# شناسایی کاربران تأثیرگذار شبکه اجتماعی توییتر: اندازه‌گیری تأثیر چندبُعدی

**منیره حسینی**[*1]
*دانشیار، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران*

**کوثر حیدری**
*دانشجوی کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران*

**مدیریت اطلاعات**
**دوره ۷، شماره ۲**
**پاییز و زمستان ۱۴۰۰**

**چکیده:** با گسترش و استفاده روزافزون از شبکه‌های اجتماعی، به محیط این شبکه‌ها به‌دلیل امکان دسترسی به تعداد زیادی از کاربران و تأثیراتی که روی افکار و اعتقادات آنها می‌گذارد، توجه شده است. با توجه به تعداد زیاد کاربرانی که در یک شبکه اجتماعی حضور دارند، شناسایی کاربران تأثیرگذار برای اهداف مختلفی مانند تبلیغات به‌منظور گسترش بیشتر اطلاعات یا مهار شیوع افکار منفی عمومی در شبکه، ضروری است. تاکنون برای شناسایی کاربران تأثیرگذار روش‌های مختلفی ارائه شده است که اغلب، از طریق معیارهای یک‌بعدی و ساده، کاربران تأثیرگذار را شناسایی کرده‌اند. در پژوهش حاضر، معیارهای مختلفی از سه بعد محتوا، فعالیت و شبکه ترکیب شده‌اند و رویکرد اندازه‌گیری جدیدی برای شناسایی کاربران تأثیرگذار در شبکه اجتماعی توییتر ارائه شده است. به کمک روش‌های اعتبارسنجی گسترش، دقت، فراخوانی، F1-Measure و درصد تکرار، نشان داده شده است که روش این پژوهش، از کارهای شناخته‌شده در این حوزه، عملکرد بهتری دارد.

**کلیدواژه‌ها**: کاربران تأثیرگذار، شبکه‌های اجتماعی، شناسایی کاربران تأثیرگذار، توییتر.

۱. نویسنده مسئول: hosseini@kntu.ac.ir

## مقدمه

شبکه‌های اجتماعی آنلاین، بستری فراهم می‌کنند که در آن افراد مختلف، می‌توانند با یکدیگر تعامل کنند، نظرهای خود را بیان کنند و به اشتراک بگذارند (Jain & Sinha, 2020 ). رشد رسانه‌های اجتماعی، نحوه تعامل، برقراری ارتباط و درگیر کردن افراد را کاملاً بازسازی کرده است و در تسهیل دسترسی و نفوذ بیشتر، نقشی کلیدی دارند (Arora et al., 2019). افراد تأثیرگذار[1]، اعضایی هستند که در شبکه اجتماعی آنلاین، در مقایسه با یک عضو عادی، تأثیر بیشتری بر دیگران دارند ( Ríos et al., 2019). شناسایی کاربران تأثیرگذار[2]، می‌تواند در بسیاری از شرایط، ارزش‌های اجتماعی و اقتصادی شایان توجهی به ارمغان آورد؛ برای مثال، شناسایی کاربران تأثیرگذار در شرایطی نظیر بازاریابی ویروسی، تجزیه و تحلیل برند، انتشار اخبار، گسترش آگاهی سلامت، تبلیغ جنبش‌های سیاسی و رهبران افکار برای توانمندسازی حکومت، کاربرد فراوان دارد (Jain & Sinha, 2020; Lü et al., 2016). از این رو، ارائه روشی دقیق برای تعیین میزان نفوذ کاربران شبکه، به یکی از موضوع‌های داغ تبدیل شده است که دانشمندان در شاخه‌های مختلف به آن پرداخته‌اند (Zareie et al., 2019).

برای شناسایی کاربران تأثیرگذار، تلاش زیادی انجام شده است. با این حال، هنوز یک معیار جهانی اندازه‌گیری تأثیر وجود ندارد (Zhuang et al., 2021). کشف کاربران تأثیرگذار مناسب، از داده‌های گسترده شبکه‌های اجتماعی آنلاین، کاری دشوار است، زیرا ویژگی‌های زیادی وجود دارد که به تأثیر کاربر مربوط می‌شود (Sun et al., 2016). علاقه پژوهش‌های قبلی، اغلب بر اساس معیارهای واحد بوده است (Jain & Sinha, 2020). در حالی که در پژوهش‌های محدودی، پژوهشگران نشان داده‌اند که ترکیب معیارها می‌تواند موجب شناسایی کاربران تأثیرگذار قدرتمندتری شود ( Jain & Sinha, 2020; Sun et al., 2016; Zhuang et al., 2021).

در پژوهش حاضر، نویسندگان به‌دنبال آن هستند تا معیارهای مختلفی از بعد شبکه، فعالیت و محتوا را با هم ترکیب کنند و برای شناسایی کاربران تأثیرگذار، معیار جدیدی ارائه دهند. نویسندگان این پژوهش تمرکز خود را بر شناسایی کاربران تأثیرگذار در شبکه اجتماعی توییتر قرار داده‌اند، زیرا اولاً، یکی از موضوع‌های پژوهشی مهمی که در توییتر به آن توجه شده، شناسایی کاربران تأثیرگذار است (Riquelme & González-Cantergiani, 2016). ثانیاً، با توجه به اینکه تمرکز در پژوهش‌های پیشین، بیشتر روی شناسایی کاربران تأثیرگذار در شبکه اجتماعی توییتر بوده، به‌منظور امکان مقایسه روش پژوهش حاضر با سایر روش‌ها، از داده‌های شبکه اجتماعی توییتر استفاده شده است.

بخش‌های بعدی این پژوهش، به این صورت سازمان‌دهی شده است. در بخش پیشینه پژوهش، ادبیات مرتبط با پژوهش حاضر مرور شده است. در بخش تشریح روش پیشنهادی، در خصوص معیارهای استفاده‌شده در این پژوهش و نحوه محاسبه آنها صحبت شده است. نحوه اجرای روش پیشنهادی روی

1. Influencer
2. Influential users

مجموعه داده توییتر، ارزیابی‌ها و نتایج آنها، در بخش اجرا و ارزیابی روش پیشنهادی نشان داده شده و بخش آخر، به بحث و نتیجه‌گیری و ارائه پیشنهادهایی برای کارهای آینده اختصاص داده شده است.

## پیشینه پژوهش

تاکنون، در زمینه شناسایی کاربران تأثیرگذار، پژوهش‌های زیادی انجام شده است. برخی از روش‌های اندازه‌گیری تأثیر، از نظریه گراف ادغام‌شده با ویژگی‌های ساختاری استفاده می‌کنند، در حالی که برخی دیگر، فقط بر اساس ویژگی‌های پروفایل فردی یا تعامل با سایر کاربران آنلاین بنا شده‌اند (Jain & Sinha, 2020).

رویکردهای مبتنی بر مرکزیت[1]، مانند بینابینی[2] (فریمن[3]، ۱۹۷۸)، بردار ویژه[4] (بوناچیچ[5]، ۱۹۷۲) و رتبه صفحه[6] (پیج و همکاران[7]، ۱۹۹۹)، برای شناسایی کاربران تأثیرگذار فقط بر اساس اطلاعات ساختاری اختصاص یافته‌اند. به‌طور کلی، یک معیار مرکزیت، مقداری واقعی را به هر گره در شبکه اختصاص می‌دهد، جایی که انتظار می‌رود مقادیر تولیدشده، رتبه‌بندی گره‌ها را با توجه به اهمیت آنها ارائه دهند (Lü et al., 2016). با این حال، روش‌های مبتنی بر مرکزیت، یک محدودیت آشکار دارند، این محدویت به این صورت است که مرکزیتی که برای یک شبکه مطلوب است، ممکن است به‌دلیل ساختار توپولوژیکی متفاوت، برای شبکه دیگری بهینه نباشد (Lü et al., 2016; Zhuang et al., 2021).

