

تشخیص ترندهای سازمان‌دهی شده توییتری: یک دیدگاه ریاضیاتی

تاریخ دریافت: ۱۴۰۰/۰۳/۰۹

تاریخ پذیرش: ۱۴۰۰/۰۴/۱۴

کد مقاله: ۵۹۲۲۸

هادی صفری^{*۱}

چکیده

در دهه اخیر توییتر به یکی از تأثیرگذارترین شبکه‌های اجتماعی در سپهر سیاست ایران تبدیل شده است. ترندهای توییتری می‌توانند بازنمایی مناسبی از نظرات کنشگران و جریان‌سازان سیاسی کشور ارائه دهند. اما کمپین‌های رسانه‌ای از ابزارهایی استفاده می‌کنند تا با ترندهای سازمان‌یافته یا دستکاری‌شده، صدای خود را بلندتر از آنچه که در واقع هست نشان دهند؛ پس تشخیص ترندهای دستکاری‌شده در تشخیص و تحلیل نظرات واقعی کنشگران اهمیتی بسیار دارد. در این مقاله پس از دسته‌بندی ترندهای توییتری از نظر نوع رشد، روشی مبتنی بر توزیع سنجه مرکزیت اقتدار در گراف بازنشرها برای تشخیص ماهیت یک ترند توییتری ارائه می‌شود.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

واژگان کلیدی: توییتر، مرکزیت، شبکه‌های اجتماعی، گراف، ریاضیات کاربردی

رشد شبکه‌های اجتماعی مجازی در سال‌های اخیر انواع جدیدی از مشارکت سیاسی را امکان‌پذیر کرده است. انتخابات سال ۱۳۸۸ و وقایع پس از آن باعث شد شبکه‌های اجتماعی و به طور خاص توئیتر نقش ویژه‌ای در سپهر سیاسی ایران پیدا کنند (برنز و التام، ۲۰۰۹: ۲۹۸)؛ این رسانه‌های مجازی در سال‌های بعد نیز اهمیت خود را در بین کنشگران سیاسی ایران حفظ کرده‌اند (میره‌بیگی و عنایتی شبکلائی، ۱۳۹۶: ۶). در دهه اخیر، شبکه‌های اجتماعی مجازی در تحولات سیاسی سایر کشورهای خاورمیانه نیز بسیار تأثیرگذار بوده‌اند؛ چنان که انقلاب سال ۲۰۱۲ مصر به انقلاب فیسبوکی معروف شد.

با توجه به ضریب نفوذ نسبتاً پایین توئیتر در ایران برخی از تحلیل‌گران معتقدند نتایج بررسی این شبکه اجتماعی به کل جامعه تعمیم‌پذیر نیست و بررسی آن اهمیتی ندارد. در مقابل، گروهی دیگر از پژوهشگران معتقدند توئیتر در بین جریان‌سازان سیاسی و کنشگران مدنی بسیار رایج است و در نتیجه می‌تواند بازنمایی نسبتاً دقیقی از عقاید، نظرات و برنامه‌های این گروه ارائه دهد. میره‌بیگی و عنایتی شبکلائی (۱۳۹۶) معتقدند در انتخابات سال ۱۳۹۶ انواع مشارکت سیاسی در توئیتر فارسی مشاهده شده است. آن‌ها همچنین مزایایی برای کنشگری سیاسی در توئیتر نسبت به کنشگری در دنیای واقعی برشمرده‌اند (میره‌بیگی و عنایتی شبکلائی، ۱۳۹۶: ۴۴). با وجود این، پژوهش‌ها درباره شبکه‌های اجتماعی فارسی در ایران و ویژگی‌های خاص آن‌ها اندک است و بسیاری از پژوهش‌های موجود نیز فقط از دیدگاه جامعه‌شناختی یا بر اساس نظریه‌های علوم سیاسی این شبکه‌ها را بررسی کرده‌اند. پژوهش خضرائی (۲۰۱۹) درباره سپهر سیاسی توئیتر فارسی در انتخابات سال ۱۳۹۲ از معدود تحقیقاتی است که روش‌های نوین ریاضیاتی و علم شبکه را نیز در تحلیل توئیتر فارسی به کار گرفته است.

اگر اهمیت بررسی عقاید و نظرات کنشگران سیاسی توئیتر فارسی را بپذیریم، باید روش‌هایی برای تشخیص وزن و عیار واقعی هر دیدگاه و تعداد موافقین و مخالفین آن بیابیم. کمپین‌های رسانه‌ای در سراسر دنیا از روش‌های تبلیغاتی مختلفی استفاده می‌کنند تا صدای خود را در شبکه‌های اجتماعی بلندتر از آنچه واقعاً هست بنمایند. استفاده از حساب‌های خودکاری که به ربات^۱ توئیتری مشهورند یا گروه‌های سازمان‌یافته از افراد واقعی که به شکل حرفه‌ای موضوعات مختلف را ترند می‌کنند و هماهنگی برای آغاز همزمان ترندی توئیتری در یک زمان خاص برخی از روش‌های مورد استفاده کمپین‌ها است. در سال ۲۰۱۸ با حدود ۲۰۰ دلار می‌شد یک هشتگ توئیتری را در جهان عربی ترند کرد (بی‌بی‌سی ترندینگ، ۲۰۱۸). تصور می‌شود بیش از ۴۰٪ کاربران شبکه‌های اجتماعی در عربستان سعودی ربات باشند (بی‌بی‌سی ترندینگ، ۲۰۱۸)؛ هرچند محققین درباره چنین آمارهایی اتفاق نظر ندارند و برخی از آن‌ها اعداد کوچک‌تری را برای جهان عربی گزارش داده‌اند (جونز، ۲۰۱۹: ۱۳۹۹). نگارنده موفق به یافتن پژوهش مشابهی درباره توئیتر فارسی نشد. عیارسنجی جریان‌ها و ترندهای توئیتری نیازمند تشخیص چنین ابزارهای تبلیغاتی‌ای است.

