

فصل ۴

مدل پیشنهادی مقابله با خبر جعلی

مطالب این فصل

۴۲	۱-۴ مقدمه
۴۳	۲-۴ مفاهیم مورد نیاز
۴۳	۳-۴ تعریف رسمی مساله
۴۴	۴-۴ مدل پیشنهادی مقابله با خبر جعلی
۴۷	۵-۴ جمع‌بندی

در این فصل ابتدا مقدمه‌ای در مورد راه‌کار پیشنهادی حل مساله‌ی مقابله با خبر جعلی و مزایای آن نسبت به بقیه‌ی رویکردها را بیان می‌کنیم. در ادامه پس از توضیح مختصری از مفاهیم مورد نیاز، مساله را به صورت رسمی تعریف می‌کنیم و در نهایت در مورد جزئیات مدل ارائه شده بحث می‌کنیم.

۴-۱ مقدمه

همانطور که در فصل ۲ گفته شد، مقابله با شایعه یک مفهوم کلی است که در سال‌های اخیر مدل‌های گوناگون با طراحی مساله به شیوه‌های مختلف آن را حل کرده‌اند. روش حل مساله تا حدی زیادی وابسته به نوع ورودی‌های مساله و رویکرد مقابله است که قبل از شروع به حل مساله باید این موارد را تعیین کرد. اکثر مدل‌های پیشین ارائه شده در این حوزه از ورودی گراف کاربران شبکه برای حل مساله استفاده کرده‌اند. اما نکته‌ی مهم این است که اطلاعات گراف مساله به سادگی قابل استخراج نیست و دسترسی کامل به آن بیشتر در سطح مدیریت شبکه امکان‌پذیر است. نکته‌ی دیگر این است که صرف داشتن گراف روابط کاربران نمی‌تواند برای مدل‌سازی پخش خبر کافی باشد زیرا طبق پژوهش‌های انجام شده مدل پخش اخبار جعلی و واقعی در شبکه‌ها متفاوت است. با این وجود اکثر کارهای پیشین از گراف روابط کاربران یک شبکه برای حل مساله مقابله استفاده کرده‌اند و در ورودی مشخصات پخش خبر جعلی و واقعی در نظر گرفته نشده‌است. به همین دلیل ما تلاش کردیم مدلی برای مقابله ارائه دهیم که اولاً از اطلاعات دنباله خبری پخش که استخراج آن نسبت به استخراج گراف روابط بسیار آسان‌تر است استفاده کنیم و دوماً با استفاده از

مجموعه داده‌ای که در همان بخش تشخیص خبر جعلی استفاده شد، مدلی ارائه دهیم که قابل یکپارچه‌سازی با مدل قبلی تشخیص خبر جعلی باشد.

یکی از راه‌های مقابله با خبر جعلی در صورت تشخیص زودهنگام آن می‌تواند این باشد که با رساندن تکذیبیه خبر جعلی یا همان نسخه‌ی واقعی خبر به کاربرانی که هنوز آلوده نشده‌اند ولی در آینده احتمال می‌دهیم آلوده به خبر جعلی خواهند شد، از پخش بیشتر خبر و باور شدن توسط افراد بیشتری جلوگیری شود. برای انجام این کار باید با داشتن دنباله‌ی اولیه خبر، کاربران محتمل بعدی دنباله را شناسایی کنیم. در ادامه بعد از شرح مفاهیم مورد نیاز و تعریف رسمی مساله، مدل ارائه‌شده را توضیح خواهیم داد.