معیار TFF[8] که بیگونها و همکاران[9] (۲۰۱۲) معرفی کرده‌اند، به حذف هرزنامه‌نویسان[10] احتمالی کمک می‌کند و به کاربرانی که به‌طور گسترده‌ای دنبال می‌شوند، اما انتخابی برای دنبال کردن سایر افراد دارند، ارزش می‌بخشد. ناگموتی و همکاران[11] (۲۰۱۰)، برای اندازه‌گیری محبوبیت کاربر، معیار FollowerRank را پیشنهاد داده‌اند. این معیار، نسبت دنبال‌کنندگان[12] به مجموع دنبال‌شوندگان[13] و دنبال‌کنندگان را محاسبه می‌کند و همین‌طور معیار TweetRank را برای اندازه‌گیری فعالیت کاربر پیشنهاد داده‌اند که از تعداد کل توییت‌های کاربر، برای تعیین میزان تأثیرگذار بودن آن استفاده می‌کند. نورو و همکاران[14] (۲۰۱۳) معیار Tweet Count را پیشنهاد داده‌اند که کمی پیچیده‌تر از معیار

1. Centrality
2. Betweenness
3. Freeman
4. Eigenvector
5. Bonacich
6. PageRank
7. Page et al
8. Twitter Follower-Followee ratio
9. Bigonha et al
10. Spammers
11. Nagmoti et al
12. Follower
13. Followee
14. Noro et al

TweetRank است. در این معیار، علاوه بر تعداد توییت‌های اصیل کاربر، تعداد ریتوییت‌ها نیز برای محاسبه میزان فعالیت کاربر در نظر گرفته می‌شود.

بشری و فضل‌ارثی[1] (۲۰۲۰)، با تمرکز روی پست‌های ایجادشده توسط کاربر، به‌عنوان منبع جایگزین اطلاعات، روشی جدید برای کشف کاربران تأثیرگذار در اینستاگرام طراحی کرده‌اند تا به‌طور بالقوه، راه‌حل‌های موجود را بر اساس توپولوژی شبکه یا اتصالات تقویت کند. این روش، برخلاف اکثریت قریب به اتفاق راه‌حل‌های موجود، فقط به تجزیه و تحلیل محتوای تولیدشده توسط کاربر و نه به میزان تعامل کاربر یا ساختار اتصال آنها متکی است.

به بیان ریوز و همکاران[2] (۲۰۱۹)، رویکردهای شناسایی تأثیرگذار، با استفاده از گراف‌های تحلیل شبکه، می‌توانند با فعالیت کاربرانی که در هدف شبکه اجتماعی مشارکت ندارند یا فعالیت تولیدکنندگان جعلی اسناد و فاقد اطلاعات مربوطه، گمراه شوند. از این رو، با استفاده از تجزیه و تحلیل معنایی برای فیلتر کردن چنین نوع تعاملی و دستیابی به یک نمایش گراف ساده که ویژگی‌های اصلی شبکه اجتماعی را حفظ می‌کند، امکان شناسایی تأثیرگذار واقعی را پیشنهاد می‌کنند.

آرورا و همکاران[3] (۲۰۱۹)، در پژوهش خود، سازوکاری برای اندازه‌گیری امتیاز تأثیرگذار در پلتفرم‌های رسانه‌های اجتماعی محبوب از جمله فیس‌بوک، توییتر و اینستاگرام ارائه می‌دهند. آنها چند ویژگی از رسانه‌های اجتماعی را طبقه‌بندی کرده و از مدل‌سازی رگرسیون، برای محاسبه امتیاز تأثیرگذار با استفاده از ساختارهای شناسایی‌شده استفاده می‌کنند. آلپ و اگیودوچی[4] (۲۰۱۸)، در پژوهش خود، روش جدید الگوریتم رتبه صفحه شخصی‌سازی‌شده[5] را پیشنهاد می‌دهند که هم از ویژگی‌های خاص کاربر و هم از ویژگی های توپولوژیکی شبکه برای شناسایی تأثیرگذار اجتماعی موضوعی در یک شبکه استفاده می‌کند.

لی و همکاران[6] (۲۰۱۱)، برای ارزیابی قدرت افراد تأثیرگذار و شناسایی وبلاگ‌نویسان تأثیرگذار در وبلاگ رچ[7]، مدل ارزش بازاریابی تأثیرگذار (MIV)[8] را توسعه داده‌اند. آنها، سه بعد ویژگی‌های وبلاگ (عوامل مبتنی بر شبکه، مبتنی بر محتوا و مبتنی بر فعال بودن) را تجزیه و تحلیل کرده و از شبکه عصبی مصنوعی (ANN)[9]، برای کشف وبلاگ‌نویسان بالقوه استفاده کرده‌اند. جین و سینها[10] (۲۰۲۰)، معیار جدید تأثیر هم‌بسته وزنی (WCI)[11] را معرفی کرده‌اند که ترکیب وزنی ویژگی‌های نمایه کاربر و ویژگی‌های ساختاری شبکه مبتنی بر موضوع (ویژگی‌های تعامل) را برای شناسایی کاربران تأثیرگذار در

1. Bashari & Fazl-Ersi
2. Ríos et al
3. Arora et al
4. Alp & Öğüdücü
5. Personalized PageRank
6. Li et al
7. Wretch
8. Marketing Influential Value
9. Artificial Neural Networks
10. Jain & Sinha
11. Weighted Correlated Influence

نظر می‌گیرد. رویکرد MSI[1] که ژانگ و همکاران[2] (۲۰۲۱) معرفی کرده‌اند، با در نظر گرفتن عوامل چندبعدی (مبتنی بر ساختار، مبتنی بر اطلاعات و مبتنی بر عمل) و ترکیب آنها تأثیر اجتماعی را در سطح موضوع و در سطح جهانی، روی شبکه اجتماعی سینا ویبو[3] اندازه‌گیری می‌کند. هریگان و همکاران[4] (۲۰۲۱)، برای دستیابی به هدف خود که نشان دادن چگونگی شناسایی تأثیرگذارها، به‌ویژه متخصصان بازار[5]، در شبکه‌های اجتماعی است، در پژوهش خود از روش دو مرحله‌ای استفاده کرده‌اند. در مرحله یک، آنها از طریق یک نظرسنجی آنلاین از کاربران رسانه‌های اجتماعی، متخصص‌های بازار خود ـ ادراک ـ شده[6] را شناسایی کردند و در مرحله دو، با خراشیدن[7] و تجزیه و تحلیل فعالیت‌های اخیر آنها در شبکه‌های اجتماعی، رفتارهای رسانه‌های اجتماعی این «متخصص‌ها» و «غیرمتخصص‌های» بازار را مقایسه کرده‌اند. در جدول ۱ به‌طور خلاصه مروری بر این پژوهش‌ها داشته‌ایم.

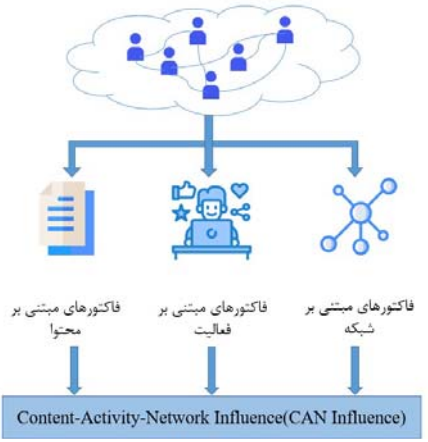
جدول ۱. خلاصه پیشینه پژوهش

| روش | نام معیار | ابعاد | | | منبع |
|---|---|---|---|---|---|
| | | شبکه | فعالیت | محتوا | |
| تک‌بعدی | Betweenness | ✓ | | | Freeman, 1978 |
| | Eigenvector | ✓ | | | Bonacich, 1972 |
| | PageRank | ✓ | | | Page et al., 1999 |
| | TFF | ✓ | | | Bigonha et al., 2012 |
| | FollowerRank | ✓ | | | Nagmoti et al., 2010 |
| | TweetRank | | ✓ | | Nagmoti et al., 2010 |
| | Tweet Count | | ✓ | | Noro et al., 2013 |
| | [8]Untitled | ✓ | | | Ríos et al., 2019 |
| | Untitled | | | ✓ | Bashari & Fazl-Ersi, 2020 |
| چندبعدی | Untitled | ✓ | ✓ | | Arora et al., 2019 |
| | Personalized PageRank | ✓ | ✓ | | Zengin Alp & Gündüz Öğüdücü, 2018 |
| | WCI | ✓ | ✓ | | Jain & Sinha, 2020 |
| | MSI | ✓ | ✓ | ✓ | Zhuang et al., 2021 |
| | MIV | ✓ | ✓ | ✓ | Li et al., 2011 |
| | Untitled | ✓ | ✓ | ✓ | Harrigan et al., 2021 |

