج گردید.

جدول ۲. مدل بومی سنجش کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر در ایران

معیار	شاخص
قابلیت اطمینان پشتیبانی (S)	استفاده از بازخوردهای کاربران
	ارائه به‌موقع خدمات در زمان تعهد داده‌شده
	اجرای تعهدات مربوط به قرارداد
پاسخگویی (R)	وجود اطلاعات تماس
	ارائه کانال ارتباطی مناسب
	در دسترس بودن کارکنان
	ساعت کاری ارائه خدمات پشتیبانی مناسب برای کاربران
	ارائه خدمات در کوتاه‌ترین زمان
	دستورالعمل کاربرپسند
	امکانات پیگیری درخواست‌های کاربران
شایستگی‌های اخلاقی و فنی کارکنان (C)	تناسب فرهنگی بین ارائه‌دهنده خدمات بازی و دریافت‌کننده خدمت
	ایجاد همدردی با کاربران
	ثبات عملکرد خدمات
قابلیت اطمینان و دسترسی نرم‌افزار (A)	اجرای تعهدات مربوط به قرارداد
	ارائه خدمات بدون خطا
	بازیابی کارآمد فاجعه

تعهدات مربوط به قرارداد» (هر دو با وزن نهایی ۵/۲۵) بوده اند؛

۲. مهم ترین زیربدها در بعد «پاسخگویی»، «ارائه خدمات در کوتاه ترین زمان» (با وزن نهایی ۶/۰۰)، «ارائه کانال ارتباطی مناسب» (با وزن نهایی ۵/۷۵)، «دستورالعمل کاربرپسند» و «امکانات پیگیری درخواست های کاربران» (هر دو با وزن نهایی ۵/۵۰)، «» (با وزن نهایی ۶/۰۰) و «وجود اطلاعات تماس»، «ساعت کاری ارائه خدمات پشتیبانی مناسب برای کاربران» و «در دسترس بودن کارکنان» (هر سه با وزن نهایی ۵/۲۵) بوده اند؛

۳. مهم ترین زیربدها در بعد «شایستگی های اخلاقی و فنی کارکنان»، «تناسب فرهنگی بین ارائه دهنده خدمات بازی و دریافت کننده خدمت» (با وزن نهایی ۶/۲۵) و «ایجاد همدردی با کاربران» (با وزن نهایی ۵/۲۵) بوده اند؛

۴. مهم ترین زیربدها در بعد «قابلیت اطمینان و دسترسی نرم افزار»، «ثبات عملکرد خدمات» (با وزن نهایی ۶/۰۰) «ارائه خدمات بدون خطا»، «بازیابی کارآمد فاجعه» و «اجرای تعهدات مربوط به قرارداد» (هر سه با وزن نهایی ۵/۵۰) و «سرعت پردازش اطلاعات و بارگذاری» (با وزن نهایی ۵/۲۵) بوده است؛

۵. مهم ترین زیربدها در بعد «انعطاف پذیری»، «مدولار بودن ویژگی های انتخاب شده» و «جذابیت انتخاب خدمات» (هر دو با وزن نهایی ۵/۲۵) و «قابلیت ادغام و همکاری بازی با زیرساخت ICT کاربران» (با وزن نهایی ۵/۰۰) بوده اند؛

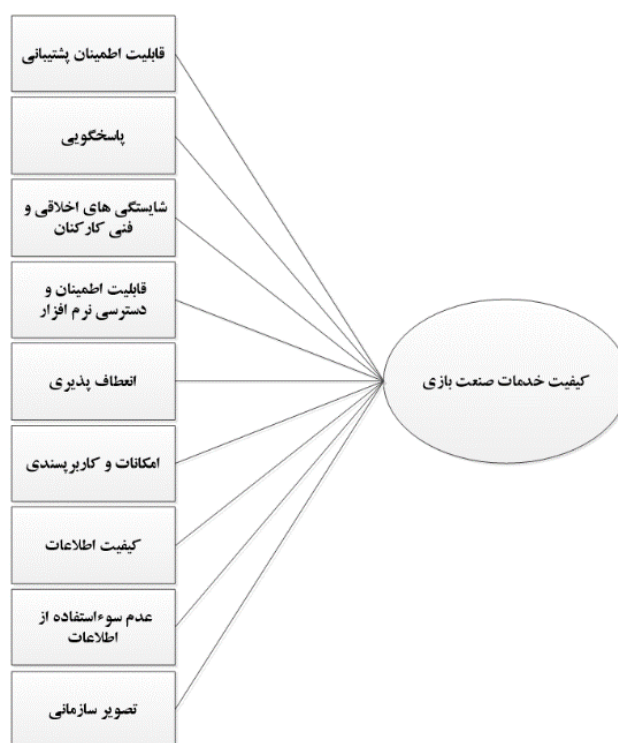
۶. مهم ترین زیربدها در بعد «امکانات و کاربرپسندی»، «استفاده از المان های (شخصیت) جذاب عجین شده با فرهنگ ایرانی» (با وزن نهایی ۶/۲۵) و «تناسب امکانات بازی با نیاز کاربران» (با وزن نهایی ۶/۰۰) بوده اند؛

۷. مهم ترین زیربدها در بعد «کیفیت اطلاعات»، «قابل فهم بودن اطلاعات» و «به روز بودن اطلاعات» (هر دو با وزن نهایی ۵/۵۰) بوده اند؛

۸. مهم ترین زیربدها در بعد «عدم سوءاستفاده از اطلاعات»، «عدم استفاده از اطلاعات شخصی» (با وزن نهایی ۵/۲۵) و «رمزگذاری صوتی و تصویری» (با وزن نهایی ۵/۰۰) بوده اند؛

۹. مهم ترین زیربدها در بعد «تصویر سازمانی»، «تصویر اخلاقی و شهرت شرکت» (با وزن نهایی ۵/۷۵)، «نگرش قبلی کاربران نسبت به استفاده از خدمات» (با وزن نهایی ۵/۵۰) و «نحوه نمایه کردن بازی» (با وزن نهایی ۵/۲۵) بوده اند؛

شاخص	معیار
قابل فهم بودن اطلاعات	(Q)
عدم استفاده از اطلاعات شخصی	عدم سوءاستفاده از اطلاعات (M)
رمزگذاری صوتی و تصویری	
نحوه نمایه کردن بازی	تصویر سازمانی (I)
نگرش قبلی کاربران نسبت به استفاده از خدمات	
تصویر اخلاقی و شهرت شرکت	



شکل ۳. مدل کیفیت خدمات صنعت بازی مبتنی بر ابر

به این ترتیب مدل ارزیابی کیفیت خدمات در صنعت بازی بر بستر ابر با رویکرد غلبه بر نقاط ضعف مدل های سنتی، ارائه شد. این مدل دارای ۹ بعد اصلی «قابلیت اطمینان پشتیبانی»، «پاسخگویی»، «شایستگی های اخلاقی و فنی کارکنان»، «قابلیت اطمینان و دسترسی نرم افزار»، «انعطاف پذیری»، «امکانات و کاربرپسندی»، «کیفیت اطلاعات»، «عدم سوءاستفاده از اطلاعات» و «تصویر سازمانی» می باشد. طبق یافته های تحقیق، نتایج زیر برای هر یک از ابعاد مدل به دست آمد:

۱. مهم ترین زیربدها در بعد «قابلیت اطمینان پشتیبانی»، «استفاده از بازخوردهای کاربران» (با وزن نهایی ۵/۷۵) و «ارائه به موقع خدمات در زمان تعهد داده شده» و «اجرای

## ۵- نتیجه‌گیری

در این پژوهش، عوامل مؤثر در «کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر» در مطالعه‌ای مروری مورد مطالعه قرار گرفت. مدل نهایی تحقیق پس از چهار گام استخراج گردید که دارای ۹ بعد اصلی و ۵۲ زیربعد می‌باشد. «امکانات و کاربرپسندی» با ۲۵ زیربعد بیشترین اجزای مدل را تشکیل داده و «تصویر سازمانی» با ۳ زیربعد کمترین اجزای مدل را تشکیل داده‌اند. نتایج این تحقیق در بسیاری از ابعاد با نتایج سایر تحقیقات صورت گرفته هماهنگی دارد. به‌طور مثال بر مبنای نتایج پژوهش، از بین بعدهای اصلی، بعد «امکانات و کاربرپسندی» بیشترین شاخص را به خود اختصاص داد که [۵۷] نیز در پژوهشی در زمینه ارائه خدمات در حوزه بانکداری به این مهم دست یافتند. در پاره‌ای از تحقیقات نیز نتایج متفاوتی به دست آمد که به‌عنوان مثال می‌توان به تحقیق [۳۹] عامل «کیفیت اطلاعات» را مهم‌ترین عامل در زمینه سنجش کیفیت خدمات در حوزه فناوری دانسته‌اند. به‌طور کلی نتایج این پژوهش آگاهی ما را نسبت به عوامل تعیین‌کننده «کیفیت خدمات صنعت بازی‌های رایانه‌ای بر بستر ابر» افزایش داد. با توجه نتایج تحقیق، پیشنهاد می‌شود با استفاده از دادن حق انتخاب به کاربر بر سهولت و انعطاف استفاده از بازی‌ها کمک نمایند. از آنجایی که کاربرپسند بودن هر محصول رایانه‌ای از جدی‌ترین معیارهای کیفیت بشمار می‌رود، می‌توان با کوتاه نمودن مراحل و فرایندها جهت دسترسی به بازی‌ها به این مهم دست یافت. همچنین پیشنهاد می‌گردد که سرعت پاسخ‌دهی را به حدی سرعت بخشند که کاربران حرفه‌ای بتوانند بدون هیچ تأخیری بازی نمایند و از خدمات ارائه شده راضی باشند. مهم‌ترین محدودیت پژوهش این است که سنجش کیفیت خدمات ابری در صنعت بازی ممکن است متفاوت با سایر صنایع باشد و قابلیت تعمیم گویه‌های به‌دست‌آمده وجود ندارد، از این رو برای انجام تحقیقات آتی به پژوهشگران پیشنهاد می‌شود مدل این تحقیق را مبنای پژوهش خود قرار داده و در دیگر صنایع فعال بررسی کنند.

## مراجع

- [۳] Cai, W., Chen, M., & Leung, V. C. Toward gaming as a service. *IEEE Internet Computing*, ۱۸(۳), ۱۲-۱۸, ۲۰۱۴.
- [۴] Claypool, M. Motion and scene complexity for streaming video games. In *Proceedings of the ۴th International Conference on Foundations of Digital Games* (pp. ۳۴-۴۱, ۲۰۰۹, April. ACM.
- [۵] Gupta, A., & Dutta, K. Cloud Gaming: Architecture and Quality of Service. *Career Point University Hamirpur Research Journal*, ۱(۲), ۱۹-۲۲, ۲۰۱۵.
- [۶] Murugesan, S., & Bojanova, I. (Eds.). *Encyclopedia of cloud computing*. John Wiley & Sons, ۲۰۱۶.
- [۷] Claypool, M., Finkel, D., Grant, A., & Solano, M. Thin to win?: network performance analysis of the OnLive thin client game system. In *Proceedings of the ۱۱th Annual Workshop on Network and Systems Support for Games* (p. ۱), ۲۰۱۲, November. IEEE Press.
- [۸] Shea, R., Liu, J., Ngai, E. C. H., & Cui, Y. Cloud Gaming: architecture and performance. *IEEE network*, ۲۷(۴), ۱۶-۲۱, ۲۰۱۳.
- [۹] Chen, K. T., Chang, Y. C., Tseng, P. H., Huang, C. Y., & Lei, C. L. Measuring the latency of cloud gaming systems. In *Proceedings of the ۱۹th ACM international conference on Multimedia* (pp. ۱۲۶۹-۱۲۷۲), ۲۰۱۱, November, ACM.
- [۱۰] Aleem, S., Capretz, L. F., & Ahmed, F. Empirical investigation of key business factors for digital game performance. *Entertainment Computing*, ۱۳, ۲۵-۳۶, ۲۰۱۶.
- [۱۱] Dabholkar, P. A. How to improve perceived service quality by increasing customer participation. In *Proceedings of the ۱۹۹th Academy of Marketing Science (AMS) Annual Conference* (pp. ۴۸۳-۴۸۷), ۲۰۱۵. Springer International Publishing.
- [۱۲] Chauhan, S., & Jaiswal, M. Exploring factors affecting service quality of ERP on cloud: a revelatory case study. *International Journal*
- [۱] Ross, P. E. Cloud computing's killer app: Gaming. *IEEE Spectrum*, ۴۶(۳), ۱۴-۱۴, ۲۰۰۹.
- [۲] Amiri, M., Osman, H. A., Shirmohammadi, S., & Abdallah, M. Toward delay-efficient game-aware data centers for cloud gaming. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, ۱۲(۵s), ۷۱, ۲۰۱۶.

- (MMSP), ۲۰۱۴ *IEEE ۱۶th International Workshop on* (pp. ۱-۶), ۲۰۱۴, September. IEEE.
- [۲۲] Chen, K. T., Chang, Y. C., Hsu, H. J., Chen, D. Y., Huang, C. Y., & Hsu, C. H. On the quality of service of cloud gaming systems. *IEEE Transactions on Multimedia*, ۱۶(۲), ۴۸۰-۴۹۵, ۲۰۱۴.
- [۲۳] Beyer, J., Varbelow, R., Antons, J. N., & Möller, S. Using electroencephalography and subjective self-assessment to measure the influence of quality variations in cloud gaming. In *Quality of Multimedia Experience (QoMEX)*, ۲۰۱۵ *Seventh International Workshop on* (pp. ۱-۶), ۲۰۱۵, May. IEEE.
- [۲۴] Cai, W. *Optimizing cloud gaming service delivery* (Doctoral dissertation, University of British Columbia), ۲۰۱۶.
- [۲۵] Metzger, F., Rafetseder, A., & Schwartz, C. A Comprehensive End-to-End Lag Model for Online and Cloud Video Gaming. In *Proc. ۵th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS ۲۰۱۶)* (pp. ۱۵-۱۹), ۲۰۱۶.
- [۲۶] Slivar, I., Skorin-Kapov, L., & Suznjevic, M. Cloud gaming QoE models for deriving video encoding adaptation strategies. In *Proceedings of the ۷th International Conference on Multimedia Systems* (p. ۱۸), ۲۰۱۶, May. ACM.
- [۲۷] Nan, X., Guo, X., Lu, Y., He, Y., Guan, L., Li, S., & Guo, B. Delay-Rate-Distortion Optimization for Cloud Gaming With Hybrid Streaming. *IEEE Transactions on Circuits and Systems for Video Technology*, ۲۷(۱۲), ۲۶۸۷-۲۷۰۱, ۲۰۱۷.
- [۲۸] Cai, W., Chi, Y., Zhou, C., Zhu, C., & Leung, V. C. UBC Gaming: Ubiquitous Cloud Gaming System. *IEEE Systems Journal*, ۲۰۱۸.
- [۲۹] Cronin Jr, J. J., & Taylor, S. A. Measuring service quality: a reexamination and extension. *The journal of marketing*, ۵۵-۶۸, ۱۹۹۲.
- [۳۰] Li, Y. N., Tan, K. C., & Xie, M. Factor analysis of service quality dimension shifts in *of Business Information Systems*, ۱۹(۱), ۸۷-۱۰۲, ۲۰۱۵.
- [۱۳] Wen, P. X., & Dong, L. Quality model for evaluating SaaS service. In *Emerging Intelligent Data and Web Technologies (EIDWT)*, ۲۰۱۳ *Fourth International Conference on* (pp. ۸۳-۸۷), ۲۰۱۳, September. IEEE.
- [۱۴] Hu, H., & Zhang, J. The evaluation system for cloud service quality based on servqual. In *Proceedings of the ۲۰۱۲ International Conference on Information Technology and Software Engineering* (pp. ۵۷۷-۵۸۴), ۲۰۱۳. Springer Berlin Heidelberg.
- [۱۵] Grandhi, V. V. S. S. G. On the Quality of Service of mobile cloud gaming using GamingAnywhere, ۲۰۱۶.
- [۱۶] Pedersen, J. M., Riaz, M. T., Junior, J. C., Dubalski, B., Ledzinski, D., & Patel, A. Assessing measurements of QoS for global cloud computing services. In *Dependable, Autonomic and Secure Computing (DASC)*, ۲۰۱۱ *IEEE Ninth International Conference on* (pp. ۶۸۲-۶۸۹), ۲۰۱۱, December. IEEE.
- [۱۷] Wu, H. C., & Hsu, F. S. A multi-dimensional and hierarchical model of service quality in the gaming industry. *International Journal of Tourism Sciences*, ۱۲(۳), ۹۰-۱۱۸, ۲۰۱۲.
- [۱۸] Lampe, U., Wu, Q., Dargutev, S., Hans, R., Miede, A., & Steinmetz, R. Assessing latency in cloud gaming. In *International Conference on Cloud Computing and Services Science* (pp. ۵۲-۶۸), ۲۰۱۳, May. Springer, Cham.
- [۱۹] Gharsallaoui, R., Hamdi, M., & Kim, T. H. A Comparative Study on Cloud Gaming Platforms. In *Control and Automation (CA)*, ۲۰۱۴ *۷th Conference on* (pp. ۲۸-۳۲), ۲۰۱۴, December. IEEE.
- [۲۰] Beyer, J., & Möller, S. Assessing the Impact of Game Type, Display Size and Network Delay on Mobile Gaming QoE. *PIK-Praxis der Informationsverarbeitung und Kommunikation*, ۳۷(۴), ۲۸۷-۲۹۵, ۲۰۱۴.
- [۲۱] Wen, Z. Y., & Hsiao, H. F. QoE-driven performance analysis of cloud gaming services. In *Multimedia Signal Processing*

- as-a-service based on service-dominant logic. *The Service Industries Journal*, ۳۴(۷), ۶۴۵-۶۵۸, ۲۰۱۴.
- [۴۱] Freitas, A. L. P., & Freitas Neto, M. M. Assessing the service quality in Software-as-a-Service from the customers' perspective: a methodological approach and case of use. *Production*, ۲۷, ۲۰۱۷.
- [۴۲] Kao, T. W. D., & Lin, W. T. The relationship between perceived e-service quality and brand equity: A simultaneous equations system approach. *Computers in Human Behavior*, ۵۷, ۲۰۸-۲۱۸, ۲۰۱۶.
- [۴۳] Sá, F., Rocha, Á., & Cota, M. P. Potential dimensions for a local e-Government services quality model. *Telematics and Informatics*, ۳۳(۲), ۲۷۰-۲۷۶, ۲۰۱۶.
- [۴۴] Manuel, P. A trust model of cloud computing based on Quality of Service. *Annals of Operations Research*, ۲۳۳(۱), ۲۸۱-۲۹۲, ۲۰۱۵.
- [۴۵] Möller, S., Pommer, D., Beyer, J., & Rake-Revelant, J. Factors influencing gaming QoE: Lessons learned from the evaluation of cloud gaming services. In *Proceedings of the 5th International Workshop on Perceptual Quality of Systems (PQS ۲۰۱۳)* (pp. ۱-۵), ۲۰۱۳, September.
- [۴۶] Kim, K. I., Bae, S. Y., Lee, D. C., Cho, C. S., Lee, H. J., & Lee, K. C. Cloud-Based Gaming Service Platform Supporting Multiple Devices. *ETRI Journal*, ۳۵(۶), ۹۶۰-۹۶۸, ۲۰۱۳.
- [۴۷] Garg, S. K., Versteeg, S., & Buyya, R. A framework for ranking of cloud computing services. *Future Generation Computer Systems*, ۲۹(۴), ۱۰۱۲-۱۰۲۳, ۲۰۱۳.
- [۴۸] Krenz, H., Terziyski, S., & Virjee, F. Consumer Acceptance of Cloud Computing Based Gaming: Bachelor's Thesis in Informatics, ۲۰۱۱.
- [۴۹] سرمد، ز؛ بازرگان، ع؛ و حجازی، ا؛ «روش‌های تحقیق در علوم رفتاری (چاپ بیست و ششم)». تهران: نشر آگه، ۱۳۹۳.
- [۵۰] Dagger, T. S., Sweeney, J. C., & Johnson, L. W. A hierarchical model of health service quality scale development and investigation of an integrated model. *Journal of Service Research*, ۱۰(۲), ۱۲۳-۱۴۲, ۲۰۰۷.
- the information age. *Managerial Auditing Journal*, ۱۸(۴), ۲۹۷-۳۰۲, ۲۰۰۳.
- [۳۱] Santos, J. E-service quality: a model of virtual service quality dimensions. *Managing Service Quality: An International Journal*, ۱۳(۳), ۲۳۳-۲۴۶, ۲۰۰۳.
- [۳۲] Parasuraman, A., Zeithaml, V. A., & Malhotra, A. ES-QUAL a multiple-item scale for assessing electronic service quality. *Journal of service research*, ۷(۳), ۲۱۳-۲۳۳, ۲۰۰۵.
- [۳۳] Kettinger, W. J., & Lee, C. C. Zones of tolerance: alternative scales for measuring information systems service quality. *MIS quarterly*, ۶۰۷-۶۲۳, ۲۰۰۵.
- [۳۴] Papadomichelaki, X., & Mentzas, G. e-GovQual: A multiple-item scale for assessing e-government service quality. *Government information quarterly*, ۲۹(۱), ۹۸-۱۰۹, ۲۰۱۲.
- [۳۵] Thaichon, P., Lobo, A., & Mitsis, A. An empirical model of home internet services quality in Thailand. *Asia Pacific Journal of Marketing and Logistics*, ۲۹(۲), ۱۹۰-۲۱۰, ۲۰۱۴.
- [۳۶] Benlian, A., Koufaris, M., & Hess, T. Service quality in software-as-a-service: Developing the SaaS-Qual measure and examining its role in usage continuance. *Journal of management information systems*, ۲۸(۳), ۸۵-۱۲۶, ۲۰۱۱.
- [۳۷] Lee, J. Y., Lee, J. W., & Kim, S. D. A quality model for evaluating software-as-a-service in cloud computing. In *Software Engineering Research, Management and Applications, ۲۰۰۹. SERA'۰۹. ۷th ACIS International Conference on* (pp. ۲۶۱-۲۶۶), ۲۰۰۹, December. IEEE.
- [۳۸] Chou, S. W., & Chiang, C. H. Understanding the formation of software-as-a-service (SaaS) satisfaction from the perspective of service quality. *Decision Support Systems*, ۵۶, ۱۴۸-۱۵۵, ۲۰۱۳.
- [۳۹] Zheng, X., Martin, P., Brohman, K., & Da Xu, L. Cloudqual: A quality model for cloud services. *IEEE transactions on industrial informatics*, ۱۰(۲), ۱۵۲۷-۱۵۳۶, ۲۰۱۴.
- [۴۰] Chou, S. W., Chang, Y. C., & Hsieh, P. H. Understanding the performance of software-

- [۵۱] Mittal, S., Gera, R., & Singhvi, S. R. Developing and testing a hierarchical model of customer perceived service quality for life insurance services. *Asia-Pacific Journal of Management Research and Innovation*, ۹(۱), ۶۳-۷۶, ۲۰۱۳.
- [۵۲] Thai, V. V. Logistics service quality: conceptual model and empirical evidence. *International Journal of Logistics Research and Applications*, ۱۶(۲), ۱۱۴-۱۳۱, ۲۰۱۳.
- [۵۳] Woo, K. S., & Ennew, C. T. Measuring business-to-business professional service quality and its consequences. *Journal of Business Research*, ۵۸(۹), ۱۱۷۸-۱۱۸۵ ۲۰۰۵.
- [۵۴] Leisen Pollack, B. The nature of the service quality and satisfaction relationship: empirical evidence for the existence of satisfiers and dis-satisfiers. *Managing Service Quality: An International Journal*, ۱۸(۶), ۵۳۷-۵۵۸, ۲۰۰۸.
- [۵۵] Mahapatra, S. S., & Khan, M. S. A neural network approach for assessing quality in technical education: an empirical study, ۲۰۰۷.
- [۵۶] DeLone, W. H., & McLean, E. R. The DeLone and McLean model of information systems success: a ten-year update. *Journal of management information systems*, ۱۹(۴), ۹-۳۰, ۲۰۰۳.
- [۵۷] Vivekanandan, L., & Jayasena, S. Facilities offered by the banks and expectations of IT savvy banking customers. *Procedia-Social and Behavioral Sciences*, ۴۰, ۵۷۶-۵۸۳, ۲۰۱۲.





# اندازه‌گیری میزان تشابه مسیرهای جهت‌دار بر روی داده‌های هندسی

زینب سعیدی\* محمد فرشی\*\*

\* دانشجوی دکتری، گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه یزد

\*\* استادیار، گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه یزد

تاریخ دریافت: ۱۳۹۹/۰۶/۰۱۰ تاریخ پذیرش: ۱۳۹۹/۱۰/۲۰

نوع مقاله: پژوهشی

## چکیده

در این مقاله به بررسی مسئله تشابه زیر در حوزه فاصله فرشه می‌پردازیم. یک مسیر جهت‌دار  $\pi$  به عنوان ورودی و یک پاره‌خط افقی  $Q$  که در لحظه پرس‌وجو توسط کاربر ارائه می‌شود، داده شده‌اند، هدف پیش‌پردازش و ذخیره مسیر جهت‌دار  $\pi$  در یک ساختمان داده است به طوری که با توجه به اطلاعات ذخیره شده در ساختمان داده بتوان زیرمسیری از مسیر جهت‌دار را گزارش کرد که فاصله فرشه میان زیرمسیر گزارش شده و پاره‌خط افقی  $Q$  بین تمام زیرمسیرهای ممکن مینیمم باشد. تا آنجایی که ما اطلاع داریم هیچ‌گونه نتیجه تئوری برای این مسئله گزارش نشده است. در این مقاله اولین الگوریتم ابتکاری برای مسئله ارائه شده است و به دلیل عدم ارائه الگوریتمی برای حل این مسئله در گذشته، صرفاً کیفیت الگوریتم ارائه شده بر روی چند پایگاه داده بررسی می‌گردد.

**واژگان کلیدی:** ساختمان داده، فاصله فرشه، فاصله هاسدورف، تشابه، مسیر جهت‌دار

## ۱. مقدمه

فاصله هاسدورف یک ابزار برای اندازه‌گیری میزان تشابه بین دو منحنی است. این فاصله برای اولین بار توسط فلیکس هاسدورف<sup>۱</sup> در سال ۱۹۱۴ میلادی در کتابش معرفی شد [۱]. در سال ۲۰۰۹، آلت<sup>۲</sup> نشان داد که فاصله هاسدورف میان دو چندضلعی با  $n$  و  $m$  رأس در زمان  $O((n+m)\log(n+m))$  قابل محاسبه است. یکی از کاربردهای فاصله هاسدورف در گرافیک کامپیوتری برای اندازه‌گیری تفاوت بین دو دید مختلف از یک شی می‌باشد [۳] [۲]. فاصله هاسدورف به صورت غیر رسمی برابر با ماکزیمم فاصله بین دو نقطه از دو منحنی است زمانی که هر نقطه از یک منحنی به

توصیف عددی دور بودن دو شی از هم فاصله نامیده می‌شود. در ریاضیات، تابع فاصله یا متر، تعمیمی از مفهوم فاصله فیزیکی است. متر، یک راه برای توصیف مفهوم نزدیک یا دور بودن عناصر یک فضا از هم است. اندازه‌گیری همسانی بین دو شی هندسی، یک مسئله اساسی در زمینه‌های بسیاری از علوم و مهندسی می‌باشد. برای مهیا کردن چنین مقایسه‌هایی، یک معیار خوب برای فرمول-بندی میزان تشابه یا همسانی لازم می‌باشد.

<sup>۱</sup> Felix Hausdorff  
<sup>۲</sup> Alt

نویسنده مسئول: محمد فرشی [mfarshi@yazd.ac.ir](mailto:mfarshi@yazd.ac.ir)

## ۲. مروری بر کارهای گذشته

دسته جالبی از مسائل که در حوزه فاصله فرشه مطرح می‌شود، ارائه ساختمان داده کارآمد برای مسائل پرس‌وجوی فرشه است. به این صورت که یک مسیر جهت‌دار داده شده است. می‌خواهیم با انجام یک پیش‌پردازش آن را در یک ساختمان داده ذخیره کنیم به صورتی که با توجه به اطلاعات ذخیره شده در ساختمان داده بتوانیم فاصله فرشه میان مسیر جهت‌دار و هر پاره‌خطی که در زمان پرس‌وجو داده می‌شود را محاسبه کنیم و یا تعداد زیرمسیرهایی را شمارش کنیم که فاصله فرشه آنها از پاره‌خطی که در زمان پرس‌وجو داده می‌شود، کوچکتر یا مساوی مقدار آستانه مشخص شده در زمان پرس‌وجو باشد.

دی برگ<sup>۳</sup> و همکارانش در سال ۲۰۱۳ مسئله پرس‌وجوی فرشه در مسیر جهت‌دار را مورد بررسی قرار دادند. آنها برای ساده‌تر کردن مسئله، پرس‌وجو را به صورت پاره‌خط افقی در نظر گرفتند و تعداد زیرمسیرهایی که فاصله فرشه آنها از پاره‌خط افقی داده شده حداکثر برابر  $\epsilon$  است را شمارش کردند [۸]. ساختمان داده‌ای که توسط دی برگ و همکارانش ارائه شد، این مسئله را به صورت تقریبی حل می‌کند. این ساختمان داده، علاوه بر تمام زیرمسیرهایی با فاصله  $\epsilon$ ، شامل تعدادی زیرمسیر اضافه است که فاصله فرشه آنها از پاره‌خط داده شده حداکثر برابر با  $\epsilon(2 + 3\sqrt{2})$  است. آنها فرض کردند که طول پاره‌خط داده شده در پرس‌وجو حداقل  $6\epsilon$  است و مسئله را در دو حالت بررسی کردند. الف) حالتی که  $\epsilon$  از قبل مشخص است. ب) حالتی که  $\epsilon$  جزء پرس‌وجو است و از قبل مشخص نمی‌باشد. آنها با طراحی یک ساختمان داده چندسطحی با زمان پیش‌پردازش  $O(n^2 + s \times \text{polylog } n)$  و فضای ذخیره‌سازی  $O(s \times \text{polylog } n)$  حالت الف) مسئله را حل کردند که  $n$  تعداد رئوس مسیر است و  $s$  پارامتری است که بین  $n$  و  $n^2$  قرار دارد. زمان پاسخ به پرس‌وجو  $O(\frac{n}{\sqrt{s}} \text{polylog } n)$  است. با افزایش زمان پیش‌پردازش به  $O(n^2 \log n)$  می‌توان مسئله را در حالت ب) حل کرد. فضای ذخیره‌سازی و زمان پاسخ به پرس‌وجو نسبت به حالت الف) تغییری نمی‌کند.

گودموندسون<sup>۴</sup> و اسمید<sup>۵</sup> در سال ۲۰۱۳ مسئله پرس‌وجوی فرشه در درخت هندسی (درختی که بر روی یک مجموعه از نقاط در صفحه ساخته شود) را مورد بررسی قرار دادند [۹]. درخت  $T$  و مقدار  $\Delta$  را به عنوان ورودی داریم. پرس‌وجو  $Q$  به صورت مسیر در نظر گرفته شده است. آنها در پی یافتن ساختمان داده‌ای برای ذخیره

نزدیک‌ترین نقطه‌اش در منحنی دیگر نگاشت گردد. یکی از مزیت‌های فاصله هاسدورف این است که به آسانی قابل توصیف است. فاصله هاسدورف جهت‌دار از منحنی  $A$  به منحنی  $B$  به صورت زیر تعریف می‌شود:

$$\delta_{\vec{h}}(A, B) = \supinf_{a \in A, b \in B} \text{dist}(a, b) \quad (1)$$

فاصله هاسدورف بین  $A$  و  $B$  نیز به صورت زیر تعریف می‌شود:

$$\delta_H(A, B) = \max\{\delta_{\vec{h}}(A, B), \delta_{\vec{h}}(B, A)\} \quad (2)$$

فاصله فرشه توسط موريس رنه فرشه<sup>۱</sup> در سال ۱۹۰۶ میلادی معرفی شد [۴]. یک مطالعه پایه روی ویژگی‌های محاسباتی فاصله فرشه توسط آلت و گودا<sup>۲</sup> در سال ۱۹۹۲ انجام شد [۵]. آنها الگوریتمی ارائه دادند که فاصله فرشه بین دو چندضلعی را در زمان  $O(nm \log^2(nm))$  محاسبه می‌کند که  $n$  و  $m$  تعداد رئوس چندضلعی‌ها هستند. در سال ۱۹۹۵، آلت و گودا نشان دادند که فاصله فرشه بین دو چندضلعی در زمان  $O(nm \log(nm))$  قابل محاسبه است و تا کنون هم نتیجه بهتری گزارش نشده است [۶]. فاصله فرشه دارای کاربردهای متنوعی از جمله تشخیص گفتار، تشخیص امضا و تشخیص دست‌خط است [۷].

فاصله فرشه به طور غیررسمی به صورت زیر تعریف می‌شود. یک شخص و سگش را در نظر بگیرید به گونه‌ای که سگ با یک قلاده به صاحبش وصل شده است. هر کدام از این دو روی یک مسیر جهت‌دار (منحنی) از نقطه ابتدایی آن تا انتهایش در حال حرکت می‌باشند. هر دوی آنها اجازه دارند که سرعت خود را کنترل کنند اما نمی‌توانند به عقب برگردند. هزینه هر پیاده‌روی برابر با حداکثر طول قلاده‌ای است که شخص را به سگش وصل می‌کند. پیاده‌روی‌های متفاوت هزینه‌ی متفاوت دارند. فاصله فرشه برابر با هزینه پیاده‌روی است که دارای هزینه مینیمم در بین تمام پیاده‌روی‌های ممکن است.

فاصله فرشه بین منحنی  $A$  و  $B$  به صورت زیر تعریف می‌شود:

$$\delta_F(A, B) = \inf_{\mu} \max_{a \in A} \text{dist}(a, \mu(a)) \quad (3)$$

که  $\text{dist}(a, \mu(a))$  فاصله اقلیدسی را نشان می‌دهد و  $\mu: A \rightarrow B$  یک تابع پیوسته غیرنزولی است که هر نقطه  $a \in A$  را به نقطه  $\mu(a)$  در  $B$  نگاشت می‌کند. تحت نگاشت  $\mu$ ، نقطه شروع  $A$  به نقطه شروع  $B$  نگاشت می‌شود که مطمئن باشیم دو تا منحنی در یک جهت پیموده می‌شوند. مزیت فاصله فرشه نسبت به معیارهای دیگر این است که فاصله فرشه ترتیب نقاط در امتداد منحنی‌ها را رعایت می‌کند.

<sup>۳</sup> de Berg

<sup>۴</sup> Gudmundsson

<sup>۵</sup> Smid

<sup>۱</sup> Maurice René Fréchet

<sup>۲</sup> Godua

کردن درخت  $T$  بودند که بتواند در زمان پرس‌وجو مشخص کند آیا درخت  $T$  شامل مسیری مانند  $P$  است که فاصله فرشه آن از مسیر داده شده در پرس‌وجو کمتر از  $\Delta$  باشد. آنها مسئله را در حالتی که درخت  $T$  به صورت  $c$ -بسته‌بندی (درختی که کل طول قرار گرفته از آن داخل هر دیسک حداکثر  $c$  برابر شعاع دیسک باشد) است و طول هر کدام از یال‌های درخت و مسیر داده شده در پرس‌وجو از  $\Omega(\Delta)$  است، بررسی کردند. ساختمان داده ارائه شده دارای زمان ساخت و فضای ذخیره‌سازی  $O(n \text{ polylog } n)$  است و برای مسیر  $Q$  داده شده در پرس‌وجو با  $m$  رأس می‌تواند در زمان  $O(m \text{ polylog } n)$  پاسخ آری یا نه را بدهد. اگر ساختمان داده پاسخ نه بدهد، حتماً مسیری در  $T$  وجود نداشته است که فاصله آن از  $Q$  کمتر از  $\Delta$  باشد. اگر پاسخ آری بدهد و  $m = 2$  باشد، به این معنی است که مسیری در  $T$  وجود دارد که فاصله فرشه آن از  $Q$  حداکثر برابر با

$\Delta \sqrt{2}(1 + \epsilon)$  است. اگر پاسخ آری بدهد و  $m > 2$  باشد، به این معنی است که مسیری در  $T$  وجود دارد که فاصله فرشه آن از  $Q$  حداکثر برابر با  $3\Delta$  است. یعنی در حالت آری پاسخ تقریبی خواهیم داشت. آنها در سال ۲۰۱۵ [۱۸] ضریب  $\sqrt{2}$  را از کران بدست آمده برای حالتی که  $m = 2$  است، حذف کردند و نتیجه بهتری را بدست آوردند. اما برای حالت  $m > 2$  کران بدست آمده جدید برابر با  $\Delta(1 + \epsilon)^3$  است.

دریمیل<sup>۱</sup> و هارپلد<sup>۲</sup> در سال ۲۰۱۳ مسئله پرس‌وجوی فرشه با کمک یال میان‌بر را بررسی کردند. آنها ساختمان داده‌ای ارائه کردند که می‌تواند دو مسئله‌ای که در ادامه مطرح می‌شود را حل کند [۱۰].

• منحنی  $\pi$  که دارای  $n$  رأس است را به عنوان ورودی داریم. پرس‌وجو، منحنی  $Q$  شامل  $k$  رأس و رئوس  $p$  و  $q$  است. هدف، ذخیره منحنی  $\pi$  در یک ساختمان داده است به صورتی که بتوانیم به ازای هر  $Q$ ،  $p$  و  $q$  که در پرس‌وجو داده می‌شود، فاصله فرشه میان زیرمنحنی از  $\pi$  که بین  $p$  و  $q$  قرار دارد و منحنی  $Q$  را محاسبه کنیم. ساختمان داده ارائه شده دارای زمان ساخت  $O(n \log^2 n)$  و فضای ذخیره‌سازی  $O(n \log n)$  است. زمان پاسخ به پرس‌وجو  $O(k^2 \log n \log(k \log n))$  است. ساختمان داده ارائه شده، فاصله فرشه تقریبی با ضریب تقریب ثابت ۲۳ را محاسبه می‌کند.

• منحنی داده شده در پرس‌وجو را به صورت پاره‌خط افقی در نظر می‌گیرد. ساختمان داده ارائه شده دارای زمان ساخت  $O(\chi^2 n \log^2 n)$  و فضای ذخیره‌سازی  $O(\chi^2 n)$  است که

دی برگ و مهرابی در سال ۲۰۱۵ مسئله یافتن زیرمسیرهای مستقیم در مسیر جهت‌دار را بررسی کردند. آنها دو معیار متفاوت برای مستقیم بودن ارائه کردند و با توجه به آن دو تعریف، مسئله را بررسی کردند. معیار اول، ضریب کشش و معیار دوم انحراف جهت است. با توجه به این دو معیار ساختمان داده‌های زیر طراحی شده است [۱۱].

• ضریب کشش: مسیر جهت‌دار  $\pi$  و مقدار  $t$  را به عنوان ورودی داریم. در پرس‌وجو دو مکان  $s$  و  $t$  که به ترتیب نشان دهنده نقطه شروع و نقطه پایان است، داده می‌شود. هدف گزارش تمام زیرمسیرهایی از  $\pi$  است به صورتی که نقطه شروع و پایان آن به ترتیب  $s$  و  $t$  باشد و ضریب کشش آن حداکثر  $t$  باشد. آنها صفحه را به سلول‌های مربعی تقسیم‌بندی کردند. به عبارتی صفحه را به صورت مشبکه  $m \times m$  در نظر گرفتند. ساختمان داده ارائه شده نیازمند فضای ذخیره‌سازی  $O(\min(n^2, nm^2, \tau nm))$  است که در زمان مورد انتظار  $O(n^{1/\Delta + \epsilon} + k)$  ساخته می‌شود و  $n$  تعداد رئوس مسیر جهت‌دار و  $k$  تعداد کل زیرمسیرهای جهت‌دار با شرایط ذکر شده است.

• انحراف جهت: مسیر جهت‌دار  $\pi$  و مقدار  $\alpha$  را به عنوان ورودی داریم. در پرس‌وجو دو مکان  $s$  و  $t$  که به ترتیب نشان دهنده نقطه شروع و نقطه پایان است، داده می‌شود. هدف گزارش تمام زیرمسیرهایی از  $\pi$  است که نقطه شروع و پایان آن‌ها به ترتیب  $s$  و  $t$  هستند و انحراف جهت آن حداکثر  $\alpha$  باشد که انحراف جهت مقدار بیشینه زاویه بین هر دو پاره‌خط متوالی مسیر جهت‌دار است. آنها صفحه را به سلول‌های مربعی تقسیم‌بندی کردند. به عبارتی صفحه را به صورت مشبکه در نظر گرفتند. ساختمان داده ارائه شده نیازمند فضای ذخیره‌سازی  $O(n + k)$  ساخته می‌شود و  $n$  تعداد رئوس مسیر جهت‌دار و  $k$  تعداد کل زیرمسیرهای جهت‌دار با شرایط ذکر شده است.

مسئله‌ی دیگری که در سال ۲۰۱۷ توسط دی برگ و همکارانش مورد بررسی قرار گرفته است، محاسبه فاصله فرشه میان مسیر جهت‌دار  $\pi$  و پرس‌وجو ورودی  $Q$  که به صورت پاره‌خط افقی فرض شده است، می‌باشد [۱۲]. هدف طراحی ساختمان داده‌هایی که برای مسیر  $\pi$  بتوانند برای هر پاره‌خطی که به عنوان پرس‌وجو داده می‌شود، فاصله فرشه میان پاره‌خط و مسیر جهت‌دار را به صورت دقیق محاسبه کنند. آن‌ها ساختمان داده‌ای ارائه دادند که فاصله فرشه میان مسیر جهت‌دار  $\pi$  را از هر پاره‌خط داده شده در

<sup>۱</sup> Driemel<sup>۲</sup> Har-Peled

پرس‌وجو با زمان پیش‌پردازش  $O(n^2 \log n)$  و فضای ذخیره‌سازی  $O(n^2)$  و زمان پرس‌وجوی  $O(\log^2 n)$  محاسبه می‌کند. دی برگ و همکارانش ساختمان داده دیگری طراحی کردند که فاصله فرشه میان زیرمسیری که توسط دو رأس داده شده در پرس‌وجو مشخص می‌شود و هر پاره‌خط دلخواهی که در زمان پرس‌وجو داده می‌شود را محاسبه کند. زمان پیش‌پردازش ساختمان داده ارائه شده  $O(n^2 \log^2 n)$  و پیچیدگی فضای پرس‌وجو برابر با  $O(\log^2 n)$  است.

بوخین و همکارانش در سال ۲۰۲۰ فضای ذخیره‌سازی ساختمان داده برای محاسبه فاصله فرشه میان مسیر جهت‌دار  $\pi$  و پاره‌خط افقی داده شده در پرس‌وجو را به  $O(n^{2/3})$  کاهش دادند [۱۳].