دستکاری روند ترندهای توئیتری به طور معمول با سه ابزار صورت می‌گیرد (نیمو، ۲۰۱۹: ۷):

- ربات‌های توئیتری که گاهی در شبکه‌های بزرگی از ربات‌ها^۲ سازمان‌دهی می‌شوند
- سایبرگ‌ها^۳ یا حساب‌های کاربری نیمه‌خودکار که با برخی الگوهای رفتار انسانی سازگارند اما با تناوبی مانند حساب‌های خودکار توئیت می‌کنند
- حساب‌های کاربری انسانی

تشخیص حساب‌های جعلی، ربات‌های توئیتری و ترندهای دستکاری‌شده دشوار است. برخی از این ابزارهای تبلیغاتی ناقص قوانین توئیترند؛ توئیتر فعالانه بسیاری از حساب‌های درگیر در چنین فعالیت‌هایی را تشخیص می‌دهد و مسدود می‌کند. به همین دلیل، بسیاری از این کاربران به تناوب مسدود می‌شوند و حساب‌های جدید می‌سازند. همچنین، تغییر نام کاربری برای ردگم‌کنی در بین کاربران بسیار شایع است. امروزه ربات‌های توئیتری هوشمندتر شده‌اند و مرز بین ربات و انسان و واقعی و جعلی نیز محو و نامعلوم شده است (سرسکی، ۲۰۲۰: ۷). گروه دیگری از ابزارهای تبلیغاتی مذکور قوانین توئیتر را نقض نمی‌کنند و بیشتر به هماهنگی برای برگزاری یک تظاهرات در دنیای واقعی شباهت دارند؛ اما تشخیص چنین جریان‌هایی نیز در سنجش اهمیت واقعی جریان‌های توئیتری نقشی اساسی دارد.

هدف دستکاری ترندهای توئیتری آن است که با هزینه کم (تعداد کمی کاربر واقعی) تعداد زیادی پست مرتبط تولید شود؛ به عبارت دیگر، چنان وانمود شود که تعداد طرفداران یک ایده بسیار بیشتر از تعداد واقعی است. برای رسیدن به این هدف معمولاً از روش‌های مشابهی استفاده می‌شود (نیمو، ۲۰۱۹: ۸):

- هر حساب کاربری تعداد زیادی پست تولید می‌کند.
- هر حساب کاربری تعداد کمی پست تولید می‌کند که به دفعات زیادی بازنشر می‌شوند.

1 bot
2 botnet
3 cyborg

- تعداد زیادی ربات یا سایبرگ پست‌های تولیدشده را بازنشر می‌کنند.
 - تعداد زیادی ربات یا سایبرگ تعداد زیادی پست با محتوای یکسان تولید می‌کنند.
- بررسی‌ها نشان می‌دهد استفاده از روش‌های دستکاری یا سازمان‌دهی ترندهای توییتری در ایران نیز رایج شده است. در قسمت ترندهای سازمان‌دهی شده در توییتر فارسی توضیحات مفصل‌تر و شواهدی از این رخداد ارائه خواهد شد.
- در این مقاله روشی مبتنی بر بازنشرهای توییتهای یک ترند و توزیع سنجهٔ مرکزیت HITS^۲ برای دسته‌بندی ترندهای توییتری ارائه می‌شود. چنین روشی می‌تواند مورد استفادهٔ تحلیلگران شبکه‌های اجتماعی قرار گیرد تا نوع ترند را تشخیص دهند و اهمیت آن را بیش از مقدار واقعی ارزیابی نکنند.

۲- پیشینه پژوهش

- اولین تلاش‌ها برای تشخیص حساب‌های کاربری خودکار در شبکه‌های اجتماعی به سال ۲۰۱۰ بازمی‌گردد (سرسکی، ۲۰۲۰: ۴). بسیاری از روش‌های تشخیص ترندهای دستکاری‌شده در توییتر بر اساس روش‌های خودکار تشخیص ربات‌ها یا اسپم‌ها عمل می‌کنند. وو و همکاران این روش‌ها را به سه دستهٔ کلی تقسیم می‌کنند (وو و دیگران، ۲۰۱۸):
- **روش‌های مبتنی بر ساختار^۳**: این روش‌ها یا بر اساس بخش‌های مهم تویییت و حساب (مانند آدرس‌های اینترنتی موجود در تویییت یا الگوهای نام کاربری حساب) عمل می‌کنند یا از تحلیل محتوایی تویییت‌ها (مانند بررسی مجموعهٔ کلمات یا تکرار کلمات خاص) استفاده می‌نمایند.
 - **تحلیل ویژگی‌ها^۴**: این روش‌ها یا بر اساس اطلاعات آماری (مانند اطلاعات آماری فعالیت‌های حساب‌های کاربری یا کل کمپین) کار می‌کنند یا به تحلیل گراف‌های اجتماعی می‌پردازند.
 - **لیست سیاه^۵**: در این روش‌های از فهرست‌های از پیش آماده‌شدهٔ حساب‌های کاربری اسپم استفاده می‌شود. روش‌های مبتنی بر تحلیل گراف اجتماعی در سال‌های اخیر کمتر مورد توجه قرار گرفته‌اند. در این روش‌ها به طور معمول از شبکهٔ دنبال کردن حساب‌های توییتری موجود در یک مجموعه‌داده یا شرکت‌کنندگان در یک کمپین توییتری استفاده می‌شود. سانگ و دیگران (۲۰۱۱) از برخی معیارهای ساختاری گراف مانند فاصلهٔ رأس‌ها و همبندی^۶ گراف مستخرج از روابط دنبال کردن برای بهبود نتایج شبکه‌های عصبی پیش‌بینی‌کنندهٔ اسپم‌ها استفاده کرده است. گائو و دیگران (۲۰۱۲) از درجهٔ گره‌ها و اندازهٔ خوشه‌ها^۷ در کنار سنجه‌های غیرگرافی برای تشخیص اسپم‌ها استفاده کرده‌اند. گاش و دیگران (۲۰۱۲) نحوهٔ گسترش ارتباطات اسپم‌ها را در شبکهٔ دنبال کردن کاربران بررسی کرده‌اند. یانگ و دیگران (۲۰۱۳) از مرکزیت بینابینی^۸ در کنار چند سنجهٔ دیگر برای تشخیص حساب‌های اسپم استفاده کرده‌اند. با گسترش سریع تعداد حساب‌های توییتری، جمع‌آوری مجموعه‌داده‌های فراگیر از کاربران و روابطشان به کاری دشوار تبدیل شده است؛ همچنین، پردازش چنین گراف‌های بزرگی بسیار زمان‌گیر است و به‌خصوص برای تحلیل‌های بی‌درنگ^{۱۰} مناسب نیست (وو و دیگران، ۲۰۱۸: ۲۷۳).
- بسیاری از روش‌های موجود با هدف تشخیص حساب‌های اسپم و با دیدگاهی صرفویکی طراحی شده‌اند. یکی از کاربردهای اصلی این روش‌ها پردازش تقریباً بی‌درنگ داده‌ها و تشخیص حساب‌های کاربری ناقص قوانین است تا شبکهٔ اجتماعی زیرساخت کمپین بتواند آن‌ها را محدود کند یا تأثیر آن‌ها را در تشخیص ترندها نادیده بگیرد؛ هرچند برخی پژوهشگرانی به روش‌هایی برای تشخیص میزان دستکاری در روند طبیعی ترندها پرداخته‌اند. نیمو (۲۰۱۹) با ترکیب چند ویژگی آماری کمپین‌ها مانند تعداد بازنشرها، سهم پنجاه کاربر برتر از کل ترافیک و میانگین تعداد توییتهای کاربران، ضریب دستکاری جریان را برای ارزیابی یک ترند توییتری معرفی کرده است.
- بر خلاف روش‌های رایج، در روش پیشنهادی هدف تشخیص ماهیت یک یا چند حساب کاربری خاص نیست که ممکن است سعی کنند یک ترند را منحرف کنند یا در آن تأثیر بگذارند؛ بلکه تشخیص ماهیت کلی ترند اهمیت دارد. در نتیجه، استفاده از مرکزیت حساب‌های کاربری مختلف و مقایسهٔ آن‌ها (مانند روش یانگ و دیگران (۲۰۱۳)) کافی نیست.