1. Multidimensional Influence
2. Zhuang et al
3. Sinaweibo
4. Harrigan et al
5. Mavens
6. Self-perceived
7. Scraping

۸. Untitled در ستون نام معیار به این معناست که نام خاصی برای این معیار در پژوهش مربوطه تعیین نشده است.

اگرچه این پژوهش‌ها، بینش مهمی در شناسایی کاربران تأثیرگذار ارائه می‌دهند، با این حال، تشخیص تأثیرگذاران در شبکه در مقیاس بزرگ هنوز فاقد رویکرد دقیق و کارآمد است و با وجود اهمیت ترکیب معیارها برای شناسایی کاربران تأثیرگذار، پژوهش‌های بسیار کمی به این موضوع پرداخته‌اند (Zhuang et al., 2021). با توجه به بینشی که از بررسی پژوهش‌های پیشین حاصل شد، نویسندگان در پژوهش حاضر، روشی ارائه دادند که بر مبنای ترکیب معیارهایی مبتنی بر شبکه، فعالیت و محتوا است که به‌صورت چارچوبی کلی در شکل ۱ نشان داده شده است. در این پژوهش، از ترکیب ۹ معیار مختلف در شبکه اجتماعی توییتر برای ارائه یک معیار ترکیبی جدید استفاده شده است که با توجه به بررسی‌ها و مرور پژوهش‌های نویسندگان، تاکنون ترکیب همه این ویژگی‌ها به‌طور یک‌جا در مقاله‌ای دیده نشده است. به‌علاوه، یکی از این ۹ معیار، معیار ابتکاری نویسندگان برای محاسبه تعامل است. در بخش بعدی، روش پژوهش، حاضر برای شناسایی کاربران تأثیرگذار شرح داده شده است.




شکل ۱. چارچوب کلی روش پژوهش حاضر

## تشریح روش پیشنهادی

پس از مرور ادبیات پیشین پژوهش و شناسایی شکاف پژوهش‌های گذشته، مدل پژوهش حاضر برای شناسایی کاربران تأثیرگذار ایجاد شده است. مدل کلی این پژوهش به این صورت است که معیارهایی از سه بعد شبکه، فعالیت و محتوا با یکدیگر ترکیب می‌شوند و معیار ترکیبی جدیدی برای شناسایی کاربران تأثیرگذار ایجاد می‌شود. همان‌طور که در ادامه به آن اشاره خواهد شد، پس از اجرای مدل پژوهش با استفاده از زبان برنامه‌نویسی پایتون[1] روی داده‌های شبکه اجتماعی توییتر، به کمک روش‌های اعتبارسنجی معتبر، اثربخشی و دقت روش پیشنهادی ارزیابی می‌شود.

1. Python

در این بخش، در گام نخست معیارهای استفاده‌شده در پژوهش حاضر معرفی خواهند شد، در گام دوم روش ترکیب این معیارها و در گام سوم روش ایجاد معیار ترکیبی جدید شرح داده خواهد شد.

### مرحله نخست: معرفی معیارها

همان‌طور که ریکلمه و گونزالس کانترجیانی[1] (۲۰۱۶) بیان کرده‌اند، می‌توان معیار را یک عبارت ساده ریاضی در نظر گرفت که به ما کمک می‌کند تا اطلاعات اساسی در خصوص شبکه اجتماعی را به‌صورت مقدار عددی ارائه دهیم. همچنین، معیارها را می‌توان برای تعریف یک معیار (رتبه‌بندی) ترکیب کرد، برای مثال، یک رابطه یا یک الگوریتم که ملاکی برای رتبه‌بندی هر کاربر در یک شبکه فراهم می‌کند.

جدول ۲. توصیف معیارهای استفاده‌شده در کار حاضر

| بعد | معیار | توضیح | استفاده‌شده در پژوهش‌ها |
|---|---|---|---|
| شبکه | Follower count (Flw) | تعداد دنبال‌کنندگان | Harrigan et al., 2021; Jain & Sinha, 2020; Ramírez-de-la-Rosa et al., 2014 |
| | Betweenness(Bet) | مرکزیت مبتنی بر مسیر | Bigonha et al., 2012; Jain & Sinha, 2020; Zhuang et al., 2021 |
| | PageRank(Pr) | مرکزیت مبتنی بر ارزش | Jain & Sinha, 2020; Li et al., 2011 |
| فعالیت | Signal strength(Ss) | توییت‌های کاربر (نه ریتوییت) / توییت‌ها + ریتوییت‌ها | Pal & Counts, 2011; Zengin Alp & Gündüz Öğüdücü, 2018 |
| | Tweet count(Tc) | تعداد کل توییت‌های کاربر | Harrigan et al., 2021; Jain & Sinha, 2020; Li et al., 2011; Noro et al., 2013 |
| | Original interaction(Oi) | نقل قول توییت / ریتوییت | معیار پیشنهادی |
| محتوا | Levenshtein(Lev) | میانگین کاربر از فاصله ویرایشی لونشتاین بین توییت اصلی و تبدیل حروف کوچک توییت تقسیم بر طول کل توییت | Harrigan et al., 2021 |
| | Status length(Sl) | میانگین طول پست | Li et al., 2011; Ramírez-de-la-Rosa et al., 2014 |
| | Hashtag mean(Hst) | میانگین مجموع هشتگ‌های (#) استفاده شده در توییت‌ها | Harrigan et al., 2021; Ramírez-de-la-Rosa et al., 2014 |

1. Riquelme & González-Cantergiani

در پژوهش حاضر، برای اندازه‌گیری تأثیر کاربران، از ۹ معیار مختلف استفاده شده است و همان‌طور که در جدول ۲ به‌طور خلاصه نشان داده شده، یکی از معیارها، حاصل نوآوری پژوهش حاضر و ۸ معیار دیگر، حاصل پژوهش‌های پیشینی هستند که در حوزه شناسایی کاربران تأثیرگذار انجام شده‌اند که در ستون ـ استفاده‌شده در پژوهش‌ها ـ به این پژوهش‌ها اشاره شده است. نویسندگان این پژوهش، ترکیب این ۹ معیار را نیز در پژوهش‌های گذشته، مشاهده نکرده‌اند. هر یک از این معیارها در ادامه به‌طور کامل معرفی شده‌اند.

### *معرفی معیارهای مبتنی بر شبکه*

تأثیر یک گره تا حد زیادی تحت‌تأثیر ساختار توپولوژیکی شبکه‌ای است که به آن تعلق دارد (Lü et al., 2016). مانند هر شبکه اجتماعی دیگر، توییتر را نیز می‌توان با یک گراف جهت‌دار $G = (V, E)$ نشان داد که V مجموعه گره‌ها و E مجموعه یال‌های[1] جهت‌داری است که نشان‌دهنده نحوه ارتباط گره‌ها است (Riquelme & González-Cantergiani, 2016). در این کار، معیارهای مبتنی بر شبکه شامل سه معیار در ارتباط با شبکه کاربر هستند: مرکزیت بینابینی، مرکزیت رتبه صفحه و تعداد دنبال‌کنندگان کاربر.