گودموندسون و همکارانش در سال ۲۰۱۷ مسئله پرس‌وجوی فرشه برای منحنی با طول یال‌های طویل را مورد بررسی قرار دادند. ورودی مسئله منحنی  $P$  شامل  $n$  رأس است [۱۴]. طول کوتاه‌ترین یال منحنی  $P$  است. در پرس‌وجو، منحنی  $Q$  با  $m$  رأس و مقدار  $\Delta$  داده می‌شود.  $l_Q$  طول کوتاه‌ترین یال منحنی  $Q$  است. منحنی  $P$  و  $Q$  دارای یال‌هایی با شرط  $l_P > 2 * \Delta$  و  $l_Q > (1 + \sqrt{2})\Delta$  هستند. شرطها در واقع نشان دهنده این هستند که مسئله برای منحنی با طول یال‌های طویل است. هدف، مشخص کردن این است که فاصله فرشه میان  $Q$  و  $P$  حداکثر مقدار  $\Delta$  است یا نه. گودموندسون و همکارانش ساختمان داده‌ای ارائه کردند که با زمان پیش‌پردازش و فضای ذخیره‌سازی  $O(n \log n)$  به پرس‌وجو در زمان  $O(m \log^2 n)$  پاسخ می‌دهد. دی برگ و همکارانش در سال ۲۰۱۷ ساختمان داده پویا برای پرس‌وجوی مجاورت در مجموعه‌ای از مسیرهای جهت‌دار ارائه کردند. برخلاف کارهای قبلی که ورودی مسئله تنها یک مسیر جهت‌دار بود، در این مقاله ورودی مسئله مجموعه  $S$  شامل  $n$  مسیر جهت‌دار و مقدار ثابت  $k$  است که تعداد رؤوس مسیر جهت‌داری است که در زمان پرس‌وجو داده می‌شود. ساختمان داده طراحی شده توسط دی برگ و همکارانش می‌تواند پرس‌وجوهای زیر را پاسخ بدهد [۱۵].

۱. نزدیک‌ترین همسایه: پرس‌وجو مسیر جهت‌دار  $Q$  است. هدف یافتن مسیری از بین مسیرهای جهت‌دار موجود در  $S$  است که فاصله فرشه آن از  $Q$  کمینه باشد.

۲.  $J$ -تا نزدیک‌ترین: پرس‌وجو، مسیر جهت‌دار  $Q$  و مقدار صحیح  $J$  است. هدف گزارش  $J$  مسیر از بین مسیرهای جهت‌دار موجود در  $S$  است که فاصله فرشه آنها از  $Q$  کمینه است.

۳. پرس‌وجوی بازه‌ای: پرس‌وجو، مسیر جهت‌دار  $Q$  و مقدار  $\delta_{Max}$  است. هدف گزارش کردن تمام مسیرهای جهت‌دار است که فاصله فرشه آنها از  $Q$  حداکثر  $\delta_{Max}$  است.

۴. تشابه: پرس‌وجو، مسیر جهت‌دار  $Q$  و مسیر جهت‌دار  $\tau$  موجود در مجموعه ورودی است. هدف محاسبه فاصله فرشه میان  $Q$  و  $\tau$  است.

ساختمان داده ارائه شده، مسائلی که در بالا بیان شد را به صورت تقریبی حل می‌کند. ضریب تقریب حداکثر برابر با  $\epsilon.reach(Q)$  است که  $reach(Q)$  حداکثر فاصله اقلیدسی میان نقطه شروع  $Q$  و سایر رؤوس دیگر  $Q$  است. فضای مورد نیاز برای ساختن ساختمان داده  $O(n/\epsilon^{2*k})$  است. زمان پاسخ به پرس‌وجو برای مسائل ذکر شده در بالا به ترتیب به صورت زیر است:

۱. نزدیک‌ترین همسایه:  $O(1)$

۲.  $J$ -تا نزدیک‌ترین:  $O(j)$

۳. پرس‌وجوی بازه‌ای:  $O(1 + K)$   $K$  تعداد مسیرهای

جهت‌دار خروجی است.

۴. تشابه:  $O(1)$

زمان درج کردن یک مسیر جهت‌دار با  $m$  رأس در ساختمان داده به صورت سرشکن برابر با  $O(1/\epsilon^{2*k}(m + \log(1/\epsilon)))$  است. زمان حذف کردن یک مسیر جهت‌دار با  $m$  رأس از ساختمان داده به صورت سرشکن برابر با  $O(1/\epsilon^{2*k})$  است.

پرس‌وجوی فاصله فرشه برای مسئله جستجوی بازه‌ای بین یک مجموعه از مسیرهای جهت‌دار و یک پرس‌وجو به صورت تجربی مورد بررسی قرار گرفته است و کیفیت الگوریتم بر روی پایگاه داده گزارش شده است [۱۶، ۱۷، ۱۸].

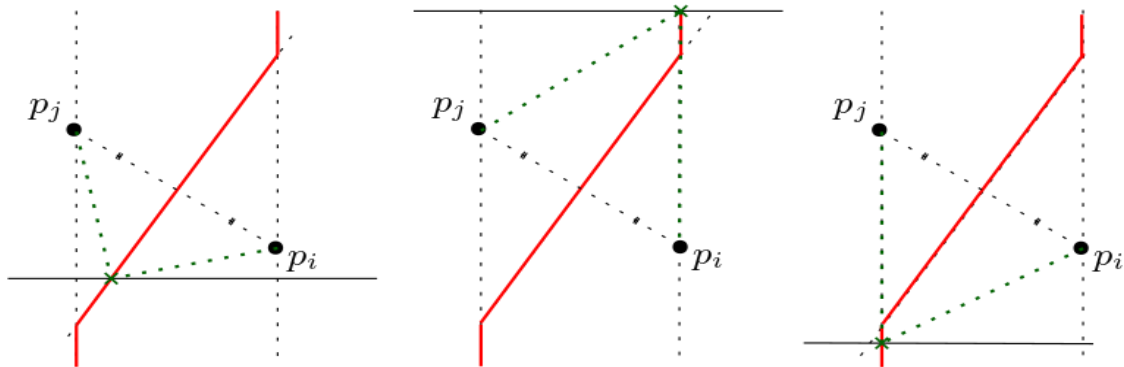
در این مقاله مسئله تشابه در حوزه فاصله فرشه را به صورت تجربی بررسی می‌کنیم. به این صورت که مسیر جهت‌دار  $\pi$  با  $n$  رأس به عنوان ورودی داده شده است. در زمان پرس‌وجو پاره‌خط افقی  $Q$  توسط کاربر مشخص می‌شود. هدف ذخیره مسیر جهت‌دار  $\pi$  در یک ساختمان داده است به صورتی که بتوانیم زیرمسیر  $\pi'$  از مسیر جهت‌دار را گزارش کنیم که فاصله فرشه میان زیرمسیر  $\pi'$  و پاره‌خط افقی  $Q$  بین تمام زیرمسیرهای ممکن مینیمم باشد. زیرمسیر یک مسیر جهت‌دار همبند از مسیر اصلی است به صورتی که نقطه ابتدا و انتهای آن جزء رؤوس مسیر اصلی هستند.

این مسئله بسیار پیچیده‌ای است و تا جایی که ما اطلاع داریم هیچ‌گونه نتیجه تئوری و تجربی برای این مسئله گزارش نشده است. راه حل اولیه برای یافتن پاسخ این است که فاصله فرشه میان تمامی زیرمسیرهای ممکن را محاسبه کنیم و سپس زیرمسیری که کمترین فاصله فرشه را دارا می‌باشد، گزارش کرد. با کمک ساختمان داده ارائه شده، فاصله فرشه میان زیرمسیر و پاره‌خط افقی در زمان  $O(\log^2 n)$  قابل محاسبه است [۱۲]. از طرفی تعداد زیرمسیرهای ممکن  $O(n^2)$  می‌باشد. بنابراین می‌توان در زمان  $O(n^2 \log^2 n)$  هر پرس‌وجویی را پاسخ داد. اما این زمان در

می‌کنیم و سپس در مرحله دوم با کمک روش‌های ارائه شده، زیرمسیر اولیه را تغییر می‌دهیم تا به پاسخ بهینه نزدیک گردد.

عمل کارآمد نمی‌باشد و نیازمند ایده خلاقانه برای بهبود زمان هستیم.

در این مقاله یک الگوریتم ابتکاری برای مسئله ارائه خواهیم داد. این الگوریتم از دو مرحله تشکیل شده است. در مرحله اول، با کمک ساختمان داده جستجوی بازه‌ای یک زیرمسیر را مشخص



شکل ۱. حالات مختلف برای محاسبه فاصله زوج برگشتی

زوج برگشتی را نشان می‌دهد. فاصله زوج برگشتی برای یک زوج برگشتی  $(p_i, p_j)$  به صورت زیر تعریف می‌گردد:

$$B_{(p_i, p_j)}(y) = \min_{x \in \mathbb{R}} \max\{\|p_i - (x, y)\|, \|p_j - (x, y)\|\}$$

عبارات داخل آکولاد فاصله بین نقطه داده شده  $(x, y)$  و نقطه  $p_i$  و نقطه  $p_j$  را محاسبه می‌کند. فاصله زوج نقاط برگشتی مربوط به زیرمسیر جهت‌دار  $\pi[u, v]$  به این گونه تعریف می‌شود:

$$B(\pi[u, v], y) = \max_{\forall p_i, p_j \in \pi[u, v]: i < j, p_i, x \geq p_j, x} B_{(p_i, p_j)}(y)$$

فاصله هاسدورف جهت‌دار بین زیرمسیر جهت‌دار  $\pi$  و پاره‌خط افقی  $Q$  به صورت زیر محاسبه می‌شود:

$$\delta_{\pi}(\pi[u, v], Q) = \max\{\max_{p_i, x \in (-\infty, x]} \|p - p_i\|, \max_{p_i, x \in [x_i, \infty)} \|q - p_i\|, \max_i \|y - p_i, y\|\}$$

طبق نتایج موجود در [۱۲]، فاصله فرشه بین زیرمسیر جهت‌دار  $\pi[u, v]$  و پاره‌خط افقی  $Q$  به صورت زیر محاسبه می‌شود:

$$\delta_F(\pi[u, v], pq) = \max\{\|up\|, \|vq\|, \delta_{\pi}(\pi[u, v], pq), B(\pi[u, v], y)\} (1 - \epsilon)$$

$\|up\|$  فاصله اقلیدسی میان نقطه  $u$  و  $p$  می‌باشد.

زیرمسیر جهت‌دار بهینه  $\pi'_{opt} = \pi[u', v']$  زیرمسیری از  $\pi$  است که فاصله فرشه آن از پاره‌خط افقی  $pq$  بین تمامی زیرمسیرهای ممکن مینیمم است. بنابراین برای هر زیرمسیر جهت‌دار  $\pi'$  از  $\pi$  خواهیم داشت:

$$\forall \pi' \subseteq \pi: \delta_F(\pi'_{opt}, pq) \leq \delta_F(\pi', pq)$$

#### ۴. الگوریتم پیشنهادی

در این قسمت یک الگوریتم ابتکاری ارائه خواهد شد که زیرمسیری از مسیر جهت‌دار را گزارش می‌کند که فاصله فرشه آن از پاره‌خط افقی میان تمام زیرمسیرهای ممکن مینیمم است. الگوریتم از دو مرحله تشکیل شده است که در مرحله اول نقطه ابتدا و انتها برای زیرمسیر مشخص می‌شود که این نقاط برآساس دو تا عبارت ابتدایی فرمول (۴) مشخص می‌گردند. در مرحله دوم

الگوریتم ارائه شده را بر روی چند پایگاه داده اجرا کرده‌ایم و کیفیت الگوریتم ارائه شده را برحسب درصد پرس‌وجوهایی که به درستی پاسخ داده می‌شوند، گزارش می‌کنیم.

### ۳. مفاهیم اولیه

مختصات هر نقطه  $p$  در صفحه را به صورت  $(x, y)$  نشان می‌دهیم. فرض می‌کنیم که  $p_1, p_2, \dots, p_n$  توالی از  $n$  نقطه در صفحه است و  $\pi = (p_1, p_2, \dots, p_n)$  یک مسیر جهت‌دار در صفحه است که توسط توالی ذکر شده تعریف شده است.  $Q = (p, q)$  یک خط افقی در صفحه است به صورتی که مختصات نقطه  $p$  و  $q$  به ترتیب برابر با  $p = (x, y)$  و  $q = (x_1, y)$  است به صورتی که  $x_1 \leq x$  و  $y \in \mathbb{R}$  فرض کنید که  $u$  و  $v$  دو رأس دلخواه از مسیر جهت‌دار  $\pi$  باشد. زیرمسیر جهت‌دار که از رأس  $u$  آغاز و به رأس  $v$  ختم می‌گردد را با عبارت  $\pi[u, v]$  نشان می‌دهند. زوج نقطه  $(p_i, p_j)$  زوج برگشتی (رو به جلو) است اگر  $j > i$  و نقطه  $p_j$  سمت چپ (رأست) نقطه  $p_i$  واقع شده باشد. یک یال که نقطه  $p_i$  را به نقطه  $p_{i+1}$  وصل می‌کند، یال برگشتی (رو به جلو) است اگر زوج  $(p_i, p_{i+1})$  یک زوج برگشتی (رو به جلو) باشد.

فاصله زوج برگشتی، ماکزیمم فاصله میان نقطه  $p_i$  و  $p_j$  و یک نقطه بر روی پاره‌خط افقی است. نقطه‌ای بر روی پاره‌خط افقی که فاصله زوج برگشتی را مشخص می‌کند، نقطه تقاطع بین عمودمنصف خط واصل بین  $p_i$  و  $p_j$  و پاره‌خط افقی است، اگر نقطه تقاطع بین نوار عمودی  $[p, x, q, x]$  واقع شده باشد. غیر اینصورت نقطه تقاطع بین خط  $x = p_i, x$  یا  $x = p_j, x$  و پاره‌خط افقی را باید محاسبه کرد. شکل ۱ حالات مختلف فاصله

$p_i$  و  $p_j$  را به عنوان اندیس جدید انتخاب می‌کنیم که ماکزیمم فاصله آن از نقاط  $p$  و  $q$  مینیمم باشد. الگوریتم ۱ جزئیات مرحله اول الگوریتم را نشان می‌دهد.

**Algorithm 1 Finding The Subtrajectory**

**Input:** The trajectory  $\pi$  and a horizontal segment  $pq$

**Output:** Candidate subtrajectory

- ۱: Let  $i$  and  $j$  be the index of the closest vertices of  $\pi$  to  $p$  and  $q$
- ۲: **if** ( $i > j$ ) **then**
- ۳: **if** ( $dist_i < dist_j$  and  $p_j \in B_i$ ) **then**
- ۴: change  $i$  into a vertex in  $B_i$  before  $p_j$
- ۵: **else**
- ۶: **if** ( $dist_i > dist_j$  and  $p_i \in B_j$ ) **then**
- ۷: change  $j$  into index of vertex in  $B_j$  after  $p_i$
- ۸: **else**
- ۹:  $i = j =$  The index of a vertex between  $j$  and  $i$  which minimize the maximum distance to  $p_i$  and  $p_j$ ;
- ۱۰: **end if**
- ۱۱: **end if**
- ۱۲: **end if**

در مرحله دوم الگوریتم، ماکزیمم عباراتی که فاصله فرشه میان زیرمسیر انتخاب شده از مرحله اول و پاره‌خط افقی را مشخص می‌کند، محاسبه می‌کنیم. در پایگاه داده‌هایی که در این مقاله بررسی می‌کنیم، عبارت آخر هیچ گاه فاصله فرشه را مشخص نمی‌کند. دلیل اصلی این است که چون طول یال‌های مسیرهای جهت‌دار در پایگاه داده‌های مورد بررسی کوچک است، نوار عمودی ایجاد شده باریک است و عمودمنصف، پاره خط افقی را خارج از نوار عمودی قطع می‌کند (حالت میانی و سمت راست شکل ۱). در این حالات، تصویر عمودی نقطه  $p_i$  یا  $p_j$  فاصله زوج برگشتی را مشخص می‌کند. بنابراین فاصله زوج برگشتی برای یک زیرمسیر جهت‌دار حداکثر برابر با فاصله هاسدورف است.

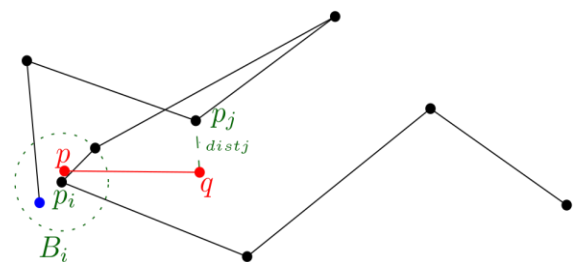
بر اساس مقدار ماکزیمم، حالات زیر را خواهیم داشت:

- اگر مقدار ماکزیمم عبارت  $\|p_i - p\|$  یا  $\|p_j - q\|$  باشد، الگوریتم زیرمسیر  $\pi[p_i, p_j]$  را گزارش می‌کند. از آنجایی که  $p_i$  و  $p_j$  نزدیک‌ترین نقاط به  $p$  و  $q$  هستند، بنابراین دو عبارت اول، کمترین مقدار ممکن خود را دارا هستند، و از طرفی ماکزیمم این دو عبارت فاصله فرشه را تعیین می‌کند، بنابراین زیرمسیر  $\pi[p_i, p_j]$  کمترین فاصله فرشه از پاره‌خط افقی را دارا است. لذا خواهیم داشت:

الگوریتم بررسی می‌شود که آیا سایر عبارات دیگر فرمول (۴) فاصله فرشه را اضافه می‌کنند یا نه. اگر سایر عبارات فرمول (۴) باعث ازدیاد فاصله فرشه گردد، زیرمسیر بدست آمده در مرحله اول تغییر داده می‌شود و در نهایت زیرمسیری گزارش می‌گردد.

نکته کلیدی برای بدست آوردن زیرمسیر  $\pi'_{opt}$  مینیمم کردن دو عبارت ابتدایی فرمول (۴) است. برای مینیمم کردن دو عبارت ابتدایی، نزدیک‌ترین نقاط مسیر جهت‌دار  $\pi$  به نقطه  $p$  و نقطه  $q$  محاسبه می‌گردد. این نقاط به ترتیب  $p_i$  و  $p_j$  نامیده می‌شود. با کمک ساختمان داده نزدیک‌ترین همسایه نقاط  $p_i$  و  $p_j$  محاسبه می‌شوند. اگر نقطه  $p_i$  بعد از نقطه  $p_j$  در مسیر  $\pi$  واقع شده باشد، اصلاحاتی بر روی زیرمسیر اعمال می‌گردد به صورتی که بعد از آن نقطه  $p_i$  قبل از نقطه  $p_j$  واقع می‌گردد. قابل ذکر است که اگر از ابتدا نقطه  $p_i$  قبل از نقطه  $p_j$  واقع شده باشد، می‌توان از مرحله اول صرفنظر کرد و به مرحله دوم رفت. در غیر اینصورت تغییراتی را اعمال می‌کنیم به صورتی که بعد از آن نقطه  $p_i$  قبل از نقطه  $p_j$  قرار خواهد گرفت.

به منظور اعمال تغییرات، فواصل  $dist_i = \|p_i - p\|$  و  $dist_j = \|p_j - q\|$  محاسبه می‌گردد. اگر  $dist_i < dist_j$  باشد، اندیس  $i$  را می‌توان به رأسی قبل از نقطه  $p_j$  منتقل کرد که فاصله آن از نقطه  $p$  کمتر از  $dist_j$  باشد. دیسک  $B_i$  را با مرکز  $p_i$  و شعاع  $dist_j$  در نظر بگیرید. همانگونه که در شکل ۲ مشخص است، اگر دیسک  $B_i$  مسیر جهت‌دار  $\pi$  را قبل از نقطه  $p_j$  قطع کند، نقطه  $p_i$  به رأسی داخل دیسک  $B_i$  قبل از  $p_j$  انتقال داده می‌شود. در غیر اینصورت رأسی بین  $p_i$  و  $p_j$  را به عنوان اندیس جدید انتخاب می‌کنیم که ماکزیمم فاصله آن از نقاط  $p$  و  $q$  مینیمم باشد.



شکل ۲: در این حالت  $i > j$  و  $dist_i < dist_j$  است. نقطه آبی رنگ داخل دیسک  $B_i$  اندیس جدید برای  $i$  می‌باشد.

اگر  $dist_i > dist_j$  باشد، اندیس  $j$  را می‌توان به رأسی بعد از نقطه  $p_i$  منتقل کرد که فاصله آن از نقطه  $q$  کمتر از  $dist_i$  باشد. دیسک  $B_j$  را با مرکز  $p_j$  و شعاع  $dist_i$  در نظر بگیرید. اگر دیسک  $B_j$  مسیر جهت‌دار  $\pi$  را بعد از نقطه  $p_i$  قطع کند، نقطه  $p_j$  به رأسی داخل دیسک  $B_j$  بعد از  $p_i$  انتقال داده می‌شود. اگر نتوانیم نقاط جدید در دیسک‌های  $B_i$  و  $B_j$  پیدا کنیم، رأسی بین

## Algorithm ۲ Reporting The Subtrajectory

Input: The subtrajectory  $[p_i, p_j]$  that is determined in Algorithm ۱

Output: Report the optimal subtrajectory

```

۱:  $m = \max\{\|p_i - p\|, \|p_j - q\|\}$ 
    $\delta_h \rightarrow (\pi[p_i, p_j], p, q), B(\pi[p_i, p_j], p, y)$ 
۲: if  $(m == \|p_i - p\| \text{ or } m == \|p_j - q\|)$  then
۳:   return  $\pi[p_i, p_j]$ 
۴: else
۵:   if  $(m == \delta_h \rightarrow (\pi[p_i, p_j]))$  then
۶:      $k =$  the index of vertex  $p_k$  between  $p_i$  and  $p_j$  that
       determines the Hausdroff distance.
۷:     if  $(p_k \cdot x < p \cdot x)$  then
۸:        $i = \min_{k \leq k' \leq j} \{\|p_{k'} - p\|\}$ 
۹:       return  $\pi[p_i, p_j]$ 
۱۰:    else
۱۱:     if  $(p_k \cdot x < q \cdot x)$  then
۱۲:        $j = \min_{i \leq k' < k} \{\|p_{k'} - q\|\}$ 
۱۳:       return  $\pi[p_i, p_j]$ 
۱۴:    else
۱۵:     if  $(p \cdot x \leq p_k \cdot x \leq q \cdot x)$  then
۱۶:        $i' = \min_{k \leq k' \leq j} \{\|p_{k'} - p\|\}$ 
۱۷:        $j' = \min_{i \leq k' < k} \{\|p_{k'} - q\|\}$ 
۱۸:        $H = \min\{\delta_h \rightarrow (\pi[p_{i'}, p_j]), \delta_h \rightarrow (\pi[p_i, p_{j'}]),$ 
          $\delta_h \rightarrow (\pi[p_{i'}, p_{j'}])\}$ 
۱۹:       if  $(H == \delta_h \rightarrow (\pi[p_{i'}, p_j]))$  then
۲۰:         return  $\pi[p_{i'}, p_j]$ 
۲۱:       else
۲۲:         if  $(H == \delta_h \rightarrow (\pi[p_i, p_{j'}]))$  then
۲۳:           return  $\pi[p_i, p_{j'}]$ 
۲۴:         else
۲۵:           return  $\pi[p_i, p_{j'}]$ 
۲۶:         end if
۲۷:       end if
۲۸:     end if
۲۹:   end if
۳۰: end if
۳۱: end if
۳۲: end if

```

الگوریتم بر روی هر کدام از مسیرهای جهت‌دار موجود در پایگاه داده به صورت میانگین چند ثانیه طول می‌کشد.

$$\delta_F(\pi'_{opt}, pq) = \max\{\|p_i - p\|, \|p_j - q\|\}$$

• اگر مقدار ماکزیمم عبارت  $\delta_h \rightarrow (\pi[p_i, p_j], pq)$  باشد، آنگاه

ابتدا رأس  $p_k$  بین  $p_i$  و  $p_j$  که فاصله هاسدورف جهت‌دار بین  $\pi[p_i, p_j]$  و پاره‌خط افقی را تعیین می‌کند، مشخص می‌کنیم. با توجه به موقعیت این رأس سه حالت زیر را خواهیم داشت:

۱- اگر رأس  $p_k$  سمت چپ پاره‌خط افقی واقع شده باشد، آنگاه اندیس  $i$  را به رأسی قبل از  $p_k$  منتقل می‌کنیم که فاصله آن از نقطه  $p$  مینیمم باشد. بنابراین،  $i = \min_{k \leq k' \leq j} \{\|p_{k'} - p\|\}$

۲- اگر رأس  $p_k$  سمت راست پاره‌خط افقی واقع شده باشد، آنگاه اندیس  $j$  را به رأسی قبل از  $p_k$  منتقل می‌کنیم که فاصله آن از نقطه  $q$  مینیمم باشد. بنابراین،  $j = \min_{i \leq k' < k} \{\|p_{k'} - q\|\}$

۳- اگر رأس  $p_k$  بین نقطه  $p$  و  $q$  واقع شده باشد، آنگاه اندیس  $i$  و  $j$  را به اندیس‌های جدید  $i'$  و  $j'$  مطابق دو مورد بالا منتقل می‌کنیم. سپس عبارات  $\delta_h \rightarrow (\pi[p_{i'}, p_j])$

و  $\delta_h \rightarrow (\pi[p_i, p_{j'}])$  را محاسبه می‌کنیم و هر کدام که مقدارش مینیمم گردید را گزارش می‌کنیم. جزئیات بیشتر این مرحله در الگوریتم ۲ نشان داده شده است.

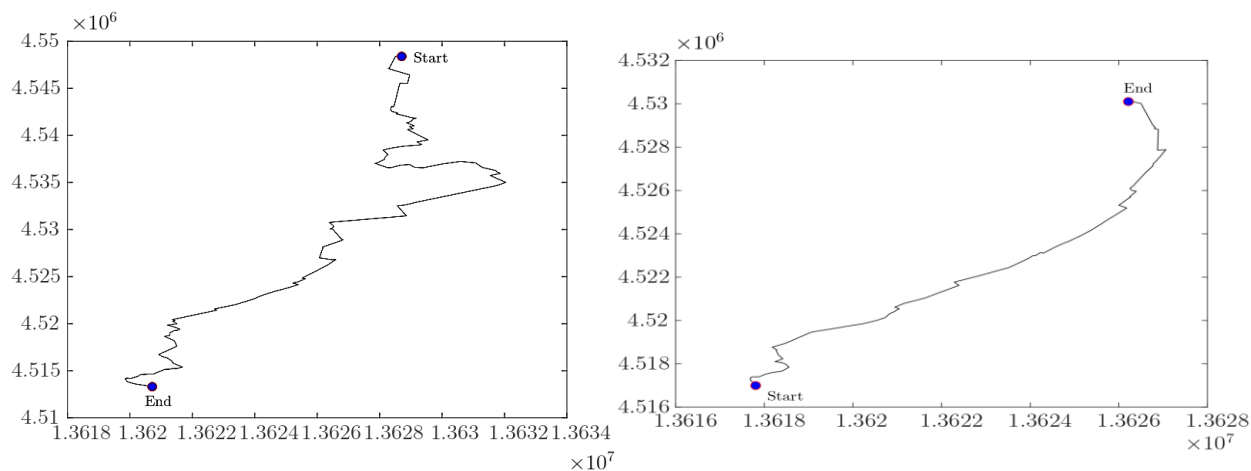
تحلیل پیچیدگی زمانی الگوریتم ارائه شده: اندیس  $i$  و  $j$  در خط اول الگوریتم ۱ با کمک ساختمان داده جستجوی نزدیک‌ترین همسایه در زمان  $O(\log n)$  محاسبه می‌گردد. خط ۴، ۷ و ۹ را می‌توان در بدترین حالت در زمان  $O(n)$  محاسبه کرد (در بدترین حالت باید  $n$  نقطه را بررسی کرد). بنابراین، الگوریتم ۱ به زمان

$O(n)$  نیاز دارد. محاسبه خط اول الگوریتم ۲ با استفاده از ساختمان داده‌های ارائه شده در زمان  $O(\log^2 n)$  محاسبه می‌گردد. خطوط ۸، ۱۲، ۱۶ و ۱۷ را می‌توان در بدترین حالت در

زمان  $O(n)$  محاسبه کرد. بنابراین، الگوریتم ۲ به زمان  $O(n)$  نیاز دارد [۱۲]. بنابراین الگوریتم ارائه شده در مقاله در زمان  $O(n)$  زیرمسیر را گزارش می‌کند. میانگین زمان صرف شده برای اجرای الگوریتم در پایگاه داده‌های استفاده شده در این مقاله در جدول ۱ آورده شده است. مقادیر جدول نشان می‌دهد که اجرای

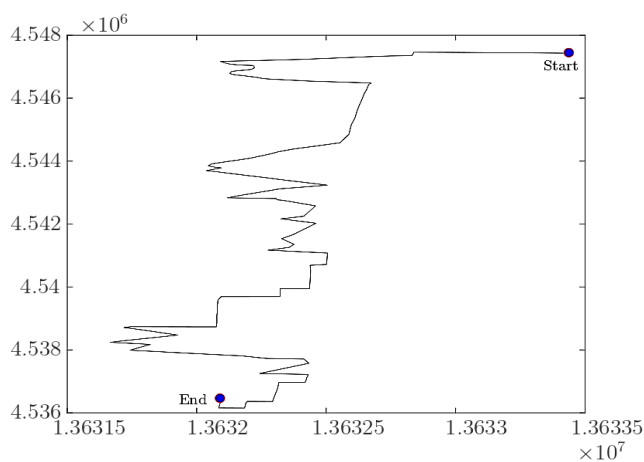
پایگاه داده	زمان بر حسب ثانیه
ECG۲۰۰	۱/۶۴۲۴
CBF	۳/۹۳۶۶
GunPoint	۶/۴۰۹۴
FaceUCR	۳/۹۷۸۷
FaceAll	۴/۱۵۲۴
Plane	۵/۳۵۰۴
SwedishLeaf	۴/۰۱۸۸
SIGSPATIAL	۳/۹۱۲۵

جدول ۱. زمان صرف شده برای اجرای الگوریتم



ب

الف



ج

شکل ۳. (الف): مسیر جهت‌دار روبه جلو (ب): مسیر جهت‌دار روبه عقب. (ج): مسیر جهت‌دار خنثی

## ۵. نتایج تجربی

در این قسمت نتایج تجربی بر روی الگوریتم ارائه شده در قسمت قبل ارائه و کیفیت الگوریتم و زمان آن گزارش می‌شود. مسائل پرس‌وجوی فرشه مسائل جدیدی هستند. نتایج تئوری و تجربی زیادی در حوزه پرس‌وجوی فرشه وجود ندارد. به طور خاص، مسئله بررسی شده در این مقاله نه از نظر تئوری و نه از نظر تجربی تا-کنون بررسی نشده است و الگوریتم ارائه شده در بخش قبلی، اولین الگوریتم ابتکاری برای این مسئله است. با توجه به عدم وجود الگوریتم‌های دیگری برای حل مساله، امکان مقایسه بین الگوریتم‌های مختلف به صورت تجربی وجود ندارد. در ادامه، کیفیت این الگوریتم با اجرای آن روی چند پایگاه داده بررسی می‌شود.

پنج پایگاه داده از مجموعه UCR که شامل ۸۵ پایگاه داده از مسیرهای جهت‌دار یک بعدی است را برای اجرای الگوریتم انتخاب کرده‌ایم [۱۹]. هم‌چنین از پایگاه داده که مربوط به سفرهای جاده‌ای در سانفرانسیسکو می‌باشد، استفاده کرده‌ایم. این پایگاه داده، پایگاه داده‌ای است که در چالش سال ۲۰۱۷ مربوط به SIGSPATIAL استفاده شده است. این پایگاه داده از ۲۰۰۰۰ مسیر جهت‌دار تشکیل شده است که تعداد رئوس مسیرها بین ۱۱ تا ۷۶۹ می‌باشد. این پایگاه داده دارای یک ویژگی خوب است، طول یال‌های مسیرهای جهت‌دار موجود در پایگاه داده کوچک است و این ویژگی باعث می‌شود که الگوریتم قادر باشد در اکثر موارد پاسخ بهینه را پیدا کند. پرس‌وجو که یک پاره‌خط افقی است به صورت تصادفی تولید شده است. برای هر مسیر جهت‌دار ده پاره‌خط افقی به صورت تصادفی تولید شده است (دو نقطه برای مولفه دو سر



$distj$  یا  $disti$  فاصله فرشه را تعیین می‌کند. تنها در ۱۰٪ پرس‌وجوها عبارت مربوط به فاصله هاسدورف، فاصله فرشه را تعیین می‌کند و الگوریتم ارائه شده علیرغم اصلاحات صورت گرفته قادر نیست که جواب بهینه را پیدا کند. کیفیت الگوریتم برای مسیرهای جهت‌دار روبه‌عقب و مسیرهای جهت‌دار خنثی در پایگاه داده ۹۳/۳٪ است. الگوریتم قادر نیست که در مرحله اول برای اکثر مسیرهای جهت‌دار روبه‌عقب، اندیس  $i$  یا  $j$  را با کمک دیسک  $B_i$  یا  $B_j$  تغییر بدهد. در واقع الگوریتم زیرمسیری را گزارش می‌کند که نقطه شروع و انتهای آن یکسان است. در ۶/۷٪ پرس‌وجوها عبارت مربوط به فاصله هاسدورف، فاصله فرشه را تعیین می‌کند و الگوریتم ارائه شده نمی‌تواند زیرمسیر بهینه را پیدا کند. در واقع اصلاحات انجام شده در فاصله هاسدورف در الگوریتم ۲ همواره به درستی عمل نمی‌کند. اگر جهت پاره‌خط افقی را تغییر بدهیم، یعنی پاره‌خط افقی از سمت راست به چپ باشد، مسیر جهت‌دار رو به عقب تبدیل به مسیر جهت‌دار رو به جلو می‌شود و نتایج هم تغییر خواهد کرد.

کیفیت الگوریتم ارائه شده برای پایگاه داده ۹۸/۸٪ است و تنها در یک درصد از پرس‌وجوها زیرمسیر بهینه گزارش نمی‌گردد. تنها علت اصلی برای خطای الگوریتم اصلاحاتی است که در الگوریتم ۲ در قسمت فاصله هاسدورف صورت گرفته است. برای حذف کردن یک درصد خطا باید اصلاحات این قسمت را بهبود داد.

**تحلیل پایگاه داده GunPoint:** نتایج نشان می‌دهد که کیفیت الگوریتم برای مسیرهای جهت‌دار روبه‌جلو در این پایگاه داده ۹۰/۳۸٪ است. در واقع الگوریتم ارائه شده تقریباً در ۱۰٪ پرس‌وجوها برای مسیرهای جهت‌دار روبه‌جلو قادر نیست که جواب بهینه را گزارش کند. کیفیت الگوریتم برای مسیرهای جهت‌دار روبه‌عقب و مسیرهای جهت‌دار خنثی در پایگاه داده تقریباً ۹۵٪ است.

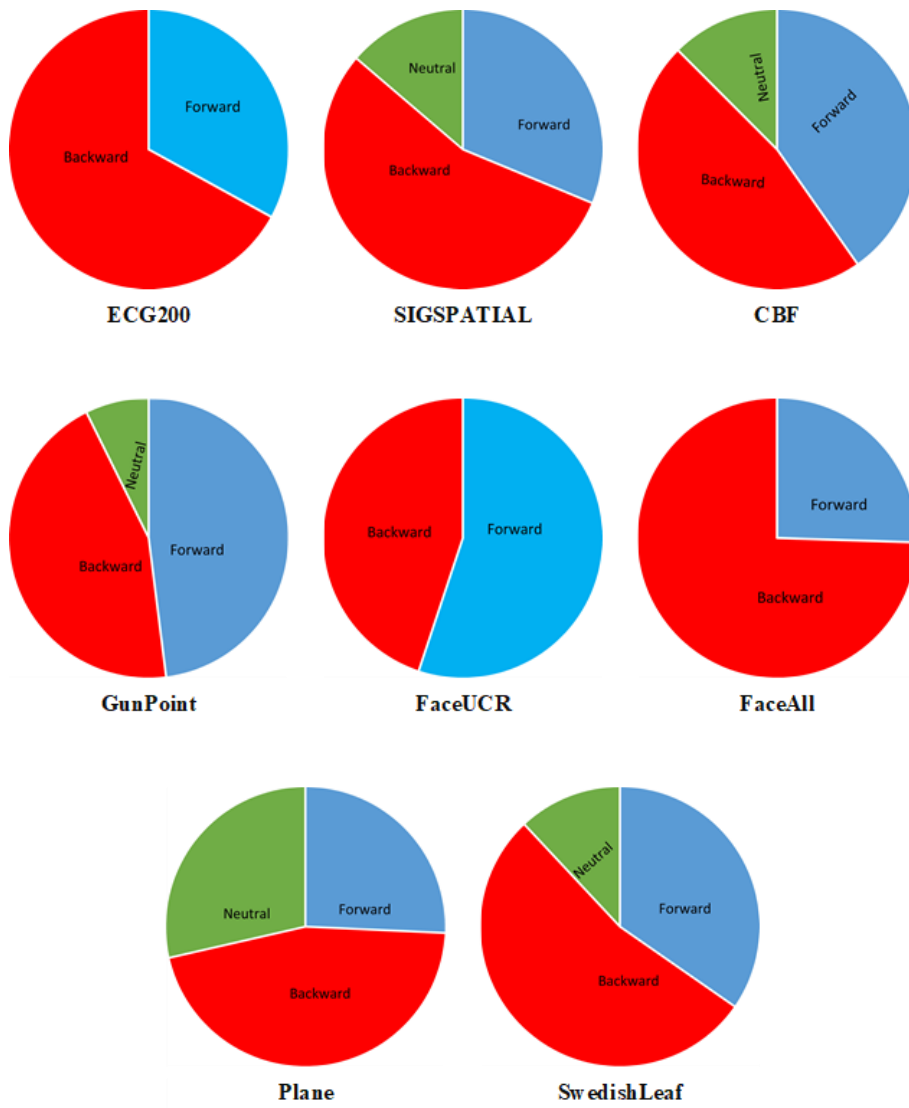
پاره‌خط و یک نقطه برای مولفه  $y$  به صورت تصادفی تولید کرده‌ایم) و الگوریتم بر روی آن اجرا شده است.

الگوریتم ارائه شده در نرم‌افزار متلب بر روی سیستم عامل ویندوز ۱۰ پیاده‌سازی شده است. الگوریتم بر روی لپ‌تاپ لنوو با پردازنده مرکزی اینتل مدل core-i۷ ۴۵۱۰ با سرعت ۲ گیگاهرتز و RAM هشت گیگابایت اجرا شده است.

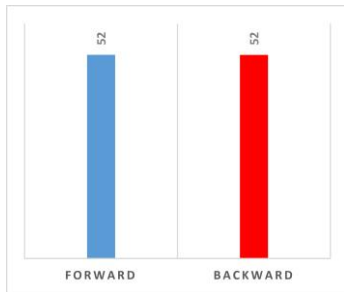
در پایگاه داده‌های استفاده شده با توجه به جهت پاره‌خط افقی که از سمت چپ به راست می‌باشد، سه نوع مسیر جهت‌دار وجود دارد. مسیر روبه‌جلو مسیری است که شی متحرک در بیشتر مواقع رو به جلو حرکت می‌کند به این معنی که تعداد یال‌های رو به جلو بیشتر از نصف کل یال‌ها است و نقطه شروع حرکت سمت چپ نقطه انتها قرار دارد. نوع بعدی، مسیر جهت‌دار روبه‌عقب می‌باشد. در این نوع مسیر جهت‌دار، شی متحرک در بیشتر مواقع رو به عقب حرکت می‌کند به این معنی که تعداد یال‌های برگشتی بیشتر از نصف کل یال‌ها است و نقطه شروع حرکت سمت راست نقطه انتها قرار دارد. نوع آخر مسیر جهت‌داری است که شی متحرک تقریباً به صورت مساوی به جلو و عقب حرکت کرده است. این نوع را مسیر جهت‌دار خنثی می‌نامیم. شکل ۳ انواع مختلف مسیر جهت‌دار را نشان می‌دهد.

شکل ۴ توزیع مسیرهای جهت‌دار مختلف در پایگاه داده‌های استفاده شده را نشان می‌دهد. همان‌طور که در شکل ۴ مشخص شده است بیشترین قسمت از هر پایگاه داده شامل مسیرهای جهت‌دار روبه‌عقب است. در پایگاه داده FaceAll، ECG۲۰۰ و FaceUCR مسیر خنثی وجود ندارد و فقط شامل مسیرهای روبه‌جلو و روبه‌عقب می‌باشد. در ادامه الگوریتم را بر روی انواع مختلف مسیرهای جهت‌دار در پایگاه داده‌های ذکر شده اجرا کرده و کیفیت الگوریتم را گزارش می‌کنیم. شکل ۵ میزان درستی الگوریتم در پایگاه داده‌های استفاده شده را نشان می‌دهد. در ادامه نتیجه چند پایگاه داده را تحلیل خواهیم کرد.