- 1 retweet
- 2 hyperlink-induced topic search (HITS)
- 3 syntax analysis
- 4 feature analysis
- 5 black list
- 6 distance
- 7 connectivity
- 8 cluster
- 9 betweenness centrality
- 10 real-time

روش پیشنهادی را می‌توان مطابق دسته‌بندی وو و دیگران (۲۰۱۸) به‌نوعی جزء روش‌های تحلیل ویژگی‌ها و به‌طور خاص‌تر روش‌های مبتنی بر گراف‌های اجتماعی دانست؛ هرچند مقایسه مستقیم این روش با روش‌های رایج فهرست‌شده، به دلیل ماهیت و کاربرد متفاوت این روش، ممکن نیست. برای کاربرد پیشنهادشده تشکیل شبکه حساب‌های شرکت‌کننده در یک ترند خاص کافی است و نیازی به تشکیل مجموعه داده‌های فراگیر بزرگ نیست. همچنین، در مقایسه با کاربردهای بی‌درنگ، منابع زمانی بیشتری در اختیار است؛ بنابراین دشواری‌های اساسی روش‌های مبتنی بر گراف اجتماعی در روش پیشنهادی مشکل‌ساز نیستند.

۳- پیش‌زمینه علمی

۳-۱- الگوریتم مرکزیت HITS

سنجه‌های مرکزیت معیاری ریاضی برای سنجش اهمیت یک رأس در گراف هستند. الگوریتم HITS را کلینبرگ (۱۹۹۹) بر اساس این ایده معرفی کرد که اهمیت هر رأس در گراف وابسته به رأس‌های همسایه است. این الگوریتم در ابتدا به عنوان روشی برای رتبه‌بندی نتایج جست‌وجوهای اینترنتی بر اساس ارجاعات بین صفحات اینترنتی طراحی شد اما بعدها کاربردهای متنوعی پیدا کرد.

در این الگوریتم برای هر رأس دو سنجه مرکزیت اقتدار^۱ (x_i) و محوریت^۲ (y_i) تعریف می‌شود. اگر رأس‌ها نماینده صفحات صفحات اینترنتی باشند، یک صفحه مقتدر در اینترنت صفحه‌ای است که اطلاعات مفیدی دارد و یک محور صفحه‌ای است که نشان می‌دهد اطلاعات مفید را از کدام صفحه‌ها می‌توان دریافت کرد. به‌طور دقیق‌تر، یک رأس با محوریت بالا رأسی است که به رأس‌هایی با اقتدار بالا اشاره می‌کند و یک رأس با اقتدار بالا نیز رأسی است که رأس‌هایی با محوریت بالا به آن اشاره می‌کنند. به شکل نظری، اقتدار و مرکزیت به کل مرحله‌به‌مرحله و مکرر محاسبه می‌شود. در ابتدا اقتدار و محوریت همه رأس‌ها برابر ۱ در نظر گرفته می‌شود و در هر مرحله این دو سنجه با روابط زیر به‌روزرسانی می‌شوند:

$$x_i = \alpha \sum_j A_{ij} y_j, \quad y_i = \beta \sum_j A_{ji} x_j \quad (1)$$

که در آن A_{ij} درایه سطر i -ام و ستون j -ام ماتریس مجاورت و نشانگر وجود یال از رأس i به رأس j و α و β اعدادی ثابت هستند. پیاده‌سازی عملی HITS با استفاده از روابط جبر خطی در کتاب «مقدمه‌ای بر شبکه‌ها» نیومن (۲۰۱۰) توصیف شده است.

۴- روش‌شناسی پژوهش

با بررسی فضای توپیتر فارسی و ترندهای آن، چند هشتگ توپیتتری ماه‌های اخیر انتخاب شد. سعی شد تا هشتگ‌هایی انتخاب شود که انواع مختلف ترندهای توپیتر فارسی را پوشش دهند. در صورت موجود بودن داده‌های قدیمی، از داده‌هایی که نگارنده در زمان انتشار هر یک از هشتگ‌ها جهت تحلیل آن‌ها جمع‌آوری کرده بود استفاده شد. در غیر این صورت همه توپیت‌های در دسترس حاوی آن هشتگ از واسط برنامه‌نویسی کاربردی^۳ توپیتر دریافت شد. پس از جمع‌آوری توپیت‌های منتشرشده با هر هشتگ، ویژگی‌های آن‌ها (از جمله تعداد بازنشرها، تعداد کاربران، تعداد توپیت‌های هر کاربر و...) بررسی گردید تا ماهیت ترند مشخص شود. در ادامه، بازنشرهای همه توپیت‌های هر یک از هشتگ‌ها از واسط برنامه‌نویسی کاربردی توپیتر دریافت شد. به دلیل محدودیت‌های توپیتر، حداکثر اطلاعات ۱۰۰ بازنشر آخر هر توپیت جمع‌آوری گردید؛ اما با توجه به تعداد اندک توپیت‌های با بیش از ۱۰۰ بازنشر در مجموعه داده مورد بررسی، این مسئله تأثیر چندانی بر نتیجه‌گیری‌ها و مراحل بعدی کار نداشت. با این اطلاعات گراف بازنشرهای میان کاربران تشکیل شد. به ازای هر کاربر منتشرکننده توپیتی که حداقل یک بار بازنشر شده است و هر کاربر بازنشرکننده یک رأس در نظر گرفته شد. ابتدا به ازای هر توپیت کاربر D که کاربر S آن را بازنشر کرده بود، یک یال جهت‌دار از رأس نظیر S به رأس نظیر D رسم شد. سپس با ادغام یال‌های چندگانه، یک گراف وزن‌دار جهت‌دار تشکیل شد که در آن وزن یال S به D نشان‌گر تعداد توپیت‌هایی است که D منتشر و S بازنشر کرده است. چنین گرافی را در ادامه گراف بازنشرهای هشتگ می‌نامیم.

