Rumor Detection and Prevention in Social Networks

نگارندگان:

محمد قهری و محمد قاسمی اریان

استاد مربوطه:

دکتر هاله امین طوسی

بهمن ۱۴۰۱


بسم الله الرحمن الرحيم

فهرست مطالب

فهرست تصاویر

چکیده

شایعات یک موضوع مهم در رسانه‌های اجتماعی هستند، زیرا می‌توانند به سرعت پخش شوند و باعث سردرگمی و وحشت شوند. چالش این است که شایعات را قبل از فراگیر شدن زودهنگام شناسایی کنیم. برای انجام این کار، ما باید پویایی انتشار شایعات را درک کنیم و روش‌هایی را برای شناسایی آن‌ها در زمان واقعی توسعه دهیم. ما همچنین باید استراتژی‌هایی را برای جلوگیری از انتشار شایعات در وهله‌ی اول توسعه دهیم. این گزارش در مورد چگونگی استفاده از تکنیک‌های هوش مصنوعی برای شناسایی و جلوگیری از شایعات در پتلفرم‌های رسانه‌های اجتماعی بحث خواهد کرد. تکنیک‌های موجود برای تشخیص شایعات مانند الگوریتم‌های پردازش زبان طبیعی (NLP) و یادگیری ماشین (ML) و همچنین رویکردهای بالقوه برای جلوگیری از جلب توجه شایعات در وهله‌ی اول را بررسی می‌کند.

۱ . مقدمه

اشتراک‌گذاری اطلاعات نادرست در شبکه‌های اجتماعی - از جمله، اما نه محدود به، "اخبار جعلی" سیاسیِ آشکارا دروغ و محتوای فراحزبی گمراه‌کننده - در سال‌های اخیر به کانون اصلی بحث‌های عمومی‌و مطالعات آکادمیک تبدیل شده است. اگرچه اطلاعات نادرست چیز جدیدی نیست، اما این موضوع در سال ۲۰۱۶ پس از انتخابات ریاست جمهوری ایالات متحده و رفراندوم برگزیت بریتانیا، که طی آن داستان های کاملا ساختگی (که به عنوان اخبار مشروع ارائه می‌شوند) از طریق شبکه‌های اجتماعی توزیع گسترده ای داشتند، اهمیت پیدا کرد - مشکلی که در طول این مدت توجه بیشتری را به خود جلب کرد. همه‌گیری COVID-19 – و یورش به ساختمان کنگره پس از انتخابات ریاست‌جمهوری ۲۰۲۰ ایالات متحده‌ی آمریکا. اطلاعات نادرست مشکل ساز است زیرا منجر به باورهای نادرست می‌شود و می‌تواند اختلافات حزبی را حتی بر سر حقایق اساسی تشدید کند. صرف خواندن پست‌های خبری کذب - از جمله پست‌های سیاسی که به شدت غیرقابل قبول هستند و با ایدئولوژی سیاسی فرد همخوانی ندارند - باعث می‌شود که بعداً آن‌ها واقعی‌تر به نظر برسند.

در بخش‌های روان‌پزشکی، شایعات می‌توانند پویایی روابط درون موسسه را تغییر دهند و می‌توانند منجر به حالت‌های مزمن درگیری با افراد یا قربانی شدن گروهی، روحیه پایین، اختلال در بهره‌وری و در بدترین موارد، فروپاشی انسجام یک بخش شوند. ۱ چنین فروپاشی ممکن است از باورهای نادرست مزمن در مورد یک یا چند همکار ناشی شود که همکاری های اعتماد را غیرممکن می‌کند.

تمام این دلایل باعث می‌شود که نیازمندِ یک سری راهکار ، برای تشخیص، جلوگیری و پیشگیری از پخش شایعات باشیم . در این گزارش به بررسی مفاهیمِ روانشناسیِ شایعات و اخبار جعلی می‌پردازیم و در ادامه‌ی اون یکی سری راهکار برای مواجه با این پدیده را بررسی میکنیم.

۲ .روانشناسیِ اخبارِ جعلی[1]

اخبار ساختگی چیز جدیدی نیست. به عنوان مثال، در سال ۱۸۳۵ روزنامه سان در نیویورک شش مقاله در مورد زندگیِ در ماه منتشر کرد که به عنوان "حقه بزرگ ماه"[2] شناخته شد.

تصویر ۱. تصویری از حقه‌ی بزرگِ ماه از مجله‌ی سان سال ۱۸۳۵

با این حال، در طول انتخابات ریاست جمهوری ۲۰۱۶ ایالات متحده و رفراندوم برگزیت بریتانیا، شکل متفاوتی از اخبار جعلی برجسته شد: داستان‌های «اخبار سیاسی» نادرست یا بسیار گمراه‌کننده، که عمدتاً از شبکه‌های اجتماعی سرچشمه می‌گیرند.

[1] Fake news
[2] Great Moon Hoax

۲–۱. انواع اخبار جعلی موردِ بررسی

به طور کلی اخبار جعلی که در تحقیقات مد نظرمون هست و از نظر اهمیتِ جلوگیری از پخش اونها در درجه‌های بالاتری قرار میگیرند و منابع احتمالی قطبی شدن سیاسی[۱]هستند به چند دسته تقسیم میشوند:

- اطلاعات غلط[۲]و اطلاعات نادرست:[۳] وقتی از misinformation صحبت میکنیم منظورمون عمدتا اطلاعات غلطی هست که به طور سهوی ایجاد می‌شود و یا فرد قصد و نیتی بدی پشت اون نداره، یا صرفا به خاطر جاهل بودن و عدم اطلاع کافی اون رو منتشر می‌کند. به عنوان مثال می‌توان به شایعات، توهین‌ها و شوخی‌ها[۴]اشاره کرد. در مقابل ما disinformation رو داریم، از لحاظ معنای بسیار شبیه misinformation هست، ولی یک فرق اساسی اونا رو کاملا از هم جدا می‌کند، اونم اینه که disinformation عمدی است و شامل محتوای مخرب مانند حقه‌بازی[۵] فیشینگ نیزه‌ای و پروپاگاندا است. موارد بسیار زیادی از این دو مورد رو در رابطه با پاندمی‌کوید-۱۹ و انتخابات ۲۰۲۰ ایالات متحده داشتیم.
- اخبار فراحزبی:[۶] محتوای خبری که کاملاً ساختگی نیست، اما رویدادهایی را پوشش می‌دهد که در واقع با یک سوگیری حزبی شدید رخ داده اند. در نتیجه، اخبار فراحزبی معمولاً گمراه کننده است و بنابراین معمولا آن را به عنوان نوعی اطلاعات نادرست در نظر میگیرند.
- روزنامه‌نگاری زرد:[۷] محتوای روزنامه‌ها، مجلات یا وب‌سایت‌هایی که تحقیقات ضعیفی دارند و هیجان‌انگیز هستند و با هدف افزایش فروش یا در اینترنت، کلیک ایجاد می‌شوند. تقریباً معادل روزنامه نگاری تبلوید است. روزنامه نگاران تثبیت شده اواخر قرن نوزدهم این اصطلاح را برای

---

[۱] political polarization
[۲] misinformation
[۳] disinformation
[۴] prank
[۵] hoax
[۶] hyperpartisan news
[۷] Yellow journalism

تحقیر تکنیک های غیر متعارف رقبای خود ابداع کردند. چنین روزنامه نگاری دارای ویژگی های زیر بود:

- استفاده از سرفصل های چند ستونی، تصاویر بزرگ و گرافیک غالب.
- داستان‌های صفحه اول که از جنجال‌انگیز تا پرخاشگر در همان شماره متفاوت بودند.
- جنگوئیسم[۱] یا برافروختن احساسات ملی از طریق اخبار کج، که اغلب مربوط به جنگ داخلی است.
- استفاده گسترده از منابع ناشناس توسط خبرنگاران بیش از حد غیور، به ویژه در داستان های تحقیقی در مورد "تجارت های بزرگ"، افراد مشهور، یا شخصیت های سیاسی؛

ما در اینجا عمدتاً بر محتوای آنلاینی تمرکز می‌کنیم که در قالب موارد بالا ارائه می‌شود. با این حال، ادعاهای نادرست و گمراه کننده به اشکال مختلف وجود دارد، و ادبیات متعددی وجود دارد که به وضوح به هم مرتبط هستند، اما خارج از محدوده بررسی ما هستند . اینها شامل کار بر روی باور توطئه [۲]، خرافات[۳]، پذیرش مزخرف[۴] و ادراک نادرست[۵] و غیره است. علاوه بر این، تمرکز ما بر روی نمونه‌های فردی اطلاعات نادرست است، و نه بر روی کمپین‌های سازمان‌یافته اطلاعات نادرست (مثلاً توسط آژانس تحقیقات اینترنتی روسیه، یا کمپین‌های مربوط به گرمایش جهانی یا تقلب در انتخابات ریاست‌جمهوری ۲۰۲۰ ایالات متحده).

۲-۲. تاثیر اخبار جعلی بر بافت اجتماعی

هنگام در نظر گرفتن عواملی که ممکن است بر باور مردم تأثیر بگذارد، ضروری است بین دو روش اساساً متفاوت برای مفهوم‌سازی باور به اخبار درست و نادرست تمایز قائل شویم.

---

[۱] jingoism
[۲] conspiracy belief
[۳] superstition
[۴] bullshit receptivity
[۵] misperceptions

- یکی از رویکردهای رایج، تمرکز بر «تشخیص» حقیقت، یا میزانی است که به اطلاعات نادرست باور می‌شود به «نسبت» به محتوای دقیق . تشخیص، که معمولاً به عنوان اعتقاد به اخبار واقعی منهای اعتقاد به اخبار نادرست محاسبه می‌شود، دقت «کلی» باورهای فرد را به تصویر می‌کشد و بنابراین بینشی را در مورد شکست‌ها در تمایز بین محتوای واقعی و نادرست («باوربه اخبار جعلی») می‌دهد.
- رویکرد دیگر این است که بر باور کلی تمرکز کنیم، یا اینکه تا چه حد یه یک اخبار - صرف نظر از صحت آن - باور دارند. ه طور انتقادی، عواملی که باور کلی را تغییر می‌دهند، نیازی به تأثیری بر توانایی افراد برای تشخیص حقیقت از دروغ ندارند : افزایش یا کاهش اعتقاد به سرفصل‌های درست و نادرست به میزانی معادل، تأثیری بر دقت کلی باورهای فرد ندارد.

حالا باید به مواردی که باعث می‌شود یک فرد یا گروهی از افراد نسبت به یک سری از اخبار غلط و اطلاعات نادرست واکنش نشون بدن و این اطلاعات رو باور کنند بپردازیم:

۲-۲-۱. انگیزه های سیاسی[1] و باورهای ایدئولوژیک

یک روایت رایج این است که عدم تشخیص اخبار درست و نادرست ریشه در انگیزه های سیاسیِ مردم دارد. برای مثال ما با مفهومی‌به اسم "محافظتِ هویتِ شناختی[2]" مواجه هستیم - محافظتِ هویتِ شناختی به تمایل افراد دارای تنوع فرهنگی به اعتبار انتخابی و رد شواهد در الگوهایی اشاره دارد که منعکس کننده باورهای غالب در گروه آنها هستند - و این باعث می‌شود که آنها نسبت به محتوایی که با هویت حزبی آنها سازگار است بیش از حد باور داشته باشند و نسبت به محتوایی که با هویت حزبی آنها سازگار نیست بیش از حد شک داشته باشند.

نظریه مرتبط استدلال می‌کند که مردم وفاداری به هویت سیاسی خود را بالاتر از حقیقت قرار می‌دهند -

[1] Political Motivations
[2] identity-protective cognition

و بنابراین نمی‌توانند حقیقت را از دروغ تشخیص دهند و صرفاً به اطلاعاتی که از نظر ایدئولوژیک منطبق هستند باور می‌کنند. این گزارش‌ها ادعا می‌کنند که تأثیر علّیِ قویِ انگیزه‌های سیاسی بر باور، عامل غالبی است که توضیح می‌دهد چرا مردم به اخبار جعلی علاقه‌مند می‌شوند.