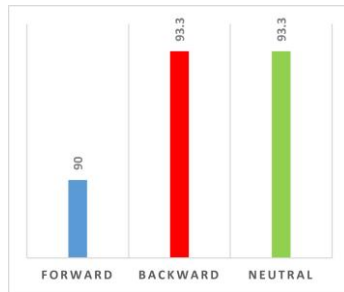
**تحلیل پایگاه داده SIGSPATIAL:** نتایج نشان می‌دهد که کیفیت الگوریتم برای مسیرهای جهت‌دار روبه‌جلو در این پایگاه داده ۹۰٪ است. در واقع الگوریتم ارائه شده فقط در ۱۰٪ پرس‌وجوها برای مسیرهای جهت‌دار روبه‌جلو قادر نیست که جواب بهینه را گزارش کند. الگوریتم اندیس  $i$  یا  $j$  را با کمک دیسک  $B_i$  یا  $B_j$  در مرحله اول الگوریتم برای اکثر مسیرهای جهت‌دار روبه‌جلو تغییر می‌دهد. همچنین در مرحله دوم الگوریتم ماکزیمم عبارات



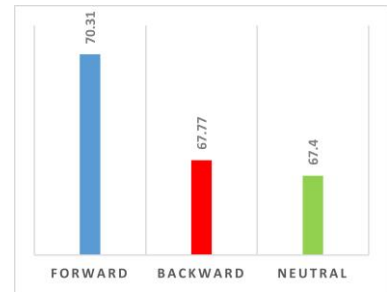
شکل ۴. توزیع مسیرهای جهت‌دار در پایگاه داده‌های استفاده شده



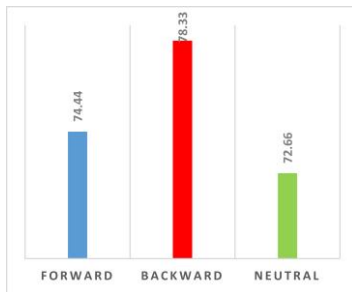
**ECG200**



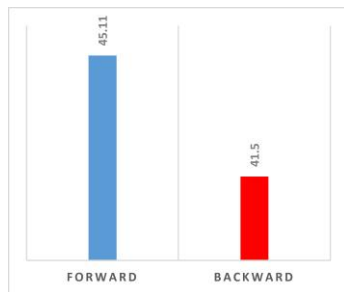
**SIGSPATIAL**



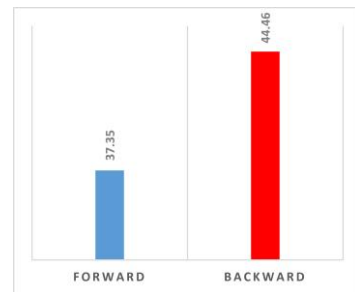
**CBF**



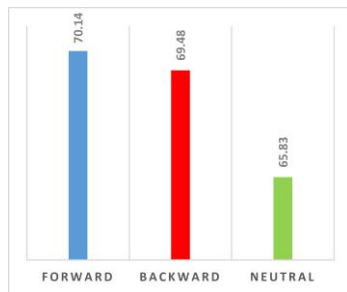
**GunPoint**



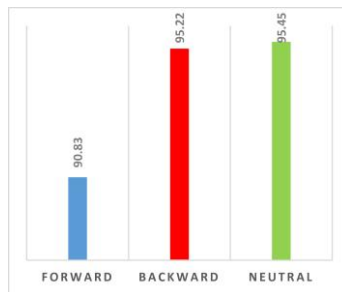
**FaceUCR**



**FaceAll**



**Plane**



**SwedishLeaf**

شکل ۵. درستی الگوریتم ارائه شده برای مسیرهای جهت‌دار مختلف در پایگاه داده‌های استفاده شده

۲. چگونه می‌توان الگوریتمی ارائه کرد که بتواند مسئله را برای هر نوع پرس‌وجو حل کند؟
۳. آیا می‌توان الگوریتم را به صورتی تغییر داد که خطا نداشته باشد و همواره پاسخ بهینه را تولید کند؟

### مراجع

- [۱] Felix Hausdorff. *Grundzuge der mengenlehre*, ۱۹۴۹
- [۲] Helmut Alt. The computational geometry of comparing shapes. In *Efficient Algorithms*, volume ۵۷۶۰ of *Lecture Notes in Computer Science*, pages ۲۳۵-۲۴۸, ۲۰۰۹.
- [۳] Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno. Metro: measuring error on simplified surfaces. In *Computer graphics forum*, pages ۱۶۷-۱۷۴, ۱۹۹۸
- [۴] Maurice Fréchet. Les ensembles abstraits et le calcul fonctionnel. *Rendiconti del Circolo Matematico di Palermo*, pages ۱-۲۶, ۱۹۱۰.
- [۵] Helmut Alt and Michael Godau. Measuring the resemblance of polygonal curves. In *Proceedings of the eighth annual symposium on Computational geometry*, pages ۱۰۲-۱۰۹, ۱۹۹۲.
- [۶] Helmut Alt and Michael Godau. Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, pages ۷۵-۹۱, ۱۹۹۵.
- [۷] E Sriraghavendra, K Karthik, and Chiranjib Bhattacharyya. Fréchet distance based approach for searching online handwritten documents. In *Ninth International Conference on Document Analysis and Recognition*, pages ۴۶۱-۴۶۵, ۲۰۰۷.
- [۸] Mark de Berg, Atlas F Cook, and Joachim Gudmundsson. Fast Fréchet queries. *Computational Geometry*, pages ۷۴۷-۷۵۵, ۲۰۱۳
- [۹] Joachim Gudmundsson and Michiel Smid. Fréchet queries in geometric trees. In *European Symposium on Algorithms*, pages ۵۶۵-۵۷۶. Springer, ۲۰۱۳
- [۱۰] Anne Driemel and Sariel Har-Peled. Jaywalking your dog: computing the Fréchet distance with shortcuts. *SIAM Journal on Computing*, pages ۱۸۳۰-۱۸۶۶, ۲۰۱۳
- [۱۱] Mark de Berg and Ali D Mehrabi. Straight-path queries in trajectory data. *Journal of Discrete Algorithms*, pages ۲۷-۳۸, ۲۰۱۶.
- [۱۲] Mark de Berg, Ali D Mehrabi, and Tim Ophelders. Data structures for Fréchet queries in trajectory data. In *Proceedings of the ۲۹th Canadian Conference on Computational Geometry (CCCG ۲۰۱۷)*, pages ۲۱۴-۲۱۹, ۲۰۱۷.

نکات زیر در مورد اجرای الگوریتم ارائه شده بر روی پایگاه داده‌های ذکر شده بدست آمده است:

- اگر جهت پاره‌خط افقی را تغییر بدهیم، یعنی پاره‌خط افقی از سمت راست به چپ باشد، مسیر جهت‌دار رو به عقب تبدیل به مسیر جهت‌دار رو به جلو می‌شود و نتایج هم تغییر خواهد کرد.
  - دلیل اصلی برای پایین بودن خطای الگوریتم ارائه شده این است که طول یال‌های مسیرهای جهت‌دار موجود در پایگاه داده‌های مورد بررسی کوچک است. بنابراین اصلاحات صورت گرفته در مرحله اول و دوم الگوریتم ارائه شده به خوبی عمل می‌کند و الگوریتم قادر است در اکثر موارد پاسخ بهینه را پیدا کند.
  - هر چقدر ماکزیمم طول یال در پایگاه داده بیشتر باشد، میزان خطای الگوریتم بالا می‌رود.
- جدول ۱ میزان اختلاف میان پاسخ الگوریتم ارائه شده و جواب بهینه، برای پایگاه داده‌های استفاده شده را نشان می‌دهد. همان‌طور که مشخص است، بیشترین اختلاف مربوط به پایگاه داده FaceAll است.

جدول ۲. میزان اختلاف با مقدار بهینه

پایگاه داده	Forward	Backward	Neutral
ECG۲۰۰	۰/۳۱	۰/۳۳	
CBF	۰/۴۸	۰/۵۱	۰/۵۶
GunPoint	۰/۱۶	۰/۰۸	۰/۰۶
FaceUCR	۱/۱۳	۱/۲	
FaceAll	۱/۳۱	۱/۱۵	
Plane	۰/۳۹	۰/۳۷	۰/۵
SwedishLeaf	۰/۵۱	۰/۵۳	۰/۶
SIGSPATIAL	۰/۲۲	۱/۵۳	۰/۴۳

### ۶. نتیجه‌گیری

در این مقاله مسئله تشابه در حوزه فاصله فرشه به صورت تجربی مورد بررسی قرار گرفت. به این صورت که مسیر جهت‌دار  $\pi$  به عنوان ورودی داده شده است. در زمان پرس‌وجو پاره‌خط افقی  $Q$  توسط کاربر مشخص می‌شود. الگوریتم ابتکاری ارائه شده زیرمسیری از مسیر جهت‌دار را گزارش می‌کند به صورتی که فاصله فرشه میان زیرمسیر گزارش شده و پاره‌خط افقی  $Q$  بین تمام زیرمسیرهای ممکن مینیمم می‌باشد. کیفیت الگوریتم ارائه شده بر روی چند پایگاه داده بررسی و گزارش شد. مقاله را با ارائه چند مسئله باز به پایان می‌رسانیم:

۱. چگونه می‌توان این مسئله را برای حالتی که پاره‌خط غیرافقی باشد، حل کرد؟

۲۰۱۵. [http://www.cs.ucr.edu/eamonn/time\\_series\\_data](http://www.cs.ucr.edu/eamonn/time_series_data).
- [۲۰] Joachim Gudmundsson and Michiel Smid. Fast algorithms for approximate Fréchet matching queries in geometric trees. *Computational Geometry*, pages ۴۷۹-۴۹۴, ۲۰۱۵
- [۲۱] Bahram Sadeghi Bigham, and Samaneh Mazaheri. Survey on Metrics for Shape Matching Based on Similarity, Scaling and Spatial Distance. In The ۷th International Conference on Contemporary Issues in Data Science, pages ۱۳-۲۳, ۲۰۱۷
- [۱۳] Maïke Buchin, Ivor van der Hoog, Tim Ophelders, Rodrigo I Silveira, Lena Schlipf, and Frank Staals. Improved space bounds for Fréchet distance queries. In Proceeding of the ۳۶th European Workshop on Computational Geometry, pages ۱-۷, ۲۰۲۰.
- [۱۴] Joachim Gudmundsson, Majid Mirzanezhad, Ali Mohades, and Carola Wenk. *International Journal of Computational Geometry & Applications*, pages ۱۶۱-۱۸۷, ۲۰۱۹
- [۱۵] Mark de Berg, Joachim Gudmundsson, and Ali D Mehrabi. A dynamic data structure for approximate proximity queries in trajectory data. In *Proceedings of the ۲۰th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages ۱-۴, ۲۰۱۷
- [۱۶] Julian Baldus and Karl Bringmann. A fast implementation of near neighbors queries for Fréchet distance (GIS cup). In *Proceedings of the ۲۰th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages ۱-۴, ۲۰۱۷.
- [۱۷] Matteo Ceccarello, Anne Driemel, and Francesco Silvestri. Fresh: Fréchet similarity with hashing. In *Workshop on Algorithms and Data Structures*, pages ۲۵۴-۲۶۸, ۲۰۱۹.
- [۱۸] Fabian Dütsch and Jan Vahrenhold. A filter-and-refinement-algorithm for range queries based on the Fréchet distance (gis cup). In *Proceedings of the ۲۰th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages ۱-۴, ۲۰۱۷
- [۱۹] Yanping Chen, Eamonn Keogh, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen, and Gustavo Batista. The UCR time series classification archive, July

اندازه‌گیری میزان تشابه مسیرهای جهت‌دار بر روی داده‌های هندسی

## بررسی رابطه بین مدیریت منابع انسانی الکترونیکی، چابکی و اثربخشی سازمانی در

### سازمان‌های زیرمجموعه وزارت ارتباطات و فناوری اطلاعات

عبدالرضا بهادری فرد\* رضا رسولی\*\*

\* دانشجوی دکتری مرکز تحصیلات تکمیلی دانشگاه پیام نور

\*\* عضو هیات علمی دانشگاه پیام نور

تاریخ دریافت: ۱۳۹۹/۰۸/۱۲ تاریخ پذیرش: ۱۳۹۹/۱۱/۲۱

نوع مقاله: پژوهشی

#### چکیده

مدیریت منابع انسانی برای سال‌ها به فعالیت‌های اداری و روزمره مانند کارمندیابی، گزینش و آموزش که تمام انرژی و زمان متخصصان منابع انسانی سازمان‌ها را صرف می‌کرد، محدود شده بود. بخش‌های منابع انسانی اغلب آنقدر درگیر چنین فعالیت‌هایی بودند که دیگر زمانی برای فعالیت‌هایی که ارزش افزوده برای سازمان دارند، مانند مدیریت دانش، مدیریت فرهنگ سازمانی و بازاریابی و بازسازی استراتژیک سازمان نداشتند. امروز برای کاهش هزینه‌های اداری و افزایش سرعت ارائه خدمات نیازمند راهکارهای بهتر، سریع‌تر و هوشمندانه‌تر مدیریت منابع انسانی هستیم. مدیریت منابع انسانی الکترونیکی<sup>۱</sup>، پاسخی به چالش‌های امروزه مدیریت منابع انسانی است. پیدایش این مفهوم، اهداف، عملکرد و اثربخشی حوزه منابع انسانی را دست‌خوش تحولات زیادی نموده است. با ظهور فناوری اطلاعات<sup>۲</sup>، نقش، جایگاه و حتی وظایف مدیریت منابع انسانی در سازمان با تغییرات عمده‌ای مواجه شد. متخصصان منابع انسانی که تا پیش از این اغلب وظایف درجه دو سازمانی و فعالیت‌های غیر ارزش‌افزا را انجام می‌داده‌اند، با بکارگیری گسترده فناوری اطلاعات در حوزه منابع انسانی به شرکای استراتژیک سازمان تبدیل شدند. جامعه آماری تحقیق کلیه کارشناسان و مدیران سازمان‌های زیرمجموعه وزارت ارتباطات و فناوری اطلاعات بوده است. حجم نمونه برابر ۳۸۴ تعیین گردید. این تحقیق را می‌توان یک تحقیق میدانی دانست که از ابزار پرسشنامه و مصاحبه استفاده می‌نماید. به این منظور ترکیبی از مصاحبه، پرسشنامه و اسناد و مدارک استفاده می‌شود. همچنین به منظور شناسایی شاخص‌ها و مولفه‌های مدیریت الکترونیکی منابع انسانی در سازمان‌های مورد بررسی در بخش کیفی از ابزار مصاحبه و پرسشنامه و گروه‌های کانونی بر اساس مدل اولیه تحقیق در روش داده بنیاد استفاده می‌شود. در بخش کمی بعد از این مرحله با استفاده از روش رتبه بندی اولویت بندی این شاخص‌ها با استفاده از پرسشنامه مقایسه زوجی تعیین می‌گردد و در نهایت نیز به منظور تایید برازش کلی مدل از پرسشنامه تعیین وضعیت شاخص‌ها استفاده می‌شود. روایی این پرسشنامه‌ها و مصاحبه‌ها از طریق روش CVR تعیین می‌گردد و پایایی نیز با استفاده از روش آلفای کرونباخ تعیین می‌گردد.

**واژگان کلیدی:** منابع انسانی، مدیریت منابع انسانی الکترونیکی، فناوری اطلاعات، چابکی، اثربخشی.

<sup>۱</sup> Electronic Human Resource Management (e-HRM)

<sup>۲</sup> Information Technology (IT)

## ۱- مقدمه

جهان امروز چه در سطح کشورهای توسعه یافته و یا در حال توسعه به سرعت در حال تغییر از اقتصادی‌های سنتی به سمت اقتصادهای مدرن می‌باشد [۱]. در دنیای امروز، دانش و نوآوری اساسی‌ترین عامل پیشرفت در عرصه‌های اقتصادی و صنعتی محسوب می‌شود.

نیروی انسانی سازمان‌ها در عصر جدید به عنوان یکی از مهم‌ترین منابع مطرح می‌باشند. نیروی انسانی می‌تواند در نهایت بر اثربخشی و کارایی سایر منابع نیز اثرگذار باشد. توجه به نیروی انسانی در جوامع پیشرفته به صورت روزافزونی گسترش یافته است به صورتیکه بسیاری از سازمان‌ها به سمت استفاده از نیروهای دانشی حرکت کرده‌اند [۲]. این امر در کنار رشد و توجه به تکنولوژی‌های جدید سبب شده است تا کشورهای توسعه یافته بتوانند زیرساخت‌های خوبی را برای توسعه کشور از بستر توسعه منابع انسانی ایجاد نمایند. اگر یک کشور را متشکل از اجتماع منابع انسانی آن در نظر بگیریم، این منابع به صورت مستمر با یادگیری و توسعه در نقش‌های کلیدی خود در قالب خانواده، جامعه، محیط کار و غیره می‌توانند در نهایت به ایجاد یک جامعه واحد و توسعه یافته کمک کنند. در این مفهوم می‌توان توسعه را مرهون منابع انسانی یک اکوسیستم دانست [۳]. به همین طریق در هرشاخه‌ای از صنعت نیز می‌توان توسعه را در قالب بهبود مستمر منابع انسانی آن صنعت مطرح دانست [۴].

چالش جدی برای مدیران در صنایع مختلف به منظور توسعه، همگام سازی برنامه‌های منابع انسانی در همین راستا می‌باشد. در واقع ایجاد همسویی بین کلیه بخش‌ها و نیروی انسانی در سازمان‌ها بخصوص زمانی که کشور در یک اوضاع نامساعد اقتصادی می‌باشد دارای چالش‌های جدی تری است [۵].

در این میان رشد استفاده از فناوری اطلاعات در کلیه سازمان‌ها سبب شده است تا استفاده از این فناوری در حوزه‌های منابع انسانی نیز به ادبیات مدیریتی این حوزه وارد شود [۶]. امروزه بسیاری از افراد با استفاده از شبکه‌هایی

همچون لینکدین اقدام به معرفی توانمندی‌های خود می‌نمایند که می‌تواند فرآیندهای جذب نیروی انسانی را با تغییرات اساسی رو به رو سازد. در برخی حوزه‌ها بخصوص در ارتباط با فعالیت‌های فناوری اطلاعات امروزه شرکت‌ها از همین کانال‌ها اقدام به شناسایی نیروهای مورد نظر خود می‌نمایند. از سویی نیروی انسانی نیز قبل از ورود به سازمان با بررسی شبکه‌های اجتماعی و اطلاعات در خصوص سازمان به نوعی خود اولین گام‌ها برای اجتماعی سازی را می‌پیماید [۷]. از سویی استفاده از فناوری‌های اطلاعات به منظور آموزش نیروی انسانی نیز به شدت رونق یافته است به صورتیکه استفاده از شبکه‌های شبیه سازتا شبیه سازی‌های عملیاتی در شرکت‌های ساختمانی دربخش آموزش الکترونیک واز راه دور موید این موضوع می‌باشد. بنابراین مدیریت منابع انسانی در عصر حاضر با استفاده از خدمات الکترونیک و فناوری اطلاعات دچار تغییرات بنیادینی شده است [۸].

امروزه کمتر کسب و کار را می‌توان مشاهده نمود که از ابزارها و امکانات فناوری اطلاعات در بهبود روش‌های تولید واریه خدمات خود استفاده نکرده باشد. از سویی صناعی که از این ابزارها استفاده نموده اند نیز رشد قابل توجهی را گزارش نموده اند [۹].

نکته قابل ملاحظه در این صنایع این می‌باشد که معمولا نیروهای انسانی مشغول در آن غالبا از افراد دانشی می‌باشند. این افراد دارای ویژگی‌های شخصیتی و اجرایی متمایز و خاص خود می‌باشند که در سال‌های اخیر بخصوص در حوزه‌های منابع انسانی کمتر مورد توجه قرار گرفته اند بصورتیکه می‌توان به وضوح خلا تحقیقاتی را در این زمینه مشاهده نمود [۱۰].

حال سوال مطرح شده این می‌باشد که چگونه می‌توان با طراحی مدلی چنین نیروهایی با مشخصه‌های منحصر به فرد را بر اساس روش‌های نوین مدیریت منابع انسانی بکارگیری نمود تادر نهایت ضمن برآورده نمودن اهداف تعیین شده سازمانی با استفاده از کاهش هزینه‌ها افزایش



اثربخشی را از طریق بکارگیری فناوری اطلاعات ایجاد نمود؟

چالش دوم و مهم این است که بعد از طراحی چنین مدلی چگ.نه می‌توان اثربخشی مناسبی را در سازمان ایجاد نمود. چراکه عملکرد یک مدل منابع انسانی مناسب را می‌توان در بهبود اثربخشی و دستیابی سازمان به اهداف نهایی جستجو کرد. این موضوع هم در سازمان‌های خصوصی و دولتی مطرح می‌باشد. در این مقاله سعی شده است علاوه بر مروری بر تحقیقات قبلی مرتبط با موضوع مقاله و بررسی تحلیلی بر ادبیات موضوع مقاله با استفاده از نتایج استخراجی از مصاحبه‌ها و تحلیل آماری نتایج پرسشنامه‌های توزیع شده علاوه بر شناسایی مولفه‌های اصلی و فرعی مدل تحقیق به تاثیر گذاری متقابل و در نهایت مدل اثر گذار بر کارایی و اثربخشی را ترسیم نمود بدیهی است که ساختار تحقیق انجام شده بر پایه اصول علمی برای شناسایی عواملی طراحی و اجرا گردیده که نهایتاً ما را به سمت اثربخش بودن مدل منابع انسانی الکترونیک در سازمانهای زیر مجموعه وزارت ارتباطات رهنمون سازد.

## ۲- پیشینه تحقیق

در این بخش سعی شده است که به عنوان پیشینه به تحقیقاتی بپردازیم که در تمام آنها محقق بدنبال شناسایی، آرایه و یا بررسی نقش فناوری اطلاعات در نگاه جدید به منابع انسانی بوده اند، در تمامی این تحقیقات محققین بدنبال آرایه راهکارهایی بوده اند که با توجه به تغییرات شگرف در فناوری اطلاعات به گونه ای چگونگی ایجاد ارتباط بین منابع انسانی و فناری اطلاعات را ترسیم نمایند.

در [۱۱] بیان شده است که امروزه مدیریت منابع انسانی الکترونیکی برای مدیریت منابع انسانی یک نو آوری است. لذا هدف از انجام این تحقیق شناسایی عوامل موثر بر درک مفید بودن مدیریت منابع انسانی الکترونیک میباشد. تحقیق حاضر از حیث هدف کاربردی و از حیث روش اجرای تحقیق توصیفی-همبستگی-پیمایشی و همچنین جامعه آماری تمامی مدیران و سرپرستان اجرایی منابع انسانی شرکت سایپا در سال ۹۲

می‌باشد. نتایج حاکی از آن است که تجربه استفاده از فناوری اطلاعات بر درک مفید بودن و درک سهولت استفاده از کامپیوتر رابطه مثبت و معناداری داشته است و فرهنگ سازمانی تاثیر مثبت و معنی داری بر درک سهولت استفاده از کامپیوتر داشته و نقش منابع انسانی تاثیر مثبت و معنی داری بر درک مفید بودن داشته است و درک سهولت استفاده از کامپیوتر نیز تاثیر مثبت و معنی داری بر درک مفید بودن و پذیرش مدیریت الکترونیک منابع انسانی داشته است و همی‌نطور درک مفید بودن تاثیر مثبت و معنی داری بر پذیرش مدیریت الکترونیک منابع انسانی داشته است.

در [۱۲] بیان شده است که در قرن بیست و یکم از یک طرف نیروی انسانی در ادبیات اقتصادی، اجتماعی و سیاسی نقش سرمایه به خود گرفته و از عامل تولید پا به عرصه سرمایه گذاشته است و از طرف دیگر تکنولوژی به عنوان زیربنای توسعه منابع انسانی در سازمانها و جوامع ایفای نقش نموده و چارچوب و محورهای اساسی آن را تعیین و محتوی و ماهیت آن را نیز متناسب با نیاز جامعه و انسان تدوین می‌نماید. در توسعه منابع انسانی همواره باید ظرفیتهای جدید انسانی شناخته شده و عملیاتی و کاربردی شود تا توسعه قابلیت‌های انسانی به صورت یک فرایند دایم انجام شود و این امر مستلزم بهره گیری از مدیریت تکنولوژی است.

در [۱۳] بیان شده است که این پژوهش با هدف بررسی نقش فناوری اطلاعات و ارتباطات در توسعه منابع انسانی کارکنان سازمان تعاون روستایی استان فارس انجام شده است. نتایج پژوهش نشان داد رابطه معناداری بین فناوری اطلاعات و ارتباطات و توسعه منابع انسانی وجود دارد؛ همچنین نتایج گویای آنست که بین فناوری اطلاعات و ارتباطات و توسعه دانش کارکنان، فناوری اطلاعات و ارتباطات و نگرش کارکنان، فناوری اطلاعات و ارتباطات و توسعه مهارت کارکنان و فناوری اطلاعات و ارتباطات و توسعه رفتار کارکنان رابطه مثبت و معناداری وجود دارد.

مناسبی از مفهوم نشده است، از طرفی به علت بی‌توجهی به عوامل مؤثر بر استقرار آن با مشکلاتی مواجه شده است. این مقاله پژوهشی با رویکردی آمیخته اقدام به شناسایی عوامل مؤثر بر استقرار EHRM نموده که این عوامل با توجه به الگوی جامع استراتژیک شناسایی گردیده، لذا آنچه مقاله را کم‌نظیر نموده، کثرت‌گرایی روش گردآوری اطلاعات است. همچنین در انتها با روش "کیو" که روشی کمی کیفی می‌باشد، مدل مفهومی استقرار مدیریت منابع انسانی الکترونیک در سازمان‌های دولتی استان خراسان رضوی ارائه گردیده است.

در تحقیقی به بررسی شاخص‌های پیاده‌سازی مدیریت منابع انسانی در سازمان‌ها و توسعه این شاخص‌ها در ابعاد مختلف پرداخته شده است. در این تحقیق که به صورت مطالعه مقالات و نشریات مختلف بوده است در نهایت شاخص‌های اصلی از زبان خبرگان صنعت اتومبیل مورد بررسی قرار گرفته است [۱۶]. نتایج تحلیل‌ها نشان داده است که سطح عوامل زمینه‌ای و محیطی در سازمان‌ها به نسبت عوامل فردی و سازمانی نقش کمتری در توسعه مدل‌های پیاده‌سازی مدیریت منابع انسانی نقش داشته اند. در این میان حمایت مدیران ارشد سازمانی به شدت در توسعه منابع انسانی سازمان‌ها مؤثر بوده است.

رابطه پذیرش EHRM و سیستم مدیریت عملکرد با اثربخشی تصمیمات در مؤسسات آموزش عالی نیجریه را بررسی کرده است؛ مقاله وی نشان می‌دهد که قبل از هر پیشرفت و موفقیت چشمگیری برای ایجاد تغییر، درک درست از یک برنامه تغییر است [۱۷]. همچنین چیزی که شانس موفقیت را در استقرار EHRM در مؤسسات آموزش عالی نیجریه افزایش می‌دهد، پذیرش برنامه‌های تغییر است. چون در مؤسسات آموزش عالی، افراد دارای تخصص فعالیت می‌کنند، اثربخشی برنامه‌های مدیریت منابع انسانی الکترونیک به پذیرش اعضای هیأت علمی آن بستگی دارد. مقاله به این مهم می‌پردازد که توفیق اجرای مدیریت منابع انسانی الکترونیک به تطابق آن با سیستم مدیریت عملکرد در مؤسسات آموزش عالی نیز

در [۱۴] بیان شده است که هدف اصلی این پژوهش پاسخ به این سؤال است که آیا توزیع ویژگی‌های سرمایه انسانی در شرکت‌های دانش‌بنیان ایران بر مبنای سبک‌های مختلف نوآوری تفاوت معنادار دارد؟ این مقاله با نظرجویی از ۱۳۲ مدیر ارشد در شرکت‌های دانش‌بنیان مستقر در پارک‌های علم و فناوری دانشگاه تهران و پردیس به تعیین سبک‌های نوآوری پرداخته، سپس به سنجش ویژگی‌های سرمایه انسانی در شرکت‌های دانش‌بنیان در چهار سبک نوآوری با طبقه‌بندی علم محور، فناوری اطلاعات محور، بازار محور و هزینه محور می‌پردازد. نتایج نشان می‌دهد که ویژگی‌های سرمایه انسانی در سبک‌های مختلف نوآوری، تفاوت معناداری ندارند و شرکت‌های دانش‌بنیان در جامعه مورد مطالعه، از سرمایه انسانی متناسب با سبک نوآوری خود، بهره‌مند نیستند یا سرمایه انسانی ما اساساً برای تطابق با سبک‌های نوآوری متداول پرورش نیافته‌اند. همچنین نتایج به‌دست‌آمده با نتایج مطالعات، پیرامون سرمایه انسانی و نوآوری در کشورهای اروپایی مقایسه شد. در میان سبک‌های نوآوری به لحاظ توصیفی سبک نوآوری بازار محور، بالاترین میانگین ویژگی یکتایی و ارزش استراتژیک را نشان می‌دهد که نشانگر تمرکز شرکت‌های دانش‌بنیان بازار محور بر جذب نیروهای یکتا و ارزش‌آفرین است.

در [۱۵] بیان شده است که در دهه‌های گذشته، فناوری، تاثیر چشم‌گیری بر فرآیندها و عملکردهای مدیریت منابع انسانی داشته است. برای مثال، فناوری، خصوصاً اینترنت، کمک کرده است تا بسیاری از فرآیندهای منابع انسانی با بهره‌وری بالاتری طی شود. سیستم‌های جدید متخصصان منابع انسانی را قادر ساخته تا خدمات بهتری برای ذینفعان سازمان خود فراهم کنند، مدیریت منابع انسانی الکترونیک بهترین مصداق این نوع از سیستم‌هاست، کاربرد آن در عصر حاضر آنقدر گسترده شده است که کاربرد آن به یک ضرورت تبدیل شده است. البته این سیستم در ایران با مسائلی روبروست؛ از طرفی تبیین

بستگی دارد و در صورتیکه برنامه ارزیابی عملکرد مرتبط با مدیریت منابع انسانی الکترونیک نباشد، خود می‌تواند یک مانع باشد. همچنین مقاله توانسته به نقش سخت افزار مناسب به عنوان یک کاتالیزور نگاه کند. مقاله با ارائه یک مدل مفهومی در راستای ارتباط EHRM با سیستم مدیریت عملکرد و رابطه پذیرش EHRM با اثربخشی تصمیمات به پایان می‌رسد.

در [۱۸] اشاره شده است که در بین صاحبان نظران، توافق و اجماع بر روی عوامل زمینه ساز موفقیت کارکنان سازمانی وجود ندارد. برخی مهارت کارکنان، برخی نگرش کارکنان و برخی رفتار کارکنان در سازمان را عامل توفیق میدانند. لیکن در عصر حاضر با وجود انقلاب در فناوری اطلاعات، نقش تخصص و مهارت‌های فنی برای کارکنان امری ناگزیر است. توانایی کارکردن با فضای دیجیتال نیز امری غیر قابل انکار برای سازمان محسوب می‌شود. آنچه این مقاله مهم می‌داند تطابق منابع انسانی با فناوری اطلاعات و پذیرش IT در سازمان است. لذا مقاله اشاره می‌کند که راه تطابق و پذیرش EHRM است. در ادامه مقاله ضمن معرفی برخی کمپانیهای موفق در حوزه فناوری اطلاعات، خاطر نشان می‌کند حتی در این قلمرو از فضای کاری نیز، اگر کارکنان مدیریت منابع انسانی الکترونیک را باب میل خود ندانسته و به عبارتی آن را نپذیرند، مدیریت منابع انسانی الکترونیک موفق نخواهد بود.

در [۱۹] عنوان شده است در عصر اطلاعات، شرکتها به طور فزایندهای به سمت فناوری اطلاعات حرکت کرده اند و اولین حوزه مورد توجه نیز منابع انسانی سازمان بوده است. لیکن چالشهایی نیز متوجه حوزه منابع انسانی است. دو چالش عمده در الکترونیک کردن منابع انسانی را تطبیق فناوری اطلاعات با روحیات و توانایی کارکنان و همچنین تطبیق و پذیرش ساختارهای مجازی معرفی کرده است. در ادامه مقاله عنوان میکند که دانش فناوری اطلاعات از سوی کارکنان عامل مهم برای پذیرش EHRM است که در صورت فقدان می‌تواند مانعی بزرگ بر اجرای

سیستم باشد. همچنین حرکت به سمت ساختارهای مجازی را نیز زمینه ساز موفقیت مدیریت منابع انسانی الکترونیک می‌داند. مقاله با راهکارهایی برای مدیران جهت توفیق اجرا برنامه‌های مدیریت منابع انسانی الکترونیک به پایان می‌رسد.

در [۲۰] اشاره شده است که تبدیل فعالیت‌های منابع انسانی به فعالیت‌های الکترونیک، می‌تواند به شکل استاندارد اتفاق بیفتد. اعتقاد نویسندگان این است که با جهانی شدن، نباید برای ایالت محلی، الگوی متفاوت برای مدیریت منابع انسانی الکترونیک ارائه نمود. بلکه باید با یک نسخه استاندارد، نسبت به اجرا و استقرار مدیریت منابع انسانی الکترونیک در شرکتهای جند ملیتی و ادارات محلی امریکا عمل نمود. لیکن اشاره میکند عوامل محیطی می‌تواند نقش مهمی در اجرای این آرمان داشته باشد. مسائل فرهنگی، اقلیمی، ساختاری و حتی ماهیت کار ادارات می‌تواند تعیین کننده یک الگوی خاص باشند. به عبارتی به تعداد ادارات، شاهد الگوها و پارادایم‌های تلوریک متفاوت برای EHRM خواهیم بود. مقاله اذعان میکند که توجه به عوامل محیطی می‌تواند یک عامل توفیق در پذیرش مدیریت منابع انسانی الکترونیک در ایالت مختلف امریکا باشد و با تبیین این که مدیریت منابع انسانی الکترونیک با سیستم‌های اطلاعات منابع انسانی یا همان HRIS متفاوت است داشتن سیستم‌های اطلاعاتی در سازمان می‌تواند در استقرار EHRM کمک کند. به عبارتی HRIS را گامی مقدماتی اما مهم برای مدیریت منابع انسانی الکترونیک می‌داند.

[۲۱] مدیریت الکترونیک منابع انسانی را به عنوان راهی برای اجرای راهبردها، سیاستها و شیوه‌های مدیریت منابع انسانی در سازمان از طریق حمایت آگاهانه و مستقیم با استفاده کامل از کانالهای مبتنی بر فناوری وب تعریف نموده اند در این رویکرد تمامی فعالیت‌های منابع انسانی توسط فناوری وب پشتیبانی می‌شوند. برخلاف سیستم‌های اطلاعات منابع انسانی که دسترسی به آنها تنها در اختیار عده محدودی از متخصصان و مشاوران

مدیریت منابع انسانی است در این رویکرد جدید نه تنها متخصصین دپارتمانهای منابع انسانی بلکه کلیه مدیران و کارکنان سازمان به راحتی میتوانند به اطلاعات پرسنلی مورد نیاز دسترسی داشته باشند و نسبت به ارزیابی و بهینه نمودن آنها اقدام نمایند. به همین علت از آن به عنوان یک نوآوری در منابع انسانی یاد می‌شود؛ اولاً به علت فرصتهایی که در روابط کارکنان و مدیران ایجاد میکند ثانیاً بعلاوه استفاده از ابزار فناوری اطلاعات تسهیلاتی را برای طراحی ابزارها و ادوات مدیریت منابع انسانی ایجاد می‌کنند که بدون فناوری اطلاعات امکانپذیر نیست.

در نهایت با جمع بندی نظرات سایر محققین در پیشینه تحقیق به این واقعیت دست پیدا میکنیم که استفاده از تکنولوژی و فناوری اطلاعات با توجه به تغییرات سیستماتیک جوامع سازمانی و روند جهانی توجه به منابع انسانی به عنوان سرمایه در سازمانها، واقعیتی انکارناپذیر بوده و رفتن به سمت الکترونیکی نمودن مدیریت منابع انسانی راه حل بهینه در ایجاد بهره وری، کارآمدی و نهایتاً اثربخشی و کارایی سازمانهای قرن حاضر هستند که البته با توجه به روند رو به رشد تکنولوژی شاید در آینده ای نزدیک این نیاز و این مدل نیز دستخوش تغییرات عمده ای گردد. اما متأسفانه علیرغم روند افزایشی استفاده از فناوری اطلاعات در زندگی شخصی و سازمانی در ایران هنوز مدیریت سنتی بر سازمانها جو غالب سازمانهای دولتی و حتی خصوصی ما می‌باشد بهمین منظور برای شناسایی راه ورود به این عرصه پاره ای جز مطالعات موردی و تحقیق پیرامون طراحی مدل های بومی مدیریت در کشور خواهیم داشت، تحقیق حاضر هم انجام این مهم را در تخصصی ترین وزارتخانه مرتبط با فناوری اطلاعات به انجام رسانده است.

### ۳- تعاریف و پیشینه مدیریت منابع انسانی الکترونیکی

مدیریت منابع انسانی الکترونیکی، یک مفهوم کاملاً جدید نمی‌باشد و از اوایل دهه ۱۹۹۰، مورد استفاده قرار گرفته است. عبارت مدیریت منابع انسانی الکترونیکی به صورت گسترده مورد استفاده قرار گرفت اما هنوز یک تعریف پذیرفته شده جهانی در مورد آن وجود ندارد [۲۲] [۲۳].

این واژه اغلب مترادف با واژه‌های مشابه، نظیر منابع انسانی مبتنی بر وب<sup>۳</sup> سیستم‌های اطلاعاتی منابع انسانی<sup>۴</sup>، مدیریت منابع انسانی مجازی<sup>۵</sup>، اینترنات منابع انسانی<sup>۶</sup>، سیستم‌های مدیریت منابع انسانی مبتنی بر کامپیوتر<sup>۷</sup> و پورتال‌های منابع انسانی<sup>۸</sup> در نظر گرفته می‌شود [۲۴] [۲۵].

این اصطلاحات دارای معنای مشابهی هستند اما به منظور دنبال کردن اهداف این تحقیق از اصطلاح مدیریت منابع انسانی الکترونیکی استفاده خواهد شد. مدیریت منابع انسانی الکترونیکی را بعنوان "راهی برای اجرای استراتژیها، خط‌مشی‌ها و شیوه‌های منابع انسانی در سازمان‌ها از طریق یک پشتیبان هدایت شده و آگاه با استفاده کامل از کانال‌های مبتنی بر فناوری وب"<sup>۹</sup> تعریف کرده‌اند [۲۶].

همین نویسندگان در تحقیق بعدی خود این تعریف را بسط داده و بعد ارتباطات منابع انسانی الکترونیکی را هم اضافه کردند، جایی که کارمند و کارفرمایان از طریق مدیریت منابع انسانی الکترونیکی می‌توانند درباره محتوای منابع انسانی به طور اثربخش‌تری با هم ارتباط برقرار نمایند.

مدیریت منابع انسانی الکترونیکی همچنین بعنوان مجموعه‌ای از دانش، قواعد و بهترین شیوه‌ها در نظر گرفته می‌شود که منجر به مدیریت منابع انسانی اثربخش می‌شود [۲۷]. در [۲۸] مدیریت منابع انسانی الکترونیکی را

<sup>۷</sup> computer-based human resource management systems

<sup>۸</sup> human resource portals

<sup>۹</sup> web-technology based channels

<sup>۳</sup> web-based human resources

<sup>۴</sup> human resource information systems (HRIS)

<sup>۵</sup> virtual human resource management

<sup>۶</sup> human resource intranet

اطلاعات، منابع عمده و اصلی سازمان می‌شود. در این بعد، فناوری اطلاعات در ابداع محصول، خدمت و استراتژیک مشارکت می‌کند، همانگونه که در افزایش انعطاف‌پذیری و مزیت رقابتی مشارکت دارد. در این حالت فناوری اطلاعات امکان برقراری ارتباط با فراسوی مرزهای جغرافیایی و سهمیم شدن در اطلاعات آنها را فراهم می‌کند. عبارتی محدودیت‌های زمان و مکان حذف می‌شود. در نتیجه مدیران منابع انسانی با یک فرآیند جدید مواجه شده و خارج از چارچوب کنونی فکر می‌کنند و نوع‌آوری‌های جدیدی را شکل می‌دهند [۳۰] [۳۱] [۳۲].

### ۲-۳ نتایج استقرار مدیریت منابع انسانی

#### الکترونیک

بیر و همکارانش نتایج چهارگانه استقرار مدیریت منابع انسانی الکترونیک در سازمان را به شرح زیر بیان می‌دارند: تعهد بالاتر: تعهد بالا به سطح بالایی از اعتماد میان مدیریت و نیروی کار اشاره دارد. به این مفهوم که به نیروی انسانی سازمان انگیزه داده می‌شود و آنها مشتاق تعامل با مدیریت در مورد تغییرات محیطی سازمان می‌شوند. برای خود کارکنان منابع انسانی بدین معنی است که به آنها بعنوان عامل تغییر نگاه می‌شود. نیروی کار با انگیزه و با درک بالاتر کار می‌کنند و اینکه آنها خود دوست دارند که با تغییرات محیط سازمان و تاثیر آن بر محیط داخلی سازمان در تعامل باشند. در مدیریت منابع انسانی نیز تعهد به معنای ایفا کردن نقش عامل تغییر است. تعهد بالا باعث بوجود آمدن اطمینان بیشتر میان مدیریت و کارکنان خواهد شد. شایستگی بالاتر: شایستگی بالا به قابلیت‌های کارکنان در یادگیری وظایف و نقش‌های جدید اشاره دارد. حرکات به سمت توانمندسازی کارکنان برای یادگیری وظایف و نقش‌های جدید مناسب با شرایط. در واقع سطح اعتماد بالاتر میان مدیریت و کارکنان بوجود می‌آید.

اثربخشی هزینه: مربوط به هزینه‌های مختلف نارضایتی کارکنان مانند مقاومت در مقابل تغییر و یا انحصارات است.

بعنوان پشتیبانی اداری از کارکرد منابع انسانی در سازمان‌ها بوسیله استفاده از فناوری اینترنت "تعریف می‌کنند. اما همچنین بر اهمیت درک این نکته تاکید دارند که معرفی مدیریت منابع انسانی الکترونیک ممکن است منجر به تغییر در محتوا و موضع‌گیری نقش منابع انسانی شود. اتخاذ فناوری‌های مبتنی بر وب برای کارکرد منابع انسانی شامل ترکیب دو جز می‌باشد: استفاده از رسانه الکترونیکی و مشارکت فعال افراد در این فرآیند [۲۹].

### ۱-۳ تاثیر فناوری اطلاعات بر مدیریت منابع

#### انسانی

تاثیر فناوری اطلاعات بر مدیریت منابع انسانی را می‌توان در سه مرحله از کاربرد آن مورد بررسی قرار داد:

#### ۱-۱-۳ اتوماسیون یا خودکار شدن

در این مرحله از فناوری اطلاعات به منظور مکانیزه کردن سیستم‌های دستی و کاهش نیاز به افراد برای انجام کارهای روزمره استفاده می‌شود که حجم کارهای روتین اداری را کاهش می‌دهد. افراد فرصت بیشتری برای فکر کردن پیدا کرده و می‌توانند قابلیت‌های شناختی خود را بکار بگیرند و بیشتر به تفسیر اطلاعات بپردازند.

#### ۲-۱-۳ اطلاع رسانی و اطلاعات

وقتی فناوری اطلاعات، اطلاعات را ارائه می‌دهد فراتر از مرحله خودکار شدن و اتوماسیون ظاهر می‌شود تا اثربخشی را افزایش دهد. اطلاع رسانی فناوری اطلاعات سطح شفافیت فعالیت‌ها، حوادث و اهداف را بالا می‌برد. زیرا افراد اطلاعات بیشتر و عمیق‌تری درباره فرآیندهای اساسی تولید و اداره سازمان کسب می‌کنند و در مورد آنچه در سازمان یا واحد آنها رخ می‌دهد، بیشتر می‌دانند. در این حالت فناوری اطلاعات با افزایش سطح کیفیت خدمات و کاهش زمان دسترسی می‌تواند بر جنبه‌های ارتباطی منابع انسانی تاثیر بگذارد.