سپس با روش کلینبرگ (۱۹۹۹) مرکزیت HITS و سنجه‌های اقتدار و محوریت هر رأس گراف بازنشرها محاسبه شد. در مرحله بعد رأس‌هایی که فقط به بازنشر توپیت‌های دیگران پرداخته بودند و خودشان توپیتی منتشر نکرده بودند یا توپیت‌هایشان هرگز بازنشر نشده بود (رأس‌هایی با درجه ورودی ۰) از گراف حذف شدند تا توزیع اقتدار رأس‌ها فقط نمایانگر اقتدار رأس‌هایی باشد که توپیت‌هایشان بازنشر شده است و در مقدار حداقل اقتدار نوپز نداشته باشد.

1 authority

2 hub

3 Application Programming Interface (API)

نمودار توزیع اقتدار گره‌های باقی‌مانده با دسته‌بندی‌هایی به طول ۰.۰۱ رسم شد. ترندها بر اساس نوع توزیع سنجۀ اقتدار به سه دسته تقسیم شدند. ویژگی‌های این دسته‌ها در ادامه مقاله بیان خواهد شد.

۵- ترندهای سازمان‌دهی‌شده در توئیتر فارسی

مجاهدین خلق از سال‌های آغازین توئیتر فارسی در این شبکه اجتماعی حضور داشته‌اند. در بسیاری از تحلیل‌های دانشگاهی و غیردانشگاهی توئیتر فارسی [مانند (خضرائی، ۲۰۱۹: ۱۰)]، دسته‌ای نسبتاً مجزا از کاربران مرتبط به این گروه دیده می‌شود. حساب کاربری حشمت علوی - که به دلیل هویت جعلی مدتی تعلیق شده بود - یکی از حساب‌های کاربری سازمانی این گروه است که به دلیل پخش اخبار نادرست و اطلاعات گمراه‌کننده مشهور است. این حساب کاربری در بسیاری از ترندهای مرتبط با مجاهدین خلق نیز شرکت کرده است (حسینی، ۲۰۱۹). اصطلاح «ربات‌های آلبانیایی» سال‌ها است که در فضای رسانه‌ای ایران برای اشاره به حساب‌های کاربری سازمانی این گروه به کار می‌رود.

در ۲۰ اسفند ۱۳۹۹، سید علی‌اکبر مطهری، کاربر انقلابی توئیتر، در توئیتی^۱ با عنوان «طریقه داغ کردن خبر در توئیتر» ویدیویی را منتشر کرد که به نظر می‌رسید بخشی از یک ویدیوی آموزشی جنگ نرم و فضای مجازی متعلق به یک نهاد انقلابی باشد. او چند ساعت بعد این توئیت را حذف کرد؛ اما در همین مدت ویدیوی موجود در توئیت^۲ بارها در توئیتر و تلگرام بازنشر شده بود. در این ویدیو یک نرم‌افزار تحت وب نشان داده می‌شود که کاربر می‌تواند با استفاده از آن همزمان برای اخبار خبرگزاری فارس پاسخ بنویسد و آن پاسخ را در توئیتر نیز منتشر کند. همچنین، این نرم‌افزار به گروه‌های مختلفی از حساب‌های توئیتری با گرایش‌های سیاسی متفاوت متصل است و کاربر نرم‌افزار می‌تواند از آن‌ها برای «عملیات لایک‌ریزی» استفاده کند. در ویدیوی منتشرشده نام ۱۹ حساب کاربری توئیتر دیده می‌شود. نگارنده بسیاری از این کاربران را چند ماه پیش از انتشار ویدیو در بررسی هشتگ سازمان‌یافته #قه‌رمان_خانواده مشاهده کرده بود. به نظر می‌رسد این کاربران را مطابق تقسیم‌بندی نیمو (۲۰۱۹) می‌توان سایبرگ به حساب آورد. بسیاری از این کاربران در زمان نگارش مقاله، با تغییر نام کاربری، همچنان در توئیتر فعالند و برخی نیز معلق شده‌اند.

گروه‌های مختلف سیاسی و غیرسیاسی از روش‌های مختلفی برای ترند کردن سازمان‌یافته هشتگ‌ها و موضوعات یا دستکاری ترندها استفاده می‌کنند. با بررسی هشتگ‌های مختلف سال‌های اخیر توئیتر فارسی، به نظر می‌رسد در ایران استفاده از سایبرگ‌ها و نیز سازماندهی کاربران واقعی رایج‌ترین ابزارهای دستکاری ترندهای توئیتری هستند. سازماندهی کاربران واقعی در توئیتر (قرار گذاشتن برای آغاز هشتگ در زمانی خاص) و گروه‌های تلگرامی و واتس‌آپی (همه‌انگهی برای آغاز هشتگ در زمانی خاص و ارائه محتوایی که همه اعضای گروه با حساب‌های خودشان آن‌ها را توئیت می‌کنند) صورت می‌گیرد. تصاویری از گروه‌های مذکور و نیز تصاویری از توئیتهای کاملاً مشابه انبوهی از کاربران به‌ظاهر بی‌ارتباط بارها در توئیتر فارسی منتشر شده است. از بین روش‌هایی که نیمو (۲۰۱۹) تشریح کرده است، در ترندهای توئیتری دستکاری‌شده یا سازمان‌یافته توئیتر فارسی استفاده از تعدادی کاربر برای بازنشر انبوه تعداد محدودی توئیت (روش ۲)، بازنشر کور توئیتهایی که با یک هشتگ خاص نوشته شده‌اند (روش ۳) و در برخی موارد انتشار تعداد زیادی پست با محتوای یکسان (روش ۴) بیشتر استفاده می‌شوند. انتشار انبوه تعداد زیادی پست از طرف تعداد محدودی کاربر (روش ۱) در توئیتر فارسی چندان رایج نیست؛ با آن که این روش در جهان عرب استفاده می‌شود.

جونز (۲۰۱۹) معتقد است در جهان عرب استفاده از ربات‌ها بسیار رایج است (جونز، ۲۰۱۹: ۱۴۰۹). او در بررسی‌های بعدی خود (جونز، ۲۰۲۱) هشتگ‌هایی را مشاهده کرده است که در آن‌ها انبوهی پست با محتوای بی‌ارتباط به هشتگ منتشر شده است.

بر این اساس ترندهای توئیتر فارسی را می‌توان به سه دسته تقسیم کرد:

- ترندهای طبیعی که در آن ترند بدون همه‌انگهی قبلی به تدریج رشد می‌کند و گسترش می‌یابد.
- ترندهای همه‌انگهی شده که در آن گروه کوچک یا بزرگی از کاربران با همه‌انگهی قبلی و با هدف ترند کردن یک موضوع یا هشتگ به طور همزمان تعدادی توئیت منتشر می‌کنند.
- ترندهای سازمان‌یافته که در آن‌ها گروه کوچکی از کاربران با هدف ترند کردن سریع یک موضوع یا هشتگ محدودی توئیت منتشر می‌کنند که بارها بازنشر می‌شود.