- **رتبه صفحه:** ایده اصلی رتبه صفحه به این صورت است که اهمیت یک صفحه وب توسط کمیت و کیفیت صفحات پیوند داده‌شده به آن تعیین می‌شود (Lü et al., 2016). رتبه صفحه، تأثیر همسایگان و همسایگان آنها را در نظر می‌گیرد. برای مثال، با تعداد کمی از دوستان تأثیرگذار، امتیازات می‌تواند بالاتر از کسانی با دوستان بسیار، اما کمتر تأثیرگذار باشد (Joshi & Mohammed, 2020).
- **بینابینی:** بینابینی یکی از معیارهای برجسته مرکزیت است و با کسری از کوتاه‌ترین مسیرها بین جفت گره‌هایی که از گره مد نظر عبور می‌کنند، تعریف می‌شود (Brandes, 2008). در گراف ریتوییت‌ها، کاربران دارای بینابینی بالا، در فرایند انتشار اطلاعات نقش مهمی دارند، زیرا به‌عنوان پل‌هایی برای جریان داده عمل می‌کنند (Bigonha et al., 2012). در واقع، گره با بیشترین مرکزیت بینابینی، قوی‌ترین کنترل را بر جریان اطلاعات دارد (Lü et al., 2016).
- **تعداد دنبال‌کنندگان کاربر:** تعداد دنبال‌کنندگان میزان توجه افراد به کاربر را نشان می‌دهد و می‌توان آن را از صفحه کاربر مشاهده کرد (Sun et al., 2016). کاربران تأثیرگذار، به‌طور شایان توجهی در مقایسه با کاربران غیرتأثیرگذار، تعداد بیشتری دنبال‌کننده دارند (Harrigan et al., 2021; Ramírez-de-la-Rosa et al., 2014).

1. Edges

### معرفی معیارهای مبتنی بر فعالیت

یک گره تأثیرگذار، باید در جامعه شبکه خود، جذابیت و فعالیت داشته باشد (Li et al., 2011). در پژوهش حاضر، برای محاسبه میزان فعالیت کاربران، از سه معیار تعداد توییت‌های کاربر[1]، قدرت سیگنال[2] و معیار ابتکاری تعامل اصیل[3] استفاده شده است.

- **تعداد توییت:** به‌اشتراک‌گذاری محتوا یکی از فعالیت‌های اصلی وبلاگ‌نویسان است و اگر به‌طور منظم باشد، می‌تواند به وبلاگ‌نویسان کمک کند تا محبوبیت بیشتری برای وبلاگ خود کسب کنند (Li et al., 2011). متخصصان بازار، در مقایسه با غیرمتخصصان تعداد پست‌های شایان توجه بیشتری دارند (Harrigan et al., 2021). در این کار، تعداد کل توییت‌های کاربر به‌عنوان یکی از معیارهای ارزیابی تأثیر کاربر در نظر گرفته شده است.
- **قدرت سیگنال:** این ویژگی، اصالت توییت‌های نویسنده را اندازه‌گیری می‌کند، اگر بیشتر توییت‌های یک کاربر از خودش نشئت گرفته باشند، مقدار قدرت سیگنال برای آن کاربر بیشتر است. به‌گونه‌ای که نویسندگی بیشتر، به‌معنای مقادیر نزدیک به ۱ است (Pal & Counts, 2011; Riquelme & González-Cantergiani, 2016; Zengin Alp & Gündüz Öğüdücü, 2018). این معیار به‌صورت رابطه ۱ محاسبه می‌شود:

$$Signal\ strength\ (Ss) = \frac{OT}{OT + RT}$$ رابطه ۱)

OT تعداد توییت‌های اصلی کاربر و RT تعداد ریتوییت‌های توییت دیگران توسط کاربر است.

- **تعامل اصیل:** یک وبلاگ‌نویس فعال‌تر، تمایل بیشتری برای برقراری مکالمه اجتماعی با سایر کاربران از خود نشان می‌دهد (Li et al., 2011). تعامل با سایر کاربران توییتر، به چند صورت امکان‌پذیر است، برای مثال ریتوییت، نقل قول[4] کردن، لایک کردن و غیره. ریتوییت، توییت اصلی و سازنده محتوای آن را در جلو و مرکز قرار می‌دهد و اجازه می‌دهد توییت آنها در تمرکز باشد. نقل قول توییت، محتوای اصیلی است که بر اساس محتوای همکاران ساخته می‌شود و تلاقی عالی‌ای بین اصالت و تعامل است که به نظر می‌رسد توییتر مشتاق تشویق آن است[5]. بر همین اساس، نویسندگان معیار «تعامل اصیل» را به‌عنوان معیاری ابتکاری برای بررسی تأثیرگذار بودن یا نبودن یک کاربر ارائه داده‌اند که به‌صورت رابطه ۲ محاسبه می‌شود:

$$Original\ interaction(Oi) = \frac{QT}{RT}$$ رابطه ۲)

QT تعداد نقل قول توییت‌هایی است که کاربر به اشتراک گذاشته است.

---

1. Tweet count(Tc)
2. Signal strength(Ss)
3. Original interaction(Oi)
4. Quote
5. https://nealschaffer.com/what-is-a-quote-tweet/ [Accessed December 5,2021].

### *معرفی معیارهای مبتنی بر محتوا*

محتوای تولیدشده توسط کاربر، معمولاً، به‌دلیل تنوع پیشینه کاربران و اهداف متفاوت آنها، بسیار ناهمگن است (Bigonha et al., 2012). تجزیه و تحلیل محتواگرا، برای دستیابی به درک بهتر تأثیر اجتماعی ضروری است (Zhuang et al., 2021). هدف این پژوهش از تجزیه و تحلیل محتوای توییت‌ها، رتبه‌بندی پست‌های بهتر (و در نتیجه نویسندگان آنها) است و در اینجا از معیارهای تعداد هشتگ‌ها، طول پست و میانگین درصد حروف بزرگ استفاده شده است.

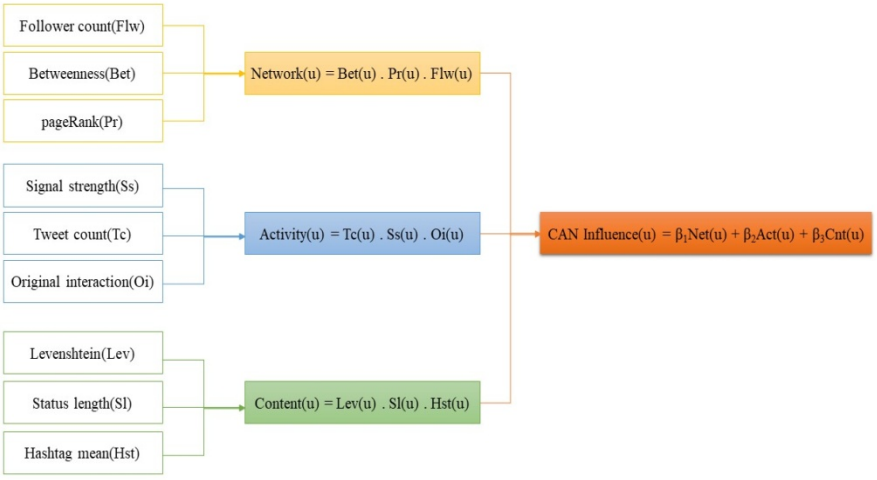
- **میانگین هشتگ‌ها:** کاربران تأثیرگذار در مقایسه با کاربران غیرتأثیرگذار بیشتر از هشتگ‌ها استفاده می‌کنند (Harrigan et al., 2021; Ramírez-de-la-Rosa et al., 2014). به همین منظور، در پژوهش حاضر از میانگین مجموع هشتگ‌های (#) استفاده‌شده در مجموعه توییت‌ها، به‌عنوان یکی از معیارهای شناسایی افراد تأثیرگذار استفاده شده است.
- **طول پست:** طول پست وبلاگ با تعداد نظرها هم‌بستگی مثبت دارد، به این معنا که پست‌های طولانی‌تر وبلاگ، توجه افراد را به خود جلب می‌کنند (Agarwal et al., 2008). توییت‌ها، محدودیت تعداد کاراکتر دارند، از این رو، احتمال بیشتری وجود دارد که هر بار کاربران تأثیرگذار، پستی می‌نویسند، از حداکثر تعداد کاراکترهای مجاز استفاده کنند (Ramírez-de-la-Rosa et al., 2014). در پژوهش حاضر نیز، میانگین طول توییت‌های یک کاربر که از تقسیم مجموع کاراکترهای توییت‌های کاربر به تعداد توییت‌ها به دست می‌آید، به‌عنوان معیاری در تعیین تأثیر کاربران در نظر گرفته شده است.
- **میانگین درصد حروف بزرگ:** کاربران تأثیرگذار، در مقایسه با کاربران غیرتأثیرگذار، به‌طور شایان توجهی در پست‌های خود، از حروف بزرگ استفاده می‌کنند (Harrigan et al., 2021). مشابه کار هریگان و همکاران[1] (۲۰۲۱)، در این پژوهش نیز از معیار لونشتاین[2] به‌عنوان یکی از معیارهای تشخیص کاربران تأثیرگذار استفاده شده است. معیار لونشتاین به این صورت محاسبه می‌شود که میانگین فاصله ویرایش لونشتاین بین توییت اصلی و توییت تبدیل‌شده با حروف کوچک، محاسبه شده و بر طول کل توییت تقسیم می‌شود. فاصله ویرایش لونشتاین تعداد کاراکترهای تکی را نشان می‌دهد که برای به دست آوردن توییت اصلی باید با حروف بزرگ نوشته شوند.