#### ۳-۱-۳ تحول آفرینی

فناوری اطلاعات شرکت را به عملیات و اقدامات تجاری جدید مجهز می‌کند و تحول آفرینی، خلاقیت، دانش و

### ۳-۳ مقایسه مدیریت منابع انسانی سنتی و

#### مدیریت منابع انسانی الکترونیکی

مدیریت منابع انسانی سنتی دارای تفاوت‌هایی با مدیریت منابع انسانی الکترونیکی می‌باشد. این تفاوت‌ها را می‌توان در فرآیندهای اصلی منابع انسانی شامل کارمندیابی، جبران خدمات و پاداش‌دهی، توسعه منابع انسانی، ایمنی و خدمات کارکنان و حفظ و نگهداری منابع انسانی خلاصه کرد. با توجه به زیر شاخصه‌های این فرآیندها که در جدول شماره ۱ آمده است می‌توان این دو را از هم تفکیک کرد.

در واقع نحوه پرداخت کارکنان رقابتی شده و رضایت کارکنان جلب می‌گردد. همانطور که آریچ بیان می‌کند، منابع انسانی خود می‌تواند نقش کارشناس اجرایی را ایفا کند تا در تناسب و تعدیل قیمت‌ها در سازمان‌ها سهیم باشند. تناسب بالاتر: مربوط به سیستم‌های پرداخت متناسب با عملکرد افراد است. در واقع سیستم پاداش دهی بر مبنای خروجی و بر اساس ورودی‌ها و بازدهی تعریف می‌شود [۳۳].

جدول شماره ۱. مقایسه مدیریت منابع انسانی سنتی و مدیریت منابع انسانی الکترونیکی

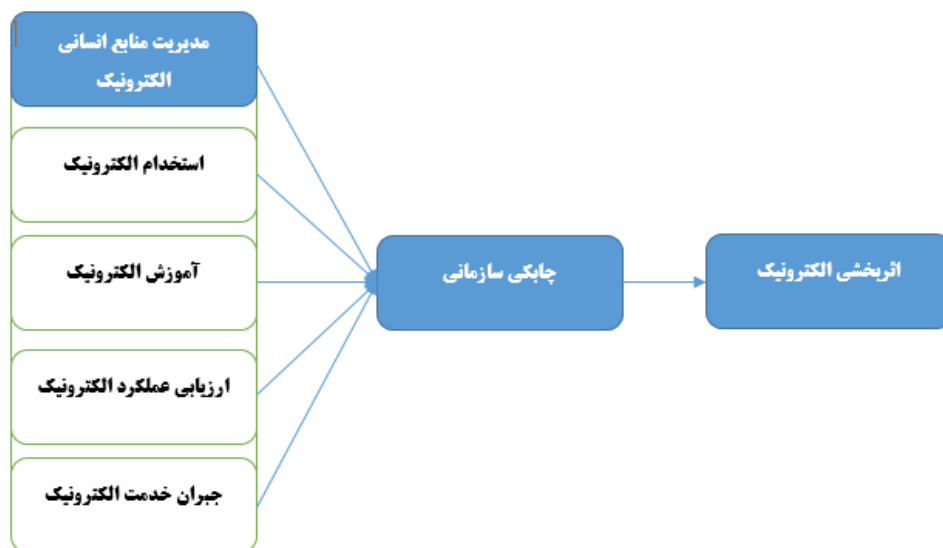
[۳۴]

مدیریت منابع انسانی الکترونیکی	مدیریت منابع انسانی -	فرآیندهای اصلی	
✓ چکیده سوابق الکترونیکی و پست الکترونیکی	✓ چکیده سوابق کاغذی	استخدام	کارمندیابی
✓ بررسی الکترونیکی چکیده سوابق	✓ بررسی دستی چکیده‌های سوابق	انتخاب	
✓ نرم افزارهای ارزیابی	✓ ارزیابی سرپرست و رو در رو	ارزیابی عملکرد	جبران خدمات و پاداش دهی
✓ تاکید بر مالکیت و کیفیت زندگی	✓ تاکید بر دستمزدها و مساعده‌ها	حقوق و دستمزد	
✓ آموزش آنلاین انعطاف‌پذیر ✓ فرآیند توسعه کارمند محور	✓ آموزش داخل کلاس ✓ فرآیند توسعه، مدیر محور	آموزش و توسعه	توسعه منابع انسانی
✓ کارکنان به کمک مسئولان منابع انسانی، مسیرهای شغلی خود را مدیریت می‌کنند.	✓ مدیران منابع انسانی مسیرهای شغلی را برای کارکنان تعیین می‌کنند.	کار راه شغلی	
✓ ملاحظات کار پژوهی ✓ خستگی روانی ✓ محرکهای تنش زای شخصی و مربوط با کار	✓ امنیت ساختمان و تجهیزات ✓ خستگی جسمی ✓ محدود به محرکهای تنش زای کار	بهداشت و امنیت	ایمنی و خدمات کارکنان
✓ تمرکز بر روابط کارمند- کارمند	✓ تمرکز بر روابط کارمند- مدیریت	روابط کار / مسائل کارکنان	

#### ۴\_ روش تحقیق

پارادایم این تحقیق می‌تواند در قالب پارادایم اثبات‌گرایی مطرح نمود، چراکه در این تحقیق به بررسی واقعیت‌های موجود بر اساس شاخص‌های مدیریت منابع انسانی الکترونیک پرداخته و سپس بر این اساس به تحلیل و اثبات واقعیت‌ها پرداخته می‌شود. در لایه دوم می‌توان این تحقیق را می‌توان یک تحقیق کاربردی مطرح دانست چراکه به دنبال ارائه یک مدل کاربردی از سازمان‌های مورد بررسی در حوزه موضوع تحقیق می‌باشیم. در سطح بعدی بر اساس توضیحات ارائه شده این تحقیق به نوعی تحقیقی استقرایی می‌باشد چرا که با توجه به خطاهای موجود در سایر تحقیقات حوزه منابع انسانی این تحقیق و مقدمات آن از نتیجه سایر تحقیقات به صورت محتمل پشتیبانی می‌کند. همچنین می‌توان این تحقیق را از انواع تحقیقات ترکیبی دانست که به صورت تومانی از روش‌های کمی و کیفی استفاده می‌نماید. در بخش بعدی این تحقیق را می‌توان یک تحقیق میدانی دانست که از ابزار

پرسشنامه و مصاحبه استفاده می‌نماید. بنابراین نوع تحقیق پیمایشی می‌باشد و در آن به اکتشاف و توصیف پدیده‌ها و روابط پرداخته می‌شود و به این منظور ترکیبی از مصاحبه، پرسشنامه و اسناد و مدارک استفاده می‌شود. در این تحقیق به منظور جمع‌بندی و ارائه مدل در ابتدا بر اساس پیشینه تحقیقات صورت گرفته شاخص‌های اصلی تهیه و سپس مولفه‌ها و زیر مولفه‌ها شناسایی و ارائه می‌شود. در بخش کمی نیز با توجه به قابلیت دسته‌بندی سازمان‌های مورد بررسی می‌توان از روش خوشه‌ای ساده استفاده نمود. تعداد نمونه تحقیق با استفاده از فرمول کوکران برابر ۳۸۴ نفر در نظر گرفته شده است. به منظور تحلیل اطلاعات نیز از مدل یابی معادلات ساختاری استفاده می‌گردد. در بخش جمع‌آوری اطلاعات از ابزار پرسشنامه استفاده گردید که ضریب آلفای کرونباخ پرسشنامه کلی برابر ۰,۷۵۶ محاسبه شده است که نشان از پایایی مناسب ابزار تحقیق دارد. همچنین روایی تحقیق با استفاده از شاخص CVR محاسبه و مورد تایید قرار گرفت.



شکل شماره ۱. مدل مفهومی اولیه بر اساس پیشینه تحقیقاتی

## ۵-۱ اهداف انجام تحقیق

### ۵-۱-۱ هدف اصلی

تبیین ارتباط بین مدیریت منابع انسانی الکترونیکی و افزایش چابکی سازمانی .  
تبیین ارتباط بین چابکی سازمانی و افزایش اثربخشی سازمانی.

### ۵-۲ هدف‌های فرعی

تبیین ارتباط بین مدیریت منابع انسانی الکترونیکی (جذب و استخدام الکترونیکی) و افزایش چابکی سازمانی .  
تبیین ارتباط بین مدیریت منابع انسانی الکترونیکی (آموزش الکترونیکی) و افزایش چابکی سازمانی .  
تبیین ارتباط بین مدیریت منابع انسانی الکترونیکی (ارزیابی عملکرد الکترونیکی) و افزایش چابکی سازمانی .  
تبیین ارتباط بین مدیریت منابع انسانی الکترونیکی (جبران خدمات الکترونیکی) و افزایش چابکی سازمانی.

چه راهکارهایی به منظور پیاده سازی الگوی مدیریت منابع انسانی الکترونیکی در سازمان‌های نوآور قابل ارائه می‌باشد؟

### فرضیه‌های تحقیق

مدیریت منابع انسانی الکترونیکی موجب افزایش چابکی سازمانی می‌شود.  
چابکی سازمانی موجب افزایش اثربخشی سازمانی می‌شود.

### فرضیه‌های فرعی

مدیریت منابع انسانی الکترونیکی (جذب و استخدام الکترونیکی) موجب افزایش چابکی سازمانی می‌شود.  
مدیریت منابع انسانی الکترونیکی (آموزش الکترونیکی) موجب افزایش چابکی سازمانی می‌شود.  
مدیریت منابع انسانی الکترونیکی (ارزیابی عملکرد الکترونیکی) موجب افزایش چابکی سازمانی می‌شود.  
مدیریت منابع انسانی الکترونیکی (جبران خدمات الکترونیکی) موجب افزایش چابکی سازمانی می‌شود.

## ۷-۱ جامعه آماری

جامعه آماری این تحقیق به دو دسته کمی و کیفی دسته بندی می‌شود. در بخش کیفی به منظور شناسایی مولفه‌ها از خبرگان سازمان‌های فناور در حوزه فناوری اطلاعات استفاده می‌شود که دارای مشخصه‌های زیر می‌باشند:

\_ دارای حداقل ۱۰ سال سابقه کار مرتبط

\_ دارای تحصیلات مرتبط

\_ تسلط به امور فنی و مدیریتی منابع انسانی

\_ همچنین در بخش کمی کلیه کارشناسان و مدیران سازمان‌های زیرمجموعه وزارت ارتباطات و فناوری اطلاعات به عنوان جامعه تحقیق مطرح می‌باشند.

## ۷-۱ روش نمونه‌گیری و تخمین حجم نمونه

در بخش کیفی با توجه به محدود بودن جامعه تحقیق روش نمونه‌گیری سرشماری می‌باشد. در بخش کمی نیز با توجه به قابلیت دسته بندی سازمان‌های مورد بررسی می‌توان از

## ۶-۱ سوالات و فرضیات تحقیق

### ۶-۱-۱ سوال اصلی

آیا مدیریت منابع انسانی الکترونیکی موجب افزایش چابکی سازمانی می‌شود؟  
آیا چابکی سازمانی موجب افزایش اثربخشی سازمانی می‌شود؟  
فرضیه‌های فرعی:

آیا مدیریت منابع انسانی الکترونیکی (جذب و استخدام الکترونیکی) موجب افزایش چابکی سازمانی می‌شود؟  
آیا مدیریت منابع انسانی الکترونیکی (آموزش الکترونیکی) موجب افزایش چابکی سازمانی می‌شود؟  
آیا مدیریت منابع انسانی الکترونیکی (ارزیابی عملکرد الکترونیکی) موجب افزایش چابکی سازمانی می‌شود؟  
آیا مدیریت منابع انسانی الکترونیکی (جبران خدمات الکترونیکی) موجب افزایش چابکی سازمانی می‌شود؟



کمتر از ۵ سال، ۲۲ درصد در گروه بین ۵ الی ۱۰ سال و ۱۰ درصد در گروه سابقه بالای ۱۰ سال بوده اند.

نتایج آمار استنباطی

در این بخش از تحقیق هدف اصلی محقق بررسی ارتباط بین متغیرهای تحقیق می‌باشد. اما در ابتدا باید وضعیت متغیرهای تحقیق مورد ارزیابی قرار گیرد و سپس برازش مرتبط با مدل‌های هر یک از متغیرها ارزیابی شود. در بررسی برازش مدل‌ها شاخص‌های مختلفی ارائه شده است. که معروفترین این شاخص‌ها در قالب RMSEA, GFI, AGFI, CFI, CMIN/DF معرفی شده است. هر یک از این شاخص‌ها معرف بخشی از برازش مدل می‌باشد و اعداد مورد قبول برای هر شاخص متفاوت می‌باشد. جدول شماره ۲ نشان دهنده [۳۵].

روش خوشه ای ساده استفاده نمود. در بخش کیفی این تحقیق حجم نمونه ۱۵ نفر می‌باشد و در بخش کمی نیز حجم نمونه برابر ۳۸۴ نفر می‌باشد.

## ۲\_۷ نتایج توصیفی تحقیق

در بررسی ویژگی‌های جمعیت شناختی تحقیق نتایج این بررسی‌ها نشان داد که مردان با تعداد ۷۰ درصد و زنان با ۳۰ درصد در توزیع جنسیتی این گروه قرار داشته اند. همچنین در بحث تحصیلات می‌توان بیان نمود که ۵۳ درصد نمونه مورد بررسی دارای تحصیلات فوق لیسانس و بالاتر و ۴۷ درصد دارای تحصیلات لیسانس می‌باشند. در بحث سمت سازمانی نیز نتایج این تحقیق نشان داده است که ۱۵,۳۱ درصد کارشناس، ۴۰,۵۵ رئیس اداره، ۳۳,۳۳ مدیر میانی و ۱۰,۸۱ درصد مدیر ارشد می‌باشند. در بخش سابقه کاری نتایج نشان داده است ۶۸ درصد افراد مورد مطالعه در گروه سابقه کاری

جدول شماره ۲. مقادیر قابل قبول شاخص‌های برازش مدل

شاخص	مقدار قابل قبول
GFI	$GFI > 90\%$
AGFI	$AGFI > 90\%$
CFI	$0.90 < CFI < 1$
CMIN/DF	مقدار کمتر از ۳
RMSEA	$RMSEA < 0.08$

می‌بایست مدل مورد ارزیابی مجدد قرار گیرد تا عامل بر هم زننده برازش مدل شناسایی و حذف یا اضافه شود. در نهایت پس از تایید مدل نهایی با استفاده از تحلیل مسیر روابط بین متغیرهای تحقیق مورد ارزیابی قرار می‌گیرد.

سپس مولفه‌هایی که دارای کمتر از ۴ سوال هستند در مدل اندازه‌گیری بررسی می‌شوند. یادآوری می‌شود که این دو مرحله پیش نیاز معادله ساختاری در مدل نهایی هستند. ۱- بررسی اعتبار مدل: بدین منظور بارهای عاملی هر سوال مد

تمامی شاخص‌های بالا به معنی شاخص‌های برازش می‌باشد که در همگی این شاخص‌ها مقادیر کای دو مورد استفاده قرار گرفته است. با توجه به جدول شماره ۲ در صورتیکه مقادیر محاسبه شده برای شاخص‌های مورد ارزیابی در حد قالب قبول باشد برازش مدل تایید می‌شود و در غیر این صورت

## ۳\_۷ تحلیل عاملی تاییدی مؤلفه‌های

### تحقیق (CFA)

برای انجام تحلیل عاملی مولفه‌های تحقیق، ابتدا مولفه‌هایی که دارای ۴ سوال یا بیشتر هستند، مورد تحلیل قرار می‌گیرند.

بار عاملی کمتر است حذف می‌گردد.

۳- نرمال بودن داده‌ها: در آزمون برازش مدل مفهومی تحقیق

لازم است تا نرمال بودن داده‌ها مورد بررسی قرار گیرد.

۴- لازم به ذکر است که تست نرمال بودن داده‌ها به صورت

کلی با انجام آزمون کلموگروف اسمیرونوف (K-S) در بخش

قبلی آورده شده است

## ۷\_۴ تحلیل عاملی تأییدی مدل در حالت اندازه

### گیری

در این بخش در ابتدا هر یک از زیر مدل‌ها به همراه مولفه‌ها

ترسیم می‌شود و سپس با ایجاد ارتباط کواریانس بین

شاخص‌ها به بررسی روایی مدل مفهومی در حالت کلی

می‌پردازیم:

نظر قرار می‌گیرد. سوالاتی که بار عاملی کمتر از ۰/۴ دارند، اثرگذاری آنها در مدل کمتر از حد پذیرفته شده است و بهتر است از مدل حذف گردند تا در برازش مدل مشکلی رخ ندهد [۳۵].

۲- بررسی برازش مدل: model fit شاخص‌های برازش مدل بر

اساس شاخص‌هایی که در فصل سوم آورده شده است بررسی می‌شود. همان‌گونه که بیان شد چنانچه ۴ شاخص مورد

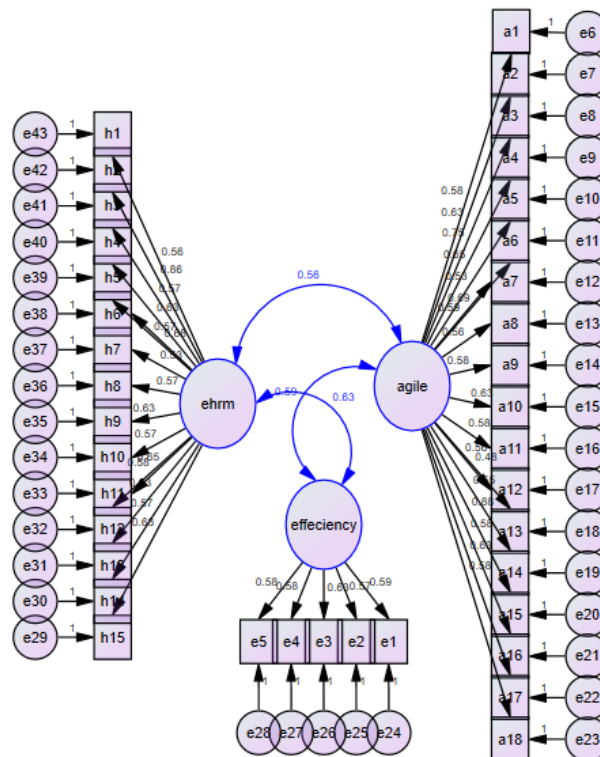
پذیرش قرار گیرند، مدل به طور مناسب برازش شده است.

در صورت عدم تحقق این امر، می‌توان از ارتباط بین سایر متغیرهای اثرگذار بر سوالات (ارتباط بین خطاهای اندازه-

گیری) استفاده کرد؛ تا این شاخص‌ها بهبود یافته و برازش

مدل تأیید گردد. البته در صورت برقراری ارتباط بین خطاها ارتباط منطقی آنها نیز از طریق مراجعه به پرسشنامه باید

مورد تأیید قرار گیرد و در صورت عدم تأیید، سوالی که دارای



شکل شماره ۲. مدل تأیید در حالت اندازه گیری

جدول شماره ۳. شاخص‌های نیکویی برازش مدل تأیید عاملی در حالت اندازه گیری

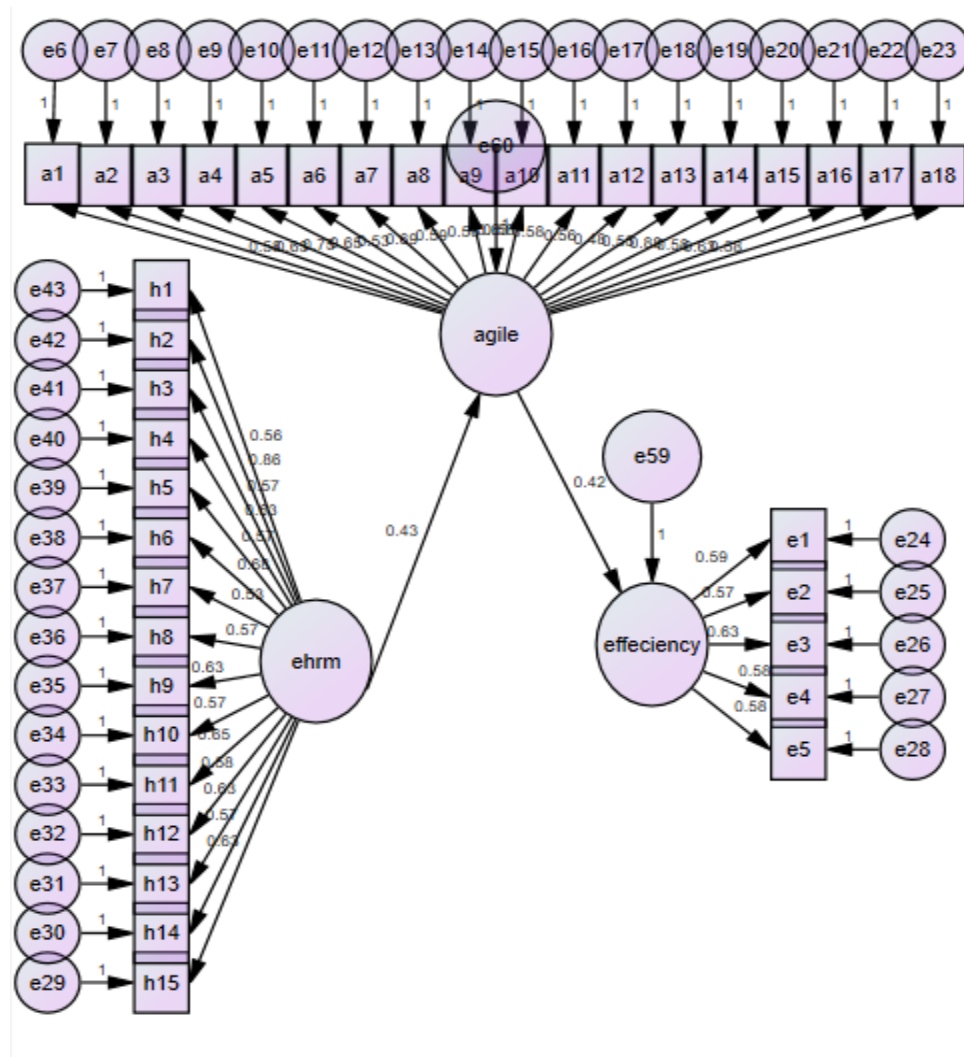
شاخص	مقدار کسب شده	مقدار قابل قبول	وضعیت
GFI	۰,۹۲۱	$GFI > 90\%$	قبول
CFI	۰,۹۳۳	$0,90 < CFI < 1$	قبول
CMIN/df	۲,۹۵	مقدار کمتر از ۳	قبول
RMSEA	۰,۰۸	$RMSEA < 0,08$	قبول

دارد. در نهایت مقدار کای دو بر درجه آزادی ۲,۹۵ محاسبه شده است. استاندارد این مقدار ۳ و کمتر از آن می‌باشد. این به آن معنی است که مدل مورد نظر هماهنگ با الگوی همپراشی بین متغیرهای مشاهده شده است. در نهایت ریشه میانگین مجذورات تقریب یا همان RMSEA مقدار ۰,۰۹ محاسبه شده است. که کم تر از معیار ۰,۱ می‌باشد و بنابراین می‌توان بیان نمود که برازش مدل به طور کلی در سطح مناسبی ارزیابی می‌شود.

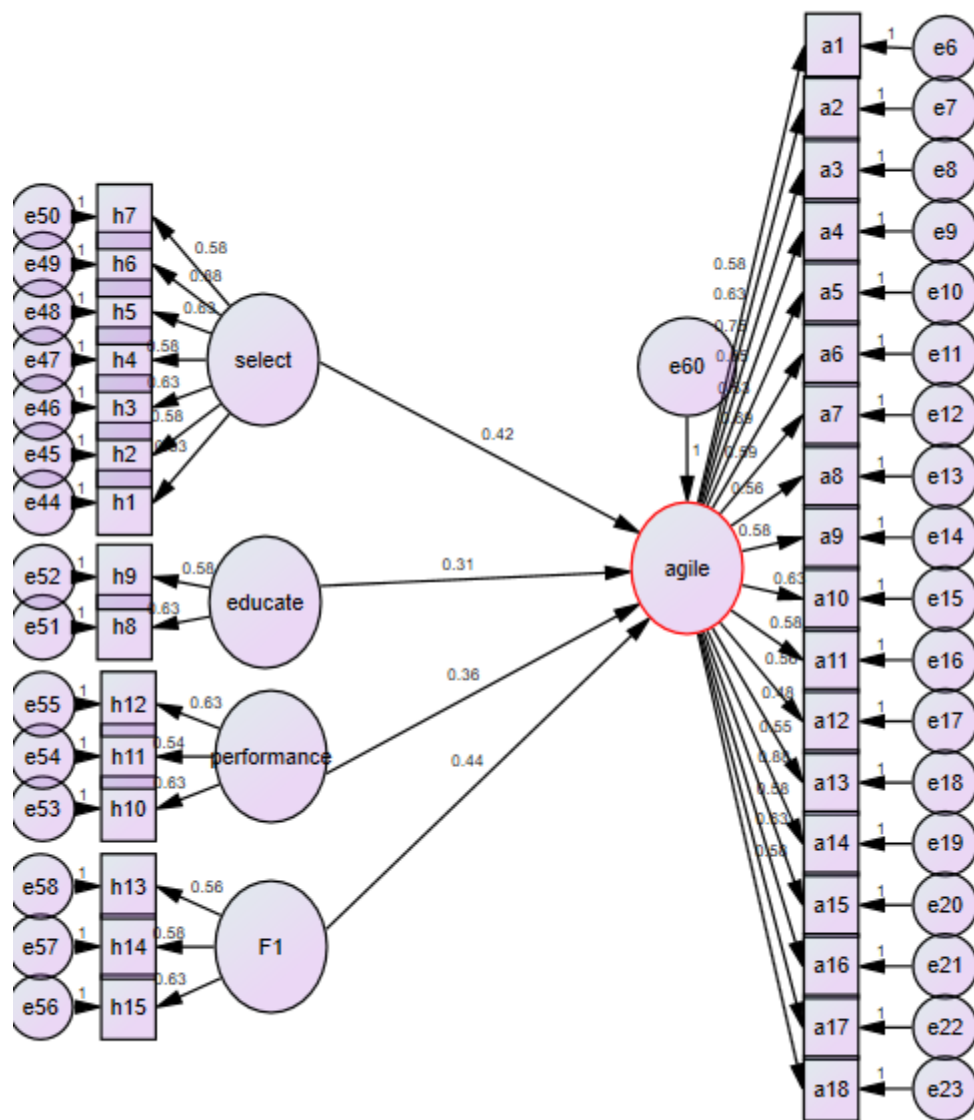
#### ۷\_۵ مدل معادلات ساختاری

پس از بررسی و تأیید مدل اندازه‌گیری، مدل تحقیق ترسیم و روابط بین متغیرهای پنهان (که در واقع همان فرضیات پژوهشی می‌باشند) مورد بررسی قرار می‌گیرند. در این مورد ابتدا از شاخص‌های برازش اطمینان حاصل می‌کنیم و سپس به سراغ بررسی روابط مفروض بین متغیرهای پنهان می‌رویم.

با توجه به جدول شماره ۶ کلیه شاخص‌های مورد بررسی به منظور ارزیابی برازش مدل عاملی تأییدی مورد پذیرش قرار گرفته اند. با توجه به این مطلب می‌توان بیان نمود که برازش این مدل تأیید شده است. همانطور که در این جدول مشاهده می‌شود مقدار GFI، ۰,۹۲۱ محاسبه شده است که بیشتر از معیار ۰,۹ می‌باشد. این بدین معنی است که مقدار نسبی واریانس‌ها و کوواریانس‌ها به صورت مشترک در سطح مناسبی ارزیابی می‌شود. همچنین مقدار شاخص CFI، ۰,۹۳۳ محاسبه شده است که این شاخص نیز بیشتر از حد نصاب ۰,۹ می‌باشد. بدین معنی که مقدار بهبود از طریق مقایسه یک مدل به اصطلاح مستقل که در آن بین متغیرها هیچ رابطه ای نیست با مدل پیشنهادی مورد نظر، در سطح مناسبی قرار



شکل شماره ۳. مدل معادله ساختاری تحقیق شاخص‌های اصلی



شکل شماره ۴. مدل یابی معادلات ساختاری مولفه‌های فرعی

همچنین شاخص‌های برازش مدل مطابق جدول شماره ۳ خواهد بود.

جدول شماره ۳ شاخص‌های نیکویی برازش مدل تأیید مدل ساختاری

شاخص	مقدار کسب شده	مقدار قابل قبول	وضعیت
GFI	۰,۸۳۹	$GFI > ۹۰\%$	قبول
CFI	۰,۹۱۴	$۰,۹۰ < CFI < ۱$	قبول
CMIN/df	۱,۶۲۲	مقدار کمتر از ۳	قبول
RMSEA	۰,۰۵۳	$RMSEA < ۰,۰۸$	قبول

شده است. استاندارد این مقدار ۳ و کمتر از آن می‌باشد. این به آن معنی است که مدل مورد نظر هماهنگ با الگوی همپراشی بین متغیرهای مشاهده شده است. در نهایت ریشه میانگین مجذورات تقریب یا همان RMSEA مقدار ۰,۰۵۳ محاسبه شده است. که کمتر از معیار ۰,۰۸ می‌باشد و بنابراین می‌توان بیان نمود که برازش مدل به طور کلی در سطح مناسبی ارزیابی می‌شود.

در جدول شماره ۸ میزان ضرایب رگرسیونی که بیانگر میزان اثرگذاری مولفه‌ها بر یکدیگر است و معناداری این ضرایب به طور خلاصه آورده شده است.

با توجه به جدول فوق کلیه شاخص‌های مورد بررسی به منظور ارزیابی برازش مدل عاملی تأییدی مورد پذیرش قرار گرفته اند. با توجه به این مطلب می‌توان بیان نمود که برازش این مدل تأیید شده است. همانطور که در این جدول مشاهده می‌شود مقدار GFI، ۰,۸۳۹ محاسبه شده است که کمتر از معیار ۰,۹ می‌باشد. این بدین معنی است که مقدار نسبی واریانس‌ها و کوواریانس‌ها به صورت مشترک در سطح مناسبی ارزیابی می‌شود. همچنین مقدار شاخص CFI، ۰,۹۱۴ محاسبه شده است که این شاخص نیز بیشتر از حد نصاب ۰,۹ می‌باشد. بدین معنی که مقدار بهبود از طریق مقایسه یک مدل به اصطلاح مستقل که در آن بین متغیرها هیچ رابطه ای نیست با مدل پیشنهادی مورد نظر، در سطح مناسبی قرار دارد. در نهایت مقدار کای دو بر درجه آزادی ۱,۷۴۴ محاسبه

جدول شماره ۴. ضرایب رگرسیونی مولفه‌های تحقیق (آزمون فرضیات)

فرضیه	تاثیر مولفه	بر مولفه	ضریب رگرسیونی	عدد معناداری	سطح معنی داری	نتیجه
۱	مدیریت منابع انسانی الکترونیک	چابکی	۰,۴۳	۸,۲۱۴	۰,۰۰۱	تایید
۲	چابکی	اثربخشی سازمانی	۰,۴۲	۸,۱۲۳	۰,۰۲۳	تایید
۳	جذب	چابکی سازمانی	۰,۴۲	۸,۵۴۳	۰,۰۱	تایید
۴	آموزش	چابکی سازمانی	۰,۳۱	۶,۲۱	۰,۰۰۰	تایید
۵	عملکرد	چابکی سازمانی	۰,۳۶	۶,۳۶	۰,۰۰۱	تایید
۶	جبران خدمت	چابکی سازمانی	۰,۴۴	۹,۲۱۴	۰,۰۲۵	تایید

این ارتباط ۰,۳۶ می‌باشد که بیانگر ارتباط مستقیم دو متغیر می‌باشد.

بررسی فرضیه ششم: در بررسی رابطه بین مدیریت منابع انسانی الکترونیک (پاداش و جبران خدمت) و چابکی سازمانی سطح معناداری ۰,۰۲۵ و عدد معناداری ۹,۲۱۴ محاسبه گردید که بیانگر وجود ارتباط بین دو متغیر می‌باشد. ضریب ارتباط دو متغیر ۰,۴۴ می‌باشد که بیانگر ارتباط مثبت و مستقیم بین دو متغیر می‌باشد.

#### ۸\_ الگوی نهایی تحقیق

تاکنون دو مرحله برای دستیابی به الگوی نهایی تحقیق انجام گرفته است. ابتدا و در مرحله اول محقق با بررسی گسترده ادبیات تحقیق، متغیرهای مرتبط با موضوع تحقیق و الگوی اولیه را شناسایی نمود که در فصل دوم ارائه گردید. سپس در مرحله دوم، محقق به مطالعه کیفی پرداخت و براساس مصاحبه‌های عمیق انجام شده با مدیران سازمان‌های تحت نظارت وزارت ارتباطات و فناوری اطلاعات مدل نهایی تدوین شد.

در این بخش از تحقیق، جمع بندی دو روش مذکور (بررسی ادبیات و مطالعه کیفی) در قالب الگوی نهایی تحقیق و کلیه متغیرهای آن در جدول شماره ۹ ارائه می‌گردد.

در این بخش نیز به منظور بررسی رابطه بین متغیرهای تحقیق از روابط رگرسیونی استفاده شده است. نتایج بررسی فرضیات تحقیق در ادامه بیان شده است:

بررسی فرضیه اول: در بررسی رابطه بین مدیریت منابع انسانی الکترونیک و چابکی سازمانی عدد معناداری ۸,۲۱۴ و سطح معنی داری ۰,۰۰۱ محاسبه گردید که نشان دهنده رابطه مثبت بین دو متغیر است. ضریب ارتباط بین دو متغیر منفی (۰,۴۳) به دست آمده است که نشان دهنده رابطه مستقیم بین دو متغیر می‌باشد.

بررسی فرضیه دوم: در بررسی رابطه اثربخشی سازمانی و چابکی سازمانی عدد معنی داری ۸,۱۲۳ محاسبه گردید که نشان دهنده وجود ارتباط بین دو متغیر می‌باشد. ضریب رگرسیونی این ارتباط ۰,۴۲ می‌باشد که بیانگر ارتباط مستقیم بین دو متغیر می‌باشد.

بررسی فرضیه سوم: در بررسی رابطه مدیریت منابع انسانی الکترونیک (جذب و استخدام) و چابکی سازمانی عدد معنی داری ۸,۵۴۳ و سطح معناداری ۰,۰۱ محاسبه گردید که نشان دهنده وجود ارتباط بین دو متغیر می‌باشد. ضریب رگرسیونی این ارتباط ۰,۴۲ می‌باشد که بیانگر ارتباط مثبت و مستقیم دو متغیر می‌باشد.

بررسی فرضیه چهارم: بررسی فرضیه مدیریت منابع انسانی الکترونیک (آموزش) و چابکی سازمانی نشان داده است که عدد معنی داری این ارتباط مقدار ۶,۲۱ محاسبه شده است. بنابراین این ارتباط تایید می‌شود. همچنین ضریب ارتباط بین این دو متغیر مقدار ۰,۳۱ محاسبه شده است که این مقدار مثبت بوده است و بنابراین رابطه از نوع مستقیم می‌باشد.

بررسی فرضیه پنجم: در بررسی رابطه مدیریت منابع انسانی الکترونیک (ارزیابی عملکرد) و چابکی سازمانی عدد معنی داری ۶,۳۶ و سطح معناداری ۰,۰۰۱ محاسبه گردید که نشان دهنده وجود ارتباط بین دو متغیر می‌باشد. ضریب رگرسیونی

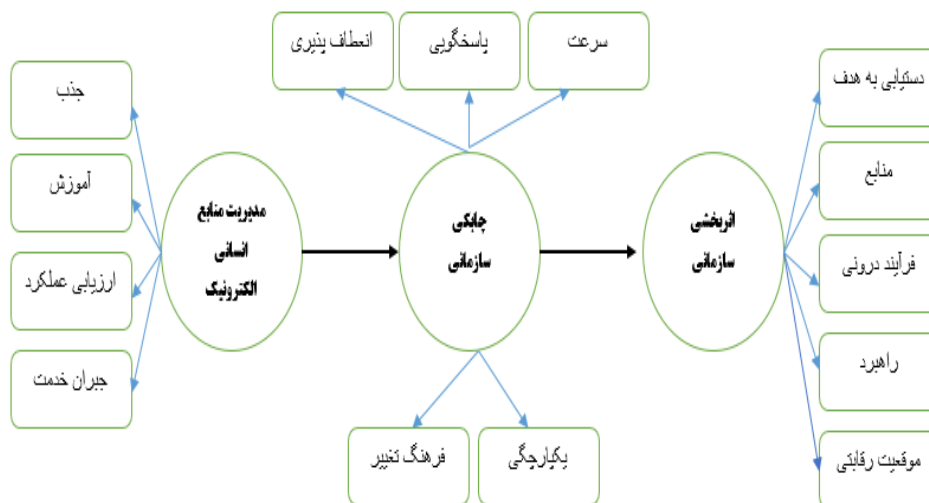
جدول شماره ۹ متغیرهای الگوی نهایی تحقیق (ترکیب دو روش ادبیات موضوعی و تحقیق کیفی)

ردیف	عنوان متغیر	مفاهیم متناظر
۱	جذب و استخدام	جذب از طریق شبکه‌های اجتماعی جذب از طریق سایت‌های استخدامی جذب از طریق ارسال ایمیل جذب از طریق سایت رسمی فرم‌های استخدام الکترونیک پرسشنامه‌های استخدامی الکترونیک داشبورد الکترونیک متقاضیان استخدام
۲	آموزش الکترونیک	استفاده از آموزش مجازی سنجش عملکرد آموزش به صورت آنلاین
۳	ارزیابی عملکرد	ارزیابی عملکرد به صورت آنلاین داشبورد عملکردی آنلاین مدیریت عملکرد آنلاین
۴	جبران خدمت	پرداخت آنلاین حقوق و مزایا محاسبه الکترونیک پاداش خدمات رفاهی آنلاین
۵	انعطاف‌پذیری	انعطاف‌پذیری در ساختار انعطاف‌پذیری در مدیریت انعطاف‌پذیری در زیرساخت‌ها
۶	پاسخگویی	پاسخگویی در برابر مشتریان پاسخگویی در برابر کارکنان پاسخگویی به بازار
۷	فرهنگ تغییر	انطباق با اهداف کسب و کار حمایت محیطی برای آزمایش کردن، یادگیری و نوآوری نگرش مثبت به تغییر بهبود مستمر یادگیری و آموزش مدیریت تغییرات سازمانی



ردیف	عنوان متغیر	مفاهیم متناظر
۸	سرعت	یادگیری و انجام کار و تغییرات در کمترین زمان زمان عملیات و تولید محصول پایین زمان یادگیری و انطباق با محیط
۹	یکپارچگی و پیچیدگی	تلفیق تکنولوژی‌های مختلف پیچیدگی کم ساختاری یکپارچگی میان فرآیندها فرآیندهای آسان و هماهنگ
۱۰	اثربخشی	میزان دستیابی به اهداف استفاده اثربخش از منابع اثربخشی فرآیندهای داخلی اثربخشی راهبردها جایگاه رقابتی

مدل نهایی این پژوهش که در واقع مهم‌ترین خروجی آن محسوب می‌شود، در شکل شماره ۶ نمایش داده شده است. در نتیجه این پژوهش، مدل زیر ارائه گردید که در واقع نوآوری اصلی این پژوهش می‌باشد چرا که در این مدل عوامل موثر بر مدیریت منابع انسانی را با ذکر تمامی ابعاد آمده است. مدل نهایی در قالب شکل زیر ارائه شده است.



شکل شماره ۶. مدل نهایی پژوهش

## ۹\_ پیشنهادهای مبتنی بر نتایج تحقیق

دارا باشد می‌توان ضمن پرداخت مزایای بموقع درخصوص جبران کارایی و اثر بخشی کارکنان با دقت و سرعت عمل نمود. همچنین می‌توان با ارائه سیستم‌های رفاهی آنلاین برای خانواده‌ها و پرسنل در این زمینه از سیستم‌های الکترونیکی استفاده کرد.

✓ در زمینه انعطاف پذیری: با توجه به گستردگی فناوری اطلاعات در زمینه استفاده بهینه از زمان و مکان و عدم وابستگی به هر دو مقوله مذکور می‌توان برای جلوگیری از اتلاف هزینه و زمان با ایجاد محیط کار مجازی و به نوعی دورکاری فضای کاری را از ساختارهای پیچیده فیزیکی به ساختارهای انعطاف پذیر بر اساس نیاز سوق داد.

✓ در زمینه پاسخگویی: در دهه گذشته و توجه به مشتریان و لزوم پاسخگویی سریع و با دقت به آنها به عنوان یکی از مزیت‌های رقابتی در بین سازمانها تبدیل شده است و به مهم لزوم پاسخگویی مناسب به همکاران سازمانی اگر اضافه گردد بیش از پیش ما را به سمت استفاده از سیستم مدیریت مشتریان سوق خواهد داد لذا پیشنهاد می‌گردد در این خصوص از فضای وب و اپلیکشن‌های موبایلی که به شدت در حال گسترش می‌باشد استفاده گردد.

✓ در زمینه فرهنگ تغییر: برای ایجاد یک سازمان کارا و اثربخش نیاز به وجود فرهنگ کارایی و اثر بخشی است که می‌توان با استفاده از آموزش و روشهای نوین آن که در بخش آموزش توضیح داده شد زمینه تغییر فرهنگ و رویکرد سازمان را ایجاد نمود، بدون برنامه برای تغییر در رویکردهای سازمان امکان ایجاد کارایی و اثربخشی وجود نخواهد داشت.

✓ در زمینه سرعت: همانگونه که مشخص است این عنوان شامل ایجاد سرعت در کلیه فرایندها، و رویکردهای سازمان خواهد بود قطعاً استفاده از سیستم‌های الکترونیکی

✓ در زمینه جذب و استخدام: استفاده از فضای وب و شبکه‌های اجتماعی مانند لینکدین و... ضمن افزایش دامنه و وسعت جامعه مخاطبین باعث تنوع مراجعین و همچنین پایین آوردن هزینه‌های مرتبط با فرایند جذب و استخدام کارکنان را در پی خواهد داشت

✓ در زمینه آموزش: استفاده از تکنولوژی‌های نوین مبتنی بر فناوری اطلاعات و فضای مجازی ضمن ایجاد تنوع در روش‌های آموزش باعث افزایش راندمان، کاهش هزینه و گستردگی جامعه مخاطبین بدون محدودیت زمانی و مکانی را در پی خواهد داشت. همچنین در بخش آموزش نیروی انسانی می‌توان از فیلم‌ها و روش‌های ترکیبی جذابیت آموزش را افزایش دهد. با استفاده از فناوری اطلاعات و برخط می‌توان با آموزش‌های بین‌المللی بدون نیاز به حضور در کشورهای دیگر و تنها از طریق آموزش مجازی امکان دسترسی به دانش بین‌المللی را برای کارکنان ایجاد نمود.