نمونه‌هایی از انواع هشتگ‌ها و ترندهای توئیتر فارسی در ادامه بررسی خواهد شد.

1 https://twitter.com/Ali_A_Motahari/status/1365948615223517184

2 <https://youtu.be/3AEYcG45Y1c>

۶- دسته‌بندی ترندهای تویتری

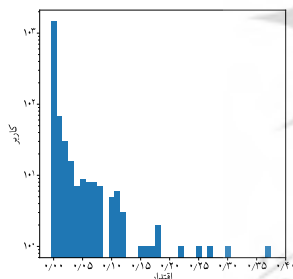
توییت‌های مرتبط با ۹ هشتگ تویتری ماه‌های اخیر تویتر فارسی با روش پیشنهادی پیشنهادی تحلیل شد. مشخصات این هشتگ‌ها در جدول ۱ آمده است. در شمارش کاربران و توییت‌ها فقط توییت‌های اصیل (شامل پاسخ‌ها) در نظر گرفته شده است و بازنشرها شمرده نشده‌اند. هشتگ #ظریف_تنها_نیست یک بار نیز پیش از این ترند شده بود اما فقط توییت‌های مرتبط با ترند سال ۱۴۰۰ در نظر گرفته شده‌اند.

جدول ۱- مشخصات هشتگ‌های تحلیل شده

هشتگ	زمان انتشار	تعداد توییت	تعداد کاربران
#جان_پدر_کجاستی	۱۲ تا ۱۹ آبان ۱۳۹۹	۷,۶۴۸	۵,۵۱۵
#بورس_محکم	۱۲ تا ۱۷ شهریور ۱۳۹۹	۱,۳۸۲	۷۶۶
#ظریف_تنها_نیست	۵ تا ۱۵ اردیبهشت ۱۴۰۰	۸۶۶	۴۴۶
#من_یک_رزیدتم	۱۵ تا ۱۷ اردیبهشت ۱۴۰۰	۵۲۶	۳۱۱
#بورس_سپرتوری	۱۰ تا ۱۴ اردیبهشت ۱۴۰۰	۲,۸۸۲	۱,۰۴۷
#بورس_مظلوم	۵ تا ۱۱ اردیبهشت ۱۴۰۰	۲,۷۰۷	۶۲۲
#شور_دارالعباده	۲۶ تا ۳۰ اردیبهشت ۱۴۰۰	۹,۳۰۰	۳۷۳
#قهرمان_خانواده	۱۴ تا ۱۶ دی ۱۳۹۹	۱۲۵	۸۲
#ته_به_ضدایرانی	۷ فروردین ۱۴۰۰	۵۱	۳۴

۱-۶- ترندهای طبیعی

چنانچه یک ترند تویتری رشدی طبیعی داشته باشد، انتظار می‌رود مانند بسیاری از پدیده‌های اجتماعی از توزیع دم‌بلند^۱ و قانون توانی^۲ پیروی کند؛ به این شکل که برای مقادیر بزرگ‌تر از یک حد آستانه $(x > x_{min})$ رابطه $p(x) = C x^{-\alpha}$ برقرار باشد. کتاب «آشفستگی سیاسی» مارگتس و دیگران (۲۰۱۶) توضیحاتی دربارهٔ چنین توزیع‌هایی در شبکه‌های اجتماعی ارائه می‌دهد.

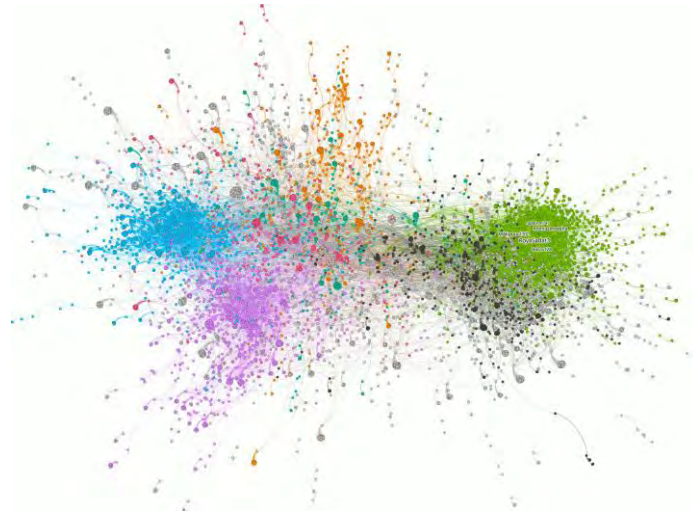


شکل ۱- توزیع اقتدار کاربران هشتگ #جان_پدر_کجاستی

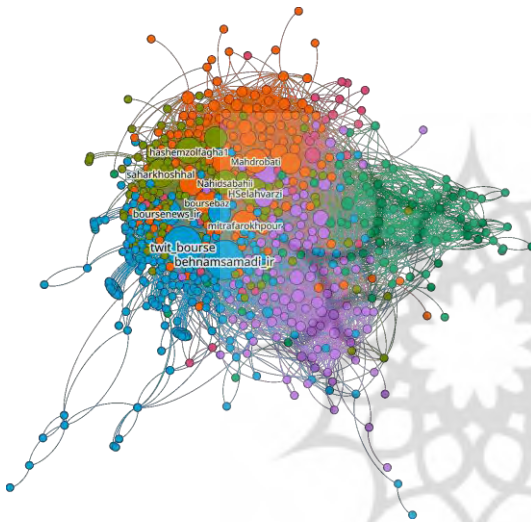
جزئیات ریاضیاتی مفصل‌تر این توزیع نیز در کتاب «مقدمه‌ای بر شبکه‌ها» نیومن (۲۰۱۰) در دسترس است. در صورت پیروی از چنین توزیعی اصل پارتو^۳ برقرار خواهد بود؛ به بیان ساده‌تر، تعداد کمی از کاربران بخش زیادی از اقتدار شبکه را در اختیار خواهند داشت. همچنین معمولاً بیشینهٔ اقتدار رأس‌ها در این گروه از ترندها بیش از سایر گروه‌ها است.

چنین ویژگی‌هایی در هشتگ‌های #جان_پدر_کجاستی (شکل ۱ و شکل ۲)، #بورس_محکم (شکل ۳ و شکل ۴) و #ظریف_تنها_نیست (شکل ۵ و شکل ۶) دیده می‌شود. با افزایش اقتدار در هر یک از این هشتگ‌ها، تعداد کاربران با روندی مشخص کاهش می‌یابد.