۲-۲-۲. استدلال

دیدگاه دیگری درباره (نا)توانایی تمایز بین صدق و کذب از حوزه استدلال می‌آید. کار در این زمینه تمرکز ویژه‌ای بر نظریه‌های فرآیند دوگانه[۱]دارد که تصریح می‌کند که تفکر تحلیلی می‌تواند پاسخ‌های خودکار و شهودی را نادیده بگیرد . سوال کلیدی این دیدگاه این است که نقش استدلال بازتابی در توانایی تشخیص اخبار جعلی از حقیقت چیست؟

نظریه‌های فرآیند دوگانه جزء اصلی تحقیق در مورد علم شناختی استدلال هستند. این نظریه‌ها استدلال می‌کنند که شناخت انسان را می‌توان به دو نوع فرآیند اساساً متفاوت تقسیم کرد که از نظر ویژگی‌هایشان متفاوت است: پردازش نوع ۱ (یا سیستم ۱) که عمدتاً با خودکار بودن مشخص می‌شود، به طوری که خروجی‌های نوع ۱ ("شهود") مستقیماً به عنوان پاسخی به محرک به ذهن می‌آیند، و پردازش نوع ۲ (یا سیستم ۲) که با مشورتی مشخص می‌شود که ممکن است با توجه به یک خروجی بصری خاص (یا مجموعه‌ای از خروجی‌ها) ایجاد شود یا هم ممکن است ایجاد نشود.

از مسائل معروف این زمینه که بیان میشود، مسئله‌ی توپ و چوب بیسبال است. یک چوب بیسبال و توپ در مجموع ۱٫۱۰ دلار قیمت دارند. چوب بیسبال ۱٫۰۰ دلار بیشتر از توپ قیمت دارد. قیمت توپ چقدر است؟ پاسخ شهودی نادرست - ۱۰ سنت - به طور شهودی برای اکثر مردم به ذهن می‌رسد، در حالی که پاسخ صحیح - ۵ سنت - (برای بیشتر) تنها با یک فرآیند مشورتیِ[۲]اضافی ظاهر می‌شود.

---

[۱]-dual-proccess model

[۲]deliberation

پس چگونه مردم صحت اخبار را تعیین می‌کنند؟ همبستگی بین بازتاب شناختی و عدم اعتقاد به اخبار جعلی در مواردی که محتوا به وضوح غیرقابل قبول است (و برعکس برای اخبار واقعی) قوی تر است. این نشان می‌دهد که در مواردی که افراد صبر میکنند و فکر می‌کنند، احتمالاً دانش قبلی مرتبط یک عامل مهم است. در واقع، دانش سیاسی به طور مثبتی با تشخیص حقیقت برای محتوای اخبار سیاسیِ مرتبط است، همانطور که سواد رسانه‌ای و سواد اطلاعاتی عمومی‌نیز همینطور است. متأسفانه این نشان می‌دهد که استدلال ممکن است دقت را در زمینه‌هایی که دانش قبلی به شدت تحریف شده است، بهبود نبخشد. بنابراین، وقتی صحبت از نقش استدلال به میان می‌آید، به نظر می‌رسد که مردم در تشخیص حقیقت از باطل ناکام می‌مانند، زیرا از تأمل کافی در دانش قبلی خود خودداری می‌کنند (یا دانش قبلی ناکافی یا نادرست دارند) - و نه به این دلیل که توانایی‌های استدلالی توسط انگیزه‌های سیاسی که دارند دچار اختلال شده است.

۲-۲-۳. فرآیندهای کاوشی[۱]

تحقیقاتِ جدید در قضاوت و تصمیم گیری نشان می‌دهد که مردم احتمالاً از هیوریستیک‌ها یا میانبرهای ذهنی هنگام قضاوت در مورد سرفصل های خبری استفاده می‌کنند. در واقع یکی از راه های کلیدی برای باور شهودی به اخبار، میزان آشنایی[۲]است. تاثیر قرار گرفتن در معرضِ قضاوت‌های قبلی از یک حقیقت یا فَکت که گاهی بهش اثر حقیقتِ وهمی[۳]هم گفته می‌شود به خوبی مستند شده است. اثر حقیقت توهم‌آمیز که به عنوان توهم حقیقت[۴]نیز شناخته می‌شود، توضیح می‌دهد که چگونه، زمانی که همان اطلاعات نادرست را بارها و بارها تکرار می‌کنیم، اغلب باور می‌کنند که درست است. این حتی زمانی اتفاق می‌افتد که مردم در ابتدا می‌دانند که این اطلاعات نادرست هستند.

---

[۱] Heuristics
[۲] familiarity
[۳] illusory truth effect
[۴] illusion of truth

در نهایت، به نظر می‌رسد یکی از ویژگی های برجسته تیترهای اخبار جعلی این است که آنها اغلب از نظر احساسی برانگیزاننده هستند. یعنی اخبار جعلی اغلب به سمت تحریک، شوک، ترس، خشم یا (به طور کلی تر) خشم اخلاقی[1] تنظیم می‌شود، یعنی بیشتر به سمت خشم، انزجار یا ناامیدی قابل توجیه نسبت به دیگرانی که ارزش ها یا استانداردهای اخلاقی را نقض می‌کنند بُرده میشوند. این مهم است زیرا افرادی که گزارش می‌دهند در ابتدای کار احساسات بیشتری (مثبت یا منفی) را تجربه کرده اند، احتمال بیشتری دارد که اخبار نادرست (اما نه واقعی) را باور کنند

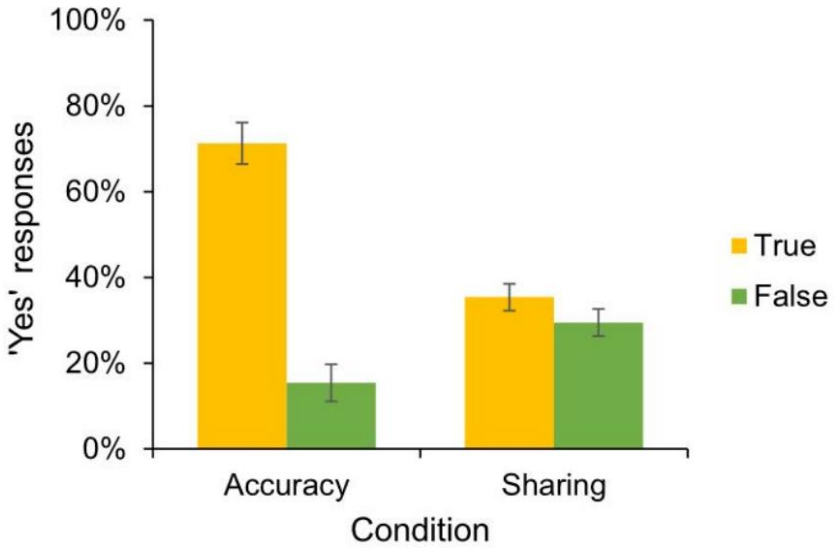
۲–۳. باور در مقابل به اشتراک گذاری اخبار جعلی

ممکن است انتظار داشته باشیم که مردم اخبار را در رسانه های اجتماعی به اشتراک بگذارند زیرا معتقدند این خبر درست است. بر این اساس، به اشتراک گذاری گسترده محتوای نادرست اغلب به عنوان شاهدی بر باورهای نادرست گسترده در نظر گرفته می‌شود. با این حال تحقیقات اخیر نشان داده است که به اشتراک گذاشتنِ قضاوت‌ها در شبکه‌های اجتماعی میتواند کاملا با دقتِ در مورد اون قضاوت در تضاد باشد.

برای مثال در تحقیقی[2] به هزار نفر شرکت کننده یک سری سرتیتر خبری که تعدادی از اونها جعلی بودند نشان داده شد و از آنها خواسته شد تا نشان دهند که آیا فکر می‌کنند تیترها صحت دارند یا اینکه آیا آنها را در نظر دارند این اخبار را در رسانه‌های اجتماعی به اشتراک بگذارند.

در نتیجه، قصد اشتراک‌گذاری برای سرفصل‌های نادرست بسیار بیشتر از ارزیابی‌های صحت آنها بود (به عنوان مثال، ۹۱٪ بیشتر در شکل ۲)، که نشان می‌دهد بسیاری از افراد ظاهراً مایل به اشتراک‌گذاری محتوایی بودند که می‌توانستند آن را نادرست تشخیص دهند.

[1] moral outrage



تصویر ۲.به اشتراک گذاری رسانه های اجتماعی لزوماً به معنای باور نیست

۲-۴. مداخلات برای مبارزه با اخبار جعلی

در اکثر تحقیقات، روش های رایج برای کنترل شایعات را می‌توان به سه دسته تقسیم کرد: حذف ارتباط بین کاربران برای جلوگیری از شایعات. مسدود کردن کاربران با نفوذ؛ انتشار حقیقت برای روشن شدن شایعات. در بخش‌های بعدی به تعدادی از این راهکارهای ارائه شده برای جلوگیری و پیشگیری از پخش شایعات و کاهش صدمات احتمالی‌ای که وارد میکنند خواهیم پرداخت.

۳ . رویکرد تشخیص شایعات مبتنی بر بلاک چین برای کووید-۱۹ [3]

وضعیت غیرقابل پیش‌بینی همه‌گیری کووید-۱۹، جهان را به سوی یک سری چالش‌های جدید سوق داد. در طول همه گیری کووید ۱۹، میلیون ها نفر برای سرگرمی، آموزش و غیره به سمت دنیای آنلاین و اینترنت سرازیر شدند. در نتیجه، طبق آمار اولیه، کل استفاده از اینترنت بین ۵۰ تا ۷۰ درصد در مقایسه با سناریوهای قبل از قرنطینه افزایش یافت، در حالی که برخی مناطق نیز شاهد افزایش ۱۰۰ درصدی استفاده از اینترنت بودیم. از یک طرف، زندگی آنلاین با حفظ فاصله اجتماعی، سبک زندگی ما را بهبود می‌بخشد. از سوی دیگر خطرات جدیدی را برای زندگی انسان به همراه دارد. شایعات عناصری هستند که عمیقاً در

ارتباطات و تعامل انسانی تاثیر دارند. اخبار نادرست اغلب به سرعت از طریق سایت‌های رسانه‌های اجتماعی مانند توییتر، یوتیوب و فیس‌بوک منتشر می‌شوند، زیرا شبکه‌های اجتماعی پلتفرم‌هایی هستند که این روزها به راحتی برای مردم قابل دسترسی هستند.

در همه‌گیریِ کووید، اطلاعات نادرست درباره «درمان‌های» COVID-19 و شایعات در مورد بهترین شیوه‌ها خطرناک بوده و حتی منجر به مرگ‌های متعدد شده است. در رابطه با بیماری همه گیر، بسیاری از اطلاعات نادرست و تئوری های توطئه کشف شده است، مانند این که ویروس وجود ندارد و صرفا یک نمای رسانه‌ای است، و این یک سلاح جنگ زیستی است که در چین توسعه یافته است، یا اینکه ویروس کرونا مسری نیست یا تجویز داروهای نادرست. همچنین اطلاعات نادرستی وجود دارد مبنی بر اینکه پوشیدن ماسک با فعال کردن ویروس، انتشار آن را افزایش می‌دهد.

شناسایی و جلوگیری از انتشار اخبار نادرست که به عنوان شایعه پنهان شده اند ضروری است. یادگیری ماشینی زمینه ای از مطالعات است که شامل تصمیم گیری و طبقه بندی نتایج بر اساس یادگیری و شناسایی الگوها از داده ها با حداقل مداخله انسانی است. چنین مدل هایی را می‌توان برای شناسایی و جلوگیری از انتشار اخبار تایید نشده، شایعات و اطلاعات نادرست در برنامه های کاربردی دنیای واقعی که از این مشکل رنج می‌برند، مانند شبکه های رسانه های اجتماعی، مقیاس بندی کرد.

در راهکاری که در این مقاله پیشنهاد شده، یک شبکه رسانه اجتماعی مجهز به بلاک چین و مدل های یادگیری عمیق قوی را برای تأیید سریع اعتبار اطلاعات و اقدامات طراحی برای جلوگیری از انتشار اطلاعات نادرست و شایعات در این مدیا تعریف شده است.