✓ در زمینه ارزیابی عملکرد: استفاده از فناوری اطلاعات در ارتباط سیستم‌های حضور و غیاب، کنترل پروژه، کارسنجی، سیستم ارزیابی کارکنان دورکار از طریق فضای وب از جمله مواردی است که ضمن سرعت در ارزیابی، دقت فراوان، مدیریت را جهت ارزیابی عملکرد کارکنان توانمند می‌سازد. همچنین بخش عملکرد می‌توان با استفاده از داشبوردهای تخصصی تولید شده توسط نرم‌افزارهای این حوزه زمینه را برای ارزیابی عملکرد اثربخش تر ایجاد نمود.

برای ارزیابی عملکرد می‌توان در بخش‌های مختلف با استفاده از سیستم‌های API کلیه عملکرد منابع انسانی را به صورت اتومات اجرایی کرد

✓ در زمینه جبران خدمت: پیشنهاد این است که با استفاده از سیستم‌های حقوق و دستمزد یکپارچه که متصل به ارزیابی عملکرد بوده و قابلیت ارزیابی آنلاین کارکنان را

بخش کمی این پژوهش تحلیل نقطه نظرات کارشناسان و مدیران بوده است. توصیه می‌شود در پژوهش‌های آتی از نگاه ارباب رجوع نیز این موضوع مورد مطالعه و بررسی قرار گیرد.

### ۱۰\_ جمع بندی و نتیجه گیری

با بررسی میدانی صورت گرفته و بر اساس نتایج تحقیق و واقعیت‌های حال حاضر دنیا در خصوص پاندمی ویروس کوید ۱۹ بایستی اذعان داشت که در حال حاضر این تغییرات تکنولوژی به تنهایی نیستند که سازمانها را وادار می‌کنند تا برای اثر بخشی بیشتر و کارایی مطلوب تر به سمت استفاده از ابزارهای الکترونیکی در مدیریت منابع انسانی حرکت نمایند بلکه نگاه ویژه ای که شروع این پاندمی به جهان بخشید بیش از پیش ما را به سمت حضور حداقلی در سازمانها و استفاده از فناوری برای دور کاری و ارایه سرویس‌های سازمانی بر پایه وب سوق داد لذا ما چاره ای جز طراحی مدل‌های جدیدی از ساختارهای سازمانی با رویکرد منابع انسانی الکترونیک نخواهیم داشت، ساختارهای که علاوه بر کارایی از سرعت و انعطاف پذیری بسیار بیشتری نسبت به ساختارهای فیزیکی برخوردار خواهند بود، ما با واقعیت جهان همراه با کرونا و ناشناخته‌های عصر پس از کرونا مواجه هستیم واقعیت‌هایی که هر یک قادر خواهند بود تا مسیر مدیریت منابع انسانی در سازمانها را دچار تغییرات شگرف نماید.

- ومبتنی بر فناوری اطلاعات و ایجاد یکپارچگی و جامعیت بخشیدن به آنها سرعت انجام امور را در سازمان بالا خواهد برد و علاوه بر سرعت از دقت کافی هم برخوردار می‌گردد.
- ✓ در زمینه چابکی سازمانی می‌توان با ایجاد ساختار ماتریسی در سازمان زمینه‌های لازم برای ارائه خدمات لازم در هر سطحی با مهارت‌های مختلف کارکنان را ایجاد کرد. با ایجاد مهارت‌های چند سطحی می‌توان زمینه مناسب برای بهبود چابکی سازمانی را ایجاد کرد. در صورتیکه کارکنان بتوانند مهارت‌های مختلف را در خود ایجاد کنند مسلماً چابکی نیز در سازمان بهبود می‌یابد. همچنین به منظور بهبود اثربخشی سازمانی استفاده از تحلیل مستمر شاخص‌های عملکردی و سنجش این شاخص‌ها با ارائه آلامر‌هایی در سیستم به منظور نشان دادن نقاط بحرانی زمینه را برای کنترل اثربخشی سازمانی ایجاد کرد.

### ۹\_۱ پیشنهاد به محققین آینده

- با توجه به اینکه این پژوهش در تعداد محدودی سازمان انجام شده است، توصیه می‌شود در سایر سازمان‌های دولتی و خصوصی نیز مدل این پژوهش مورد بررسی و مقایسه قرار گیرد.
- اهمیت نقش امور فرهنگی در مدیریت منابع انسانی الکترونیک در مبانی نظری مورد بررسی قرار گرفت. توصیه می‌شود پژوهشگران علاقه مند این متغیر را به طور خاص در مبحث ارزیابی این سیستم مطالعه و مورد بررسی قرار دهند.
- برای بررسی عمیق تر و شناسایی دقیق تر ابعاد مختلف مدل این پژوهش می‌توان هر کدام از اجزای مدل را به طور جداگانه مورد مطالعه و سنجش قرار داد.

### مراجع

Brown, K. (۲۰۱۸). HRM development in post-colonial societies: The challenges of advancing HRM practices in Ghana. *International Journal of Cross Cultural Management*, ۱۸(۲), ۱۲۵-

۱. Graham, M., Ojanpera, S., Anwar, M., & Friederici, N. (۲۰۱۸). Digital connectivity and African knowledge economies.
۲. Ayentimi, D. T., Burgess, J., &

- Amuna, Y. M. A., & El Talla, S. A. (۲۰۱۷). The Efficiency of Information Technology and its Role of e-HRM in the Palestinian Universities.
۱۰. Galanaki, E., Lazazzara, A., & Parry, E. (۲۰۱۹). A cross-national analysis of e-HRM configurations: integrating the information technology and HRM perspectives. In *Organizing for Digital Innovation* (pp. ۲۶۱-۲۷۶). Springer, Cham.
۱۱. خورسندی، سمیه و فرشاد حاج علیان، ۱۳۹۸. شناسایی عوامل پذیرش مدیریت منابع انسانی الکترونیک با رویکرد معادلات ساختاری، کنفرانس بین‌المللی پیشرفت در مدیریت، کسب و کار تجارت اقتصاد و امور مالی، ترکیه، شرکت همایش آروین البرز
۱۲. وخشوری، اسماعیل؛ محمود حقانی و حسن شیبانی، ۱۳۹۷، *ارایه الگوی توسعه منابع انسانی بر مبنای مدیریت تکنولوژی، کنفرانس ملی نقش مدیریت در چشم انداز ۱۴۰۴، رشت، سازمان بسیج مهندسی صنعتی استان گیلان*
۱۳. حسینی، سیدعبدالرسول، قاسمی، محمد (۱۳۹۷)، *نقش فناوری اطلاعات و ارتباطات (مطالعه موردی: کارکنان سازمان تعاون روستایی استان فارس)*، تهران: نخستین کنفرانس ملی تحقیق و توسعه در مدیریت و اقتصاد مقاومتی.
۱۴. قربانی زاده؛ وجه الله، میرعلی سید تقوی؛ داوود حسین پور (۱۳۹۵)، *وضعیت سرمایه انسانی در شرکت‌های دانش‌بنیان بر مبنای سبک‌های نوآوری، انجمن علوم مدیریت ایران، دوره ۱۱*
- ۱۴۷.
۳. Landolt, P., Stephan, R., & Scherrer, S. (۲۰۱۹). Development of a new High Resolution Melting (HRM) assay for identification and differentiation of Mycobacterium tuberculosis complex samples. *Scientific reports*, ۹(۱), ۱۸۵۰.
۴. Boxall, P. (۲۰۱۸). The development of strategic HRM: reflections on a ۳۰-year journey. *Labour & Industry: a journal of the social and economic relations of work*, ۲۸(۱), ۲۱-۳۰.
۵. Malik, A., Pereira, V., & Tarba, S. (۲۰۱۹). The role of HRM practices in product development: Contextual ambidexterity in a US MNC's subsidiary in India. *The International Journal of Human Resource Management*, ۳۰(۴), ۵۳۶-۵۶۴.
۶. Burbach, R. (۲۰۱۹). ۱۴ Strategic evaluation of e-HRM. *e-HRM: Digital Approaches, Directions & Applications*.
۷. Islam, M., & Azad, M. (۲۰۱۸). Emerging Role of Electronic Human Resources Management Functions in Bangladesh-A Critical Review. *Emerging Role of Electronic Human Resources Management Functions in Bangladesh-A Critical Review* (December ۳۱, ۲۰۱۸).
۸. Berber, N., Đorđević, B., & Milanović, S. (۲۰۱۸). Electronic human resource management (e-HRM): A new concept for digital age. *Strategic Management*, ۲۲(۲), ۲۲-۳۲.
۹. Al Shobaki, M. M., Naser, S. S. A.,

- (Eds), Evidence-Based e-HRM? On the way to rigorous and relevant research, Proceedings of the Third European Academic Workshop on electronic Human Resource Management, Bamberg, Germany, May ۲۱-۲۰, CEUR-WS. org, ISSN ۰۱۰۸-۱۱۴۸, Vol. ۱۴۱, online: CEUR-WS.org/Vol-۱۴۱/, pp. ۲۳۰-۸۱۱).
۲۱. Huub J.M. Ruel & Tanya V. Bondarouk & Mandy Van der Velde, (۲۰۰۷), The contribution of e-HRM to HRM effectiveness: Results from a quantitative study in a Dutch Ministry, Employee Relations Vol. ۲۹ No. ۳, pp.
۲۲. Lepak, D.P. & Snell, S.A. (۱۹۹۸). Virtual HR: Strategic human resource management in the ۲۱st Century. Human Resource Management Review. ۸(۳), ۲۱۵-۲۳۴.
۲۳. Olivas-Lujan, M., Ramirez, J., & Zapata-Cantu, L. (۲۰۰۷). E-HRM in Mexico: Adapting innovativeness for global competitiveness. International Journal of Manpower, ۲۸(۵), ۴۱۸-۴۳۴.
۲۴. Walker, A.J. (Ed.). (۲۰۰۱). Web-based human resources: The technologies and trends that are transforming HR. New York: McGraw-Hill.
۲۵. Ruël, H., Bondarouk, T., & Looise, J. K. (۲۰۰۴). E-HRM: Innovation or Irritation? An Exploration of Web-based Human Resource Management in Large Companies. Utrecht: Lemma Publishers. ۱ Micro-Political Conflicts and Institutional Issues ۱۹
۱۵. عرفانیان، حمید، نجف بیگی، رضا، دانش فرد، کرم الله (۱۳۹۴)، مدیریت منابع انسانی الکترونیک ابزار تحول در بهره‌وری سازمان، نشریه مدیریت بهره‌وری.
۱۶. Turner, M. R., McIntosh, T., Reid, S. W., & Buckley, M. R. (۲۰۱۹). Corporate implementation of socially controversial CSR initiatives: Implications for human resource management. Human Resource Management Review, ۲۹(۱), ۱۲۵-۱۳۶.
۱۷. Abubakar Allumi Nura & Nor Hasni Osman. (۲۰۱۲), The Proposed Relationship connecting e-HRM adoption, Performance Management System and Effective Decision making in Higher Educational Institutions in Nigeria, European Journal of Business and Management www.iiste.org ISSN ۲۲۲۲-۱۹۰۵ (Paper) ISSN ۲۲۲۲-۲۸۳۹ (Online) Vol ۴, No. ۱۸.
۱۸. Chamaru De Alwis, A. (۲۰۱۰), The impact of Electronic Human Resource Management on the role of Human Resource Managers, E + M EKONOMIE A MANAGEMENT Journal – volume ۴- pages ۴۷-۶۱
۱۹. Liang-Hung Lin. (۲۰۱۱), Electronic human resource management and organizational innovation: the roles of information technology and virtual organizational structure, The International Journal of Human Resource Management, Vol. ۲۲, No. ۲, January ۲۰۱۱, ۲۳۵-۲۵۷
۲۰. Ralf Burbach & Tony Royle, Strohmeyer, S. Diederichsen, A

- strategies for effective human resource management (p. ۳۳۳). NJ: Jossey-Bass.
۳۲. Boudreau, J. W. (۲۰۰۳). Strategic knowledge measurement and management. In S. E. Jackson, A. DeNisi & M. A. Hitt (Eds.), *Managing knowledge for sustained competitive advantage* (pp. ۳۶۰-۳۹۶).
۳۳. Seyed Mehdi Mousavi Davoudi; Kiarash Fartash., (۲۰۱۱) *Electronic Human Resource Management : New Avenues which Leads to Organizational Success..A Journal of Multidisciplinary Research Vol.۱ Issue ۲*, pp. ۷۵-۸۷.
۳۴. درویش، حسن، کریم زادگان، داوود، میرزائزاد، رضا (۱۳۸۹)، ارزیابی آمادگی الکترونیکی سازمان جهت استقرار سیستم‌های مدیریت منابع انسانی الکترونیکی (e-HRM)، شیراز: اولین کنفرانس سالانه مدیریت، نوآوری و کارافرینی.
۳۵. هومن، حیدر علی (۱۳۹۰) مدل یابی معادلات ساختاری با کاربرد نرم افزار لیزرل. تهران: انتشارات سمت.
۲۶. Bondarouk, T. & Ruel, H.J.M. (۲۰۰۴). *Discourse Analysis: Making Complex Methodology Simple*. Paper presented at the European Conference on Information Systems, Turku, Finland.
۲۷. Bondarouk, T. & Ruel, H.J.M. (۲۰۰۶, June). Does e-HRM contribute to HRM Effectiveness? Results from a quantitative study in a Dutch Ministry. Paper presented at the ۴th International Conference of the Dutchs HRM Network: The Netherlands.
۲۸. Evans, E.M. (۲۰۰۱). Internet-age performance management: lessons from high-performing organizations. In A.J. Walker (Ed.), *Web-Based Human Resources: The technologies and trends that are transforming HR* (pp. ۶۵-۸۲). New York: McGraw-Hill.
۲۹. Voermans, M., & Van Veldhoven, M. (۲۰۰۷). Attitude towards E-HRM: An empirical study at Philips. *Personnel Review*, ۳۶(۵/۶), ۸۸۷-۹۰۲.
۳۰. Wongrassam, Gardiner P.D and Simmons J.E.L. (۲۰۰۳), "Performance Measurement Tools: The Balanced Scorecard And The EFQM Excellence Model", *Journal: Measuring Business Excellence*, ۷ (۱) : ۱۴-۲۸.
۳۱. Lepak, D. P., & Snell, S. A. (۲۰۰۳). Managing the human resource architecture for knowledge-based competition. In S. E. Jackson, A. DeNisi, & M. A. Hitt (Eds.), *Managing knowledge for sustained competitive advantage: designing*



## ارائه روشی جدید مبتنی بر مدل کوکومو بمنظور افزایش دقت تخمین تلاش در پروژه های نرم افزاری

مهديه سالاری\* و حید خطیبی بردسیری\*\* عمید خطیبی بردسیری\*\*  
\*گروه علمی مهندسی کامپیوتر، موسسه آموزش عالی غیرانتفاعی غیر دولتی جاوید  
\*\*استادیار گروه مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد بردسیر  
تاریخ دریافت: ۱۳۹۹/۰۱/۰۵ تاریخ پذیرش: ۱۳۹۹/۰۹/۲۱  
نوع مقاله: پژوهشی

### چکیده

تخمین و برآورد معیارها یک فعالیت حیاتی در پروژه‌های نرم‌افزاری محسوب می‌شود. به طوری که تخمین تلاش در مراحل اولیه توسعه نرم‌افزار، یکی از مهم‌ترین چالش‌های مدیریت پروژه‌های نرم‌افزاری است. تخمین نادرست می‌تواند منجر به شکست پروژه گردد. لذا یکی از فعالیت‌های اصلی و کلیدی در توسعه مؤثر و کارآمد پروژه‌های نرم‌افزاری تخمین دقیق هزینه‌های نرم‌افزار است. از این رو در این پژوهش دو روش به منظور تخمین تلاش در پروژه‌های نرم‌افزاری ارائه شده است، که در این روش ها سعی شده با تجزیه و تحلیل محرک‌ها و استفاده از الگوریتم‌های فرا ابتکاری و ترکیب با شبکه عصبی راهی برای افزایش دقت در تخمین تلاش پروژه های نرم افزاری ایجاد شود. روش اول تأثیر الگوریتم فاخته جهت بهینه‌سازی ضرایب تخمین مدل کوکومو و روش دوم به صورت ترکیبی از شبکه عصبی و الگوریتم بهینه‌سازی فاخته جهت افزایش دقت برآورد تلاش توسعه نرم‌افزار ارائه شده است. نتایج بدست آمده روی دو پایگاه داده واقعی نشان دهنده عملکرد مطلوب روش ارائه شده در مقایسه با سایر روشهاست.

**واژگان کلیدی:** الگوریتم فاخته، تخمین هزینه، شبکه عصبی، کوکومو.

### ۱- مقدمه

یکی از روش‌های تخمین هزینه‌های نرم‌افزار، پیشگویی میزان تلاش لازم برای ساخت نرم‌افزار و زمان موردنیاز جهت توسعه آن است؛ که به شکل ساده می‌توان تلاش لازم را برحسب تعداد افرادی که در روز/هفته/ماه و یا حتی سال بر روی پروژه کار می‌کنند را تخمین زد [۱]. مدل‌های تخمین زمان و هزینه نرم‌افزار را به دودسته الگوریتمی و غیر الگوریتمی تقسیم می‌کنند [۲-۴]. مدل‌های الگوریتمی مدلهایی هستند که بر مدل‌های ریاضی استوارند و از نظر محاسبات ریاضی پیچیده می‌باشند. این مدل‌ها تلاش می‌کنند تا

تخمین زمان و هزینه نرم‌افزاری از فعالیت‌های علمی موردنیاز در زمینه مدیریت پروژه‌های نرم‌افزاری محسوب می‌شود. دقت تخمین نرم‌افزار دارای اهمیت بالایی است، زیرا می‌تواند به طبقه‌بندی و اولویت‌بندی پروژه به منظور معین کردن منابعی که در آینده موردنیاز پروژه خواهند بود، کمک به سزایی کند.

نویسنده مسئول: مهديه سالاری mahdieh.salarie@gmail.com



مدل تخمین کوکومو نتایج دقیق تری در مقایسه با مدل کوکوموی اولیه با دو معیار عملکرد متوسط شدت خطای نسبی ۱/۲۶ و درصد پیش‌بینی ۱/۵۴ را ایجاد نموده است. همچنین این مدل در مقایسه با شبکه عصبی مصنوعی، شبکه عصبی فازی<sup>۳</sup> و فازی تجزیه و تحلیل رابطه‌ای گری<sup>۴</sup> نتایج بهتری را ارائه داده است [۶]. در مقاله [۷] نسخه جدیدی از مدل کوکومو با تکیه بر انتخاب محرک‌های هزینه و بهینه‌سازی ضرایب با استفاده از الگوریتم کلونی زنبور پرداخته است. در روش ارائه شده ضرایب کارآمدتری نسبت به حالت اولیه مدل کوکومو به دست آمده و انتخاب بهترین ضرایب بالاترین هدف این روش می‌باشد و نتایج شبیه‌سازی نمایانگر این برتری می‌باشد. در مقاله دیگری از الگوریتم رقابت استعماری برای تخمین هزینه نرم‌افزار استفاده شده است که روی پایگاه داده کوکومو ۸۱ صورت گرفت. نتایج تجربی نشان می‌دهند که در حالت کوکوموی مجزا، الگوریتم رقابت استعماری طوری پارامترهای مدل کوکومو ۸۱ را مقداردهی می‌کند که MMRE برابر با ۰,۲۸۶۳ و PRED(۰,۲۵) برابر با ۰,۳۵۷۱ می‌گردد. واضح است که در مقایسه با مدل کوکومو ۸۱ که MMRE برابر با ۰,۳۱۸۰ و PRED(۰,۲۵) برابر با ۰,۳۴۹۲ می‌باشد، الگوریتم رقابت استعماری توانسته است، با دقت بیشتری عملیات تخمین هزینه را انجام دهد [۸]. در یک مقاله، ACOT برای پیش‌بینی برآورد هزینه نرم‌افزار با استفاده از داده‌های تخمین تلاش مختلف، استفاده می‌کند. در طول مطالعه، اعتبار سنجی متقابل ده برابر را در مجموعه داده‌های ISBSG و اعتبار سنجی متقابل سه‌گانه در مجموعه داده‌های IBM DP و COCOMO ۸۱ انجام شده است؛ و عملکردهای ACOT را با SVR، GP، MLP، DENFIS، MARS و PSO را مقایسه کردند نتایج تجربی با توجه به ارزش RMSE<sup>۵</sup> بهتر از سایر تکنیک‌های مستقل بوده و از این رو، نتیجه می‌گیریم که پس از آزمایش‌های گسترده‌ای که مدل ACOT نسبتاً بهترین پیش‌بینی کننده در میان تمام تکنیک‌های دیگر است [۹]. در مقاله ی دیگر الگوریتم ژنتیک برای انتخاب ویژگی در زمینه تخمین تلاش‌های پروژه نرم‌افزار مورد استفاده قرار گرفته است این تکنیک بر روی مجموعه داده‌های شناخته شده ماکسول، کوکومو ۸۱ و مجموعه داده Desharnais آزمایش شده است. نتایج پیاده‌سازی نشان می‌دهد که زیرمجموعه حاصل،

ارتباطی بین تلاش و یک یا دو مشخصه پروژه برقرار کنند. شکل کلی این مدل‌ها به صورت رگرسیون خطی یا غیرخطی می‌باشد. در این‌گونه مدل‌ها میزان کار مورد نیاز برای انجام پروژه به صورت تابعی از متغیرهای متعدد تعریف می‌شود این متغیرها در واقع فاکتورها یا عوامل هزینه هستند که به صورت خطی<sup>۱</sup>، ضربی و نمایی در محاسبه زمان و هزینه تأثیر می‌گذارند. معادله تخمین استفاده شده به شکل زیر است.

$$\text{Effort} = f(x_1, x_2, \dots, x_n) \quad (1)$$

که در آن  $x_1$  تا  $x_n$  بردار فاکتورهای هزینه است. از سوی دیگر به دلیل عدم توانایی روش‌های الگوریتمی در مهار کردن رفتار پویای پروژه‌های نرم‌افزاری و نبود اطلاعات کامل از یک پروژه در مراحل اولیه، متدهای غیر الگوریتمی ارائه شدند. که این متدها بر اساس تجزیه و تحلیل پروژه‌های نرم‌افزاری بنا شده‌اند. در این مقاله با تمرکز بر روش کوکومو به عنوان متد پایه و الگوریتم بهینه‌سازی فاخته و شبکه عصبی به عنوان ابزار کمکی، یک مدل پویا برای تخمین مقدار تلاش ارائه خواهیم کرد. ادامه این مقاله در پنج قسمت سازمان‌دهی شده است: بخش دوم به مرور کارهای مرتبط می‌پردازد. در بخش سه به ترتیب مدل کوکومو، الگوریتم فاخته و شبکه عصبی و روش ارائه شده آورده شده و بخش چهارم نیز به ترتیب نحوه ارزیابی و نتایج به دست آمده را نشان می‌دهند. در نهایت بخش پنجم به نتیجه‌گیری و پیشنهاد کارهای آینده اختصاص دارد.

## ۲- پیشنهاد پژوهش

جهت تخمین تلاش نرم‌افزار، مدل‌های مختلفی معرفی شده است که کوکومو<sup>۲</sup> یکی از مشهورترین آن‌ها می‌باشد. کوکوموی ۱ یا کوکوموی ۸۱ یک مدل مبتنی بر رگرسیون می‌باشد که در سال ۱۹۸۱ توسط بوهم پیشنهاد شد [۵]. این مدل جزو دسته مدل‌های محسوب می‌شود که می‌تواند علاوه بر میزان تلاش، میزان زمان‌بندی یک پروژه نرم‌افزاری را نیز محاسبه کند. در مدل تخمین تلاش ترکیبی، بر پایه ترکیب تکنیک‌های هوش محاسباتی، مدل کوکومو و الگوریتم بهینه‌سازی ازدحام ذرات نتایج آزمایش نشان داد که ترکیب الگوریتم‌های خوشه‌بندی با الگوریتم بهینه‌سازی ازدحام ذرات در ارزیابی

<sup>۴</sup> Fuzzy Gray Relational Analysis (FGRA)

<sup>۵</sup> Root Mean Square Error

<sup>۱</sup> Linear Models

<sup>۲</sup> Cocomo

<sup>۳</sup> Fuzzy Neural Network (FNN)

بینی) بهبود یافته است. در مقاله [۱۵] یک مدل ترکیبی از الگوریتم جستجوی کوکو و الگوریتم جستجوی هارمونی برای بهینه‌سازی چهار ضریب مدل کوکومو<sup>۲</sup> برای برآورد مطلوب پیشنهاد شده است. رویکرد پیشنهادی در مجموعه داده‌های ناسا ۹۳ و با استفاده از مقادیر خطای نسبی [۱۶] و اندازه خطای نسبی (MMRE) مورد ارزیابی قرار گرفته است. نتایج تجربی نشان می‌دهد که روش ارائه شده در برآورد تلاش و زمان توسعه پروژه نرم افزاری نسبت به کوکومو<sup>۲</sup> و الگوریتم جستجوی کوکو موثرتر است. امروزه برآورد هزینه نرم افزار (SCE) با تکنیک‌های یادگیری ماشین عملکرد بیشتری نسبت به سایر تکنیک‌های سنتی دارند در این مقاله، یک مدل ترکیبی جدید از چندین مدل ارائه شده که مدل‌ها عبارت‌اند از چندلایه پرسپترون (MLP)، شبکه عصبی مصنوعی (ANN) و بهینه‌سازی کلونی مورچه<sup>۲</sup> (ACO) و برای بالا بردن دقت در برآورد هزینه نرم افزار به نام بهینه‌سازی کلونی مورچه چندلایه پرسپترون (MLPACO) نامیده می‌شود. در این مقاله با استفاده از الگوریتم کلونی مورچه (ACO) از بین ویژگی‌های پروژه‌های نرم افزاری ویژگی‌هایی که باعث افزایش دقت تخمین می‌شود فیلتر می‌شوند و با استفاده از شبکه عصبی چندلایه پرسپترون بهینه‌سازی تخمین انجام می‌شود. نتایج نشان می‌دهد که در این رویکرد جدید افزایش دقت برای بیش از ۸۰٪ موارد بهتر از الگوریتم سازنده هزینه کوکومو می‌باشد همچنین نتایج الگوریتم پیشنهاد شده نشان می‌دهد که میانگین مقدار خطای نسبی (MMRE) در الگوریتم پیشنهاد شده است پایین‌تر از مدل کوکومو می‌باشد [۱۷]. در پژوهشی دیگر رویکردی پیشنهاد شد که شامل شبکه عصبی مصنوعی عملکردی و الگوریتم کلونی زنبور عسل مصنوعی<sup>۳</sup> که به عنوان الگوریتم آموزش شبکه و برای ارائه دقیق‌ترین تخمین هزینه نرم افزار استفاده شده است. FLANN پیچیدگی محاسباتی را در شبکه عصبی چندلایه کاهش می‌دهد و لایه پنهان ندارد و بنابراین یادگیری سریع‌تری دارد؛ و پس از ارزیابی نتایج نشان داده شده که آموزش FLANN با الگوریتم ABC برای رفع مسئله هزینه نرم افزار بسیار بهبود یافته است و علاوه بر این مدل به دلیل نداشتن لایه پنهان و محاسبات کمتر در طول آموزش از لحاظ ساختاری بسیار ساده می‌باشد؛ و یکی از اهداف مهم این تحقیق این است که آموزش FLANN با یک

در مقایسه با مجموعه داده‌های اصلی، نتیجه‌های بهتر را از لحاظ دقت برآورد تلاش به ارمغان آورده است. این مقاله نشان داد که الگوریتم‌های ژنتیک روش‌های ایده آل برای انتخاب یک زیرمجموعه از ویژگی‌ها و بهبود دقت تخمینی تلاش هستند [۱۰]. یکی از مشهورترین روش‌های تخمین تلاش، استفاده از قیاس و مقایسه یک سرویس با موارد مشابه قبلی است. متأسفانه روش قیاس بدون استفاده از وزن‌های مناسب و ارزشدهی به ویژگی‌های یک سرویس، نتایج خوبی نخواهد داشت. بنابراین در این مقاله سعی شده تا با ترکیب روش قیاس و الگوریتم تکامل تفاضلی یک مدل کارا و قابل اطمینان برای برآورد تلاش لازم جهت توسعه سرویس‌های نرم افزاری ایجاد شود. روش ارائه شده بر روی داده‌های واقعی مستخرج از پایگاه داده ISBSG و دو پایگاه داده مصنوعی مورد ارزیابی قرار گرفت و نتایج با روش‌های مشهور تخمین تلاش مقایسه گردید؛ مقادیر به دست آمده برای مخازن داده‌های ISBSG<sup>۱</sup>، همگن و ناهمگن به ترتیب و به طور میانگین بهبود ترکیبی برای تخمینی تلاش‌های پروژه‌های نرم افزاری پیشنهاد شده است. روش ارائه شده ترکیبی از الگوریتم بهینه سازی ذرات و یک روش رگرسیون خطی است که در آن نتایج ضرایب بهینه انجام می‌شود. علاوه بر این، معادله برآورد با استفاده از اندازه پروژه اندازه گیری می‌شود به طوری که دقیق ترین برآورد به دست آمده است. مجموعه داده ای نسبتاً واقعی برای ارزیابی عملکرد مدل پیشنهاد شده و نتایج با مدل های دیگر مقایسه می‌شود. نتایج به دست آمده نشان داد که مدل هیبرید پیشنهادی می‌تواند صحت برآورد ها را بهبود بخشد. در مقاله [۱۳] یک روش جدید بر اساس ترکیبی از مدل های برآورد مختلف و ریاضی پاسخ خود را ارائه شده که این روش جدید در دو داده مجموعه کاملاً متفاوت واقعی مورد بررسی قرار گرفت. در مقاله [۱۴] یک مدل جدید برای برآورد هزینه های نرم افزاری ارائه می‌دهد که با استفاده از الگوریتم جستجوی کوکو (CS) برای پیدا کردن پارامتر مطلوب مدل تخمینی هزینه است. روش ارائه شده در مجموعه داده های پروژه نرم افزار ناسا مورد آزمایش قرار گرفته است. نتایج تجربی نشان می‌دهد که روش ارائه شده عملکرد نتایج برآورد شده با توجه به MMRE (میانگین خطای نسبی) و PRED (پیش

<sup>۲</sup> Bee Colony optimization<sup>۱</sup> International Software Benchmarking Standards Group<sup>۳</sup> Ant Colony Optimization

وزن المان ها توسط الگوریتم رقابته امپریالیست بهینه می شوند. برای مقایسه روش ارائه شده برای برآورد تلاش با مدل های پیشرفته ، سه مجموعه داده به عنوان معیار ، یعنی COCOMO، Maxwell و Albrecht انتخاب شده اند. مجموعه داده ها استاندارد و در دسترس عموم برای ارزیابی هستند. آزمایش ها نتایج امیدوار کننده را نشان می دهد و عملکرد متوسط با الگوی پیشنهادی معیار عملکرد MMRE در مجموعه داده ها به ترتیب ۲۳٪ ، ۳۸٪ و ۳۵٪ بهبود یافته است.

### ۳- روش شناسی پژوهش

#### ۳-۱- کوکومو

واژه کوکومو به طور اختصار از دو حرف اول کلمات جمله‌ی *Constructive Cost Modle* که به معنای مدل هزینه سودمند است گرفته شده است. این مدل ابتدا در سال ۱۹۷۰ توسط بری بوهم ارائه و بعداً در سال ۱۹۸۱ در کتاب ایشان به نام اقتصاد مهندسی نرم افزار چاپ و در اختیار عموم قرار گرفت [۲۲، ۲۳] این مدل در سه سطح پایه، متوسط، با جزئیات ارائه شده است که در ادامه به تحلیل هر یک از این ۳ سطح می پردازیم.

#### • کوکومو پایه<sup>۱</sup>

کوکومو پایه میزان تلاش<sup>۲</sup> (و هزینه) مورد نیاز جهت توسعه نرم افزار را به صورت تابعی از اندازه‌ی برنامه محاسبه می کند. اندازه برنامه برآوردی از تعداد خطوط برنامه بر پایه ۱۰۰۰ است که ان را به اختصار *KLOC* و یا *SLOC* می نامند.

معادله کوکومو پایه به طور کلی:

تلاش کاربردی (بر اساس فرد در ماه)

$$Effort Applied (E) = a (KLOC)b \quad (2)$$

#### • کوکومو متوسط

کوکومو متوسط میزان تلاش (و هزینه) مورد نیاز جهت توسعه نرم افزار را به صورت تابعی از اندازه‌ی برنامه و مجموعه‌ای از

برنامه آموزش بهبود یافته جایگزین الگوریتم یادگیری استاندارد BP برای ارزیابی هزینه دقیق توسعه نرم افزار شده است و مزیت عمده روش ارائه شده این است که پیچیدگی محاسباتی را بدون هیچ گونه نقص در عملکرد آن برطرف می کند. نتایج به دست آمده از روش پیشنهادی با استفاده از معیارهایی مانند MRE، MMRE و MdMRE ارزیابی شده است با توجه به نتایج به دست آمده می توان گفت که FLANN دقت تخمین هزینه را افزایش می دهد و به ویژه هنگامی که این الگوریتم با الگوریتم ABC آموزش داده می شود دقت هزینه برآورد شده بسیار بهتر می شود [۱۸]. در مقاله [۱۹]، اجرای مدل سازی غیر الگوریتمی با استفاده از تکنیک های محاسبات نرم مانند منطق فازی و الگوریتم های ژنتیکی انجام می شود. روش فازی برای طراحی یک مدل فازی برای هر درایور هزینه پیاده سازی شده است. مدل فازی دارای تعریف نادرست و مبهم از محدوده ورودی درایورهای هزینه است. انتخاب پارامترهای مشخص کننده مجموعه های فازی در مدل فازی پیشنهادی با استفاده از الگوریتم های ژنتیکی بیشتر بهینه می شود. مدل ارائه شده بر روی مجموعه داده های COCOMO NASA و مجموعه داده COCOMO NASA۲ با استفاده از MATLAB آزمایش شده است. بهبود عملکرد مدل بهینه سازی شده پیشنهادی از نظر میانگین بزرگی خطای نسبی (MMRE) و Pred (۲۵٪) اندازه گیری می شود. در مقاله [۲۰] شبکه عصبی مصنوعی با استفاده از الگوریتم بهینه سازی فاخته برای پیش بینی برآورد هزینه نرم افزار آموزش داده شده است. هدف اصلی در این مقاله استفاده از یک روش یادگیری جدید در شبکه عصبی برای پیش بینی بهتر تخمین پروژه های نرم افزاری می باشد. روش ارائه شده با مجموعه داده ISBSG تأیید شده و نتایج با مدل های موجود مقایسه شده است. نتایج نشان داده شده از نظر خطای میانگین مربع (RMSE) و اندازه متوسط خطای نسبی (MMRE) است. در مقاله [۲۱]، یک رویکرد ترکیبی سه فاز برای غلبه بر مشکل تخمین هزینه ارائه شده است. در مرحله اول ، ویژگی ها با استفاده از ترکیبی از الگوریتم ژنتیکی و شبکه عصبی پرسپترون انتخاب می شوند. در مرحله دوم ، عوامل ضربه با استفاده از روش های رگرسیون خطی چندگانه که به عنوان ضرایب تأثیر برای هر ویژگی عمل می کنند ، به هر ویژگی انتخابی مرتبط می شوند. در مرحله آخر و سوم ،

<sup>۲</sup> Effort

<sup>۱</sup> Basic COCOMO

تخم‌ها نجات یابند پارامتری خواهد بود که  $COA$  قصد بهینه‌سازی آن را دارد [۲۵].

فاخته‌ها برای بیشینه کردن نجات تخم‌های خود به دنبال بهترین منطقه می‌گردند. پس از آنکه جوجه‌ها از تخم درآمدند و به فاخته بالغ تبدیل شدند، جوامع و گروه‌هایی تشکیل می‌دهند. هر گروه منطقه سکونت خود را برای زیست دارد. بهترین منطقه سکونت تمام گروه‌ها مقصد بعدی فاخته‌ها در سایر گروه‌ها خواهد بود. تمام گروه‌ها به سمت بهترین منطقه موجود فعلی مهاجرت می‌کنند. هر گروه در منطقه‌ای نزدیک بهترین موقعیت فعلی ساکن می‌شود. با در نظر گرفتن تعداد تخمی که هر فاخته خواهد گذاشت و همچنین فاصله فاخته‌ها از منطقه بهینه فعلی برای سکونت تعدادی شعاع تخم‌گذاری محاسبه شده و شکل می‌گیرد. سپس فاخته‌ها شروع به تخم‌گذاری تصادفی در لانه‌هایی داخلی شعاع تخم‌گذاری خود می‌کنند. این پروسه تا رسیدن به بهترین محلی برای تخم‌گذاری (منطقه با بیشترین سود) ادامه می‌یابد. این محلی بهینه جایی است که بیشترین تعداد فاخته‌ها در آن گرد می‌آیند.

حل یک مسئله بهینه‌سازی لازم است تا مقادیر متغیرهای مسئله به فرم یک آرایه شکل گیرند. در  $GA$  و  $PSO$  این آرایه‌ها بانام‌های «کروموزوم» و «موقعیت ذرات» مشخص می‌شوند. ولی در الگوریتم بهینه‌سازی فاخته این آرایه  $habitat$  یا «محل سکونت» نام دارند.

در یک مسئله بهینه‌سازی  $Nvar$  بعدی یک  $habitat$  یک آرایه  $Nvar \times 1$  خواهد بود که موقعیت فعلی زندگی فاخته‌ها را نشان می‌دهد. این آرایه به شکل زیر تعریف می‌شود:

$$Habitat = [x_1, x_2, \dots, xvar] \quad (4)$$

میزان مناسب بودن (مقدار سود) در  $habitat$  فعلی با ارزیابی تابع سود ( $fp$ ) در  $habitat$  به دست می‌آید؛ بنابراین:

$$Profit = fp(habitat) = f(x_1, x_2, \dots, xvar) \quad (5)$$

همان‌طور که دیده می‌شود  $COA$  الگوریتمی است که تابع سود را ماکزیمم می‌کند. برای استفاده از  $COA$  برای حل مسائل کمینه‌سازی کافی است یک علامت منفی در تابع هزینه ضرب

محرک‌ها که شامل ارزیابی محصول، سخت‌افزار، کارکنان و ویژگی‌های پروژه هستند را محاسبه می‌کند. این محرک‌ها هر یک دارای چند سطح با مقادیر مخصوص به خود هستند.

هر یک از این ۱۵ ویژگی، از خیلی کم تا بسیار زیاد در ۶ سطح امتیازدهی می‌شوند (بر اساس ارزش و یا درجه اهمیت). با استفاده از افزایشده‌ها می‌توان میزان مقیاس‌ها و اندازه‌ها را در جدول زیر تغییر داد. محصول همه‌ی این افزایشده‌ها منجر به فاکتور تعدیل و تنظیم تلاش می‌شود ( $EAF$ ). ارزش‌های به‌دست‌آمده معمولاً مقداری بین ۰٫۹ تا ۱٫۴ می‌باشند.

فرمول محاسبه تلاش کوکومو متوسط

$$E = ai(KloC)(bi)(EAF) \quad (3)$$

برای محاسبه زمان توسعه و افراد موردنیاز می‌توان از همان فرمول‌های کوکومو پایه استفاده کرد.

#### • کوکومو با جزئیات

این مدل از کوکومو دارای تمامی ویژگی‌های مدل کوکومو متوسط به‌اضافه یک ارزیابی از تأثیر محرک‌های هزینه در هر مرحله از فرایند مهندسی نرم‌افزار است.

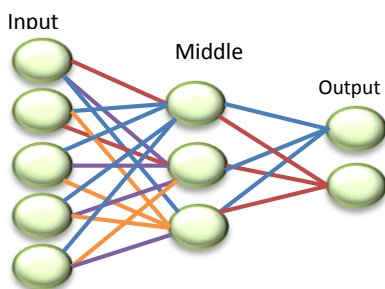
کوکومو با جزئیات میزان تلاش [۲۴] موردنیاز جهت توسعه نرم‌افزار را به‌صورت تابعی از اندازه‌ی برنامه و مجموعه‌ای از محرک‌ها بسته به چرخه حیات نرم‌افزار به دست می‌آورد.

#### ۳-۲- الگوریتم فاخته

الگوریتم بهینه‌سازی فاخته<sup>۱</sup> ( $COA$ ) همانند سایر الگوریتم‌های تکاملی هم با یک جمعیت اولیه کار خود را شروع می‌کند. جمعیتی متشکل از فاخته‌ها. این جمعیت از فاخته‌ها تعدادی تخم‌دارند که آن‌ها را در لانه تعدادی پرندگی میزبان خواهند گذاشت. تعدادی از این تخم‌ها که شباهت بیشتری به تخم‌های پرندگی میزبان دارند شانس بیشتری برای رشد و تبدیل شدن به فاخته بالغ خواهند داشت. سایر تخم‌ها توسط پرندگی میزبان شناسایی شده و از بین می‌روند. میزان تخم‌های رشد کرده مناسب بودن لانه‌های آن منطقه را نشان می‌دهند. هرچه تخم‌های بیشتری در یک ناحیه قادر به زیست باشند و نجات یابند به همان اندازه سود (تمایل) بیشتری به آن منطقه اختصاص می‌یابد؛ بنابراین موقعیتی که در آن بیشترین تعداد

<sup>۱</sup> Cuckoo optimization algorithm

است. بخش اول تابع ترکیب نام دارد که تمام ورودی ها را جمع می کند و نتیجه را به صورت یک مجموع وزن دار نشان می دهد. طبق این شکل برای هر نرون  $i$  تمام ورودی های  $I_j$  توسط فاکتور  $W_j$  وزن دار می شوند. سپس هر ورودی در وزن مربوطه ضرب شده و حاصل ضرب ها با هم جمع شده تا مجموع وزن دار  $u_i$  را نتیجه دهند:



شکل ۱. سلول ولایه های یک شبکه عصبی

شکل (۱). مدل ساده یک شبکه عصبی مصنوعی که از یک لایه ورودی، یک لایه میانی و یک لایه پنهان تشکیل شده است. شکل (۱) مدل ریاضی از فرایند پردازش در یک نرون.

$$u_i = \sum_{j=1}^i w_j I_j \quad (8)$$

بخش دوم تابع انتقال یا تابع تحریک نام دارد. وقتی مجموع وزندار به حد آستانه های رسید، تابع انتقال، تحریک شده تا خروجی را نتیجه دهد. توابع انتقال انواع مختلفی دارند که برخی از آنها عبارتند از: توابع انتقال *Tansig*, *Pureline* و *Logsig*.