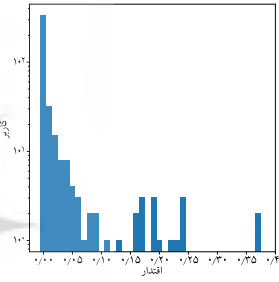
1 long tail distribution
2 power law
3 Pareto principle



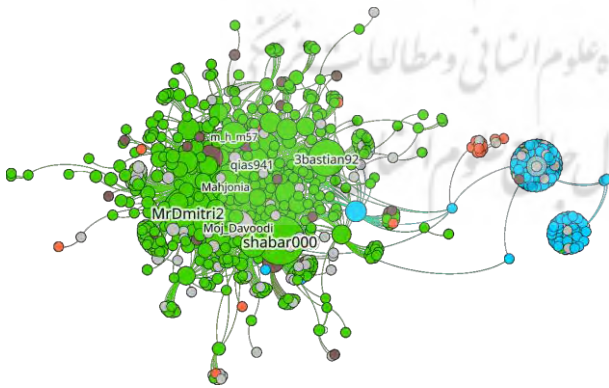
شکل ۲- بزرگ‌ترین مؤلفه همبندی هشنگ #جان_پدر_کجاستی



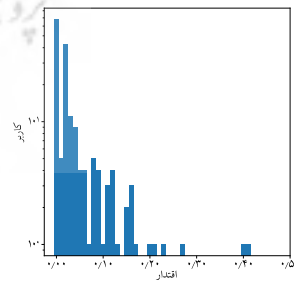
شکل ۴- بزرگ‌ترین مؤلفه همبندی هشنگ #بورس_محکم



شکل ۳- توزیع اقتدار کاربران هشنگ #بورس_محکم



شکل ۶- بزرگ‌ترین مؤلفه همبندی هشنگ #ظریف_تنها_نیست



شکل ۵- توزیع اقتدار کاربران هشنگ #ظریف_تنها_نیست

۲-۶- ترندهای هماهنگ‌شده

برخی از ترندها با هماهنگی قبلی و با هدف ترند کردن یک موضوع در توییتر ایجاد می‌شوند. کاربران هشتگی را مشخص می‌کنند تا در یک ساعت خاص به نوشتن درباره آن و بازنشر توییت‌های مرتبط پردازند. چنین فعالیت‌هایی ناقص قوانین توییتر

نیست و بیشتر می‌توان آن را شبیه هماهنگی برای برگزاری یک تظاهرات در دنیای واقعی دانست؛ اما به هر حال باعث می‌شود تعداد موافقین آن بیش از تعداد واقعی به نظر بیاید و در تحلیل این ترندها باید به نکته مذکور توجه کرد.

معمولاً در این ترندها میانگین تعداد بازنشرهای توییت بیشتر از ترندهای طبیعی است. الگوی بازنشرها توزیع اقتدار گراف بازنشرها را تحت تأثیر قرار می‌دهد و باعث می‌شود روند کاهش تعداد کاربران با افزایش اقتدار تا یک حد آستانه کندتر از ترندهای دارای رشد طبیعی شود و سپس ناگهان تعداد کاربران کاهش یابد. به بیان دیگر، تعداد کاربران نسبتاً مقتدر در این نوع ترندها بسیار بیشتر از ترندهای طبیعی است.

هشتگ #شور_دارالعباده (شکل ۷ و شکل ۸) یکی از هشتگ‌هایی بود که یک گروه حدوداً سیصدنفره از کاربران با تعداد بسیار زیادی توییت و بازنشر به همین روش ترند کردند. هشتگ‌های #بوس_مظلوم (شکل ۹ و شکل ۱۰) و #بوس_سپرتومی (شکل ۱۱ و شکل ۱۲) را گروه‌هایی از معامله‌گران خرد بازار سرمایه پس از هماهنگی در کانال‌ها و گروه‌های تلگرامی ترند کردند. هشتگ #بوس_مظلوم پس از دعوت کانال و گروه تلگرامی کمپین اعتراضات بوس^۱ آغاز شد. هشتگ #بوس_سپرتومی نیز با دعوت یکی از فعالین رسانه‌ای بوسی در کانال تلگرامی اش^۲ و حمایت کمپین اعتراضات بوس^۳ ترند شد. هشتگ #من_یک_رزیدنتم (شکل ۱۳ و شکل ۱۴) با دعوت کانال تلگرامی صدای حقوق بالینی^۴ آغاز شد؛ هرچند بعدتر با پیوستن تدریجی تعدادی دیگر از پزشکان و مردم عادی، به شکل نسبتاً طبیعی رشد کرد. توزیع اقتدار رأس‌ها در گراف بازنشرهای این هشتگ نیز علاوه بر برخی ویژگی‌های ذکرشده برای ترندهای هماهنگ‌شده، شباهت‌هایی به ترندهای طبیعی دارد.



شکل ۷- توزیع اقتدار کاربران هشتگ #شور_دارالعباده

شکل ۸- بزرگ‌ترین مؤلفه همبندی هشتگ #شور_دارالعباده

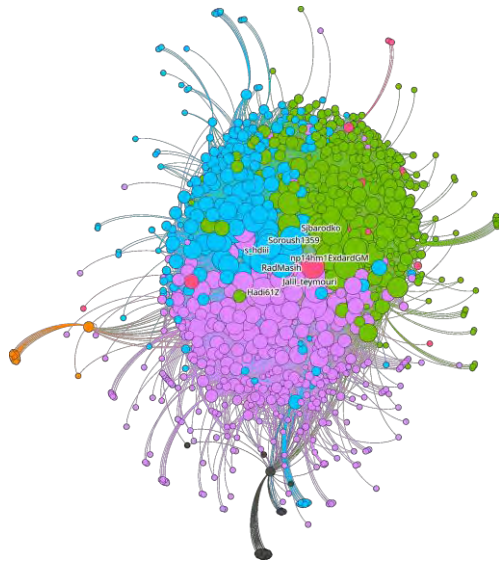
پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1 <https://t.me/boursecampaign1/2518>