### مرحله دوم: روش محاسبه تأثیر در هر بعد در روش پیشنهادی

پس از محاسبه معیارها، در هر بعد باید آنها با هم ترکیب شوند تا بتوان معیار بهتری برای سنجش تأثیر کاربر ارائه داد. به همین منظور در *پژوهش* حاضر، مشابه روش ژانگ و همکاران[3] (۲۰۲۱) ویژگی‌های

1. Harrigan et al
2. Levenshtein
3. Zhuang et al

محاسبه‌شده با هم ترکیب می‌شوند که روش ترکیب این معیارها، به‌طور خلاصه در شکل ۲ آورده شده است.




شکل ۲. نحوه ترکیب معیارها در رویکرد پیشنهادی CAN Influence

### *تأثیر شبکه*

تأثیر شبکه، نشان‌دهنده اهمیت یک فرد در شبکه است که می‌تواند به‌صورت رابطه ۳ محاسبه شود:

$$Net(u) = Bet(u) . Pr(u) . Flw(u)$$ رابطه ۳)

«.» عملگر حاصل‌ضرب داخلی است که برای ترکیب اثر عوامل بر اندازه‌گیری تأثیر استفاده می‌شود.

### *تأثیر فعالیت*

مشابه تأثیر شبکه، برای به دست آوردن میزان تأثیر فعالیت کاربر که نشان‌دهنده میزان فعالیت کاربر در شبکه اجتماعی است از رابطه ۴ استفاده شده است:

$$Act(u) = Tc(u) . Ss(u) . Oi(u)$$ رابطه ۴)

### *تأثیر محتوا*

تأثیر محتوای کاربر را می‌توان از نحوه به‌اشتراک‌گذاری محتوای آن در شبکه اجتماعی به دست آورد که در اینجا به‌صورت رابطه ۵ محاسبه شده است:

$$Cnt(u) = Lev(u) . Sl(u) . Hst(u)$$ رابطه ۵)

بـــر ایـــن اســـاس $N = (\text{Net(u)})_{n\times1}$ و $A = (\text{Act(u)})_{n\times1}$ و $C = (\text{Cnt(u)})_{n\times1}$ بـــه‌عنـــوان ماتریس‌هایی $n \times 1$ از تأثیر شبکه، فعالیت و محتوا ارائه شده است.

**مرحله سوم: روش محاسبه معیار ترکیبی جدید**

حال می‌توان با ترکیب این سه تأثیر محاسبه‌شده در رابطه‌های ۳، ۴ و ۵، تأثیر کلی هر یک از کـاربران را از طریق رابطه ۶ به دست آورد و آنها را بر اساس این معیار محاسبه‌شده رتبه‌بندی کرد تا کاربرانی کـه در چند جنبه تأثیرگذاری بیشتری دارند، شناسایی شوند.

$$CAN\ Influence(u) = \beta1Net(u) + \beta2Act(u) + \beta3Cnt(u)$$ رابطه ۶)

β1, β2, β3 وزن‌هایی هستند که می‌توانند برای تنظیم میـزان تـأثیر هـر یـک از ابعـاد تـأثیر شـبکه، فعالیت و محتوا استفاده شود. در این کار مشابه روش ژانگ و همکـاران (۲۰۲۱) از «روش وزنـی آنتروپـی اطلاعات» برای محاسبه این وزن‌ها استفاده شده است که $\sum_{l=1}^{3}\beta_l = 1$.

اصولاً، برای به دست آوردن وزن‌ها سه مرحله وجود دارد. در مرحله نخست، استانداردسـازی داده‌هـا، در مرحله دوم، محاسبه آنتروپی هر عامل بر اساس تعریف آنتروپی اطلاعات و در مرحله سوم، تعیـین وزن هر عامل با توجه به آنتروپی اطلاعاتی که دارد (Zhuang et al., 2021). در اینجا از روش آنتروپـی شـانون $E_1 = -\frac{1}{\ln(n)}\sum_{i,j} q_{i,j} \ln(q_{ij})$ (در ایـــن پـــژوهش $q_{ij} = Net(i,j)$ و $q_{ij} = Act(i,j)$ و یـــا $q_{ij} = Cnt(i,j)$ است.)، برای محاسبه آنتروپی عوامل و از $\beta_l = \frac{1-E_j}{\sum l(1-E_l)}$، برای تعیـین وزن‌هـا اسـتفاده شده است.

## اجرا و ارزیابی روش پیشنهادی

### مجموعه داده

برای پیاده‌سازی و ارزیابی روش پیشنهادی، به استفاده از مجموعه داده‌های واقعی مربوط بـه یـک شـبکه اجتماعی نیاز است. بـرای ایـن منظـور، در پـژوهش حاضـر، از مجموعـه داده تـوییتر پـژوهش ریبیـرو و همکاران[1] (۲۰۱۸) استفاده شده است. این مجموعه داده، شامل ۱۰۰۳۸۶ کاربر و ۲۲۸۶۵۹۲ یـال و ۲۰۰ توییت اخیر برای هر کاربر است که در بازه زمانی بین یکم تا هفتم اکتبر ۲۰۱۷ جمع‌آوری شده است. در این مجموعه داده، ارتباطات بین کاربران در توییتر با استفاده از شـبکه ریتوییـت نشـان داده شـده اسـت (Ribeiro et al., 2018).

1. Ribeiro et al

### ابزار پیاده‌سازی روش

روش پیشنهادی پژوهش حاضر، با بهره‌گیری از زبان برنامه‌نویسی پایتون نسخه ۳/۸ در محیط اسپایدر[1] روی مجموعه داده معرفی‌شده در بخش قبل، پیاده‌سازی شده است. برای پیاده‌سازی از یک رایانه شخصی با پردازنده اینتل Core i7، ۲/۶۰ گیگاهرتز و ۱۲ گیگابایت رم استفاده شده است. در کار حاضر، مشابه کار بیگونها و همکاران[2] (۲۰۱۱) مرکزیت‌های بینابینی، رتبه صفحه و بردار ویژه با استفاده از پکیج پایتون نتورک ایکس[3] محاسبه شده‌اند. نتورک ایکس یک بسته پایتون برای ایجاد، دستکاری و مطالعه ساختار، دینامیک و عملکرد شبکه‌های پیچیده است[4].

### ارزیابی روش

ارزیابی تأثیرگذاران شناسایی‌شده، کاری چالش‌برانگیز است، زیرا، این یک موضوع کاملاً ذهنی است و از رویکردهای مختلفی در ادبیات استفاده شده است (Zengin Alp & Gündüz Öğüdücü, 2018). در این پژوهش، از پنج روش «گسترش»[5] (Zengin Alp & Gündüz Öğüdücü, 2018)، «دقت»[6]، «فراخوانی»[7] و «F1-measure» (Jain & Sinha, 2020) و «درصد تکرار»[8] (Wang et al., 2013) برای ارزیابی اثربخشی و دقت روش پیشنهادی به کمک داده‌های توییتر، استفاده شده است. برای مقایسه عملکرد روش پیشنهادی با کارهای پیشین انجام‌شده در این حوزه، برای شناسایی گره‌های تأثیرگذار، پنج روش در نظر گرفته شده است: Tweet Count ، TFF، PageRank، Signal strength و Eigenvector.