معمولاً در یک سلول شبکه عصبی یک ورودی اضافی نیز وجود دارد که بایاس نامیده می شود. نقش بایاس افزایش یا کاهش مجموع وزن دار است [۲۶] بایاس به عنوان یک جبران کننده عمل می کند و به شبکه کمک می کند تا الگوهای موجود را بهتر بشناسد. یادگیری در یک شبکه عصبی مصنوعی توسط تعدیل (کموزیاد کردن) وزن ها صورت می گیرد [۲۷] الگوریتمهای آموزشی به دودسته نظارتی و غیر نظارتی تقسیم بندی میشوند. الگوریتم پس انتشار خطا نوعی از یادگیری نظارتی است که طی آن ورودیها وارد شبکه میشوند. وقتی که خروجی مطلوب به دست آید، خطای

کنیم. دیگر عادت هر فاخته حقیقی این است که آن ها در یک دامنه مشخص تخم های خود را می گذارند که با آن حداکثر دامنه تخم گذاری  $ELR$  <sup>۱</sup> گفته می شود. در یک مسئله بهینه سازی هر متغیر دارای حد بالا  $Var_{hi}$  و حد پایین  $Var_{low}$  است که هر  $ELR$  با استفاده از این حدود قابل تعریف خواهد بود.  $ELR$  متناسب است با تعداد کل تخم ها، تعداد تخم های فعلی فاخته و همچنین حد بالا و پایین متغیرهای مسئله؛ بنابراین  $ELR$  به صورت رابطه ۶ تعریف می شود:

$$(6)$$

$$ELR = \alpha \times \frac{\text{number of current cuckoo's eggs}}{\text{total number of eggs}} \times (Var_{hi} - Var_{low})$$

با توجه به تعداد تخم هر فاخته یک  $ELR$  برای آن مشخص می شود و سپس تخم گذاری شروع می گردد. فرمول عملگر مهاجرت در الگوریتم بهینه سازی فاخته به صورت رابطه است [۲۵]:

$$(7)$$

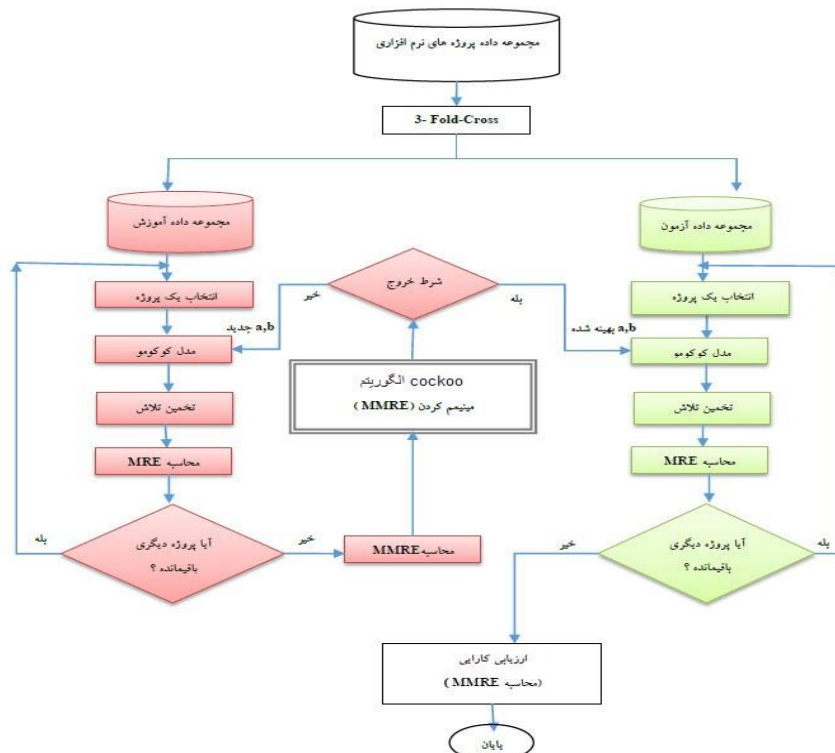
$$X_{NextHabitat} = X_{currentHabitat} + F(X_{GoalPoint} - X_{currentHabitat})$$

$F$  پارامتری است که باعث انحراف می شود.

### ۳-۳- شبکه عصبی

شبکه های عصبی یکی از کارآمدترین و گسترده ترین سیستم های هوشمند به شمار می آیند. شبکه های عصبی مصنوعی در واقع تقلید بسیار ساده ای از رفتار سلول های بیولوژیکی می باشد. اجزای تشکیل دهنده شبکه های عصبی مصنوعی شامل نرون ها و لایه ها می باشند. در حالت کلی یک شبکه عصبی از ۳ لایه تشکیل شده است شکل (۱): لایه ورودی که اطلاعات را دریافت می کند، لایه میانی یا پنهان که خود می تواند شامل چندین لایه باشد، یکی از مهم ترین پردازشگرهای اطلاعات است و لایه خروجی که این پردازش را دریافت کرده و خروجی نهایی را نتیجه می دهد. شکل (۱) مدل ساده ای از یک سلول شبکه عصبی را که نرون نامیده میشود نشان می دهد. بدنه این سلول از دو بخش تشکیل شده

<sup>۱</sup> Egg Laying Radius



شکل ۲. روش COA تخمین هزینه نرم افزار

بهترین مقادیر را برای ضرایب  $a$  و  $b$  یافته و در مدل کوکومو (پایه) جایگزین می‌نماید. در روش دوم (ANN-COA) الگوریتم فاخته از بین ۱۵ محرک هزینه مدل کوکومو محرک‌های مؤثرتر را پیشنهاد داده و به‌عنوان ورودی به همراه مقدار بهینه  $a, b$  به شبکه عصبی ارسال می‌کند و شبکه عصبی به‌عنوان خروجی تخمین تلاش نرم‌افزار را با توجه به محرک‌های مؤثر به دست می‌آورد.

۳-۴-۱ روش COA پیدا کردن ضرایب  $a$  و  $b$  بهینه برای مدل کوکومو با استفاده از الگوریتم‌های فرا ابتکاری

در این قسمت جزئیات مربوط به روش ارائه شده COA که در شکل (۲) نشان داده شده، بیان شده‌اند. این روش سعی بر آن دارد که تا حد امکان از طریق بهینه‌سازی ضرایب کوکومو بهترین تخمین‌ها را به دست آورد. فرایند بخش‌های روش ارائه شده به‌گونه‌ای است که میزان خطا در تخمین‌ها حداقل شود. این روش شامل ۲ بخش اصلی است که در ادامه به

محاسبه‌شده به‌صورت پس‌رو در شبکه منتشر میشود تا وزنها با توجه به آن تعدیل شوند. وقتی که خروجی محاسبه‌شده با خروجی واقعی تطابق خوبی داشته باشد، آموزش متوقف میشود. استفاده از شبکه عصبی در مهندسی هزینه و تخمین هزینه از سال ۱۹۹۳ شروع شده و نتایج مطلوبی را به همراه داشته است شبکه عصبی مصنوعی، معمول ترین روش مبتنی بر یادگیری مورد استفاده برای تخمین تلاش توسعه نرم افزار است؛ بنابراین روش ارائه شده بر اساس شبکه عصبی معرفی شده است.

۳-۴-۲ روش های ارائه شده

در توسعه نرم‌افزار ریسک‌های بالایی از خطا در تخمین هزینه پروژه‌ها با استفاده از مدل الگوریتمی کوکومو وجود دارد. در مدل های الگوریتمی مقادیر ثابت تخمین هزینه مقادیر تعریف‌شده‌ای نیستند و بنابراین پیدا کردن پاسخ‌های قابل‌اطمینان آسان نیست. روش ارائه شده در این پژوهش ترکیبی از الگوریتم فاخته و شبکه عصبی و مدل کوکومو (پایه) را جهت بهبود تخمین هزینه به کار گرفته است. دو روش ارائه شده عبارت‌اند از: در روش اول (COA)، الگوریتم فاخته

توضیح و تشریح هر یک از آن‌ها می‌پردازیم. بخش اول شامل بخش آموزش و بخش دوم شامل بخش آزمون می‌باشد.

### ۳-۴-۱-۱ آموزش در روش COA

روش ارائه شده در بخش آموزش ساخته شده و در واقع پارامترهای روش ارائه شده در این بخش تنظیم می‌شود. در این بخش سعی شده تا با تولید پارامترهای  $a$  و  $b$  جدید ضرایبی جدید جهت بهینه کردن مدل کوکومو ارائه دهیم. در ابتدا کل پروژه‌ها به دودسته آموزش و آزمون تقسیم می‌شوند. این تقسیم‌بندی کاملاً تصادفی بوده و این عمل با نسبت ۸۰ به ۲۰ انجام خواهد گرفت؛ که ۸۰ درصد از پروژه‌ها برای بخش آموزش و بقیه (۲۰٪) برای بخش آزمون انتخاب می‌شوند. در ادامه داده‌های آموزشی باید طبق رابطه کوکومو به ۳ دسته تقسیم شوند. این ۳ دسته شامل داده‌های ارگانیک، نیمه منفصل و تعبیه شده می‌باشند. این دسته‌بندی در مدل‌ها به دلیل درجه بالای ناهمگنی در پروژه‌های کوکومو است. حال با استفاده از الگوریتم‌های فرا ابتکاری سعی می‌کنیم تا ضرایب  $a$  و  $b$  را بهینه کنیم. برای بهینه کردن این ضرایب در هر یک از الگوریتم‌های فرا ابتکاری نیاز به یک تابع شایستگی داریم. تابعی که در این پژوهش مورد استفاده قرار گرفته شده در رابطه ۹ بیان شده است. منظور از تابع شایستگی (فیتنس) مقداری است که سعی داریم آن را حداقل کنیم.

$$\text{effort} = |\text{RealCost} - (a_i \times (K \text{loc}) b_i \times \text{EAF})| \quad (9)$$

منظور از RealCost در تابع بالا میزان هزینه واقعی پروژه است. در این پژوهش سعی شده با استفاده از تابع بالا میزان فاصله بین هزینه تولید شده توسط مقدار پیشنهادی الگوریتم فرا ابتکاری فاخته را با هزینه واقعی نرم‌افزار به حداقل برسانیم. هنگام اجرای این الگوریتم محدوده ضرایب  $a$  و  $b$  بین ۰ تا ۵ قرار گرفته‌اند؛ زیرا تعداد نقاط بهینه جواب در بازه‌های بالاتر از ۵ و کمتر از ۰ زیاد است (باید در نظر داشت به دلیل ساختار فرمول کوکومو در صورت مشخص نمودن یک بازه امکان وجود یک جواب بهینه برای تمامی اعداد حقیقی وجود خواهد داشت زیرا همواره درازای هر عدد در ضریب  $a$  یک عدد در اعداد حقیقی وجود خواهد داشت که با استفاده از آن بتوان فاصله بین جواب به دست آمده توسط مدل را با هزینه واقعی به حداقل رساند و یا حتی صفر کرد) بعد از اجرای الگوریتم فرا

ابتکاری فاخته بر روی هر یک از پروژه‌ها درازای هر پروژه یک دسته از ضرایب  $a$  و  $b$  خواهیم داشت. حال باید از هر دسته آن ضرایبی که دارای فاصله کمتری تا هزینه واقعی نسبت به بقیه هستند استخراج و در آخرین ضرایبی که در یک دسته (ارگانیک - نیمه منفصل - تعبیه شده) مشترک هستند به طور مجزا میانگین بگیریم. وقتی که تخمین برای پروژه‌های یک دسته به پایان رسید باید پارامترهای عملکردی برای آن دسته محاسبه شوند. پارامترهای عملکردی مورد نظری که در اینجا استفاده شده‌اند MMRE می‌باشند. برای محاسبه MMRE ابتدا لازم است تا MRE طبق رابطه (۱۲) محاسبه شود. پس از محاسبه MRE برای همه پروژه‌ها مقدار MMRE محاسبه می‌شود. مقدار MMRE برابر است با میانگین مقادیر به دست آمده برای MRE در دسته مورد نظر که طبق رابطه (۱۳) محاسبه می‌شود. در ادامه MMRE به دست آمده درازای اجرای الگوریتم فرا ابتکاری استفاده شده در این پژوهش ارائه شده و در انتها بهترین آن‌ها انتخاب و به همراه ضرایب جدید ارائه خواهد شد.

### ۳-۴-۱-۲ آزمون در روش COA

در این بخش از نتایج حاصل از مرحله آزمون به منظور ارزیابی روش ارائه شده استفاده خواهد شد. داده‌های مورد استفاده در این بخش داده‌های آزمون هستند و همچنین از ضرایب بهینه استفاده خواهد شد. در ابتدا تمامی پروژه‌های آزمون به سه دسته (ارگانیک - نیمه منفصل و تعبیه شده) تقسیم می‌شوند سپس ضرایب مربوطه را به هر دسته استخراج و معادله کوکومو برای تک تک این پروژه‌ها اعمال می‌شود. به طور مثال اگر پروژه‌ای ارگانیک باشد ضرایب ارگانیک به آن اختصاص می‌یابد. سپس میزان تلاش پروژه انتخاب شده طبق معادله کوکومو محاسبه می‌شود (این کار برای تمامی ضرایب ایجاد شده توسط الگوریتم‌های فرا ابتکاری و روش ارائه شده در این مرحله انجام می‌شود). این فرایند برای تمامی پروژه‌های موجود در مرحله آزمون انجام می‌شود. سپس MRE مربوط به هر پروژه محاسبه می‌شود. به عبارت دیگر به تعداد تمامی پروژه‌ها MRE وجود خواهد داشت که میانگین

آزمون)؛ که در بخش آموزش با استفاده از الگوریتم تکاملی فاخته از بین ۱۵ فاکتور تخمین تلاش مدل کوکومو محرک‌های مؤثرتر را بر اساس تابع برازش الگوریتم فاخته که میزان دقت تخمین (MRRE) می‌باشد انتخاب می‌شوند؛ و محرک‌های مؤثر به‌عنوان ورودی شبکه عصبی پیشنهاد داده می‌شود و خروجی شبکه عصبی مقدار تلاش توسعه می‌باشد؛ و پس از محاسبه MRE های همه پروژه ها، معیار کارایی MMRE محاسبه و به الگوریتم فاخته فرستاده می‌شود. بر پایه مقدار MMRE الگوریتم، ضرایب a و b و فاکتورهای انتخاب‌شده تخمین تلاش را به‌گونه‌ای تنظیم می‌کند که MMRE حداقل گردد و تخمین مجدداً بر اساس ضرایب جدید آغاز می‌شود. این فرآیند تا زمانی که معیار توقف الگوریتم فاخته رضایت‌بخش باشد ادامه می‌یابد؛ و در بخش آزمون مراحل بالا با یک تفاوت که دیگر به الگوریتم فاخته

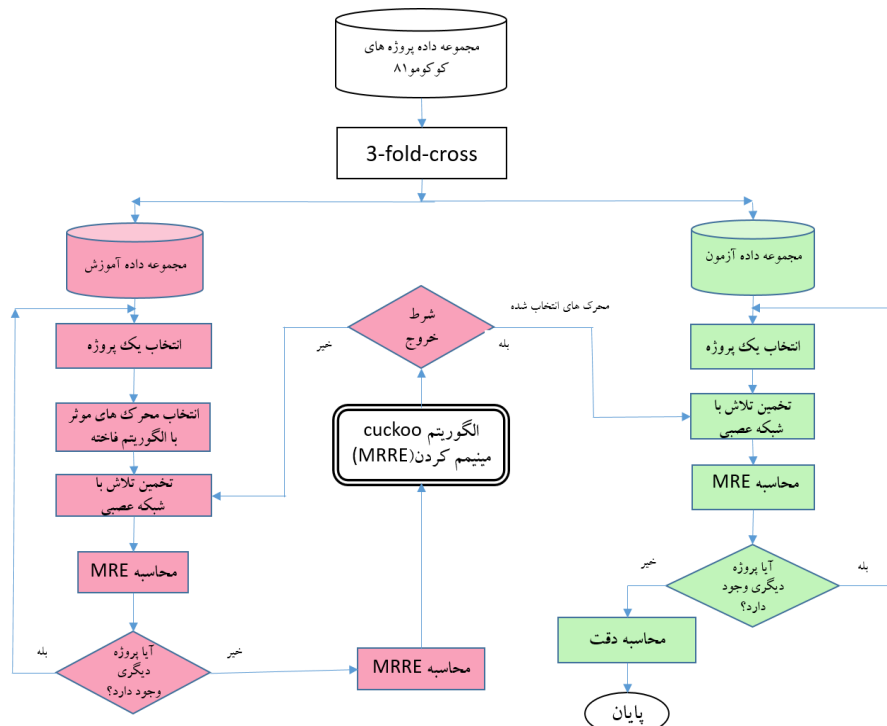
آن‌ها MMRE را تولید می‌کند. در انتها نتایج با مدل کوکومو اصلی مقایسه می‌شوند.

### ۳-۴-۲ روش ANN-COA انتخاب کردن محرک‌های مؤثرتر مدل کوکومو با استفاده از الگوریتم‌های فرا ابتکاری فاخته و شبکه عصبی

در این قسمت جزئیات مربوط به روش ارائه شده ANN-COA که در شکل (۳) نشان داده شده، بیان شده‌اند. این روش سعی بر آن دارد که تا حد امکان از طریق الگوریتم بهینه‌سازی فاخته محرک‌های مؤثر مدل کوکومو را انتخاب کنیم تا بهترین تخمین‌ها به کمک شبکه عصبی را به دست آورد. فرایند بخش‌های روش ارائه شده به‌گونه‌ای است که میزان خطا در تخمین‌ها حداقل شود. این روش شامل ۲ بخش اصلی است که در ادامه به توضیح و تشریح آن‌ها می‌پردازیم. بخش اول شامل بخش آموزش و بخش دوم شامل بخش آزمون می‌باشد.

### ۳-۴-۲-۱ آموزش و آزمون در روش ANN-COA

در این روش ابتدا داده‌ها را به دو بخش آموزش و آزمایش تقسیم می‌شوند (به‌صورت تصادفی ۸۰٪ آموزش و ۲۰٪



شکل ۳. روش ANN-COA تخمین هزینه نرم افزار



هدف اصلی همه تکنیک‌های تخمین استفاده شده در زمینه تلاش توسعه نرم‌افزار است.

برای انتخاب محرک مؤثرتر نیازی نیست چون در مرحله آموزش محرک‌ها شناسایی شده‌اند، تکرار می‌شوند.

#### ۴- بحث و نتایج

#### ۳-۵ پارامترهای عملکردی در روش ارائه شده

##### ۴-۱- نتایج در روش COA

دقت تخمین تلاش یک اصل مهم در موفقیت پروژه‌های نرم‌افزاری است رایج‌ترین روش‌ها که در روش ارائه شده مورد استفاده قرار گرفته است عبارت‌اند از [۲۸]:

روش ارائه شده COA بر روی دو مدل کوکوموی پایه و کوکوموی میانه به کار گرفته شد. در هر دو مدل جهت ارزیابی نتایج از دو مجموعه داده کوکوموی ۸۱ و ناسا ۹۳ استفاده شده است.

$$MRE = \frac{|Estimated_i - Actual_i|}{Actual_i} \quad (10)$$

$$MMRE = \frac{\sum_{i=1}^N MRE_i}{N} \quad (11)$$

##### ۴-۱-۱- نتایج با در نظر گرفتن مدل کوکوموی پایه

جدول (۱) نتایج روش ارائه شده بر روی مجموعه داده کوکوموی ۸۱ را نشان می‌دهد این جدول ضرایب a، b و همچنین معیار عملکرد MMRE برای سه دسته مختلف را نشان می‌دهد. همان‌طور که مشاهده می‌شود بهترین MMRE برای دسته ۲ حاصل شده است و میانگین آن‌ها ۰/۱۳۴۳ می‌باشد.

که  $Actual_i$ : تلاش واقعی پروژه i ام،  $Estimate_i$ : تلاش تخمین زده شده که با دو معادله کوکومو (پایه، میانه) زیر، به صورت جداگانه مقاداردهی می‌شود.

$$Estimated_i = a \times size^b \quad (12)$$

جدول ۱. نتایج مدل کوکوموی پایه، با مجموعه داده کوکوموی ۸۱

	A	B	MMRE	PRED
ارگانیک	۵	۰/۴۳۷۵	۰/۱۸۱۸	۰/۶۶۶۶
تعبیه شده	۰/۰۲۶۴	۱/۴۹۵۹	۰/۰۹۶۳	۰/۹۲۸۵
نیمه منفصل	۰/۴۹۰۶	۰/۹۵۰۸	۰/۱۲۴۸	۰/۹۰۹۰
میانگین			۰/۱۳۴۳	۰/۸۳۴۷

$$Estimated_i = a \times size^b \times \prod_{i=1}^{15} EAF_i \quad (13)$$

MRE: میزان خطای نسبی، MMRE متوسط شدت خطای نسبی، N تعداد پروژه‌ها، a و b: ضرایب ثابت می‌باشند.

EAF: ۱۵ فاکتور تنظیم تلاش که در هم ضرب می‌شوند.

جدول (۲) نتایج روش ارائه شده بر روی مجموعه داده ناسا ۹۳ را نشان می‌دهد این جدول ضرایب a، b و همچنین معیار عملکرد MMRE برای سه دسته مختلف را نشان می‌دهد. همان‌طور که مشاهده می‌شود بهترین MMRE برای دسته ۳ حاصل شده است و میانگین آن‌ها ۰/۳۴۷۲ می‌باشد.

پارامتر دیگری که برای ارزیابی عملکرد، استفاده می‌شود درصد پیش‌بینی، مطابق فرمول زیر می‌باشد.

$$PRED(x) = A/N \quad (14)$$

جدول ۲. نتایج مدل کوکوموی پایه، با مجموعه داده ناسا

	A	B	MMRE	PRED
ارگانیک	۴/۹۹۳۳	۰/۱۶۶۳	۰/۸۱۹۱	.
تعبیه شده	۵	۱/۱۷۴۵	۰/۱۴۸۷	۰/۸۰۹۵
نیمه منفصل	۰/۵۹۰۶	۱/۰۲۵۷	۰/۰۷۳۹	۰/۹۲۵۶
میانگین			۰/۳۴۷۲	۰/۵۷۸۶

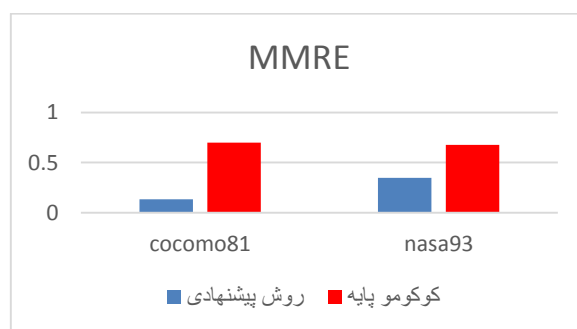
که A تعداد پروژه‌های با  $MRE \leq x$  و N تعداد پروژه‌های در نظر گرفته شده می‌باشد. معمولاً سطح قابل قبول x در متدهای تخمین هزینه نرم‌افزار [۲۹] ۰,۲۵ [متدهای مختلفی بر اساس این سطح مقایسه شده‌اند. کاهش MRE و افزایش PRED

#### ۴-۱-۲- تحلیل نتایج مدل کوکوموی پایه

جدول (۳) و شکل (۴) مقادیر مربوط MMRE و PRED حاصل از مدل کوکوموی پایه و روش ارائه شده را نشان می‌دهد. نتایج از آزمایش بر روی دو مجموعه داده ناسا ۹۳ و کوکوموی ۸۱ حاصل شده است. همان‌طور که نتایج نشان می‌دهد بر روی مجموعه داده ناسا در روش ارائه شده متوسط شدت خطای نسبی: ۰٫۳۴۷۲ به دست آمده که نسبت به مدل کوکوموی پایه با مقدار ۰٫۶۷۷۵ بهبود ۵۱٫۲۴٪ را ایجاد نموده است و بر روی مجموعه داده کوکومو ۸۱ در روش ارائه شده متوسط شدت خطای نسبی ۰٫۱۳۴۳ به دست آمده که نسبت به مدل کوکوموی پایه با مقدار ۰٫۶۹۸۴ بهبود ۱۹٫۲۲٪ را ایجاد نموده است. همچنین بهبود عملکرد بر روی مجموعه داده ناسا در مقایسه با مجموعه داده کوکوموی ۸۱ بیشتر بوده است که یکی از دلایل آن را می‌توان تعداد بیشتر پروژه‌های ناسا در مجموعه داده آزمایش دانست.

جدول ۳: نتایج کلی دو مجموعه داده کوکومو و ناسا بر روی مدل کوکوموی ۸۱

	مجموعه داده کوکومو ۸۱		مجموعه داده ناسا ۹۳	
	MMRE	PRED	MMRE	PRED
کوکومو پایه	۰٫۶۹۸۴	۰٫۵۶	۰٫۶۷۷۵	۰٫۵
روش ارائه شده	۰٫۱۳۴۳	۰٫۸۳	۰٫۳۴۷۲	۰٫۵۷



شکل ۴: نمودار PRED, MMRE مدل کوکوموی پایه

#### ۴-۱-۳- نتایج با در نظر گرفتن مدل کوکوموی میانه

جدول (۴) نتایج روش ارائه شده بر روی مجموعه داده کوکوموی ۸۱ را نشان می‌دهد این جدول ضرایب a، b و همچنین معیار عملکرد MMRE برای سه دسته مختلف را نشان می‌دهد. همان‌طور که مشاهده می‌شود بهترین MMRE برای دسته ۲ حاصل شده است و میانگین آن‌ها ۰٫۱۴۱۵ می‌باشد.

جدول ۴: نتایج مدل کوکوموی میانه، با مجموعه داده کوکومو

	A	B	MMRE	PRED
ارگانیک	۴/۹۹۸۹	۰/۵۶۷۷	۰/۱۴۸۸	۰/۸۳۳۳
تعبیه شده	۲/۵۵۰۶	۰/۴۰۰۰	۰/۰۹۹۳	۰/۹۲۸۵
نیمه منفصل	۰/۰۰۷۷	۱/۱۷۷۲	۰/۱۳۲۲	۰/۹۰۹۰
میانگین			۰/۱۴۱۵	۰/۸۹۰۲

جدول ۵: نتایج مدل کوکوموی میانه، با مجموعه داده ناسا

	A	B	MMRE	PRED
ارگانیک	۵	۰/۰۰۸۶	۰/۸۶۰۲	۰
تعبیه شده	۴/۹۴۲۱	۰/۸۲۹۲	۰/۱۲۷۱	۰/۹۰۴۷
نیمه منفصل	۴/۹۹۹۹	۰/۲۴۹۹	۰/۰۸۳۳	۰/۸۸۲۳
میانگین			۰/۳۵۶۸	۰/۵۹۵۶

جدول (۵) نتایج روش ارائه شده بر روی مجموعه داده ناسا ۹۳ را نشان می‌دهد این جدول ضرایب a، b و همچنین معیار عملکرد MMRE برای سه دسته مختلف را نشان می‌دهد. همان‌طور که مشاهده می‌شود بهترین MMRE برای دسته ۳ حاصل شده است و میانگین آن‌ها ۰٫۳۵۶۸ می‌باشد.

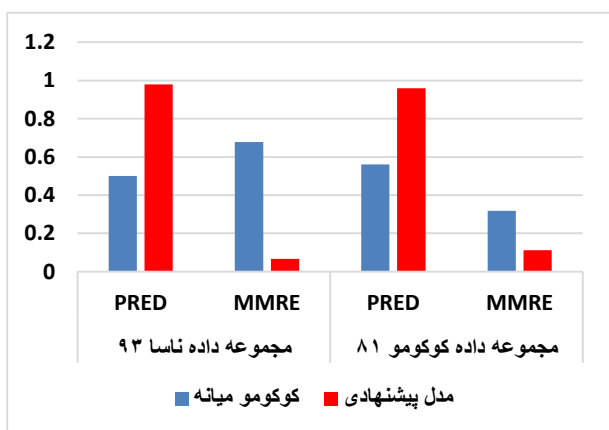
#### ۴-۱-۴- تحلیل نتایج مدل کوکوموی میانه

جدول (۶) و شکل (۵) مقادیر مربوط به MMRE و PRED حاصل از مدل کوکوموی پایه و روش ارائه شده را نشان می‌دهد. نتایج از آزمایش بر روی دو مجموعه داده ناسا ۹۳ و کوکوموی ۸۱ حاصل شده است. همان‌طور که نتایج نشان می‌دهد بر روی مجموعه داده ناسا در روش ارائه شده متوسط شدت خطای نسبی: ۰٫۳۵۶۸ به دست آمده که نسبت به مدل کوکوموی پایه با مقدار ۰٫۶۷۷۵ بهبود ۵۲٫۶۶٪ را ایجاد نموده است و بر روی مجموعه داده کوکومو ۸۱ در روش ارائه شده متوسط شدت خطای نسبی: ۰٫۱۴۱۵ به دست آمده

#### ۴-۱-۶- مقایسه روش COA با سایر روش‌ها

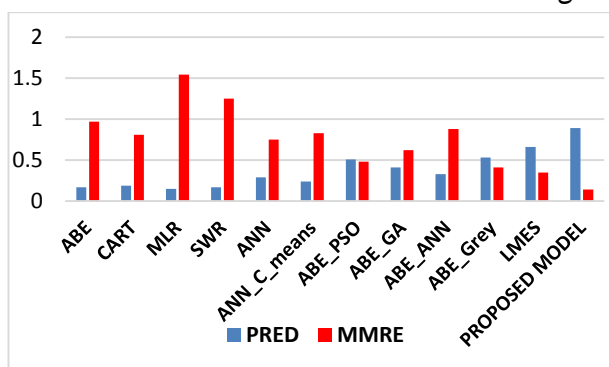
به منظور ارزیابی بهتر عملکرد روش ارائه شده، روش مورد نظر با نتایج یکی از مقالات جامع و معتبر که از همه روش‌های پیش‌بینی تلاش و از چندین دیتاست مختلف در آن استفاده شده و همچنین از نظر ارزیابی با مقاله حاضر یکسان هست، مورد مقایسه قرار گرفته است [۳۰].

شکل (۷) مقایسه عملکرد روش‌های مختلف را نشان می‌دهد. همان‌طور که مشاهده می‌شود بهترین MMRE به دست آمده



شکل ۶: مقایسه روش COA با سایر الگوریتم‌ها

مربوط به روش ارائه شده می‌باشد. کمترین درصد بهبود نسبت به مدل LMES با مقدار ۴۰/۴۲٪ حاصل شده و بیشترین درصد بهبود نسبت به مدل MLR با مقدار ۹/۱۸٪ حاصل شده است.



شکل ۷: نمودار مقایسه MMRE مدل‌های دیگر با روش ارائه شده

و همچنین به منظور ارزیابی روش ارائه شده از نتایج مقاله معتبر که در سال‌های اخیر به چاپ رسیده است استفاده شده است. در این مقاله از چندین روش برای تخمین تلاش نرم افزارهای تلفن همراه و معیارهای عملکردی (MMRE) و

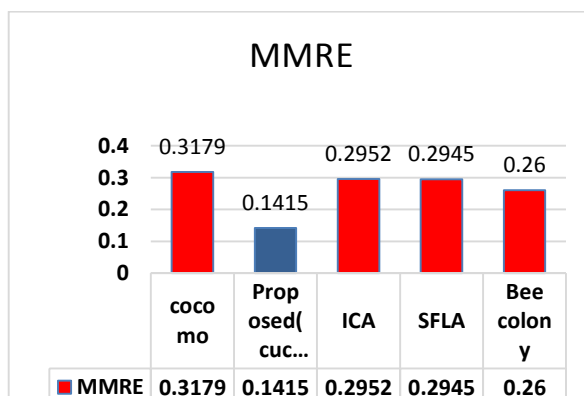
که نسبت به مدل کوکوموی پایه با مقدار ۰/۳۱۷۹ بهبود ۴۴/۵۱٪ را ایجاد نموده است. همچنین بهبود عملکرد بر روی مجموعه داده ناسا در مقایسه با مجموعه داده کوکوموی ۸۱ بیشتر بوده است که یکی از دلایل آن را می‌توان تعداد بیشتر پروژه‌های ناسا در مجموعه داده آزمایش دانست.

جدول ۶: نتایج کلی دو مجموعه داده کوکومو و ناسا بر روی مدل کوکوموی میانه

	مجموعه داده کوکومو ۸۱		مجموعه داده ناسا ۹۳	
	MMRE	PRED	MMRE	PRED
کوکومو میانه	۰/۳۱۷۹	۰/۵۶	۰/۶۷۷۵	۰/۵
روش ارائه شده	۰/۱۴۱۵	۰/۸۹	۰/۳۵۶۸	۰/۵۹

#### ۴-۱-۵- مقایسه روش COA با نتایج سایر الگوریتم‌ها و با مجموعه داده کوکومو ۸۱

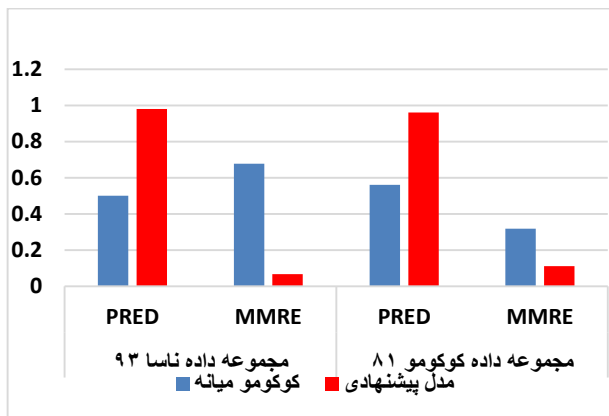
به منظور ارزیابی روش ارائه شده از نتایج مقالات معتبر که در سال‌های اخیر به چاپ رسیده است استفاده شده است. در این مرحله نیز از معیارهای عملکردی (MMRE) و (PRED) جهت ارزیابی استفاده شده است، که در شکل (۶) نشان داده شده است.



شکل ۵: نمودار PRED, MMRE مدل کوکوموی میانه

(PRED) جهت ارزیابی استفاده شده است، که در شکل (۸)

نشان داده شده است [۳۱].



شکل ۹: نمودار نتایج مربوط به روش ANN-COA بر روی مجموعه داده کوکومو و ناسا ۹۳

#### ۴-۲-۱- تحلیل نتایج در مدل ANN-COA

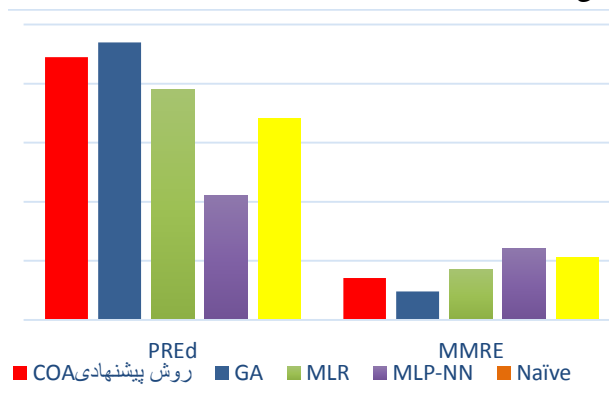
نتایج به دست آمده بر روی هر دو مجموعه داده بهبود داشته است و ما در این قسمت به تحلیل نتایج به دست آمده می‌پردازیم.

جدول (۸) و شکل (۷، ۸) مقادیر مربوط به MMRE حاصل از مدل کوکومو میانه و روش ارائه شده را نشان می‌دهد. نتایج از آزمایش بر روی مجموعه داده ناسا ۹۳ و کوکومو ۸۱ حاصل شده است. همان‌طور که نتایج نشان می‌دهد بر روی مجموعه داده ناسا در روش ارائه شده متوسط شدت خطای نسبی:  $0/0669$  به دست آمده که نسبت به مدل کوکومو میانه با مقدار  $0/6775$  بهبود  $9/87\%$  را ایجاد نموده است؛ و بر روی مجموعه داده کوکومو ۸۱ در روش ارائه شده متوسط شدت خطای نسبی:  $0/1118$  به دست آمده که نسبت به مدل کوکومو میانه با مقدار  $0/3179$  بهبود  $35/16\%$  را ایجاد نموده است.

#### ۴-۲-۲- مقایسه روش ANN-COA با سایر مدل‌ها

##### با مجموعه داده ناسا ۹۳

به منظور ارزیابی روش ارائه شده از نتایج مقالات معتبر که در سال‌های اخیر به چاپ رسیده است استفاده شده است. در این مرحله نیز از معیارهای عملکردی (MMRE) و (PRED)



شکل ۸: مقایسه روش COA با سایر روش‌ها

همان‌طور که مشاهده می‌شود بهترین MMRE به دست آمده مربوط به روش الگوریتم ژنتیک می‌باشد. و روش ارائه شده در مرتبه دوم از نظر بهترین MMRE قرار گرفته است.

#### ۴-۲-۲- نتایج در روش ANN-COA

در روش ارائه شده ابتدا داده‌ها را به دو بخش آموزش و آزمایش تقسیم می‌شوند (به صورت تصادفی و معمولاً  $80\%$  آموزش و  $20\%$  آزمون)؛ که در بخش آموزش با استفاده از الگوریتم تکاملی فاخته از بین ۱۵ فاکتور تخمین تلاش مدل کوکومو محرک‌های مؤثرتر را بر اساس تابع برازش الگوریتم فاخته که میزان دقت تخمین (MRRE) می‌باشد انتخاب می‌شوند؛ و محرک‌های مؤثر به عنوان ورودی شبکه عصبی پیشنهاد داده می‌شود و خروجی شبکه عصبی مقدار تلاش توسعه می‌باشد. این روش بر روی دو مجموعه داده کوکومو و ناسا پیاده‌سازی شده است. جدول (۷) و شکل (۸) نتایج حاصل از روش ارائه شده را بر روی مجموعه داده ناسا ۹۳ و کوکومو ۸۱ را نشان می‌دهند.

	مجموعه داده کوکومو ۸۱		مجموعه داده ناسا ۹۳	
	MMRE	PRED	MMRE	PRED
کوکومو میانه	$0/3179$	$0/56$	$0/6775$	$0/5$
روش ارائه شده	$0/1118$	$0/96$	$0/0669$	$0/98$

جدول ۷: نتایج روش ANN-COA بر روی دو مجموعه

### ۴-۳- نتیجه گیری و کارهای آینده

توسعه پروژه‌های نرم‌افزاری به دلیل ناملموس بودن، نیازمندی‌های مبهم، تولید انعطاف‌پذیر، دشوار پیچیده می‌باشد. مدیریت پروژه به‌شدت به تخمین دقیق و قابلیت اطمینان وابسته است. ماهیت خلاق پروژه‌های نرم‌افزاری و انتزاعی بودن آن تخمین هزینه و زمان انجام آن را بی‌نهایت مشکل می‌کند. در این تحقیق ابتدا مدل‌های استاندارد ارائه‌شده برای تخمین تلاش پروژه‌های نرم‌افزاری تشریح شده و سپس مدل‌های مبتنی بر الگوریتم‌های فراابتکاری مورد بحث و بررسی قرار گرفته است و درنهایت یک روش جدید تلاش لازم برای توسعه نرم‌افزار با استفاده از شبکه عصبی و الگوریتم بهینه‌سازی فاخته را ارائه شده است. مقاله حاضر تأثیر ترکیب الگوریتم بهینه‌سازی فاخته و شبکه عصبی را با مدل کوکومو پایه و میانه و با دو مجموعه داده واقعی کوکوموی ۸۱ و ناسا ۹۳ مورد بحث و بررسی قرار داده است. به‌منظور بررسی عملکرد روش ارائه شده معیار عملکرد MMRE محاسبه شده است. نتایج روش ارائه‌شده در کل باعث

جدول ۱۰: نتایج به‌دست‌آمده از رویکرد ترکیبی شبکه‌های عصبی، الگوریتم بهینه‌سازی فاخته و الگوریتم رقابتی امپریالیستی [۲۱] ژنتیک، رگرسیون خطی چندگانه و الگوریتم رقابتی امپریالیستی [۲۱] جدول ۱۰: نتایج به‌دست‌آمده از رویکرد ترکیبی شبکه‌های عصبی، الگوریتم بهینه‌سازی فاخته و الگوریتم رقابتی امپریالیستی [۲۱] ژنتیک، رگرسیون خطی چندگانه و الگوریتم رقابتی امپریالیستی [۲۱]

جهت ارزیابی استفاده شده است، که در جدول‌های (۹ و ۸) و شکل (۹) نشان داده شده است.