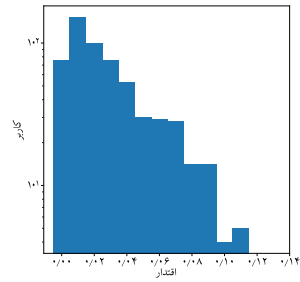
2 https://t.me/BehnamSamadi_ir/2759

3 <https://t.me/boursecampaign1/2450>

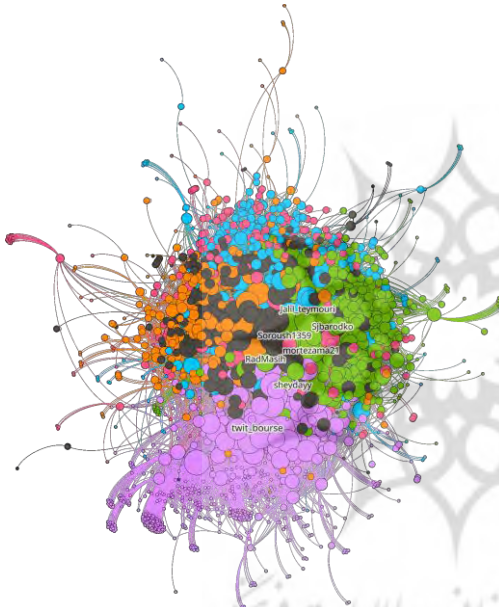
4 <https://t.me/lawmed/24124>



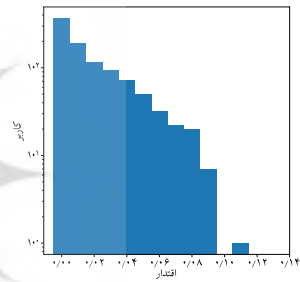
شکل ۱۰- بزرگ‌ترین مؤلفه همبندی هشگ #بورس_مظلوم



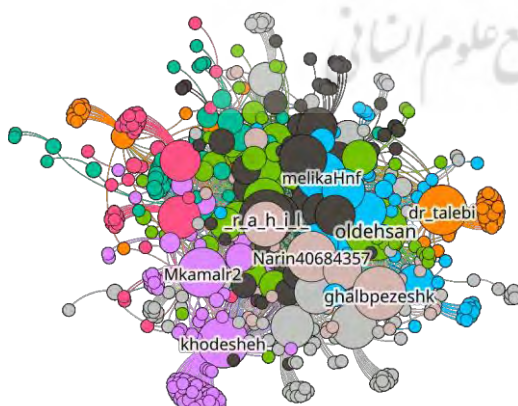
مظلوم_شکل ۹- توزیع اقتدار کاربران هشگ #بورس



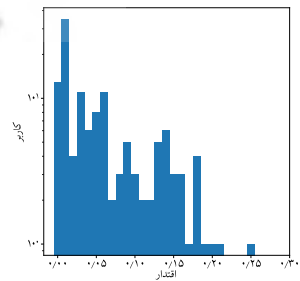
شکل ۱۲- بزرگ‌ترین مؤلفه همبندی هشگ #بورس_سپرتورمی



شکل ۱۱- توزیع اقتدار کاربران هشگ #بورس_سپرتورمی



شکل ۱۴- بزرگ‌ترین مؤلفه همبندی هشگ #من_یک_رزیدنتم

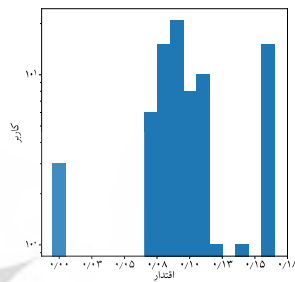
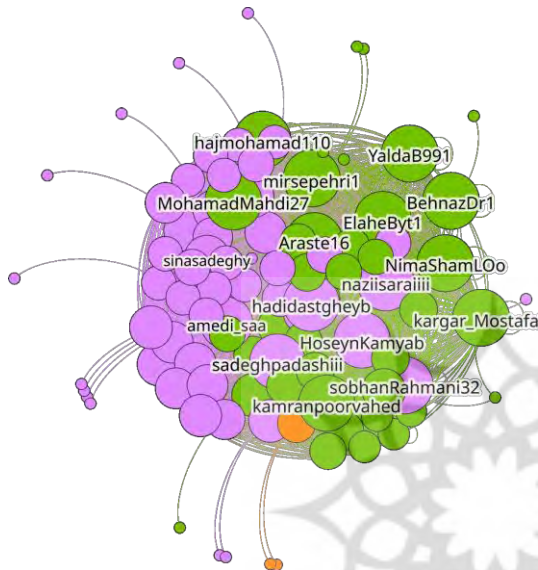


شکل ۱۳- توزیع اقتدار کاربران هشگ #من_یک_رزیدنتم

۳-۶- ترندهای سازمان یافته

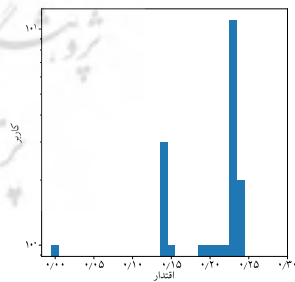
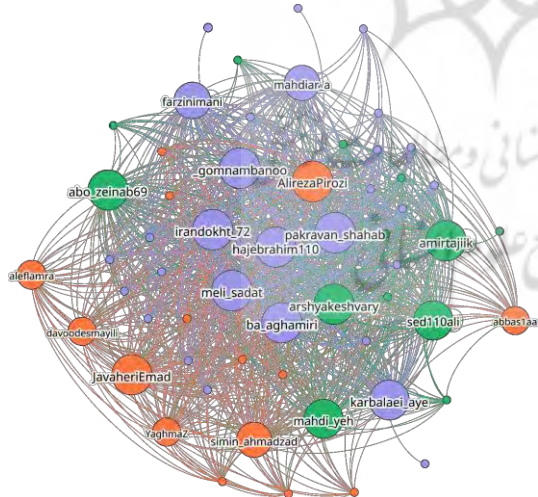
در این ترندها گروه کوچکی از کاربران که معمولاً هم را می‌شناسند و یک باند تشکیل می‌دهند، با انتشار تعدادی توییت و بازنشر فراوان توییت‌های همدیگر، به سرعت یک ترند توییتی می‌سازند. در این ترندها همه کاربران تقریباً همه توییت‌های همدیگر را بازنشر می‌کنند؛ به همین دلیل توزیع اقتدار رأس‌های گراف بازنشرها به کلی به هم می‌ریزد و با حذف بخش‌های میانی توزیع، معدودی رأس با اقتدار بسیار پایین و تعداد زیادی رأس با اقتدار بالا باقی می‌ماند.

هشتگ‌های #قهрман_خانواده (شکل ۱۵ و شکل ۱۶) و #نه_به_ضدایرانی (شکل ۱۷ و شکل ۱۸) را یک گروه از کاربران که به شکل تخصصی در حوزه دستکاری ترندهای توییتی فعالیت می‌کنند ایجاد کرده‌اند. به دلیل تغییر متناوب نام کاربری این کاربران، اسامی یکسان نیست. همچنین به دلیل تعلیق شدن مکرر این کاربران، تعداد کاربران شرکت‌کننده در هشتگ‌ها نوسانات زیادی دارد.