۳-۱. شبکه‌ی اجتماعی

در این مقاله یک شبکه به صورت گراف جهت دار در نظر گرفته شده است که $GN = (N,C)$ که N گره‌ها و C ارتباط بین گره‌ها هست. هر گره دارای یک درجه D مرتبط با خود است که نشان دهنده تعداد اتصالات

گره با همسایگان خود است. گره ها از طرح توزیع درجه‌ای پیروی می‌کنند که در فرمول ۱ ارائه شده است.

فرمول ۱ به شرح زیر

$$F(k) = e^{-kc} * ((Kc^{k}) / K!) \qquad (1)$$

که در آن F تابعی از k است که نشان دهنده احتمال گره nم دارای درجه k است. فرض بر این است که ماهیت همه گره ها بی طرفانه است، یعنی گره ها ممکن است سعی در انتقال اطلاعات نادرست از طریق شبکه داشته باشند یا نداشته باشند. به همین ترتیب، یک پیام می‌تواند به دو دسته باشد: پیام صحیح و پیام نادرست (شایعه). شبکه از پروتکل مسیریابی اپیدمی‌پیروی می‌کند. هنگامی‌که دو گره همسایه با هم تعامل دارند، گره اطلاعاتی را که ندارد را از همتای خود بررسی می‌کند. داده هایی که موجود نیستند را به گره های دیگر منتقل می‌کند و اطلاعات کامل را به عنوان یک تراکنش در بلاک چین ذخیره می‌کند. این فرآیند برای همه همسایگان یک گره میزبان در یک تلاش انتقال تکرار می‌شود تا زمانی که پیام تقریباً به تمام گره های دیگر برسد.

۳-۲. بلاکچین[۱]

یک بلاک‌چین مانند یک دفتر کل توزیع شده است که داده های آن بین شبکه ای از همتایان[۲]به اشتراک گذاشته می‌شود. فناوری بلاک چین روشی کارآمد برای ثبت تراکنش ها یا هر گونه تعامل دیجیتالی ارائه می‌دهد به گونه ای که آن را ایمن، شفاف، بسیار مقاوم در برابر قطعی ها و قابل بازرسی می‌کند. این فناوری هنوز جدید است و به سرعت در حال تغییر. در سال ۲۰۰۸، ساتوشی ناکاموتو مفهوم بیت کوین را معرفی کرد. این با انتشار مقاله محبوب "بیت کوین: یک سیستم نقدی الکترونیکی همتا به همتا[۳]" بود. این مقاله

[۱] Blockchain
[۲] peers
[۳] Peer to peer

پیشنهادی را برای توزیع تراکنش های الکترونیکی به جای حفظ وابسته به موسسات متمرکز برای مبادله ارائه کرد.

علاوه بر این، الگوریتم اجماع بخشی اساسی از بلاک چین است که تأثیر زیادی بر مقیاس پذیری بلاک چین دارد. الگوریتم‌های اجماع رایج عبارتند از: اثبات کار (PoW)، اثبات سهام (PoS) ، تحمل خطای بیزانسی عملی[1](PBFT) . در یک بلاک چین با اجماع PoW، اعضا از قدرت محاسباتی خود برای رقابت برای عملیات هش استفاده می‌کنند، اما منابع محاسباتی زیادی مصرف می‌کند و به طور کلی توان عملیاتی پایینی دارد . اجماع PoS به نفع شرکت کنندگان با توکن های بیشتر است که خطر انحصار و تمرکز را افزایش می‌دهد . PBFT عملکرد بیشتری دارد و منابع محاسباتی زیادی مصرف نمی‌کند، اما پیچیدگی ارتباطی آن به اندازه $O(N^2)$ است. با گسترش مقیاس گره، سربار ارتباط آن به سرعت افزایش می‌یابد و عملکردی مانند راندمان و توان عملیاتی به طور قابل توجهی بدتر می‌شود. رویکرد پیشنهادی از یک استراتژی مشترک PoW-PoS برای کنترل محاسبات سنگین و در عین حال حفظ تصادفی‌سازی کافی برای امنیت استفاده می‌کند.

یک گره مخرب مقید است به انجام تراکنش های اشتباه و انتقال مکرر پیام های جعلی. بنابراین، برای حفاظت از امنیت و منابع سیستم، هر گره دارای یک رتبه اعتبار است که بر اساس مشروعیت تراکنش تنظیم می‌شود. این رتبه بندی با استفاده از فرمول ۲ تنظیم می‌شود. که در آن $\lambda \ni [۰, ۱]$ ثابت اعتبار است و $\omega \ni [۱, ۱]$ یک عامل جریمه کننده است که برای افزایش یا کاهش رتبه اعتبار هر گره استفاده می‌شود.

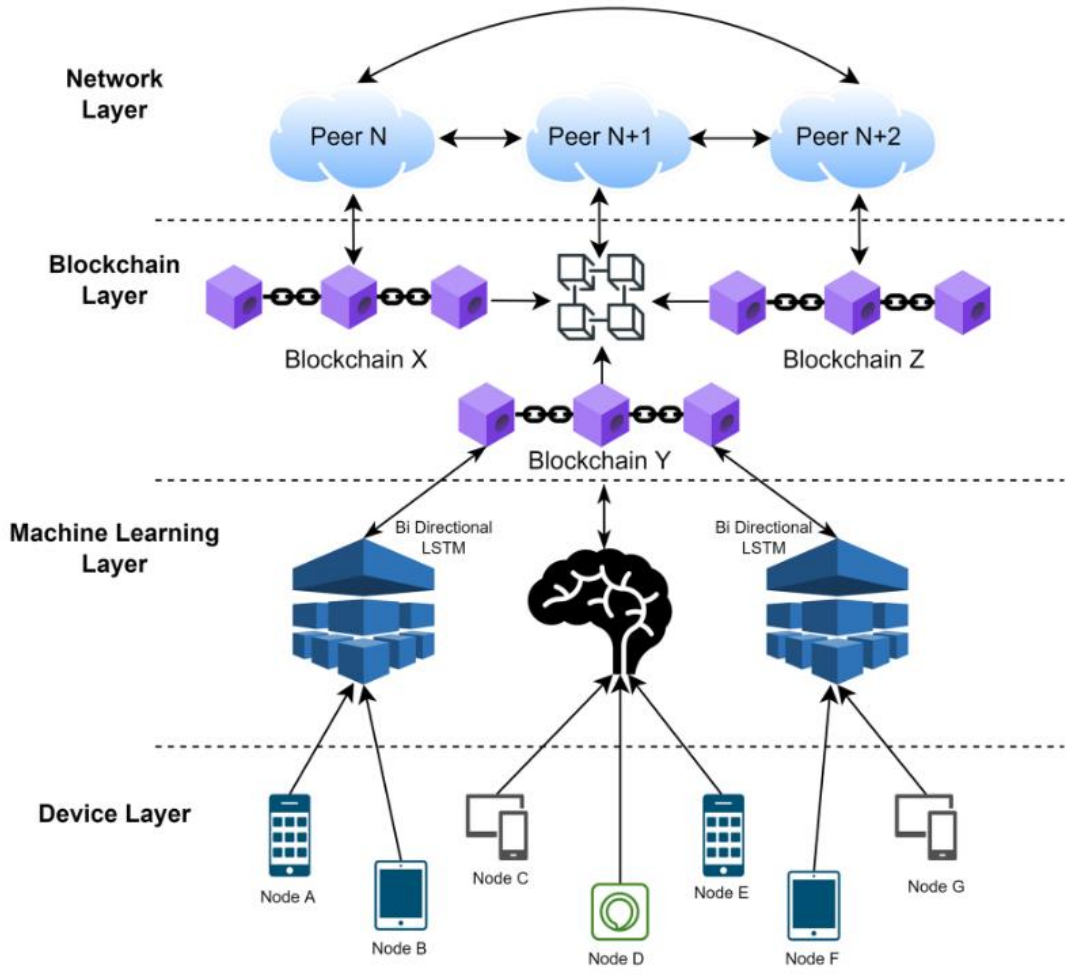
$$C(X)=\lambda(C(X))_{old} +(1 - \lambda)\omega \qquad (2)$$

اگر این رتبه اعتبار از یک آستانه از پیش تعیین شده (در مورد ما ۰,۲۵) کمتر شود، آنگاه گره مجاز به انتقال پیام از طریق شبکه نیست. در شروع اولیه شبکه، امتیاز هر گره را روی ۰,۴ قرار می‌دهیم.

[1] Practical BFT

۳–۳. لایه‌ی مربوط به یادگیری ماشین

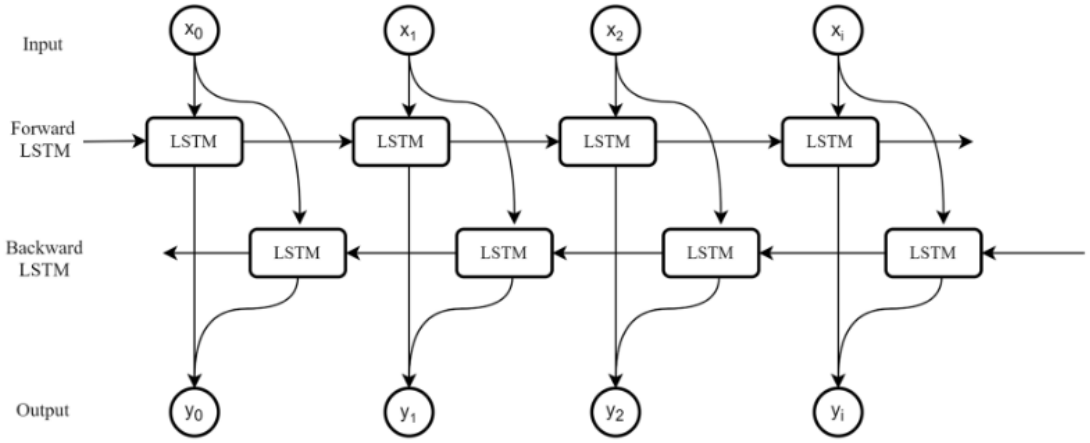
هدف اصلی این مقاله تأیید اعتبار محتوای دیجیتال و اخبار جعلی و جلوگیری از هرگونه کلاهبرداری است که ممکن است از طریق یادگیری ماشینی و بلاک چین در شبکه اجرا شود. در اینجا یک شبکه اجتماعی مبتنی بر گراف را در نظر میگیرد که در آن چندین گره متصل می‌توانند اطلاعات را به یکدیگر منتقل کنند. علاوه بر این، از بلاک چین برای اطمینان از اعتبار گره‌ها در شبکه بهره‌برداری می‌کند و یک تحلیل مقایسه‌ای از مدل‌های مختلف یادگیری ماشین برای ایجاد یک مدل بهینه برای تشخیص شایعه و تأیید اخبار انجام می‌دهد. همانطور که در شکل ۳ مشخص است یک پیام در ابتدا با استفاده از لایه دستگاه به شبکه معرفی می‌شود. در مرحله بعد، از لایه یادگیری ماشین برای بررسی مشروعیت پیام استفاده می‌شود. پس از تایید، پیام به یک بلوک در لایه بلاک چین تبدیل می‌شود. در نهایت با استفاده از لایه شبکه در سراسر شبکه منتشر می‌شود.



تصویر ۳. معماری سیستم

مدلی که برای لایه‌ی یادگیری ماشین استفاده می‌شود مدل LTSM است.LSTM نوعی شبکه عصبی بازگشتی (RNN) است که در درجه اول برای پیش‌بینی داده‌های سری زمانی / ترتیبی استفاده می‌شود. مشکل RNN سنتی حافظه کوتاه مدت تعبیه شده است. در RNN، تداوم در توالی طولانی از مرحله اولیه تا در نهایت آخرین مرحله دشوار است. RNN همچنین از مشکل گرادیان ناپدید شدن در طول انتشار پس از آن رنج می‌برد. بنابراین، لایه های مرتبط با گرادیان کوچک به درستی به روز نمی‌شوند. LSTM برای غلبه بر چنین مشکلات RNN با استفاده از گیت های مختلف با لایه های پنهان طراحی شده است. مدل‌های LSTM مزایای مختلفی نسبت به شبکه‌های معمولی مانند CNN و شبکه‌های عصبی feed-forward

دارند. ساختار منحصربه‌فرد ۳ دروازه‌ایِ LSTM به آنها توانایی یادگیری از داده‌های سری زمانی[۱] را می‌دهد، زیرا بازخورد اجازه می‌دهد تا ورودی‌های گذشته ردپایی روی مدل بگذارند. این مدل LSTM را برای کارهای طبقه بندی متنی که در آن اطلاعات زمانی باید حفظ شود، مناسب می‌کند. برای حفظ داده های زمانی، از یک مدل Bi-LSTM استفاده می‌کنیم زیرا اطلاعات متن داده شده می‌تواند در هر دو جهت (به جلو و عقب) جریان داشته باشد. این ویژگی مدل پیشنهادی را برای کارهای طبقه بندی متن مانند تشخیص اخبار جعلی بسیار مناسب می‌کند.