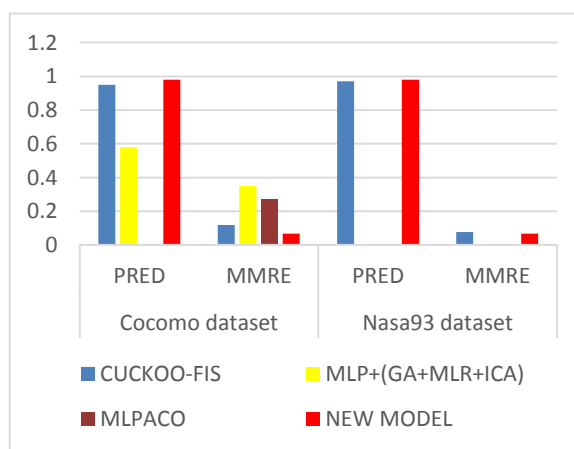
جدول ۸: نتایج به‌دست‌آمده از ترکیب شبکه عصبی و الگوریتم بهینه‌سازی فاخته [۲۰]

	MMRE
ANN-Cuckoo	۰/۱۴۸۲
روش ارائه شده	۰/۱۱۱۸

جدول ۹: نتایج به‌دست‌آمده از ترکیب شبکه عصبی MLP و الگوریتم بهینه‌سازی کلونی مورچگان [۱۷]

	مجموعه داده ناسا ۹۳	
	MMRE	PRED
MLPACO	۰/۲۶۹۱	-
روش ارائه شده	۰/۰۶۶۹	۰/۹۸

	مجموعه داده کوکومو	
	MMRE	PRED
MLP+GA+MLR+ICA	۰/۳۵	۰/۵۸
روش ارائه شده	۰/۱۱۱۸	۰/۹۶



شکل ۹: نمودار مقایسه روش ارائه شده ANN-COA با سایر مقالات

### مراجع

- [۱] Huang, G., "Cost Modeling Based on Support Vector Regression for Complex products During the Early Design Phases". Requirements for the Degree of Doctor of Philosophy in

- [۱۱] Khatibi Bardsiri, A., S.m. hashemi, and M. Razzazi, *A Novel Model for Software Services Development Effort Estimation*. Journal of Modeling in Engineering, ۲۰۱۷. ۱۵(۴۹): p. ۲۴۵-۲۶۱.
- [۱۲] Khatibi, E. and V. Khatibi Bardsiri, *An Improved Algorithmic Method for Software Development Effort Estimation*. Journal of Advances in Computer Research, ۲۰۱۸. ۹(۱): p. ۴۱-۴۹.
- [۱۳] Khatibi Bardsiri, A., *A new combinatorial framework for software services development effort estimation*. International Journal of Computers and Applications, ۲۰۱۸. ۴۰(۱): p. ۱۴-۲۴.
- [۱۴] Kumari, S. and S. Pushkar. *Software Cost Estimation Using Cuckoo Search*. in *Advances in Computational Intelligence*. ۲۰۱۷. Singapore: Springer Singapore.
- [۱۵] Puspaningrum, A. and R. Sarno, *A Hybrid Cuckoo Optimization and Harmony Search Algorithm for Software Cost Estimation*. Procedia Computer Science, ۲۰۱۷. ۱۲۴: p. ۴۶۱-۴۶۹.
- [۱۶] Çelik, E., et al. *Software test automation and a sample practice for an enterprise business software*. in *Computer Science and Engineering (UBMK)*, ۲۰۱۷ *International Conference on*. ۲۰۱۷. IEEE.
- [۱۷] Ebrahimpour, N., F. Soleimanian Gharehchopogh, and Z. Abbasi Khalifehlou, *New Approach with Hybrid of Artificial Neural Network and Ant Colony Optimization in Software Cost Estimation*. Journal of Advances in Computer Research, ۲۰۱۶. ۷(۴): p. ۱-۱۲.
- [۱۸] Wani, Z.H. and S. Quadri. *Artificial Bee Colony-Trained Industrial and Systems Engineering*, ۲۰۰۷.
- [۲] Khatibi, v. and D.N.A. awawi, *Software Cost Estimation Methods: A Review*. Journal of Emerging Trends in Computing and Information Sciences, ۲۰۱۱ ۲: p. ۲۱-۲۹.
- [۳] Abdullah, T., et al., *ANALYSIS OF SOFTWARE COST ESTIMATION MODELS*. IJEIR, ۲۰۱۲ ۱: p. ۲۰۶-۲۱۲.
- [۴] Binish Zahra, S. and M. Nazir, *A Review of Comparison among Software Estimation Techniques*. Bahria University Journal of Information & Communication Technology, ۲۰۱۲ ۵: p. ۳۹-۴۵.
- [۵] Boehm, B., "Software Engineering Economics". IEEE Transaction on Software Engineering, vol. SE-۱۰, ۱۹۸۴ p. ۴-۲۱.
- [۶] Lin, J.-C., et al., *Using Computing Intelligence Techniques to Estimate Software Effort*. International Journal of Software Engineering & Applications, ۲۰۱۳ ۴(۱): p. ۴۳.
- [۷] khatibi bardsiri, v. and m. dorosti, *An Improved COCOMO based Model to Estimate the Effort of Software Projects*. ۲۰۱۶.
- [۸] Sadeghi, B., et al., *A Novel ICA-based Estimator for Software Cost Estimation*. Journal of Advances in Computer Engineering and Technology, ۲۰۱۵. ۱(۴): p. ۱۵-۲۴.
- [۹] Venkataiah, V., et al., *Application of ant colony optimization techniques to predict software cost estimation*, in *Computer Communication, Networking and Internet Security*. ۲۰۱۷, Springer. p. ۳۱۵-۳۲۵.
- [۱۰] Shahpar, Z., et al., *Improvement of effort estimation accuracy in software projects using a feature selection approach*. Journal of Advances in Computer Engineering and Technology, ۲۰۱۶. ۲(۴): p. ۳۱-۳۸.

- [۲۵] Rajabioun, R., *Cuckoo optimization algorithm*. Applied soft computing, ۲۰۱۱. ۱۱(۸): p. ۵۵۰۸-۵۵۱۸.
- [۲۶] Karaboga, D., B. Akay, and C. Ozturk, *Artificial bee colony (ABC) optimization algorithm for training feed-forward neural networks*. MDAI, ۲۰۰۷. ۷: p. ۳۱۸-۳۱۹.
- [۲۷] Hippert, H.S. and J.W. Taylor, *An evaluation of Bayesian techniques for controlling model complexity and selecting inputs in a neural network for short-term load forecasting*. Neural networks, ۲۰۱۰. ۲۳(۳): p. ۳۸۶-۳۹۵.
- [۲۸] Khatibi, E. *Investigating the effect of software project type on accuracy of software development effort estimation in COCOMO model*. in *Fourth International Conference on Machine Vision (ICMV ۱۱)*. ۲۰۱۱ International Society for Optics and Photonics.
- [۲۹] Khatibi, V. and D.N. Jawawi, *Software cost estimation methods: A review ۱*. ۲۰۱۱
- [۳۰] Bardsiri, V.K., et al., *LMES: A localized multi-estimator model to estimate software development effort*. Engineering Applications of Artificial Intelligence, ۲۰۱۳. ۲۶(۱۰): p. ۲۶۲۴-۲۶۴۰.
- [۳۱] Pandey, M., R. Litoriya, and P. Pandey, *Validation of existing software effort estimation techniques in context with mobile software applications*. Wireless Personal Communications, ۲۰۲۰. ۱۱۰(۴): p. ۱۶۵۹-۱۶۷۷.
- [۲۴] جهانشاهی, س. و. خطیبی بردسیری, و همکاران. کاربرد روشهای محاسبات نرم در بهبود عملکرد مدل کوکومو بمنظورتخمین هزینه پروژه های نرم افزار. همایش ملی مهندسی کامپیوتر و فناوری اطلاعات, اسفند ۱۳۹۲
- [۱۹] Chhabra, S. and H. Singh, *Optimizing design parameters of fuzzy model based COCOMO using genetic algorithms*. International Journal of Information Technology, ۲۰۱۹.
- [۲۰] Desai, V.S. and R. Mohanty. *ANN-Cuckoo Optimization Technique to Predict Software Cost Estimation*. in *۲۰۱۸ Conference on Information and Communication Technology (CICT)*. ۲۰۱۸. IEEE.
- [۲۱] Khazaiepoor, M., A. Khatibi Bardsiri, and F. Keynia, *A Hybrid Approach for Software Development Effort Estimation using Neural networks, Genetic Algorithm, Multiple Linear Regression and Imperialist Competitive Algorithm*. International Journal of Nonlinear Analysis and Applications, ۲۰۲۰. ۱۱(۱): p. ۲۰۷-۲۲۴.
- [۲۲] Boehm, B., C. Abts, and S. Chulani, *Software development cost estimation approaches—A survey*. Annals of software engineering, ۲۰۰۰. ۱۰(۱-۴): p. ۱۷۷-۲۰۵.
- [۲۳] Stutzke, R.D. and M. Crosstalk, *Software estimating technology: A survey*. ۱۹۹۷: Los. Alamitos, CA: IEEE Computer Society Press.





## بهبود الگوریتم رقابت استعماری برای حل مسئله جایگذاری نودها در شبکه

### های حسگر بی سیم گرید سه بعدی

سیدوفا بارخدا\* همت شیخی\* سودابه محمدی\*

\* عضو هیئت علمی گروه کامپیوتر، دانشکده فناوری اطلاعات، دانشگاه صنعتی کرمانشاه

تاریخ پذیرش: ۱۳۹۹/۱۰/۰۵

تاریخ دریافت: ۱۳۹۹/۰۳/۳۰

نوع مقاله: پژوهشی

#### چکیده

یکی از زمینه‌های تحقیقاتی اساسی و مهم در شبکه‌های حسگر بی سیم نحوه جایگذاری نودهای حسگر است به گونه‌ای که با کمترین تعداد نود تمامی نقاط هدف پوشش داده شوند و اتصال میان تمام نودها و نود چاهک برقرار باشد. در این مقاله از یک روش جدید که بر اساس الگوریتم رقابت استعماری است برای حل مسئله ذکر شده استفاده شده است. در روش پیشنهاد شده امکان مهاجرت مستعمره‌ها از امپراطوری‌های ضعیف به امپراطوری‌های قوی‌تر به الگوریتم رقابت استعماری اضافه شده است. ایده مهاجرت از جوامع انسانی الهام گرفته شده است که انسان‌ها در برخی شرایط تصمیم به مهاجرت از یک کشور به کشور دیگر می‌کنند. شبکه حسگر بی سیم به صورت سه بعدی و گرید در نظر گرفته شده است و نودهای حسگر فقط می‌توانند در نقاط تقاطع گرید قرار بگیرند. این در حالیست که نقاط هدف ممکن است در هر مکانی از فضای سه بعدی پراکنده باشند. نتایج شبیه‌سازی نشان می‌دهد که الگوریتم پیشنهادی نسبت به الگوریتم‌های مشابه از تعداد نود حسگر کمتری برای حل مسئله استفاده می‌کند و همچنین دارای زمان اجرای بسیار کمتری است.

**واژگان کلیدی:** شبکه حسگر بی سیم، شبکه گرید سه بعدی، الگوریتم رقابت استعماری، مهاجرت، جایگذاری نود.

#### ۱. مقدمه

غیرقابل دسترس و غیره اشاره نمود [۱]. همچنین بسیاری از تحقیقات در حال تجمیع شبکه‌های حسگر بی سیم با اینترنت اشیا هستند، به نحوی که نودهای حسگر بتوانند به راحتی به اینترنت متصل شده و وظایف خود را انجام دهند [۲]. با توجه به نوع کاربرد، محیط عملیاتی برای قرارگیری حسگرها را می‌توان به صورت دو بعدی و یا سه بعدی در نظر گرفت؛ لازم به ذکر است که در بیشتر تحقیقات انجام شده بر روی شبکه‌های حسگر در سالیان گذشته، آن‌ها را به صورتی فرض کرده‌اند که حسگرها در یک محیط مسطح زمینی قرار می‌گیرند. اما در سال‌های اخیر کاربردهای فراوانی

امروزه یکی از حوزه‌های مورد توجه محققان کامپیوتر، نحوه ایجاد شبکه‌های حسگر بی سیم است و پژوهش‌های زیادی بر روی جنبه‌های مختلف طراحی و پیاده‌سازی این شبکه‌ها در حال انجام هستند. از جمله مزیت‌های این نوع از شبکه‌ها نسبت به شبکه‌های کامپیوتری سنتی می‌توان به مقیاس‌پذیری، انعطاف‌پذیری و تنوع فراوان کاربردهای آن‌ها از قبیل پیگیری نقاط هدف، نظارت بر محیط‌های

نویسنده مسئول: همت شیخی [h.sheikhi@kut.ac.ir](mailto:h.sheikhi@kut.ac.ir)

که هیچ رخدادی را مانیتور نمی‌کنند و فقط داده‌های سایر نودها را دریافت و ارسال می‌کنند.

موضوع بررسی شده در این مقاله بدین شرح است که نود چاهک به همراه تعدادی هدف نقطه‌ای با موقعیت تصادفی در فضای یک شبکه حسگر سه‌بعدی قرار دارند. مساله انتخاب رئوسی در سلول‌ها برای جایگذاری نودهای حسگر است که علاوه بر اینکه تمامی نقاط هدف پوشش داده شوند، ارتباط تمامی نودهای حسگر با چاهک نیز برقرار شود. بدیهی است که به‌منظور کاهش هزینه، این کار باید با کمترین تعداد ممکن نود انجام شود. در [۹] حل مسائل پوشش و اتصال برای شبکه‌های حسگر بی‌سیم به عنوان مسئله‌ای NP-کامل اثبات شده و از این رو برای حل این مسائل، عموماً از روش‌های اکتشافی و یا متا اکتشافی استفاده می‌شود [۱۰].

با توجه به ماهیت بهینه‌سازی در حل مسئله جایگذاری حسگر در شبکه‌های حسگر بی‌سیم، استفاده از روش‌های تکاملی بسیار مورد استقبال بوده و نتایج قابل توجهی به دست آمده است. در این میان به ویژه استفاده از الگوریتم‌های ژنتیک، کلونی مورچگان و کلونی زنبورها مورد توجه بوده است که در همه آنها به نوعی از طبیعت الهام گرفته شده است [۱۲][۱۱]. یکی از روش‌های تکاملی جدیدتر که از جوامع انسانی الهام گرفته و علیرغم داشتن پتانسیل زیاد برای حل این مسئله، تا کنون به صورت جدی مورد توجه قرار نگرفته است، الگوریتم رقابت استعماری می‌باشد.

این الگوریتم همانند دیگر الگوریتم‌های تکاملی، نیازی به دانستن اطلاعات گرادیان مرتبط با تابع هدف نداشته و بنابراین به سادگی می‌تواند در مسائلی که تعیین اطلاعات گرادیان سخت و یا غیر ممکن است، مورد استفاده قرار گیرد. الگوریتم رقابت استعماری مزیت‌های قابل توجهی نسبت به سایر الگوریتم‌های تکاملی دارد و بنابراین در این مقاله به عنوان الگوریتم پایه در روش پیشنهادی در نظر گرفته شده است. زمان اجرای این الگوریتم به صورت قابل ملاحظه‌ای کمتر از سایر الگوریتم‌ها می‌باشد. همچنین از این الگوریتم می‌توان در فضاهای مسئله ناپیوسته و یا با تغییرات بسیار شدید استفاده کرد. بعلاوه، به جای اینکه در پایان فقط یک جواب برای مسئله ارائه دهد، مجموعه‌ای از راه‌حل‌های بهینه در اختیار طراح قرار می‌دهد [۱۳].

الگوریتم رقابت استعماری، کاملاً مقیاس‌پذیر بوده و مؤلفه‌های جدید را می‌توان به آن اضافه کرد. بنابراین، در این مقاله ایده مهاجرت به عنوان یک مؤلفه پیشنهادی مورد مطالعه قرار گرفته و با اصلاح و افزودن آن به الگوریتم رقابت استعماری، از یک روش بهبود یافته برای حل مساله

برای این شبکه‌ها پیشنهاد شده است که نیازمند قرارگیری نودها در ارتفاع هستند [۲].

به عنوان مثال در شبکه‌های حسگر بی‌سیم زیرآبی، گره‌ها در یک محیط سه‌بعدی قرار دارند و از سیگنال‌های صوتی برای برقراری ارتباط استفاده می‌کنند. در سالهای اخیر این شبکه‌ها بسیار مورد توجه قرار گرفته‌اند و از آنها برای نظارت و اکتشاف در اقیانوس‌ها استفاده شده است. کاربردهایی از قبیل هشدار سونامی، کمک به نوابری و جمع‌آوری اطلاعات اقیانوس‌شناسی از جمله کاربردهای شبکه‌های حسگر زیرآبی است [۴] [۳]. همچنین می‌توان به گره‌هایی که بر روی درختان یک جنگل قرار می‌گیرند و یک ویژگی خاص مانند نحوه رشد برگ‌ها را نظارت می‌کنند اشاره نمود. با توجه به این کاربردها و همچنین این نکته که محیط‌های دوبعدی اساساً زیرمجموعه محیط‌های سه‌بعدی هستند، در این مقاله محیط عملیاتی به‌صورت سه‌بعدی در نظر گرفته شده است. به علاوه، نحوه قرارگیری نودهای حسگر از استراتژی گرید تبعیت می‌کند که در آن شبکه سه‌بعدی به سلول‌های مکعبی تقسیم شده و هر نود حسگر می‌تواند در رئوس این سلول‌ها قرار بگیرد.

یکی از چالش‌های اساسی در شبکه‌های حسگر بی‌سیم این است که نودهای حسگر باید طوری در محیط عملیاتی جایگذاری شوند که بتوانند کل محیط یا نقاط و بخش‌های مورد نیاز را پوشش دهند و رخدادهای اتفاق افتاده در آن مناطق را مانیتور کنند [۵]. بدین منظور هر نود دارای یک شعاع حسی است و قادر است تمام رویدادهای اتفاق افتاده در این شعاع را مانیتور کند. پوشش یک محیط عملیاتی به سه روش پوشش کلی، پوشش نقطه‌ای و پوشش حصاری قابل انجام است که در این میان پوشش نقطه‌ای کاربردهای واقعی بیشتری دارند [۷] [۶]. در این نوع پوشش، تنها باید مکان‌های خاصی از شبکه پوشش داده شوند؛ به عنوان مثال در یک جنگل به جای پوشش کل فضا، فقط مکان‌های نزدیک و روی درختان پوشش داده می‌شود. این کار باعث می‌شود که از جایگذاری نودهای اضافی در شبکه اجتناب شود و هزینه ایجاد شبکه، کاهش چشمگیری داشته باشد. از این رو یکی از مسائلی که در این مقاله مورد بررسی قرار می‌گیرد، مسئله پوشش نقطه‌ای در شبکه‌های حسگر است.

در جایگذاری نودها علاوه بر مسئله پوشش باید ویژگی متصل بودن شبکه نیز در نظر گرفته شود. در شبکه‌های حسگر بی‌سیم فراهم کردن پوشش بدون برقراری اتصال باعث کاهش کیفیت شبکه می‌شود زیرا تضمینی وجود ندارد که رخدادهای کشف شده در محیط عملیاتی، به‌درستی به نود چاهک گزارش شوند [۸]. برای متصل کردن نودهای حسگر به چاهک، ممکن است نیاز به جایگذاری نودهای اضافی باشد

چاهک‌ها متصل باشند. در برخی دیگر از پژوهش‌ها شرط پوشش بررسی نشده و صرفاً اتصال نودهایی که حسگر در آن‌ها قرار دارد مد نظر بوده است [۱۶][۱۵].

در [۱۸][۱۷] مسئله به صورت  $k$ -پوشش و  $m$ -متصل حل شده است و همه حسگرها هنوز دارای شعاع پوشش یکسان می‌باشند در حالیکه در [۱۹] مسئله  $m$ -متصل به صورت غیرهمگن حل شده است. نکته شایان توجه این است که همه الگوریتم‌های ذکر شده فضای مسئله را به صورت دو بعدی در نظر گرفته‌اند. در [۱۴] با استفاده از الگوریتم کلونی مورچگان مسئله جایگذاری حسگرها در یک شبکه حسگر بی-سیم سه بعدی حل شده است.

الگوریتم‌های تکاملی، دسته‌ای از الگوریتم‌ها هستند که عموماً از طبیعت و سیستم‌های بیولوژیکی طبیعی الهام گرفته‌اند. برخی از این الگوریتم‌ها همچون الگوریتم کلونی زنبورعسل [۲۰] و PSO [۲۱] برای حل مسئله WSN مورد استفاده قرار گرفته‌اند. یکی از الگوریتم‌های تکاملی جالب توجه که بیشتر برای حل این مسئله به کار گمارده شده است، الگوریتم کلونی مورچگان است. در [۲۲] با استفاده از این الگوریتم، روشی با عنوان EasiDesign برای حل مسائل مربوط به شبکه‌های گرید ارائه شده است. این رویه در [۲۳] با استفاده از افزایش تعداد مورچگان بهبود یافته و در [۲۴] هم یک روش حریصانه بر مبنای همین الگوریتم برای بهبود کارایی ارائه شده است.

هنگامی که مسئله  $k$ -پوشش مورد بررسی قرار می‌گیرد تلاش داریم تا همه نقاط هدف در شعاع پوشش  $k$  حسگر قرار گیرند. اما ممکن است در برخی از قسمت‌های شبکه، تعدادی از نقاط هدف توسط بیش از  $k$  حسگر پوشش داده شوند. در واقع در این بخش‌ها تعدادی حسگر قرار داده شده‌اند که در صورت حذف مشکلی در صحت و کلیت جواب پیش نخواهد آمد. این حسگرها که حسگرهای افزوده نامیده می‌شوند، موجب تحمیل هزینه غیرضروری و اضافی به سیستم شده و تشخیص و حذف آنها موجب افزایش کارایی الگوریتم خواهد شد. همین مشکل در مورد اتصال حسگرها هم وجود دارد. در [۲۵] راهکاری برای تشخیص و حذف حسگرهای افزوده با استفاده از الگوریتم کلونی مورچگان ارائه شده است.

تا کنون تحقیقات زیادی برای استفاده از روش رقابت استعماری برای حل مسائل پوشش و اتصال در شبکه‌های حسگر بی‌سیم انجام نشده است. در [۲۶] با استفاده از الگوریتم رقابت استعماری، یک الگوریتم به نام ICABC برای مسئله ۱-پوشش حصارها و ۱-متصل نودها در هر حصار پیشنهاد شده است. در [۲۷] نیز با استفاده از همین الگوریتم،

پوشش و اتصال در شبکه‌های حسگر بی‌سیم گرید سه‌بعدی استفاده شده است.

نوآوری‌های انجام شده در این مقاله بدین شرح است:

- برای اولین بار از الگوریتم رقابت استعماری به عنوان پایه برای حل مسائل پوشش و اتصال در شبکه‌های حسگر سه بعدی بر مبنای نقاط هدف استفاده شده است.
- ایده مهاجرت برای اولین بار ارائه شده و به عنوان یک مولفه اصلی به الگوریتم رقابت استعماری اضافه شده است.
- فرمول‌بندی ریاضی هم برای مولفه مهاجرت و هم برای جایگذاری گره‌ها، ارائه شده است.

در ادامه این مقاله، ابتدا کارهای قبلی انجام گرفته در بخش ۲ مورد بررسی قرار خواهند گرفت. در بخش ۳ پیش‌زمینه‌ای از الگوریتم رقابت استعماری شرح داده خواهد شد. سپس در بخش ۴ روش پیشنهادی شامل بهبود الگوریتم رقابت استعماری و همچنین فرمول‌بندی مسئله بیان شده و در قسمت ۵ نتایج شبیه‌سازی و مقایسه کارایی آن با الگوریتم‌های دیگر بررسی خواهد شد. نهایتاً جمع‌بندی و نتیجه‌گیری در بخش ۶ ارائه خواهد شد.

## ۲. مرور کارهای انجام گرفته

تا کنون، مسئله شبکه حسگر بی‌سیم در پژوهش‌های زیادی مورد بررسی قرار گرفته است. با توجه به اینکه هر حسگر هزینه‌ای را به سیستم تحمیل می‌کند، به صورت خاص مهم‌ترین نکته جایگذاری حسگرها در داخل شبکه به گونه‌ای است که با استفاده از کمترین تعداد حسگر، کل نقاط مورد نظر در داخل شبکه پوشش داده شوند. علاوه بر شرط پوشش، مسئله اتصال هم در بیشتر موارد مورد توجه بوده است. در هر پژوهش فرضیات خاصی برای نوع شبکه، ساختار نودها و نحوه ارتباط در نظر گرفته شده است. به عنوان مثال نودها می‌توانند همگن یا غیرهمگن باشند و یا اینکه هر نود حسگری می‌تواند داده‌های دیگر نودها را منتقل کند یا نه. بحث تحمل‌پذیری در برابر خرابی و از کار افتادن نودها هم در برخی تحقیقات مورد توجه قرار گرفته شده است. در [۱۴] همه حسگرها دارای شعاع پوشش همسان در نظر گرفته شده و همچنین مسئله به صورت ۱-پوشش و ۱-متصل حل شده است؛ بدین معنا که هر نقطه هدف لزوماً باید توسط حداقل یک حسگر پوشش داده شده و همچنین هر حسگر حداقل دارای یک مسیر به نود چاهک باشد. همه این شرایط در [۹] در نظر گرفته شده است با این تفاوت که در آن چند نود چاهک موجود بوده و حسگرها می‌توانند به هر کدام از این

راستای ابعاد مختلف اجتماعی-سیاسی به خود نزدیک کند. همچنین در طی مراحل رقابت استعماری، هر کدام از کشورهای مستعمره ممکن است دچار انقلاب شوند. انقلاب یک تغییر بنیادی در قدرت یا سازماندهی یک کشور است که در یک بازه کوتاه زمانی اتفاق می‌افتد. این تغییرات جدای از تغییرات مرحله همگون‌سازی بوده که توسط استعمارگر اعمال می‌شوند و به صورت تغییرات تصادفی در موقعیت یک کشور در راستای محورهای سیاسی-اجتماعی مدل می‌شود. این فرایند قدرت جستجوی الگوریتم را توسعه داده و از همگرا شدن به کمینه‌های محلی جلوگیری می‌نماید [۲۹].

سیاست جذب در عین نابودی ساختارهای اجتماعی-سیاسی کشور مستعمره، در بعضی موارد نتایج مثبتی را نیز برای آن‌ها در پی دارد. بدین صورت که در حین حرکت مستعمرات به سمت کشور استعمارگر، ممکن است بعضی از این مستعمرات به موقعیتی بهتر از استعمارگر برسند؛ این اتفاق در طی فرایند انقلاب هم ممکن است رخ دهد. در این حالت، کشور استعمارگر و کشور مستعمره، جای خود را با هم عوض کرده و الگوریتم با کشور استعمارگر در موقعیت جدید ادامه یافته و این بار این کشور استعمارگر جدید است که شروع به اعمال سیاست همگون‌سازی بر مستعمرات خود می‌کند.

در حین جابجایی استعمارگران و مستعمراتشان به سمت کمینه سراسری، ممکن است برخی استعمارگران به نقاط مشابهی برسند. اگر فاصله دو استعمارگر از یک حد آستانه کمتر شود، آنگاه با همدیگر ادغام شده و یک امپراطوری جدید را شکل می‌دهند. همچنین همه مستعمرات دو استعمارگر هم در این امپراطوری جدید حضور خواهند داشت.

هر امپراطوری که نتواند بر قدرت خود بیفزاید و قدرت رقابت خود را از دست بدهد در جریان رقابت‌های استعماری حذف خواهد شد. این حذف شدن به صورت تدریجی صورت گرفته بدین معنی که به مرور زمان امپراطوری‌های ضعیف مستعمرات خود را از دست داده و امپراطوری‌های قوی‌تر آنها را تصاحب می‌کنند. در واقع در هر تکرار از الگوریتم، ضعیف‌ترین مستعمره از ضعیف‌ترین امپراطوری توسط سایر امپراطوری‌ها تصرف می‌شود. استعمارگری که همه مستعمرات خود را از دست بدهد حذف شده و به عنوان مستعمره به بقیه استعمارگران اختصاص می‌یابد. نهایتاً پس از تعداد مناسب تکرار، همه امپراطوری‌ها سقوط کرده و تنها یک امپراطوری باقی می‌ماند. در چنین موقعیتی رقابت استعماری به پایان رسیده و استعمارگر باقی‌مانده به عنوان جواب نهایی مسئله ارائه می‌گردد.

رویه‌ای با نام MOICA برای ۱-پوشش تمام محیط پیشنهاد شده است. هر دوی این متدها مسائل را در فضای دو بعدی مدل‌سازی کرده‌اند. در این مقاله با افزودن ایده مهاجرت به الگوریتم رقابت استعماری و بهبود آن، یک روش جدید برای حل مسئله جایگذاری حسگرها در شبکه حسگر بی‌سیم سه بعدی ارائه خواهد شد.

### ۳. پیش‌زمینه: الگوریتم رقابت استعماری

الگوریتم رقابت استعماری [۲۸] یکی از الگوریتم‌های تکاملی بوده که از جوامع بشری الهام گرفته و برای حل بسیاری از مسائل بهینه‌سازی قابل استفاده است. این الگوریتم همانند دیگر الگوریتم‌های تکاملی با تعدادی جمعیت اولیه تصادفی، که هر کدام از آن‌ها یک کشور نامیده می‌شوند، شروع می‌شود. این کشورها در واقع جواب‌های ممکن مسئله می‌باشند. تعدادی از بهترین عناصر جمعیت به عنوان استعمارگر و باقیمانده جمعیت به عنوان مستعمره در نظر گرفته می‌شود. مستعمرات با توجه به میزان قدرت استعمارگران، میان آنها تقسیم شده و تشکیل امپراطوری‌ها را می‌دهند. سپس رقابت استعماری میان امپراطوری‌ها شروع می‌شود.

هر امپراطوری که نتواند در رقابت استعماری، موفق عمل کرده و بر قدرت خود بیفزاید به تدریج از صحنه رقابت حذف خواهد شد. بنابراین بقای یک امپراطوری وابسته به قدرت آن در جذب مستعمرات امپراطوری‌های رقیب و به سيطرة در آوردن آن‌ها خواهد بود. در نتیجه در جریان رقابت‌های استعماری به تدریج بر قدرت امپراطوری‌های بزرگتر افزوده شده و امپراطوری‌های ضعیف‌تر حذف خواهند شد. استعمارگران برای افزایش قدرت مجبور خواهند شد تا مستعمرات خود را نیز پیشرفت دهند. با گذشت زمان مستعمرات از لحاظ قدرت به استعمارگرها نزدیک‌تر شده و شاهد یک نوع همگرایی خواهیم بود. حد نهایی رقابت استعماری زمانی است که یک امپراطوری واحد با مستعمراتی که از لحاظ موقعیت به خود کشور استعمارگر خیلی نزدیک هستند، باقی بماند.

یکی از مهمترین مراحل رقابت استعماری، حرکت مستعمره‌ها به سمت استعمارگر است. در واقع سیاست همگون‌سازی (جذب) با هدف تحلیل فرهنگ و ساختار اجتماعی مستعمرات در فرهنگ حکومت مرکزی انجام می‌گیرد. کشورهای استعمارگر برای افزایش نفوذ خود، شروع به ایجاد تغییرات در کشورهای مستعمره خود می‌کنند. با در نظر گرفتن شیوه نمایش یک کشور در حل مسئله بهینه‌سازی، در حقیقت حکومت مرکزی با اعمال سیاست جذب سعی دارد تا کشور مستعمره را در

#### ۱.۴ ایده پیشنهادی

در این مقاله با استفاده از الگوریتم رقابت استعماری، مسئله جایگذاری حسگرها در شبکه حسگر بی‌سیم سه بعدی با شرایط ۱- پوشش و ۱- متصل حل شده است. برای رسیدن به نتایج بهتر، بهبودهایی در الگوریتم رقابت استعماری انجام گرفته است. بدین منظور ایده مهاجرت مدل سازی شده و به این الگوریتم اضافه شده است. مهاجرت باعث ایجاد تنوع بیشتر در امپراطوری‌ها شده و از همگرا شدن آن‌ها به بهینه محلی جلوگیری می‌کند. در ادامه، ابتدا این ایده بررسی شده و سپس فرمول بندی مسئله و الگوریتم پیشنهادی ارائه می‌شود.

#### ۱.۴ مهاجرت: بهبود الگوریتم رقابت استعماری

همانطور که قبلاً ذکر شد، الگوریتم رقابت استعماری از جوامع انسانی الهام گرفته شده است. از این رو رویدادهایی که در دنیای واقعی در این جوامع رخ می‌دهد را می‌توان به این الگوریتم اضافه نمود و نتایج آن را تحلیل و بررسی نمود. در کارهای انجام شده قبلی اتفاقاتی نظیر انقلاب در یک مستعمره، اتحاد امپراطوری‌ها و سقوط یک امپراطوری بررسی شده است. بدیهی است که هر رویداد می‌تواند نتایج مثبت و منفی را برای کشورها، تمدن‌ها و دیگر عوامل تاثیرگذار در جوامع انسانی داشته باشد. بنابراین انجام شدن یک رویداد باید تحت شرایط خاص و به صورت حساب شده صورت پذیرد تا اثرات منفی آن کمتر و اثرات مثبت بیشتر شود.

یکی از اتفاقاتی که در جوامع بشری به وفور رخ می‌دهد اما در الگوریتم رقابت استعماری در نظر گرفته نشده است، مساله مهاجرت از یک سرزمین به سرزمین دیگر است. از نظر تاریخی تاکنون مهاجرت‌های زیادی صورت گرفته که بسیاری از آنها مبدا پیدایش تمدن‌های بزرگ بشری بوده‌اند، هر چند که در برخی موارد نتایج منفی به همراه داشته است. در این مقاله رویداد مهاجرت به الگوریتم رقابت استعماری اضافه شده است و نتایج آن در حل مسائل پوشش و اتصال در شبکه‌های حسگر بی‌سیم مورد بررسی قرار گرفته شده است.

می‌دانیم که در الگوریتم رقابت استعماری هر مستعمره در هر لحظه فقط عضو یک امپراطوری است و یا به عبارتی مستعمره یک استعمارگر است. در فاز اولیه امپراطوری‌ها تشکیل می‌شوند و سپس رقابت بین امپراطوری‌ها برای کسب قدرت بیشتر و تصرف مستعمرات سایر امپراطوری‌ها آغاز می‌شود. این رقابت به صورت تکراری است که با تعداد دوره‌های مشخص انجام می‌گیرد. هر مستعمره در انتهای هر دور،

با یک احتمال، می‌تواند از امپراطوری خود مهاجرت کرده و به یک امپراطوری جدید وارد شود. این احتمال که یک عدد در بازه [۰, ۱] است باید در هر دور و برای تمام مستعمرات محاسبه شود. هر مقدار که شرایط یک مستعمره در دوره‌های قبلی وخیم‌تر باشد احتمال مهاجرت آن مستعمره افزایش می‌یابد. در جدول ۱ نمادها و تعاریف مورد نیاز برای فرمول بندی و انجام محاسبات نشان داده شده است. هر کشور یکی از جواب‌های مسئله است که با توجه به قدرتش می‌تواند در یک امپراطوری به عنوان استعمارگر و یا مستعمره حضور داشته باشد. بنابراین هر کشور در هر لحظه می‌تواند در یکی از این دو وضعیت قرار داشته باشد که این موضوع در رابطه ۱ مشخص شده است.

$$S_i C_e E_e = \begin{cases} 1, & \text{if country } i \text{ is an imperialist} \\ 0, & \text{if country } i \text{ is an colony} \end{cases} \quad (1)$$

جدول ۱. نمادهای مورد نیاز برای فرمول بندی مهاجرت

نماد	تعریف
$E_e$	امپراطوری شماره e شامل تعدادی کشور
$P_i E_e$	قدرت امپراطوری شماره e در تکرار i
$P_{ini} E_e$	قدرت امپراطوری شماره e در مرحله اولیه تشکیل امپراطوری‌ها
$N_i E_e$	تعداد کشورها در امپراطوری شماره e در تکرار i
$G_i E_e$	میزان رشد امپراطوری شماره e در تکرار i
$C_e E_e$	کشور شماره c در امپراطوری شماره e
$P_i C_e E_e$	قدرت کشور شماره c در امپراطوری شماره e در تکرار i
$P_{ini} C_e E_e$	قدرت کشور شماره c در امپراطوری شماره e در مرحله اولیه تشکیل امپراطوری‌ها
$S_i C_e E_e$	وضعیت کشور شماره c در امپراطوری شماره e در تکرار i
$N_i C_e E_e$	تعداد نودهای حسگر در کشور شماره c در امپراطوری شماره e در تکرار i
$G_i C_e E_e$	میزان رشد کشور شماره c در امپراطوری شماره e در تکرار i
$H_i C_e E_e$	امید به مهاجرت کشور شماره c از امپراطوری شماره e در تکرار i
$I_i C_e E_e$	احتمال مهاجرت کشور شماره c از امپراطوری شماره e در تکرار i

در پایان احتمال مهاجرت هر مستعمره به یک امپراطوری جدید برای دور بعدی بر اساس رابطه ۸ محاسبه می‌شود که برابر میانگین وزنی امید به مهاجرت مستعمره در تکرار جاری و احتمال مهاجرت در تکرارهای قبلی است.

$$I_i C_c E_e = \begin{cases} \beta \times H_{i-norm} C_c E_e + (1 - \beta) \times I_{i-1} C_c E_e & \text{if } i > 1 \\ \beta \times H_{i-norm} C_c E_e & \text{if } i = 1 \end{cases} \quad (۸)$$

## ۲.۴ مدل شبکه

در این مقاله محیط عملیاتی به صورت سه بعدی و گرید در نظر گرفته شده که تعدادی هدف به صورت تصادفی در تمام فضای محیط پخش می‌شوند. نودهای حسگر را فقط می‌توان در نقاط تقاطع گریدها جایگذاری کرد. هر نقطه هدف باید در شعاع حسی حداقل یک نود حسگر قرار داشته باشد و به عبارتی توسط آن حسگر پوشش داده شود. نودهای حسگر به صورت همگن در نظر گرفته شده‌اند، به این صورت که هم شعاع حسی همه نودها و هم شعاع انتشار آن‌ها با هم یکسان است هر چند که این دو شعاع لزوماً با هم برابر نیستند.

در شبکه یک چاهک وجود دارد که داده تمام حسگرها، برای آن ارسال می‌شود. بین هر نود حسگر و چاهک باید حداقل یک مسیر وجود داشته باشد. نقاط تقاطع گریدها به عنوان مکان‌های بالقوه برای قرارگیری نودهای حسگر و چاهک معرفی می‌شوند. هدف تعیین کمترین تعداد نودهای حسگر و مکان آنها در شبکه است به نحوی که هر نقطه هدف حداقل توسط یک نود حسگر پوشش داده شود و هر نود حسگر حداقل با یک مسیر به چاهک متصل باشد. جدول‌های ۲ و ۳ نمادهای استفاده شده در مورد شبکه و همچنین نمادهای مربوط به الگوریتم رقابت استعماری را نشان می‌دهند.

جدول ۲. نمادهای استفاده شده برای فرمول‌بندی شبکه

نماد	تعریف
TP	مجموعه نقاط هدف
PP	مجموعه نقاط بالقوه برای جایگذاری نودهای حسگر
tr	شعاع انتشار نودها
sr	شعاع حسی نودها
Cov(TP <sub>i</sub> )	مجموعه نودهایی که نقطه هدف TP <sub>i</sub> را پوشش می‌دهند
CovTP <sub>i</sub>	وضعیت پوشش TP <sub>i</sub>
X	تعداد عناصر مجموعه‌ای به نام X

همه کشورها از تعدادی نود حسگر که در محل‌های تقاطع شبکه گرید سه بعدی جایگذاری شده‌اند، تشکیل یافته‌اند. هر اندازه که تعداد این نودهای حسگر در یک کشور کمتر باشد از دیدگاه بهینه‌سازی آن کشور قدرتمندتر است. از این رو قدرت هر کشور در تکرار  $i$  ام الگوریتم توسط رابطه ۲ محاسبه می‌شود که همواره یک عدد در بازه  $[0, 1]$  است. همچنین قدرت کلی یک امپراطوری از مجموع قدرت تمامی کشورهای تشکیل دهنده آن به دست می‌آید که این موضوع برای تکرار  $i$  ام الگوریتم در رابطه ۳ نشان داده شده است.

$$P_i C_c E_e = \frac{1}{N_i C_c E_e} \quad (۲)$$

$$P_i E_e = \sum_{j=1}^{N_i E_e} P_j C_c E_e \quad (۳)$$

در انتهای هر دور به ترتیب میزان رشد تمام مستعمرات و تمام امپراطوری‌ها بر اساس روابط ۴ و ۵ محاسبه می‌شود. این مقدار برابر اختلاف قدرت یک مستعمره یا امپراطوری در تکرار  $i$  نسبت به تکرار  $i-1$  است که می‌تواند یک عدد مثبت، منفی و یا صفر باشد.

$$G_i C_c E_e = \begin{cases} P_i C_c E_e - P_{i-1} C_c E_e & \text{if } i > 1 \\ P_i C_c E_e - P_{ini} C_c E_e & \text{if } i = 1 \end{cases} \quad (۴)$$

$$G_i E_e = \begin{cases} P_i E_e - P_{i-1} E_e & \text{if } i > 1 \\ P_i E_e - P_{ini} E_e & \text{if } i = 1 \end{cases} \quad (۵)$$

در ادامه امید به مهاجرت یک مستعمره از یک امپراطوری بر اساس رابطه ۶ محاسبه می‌شود. در این معادله رشد یک امپراطوری با وزن  $1 - \alpha$  از قدر مطلق رشد یک مستعمره با وزن  $\alpha$  کم شده و امید به مهاجرت مستعمره به دست می‌آید. اگر رشد مستعمره و امپراطوری اعداد منفی شود، امید به مهاجرت برای مستعمره بیشتر می‌شود و اگر مثبت شود این امید کاهش می‌یابد. سپس بر اساس رابطه ۷ امید به مهاجرت برای تمام مستعمرات هر امپراطوری نرمالیزه می‌شود.

$$H_i C_c E_e = |\alpha \cdot G_i C_c E_e| - (1 - \alpha) \cdot G_i E_e \quad (۶)$$

$$H_{i-norm} C_c E_e = \frac{H_{i-ola} C_c E_e - \min\{H_j C_c E_e | 1 \leq j \leq N_i E_e\}}{\max\{H_j C_c E_e | 1 \leq j \leq N_i E_e\} - \min\{H_j C_c E_e | 1 \leq j \leq N_i E_e\}} \quad (۷)$$

تعریف ۳: اگر بین هر نود حسگر و چاهک حداقل یک مسیر وجود داشته باشد، آن نود حسگر را متصل شده می‌نامیم.

تعریف ۴: اگر در یک شبکه حسگر بی‌سیم، تمامی نودهای حسگر متصل شده باشند، آن شبکه را متصل می‌نامیم.

در یک شبکه حسگر بی‌سیم اگر در شعاع انتشار هر نود حسگر مانند  $S_i$  حداقل یک نود دیگر مانند  $S_j$  وجود داشته باشد که به چاهک نزدیکتر است آنگاه  $S_i$  می‌تواند از طریق  $S_j$  داده‌هایش را برای چاهک ارسال کند. به همین ترتیب  $S_j$  از طریق یک نود دیگر در شعاع انتشار خود مانند  $S_k$  که به چاهک نزدیکتر است می‌تواند داده‌هایش را برای چاهک ارسال کند. بنابراین اگر تمامی نودها حداقل یک نود نزدیکتر به چاهک در شعاع انتشار خود داشته باشند آنگاه کل شبکه متصل است. بدیهی است که نودهای همسایه چاهک به صورت مستقیم با چاهک ارتباط دارند و نیازی به نود واسطه ندارند. رابطه ۱۳ برای یک نود دلخواه  $S_i$  شرایط فرارگیری نودهای دیگر در مجموعه  $Con(S_i)$  را نشان می‌دهد و رابطه ۱۴ وضعیت یک نود را از نظر متصل بودن به چاهک بیان می‌کند.

$$Con(S_i) = \{S_j \mid distance(S_i, S_j) \leq tr \text{ and } distance(S_j, Sink) < distance(S_i, Sink), \forall j, 1 \leq j \leq ||S|| \text{ and } j \neq i\} \quad (13)$$

$$ConS_i = \begin{cases} 1, & \text{if } ||Con(S_i)|| \geq 1 \\ 1, & \text{if } distance(S_i, Sink) \leq tr \\ 0, & \text{if } ||Con(S_i)|| = 0 \end{cases} \quad (14)$$

بنابراین شرط متصل بودن کل شبکه به صورت معادله ۱۵ خواهد بود.

$$\sum_{i=1}^{||S||} ConS_i = ||S|| \quad (15)$$

#### ۴.۴ الگوریتم پیشنهادی

در این بخش الگوریتم ۱ برای حل مسائل پوشش و اتصال در شبکه حسگر بی‌سیم سه‌بعدی گرید پیشنهاد شده است. مبنای الگوریتم بر اساس الگوریتم رقابت استعماری است که امکان مهاجرت کشورها در انتهای هر دور به آن اضافه شده است. ورودی‌های الگوریتم عبارتند از: مجموعه نقاط هدف (TP)، مجموعه نقاط بالقوه برای فرارگیری حسگرها (PP)، نود چاهک و تعداد اولیه امپراطوری‌ها (emp). همچنین الگوریتم در انتها کشور استعمارگر مربوط به قوی‌ترین

Con(S <sub>i</sub> )	مجموعه نودهایی که در شعاع انتشار نود S <sub>i</sub> بوده و به چاهک نزدیکتر هستند
ConS <sub>i</sub>	وضعیت اتصال نود S <sub>i</sub>

جدول ۳. نمادهای استفاده شده برای فرمول بندی ICA

نماد	تعریف
E	مجموعه امپراطوری‌ها
CE <sub>e</sub>	مجموعه کشورهای امپراطوری E <sub>e</sub>
COL <sub>w</sub>	ضعیف ترین مستعمره از ضعیف ترین امپراطوری
Run	تعداد نسل در الگوریتم رقابت استعماری
Emp	تعداد امپراطوری‌ها
Prev	احتمال انقلاب
IMP-BEST	کشور استعمارگر در قوی ترین امپراطوری

#### ۳.۴ فرمول بندی مسئله

تعریف ۱: اگر یک نقطه هدف توسط حداقل یک نود حسگر پوشش داده شود، آن هدف را پوشش داده شده می‌نامیم.