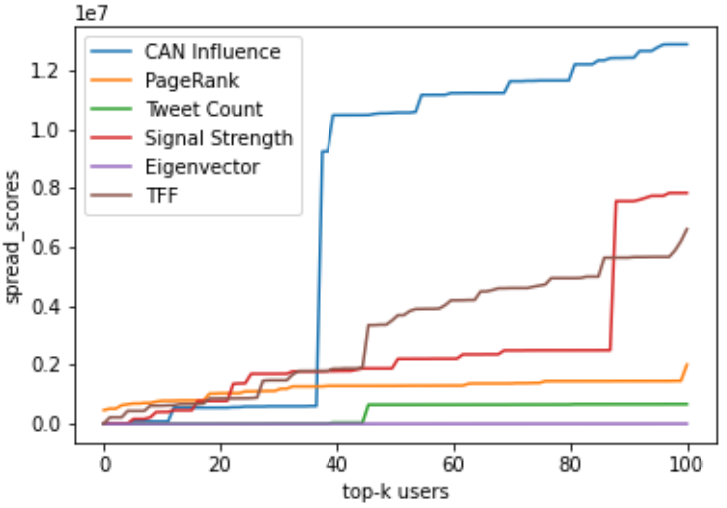
در ابتدا، عملکرد روش CAN Influence با روش اعتبارسنجی «گسترش» که بر اساس نرخ بازتوییت تأثیرگذاران محاسبه می‌شود، ارزیابی می‌شود تا پتانسیل آنها در انتشار اطلاعات مشخص شود. به این منظور، نرخ بازتوییت نرمال‌شده برای هر کاربر محاسبه می‌شود و نتایج برای محاسبه تخمین انتشار کلی اطلاعات k تأثیرگذار برتر جمع‌بندی می‌شود. رابطه ۷، نحوه محاسبه گسترش اطلاعات بالقوه را نشان می‌دهد.

$$Spread = \sum \frac{RTT}{T} \sum RT \qquad \text{رابطه ۷)}$$

$RTT$ تعداد توییت‌های یک کاربر که ریتوییت شده‌اند، $T$ تعداد توییت‌های کاربر و $RT$ تعداد ریتوییت‌ها است. امتیازات گسترش بالاتر نشان می‌دهد که کاربران تأثیرگذار شناسایی‌شده پتانسیل بیشتری برای انتشار اطلاعات دارند (Zengin Alp & Gündüz Öğüdücü, 2018).

1. Spyder
2. Bigonha
3 NetworkX
4 https://networkx.org/ [Accessed January 21,2022].
5 Spread
6. Precision
7. Recall
8. Duplication percentage

شکل ۳ مجموع امتیاز گسترش CAN Influence را برای تعداد مختلف کاربران تأثیرگذار، در مقابل سایر کارهای شناخته‌شده نشان می‌دهد. همان‌طور که در شکل ۳ نشان داده شده است، روش پیشنهادی این پژوهش، از نظر روش اعتبارسنجی گسترش در تعداد بالای کاربران تأثیرگذار شناسایی‌شده، در مقایسه با پژوهش‌های مقایسه‌شده قبلی، امتیاز گسترش بالاتری کسب کرده است.




شکل ۳. مقایسه معیار پخش k کاربر برجسته

در ادامه، برای نشان دادن اثربخشی روش پیشنهادی، از روش‌های اعتبارسنجی دقت، فراخوانی و F1-measure استفاده می‌شود. مشابه کار جین و سینها[1] (۲۰۲۰) روش‌های اعتبارسنجی روی یک مجموعه مجزا از k گره تأثیرگذار که توسط هر روش شناسایی شده‌اند، محاسبه می‌شوند. در ابتدا، فهرست مرجع یا مجموعه‌ای از کاربران تأثیرگذار ایجاد می‌شود که حاوی مقادیر منحصر به فردی برای تطبیق مجموعه top-k بازیابی‌شده توسط هر اندازه‌گیری است. فهرست مرجع جدید، هر بار برای مقدار k متمایز تولید شده و به‌صورت $R_k$ نمایش داده می‌شود. برای مثال، k کاربر با نفوذی که با معیار CAN Influence پیشنهادی شناسایی می‌شوند، با $N_k(CANI)$ نشان داده می‌شوند. به‌طور مشابه، برای روش‌های دیگر به‌طور جداگانه به‌عنوان $N_k(TR)$، $N_k(TFF)$، $N_k(PR)$، $N_k(SS)$ و $N_k(E)$ تعریف می‌شود. بنابراین، برای انجام اعتبارسنجی، فهرست مرجع به‌صورت ریاضی در رابطه ۸ تعریف شده است:

$$R_k = N_k(\text{CANI}) \cap N_k(TR) \cap N_k(PR) \cap N_k(SS) \cap N_k(TFF) \cap N_k(E) \quad \text{رابطه ۸)}$$

1. Jain & Sinha

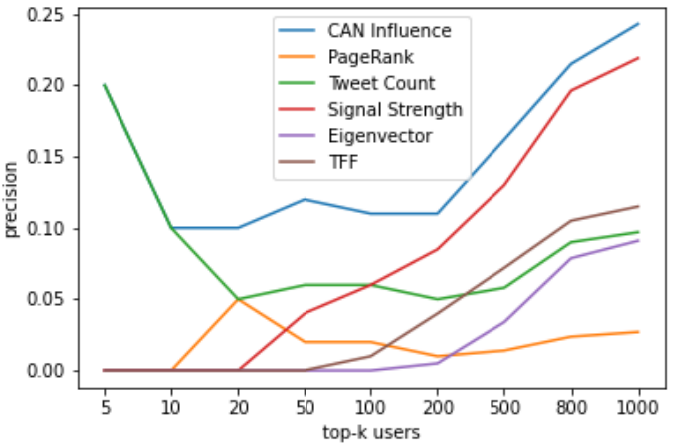
شاخص دقت، به‌عنوان نسبت تعداد کاربران تأثیرگذار که به‌درستی شناسایی شده‌اند (در صورت تطبیق با فهرست مرجع) به تعداد کل کاربران تأثیرگذار در مجموعه top-k تعریف می‌شود. به همین ترتیب، شاخص فراخوانی، نسبت تعداد کاربران تأثیرگذاری که به‌درستی شناسایی شده به تعداد کل کاربرانی است که در یک فهرست مرجع خاص هستند. شاخص F1 یا F1-measure، به میانگین وزنی شاخص دقت و شاخص فراخوانی اشاره می‌کند (Jain & Sinha, 2020). روش‌های اعتبارسنجی تعریف‌شده در زمینه کشف کاربران تأثیرگذار برای CAN Influenc در رابطه‌های ۹ تا ۱۱ ارائه شده است.

$$Precision_k = \frac{N_k(\text{CANI}) \cap R_k}{N_k(\text{CANI})}$$ رابطه ۹)

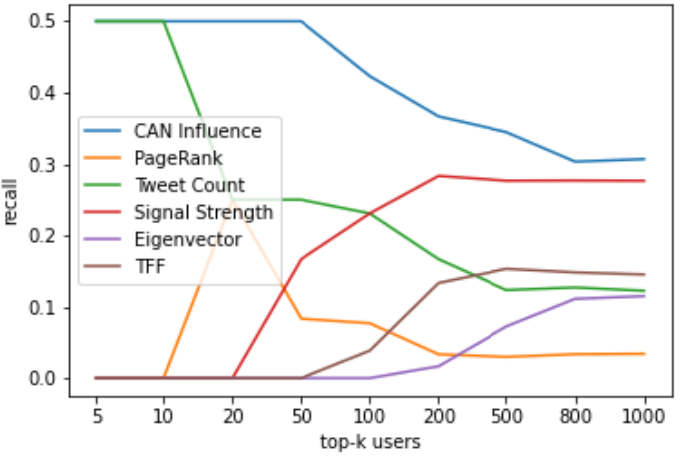
$$Recall_k = \frac{N_k(\text{CANI}) \cap R_k}{R_k}$$ رابطه ۱۰)

$$F1_k = \frac{2 \times Precision_k \times Recall_k}{Precision_k + Recall_k}$$ رابطه ۱۱)