تصویر ۴.مدل LTSM

۳-۴. نتیجه‌گیری

نتایج نشان می‌دهد که مدل‌های یادگیری عمیق نسبت به الگوریتم‌های یادگیری ماشین پیشرفته‌تر دقیق‌تر و پیشرفت‌های قابل‌توجهی دارند. تحقیقات فعلی با ارائه اقدامات واکنشی برای شناسایی شایعات محدود شده است. در این مقاله یک بلاک چین یکپارچه و رویکرد یادگیری عمیق را برای تشخیص و پیشگیری از شایعات در کار پیشنهادی مدل‌سازی شده است که راه‌حل‌های واکنشی و فعال را ارائه می‌دهد. معرفی شایعات جدید یا اخبار نادرست به سیستم با استفاده از مدل Bi-LSTM شناسایی می‌شود که به طور مداوم پیام های دریافتی را رصد کرده و پیام های نادرست را تشخیص می‌دهد.

[۱] time-series data

از نتایج، مشهود است که کار پیشنهادی به دقت بالایی با ۹۹,۶۳ درصد دست یافته و از سایر تحقیقات عملکرد بهتری داشته است.

اما از مشکلات این رویه می‌شود به این اشاره کرده که اخبار رو فقط به دو دسته‌ی صحیح و غیرصحیح(شایعه) تقسیم میکند، همچنین بین کاربرهای مختلف تفاوتی قائل نیست.

۴ .کنترل سریع شایعات با هزینه محدود در شبکه های اجتماعی[4]

همونطور قبل‌تر هم گفته شد در تحقیقات مختلف، روش های رایج برای کنترل شایعات را عمدتا می‌توان به سه دسته تقسیم کرد: (۱) حذف ارتباط بین کاربران برای جلوگیری از شایعات . (۲) مسدود کردن کاربران با نفوذ و (۳) انتشار حقیقت برای روشن شدن شایعات. این حال، همه این مطالعات فقط یک معیار را در کنترل شایعات در نظر میگیرند. در واقعیت، میزان تأثیرپذیری کاربران از شایعات می‌تواند کاملاً متفاوت باشد و بهترین راه برای کنترل شایعات، اتخاذ تدابیر متفاوت برای کاربران است. به عنوان مثال، برای آن دسته از کاربرانی که تحت تأثیر شایعات هستند یا حتی می‌خواهند به شایعات ادامه دهند، باید با حذف اکانت آنها، این روند را متوقف کنیم. برای آن دسته از کاربرانی که به احتمال زیاد تحت تاثیر شایعات هستند، باید حقیقت را به آنها منتقل کنیم یا دسترسی آنها را به اطلاعات مسدود کنیم.

آن دسته از کاربرانی که احتمال کمی‌دارند تحت تاثیر شایعات قرار بگیرند، فقط باید آنها را نشانه گذاری کرده و ردیابی کنیم. با توجه به این موضوع، در این مقاله پیشنهاد میشود که کاربران را به گروه‌های مختلف با معیارهای کنترلی مختلف طبقه‌بندی کنیم. برای عملی بودن، کاربران را از کمترین تا بالاترین احتمال تحت تاثیر قرار گرفتن از شایعات رتبه بندی می‌کنیم و در نظر داریم کاربران را به ۵ گروه ($H1$، $H2$، $H3$،

$H4$، $H5$) با اقدامات عدم اقدام[1]، برچسب گذاری کاربر[2]، مسدود کردن دسترسی به اطلاعات، انتشار حقیقت[3] و حذف حساب کاربری،

بسیاری از محققان از مدل Independent Cascade (IC) برای شبیه سازی گسترش شایعات استفاده می‌کنند. در این مدل، یک شبکه‌ی اجتماعی مبتنی بر گرافِ G(V,E) داریم، که در آن V مجموعه گره ها و E نشان دهنده مجموعه یال ها است. هر گره نشان دهنده یک کاربر و هر لبه نشان دهنده این است که بین دو کاربر ارتباطی وجود دارد و با احتمال خاصی بر گره های همسایه خود تأثیر می‌گذارد. هر گره می‌تواند حداکثر یک بار در کل فرآیند انتشار تحت تأثیر شایعات قرار گیرد .

در این مقاله یک مدلMPIC[4] معرفی شده است برای شبیه‌سازیِ نحوه‌ی انتشارِ شایعات در این شبکه‌ی اجتماعیِ فرضی. $|V| = n$ تعداد کاربران در این شبکه است و $p_{(u,v)}$ نشان دادن احتمال تأثیر از گره u به گره v برای هر یال $(u, v) \in E$ نشان میدهد.که در ایم مدل این فرض در نظر گرفته شده است که هر گره میتواند ، هر همسایه‌ی خودش رو فقط یکبار تحت تاثیر قرار دهد.

به عنوان مثال همانطور که تصویر ۵ مشاهده می‌شود، ۹ گره داریم که در ابتدا و در زمان T=0 گره‌های ۵ و ۹ تحت تاثیر شایعه هستند. در اینجا فرض می‌شود که در زمان اولیه احتمال تحت تاثیر قرار گرفتن همه‌ی نودها برابر ۱ است و همه‌ی نودها تحت دسته‌بندیِ $Ht,i(i \in \{1, 2, 3, 4, 5\})$ قرار گرفته‌اند.

در ادامه شاهد هستیم که در زمان T=1، گره‌های ۴ و ۶ تحت تاثیر قرار گرفته‌اند، ولی گره‌های ۵ و ۹ از شبکه حذف شدند. همینظور ادامه پیدا می‌کند تا در T=3 گره‌های ۱،۲،۳ و ۷ باقی میمونند که تحت تاثیر نیستند.

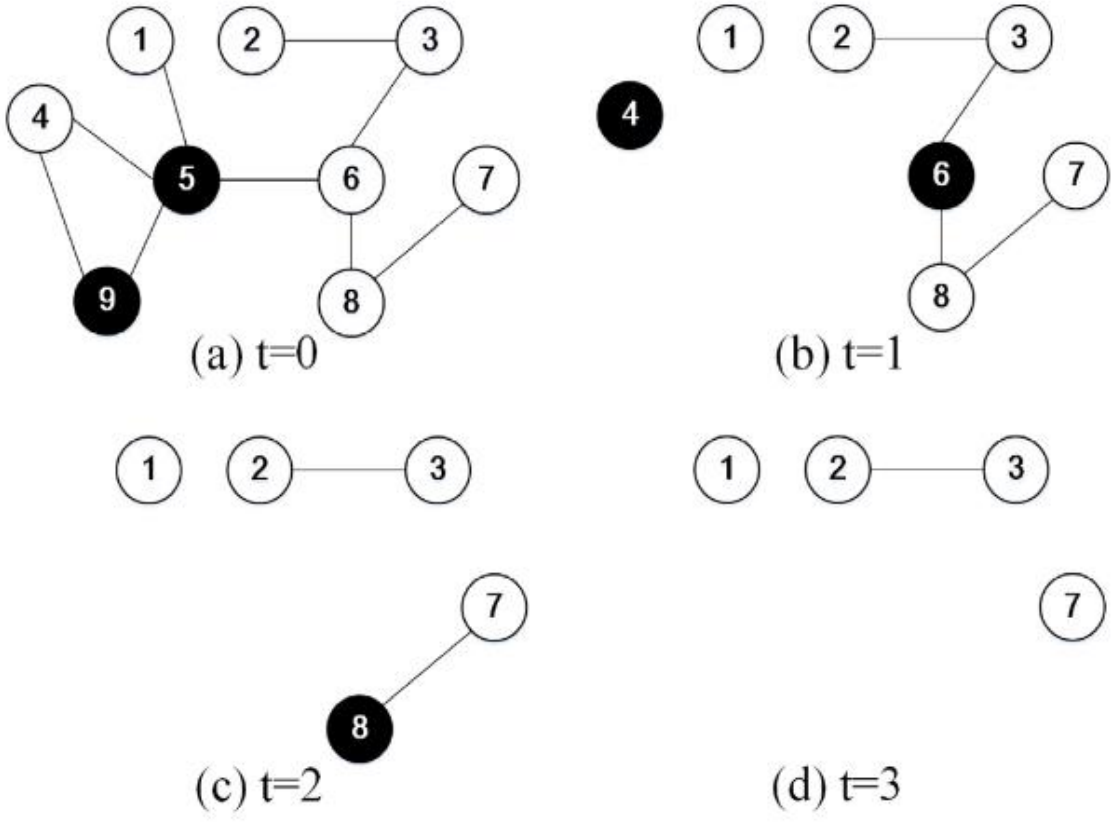
---

[1] No action

[2] Tagging the user

[3] spreading the truth

[4] Multi-probability independent cascade



تصویر ۵. مثالی از پخش شایعات تحتِ مدل MPIC

۵ .وقفه در انتشار شایعه برای گیرندگان هدف در شبکه های اجتماعی[5]

تا به اینجا چند راه حل برای مقابله با شایعات رو بررسی کردیم، . با این حال، اکثر مطالعات در مورد کنترل شایعات فرض می‌کنند که همه کاربران در شبکه‌های اجتماعی گیرنده هستند، که این یک فرض بسیار محتاطانه‌ای است. همانطور که می‌دانیم بسیاری از کاربران شبکه های اجتماعی از شایعات مصون هستند که ممکن است حقیقت را بدانند یا به شایعات اهمیتی ندهند و این شایعات برای آنها بی معنی است. فقط کسانی که شایعات را دریافت کردند و رفتارهای غیرمنطقی انجام دادند (مثلاً مصرف داروهای غیر ضروری برای COVID-19) قربانیان مورد نظر (یعنی گیرندگان هدف[۲] شایعات هستند. این با مدل‌های سنتی

[۱] Rumor Dissemination Interruption for Target Recipients in Social Networks
[۲] target recipients

انتشار شایعات که همه‌ی افراد رو گیرنده در نظر میگیره متفاوت است. اولاً، مدل‌های سنتی فرآیندی بسیار زمان‌بر برای توقف انتشار شایعات هستند، زیرا این مدل همه کاربران اجتماعی را دریافت‌کنندگان هدف می‌داند. در این میان، هدف پژوهشی این مقاله مشخص‌تر بوده و تنها به دریافت‌کنندگان هدف می‌پردازد. بنابراین روش تحقیق در این مقاله ناگزیر به صرف زمان کمتری است. دوم، هیچ راه‌حلی برای مشکل مسدود کردن شایعه گیرندگان هدف وجود ندارد که برای آن تعریف مسئله ریاضی، فرمول‌بندی مدل و اثبات نظری نیز وجود ندارد. تا آنجا که ما می‌دانیم، موضوع مسدود کردن شایعات گیرندگان هدف هنوز چالش برانگیز و باز است.

۶ .انتشار اطلاعات هدف در شبکه های اجتماعی[۱]

قبل از بررسیِ انتشارِ شایعات در شبکه‌های اجتماعی، اول باید مسئله TID را بررسی کنیم که در آن مجموعه ای از گره ها گیرندگان هدف هستند. در اینجا بر روی پویایی انتشار اطلاعات در کنترل اطلاعات تمرکز می‌شود و سعی می‌شود مدلی پیدا کنیم که انتشار شایعات و کنترل شایعات را بیان کند. یک شبکه‌ی گراف با نماد $G=(V,E)$ در نظر گرفته می‌شود، دارایی گیرندگان هدف با T و مجموعه ای از گره های منبع اطلاعات با S نشان داده می‌شوند. تعداد گره ها در S و T به ترتیب با n و m نشان داده می‌شوند. تمام مسیرهای جریان اطلاعات هدف ممکن بین S و T به عنوان RST تعریف می‌شوند. برای هر گره منبع[۲] s $(s \in S)$ و هر گره هدف t $(t \in T)$، مسیر انتشار بین s و t با R_{st} نشان داده می‌شود.