تعریف ۲: اگر در یک شبکه حسگر بی‌سیم تمام نقاط هدف پوشش داده شده باشند، آن شبکه را یک شبکه حسگر پوشش داده شده می‌نامیم.

در معادله ۱۰ مجموعه  $Cov(TP_i)$  نشان‌دهنده نودهایی است که نقطه هدف  $TP_i$  را پوشش می‌دهند و در معادله ۱۱،  $CovTP_i$  وضعیت یک نقطه هدف را از نظر پوشش نشان می‌دهد.

$$distance(p_i, p_i) = \frac{1}{\sqrt{(x_{p_i} - x_{p_i})^2 + (y_{p_i} - y_{p_i})^2 + (z_{p_i} - z_{p_i})^2}} \quad (9)$$

$$Cov(TP_i) = \{S_j \mid distance(S_j, TP_i) \leq sr, \forall j, 1 \leq j \leq ||S||\} \quad (10)$$

$$CovTP_i = \begin{cases} 1, & \text{if } ||Cov(TP_i)|| \geq 1 \\ 0, & \text{if } ||Cov(TP_i)|| = 0 \end{cases} \quad (11)$$

همچنین رابطه ۱۲ شرط پوشش داده شدن کل یک شبکه را نشان می‌دهد.

$$\sum_{i=1}^{||TP||} CovTP_i = ||TP|| \quad (12)$$

---

**Algorithm ۱: Proposed approach based on improved ICA**

---

**Input:** TP, PP, sink, emp.

**Output:** IMP-BEST;

```

۱: generate initial countries which satisfy coverage and connectivity conditions based on Equations ۱۲ and ۱۵ respectively;
۲: create emp empires and put them in E;
۳: calculate the power of all countries;
۴: put emp most powerful countries as imperialists of empires  $E_1$  to  $E_{emp}$ ;
۵: distribute the remaining countries as colonies among  $E_1$  to  $E_{emp}$  using Roulette Wheel;
۶: for  $i = 1$  to run
۷:     imperialists converge their colonies with themselves;
۸:     similar empires are united together;
۹:     for each colony do
۱۰:         revolution with probability  $p_{rev}$ ;
۱۱:     end for each
۱۲:     for each colony do
۱۳:         if the power of the colony is higher than its relevant imperialist then
۱۴:             exchange positions of the colony and the imperialist;
۱۵:         end if
۱۶:     end for each
۱۷:     calculate the power of all empires in E, all imperialists and colonies;
۱۸:     determine the colony with minimum power from the empire with minimum power as the  $COL_w$ ;
۱۹:     recapture  $COL_w$  and assign it to another empire according to the power of the imperialists using Roulette Wheel;
    //immigration module
۲۰:     for  $e=1$  to emp
۲۱:         calculate  $P_i E_e$  using Equation ۳;
۲۲:         calculate  $G_i E_e$  using Equation ۵;
۲۳:         for  $c=1$  to  $\|CE_e\|$ 
۲۴:             if  $S_i C_e E_e = 0$  then
۲۵:                 calculate  $P_i C_e E_e$  using Equation ۶;
۲۶:                 calculate  $G_i C_e E_e$  using Equation ۴;
۲۷:                 calculate  $H_i C_e E_e$  using Equation ۷;
۲۸:                 normalize  $H_i C_e E_e$  using Equation ۹ as  $H_{i-norm} C_e E_e$ ;
۲۹:                 calculate  $I_i C_e E_e$  using Equation ۸;
۳۰:                 remove  $C_c$  from  $E_e$  with probability  $I_i C_e E_e$  and add to another empire using Roulette Wheel;
۳۱:             end if
۳۲:         end for
۳۳:     end for
۳۴:     destroy empires with no colony;
۳۵: end for
۳۶: return the imperialist of the most powerful empire as IMP-BEST;

```

---



شعاع ارتباطی نودها	۲۰ متر
شعاع حسی نودها	۲۰ متر
تعداد جمعیت اولیه در روش رقابت	۱۰۰
استعماری	

## ۵. شبیه‌سازی و ارزیابی کارایی

الگوریتم پیشنهادی در محیط نرم‌افزار MATLAB به صورت کامل پیاده‌سازی شده و در این بخش کارایی آن با جدیدترین رهیافت‌های موجود به نام‌های ACO-MCC<sup>3D</sup> [۱۴] و NDSBA [۳۰] مقایسه و بررسی می‌شود. در هر دو روش مورد مقایسه برای حل مسئله پوشش و اتصال در شبکه حسگر بی‌سیم از الگوریتم کلونی مورچه‌ها در دو مرحله کلی استفاده شده است. در فاز اول نودهای مورد نیاز برای حل مسئله تعیین شده و سپس در مرحله دوم سعی می‌شود نودهای حسگر افزوده تشخیص و حذف شوند. دو معیار اساسی برای مقایسه عبارت است از زمان اجرای الگوریتم و تعداد نودهای سنسور مورد نیاز برای پوشش تمام نقاط هدف و اتصال تمام سنسورها به گره چاهک. بدیهی است الگوریتمی که در زمان کمتری اجرا می‌شود و از تعداد نود کمتری استفاده کند دارای کارایی بالاتری است. تمامی شبیه‌سازی‌ها بر روی یک سیستم با پردازنده Intel i<sup>3</sup> processor, ۳,۳ GHz حافظه ۴ GB و سیستم‌عامل Windows ۷ اجرا شده است. محیط سه‌بعدی شبکه ۳۰۰×۳۰۰×۳۰۰ متر مکعب است که به صورت گرید تقسیم‌بندی شده است. نودهای حسگر فقط می‌توانند در نقاط تقاطع گرید قرار بگیرند در حالی که نقاط هدف و نود چاهک به صورت تصادفی در تمام فضای سه‌بعدی پخش می‌شوند. پارامترهای شبیه‌سازی در جدول ۴ نمایش داده شده و نتایج از میانگین ۵۰ بار اجرای شبیه‌سازی‌ها به دست آمده است.

### ۱,۵ تعداد نودهای مورد نیاز

در اولین آزمایش تعداد نقاط هدف پوشش برابر ۲۰، ۱۰۰ و ۲۰۰ در نظر گرفته شده و تعداد نودهای حسگر برای پوشش آنها بررسی شده است. جدول ۵ این تعداد حسگر را بر مبنای اندازه شبکه نشان می‌دهد. همانگونه که قابل انتظار است با افزایش ابعاد شبکه، به جهت افزایش پیچیدگی تعداد نودهای حسگر استفاده شده اندکی افزایش یافته است. نتایج ارائه شده، مربوط به میانگین ۵۰ بار اجرای الگوریتم هستند.

جدول ۴. لیست پارامترهای شبیه‌سازی

پارامتر	مقدار
ابعاد محیط	۳۰۰ متر در ۳۰۰ متر
مکان چاهک	تصادفی
تعداد نقاط هدف	۲۰ - ۲۰۰
تعداد نقاط بالقوه	۱۰۰۰۰۰ - ۱۰۰۰

جدول ۵. تعداد نودهای حسگر برای پوشش نقاط هدف در شبکه‌ها با ابعاد

متفاوت

TP	PP	۱۰×۱۰×۱۰	۲۰×۲۰×۲۰	۳۰×۳۰×۳۰	۴۰×۴۰×۴۰	۵۰×۵۰×۵۰
۲۰		۸/۱	۹/۶	۱۱/۷	۱۵/۸	۱۹/۱
۱۰۰		۱۰/۴	۱۲/۲	۱۵/۶	۲۱/۴	۲۳/۲
۲۰۰		۱۳/۷	۱۶/۱	۱۹/۵	۲۵/۶	۲۸/۴

در آزمایشی دیگر به جهت مقایسه‌پذیری با دو الگوریتم دیگر، تعداد نقاط گرید به صورت متغیر در نظر گرفته شده به صورتی که برای هر تعداد نقاط گرید، تعداد نقاط هدف مشخصی در شبکه وجود دارد. به عنوان مثال زمانی که تعداد نقاط گرید ۱۰×۱۰×۱۰ است، تعداد نقاط هدف ۲۰ عدد و زمانی که تعداد نقاط گرید ۲۰×۲۰×۲۰ است، ۴۰ نقطه هدف در نظر گرفته شده است. برای هر حالت تعداد نود حسگر مورد نیاز برای حل مسائل پوشش و اتصال در هر سه الگوریتم محاسبه شده و در شکل‌های ۱ و ۲ نمایش داده شده است. شکل ۱ مربوط به نتایج الگوریتم‌ها بعد از ۲۰ تکرار و شکل ۲ مربوط به ۴۰ تکرار است. همانطور که در شکل مشخص است روش پیشنهادی همواره نسبت به دو روش دیگر از تعداد نود کمتری استفاده می‌کند. دلیل این نتایج به استراتژی کلی الگوریتم‌ها مربوط است. دو الگوریتم ACO-MCC<sup>3D</sup> و NDSBA تلاش می‌کنند تا در مرحله اول یک جواب نسبتاً قابل قبول یافته و آنگاه در مرحله دوم آن را بهینه نمایند؛ اما بعد از مرحله دوم همچنان تعدادی نود افزوده باقی می‌مانند. در صورتی که الگوریتم پیشنهادی از همان ابتدا تلاش می‌کند بهترین جواب ممکن را بدون نود حسگر افزوده‌ای بیابد.

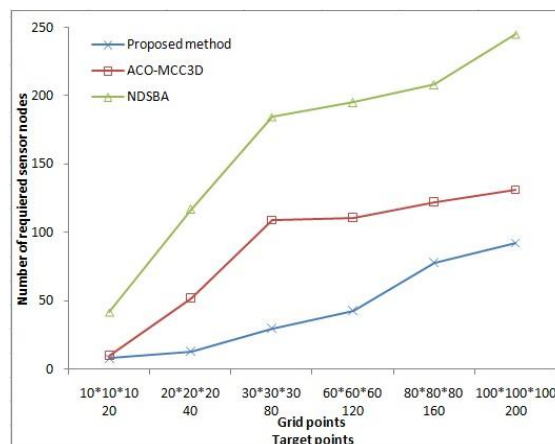
موقعیت‌هایی در محیط برای جایگذاری گره‌های حسگر است به گونه‌ای که دو شرط پوشش و اتصال با نود چاهک برآورده شود. بدیهی است با توجه به عدم اعمال محدودیت روی شناسایی موقعیت‌ها در این مرحله، تعداد زیادی موقعیت اضافی پیشنهاد می‌شود که آنها را موقعیت‌های افزونه می‌نامند. برای رسیدن به نتیجه بهتر نیاز است که این موقعیت‌ها شناسایی و حذف شوند. به همین علت هر دو روش ACO-MCC3D و NDSBA شامل یک فاز اضافی هستند که در آن مجدداً با اجرای الگوریتم کلونی مورچگان حتی المقدور سعی می‌شود که موقعیت‌های افزونه حذف شده و به جواب بهینه نزدیکتر شوند. در واقع در این روشها دو بار الگوریتم کلونی مورچگان اجرا می‌شود و همین امر سبب افزایش زمان اجرای این دو روش است. از آنجا که الگوریتم پیشنهادی تنها در یک مرحله انجام می‌گیرد زمان اجرای آن به صورت قابل ملاحظه‌ای کمتر است.

جدول ۶. مقایسه زمان اجرای الگوریتم پیشنهادی با سایر الگوریتم‌ها

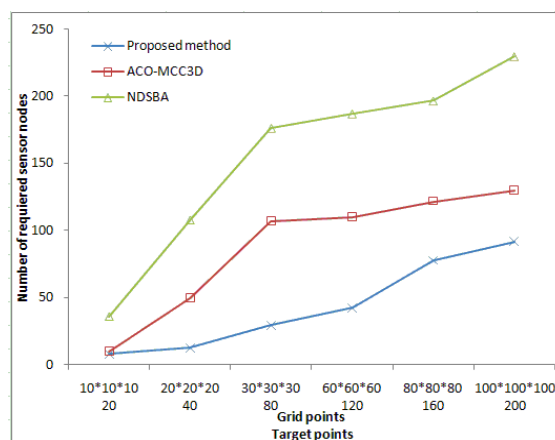
روش پیشنهادی	ACO-MCC3D	NDSBA	الگوریتم
۶	۵۰۱۱	۴۸۹۷	۲۰×۲۰×۲۰ ۴۰
۹۵۵	۱۱۴۸۵	۱۱۴۸۱	۱۰۰×۱۰۰×۱۰۰ ۲۰۰

## ۶. جمع‌بندی و نتیجه‌گیری

در این مقاله مساله جایگذاری نودها در شبکه‌های حسگر بی‌سیم گرید و سه‌بعدی مورد مطالعه قرار گرفت. در این نوع شبکه‌ها، حسگرها تنها می‌توانند در نقاط تقاطع گریدها قرار داده شوند؛ این در حالی است که نقاط هدف به صورت تصادفی در سرتاسر شبکه توزیع می‌شوند. هر نقطه هدف باید در شعاع پوششی حداقل یک حسگر قرار گیرد تا پوشش یافته نامیده شود. اگر همه نقاط هدف پوشش یافته باشند آنگاه کل شبکه پوشش یافته است. همچنین هر حسگر باید حداقل یک مسیر تا رسیدن به نود چاهک در اختیار داشته باشد تا شبکه را متصل بنامیم. الگوریتم پیشنهادی هر دو شرط پوشش و اتصال را برآورده کرده است. لازم به ذکر است که نود چاهک به صورت تصادفی در شبکه قرار داده شده است.



شکل ۱. مقایسه الگوریتم پیشنهادی با سایر الگوریتم‌ها در تکرار ۲۰



شکل ۲. مقایسه الگوریتم پیشنهادی با سایر الگوریتم‌ها در تکرار ۴۰

همانگونه که از شکل‌های ۱ و ۲ پیداست دو الگوریتم پیشنهادی و ACO-MCC3D در تکرار ۲۰ ام به نتیجه نهایی خود نزدیک شده و بنابراین نتایج آنها برای تکرارهای ۲۰ و ۴۰ تقریباً مشابه می‌باشد؛ این در حالی است که الگوریتم NDSBA همچنان نیازمند تکرارهای بیشتری می‌باشد تا به بهترین نتیجه خود برسد.

## ۲.۵ زمان اجرا

مقایسه زمان اجرای سه الگوریتم در جدول ۶ بر حسب ثانیه نشان داده شده است. زمان اجرا برای دو حالت محاسبه شده است: شبکه گرید ۲۰×۲۰×۲۰ با ۴۰ نقطه هدف و شبکه گرید ۱۰۰×۱۰۰×۱۰۰ با ۲۰۰ نقطه هدف. با توجه به جدول مشخص است که الگوریتم پیشنهادی در زمان بسیار کمتری نسبت به دو الگوریتم دیگر اجرا می‌شود. هر دو روش ACO-MCC3D و NDSBA الگوریتم کلونی مورچگان را در دو فاز به کار می‌گیرند. در فاز اول هدف اصلی انتخاب

[۵] F. Delavernhe et al., "Robust scheduling for target tracking using wireless sensor networks," *Computers & Operations Research*, vol. ۱۱۶, pp. ۱-۱۴, ۲۰۲۰.

[۶] A. Sangwan and R. P. Singh, "Survey on coverage problems in wireless sensor networks," *Wireless Personal Communications*, vol. ۸۰, no. ۴, pp. ۱۴۷۵-۱۵۰۰, ۲۰۱۵.

[۷] A. Tripathi et al., "Coverage and connectivity in WSNs: a survey, research issues and challenges," *IEEE Access*, vol. ۶, pp. ۲۶۹۷۱-۲۶۹۹۲, ۲۰۱۸.

[۸] I. Khoufi, P. Minet, A. Laouti, S. Mahfoudh, "Survey of deployment algorithms in wireless sensor networks: coverage and connectivity issues and challenges," *International Journal of Autonomous and Adaptive Communications Systems*, vol. ۱۰, no. ۱۴, pp. ۳۴۱-۳۹۰, ۲۰۱۷.

[۹] Y. Wang et al., "Coverage problem with uncertain properties in wireless sensor networks: A survey," *Computer Networks*, vol. ۱۲۳, pp. ۲۰۰-۲۳۲, ۲۰۱۷.

[۱۰] B. Wang, "Coverage problems in sensor networks: a survey," *ACM Computing Surveys (CSUR)*, vol. ۴۳, no. ۴, p. ۳۲, ۲۰۱۱.

[۱۱] B. Peng and L. Li, "An improved localization algorithm based on genetic algorithm in wireless sensor networks," *Cognitive Neurodynamics*, ISSN ۱۸۷۱-۴۰۸۰, vol. ۹, Issue ۲, pp. ۲۴۹-۲۵۶, ۲۰۱۵.

[۱۲] G. Gajalakshmi and G. U. Srikanth, "A survey on the utilization of Ant Colony Optimization (ACO) algorithm in WSN," *International Conference on Information Communication and Embedded Systems (ICICES)*, ۲۰۱۶.

[۱۳] H. Sheikhi and W. Barkhoda, "Solving the k-Coverage and m-Connected Problem in Wireless Sensor Networks through the Imperialist Competitive Algorithm," *Journal of Interconnection Networks*, vol. ۲۰, no. ۱, pp. ۱-۱۸, ۲۰۲۰.

[۱۴] T. Qasim et al., "An ant colony optimization based approach for minimum cost coverage on ۳-D grid in wireless sensor networks," *IEEE Communications Letters*, vol. ۲۲, no. ۶, pp. ۱۱۴۰-۱۱۴۳, ۲۰۱۸.

برای حل این مسئله، الگوریتم رقابت استعماری بهبود یافته پیشنهاد شده است. الگوریتم ICA یکی از الگوریتم‌های تکاملی است که از جوامع بشری الهام گرفته است. در این مقاله برای افزایش تنوع و جلوگیری از همگرا شدن به بهینه‌های محلی، ایده مهاجرت پیشنهاد شده است. بدین منظور در هر مرحله برای تمامی کشورهای مستعمره امید به مهاجرت و احتمال مهاجرت محاسبه می‌شود که در واقع نمایانگر میزان رشد کشور در طی مراحل تکامل الگوریتم است. با توجه به احتمال مهاجرت، هر کشور می‌تواند از یک امپراطوری به امپراطوری دیگر برود.

نتایج شبیه‌سازی کارایی بالای الگوریتم پیشنهادی را نسبت به بهترین الگوریتم‌های موجود نشان می‌دهد. با توجه به این واقعیت که الگوریتم پیشنهادی در یک مرحله بهترین نتایج را به دست می‌آورد، زمان اجرای کمتری نسبت به سایر الگوریتم‌ها دارد که معمولاً در دو مرحله نتایج را یافته و سپس سعی می‌کنند تا آنها را بهبود ببخشند. همچنین تعداد نودهای انتخاب شده برای برآورده کردن پوشش و اتصال در الگوریتم پیشنهادی به صورت قابل ملاحظه‌ای کمتر از الگوریتم‌های دیگر است.

## مراجع

[۱] L. Sathyapriya and A. Jawahar, "Clustering Algorithms for Wireless Sensor Networks Survey," *Sensor Letters*, vol. ۱۸, no. ۲, pp. ۱۴۳-۱۴۹, ۲۰۲۰.

[۲] K. Laubhan et al., "A low-power IoT framework: from sensors to the cloud," *IEEE International Conference on Electro Information Technology (EIT)*, Grand Forks, ND, USA, ۲۰۱۶.

[۳] X. Su et al., "A Review of Underwater Localization Techniques, Algorithms, and Challenges," *Journal of Sensors*, vol. ۲۰۲۰, pp. ۱-۲۴, ۲۰۲۰.

[۴] S. Fattah et al., "A Survey on Underwater Wireless Sensor Networks: Requirements, Taxonomy, Recent Advances, and Open Research Challenges," *Sensors*, vol. ۲۰, no. ۱۸, pp. ۱-۳۰, ۲۰۲۰.

- Network and Computer Applications*, vol. ۳۹, pp. ۳۱۰-۳۱۸, ۲۰۱۴.
- [۲۵] T. Qasim et al., "ACO-Discreet: An efficient node deployment approach in wireless sensor networks," in *Information Technology-New Generations Springer*, pp. ۴۳-۴۸, ۲۰۱۸.
- [۲۶] H. Mostafaei, M. Shojafar, B. Zaher, M. Singhal, "Barrier coverage of WSNs with the imperialist competitive algorithm," *the Journal of Supercomputing*, vol. ۲۳, no. ۱۱, pp. ۴۹۵۷-۴۹۸۰, ۲۰۱۷.
- [۲۷] R. Enayatifar, M. Yousefi, A. H. Abdullah, A. N. Darus, "A novel sensor deployment approach using multi-objective imperialist competitive algorithm in wireless sensor networks," *Arabian Journal for Science and Engineering*, vol. ۳۹, no. ۶, pp. ۴۶۳۷-۴۶۵۰, ۲۰۱۴.
- [۲۸] E. A. Gargari and C. Lucas, "Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition," *IEEE Congress on Evolutionary Computation*, Singapore, pp. ۲۵-۲۷, ۲۰۰۷.
- [۲۹] S. N. Shirkouhi et al., "Solving the integrated product mix-outsourcing problem using the Imperialist Competitive Algorithm," *Expert System With Applications*, vol. ۳۷, no. ۱۲, pp. ۷۶۱۵-۷۶۲۶, ۲۰۱۰.
- [۳۰] G. Huang, D. Chen, and X. Liu, "A node deployment strategy for blindness avoiding in wireless sensor networks," *IEEE Communications Letters*, vol. ۱۹, no. ۶, pp. ۱۰۰۵-۱۰۰۸, ۲۰۱۵.
- [۳۱] A. Chelli et al., "One-Step approach for two-tiered constrained relay node placement in wireless sensor networks," *IEEE Wireless Communications Letters*, vol. ۵, no. ۴, pp. ۴۴۸-۴۵۱, ۲۰۱۶.
- [۳۲] M. Bagaa et al., "Optimal placement of relay nodes over limited positions in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. ۱۶, no. ۴, pp. ۲۲۰۵-۲۲۱۹, ۲۰۱۷.
- [۳۳] S. K. Gupta, P. Kuila, and P. K. Jana, "Genetic algorithm approach for k-coverage and m-connected node placement in target based wireless sensor networks," *Computers & Electrical engineering*, vol. ۵۶, pp. ۵۴۴-۵۵۶, ۲۰۱۶.
- [۳۴] G. P. Gupta and S. Jha, "Biogeography-based optimization scheme for solving the coverage and connected node placement problem for wireless sensor networks," *Wireless Networks*, vol. ۲۵, no. ۶, pp. ۳۱۶۷-۳۱۷۷, ۲۰۱۹.
- [۳۵] X. Han, X. Cao, E. L. Lioyd, C. C. Shen, "Fault-Tolerant relay node placement in heterogeneous wireless sensor networks," *IEEE TRANSACTIONS ON MOBILE COMPUTING*, vol. ۹, no. ۵, pp. ۶۴۳-۶۵۶, ۲۰۰۹.
- [۳۶] R. Jena, "Artificial bee colony algorithm based multi-objective node placement for wireless sensor network," *International Journal of Information Technology and Computer Science*, vol. ۶, no. ۶, ۲۰۱۴.
- [۳۷] H. Huang, J. Zhang, R. Wang, and Y. Qian, "Sensor node deployment in wireless sensor networks based on ionic bond-directed particle swarm optimization," *Appl. Math*, vol. ۸, no. ۲, pp. ۵۹۷-۶۰۵, ۲۰۱۴.
- [۳۸] D. Li et al., "EasiDesign: an improved ant colony algorithm for sensor deployment in real sensor network system," in *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. ۱-۵, ۲۰۱۰.
- [۳۹] X. Liu, "Sensor deployment of wireless sensor networks based on ant colony optimization with three classes of ant transitions," *IEEE Communications Letters*, vol. ۱۶, no. ۱۰, pp. ۱۶۰۴-۱۶۰۷, ۲۰۱۲.
- [۴۰] X. Liu and D. He, "Ant colony optimization with greedy migration mechanism for node deployment in wireless sensor networks," *Journal of*



## **The role of organizational factors on consumer buying behavior in social networks**

Zohreh Dehdashti<sup>1\*</sup>, Mitra Daneshparvar<sup>\*\*</sup>, Vahid Nasehifar<sup>\*\*</sup>, Vahid Khashei<sup>\*\*</sup>

### **Abstract:**

The clothing industry has seen great growth in social media in recent years. Many factors affect the purchase of clothing from social networks, and one of these important and influential factors is organizational factors. Therefore, the purpose of this study is to investigate the organizational factors affecting the consumer in shopping through social networks. In this research, a combined method has been used. In the qualitative section, by reviewing the literature and interviewing vendors active in social networks, texts were prepared and coded. In the quantitative part, based on the initial model, a questionnaire was developed and distributed to 385 clothing buyers in networks. The structural equation method was used to analyze the data and the Sobel test was used to examine the mediating role. This study examines the independent and combined effect of organizational variables affecting people's trust in buying clothes through social networks. The results showed that information quality, transaction security, company reputation and company location have a significant effect on people's trust and consumer willingness to buy clothing through social networks.

**Keywords:** Social networks, consumer shopping behavior, clothing industry, organizational factors.

---

<sup>1</sup> Corresponding author(daneshparvar.basir@yahoo.com)

## **A Combination Method of DEA, DEMATEL and ANP for Evaluation of ERP Systems**

\*Amir Amini<sup>1</sup>    \*\* A.Reza Alinezhad

### **Abstract:**

In this study Data envelopment analysis is used for the assessment of enterprise resources planning systems of 18 manufacturing companies, to determine whether the defined goals for ERP systems have been able to affect the performance after the implementation of the system. Considering the identification of effective factors in implementing ERP systems and using the previous researches, the performance evaluation criteria of this system were identified. Then, using experts' views the most important input indicators were ranked by fuzzy DEMATEL and output indicators by ANP. In this ranking, time spent on implementation, the implementation infrastructure, training and user support were identified as top input indicators, and three indicators of productivity increase, proper resource management and user satisfaction were selected as top output indicators. Using selected indicators, the performance of ERP systems of selected companies was evaluated. The results of this research will be useful in identifying the strengths and weaknesses of companies compared to top ones and making their ERP system even better.

**Keywords:** ERP, DEMATEL, ANP, DEA, Efficiency Evaluation

---

<sup>1</sup> Corresponding author(A.Amini@ine.uut.ac.ir)

## **Joint Power Allocation Optimization of cooperative communication systems with Non-Orthogonal Multiple Access**

\*Hamid Amiri Ara

\*\* Mohannad bagher Masror

\*\*\* Mohammad Reaz zahabi<sup>1</sup>

### **Abstract:**

In this paper, a downlink network with two users as transmitter and relay, respectively, and a central station as a receiver is considered. The aim is to determine the optimal coefficient of non-orthogonal signal symbols and the optimal power allocation in the source-relay in order to maximize the average total rate in a cooperative communication system using the non-orthogonal multiple access technique. To achieve these goals, the average total rate of the cooperative communication system with non-orthogonal multiple access with decode and forward relay in the independent Rayleigh channel was calculated. Then, in the first step, the optimization problem of the non-orthogonal symbols coefficient is mathematically expressed for each power allocation and a closed form solution is proposed. In the second step, the power allocation optimization for the source-relay was introduced and solved. Finally, the joint optimization problem of the non-orthogonal symbols coefficient and power allocation is investigated and an algorithm proposed for the joint optimization problem. The proposed algorithm shows that the joint optimization of the non-orthogonal symbols coefficient and power allocation achieve a higher average rate than the separate optimization of each of these parameters. Also, simulations and numerical results are presented to confirm the theoretical equation, where the simulations show the 3 dB gain for the optimized system using the proposed algorithm compared to the non-optimized system.

**Keywords:** Joint Power Allocation Optimization of cooperative communication systems with Non-Orthogonal Multiple Access.

---

<sup>1</sup> Corresponding author(zahabi@nit.ac.ir)



## **An access control model for online social networks using user-to-user relationships**

\*MohamadJavad Piran<sup>1</sup> \*\*A. ghodselahi \*\*\*Mahmood Deypir

### **Abstract:**

With the pervasiveness of social networks and the growing information shared on them, users of these networks are exposed to potential threats to data security and privacy. The privacy settings included in these networks do not give users complete control over the management and privatization of access to their shared information by other users. In this article, using the concept of social graph, a new model of user access control was proposed to the user, which allows the expression of privacy policies and more accurate and professional access control in terms of pattern and depth of relationships between users in social networks. In this article, by using the regular index method, indirect relationships among users are examined and analyzed, and more precise policies than previous models are presented. The evaluation of the results showed that for 10 neighbors for each user, the probability accumulation of a qualified path for the first three counter loops is 1, 10.5 and 67.3%, respectively, and finally for the fourth counter it reaches 100%. As the defined counting characteristic increases, the average execution time of the proposed algorithm and previously proposed algorithms increases. However, for the higher limits of the counting characteristic, the proposed algorithm performs better than the previous ones.

**Keywords:** Social networks, network security, privacy, counter loop, data security.

---

<sup>1</sup>. Corresponding author (mohamadjavadpiran@gmail.com)

## Proposing an Information Retrieval Model Using Interval Numbers

\*Hooman Tahayori<sup>1</sup>    \*\*Farzad Ghahramani

### Abstract:

Recent expansions of web demands for more capable information retrieval systems that more accurately address the users' information needs. Weighting the words and terms in documents plays an important role in any information retrieval system. Various methods for weighting the words are proposed, however, it is not straightforward to assert which one is more effective than the others. In this paper, we have proposed a method that calculates the weights of the terms in documents and queries as interval numbers. The interval numbers are derived by aggregating the crisp weights that are calculated by exploiting the existing weighting methods. The proposed method, calculates an interval number as the overall relevancy of each document with the given query. We have discussed three approaches for ranking the interval relevancy numbers. In the experiments we have conducted on Cranfield and Medline datasets, we have studied the effects of weight normalization, use of variations of term and document frequency and have shown that appropriate selection of basic term weighting methods in conjunction with their aggregation into an interval number would considerably improve the information retrieval performance. Through appropriate selection of basic weighting methods we have reached the MAP of 0.43323 and 0.54580 on the datasets, respectively. Obtained results show that the proposed method, outperforms the use of any single basic weighting method and other existing complicated weighting methods.

**Keywords:** Textual information retrieval, Document ranking, Term weighting, Interval numbers, Interval weight

---

1 . Corresponding author (tahayori@shirazu.ac.ir)

## Improving Efficiency of Finding Frequent Subgraphs in Graph Stream Using gMatrix Summarization

\* Masoud Kazemi    \*\*Hossein Khasteh<sup>1</sup>    \*\*\*H.Reza Rokhsati

### Abstract:

In many real-world frameworks, dealing with huge domains of nodes and online streaming edges are unavoidable. Transportation systems, IP networks and developed social medias are quintessential examples of such scenarios. One of the most important open problems while dealing with massive graph streams are finding frequent sub-graph. There are some approaches such as count-min for storing the frequent nodes, however performing these methods will result in inaccurate modelling of structures based on the main graph. Having said that, gMatrix is one of the recently developed approaches which can fairly save the important properties of the main graph. In this approach, different hash functions are utilized to store the basis of streams in the main graph. As a result, having the reverse of the hash functions will be extremely useful in calculation of the frequent subgraph. Though gMatrix mainly suffer from two problems. First, they are not really accurate due to high compression rate of the main graph and second, the complexity of returning a query is high. In this thesis, we have presented a new approach based on gMatrix which can reduce the amount of memory usage as well as returning the queries in less amount of time. The main contribution of the introduced approach is to reduce the dependency among the hash functions. This will result in less conflicts while creating the gMatrix later. In this study we have used Cosine Similarity in order to estimate the amount of dependency and similarity among hash functions. Our experimental results prove the higher performance in terms of algorithm and time complexity.

**Keywords:** Graph Stream, Sketch Based Summarization, gMatrix, Hash Functions, Cosine Similarity

---

1 . Corresponding author (tahayori@shirazu.ac.ir)

## **Increasing the lifetime of underwater acoustic sensor networks by optimal relay node placement**

\* Zahra Mohammadi

\*\*Mohadeseh Soleimanpour Moghadam<sup>1</sup>

\*\*\* Daryoush Abbasi Moghaddam

\*\*\*Siamak Talebi

### **Abstract:**

Underwater acoustic sensor networks (UASNs) have gained growing importance due to their desirable features and wide spread practical applications in many communication fields. Due to the high cost of underwater sensor nodes as well as implementation complexity, increasing the lifetime of UASNs is an important issue. Although relay nodes have an important role in reducing the transmission distance and energy consumption. But the efficient RNP (Relay Node Placement) to avoid the critical sensor nodes' elimination is the main problem, i.e., to preserve the connected network. For this aim this paper presents an innovative solution called an Efficient Relay node Setting (ERS) algorithm, which involves formulating the Relay Node Placement (RNP) as a non-convex optimization problem. Actually, due to the Difference Convex (DC) constraints the proposed RNP problem is a non-convex problem and finding an optimal solution is complicated. However, a novel transformation can be applied to DC constraints which converts the problem into its convex programming equivalent. Application of the convex programming offers the advantage of readily computing a global optimal solution. Simulation results confirm the superiority of the proposed scheme over the competing RA method in terms of network lifetime and efficiency.

**Keywords:** Underwater acoustic sensor networks, Relay nodes, Network lifetime, Multi-hop communication, Convex programming.

---

1 . Corresponding author (tahayori@shirazu.ac.ir)

## **Identification of the Factors and Indicators of the Model of Quality Assessment of the Game Industry Services on the Cloud Platform**

\*Mohammad Taghi Taghavifard<sup>1</sup>

\*\*\*Kamran Feyzi

\*\*Mohammad Ali Abbaspour

\*\*\*Rasoul Nooralsena

### **Abstract:**

In today's world, one of the most important media and entertainment sectors is the gaming industry. The number of users of this type of games is constantly increasing and Businesses in this industry have grown dramatically. This has led to a dollars million turnover boost for developer companies of this kind of games. Therefore, the main aim of this study is to provide a model for measuring the quality of services on the cloud platform in the gaming industry. The present research is a developmental-applied objective and is a part of qualitative research. This research is also a survey research in terms of data gathering. In this study, 52 indicators were identified in 9 main dimensions after evaluating theoretical literature of the research and existing models and Delphi method for assessing the quality of the cloud computing services industry. The results of the research provide a comprehensive model for assessing the quality of cloud computing services industry services that can be used by industry and researchers.

**Keywords:** Service Quality, Service Quality Measurement Models, Cloud Based Services, Computer Games Industry.

---

<sup>1</sup> Corresponding author (dr.taghavifard@gmail.com)

## Measuring Similarity for Directed Path in Geometric Data

\*Zeynab Saeidi

\*\*Mohammad Farshi<sup>1</sup>

### Abstract:

We consider the following similarity problem concerning the Fréchet distance. A directed path  $\pi$  is given as input and a horizontal segment  $Q$  is defined at query time by the user. Our goal is to preprocess and save the directed path  $\pi$  into a data structure such that based on the information saved in the data structure, one sub-path of the directed path can be reported which Fréchet distance between the sub-path and the horizontal query segment  $Q$  is minimum between all possible sub-paths. To the best of our knowledge, no theoretical results have been reported for this problem. In this paper, the first heuristic algorithm is proposed. We only experimentally show the quality of the algorithm in several datasets due to no existing algorithm.

**Keywords:** Data structure, Fréchet distance, Hausdorff distance, Similarity, Directed Path.

---

<sup>1</sup> Corresponding author (mfarshi@yazd.ac.ir)

## **Investigating the Relationship between Electronic Human Resource Management, Agility and Organizational Effectiveness in Organizations Under the Ministry of Communications and Information Technology**

\*Abdolreza Bahadori Fard<sup>1</sup>

\*\*Reza Rasouli

### **Abstract:**

For years, human resource management was limited to administrative and day-to-day activities such as recruitment, selection, and training that devoted all the energy and time of the organization's human resources professionals. The human resources departments were often so involved in such activities that they no longer had time for activities that add value to the organization, such as knowledge management, organizational culture management, and strategic review and reconstruction of the organization. Today we need better, faster and smarter human resource management solutions to reduce administrative costs and increase the speed of service delivery. Electronic human resource management is the answer to today's human resource management challenges. The emergence of this concept has changed the goals, performance and effectiveness of the human resources field. With the advent of information technology, the role, position and even tasks of human resource management in the organization underwent major changes. Human resource professionals, who have previously performed most of the second-tier organizational tasks and non-value-added activities, have become strategic partners of the organization with the widespread use of information technology in the field of human resources. The statistical population of the study was all experts and managers of organizations under the Ministry of Communications and Information Technology. The sample size was determined to be 384. This research can be considered as a field research that uses questionnaire and interview tools. For this purpose, a combination of interviews, questionnaires and documents are used. Also, in order to identify the indicators and components of electronic human resource management in the studied organizations in the qualitative section, interview tools and questionnaires and focus groups based on the initial research model in the data method of the foundation are used. In the short section after this stage, using the ranking method, the priority of these indicators is determined using a pairwise comparison questionnaire, and finally, in order to confirm the overall fit of the model, the questionnaire of determining the status of indicators is used. . The validity of these questionnaires and interviews is determined by CVR method and the reliability is determined by Cronbach's alpha method.

**Keywords:** Human resources, electronic human resource management, information technology, agility, effectiveness.

---

<sup>1</sup> Corresponding author (rzbaha@yahoo.com)

## **A Novel Method based on the Cocomo model to increase the accuracy of software projects effort estimates**

\*Mahdiah Salari<sup>1</sup>

\*\*Vahid Khatibi

\*\*\*Amid Khatibi

### **Abstract:**

It is regarded as a crucial task in a software project to estimate the criteria, and effort estimation in the primary stages of software development is thus one of the most important challenges involved in management of software projects. Incorrect estimation can lead the project to failure. It is therefore a major task in efficient development of software projects to estimate software costs accurately. Therefore, two methods were presented in this research for effort estimation in software projects, where attempts were made to provide a way to increase accuracy through analysis of stimuli and application of metaheuristic algorithms in combination with neural networks. The first method examined the effect of the cuckoo search algorithm in optimization of the estimation coefficients in the COCOMO model, and the second method was presented as a combination of neural networks and the cuckoo search optimization algorithm to increase the accuracy of effort estimation in software development. The results obtained on two real-world datasets demonstrated the proper efficiency of the proposed methods as compared to that of similar methods.

**Keywords:** Cocomo, Cost estimation, Cuckoo algorithm, neural network.

---

<sup>1</sup> Corresponding author (rzbaha@yahoo.com)



## **Improving imperialist competitive algorithm for solving the nodes placement problem in three-dimensional grid wireless sensor networks**

\*Vafa Barkhoda

\*\*Hemmat Sheikhi<sup>1</sup>

\*\*\*Suodabeh Mohamadi

### **Abstract:**

One of the basic and important research fields in wireless sensor networks is how to place sensor nodes where by using minimum number of sensor nodes all target points are covered and all sensor nodes are connected to the sink. In this paper, a novel method based on imperialist competitive algorithm is used for solving the mentioned problem. In the proposed method, a colony can immigrate from a weak empire to more powerful empire. The idea of immigration is inspired from human society in which a human can emigrate from a country to another country. The network is supposed to be a three-dimensional grid network and the sensor nodes can be only placed at cross-points of the grids while the target points can be deployed at each point of three-dimensional space. The simulation results show that the proposed method uses fewer number of sensor nodes than other similar algorithms and has the less running time.

**Keywords:** Wireless sensor networks, Three-dimensional grid network, Imperialist competitive algorithm, Immigration, Node placement.

---

<sup>1</sup> Corresponding author (h.sheikhi@kut.ac.ir)



---

# Contents

**Iranian Journal of  
Information Technology & Communication**  
No. 45-46, Vol.12, September 2020- March 2021

---

- **The role of organizational factors on consumer buying behavior in social networks** 1  
Zohreh Dehdashti, Mitra Daneshparvar, Vahid Nasehifar, Vahid Khashei
  - **A Combination Method of DEA, DEMATEL and ANP for Evaluation of ERP Systems** 21  
Amir Amini, A.Reza Alinezhad
  - **Joint Power Allocation Optimization of cooperative communication systems with Non-Orthogonal Multiple Access** 43  
Hamid Amiri Ara, Mohammad bagher Masror, Mohammad Reaz zahabi
  - **An access control model for online social networks using user-to-user relationships** 53  
MohamadJavad Piran, A. ghodselahi, Mahmood Deypir
  - **Proposing an Information Retrieval Model Using Interval Numbers** 71  
Hooman Tahayori, Farzad Ghahramani
  - **Improving Efficiency of Finding Frequent Subgraphs in Graph Stream Using gMatrix Summarization** 95  
Masoud Kazemi, Hossein Khasteh, H.Reza Rokhsati
  - **Increasing the lifetime of underwater acoustic sensor networks by optimal relay node placement** 115  
Zahra Mohammadi, Mohadeseh Soleimanpour Moghadam, Daryoush Abbasi Moghaddam, Siamak Talebi
  - **Identification of the Factors and Indicators of the Model of Quality Assessment of the Game Industry Services on the Cloud Platform** 130  
Mohammad Taghi Taghavifard, Mohammad Ali Abbaspour, Kamran Feyzi, Rasoul Nooralsena
  - **Measuring Similarity for Directed Path in Geometric Data** 145  
Zeynab Saeidi, Mohammad Farshi
  - **Investigating the Relationship between Electronic Human Resource Management, Agility and Organizational Effectiveness in Organizations Under the Ministry of Communications and Information Technology** 159  
Abdolreza Bahadori Fard, Reza Rasouli
  - **A Novel Method based on the Cocomo model to increase the accuracy of software projects effort estimates** 185  
Mahdieh Salari, Vahid Khatibi, Amid Khatibi
  - **Improving imperialist competitive algorithm for solving the nodes placement problem in three-dimensional grid wireless sensor networks** 201  
Vafa Barkhoda, Hemmat Sheikhi, Suodabeh Mohamadi
  - **Abstracts in English** 213
-