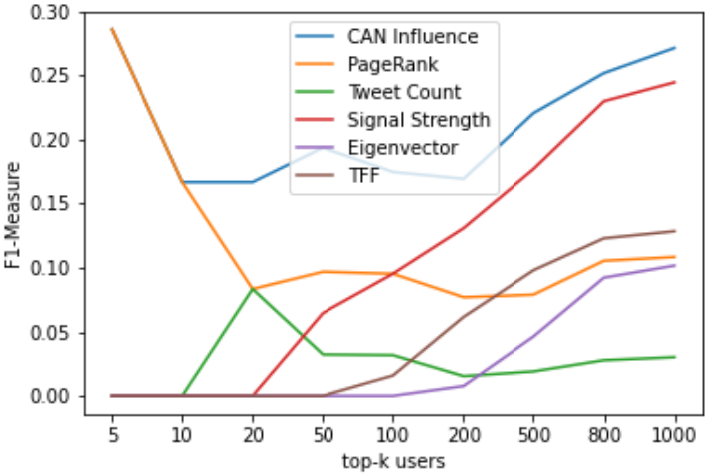
ارزیابی به‌کمک سه شاخص دقت، فراخوانی و F1-measure برای ۹ مقدار مختلف k یعنی ۵، ۱۰، ۲۰، ۵۰، ۱۰۰، ۲۰۰، ۵۰۰، ۸۰۰ و ۱۰۰۰ در مجموعه داده توییتر انجام شده است. تصاویری از نتایج ارزیابی در شکل‌های ۴، ۵ و ۶ نشان داده شده است.




شکل ۴. مقایسه معیار دقت k کاربر برجسته




شکل ۵. مقایسه معیار فراخوانی k کاربر برجسته




شکل ۶. مقایسه معیار F1-Measure برای k کاربر برجسته

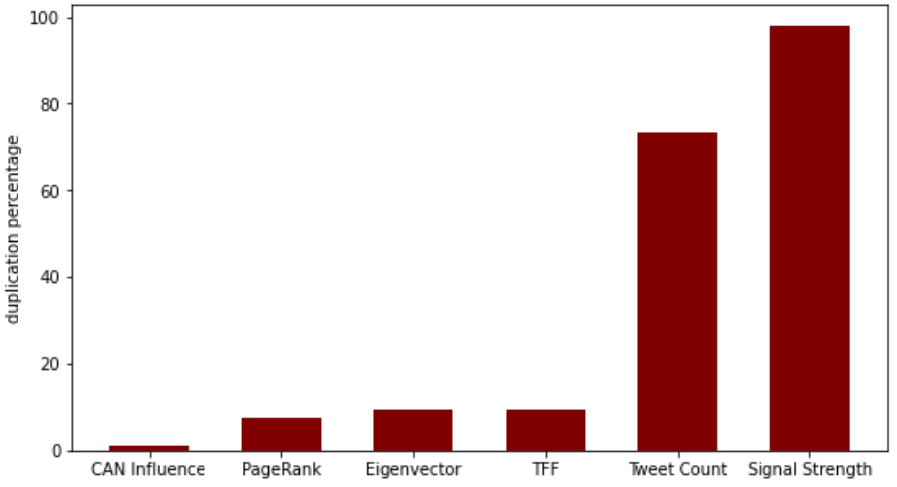
بالاترین مقدار دقت، فراخوانی و F1-measure توسط روش پیشنهادی این پژوهش، یعنی CAN Influence، روی مجموعه داده در نظر گرفته‌شده به دست آمده است. مقدار فراخوانی، تحت‌تأثیر تعداد کاربران موجود در فهرست مرجع است. بنابراین، با افزایش مقدار k، تقریباً در اغلب معیارها روند نزولی مشاهده می‌شود. در همه نتایج ارائه‌شده در شکل‌های ۴، ۵ و ۶ برای دقت، فراخوانی و F1-measure معیار

Signal strength با افزایش مقدار k، در جایگاه دوم قرار گرفته است و Eigenvector در مقایسه با سایر روش‌ها، نتایج ضعیف‌تری ارائه داده است. نتایج ارائه‌شده در شکل‌های ۴، ۵ و ۶ نشان می‌دهد که روش ترکیبی ارائه‌شده در این پژوهش در مقایسه با سایر روش‌های مقایسه‌شده در این شکل‌ها، برای شناسایی کاربران تأثیرگذار توانایی بیشتری دارد.

به‌عنوان پنجمین و آخرین روش ارزیابی، از روش اعتبارسنجی «درصد تکرار» استفاده شده است. یک رویکرد تشخیص مناسب، همه تأثیرات کاربران را به بیشترین میزان متمایز کند، به بیان دیگر، تکرار کمتر به‌معنای اثربخشی بیشتر است (Wang et al., 2013). با فرض اینکه $\xi$ تعداد گره‌هایی با مقادیر تأثیر یکسان است، آنگاه، «درصد تکرار»، $\eta$، به‌صورت رابطه ۱۲ تعریف می‌شود:

$$\eta = \begin{cases} 0\% & IF\ \xi = 0 \\ \frac{\xi}{N}.100\% & ELSE \end{cases}$$

رابطه ۱۲)




شکل ۷. مقایسه معیار درصد تکرار

با توجه به نتایج نشان داده‌شده در شکل ۷، روش پیشنهادی این پژوهش، با ۰/۰۱درصد تکرار یعنی برای ۱۰۰/۳۸۶ کاربر فقط ۱۱۱۷ رتبه تکراری، کمترین مقدار تکرار را در بین سایر روش‌ها دارد.

به‌طور خلاصه، با توجه به نتایج تجزیه و تحلیل ارائه‌شده در این بخش، می‌توان گفت که روش پیشنهادی پژوهش حاضر در روش‌های اعتبارسنجی گسترش، دقت، فراخوانی، F1-measure و درصد تکرار نتایج بهتری در مقایسه با Tweet Count، TFF، PageRank، Eigenvector و Signal strength به دست آورده است.

## نتیجه‌گیری

در سال‌های اخیر، شناسایی کاربران تأثیرگذار در شبکه‌های اجتماعی، به یکی از موضوع‌هایی تبدیل شده که پژوهشگران به آن توجه نشان داده‌اند. در میان روش‌های پیشنهادی پژوهش‌های پیشین، با وجود اهمیت این موضوع، تعداد کمی از پژوهش‌ها به شناسایی افراد تأثیرگذار از طریق بررسی ویژگی‌های مختلف این افراد پرداخته‌اند. روش پیشنهادی پژوهش حاضر، ویژگی‌های کاربران شبکه اجتماعی توییتر را از سه دیدگاه شبکه، فعالیت و محتوا بررسی می‌کند. نتایج حاصل از ارزیابی روش پیشنهادی، نشان می‌دهد که به کمک این روش، می‌توان کاربرانی را به‌عنوان تأثیرگذار معرفی کرد که قوی‌تر از کاربرانی هستند که فقط از یک دیدگاه بررسی شده‌اند.

در این مقاله، معیار ترکیبی جدیدی مبتنی بر توییتر برای شناسایی افراد تأثیرگذار پیشنهاد شده است که به‌طور کلی ۹ معیار مختلف مرکزیت بینابینی، رتبه صفحه، تعداد دنبال‌کنندگان، تعداد توییت‌ها، قدرت سیگنال، تعامل اصیل، تعداد هشتگ‌ها، طول پست و تعداد حروف بزرگ به‌کاربرده‌شده در توییت‌ها را ترکیب می‌کند که معیار تعامل اصیل معیار ابتکاری پژوهش حاضر است.