۶–۱. مسئله‌ی تک منبعی – تک هدفی[۳]

فرض کنید فقط یک گره منبع s و یک گره هدف t در شبکه وجود دارد. اطلاعات هدف را می‌توان با هر مسیری در R_{st} از s به t منتشر کرد. احتمال اینکه گره هدف t گیرنده نهایی اطلاعات از s باشد با (۱) نشان

---

۱ TARGET INFORMATION DISSEMINATING IN SOCIAL NETWORKS

۲ source

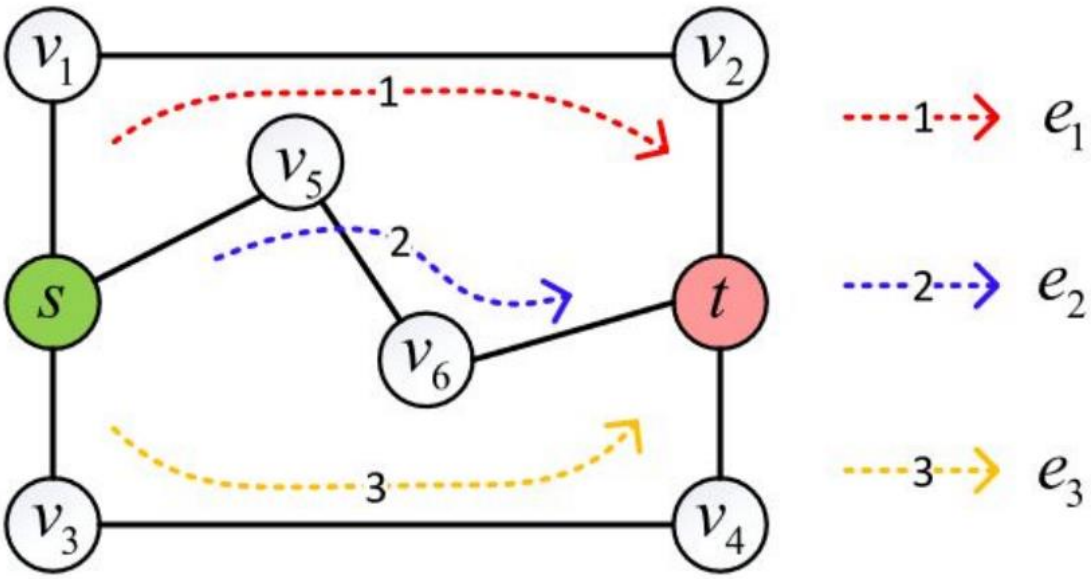
۳ Single-Source-Single-Target Problem

داده می‌شود، که در آن P^{e}_{st} احتمال جریان یافتن اطلاعات از s به t از طریق مسیر e است. ما (۲) را به عنوان احتمال اینکه t دریافت کننده اطلاعات از s از طریق همه مسیرها در Rst نباشد تعریف می‌کنیم.

$$H_{st} = 1 - \prod_{e \in R_{st}} \left(1 - p^{e}_{st}\right) \qquad (1)$$

$$F_{st} = \prod_{e \in R_{st}} \left(1 - p^{e}_{st}\right). \qquad (2)$$

هدف در این مسئله(SS-MinTID) اینه که H_{st} را با حذف کردن یا از کار انداختن یک سری از لینک‌ها با در نظر گرفتن هزینه، به حداقل برسونیم.



تصویر ۶. مثالی از نمایشِ مسیر از منبع تا هدف که از ۳ مسیر امکان پذیره

حالا با استفاده از یک الگوریتم حریصانه، لینک‌هایی که امکانش حذفشون هست رو پیدا میکنیم. برای به دست آوردن حداکثر سود، باید پیوند x را زمانی که G_x حداکثر است پیدا کنیم و آن را از شبکه G حذف کنیم.

[۱]Target Information Dissemination Minimization for Single Source Node Single Target Node

$$
\begin{aligned}
G_x^{\#} &= F_{st}^{\#}\left(E^- \cup \{x\}\right) - F_{st}^{\#}\left(E^-\right) \\
&= \frac{\prod_{e\in R_{st}}\left(1-p_{st}^e\right)}{\prod_{e\in\left\{R_{\left(E^-\cup\{x\}\right)}\cap R_{st}\right\}}\left(1-p_{st}^e\right)} - \frac{\prod_{e\in R_{st}}\left(1-p_{st}^e\right)}{\prod_{e\in\{R_{E^-}\cap R_{st}\}}\left(1-p_{st}^e\right)} \\
&= \frac{F_{st}}{\prod_{e\in\{R_{E^-}\cap R_{st}\}}\left(1-p_{st}^e\right)} \\
&\quad \times \left[\frac{1}{\prod_{e\in\{R_x\cap(R_{st}\setminus(R_{st}\cap R_{E^-}))\}}\left(1-p_{st}^e\right)} - 1\right].
\end{aligned}
$$

$$
\max \quad \left(G_x^{\#}\right) \sim \min\left\{\prod_{e\in\{R_x\cap(R_{st}\setminus(R_{st}\cap R_{E^-}))\}}\left(1-p_{st}^e\right)\right\}.
$$

تصویر ۷. الگوریتم حریصانه‌ی حذفِ لینک

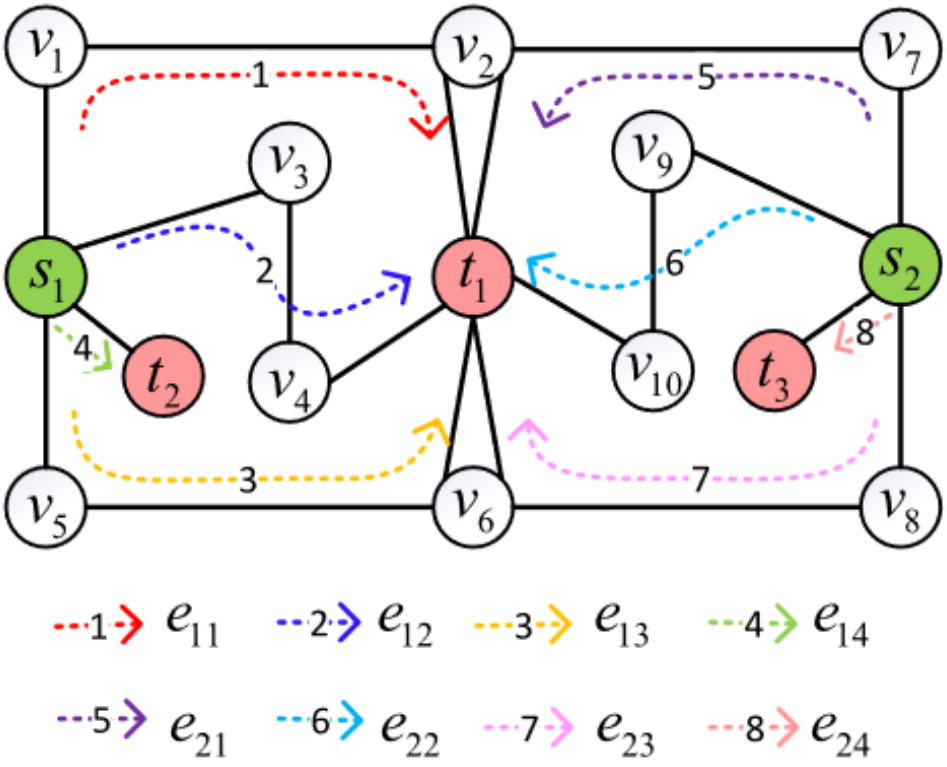
۶–۲. مسئله‌ی چند منبع – چند هدف

گره های منبع اطلاعات با S ={s1, s2,...,sn} مشخص می‌شوند که n تعداد گره های منبع است. به طور مشابه، گیرندگان هدف اطلاعات با T ={t1, t2,...,tm} نشان داده می‌شوند که m تعداد گیرندگان هدف است. تمام مسیرهای جریان اطلاعات بین S و T به عنوان R_{ST} تعریف می‌شوند. مشابه با حالت تک‌منبع-تک‌هدف اول دو مفهموم رو در نظر میگیریم:

$$
F_{ST} = \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1-p_{s_it_j}^e\right) = F_{s_1t_1} * \cdots * F_{s_nt_m} \qquad 3
$$

$$
H_{ST} = 1 - \prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in R_{s_it_j}}\left(1-p_{s_it_j}^e\right) = 1 - F_{ST}. \qquad 4
$$

F_{ST} و H_{ST} که به ترتیب در (3) و (4) نشان داده شده اند. F_{ST} احتمال این است که T گیرنده اطلاعات هدف از S از طریق همه مسیرها در R_{ST} نباشد. در مقابل H_{ST} احتمال این است که T گیرندگان نهایی اطلاعات هدف از S از طریق تمام مسیرها در R_{ST} هستند.



تصویر ۸ . حالت چند منبع- چند هدف

در این مسئله[۱](MM-MinTID) همانندِ حالت SS-MinTID به حداقل رسوندن H_{st} با حذف یک سری لینک است.که الگوریتم حریصانه‌ی مورد استفاده رو در ادامه میبینیم:

[۱] Target Information Dissemination Minimization for Multisource Nodes Multitarget Nodes

$$
\begin{aligned}
G_x^{\#} &= F_{ST}^{\#}\left(E^- \cup \{x\}\right) - F_{ST}^{\#}\left(E^-\right) \\
&= \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{\left(E^-\cup\{x\}\right)}\cap R_{s_it_j}\right\}}\left(1-p_{s_it_j}^{e}\right)} \\
&\quad - \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{E^-}\cap R_{s_it_j}\right\}}\left(1-p_{s_it_j}^{e}\right)} \\
&= \frac{F_{ST}}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_{E^-}\cap R_{s_it_j}\right\}}\left(1-p_{s_it_j}^{e}\right)} \\
&\quad \times \frac{1-\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1-p_{s_it_j}^{e}\right)}{\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1-p_{s_it_j}^{e}\right)}.
\end{aligned}
$$

حالا با استفاده از یک الگوریتم حریصانه، لینک‌هایی که امکانش حذفشون هست رو پیدا میکنیم. برای به دست آوردن حداکثر سود، باید پیوند x را زمانی که G_x حداکثر است پیدا کنیم و آن را از شبکه G حذف کنیم.

$$
\max\ \left(G_x^{\#}\right) \sim \min\left[\prod_{i=1}^{n}\prod_{j=1}^{m}\prod_{e\in\left\{R_x\cap\left(R_{s_it_j}\setminus\left(R_{s_it_j}\cap R_{E^-}\right)\right)\right\}}\left(1-p_{s_it_j}^{e}\right)\right].
$$

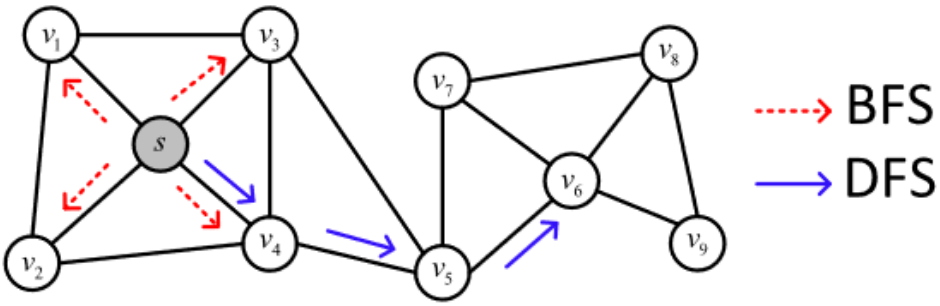
۶–۳. الگوریتم‌های استفاده شده برای پیدا کردن مسیرها

مدل TID، ابتدا باید تمام مسیرهای انتشار شایعه R_{ST} را بین گره های منبع S تولیدکنندگان شایعه و گیرندگان هدف T محاسبه کنیم. بنابراین، راه حل برای به دست آوردن تمام مسیرهای R_{ST} اولین مشکلی

است که باید انجام شود. در اینجا از ۲ روش برای پیدا کردن مسیرها[۱]استفاده می‌شود. یکی الگوریتم جستوجوی اول عمق و دیگری الگوریتم قدم زدن تصادفی.[۲]

- جستوجوی اول عمق(DFS):

DFS یک فرآیند شمارش اولیه است که ابتدا کل گراف را به شکل عمقی پیمایش می‌کند.. در بدترین حالت، پیچیدگی زمانی این الگوریتم O(N2) است. این الگوریتم مسیرها را باد دقت بالا به دست میاورد ولی زمان‌بر است.