۲-۴ مفاهیم مورد نیاز

۱۰۲-۴ جستجوی پرتو

جستجوی پرتو^۱ یک نوع جستجوی حریصانه در گراف است که در هر مرحله m (اندازه‌ی پرتو) گره را به صورت جستجوی عمق اول^۲ در حافظه نگه‌داری می‌کند و از هر کدام s گره را برای مرحله بعد مشاهده می‌کند. در نهایت با توجه به امتیاز هر گره از بین $s \times m$ گره مشاهده‌شده m گره را نگه می‌دارد. مزیت این الگوریتم استفاده‌ی کمتر از حافظه و همچنین سرعت بالاتر است. تصویر ۱-۴ روند کلی این الگوریتم را نشان می‌دهد. کاربرد این الگوریتم می‌تواند در تولید دنباله‌ها نیز باشد. برای تولید یک دنباله اگر بخواهیم به صورت کاملاً حریصانه عمل کنیم باید در هر مرحله بهترین عضو بعدی را انتخاب کنیم و به همین ترتیب ادامه دهیم. اما این روش موجب انتشار خطا می‌شود. برای کاهش این انتشار خطا استفاده با از الگوریتم جستجوی پرتو به جای تولید یک دنباله m دنباله‌ی محتمل را تولید می‌کنیم.

۳-۴ تعریف رسمی مساله

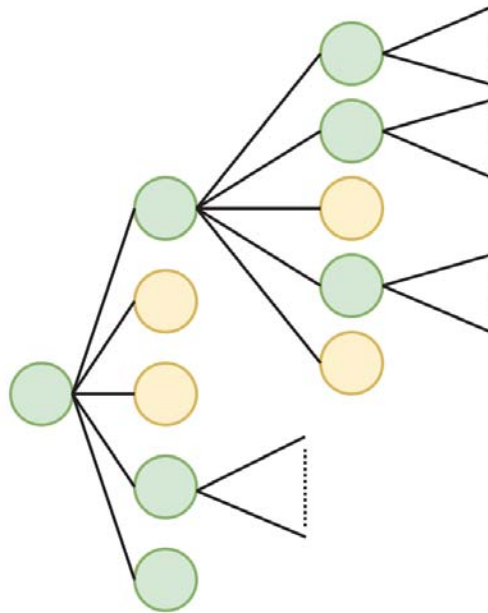
همان‌طور که در فصل ۳ در بخش ۲-۳ مساله پخش خبر جعلی را به صورت رسمی تعریف کردیم، یک مجموعه‌ی $E = \{e_1, e_2, e_3, \dots, e_N\}$ از اخبار منتشر شده در شبکه داریم که هر $e_i(t)$ به صورت $e_i(t) = \{m_{i1}, m_{i2}, \dots, m_{it}\}$ است. حال همان‌طور که گفته شد قصد داریم با شناسایی کاربرانی که احتمالاً در آینده به خبر جعلی آلوده خواهند شد، از این اتفاق جلوگیری کنیم و به همین ترتیب با پخش خبر جعلی مقابله کنیم. پس تعریف می‌کنیم مجموعه‌ی

$$U = \{u_1, u_2, \dots, u_{|U|}\} \quad (1.4)$$

مجموعه‌ی کاربران فعال در شبکه باشد. پس هر پیغام m_{ij} توسط یک کاربر $u_{ij} \in U$ منتشر شده است. همچنین می‌دانیم که کل دنباله‌ی پخش خبر به صورت $e_i = \{m_{i1}, m_{i2}, \dots, m_{iT}\}$ است. اگر تشخیص

^۱ Beam Search

^۲ Breadth First Search(BFS)



شکل ۴-۱: جستجوی پرتو با $s = 5$ و $m = 3$

برچسب در نقطه‌ی t انجام شده‌باشد، می‌خواهیم کاربرانی که پس از این نقطه در روند پخش خبر درگیر خواهند شد را شناسایی کنیم پس هدف یادگیری تابع F به صورت زیر است:

$$F(e_i(t)) = \{u_{it+1}, \dots, u_{it\tau}\} \quad (2.4)$$

۴-۴ مدل پیشنهادی مقابله با خبر جعلی

۴-۴-۱ کلیات مدل

از آنجایی که ورودی دنباله‌ای از خبر منتشرشده است، ساختار اصلی شبکه بر مبنای یک GRU طراحی شده‌است. شمای کلی مدل در تصویر ۴-۲ نمایش داده شده است. ابتدا به عنوان ورودی دنباله‌ی کاربران $\{u_1, u_2, \dots, u_t\}$ به مدل داده می‌شود. اما از آنجایی که تعداد کاربران زیاد است و استفاده از بردار تک روشن^۱ برای آن عملاً غیر ممکن است. پس نیاز به یک روش برای تعبیه کاربران در بردارهایی با طول مشخص داریم. برای این کار از ویژگی‌های کاربر که در فصل ۳ بخش ۳-۱.۳ توضیح داده‌شد، استفاده می‌کنیم. دلیل این کار این است که احتمالاً ویژگی‌های کاربران در یک دنباله با هم ارتباط دارند. که در ادامه این موضوع را با انجام آزمایش ثابت خواهیم کرد.