به‌منظور مقایسه روش پیشنهادی پژوهش حاضر با سایر کارهای شناخته‌شده در این حوزه، امتیاز تأثیر کاربر با استفاده از پنج معیار مختلف Tweet Count، TFF، PageRank، Eigenvector و Signal strength محاسبه شده و عملکرد این پنج معیار، در مقایسه با معیار ترکیبی ارائه‌شده در پژوهش حاضر، با استفاده از پنج روش اعتبارسنجی امتیاز گسترش، دقت، فراخوانی، F1-measure و درصد تکرار سنجیده شده است. نتایج به‌دست‌آمده از روش‌های اعتبارسنجی گسترش، دقت، فراخوانی و F1-measure نشان‌دهنده بالاتر بودن مقدار امتیاز محاسبه‌شده برای روش پیشنهادی پژوهش حاضر، نسبت به سایر روش‌های مقایسه‌شده است. همچنین از نظر تکراری بودن رتبه‌های پیشنهادی برای کاربران، روش پیشنهادی توانسته کمترین مقدار تکرار که ۰/۰۱ درصد است را در میان سایر کارهای مقایسه‌شده به دست آورد.

یکی از محدودیت‌هایی این پژوهش، شناسایی کاربران تأثیرگذار فقط در شبکه اجتماعی توییتر بوده است که با توجه به تنوع شبکه‌های اجتماعی، می‌توان هدف بعدی را ایجاد معیاری که بتواند افراد تأثیرگذار را در شبکه‌های اجتماعی مختلف مانند اینستاگرام، فیسبوک و غیره با دقت بالا شناسایی کند، تعیین کرد. در کارهای آینده، می‌توان معیارهای بیشتری را بررسی کرد (مانند میزان خوانایی پست‌های کاربر، میزان استفاده از تصاویر در پست‌ها و غیره) و فهرست جامع بهتری از معیارها را در هر سه بعد برای محاسبه تأثیر چندبعدی ایجاد کرد. همچنین، استفاده از تحلیل احساسات، برای تحلیل نظر دنبال‌کنندگان کاربر، در کنار معیارهای ترکیبی، می‌تواند در افزایش دقت و اثربخشی شناسایی کاربران تأثیرگذار مؤثر باشد. افزون بر این، می‌توان با ایجاد شبکه‌های موضوعی، کاربران تأثیرگذار را در حوزه‌ای که در آن تأثیرگذارتر هستند، شناسایی کرد.

## فهرست منابع


Agarwal, N., Liu, H., Tang, L., & Yu, P. S. (2008). Identifying the influential bloggers in a community. *Proceedings of the 2008 international conference on web search and data mining.*

Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y .(2019) .Measuring social media influencer index- insights from facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*, *49*, 86-101. https://doi.org/https://doi.org/10.1016/j.jretconser.2019.03.012

Bashari, B., & Fazl-Ersi, E. (2020). Influential post identification on Instagram through caption and hashtag analysis. *Measurement and Control*, *53*(3-4), 409-415. https://doi.org/10.1177/0020294019877489

Bigonha, C., Cardoso, T. N., Moro, M. M., Gonçalves, M. A., & Almeida, V. A. (2012). Sentiment-based influence detection on Twitter. *Journal of the Brazilian Computer Society*, *18*(3), 169-183.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, *2*(1), 113-120. https://doi.org/10.1080/0022250X.1972.9989806

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, *30*(2), 136-145.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215-239. https://doi.org/https://doi.org/10.1016/0378-8733 (78)90021-7

Harrigan, P., Daly, T. M., Coussement, K., Lee, J. A., Soutar, G. N., & Evers, U. (2021). Identifying influencers on social media. *International Journal of Information Management*, *56*, 102246. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2020.102246

Jain, S., & Sinha, A. (2020). Identification of influential users on Twitter: A novel weighted correlated influence measure for Covid-19. *Chaos, Solitons & Fractals*, *139*, 110037. https://doi.org/https://doi.org/10.1016/j.chaos.2020.110037

Joshi, P., & Mohammed, S. (2020). *Identifying Social Media Influencers using Graph Analytics*. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.12061662.v1

Li, Y.-M., Lai, C.-Y., & Chen, C.-W. (2011). Discovering influencers for marketing in the blogosphere. *Information Sciences*, *181*(23), 5143-5157. https://doi.org/https://doi.org/10.1016/j.ins.2011.07.023

Lü, L., Chen, D., Ren, X.-L., Zhang, Q.-M., Zhang, Y.-C., & Zhou, T. (2016). Vital nodes identification in complex networks. *Physics Reports*, *650*, 1-63. https://doi.org/https://doi.org/10.1016/j.physrep.2016.06.007

Nagmoti, R., Teredesai, A., & De Cock, M. (2010). Ranking approaches for microblog search. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*.

Noro, T., Ru, F., Xiao, F., & Tokuda, T. (2013). Twitter user rank using keyword search. *Information Modelling and Knowledge Bases XXIV. Frontiers in Artificial Intelligence and Applications*, *251*, 31-48 .

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web*.

Pal, A., & Counts, S. (2011). Identifying topical authorities in microblogs. *Proceedings of the fourth ACM international conference on Web search and data mining*.

Ramírez-de-la-Rosa, G., Villatoro-Tello, E., Jiménez-Salazar, H., & Sánchez-Sánchez, C. (2014). Towards automatic detection of user influence in twitter by means of stylistic and behavioral features. *Mexican International Conference on Artificial Intelligence.*

Ribeiro, M. H., Calais, P. H., Santos, Y. A., Almeida, V. A., & Meira Jr, W. (2018). Characterizing and detecting hateful users on twitter. *Twelfth international AAAI conference on web and social media.*

Ríos, S. A., Aguilera, F., Nuñez-Gonzalez, J .D., & Graña, M. (2019). Semantically enhanced network analysis for influencer identification in online social networks. *Neurocomputing*, *326-327*, 71-81. https://doi.org/https://doi.org/10.1016/j.neucom.2017.01.123

Riquelme, F., & González-Cantergiani, P .(2016) Measuring user influence on Twitter: A survey. *Information Processing & Management*, *52*(5), 949-975. https://doi.org/https://doi.org/10.1016/j.ipm.2016.04.003

Sun, Q., Wang, N., Zhou, Y., & Luo, Z. (2016). Identification of Influential Online Social Network Users Based on Multi-Features. *International Journal of Pattern Recognition and Artificial Intelligence*, *30*(06), 1659015. https://doi.org/10.1142/s0218001416590151

Wang, G., Jiang, W., Wu, J., & Xiong, Z. (2013). Fine-grained feature-based social influence evaluation in online social networks. *IEEE Transactions on parallel and distributed systems*, *25*(9), 2286-2296 .

Zareie, A., Sheikhahmadi, A., & Jalili, M. (2019). Identification of influential users in social networks based on users' interest. *Information Sciences*, *493*, 217-231. https://doi.org/https://doi.org/10.1016/j.ins.2019.04.033

Zengin Alp, Z., & Gündüz Öğüdücü, Ş. (2018). Identifying topical influencers on twitter based on user behavior and network topology. *Knowledge-Based Systems*, *141*, 211-221. https://doi.org/https://doi.org/10.1016/j.knosys.2017.11.021

Zhuang, Y.-B., Li, Z.-H., & Zhuang, Y.-J. (2021). Identification of influencers in online social networks: measuring influence considering multidimensional factors exploration. *Heliyon*, *7*(4), e06472. https://doi.org/https://doi.org/10.1016/j.heliyon.2021.e06472

# Identifying Influential Users in Social Networks: Measuring Multidimensional Influence

**Monireh Hosseini**[*1]
*Associate Prof., Department of Information Technology Engineering, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran*

**Kowsar Heidari**
*MSc. Student, Department of Information Technology Engineering, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran, Iran*

## Abstract

With the expansion and increasing application of social networks, the environment of these networks has been considered due to the possibility of accessing a large number of users and its impact on their thoughts and beliefs. By the great number of users on a social network, it is essential to detect the most influential ones for various purposes such as advertising to further disseminate information or curb the prevalence of negative public thoughts on the network. So far, different solutions have been created to recognize influential users, often through one-dimensional and simple criteria to indicate influential users. In the current study, we combine various criteria from three dimensions of content, activity, and network and propose a new measurement approach to identify influential users in the Twitter social network, with the aid of validation methods of spread, precision, recall, F1-Measure, and duplication percentage, we demonstrate that our approach performs more efficient than the existing methods in this scope.

**Keywords:** Influential users, Social networks, Identifying influential Users, Twitter.

1. Corresponding Author: hosseini@kntu.ac.ir