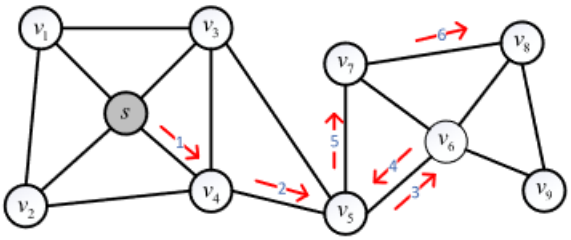
تصویر ۹. شماتیک الگوریتم bfs و dfs

- قدم زدن تصادفی(RW):

قدم زدن تصادفی یک مدل آماری است که از یک سری گام تشکیل شده است. برای نمونه‌برداری از مسیر انتشار شایعه، قدم زدن تصادفی راه‌حل بسیار خوبی است. پیچیدگی زمانی الگوریتم RW O(N) است. این الگوریتم دقت پایین‌تری نسبت به DFS دارد ولی از از لحاظی زمانی عملکرد

[۱] Path Sampling
[۲] Depth-First Search & Random Walk

بهتری نشان میدهد که برای نمونه برداری از مسیر شایعه مناسب‌تر است.



تصویر ۱۰.شماتیک الگوریتم قدم زدن تصادفی

۶-۴. تشخیص لینک‌های مهم

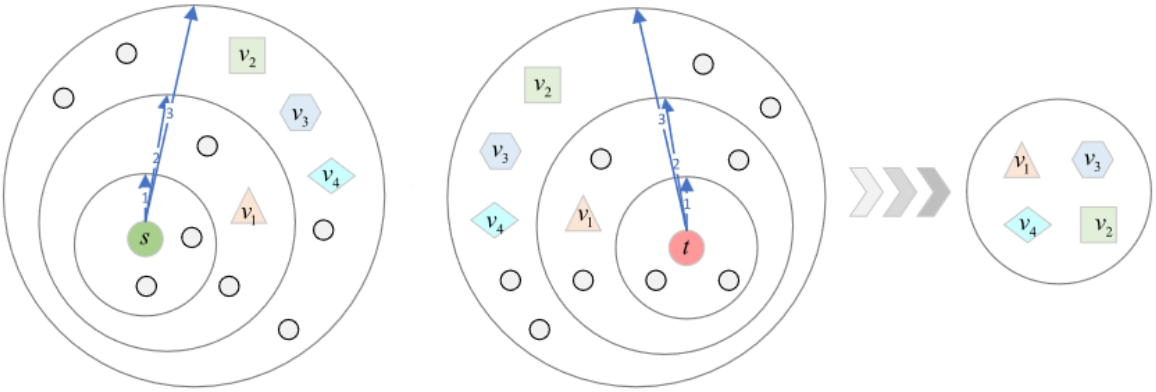
برای وقفه در انتشار شایعات در G، می‌توان با حذف برخی از حیاتی‌ترین پیوندهای اجتماعی، مسیر انتشار شایعات را از بین برد تا از انتشار موثر شایعات جلوگیری شود. البته، حذف همه پیوندها در EST توصیه نمی‌شود، زیرا باعث آسیب جدی به بهروه‌وری شبکه‌ی G می‌شود. عملیاتی که قابل انجام است حذف تعداد معینی از پیوندهای حیاتی است و تعداد آنها را β تعریف می‌کنیم. حالا چگونه این β مشخص می‌شود؟

- RumorDecay k Hop Nearest Neighbor(RumorDecay-kHNN):

روش RumorDecaykHNN یک الگوریتم انتخاب پیوند است و می‌تواند برخی از مهمترین پیوندها را برای انتشار شایعه در G انتخاب کند. این پیوندهای انتخاب شده پلی از انتشار اطلاعات بین S و T در G را تشکیل می‌دهند.

- RumorDecay k Hop Random Walk(RumorDecay-kHRW):

RumorDecaykHRW همچنین یک الگوریتم انتخاب پیوند است و می‌تواند برخی از مهمترین پیوندها را برای انتشار شایعه در G انتخاب کند. RumorDecay-kHRW همسایگان ۱ تا k هاپ گره s را با راه رفتن تصادفی پیدا می‌کند



تصویر ۱۱.شماتیک الگوریتم RumorDecaykHNN

۶–۵. نتیجه

این کار مشکل وقفه انتشار شایعه را برای گیرندگان هدف در شبکه‌های اجتماعی مطالعه می‌کند و یک استراتژی کنترل شایعه RumorDecay (یعنی RumorDecay-kHNN و RumorDecay-kHRW) پیشنهاد می‌کند. ایده اصلی RumorDecay حذف لینک با آستانه محدود است. حذف لینک های کلیدی در شبکه های اجتماعی می‌تواند به طور موثری توانایی انتشار شایعات را تضعیف کند. در زندگی واقعی، هم کنترل شایعات و هم کنترل بیماری های عفونی را می‌توان با عملیات حذف لینک حیاتی به دست آورد.بنابراین، روش انتخاب لینک بحرانی محتوای اصلی این مقاله است.

البته در این مقاله به خودِ شایعات و اینکه به چه شکل این محتوا را تشخیص می‌دهد نپرداخته است.

۷ .مقابله با شایعه به کمک روش کنترل نرم [6]

روشی دیگر که برای مقابله با شایعات در فضای مجازی ارائه شده است استفاده از روش کنترل نرم است. روش کنترل نرم روشی است که در آن برخلاف کنترل سخت که در آن کاربرانی که شایعه‌پراکنی می‌کنند با مسدود شدن، خارج کردن آن شخص از یک اجتماع و... روبرو می‌شوند، به مقابله با شایعات و افراد انتشاردهنده‌ی شایعه می‌پردازد.

ما یک مدل در کنترل شایعه‌ی نرم داریم که از دو بخش تشکیل شده است. یک بخش دوستان قابل اعتماد ما هستند و بخش دیگر افراد معتبر می‌باشند. ما با کمک گرفتن از این دو به جنگ با شایعات و افراد منتشرکننده‌ی شایعه می‌رویم. این دو منبع آگاهی‌ساز در این گزارش به اختصار با TF[1] (دوستان قابل اعتماد) و RA[2] (افراد معتبر) مشخص شده‌اند.

۷-۱. مدل‌های کنترل نرم

دو مدل دیگر که به مدل کنترل نرم ما کمک می‌کند مدل‌های اعتماد و مدل بازی تکاملی هوشمند است که به توضیح هر کدام خواهیم پرداخت.

۷-۱-۱. مدل اعتماد

مهم‌ترین بخش در مدل اعتماد برای کنترل شایعات انتخاب دوستان قابل اعتماد ما هستند. ما باید با دقت و بررسی پارامترهایی بهترین TFها را انتخاب کنیم. این پارامترها در مدل اعتماد عبارتند از: علاقه‌ی کاربر، میزان محبوبیت و صمیمیت کاربر و در نهایت انتخاب مورد اعتمادترین مشاوران در زمینه‌ی کنترل شایعات.

---

[1] Trusted Friends

[2] Reputable Authorities

۷-۱-۲. علاقه‌ی کاربر

ما برای انتخاب بهترین TFها باید توجه کنیم که شخص مورد نظر ما در چه زمینه‌هایی تخصص و مهارت دارد. به عنوان مثال اگر ما می‌خواهیم در زمینه‌ی اخبار سیاسی اطلاعات صحیح دریافت کنیم و از شایعات دور باشیم باید به سراغ افرادی برویم که در زمینه‌ی سیاسی تخصص داشته باشد و اطلاعات کافی برای آگاه کردن ما داشته باشد. اما از آنجا که در بسیاری از شبکه‌های اجتماعی اطلاع از تخصص افراد کار پیچیده و دشواری است ما باید به سراغ مورد دیگری برویم.

وقتی ما در تشخیص تخصص کاربر با مشکل مواجه شدیم باید به سراغ علاقه‌ی کاربر برویم. در این مورد ما باید فعالیت‌های کاربر مورد نظر خود را بررسی کنیم و در خصوص علاقه‌مندی‌های آن شخص اطلاعاتی کسب کنیم. به عنوان مثال ما می‌توانیم بررسی کنیم که این شخص چه پست‌هایی را Like کرده است، چه نظراتی تا به حال در این شبکه‌های اجتماعی ثبت کرده و یا اینکه تا به حال چه پست‌هایی را منتشر کرده و یا اگر کاربر در شبکه‌ی اجتماعی Twitter حضور دارد می‌توانیم بررسی کنیم که او چه Tweetهایی قرار داده یا چه Tweetهایی را Retweet یا بازنشر کرده است. ما با بررسی این موارد می‌توانیم تشخیص بدهیم که شخص مورد نظر ما در چه زمینه‌ها و موضوعاتی فعالیت بیشتری داشته است. اگر کاربر مورد نظر بیشتر فعالیتش حول موضوعات سیاسی است می‌توانیم به این نتیجه برسیم که این شخص یا در زمینه‌ی اخبار سیاسی صاحب نظر است و تخصص دارد و یا اینکه با توجه به فعالیت زیاد این شخص در این زمینه به مرور زمان تخصص و اطلاعات کافی در این خصوص را به دست آورده است.

پس ما می‌توانیم با توجه به علاقه‌مندی‌های کاربر مد نظر خود به تخصص آن شخص دست پیدا کنیم. به دست آوردن علاقه‌مندی یک کاربر از طریق فرمول زیر به دست می‌آید:

$$IN(f,q) = \sum_{r_i \in R_f} \sum_{t_j \in q} \frac{tf(t_j, r_i) \cdot irf(t_j)}{|R_f|}$$

۷–۲. صمیمیت و محبوبیت اجتماعی

معیار بعدی برای بررسی میزان اعتماد به یک کاربر بررسی میزان صمیمیت و محبوبیت اجتماعی آن شخص است. در ادامه به توضیح این دو مورد خواهیم پرداخت.

۷–۲–۱. صمیمت

در راه انتخاب دوستان قابل اعتماد برای مقابله با شایعات ممکن است ما سراغ افرادی برویم که تمایل ندارند در خصوص هر موضوعی نظر بدهند و به ما در زمینه‌ی شایعات کمک کنند. به عنوان مثال افراد مشهور به دلیل شهرت و میزان توجه بالایی که از سوی دیگران دارند شاید تمایل نداشته باشند در خصوص هر موضوعی اطلاع‌رسانی کنند و نظر خود را بیان کنند. انتخاب این افراد به عنوان یک TF مناسب نیست زیرا به تمام ابهامات و سوالات ما پاسخ نمی‌دهند.

در این شرایط ما باید به سراغ افرادی برویم که با آن‌ها صمیمی‌تر هستیم. هر چه ما با یک شخص رابطه‌ای نزدیک‌تر و صمیمانه‌تر داشته باشیم، احتمال این که آن شخص به خواسته‌ها و سوالات ما پاسخ بدهد بیشتر خواهد بود.

میزان صمیمیت اجتماعی یک کاربر از فرمول زیر به دست می‌آید:

$$SI\left(u, f_i\right) = \frac{\left(Sim\left(ol\left(u, A\right), ol\left(f_i, A\right)\right) + Sim\left(il\left(u, A\right), il\left(f_i, A\right)\right)\right)}{2}$$

۷–۲–۲. محبوبیت

معیار بعدی میزان محبوبیت اجتماعی کاربر مورد نظر است. هر چه یک کاربر در یک شبکه‌ی اجتماعی و یک اجتماع از محبوبیت بیشتری میان افراد مختلف بهره‌مند باشد و کاربران مختلف برای کسب اطلاعات در خصوص یک موضوع به آن شخص مراجعه کنند، احتمال این که نسبت به سوالات افراد پاسخگوتر از دیگران باشد هم بیشتر است.