شکل ۱۶- بزرگ‌ترین مؤلفه همبندی هشتگ #قهрман_خانواده

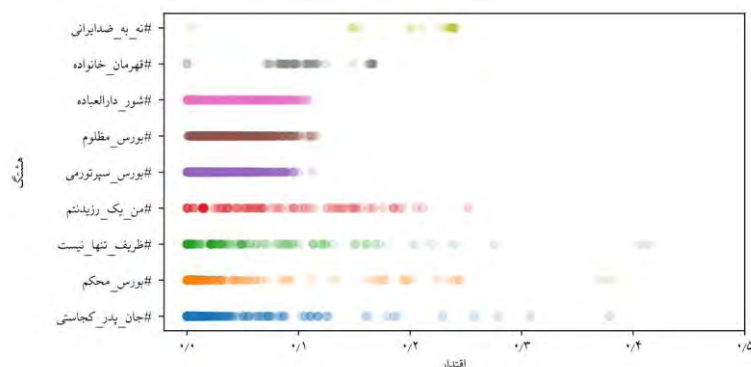
شکل ۱۵- توزیع اقتدار کاربران هشتگ #قهрман_خانواده



شکل ۱۸- بزرگ‌ترین مؤلفه همبندی هشتگ #نه_به_ضدایرانی

شکل ۱۷- توزیع اقتدار کاربران هشتگ #نه_به_ضدایرانی

شکل ۱۹ اقتدار رأس‌های گراف بازنشرهای هشتگ‌های بررسی شده را در کنار هم نشان می‌دهد. تفاوت الگوی توزیع اقتدار رأس‌ها در بین هشتگ‌های دسته‌های مختلف مبتنی بر روش پیشنهادی در این نمودار مشخص است.



شکل ۱۹- اقتدار رأس‌های گراف بازنشرها

۷- نتیجه‌گیری

در مقاله حاضر پس از بررسی اهمیت تشخیص ماهیت ترندهای توییتری و دسته‌بندی آن‌ها برای سنجش تعداد واقعی موافقین نظرات مختلف و تحلیل عقاید کنشگران سیاسی فضای مجازی و توییتر فارسی، سه گروه ترندهای طبیعی، ترندهای هماهنگ‌شده و ترندهای سازمان‌یافته از هشتگ‌های ماه‌های اخیر توییتر فارسی معرفی شد. در ادامه، روشی برای دسته‌بندی ترندهای توییتر فارسی با استفاده از توزیع سنجه مرکزیت اقتدار رأس‌ها در گراف بازنشرها ارائه گردید.

روش ارائه‌شده به دلیل وابستگی به گراف بازنشرها - به جای گراف روابط دنبال کردن - به نسبت کم‌هزینه است و برای هر ترند بدون نیاز به دسترسی به مجموعه داده‌های بزرگ از کاربران توییتری قابل استفاده است. همچنین، سنجه مرکزیت اقتدار با الگوریتم HITS محاسبه می‌شود که الگوریتمی شناخته‌شده است و بسیاری از ابزارهای تحلیل شبکه‌های اجتماعی امکان محاسبه آن را فراهم می‌کنند. این ویژگی‌ها باعث سهولت استفاده تحلیلگران شبکه‌های اجتماعی که به منابع محاسباتی قدرتمند دسترسی ندارند از روش پیشنهادی می‌گردد.

منابع

- میره‌بیگی، سید وحید و علی عنایتی شبکلائی، (۱۳۹۶)، «کنش سیاسی در توییتر: بررسی توییتهای سیاسی در جریان انتخابات ریاست جمهوری ایران در سال ۱۳۹۶»، مطالعات انتخابات، شماره ۱۴ و ۱۵، صص ۵ - ۴۸.
- BBC Trending (2018). How much to fake a trend on Twitter? in one country, about £150. BBC News. Available at <https://www.bbc.com/news/blogs-trending-43218939>. Accessed 21 May 2021.
- Burns, A. and B. Eltham (2009, November). Twitter free Iran: an evaluation of Twitter 's role in public diplomacy and information operations in Iran 's 2009 election crisis. In Communications Policy & Research Forum 2009, pp. 322-334.
- Cresci, S. (2020). A decade of social bot detection. Communications of the ACM 63(10), 72-83.
- Gao, H., Y. Chen, K. Lee, D. Palsetia, and A. N. Choudhary (2012). Towards online spam filtering in social networks. In NDSS, Volume 12, pp. 1-16.
- Ghosh, S., B. Viswanath, F. Kooti, N. K. Sharma, G. Korum, F. Benevenuto, N. Ganguly, and K. P. Gummadi (2012). Understanding and combating link farming in the Twitter social network. In Proceedings of the 21st International Conference on World Wide Web, WWW '12, New York, NY, USA, pp. 61-70. Association for Computing Machinery.
- Hussain, M. (2019). An Iranian activist wrote dozens of articles for right-wing outlets. but is he a real person? The Intercept. Available at <https://theintercept.com/2019/06/09/heshmat-alavi-fake-iran-mek>. Accessed 21 May 2021.
- Jones, M. (2019). Propaganda, fake news, and fake trends: The weaponization of twitter bots in the gulf crisis. International Journal of Communication 13.(۰)
- Jones, M. (2021). Tweet. Available at <https://twitter.com/marcowenjones/status/1379806742666752002>. Accessed 21 May 2021.
- Khazraee, E. (2019). Mapping the political landscape of persian twitter: The case of 2013 presidential election. Big Data & Society 6(1), 2053951719835232.
- Kleinberg, J. M. (1999, September). Authoritative sources in a hyperlinked environment. J. ACM 46(5), 604-632.
- Margetts, H., P. John, S. Hale, and T. Yasseri (2016). Political Turbulence. Princeton, New Jersey: Princeton University Press.

13. Newman, M. E. J. (2010). Networks: An Introduction. OUP Oxford.
14. Nimmo, B. (2019). Measuring traffic manipulation on Twitter. tech. rep., Oxford: Project on Computational Propaganda. Working Paper 2019.1. Available at <https://comprop.oii.ox.ac.uk>.
15. Song, J., S. Lee, and J. Kim (2011). Spam filtering in Twitter using sender-receiver relationship. In R. Sommer, D. Balzarotti, and G. Maier (Eds.), Recent Advances in Intrusion Detection, Berlin, Heidelberg, pp. 301–317. Springer Berlin Heidelberg.
16. Wu, T., S. Wen, Y. Xiang, and W. Zhou (2018). Twitter spam detection: Survey of new approaches and comparative study. Computers & Security 76, 265–284.
17. Yang, C., R. Harkreader, and G. Gu (2013). Empirical evaluation and new design for fighting evolving Twitter spammers. IEEE Transactions on Information Forensics and Security 8(8), 1280–1293.