در نهایت دنباله‌ی حاصل بردارهای تعبیه‌شده کاربران یعنی $\{e(u_1), e(u_2), \dots, e(u_t)\}$ به عنوان ورودی GRU داده می‌شود و بردارهای نهان $\{h(e_1), h(e_2), \dots, h(e_t)\}$ تولید می‌شود. در ادامه برای تشخیص کاربر

^۱ One Hot

بعدی دنباله نیاز است که تاثیر کاربران مختلف در اثرگذاری بر روی کاربر بعدی بررسی شود. به این منظور از یک ساختار توجه استفاده می‌شود. در واقع هر کاربر در لایه‌ی بعدی به صورت جمع وزن‌دار کاربران قبلی خودش محاسبه می‌شود یعنی:

$$p(u_{k+1}) = \sum_{j=0}^k w_{kj} h(e_j) \quad (۳.۴)$$

که برای محاسبه وزن‌های توجه یا همان w_{kj} ها از فرمول زیر استفاده می‌شود.

$$w_{kj} = \frac{\exp(h(u_j)h(u_k)^T)}{\sum_j \exp(h(u_j)h(u_k)^T)} \quad (۴.۴)$$

سپس از یک لایه CNN برای تبدیل ابعاد $p(u_k)$ به ابعاد تعداد کاربران و همچنین محاسبه‌ی تاثیرگذاری‌های محلی کاربران بر روی هم استفاده می‌شود و بعد از آن برای محاسبه احتمال حضور هر کاربر از تابع فالتسازي بیشینه‌هموار استفاده می‌شود، در نتیجه:

$$pr(u_k) = \text{softmax}(\sum_{i=0}^{FilterSize} W_i p(u_{k-i})) \quad (۵.۴)$$

که نشان‌دهنده احتمال حضور کاربر u_k به عنوان عضو بعدی دنباله است.

۲.۴-۴ فاز یادگیری

در این مرحله هدف این است که مقادیر پارامترهای Θ که همان پارامترهای مدل هستند یادگرفته‌شوند. به این منظور یک مجموعه داده از مجموعه‌ی اخبار E به این صورت تولید می‌شود که از هر e_i یک دنباله seq_i به صورت زیر ساخته می‌شود:

$$seq_i = \{u_{i1}, u_{i2}, \dots, u_{iT}\} \quad (۶.۴)$$

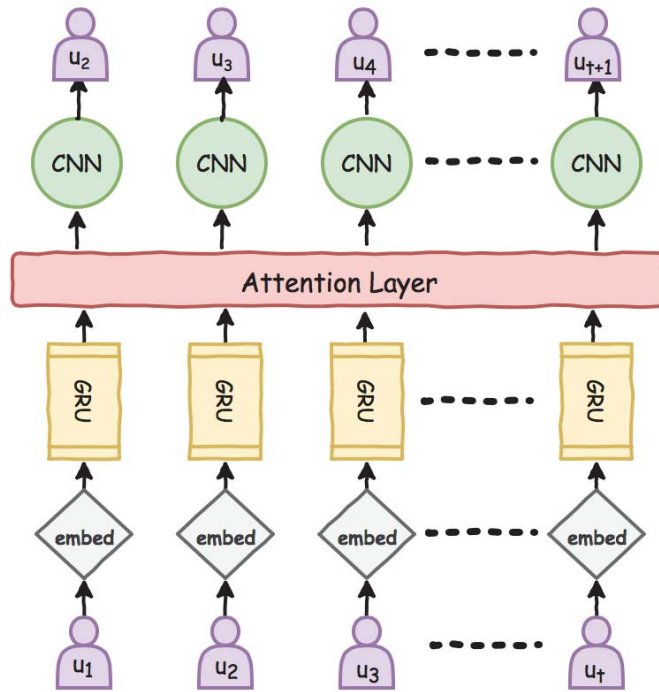
سپس دنباله‌ها به مدل داده می‌شود و مدل تلاش می‌کند در هر دنباله هر کاربر را با توجه به کاربران قبلی به دست آورد. برای این هدف تابع خطای زیر طراحی شده است که هدف کمینه کردن آن است.