میزان محبوبیت اجتماعی یک کاربر از فرمول زیر به دست می‌آید:

$$SP\left(u, f_i\right) = \frac{\sum_{k=1}^{n}\left(\frac{Retweet(t_{ik})}{max\,Retweet(A)} + \frac{Reply(t_{ik})}{max\,Reply(A)}\right)}{2}$$

۷–۳. انتخاب مشاوران قابل اعتماد

پس از بررسی معیارهای علاقه، صمیمیت و محبوبیت کاربر مورد نظر می‌توانیم مشاوران مورد اعتماد خود را انتخاب کنیم. این انتخاب با ترکیب معیارهای گفته شده و با کمک فرمول زیر انجام می‌شود.

$$Trust(u, f_i, q) = \alpha * IN(f_i, q) + \beta * SI(u, f_i) + \gamma * SP(u, f_i)$$

۷–۴. مدل بازی تکاملی هوشمند

مدل بعدی، مدل بازی تکاملی هوشمند است که در این مدل هم ما دو بخش کنترل شایعه نرم به کمک RA و TF داریم. به طور کلی در مدل بازی تکاملی هوشمند افراد یک فرد مورد اعتماد را انتخاب و از رفتارهای او تقلید می‌کنند.

۷–۴–۱. کنترل شایعه نرم به کمک RA

یکی از راه‌های کنترل شایعات دریافت پیام از سوی RAها است. این RAها هر کدام در وضعیت‌های مختلف دارای قدرت‌های متفاوتی در زمینه‌ی پخش ضد شایعه هستند. این پارامتر قدرت به این شکل در هر کاربر محاسبه می‌شود. هر چقدر یک RA در زمینه پخش ضد شایعه و انتشار اخبار صحیح بهتر عمل کند و رابطه‌ای نزدیک‌تر با افراد داشته باشد، مردم هم به او بیشتر اعتماد می‌کنند و پارامتر قدرت این شخص هم بیشتر می‌شود و دیگران بیشتر به افرادی که دارای قدرت بالا هستند اعتماد و از رفتار آن‌ها تقلید می‌کنند.

۷–۴–۲. کنترل شایعه نرم به کمک TF

در این روش هر کاربر میزان سازگاری یا به اصطلاح Fitness خودش را با TFهای مورد نظرش بررسی می‌کند و سپس یکی از TFها که بیشترین سازگاری را با او دارد به عنوان دوست قابل اعتماد خود انتخاب می‌کند و سپس از استراتژی‌های او در زمینه‌ی کنترل شایعات تقلید می‌کند.

۸ .مقابله با شایعات استفاده از یک رویکرد مبتنی بر مصونیت اجتماعی[1][7]

روش دیگر برای مقابله با شایعات استفاده از یک رویکرد مبتنی بر مصونیت اجتماعی است. این روش برای جلوگیری از شایعه، از راه مصونیت اجتماعی زیستی الگوبرداری کرده است. ارائه‌ی این راهکار و مدل برای جلوگیری از شایعات برگرفته از زندگی حشراتی مانند زنبورها می‌باشد.

زنبورها به صورت کلنی زندگی می‌کنند و همانند مورچه‌ها از جمله حشراتی هستند که به صورت گروهی زندگی می‌کنند و به همین دلیل اگر بیماری یا مشکلی هم پیش بیاید به صورت گروهی به مقابله با آن بیماری یا مشکل می‌پردازند. به این نوع رفتار محافظتی در برابر بیماری و انگل‌های مصونیت اجتماعی گفته می‌شود. بنابراین، مصونیت اجتماعی را می‌توان به عنوان یک رویکرد جمعی ضد انگل تعریف کرد که توسط حشرات با زندگی اجتماعی برای کاهش شیوع بیماری در کلنی انجام می‌شود.

به عنوان مثال، یک بیماری به نام Chalkbrood وجود دارد که برروی لاروهای زنبور عسل تاثیر می‌گذارد. هنگامی‌که لاروها با این بیماری مواجه می‌شوند می‌میرند و به صورت توده‌های سفیدی شبیه گچ می‌شوند. سایر اعضای کلنی با مشاهده‌ی همین علائم بیماری و کشته شدن لاروها متوجه وجود بیماری می‌شوند. راه مقابله با این بیماری افزایش دمای محیط هست و سایر زنبورها بعد از این که متوجه حضور این بیماری می‌شوند با روش‌هایی اقدام به افزایش دمای کلنی می‌کنند. بعد از افزایش دمای کلنی بررسی می‌کنند که آیا بیماری از بین رفته است یا خیر. اگر باز هم با مرگ لاروها روبرو شدند دوباره عملیات بالا بردن دما با شدت بیشتر تکرار می‌کنند تا زمانی که کاملا جلوی این بیماری گرفته بشود.

روشی که این مقاله برای جلوگیری از شیوع شایعات پیش گرفته است مشابه همین روشی است که زنبورها برای جلوگیری از بیماری‌ها انجام می‌دهند.

---

[1] Social Immunity

۸–۱. مراحل و اجزای مدل جلوگیری از شایعات مبتنی بر مصونیت اجتماعی

برای اجرای این مدل جلوگیری از شایعات ما سه مرحله و جزء داریم که به شرح زیر است.

۱- شناسایی کسی که در یک جامعه از همه‌ی اعضای جامعه تاثیرگذارتر است تا بتواند فرایند جلوگیری از شایعات را آغاز کند.

۲- انتخاب یک شخص به عنوان پخش‌کننده‌ی ضد شایعه میان جوامع مختلف که با نام Gateway او را در این مدل می‌شناسیم. این شخص وظیفه دارد تا اطلاعات صحیح و ضد شایعات را در بین تمام جوامع پخش کند. از آنجا که اعضای هر جامعه به اعضای خارج از جامعه‌ی خود اعتماد ندارند پس یک شخص واسط که مورد اعتماد تمام جوامع است باید باشد تا این اطلاعات را بین همه پخش کند.

۳- در مرحله‌ی آخر هم با پخش اطلاعات صحیح حساسیت افراد جامعه‌ها نسبت به شایعات کاهش پیدا می‌کند و شایعات تا حد زیادی مهار می‌شوند.

۸–۲. شدت شایعه

برای شکل‌گیری مدل کنترل شایعه مبتنی بر مصونیت اجتماعی نیاز به یک معیار به نام شدت شایعه داریم. هر چه یک شایعه قدرت گسترش‌پذیری بیشتری داشته باشد، شدت شایعه‌ی آن هم بالاتر است. ما به کمک این معیار تشخیص می‌دهیم که به چه شکلی باید با یک شایعه مقابله کنیم. شدت یک شایعه از تقسیم تعداد گره‌هایی که تحت تاثیر شایعه قرار گرفته‌اند بر تعداد کل گره‌ها به دست می‌آید.

$$r_d(t) = \frac{Number\ of\ rumor\ affected\ nodes}{Number\ of\ total\ nodes}$$

۸–۳. فرد بانفوذ گروه[1]

یکی از اجزای مدل ما Herding influencer نام دارد. به کسی Herding influencer گفته می‌شود که تاثیرگذارترین فرد در یک جامعه باشد و بتواند روی افراد آن اجتماع تاثیرگذار باشد. این شخص کسی است که فرایند مقابله با شایعات را آغاز می‌کند و افراد دیگر مطابق با کارها و نظرات او برای مقابله با شایعه با هم همکاری می‌کنند.

Algorithm 1 Herding Influence Finder

1: INPUT: $G = (V, E, B)$, a n set of communities $C = \{C_1, C_2, C_3, \dots, C_n\}$, rumor depth $r_d(t)$
2: OUTPUT: Herding Influencer set S.
3: $S \leftarrow \theta, U' \leftarrow V - S,$
4: foreach $Comm$ in C do:
5: $Seed_{Comm} = Round(rd(t) * \gamma(|Pros(t) + Ign(t)|))$
6: for $range(Seed_{Comm})$):
7: $v = \arg max_{u \in Comm/S}\{(BC_u)|u \in [Prosocial, Ignorant]\};$
8: $S = S \cup v;$
9: endfor
10: endforeach
11: return set S

تصویر ۱۲. الگوریتم Herding Influencer

۸–۴. فرد بانفوذ دروازه [2]

همانطور که پیش از این هم اشاره شد، ما نیاز به یک شخص داریم تا اطلاعات صحیح و ضدشایعات را بین جوامع مختلف منتشر کند. این شخص باید بین همه‌ی جامعه‌ها مورد اعتماد باشد. علت نیاز به وجود چنین شخصی این است که افراد داخل یک جامعه فقط به افراد داخلی اعتماد دارند و به افرادی که در جوامع دیگر حضور دارند اعتماد ندارند. پس باید یک شخص به عنوان واسط وجود داشته باشد تا بتواند بین این جوامع یک ارتباط برقرار کند و اخبار و اطلاعات صحیح را بین آن‌ها منتشر کند.

[1] Herding influencer
[2] Gateway Influencer

Algorithm 2 Gateway Influencer Finder

1: INPUT: $G = (V, E, B)$, a n set of communities $C = \{C_1, C_2, C_3, \dots, C_n\}$,
2: OUTPUT: Gateway influence set GI
3: $GI \leftarrow \theta, U' \leftarrow V - S$,
4: foreach $Comm$ in C do
5: foreach node i in $Comm$:
6: if $\exists\ e \mid (i, u)$ & $u \nexists$ $in\ Comm$:
7: $PN = PN \cup i$;
8: endif
9: endforeach
10: $w = \arg max_{u \in PN}\{(BC_u) | u \in [Prosocial, Ignorant]\}$;
11: $GI = GI \cup w$;
12: endforeach
13: return set GI

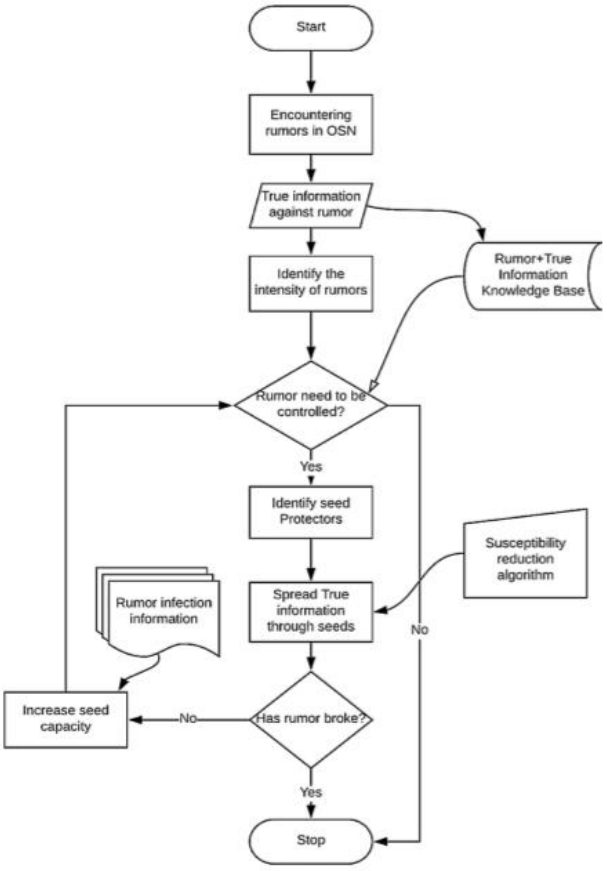
تصویر ۱۳. الگوریتم Gateway Influencer

در نهایت با انجام مراحل مدل مصونیت اجتماعی، حساسیت افراد نسبت به شایعات کمتر می‌شود و تاثیر شایعات کاهش پیدا می‌کند. همچنین لازم به ذکر است که باید اطلاعات صحیح سریع‌تر از شایعات در هر کلاستر پخش بشود تا از سرکوب درست شایعات اطمینان حاصل کنیم و این مدل بتواند بهترین نتیجه را در پی داشته باشد.

۸-۵. فلوچارت[۱] مدل جلوگیری از شایعه مبتنی بر مصونیت اجتماعی

فلوچارت مدل پیشنهادی برای جلوگیری از شایعات مبتنی بر مصونیت اجتماعی در شکل زیر آورده شده است.

[۱] Flowchart



تصویر ۱۴. فلوچارت مدل جلوگیری از شایعه مبتنی بر مصونیت اجتماعی

شروع این فلوچارت و مدل پیشنهادی با انتشار یک شایعه اتفاق می‌افتد. بعد از تشخیص پخش شایعه، اطلاعات صحیح برای مقابله با این شایعه مشخص می‌شود و وارد فلوچارت می‌شود. سپس معیار شدت شایعه محاسبه می‌شود و سپس با اطلاعاتی که تا به اینجا به دست آمده است تشخیص داده می‌شود که آیا این شایعه نیاز به کنترل شدن و مقابله دارد یا خیر. اگر جواب منفی بود که عملیات به پایان می‌رسد، در غیر این این صورت شخص تاثیرگذار آن جامعه مشخص می‌شود و اطلاعات صحیح را بین تمام کاربران منتشر می‌کند. سپس بررسی می‌کنیم که آیا شایعه مهار شده است یا خیر. اگر همچنان شایعه در حال تاثیرگذاری برروی افراد جامعه هست این عملیات با شدت بیشتر دوباره تکرار می‌شود تا زمانی که شایعه به طور کامل مهار بشود.

۹ . استفاده از ویژگی‌های روانی و زبانی برای تشخیص شایعات [8]

راه دیگری که می‌توانیم شایعات را تشخیص بدهیم استفاده از ویژگی‌های روانی و زبانی شایعات است. در این روش ما فقط به محتوای یک پیام برای تشخیص درست یا غلط بودن آن توجه نمی‌کنیم بلکه به ویژگی‌های دیگری مانند لحن نوشته، کلمات به کار برده، احساسات نهفته در متن و... هم دقت می‌کنیم. برای بررسی این روش از داده‌های برچسب‌گذاری شده با موضوع بیماری COVID-19 و به دست آمده از توییتر استفاده می‌شود.

۹–۱. داده‌های مورد بررسی

داده‌های ما شامل توییت‌ها[۱]و نظرهای[۲]مرتبط با موضوع COVID-19 و احساسات موجود در هر توییت و نظر می‌باشد. همچنین لازم به ذکر است که تحلیل احساسات موجود در هر توییت براساس ابزار MonkeyLearn به دست آمده است و برچسب‌گذاری شده است.

۹–۲. ویژگی‌های مبتنی بر محتوا، زمینه‌ای و روانی و زبانی

ما برای بررسی شایعه بودن یا نبودن یک توییت باید براساس سه ویژگی عمل کنیم. ویژگی‌های مبتنی بر محتوا، ویژگی‌های زمینه‌ای و ویژگی‌های روانی و زبانی یک توییت. در ادامه به توضیح هر کدام از این ویژگی‌ها می‌پردازیم.

۹–۲–۱. ویژگی‌های مبتنی بر محتوا

ویژگی‌های مبتنی بر محتوا چیزهایی است که مستقیما از داخل متن توییت به دست می‌آید. به عنوان مثال تعداد شکلک‌هایی[۳]که در توییت به کار رفته است، تعداد هشتگ‌ها[۴]و طول توییت و تعداد حرف‌ها و

---

[۱] Tweets
[۲] Comment
[۳] Emoji
[۴] Hashtag

کلمات به کار رفته در یک توییت از جمله ویژگی‌های مبتنی بر محتوای یک توییت به شمار می‌روند. از این ویژگی‌ها می‌توان برای بررسی درست یا غلط بودن یک توییت استفاده کرد.

۹-۲-۲. ویژگی‌های زمینه‌ای

در ویژگی‌های زمینه‌ای از محتوای متنی توییت جدا می‌شویم و بخش‌های دیگر یک توییت مورد بررسی قرار می‌گیرد. به عنوان مثال تعداد لایک‌های یک توییت، تعداد دفعاتی که یک توییت بازنشر یا به عبارتی ریتوییت شده است و تعداد نظرات و پاسخ‌هایی که برای یک توییت ثبت شده است از جمله ویژگی‌های زمینه‌ای است که باید بررسی کرد تا به میزان اعتبار یک توییت پی ببریم.

۹-۲-۳. ویژگی‌های روانی و زبانی

ویژگی‌های روانی و زبانی، سومین ویژگی است که به دو ویژگی قبل اضافه شده است و بحث اصلی ما نیز هست و در تشخیص شایعات به ما کمک می‌کند. در این ویژگی ما باید احساسات، نحوه‌ی بیان و ویژگی‌هایی از این قبیل را مورد بررسی قرار بدهیم. به عنوان مثال براساس تحقیقاتی که صورت گرفته است این نتیجه حاصل شده است که افرادی که در توییت‌های خود از ضمایر اول شخص و دوم شخص برای بیان یک موضوع و یا خبر استفاده می‌کنند، اغلب در حال انتشار محتوایی خیالی و اشتباه هستند. یا به عنوان مثالی دیگر, معمولا افرادی که از الفاظ زشت و فحش در توییت‌های خود استفاده می‌کنند نیز در حال انتشار اطلاعاتی غلط به افراد هستند. همچنین استفاده از افعال کُنشی نیز معمولا نشان‌دهنده‌ی نادرست بودن مطلب است.

در مجموع با ترکیب سه ویژگی گفته شده می‌توانیم در زمینه‌ی تشخیص شایعات به خوبی عمل کنیم و از اخبار و مطالب غلط دوری کنیم.

در جدول زیر خلاصه‌ای از ویژگی‌های هر یک از این سه مورد مشخص شده است.

Feature Class	Features
Content-based Features	noOfSmiliesInTweet
	noOfQuestionMarksInTweet
	lengthOfTweet
	noOfSmiliesInComments
	noOfMentionsInComments
	noOfHashTagsInComments
	presenceOfURLsInComments
	noOfQuestionMarksInComments
	lengthOfComments
Contextual Features	noOfLikesTweet
	noOfCommentsTweet
	noOfRetweets
	noOfRepliesTweet
	noOfLikesComments
	noOfRetweetsOfComments
	noOfRepliesOfComments
	noOfSupportStanceOfTweet
	noOfDenyStanceOfTweet
	noOfCommentStanceOfTweet
	noOfQueryStanceOfTweet
Psycho-linguistic Features	hasSwearWords
	hasModalAdverb
	hasActionAdverb
	has1stPersonPronoun
	has2ndPersonPronoun
	hasMannerAdverb
	hasHedgeWords
	hasSuperlatives
	hasComparatives
	sentimentScore
	emotionScore

تصویر ۱۵. جدول مشخصات سه ویژگی مورد بررسی در توییت‌ها

۹–۳. نتایج استفاده از ویژگی‌های بررسی توییت

در انتها آزمایش‌هایی به کمک چند الگوریتم و ویژگی‌های مبتنی بر محتوا، زمینه‌ای و روانی زبانی توییت‌ها برروی داده‌های برچسب‌گذاری انجام شده است که نشان می‌دهد اضافه شدن ویژگی روانی و زبانی دقت تشخیص شایعه را افزایش می‌دهد.

همانطور در جدول زیر قابل مشاهده است، در آزمایش اول فقط ویژگی‌های روانی و زبانی یک توییت برای تشخیص صحت استفاده شده است. در آزمایش دوم فقط ویژگی‌های مبتنی بر محتوا و زمینه‌ای توییت بررسی شده است و در آزمایش سوم هم از هر سه ویژگی برای بررسی درست یا نادرست بودن محتوای یک توییت استفاده شده است.

Algorithms	**Experiment A** (Psycho-linguistic Features Only)				**Experiment B** (Content-based and Contextual Features)				**Experiment C** (Content-based + Contextual + Psycho-linguistic features)			
	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure	Accuracy(%)	Precision	Recall	F-Measure
J48	63.77	0.692	0.502	0.582	76.34	0.767	0.760	0.763	77.80	0.768	0.800	0.784
JRip	62.18	0.664	0.499	0.570	73.98	0.731	0.761	0.746	75.50	0.740	0.787	0.763
Random Forest	66.65	0.738	0.520	0.610	80.86	0.775	0.871	0.820	81.00	0.775	0.871	0.820
Naive Bayes	62.41	0.674	0.487	0.565	61.30	0.676	0.441	0.534	66.59	0.705	0.575	0.634

تصویر ۱۶. نتایج آزمایش بر روی توییت‌ها به کمک ویژگی‌های معرفی شده

نتایج نشان می‌دهد دقت تشخیص صحت یک توییت وقتی از هر سه ویژگی با هم استفاده می‌شود از دو حالت دیگر، یعنی استفاده از ویژگی روانی و زبانی به تنهایی و استفاده از دو روش مبتنی بر محتوا و زمینه‌ای، بیشتر است و این ویژگی ارائه شده موفق شده است دقت تشخیص شایعه را برای ما افزایش بدهد.

۱۰ . نتیجه‌گیری نهایی

شبکه‌های اجتماعی آنلاین، شمشیرهای دولبه هستند. به‌عنوان مثال، در حالی که به انتشار اطلاعات جدید در یک فضای جهانی و باز کمک می‌کنند، این شبکه‌های اجتماعی همچنین به عنوان یک پلتفرم برای انتشار گسترده شایعات نیز عمل می‌کنند. شایعات در شبکه‌های اجتماعی آنلاین سریع‌تر از هر رسانه‌ی دیگری پخش می‌شوند. این امر اهمیت ابداع یک رویکرد مقرون به صرفه و سریع‌تر برای مهار شایعات را به عنوان یک ضرورت فوری نشان می‌دهد. انتشار شایعه را می‌توان با مسدود کردن شایعه در برخی گره‌های شبکه یا سرکوب شایعه با انتشار موثر اطلاعات ضد شایعه کنترل کرد. در این گزارش به روش‌های مختلفی چون بلاک‌چین، روانشناسی اخبار، بررسی مسائل روانی و زبانی، روش‌های مصونیت اجتماعی و... جهت مقابله با شایعاتی که در فضای مجازی منتشر می‌شود اشاره شد. نتایج به دست آمده از این روش‌ها نشان داده است که هر کدام، به نحوی توانسته‌اند درصد تاثیرگذاری شایعات بر کاربران را کاهش بدهند. اما با وجود ارائه‌ی این راه‌ها، هنوز هم حفره‌هایی در زمینه‌ی جلوگیری از نشر شایعات وجود دارد و بهره بردن از روش‌های گفته شد برای مقابله با اخبار نادرست امری لازم هستند اما کافی نیستند.

منابع و مراجع

[1] G. Pennycook and D. G. Rand, “The Psychology of Fake News,” *Trends Cogn. Sci.*, vol. 25, no. 5, pp. 388–402, May 2021, doi: 10.1016/j.tics.2021.02.007.

[2] “Shifting attention to accuracy can reduce misinformation online | Nature.” https://www.nature.com/articles/s41586-021-03344-2 (accessed Dec. 11, 2022).

[3] P. Rani, V. Jain, J. Shokeen, and A. Balyan, “Blockchain-based rumor detection approach for COVID-19,” *J. Ambient Intell. Humaniz. Comput.*, May 2022, doi: 10.1007/s12652-022-03900-2.

[4] X. Yao, Y. Gu, C. Gu, and H. Huang, “Fast controlling of rumors with limited cost in social networks,” *Comput. Commun.*, vol. 182, pp. 41–51, Jan. 2022, doi: 10.1016/j.comcom.2021.10.041.

[5] Z. Jiang, X. Chen, J. Ma, and P. S. Yu, “RumorDecay: Rumor Dissemination Interruption for Target Recipients in Social Networks,” *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 52, no. 10, pp. 6383–6395, Oct. 2022, doi: 10.1109/TSMC.2022.3144141.

[6] M. Askarizadeh and B.T. Ladani, "Soft rumor control in social networks: Modeling and analysis", *Engineering Applications of Artificial Intelligence*, 100, 22, 2021, doi: 10.1016/j.engappai.2021.104198.

[7] S. Srinivasan and D. Babu L D, "A social immunity based approach to suppress rumors in online social networks", *International Journal of Machine Learning and Cybernetics*, 12, 1281–1296, 2021, doi: 10.1007/s13042-020-01233-0.

[8] S. Mahbub, E. Pardede and A. S. M. Kayes, " COVID-19 Rumor Detection Using Psycho-Linguistic Features", *IEEE Access*, 10, 117530 - 117543, 2022, doi: 10.1109/ACCESS.2022.3220369.