$$L(\Theta) = - \sum_{seq_i} \sum_{k=2}^{|seq_i|} \log(pr_k^i(u_{ik})) \quad (۷.۴)$$

که در آن $pr_k^i(u_{ik})$ احتمال انتخاب کاربر درست توسط مدل در جایگاه k دنباله‌ی seq_i است.

۳.۴-۴ فاز آزمون

پس از انجام فرآیند آموزش مدل، با دادن دنباله کاربری به عنوان ورودی به مدل و گرفتن خروجی آخرین کاربر در واقع مدل احتمال هر وجود هر کدام از کاربران به عنوان عضو بعدی دنباله‌ی ورودی را حساب می‌کند یا به عبارتی :



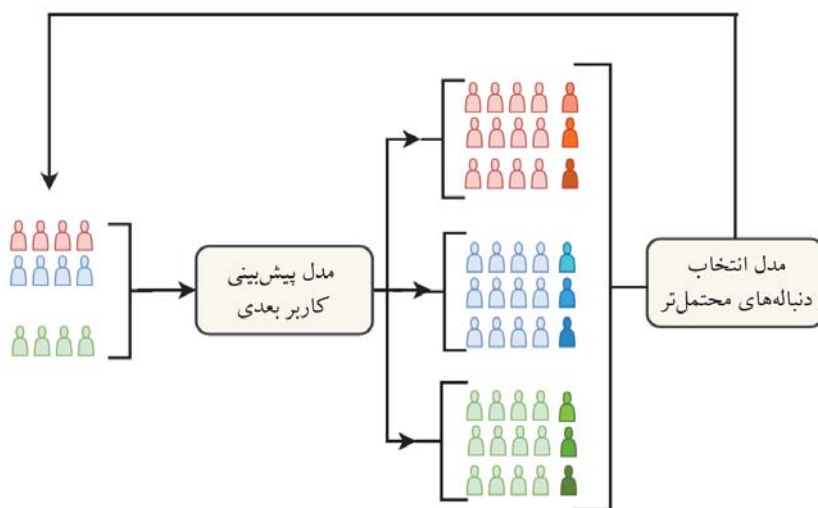
شکل ۴-۲: مدل پیشنهادی مقابله با خبر جعلی

$$pr_k^i(u_{ik}) = P(u_{ik} | \{u_{i1}, u_{i2}, \dots, u_{ik-1}\}) \quad (۸.۴)$$

یعنی احتمال اینکه عضو k ام دنباله‌ی $\{u_{i1}, u_{i2}, \dots, u_{ik-1}\}$ برابر u_{ik} باشد. حال اگر بخواهیم طبق تعریف مساله در بخش .. دنباله‌ی $\{u_{it+1}, \dots, u_{i\tau}\}$ را برای $e_i(t)$ بدست آوریم یعنی باید احتمال زیر را بیشینه کنیم:

$$P(\{u_{it+1}, \dots, u_{i\tau}\} | \{u_{i1}, u_{i2}, \dots, u_{it-1}\}) \approx \prod_{k=t}^{\tau} P(u_{ik} | \{u_{i1}, u_{i2}, \dots, u_{ik-1}\}) \quad (۹.۴)$$

که برای این کار می‌بایست در هر مرحله پس از انتخاب کاربری که احتمال $P(u_{ik} | \{u_{i1}, u_{i2}, \dots, u_{ik-1}\})$ را بیشینه می‌کند، آن کاربر را به انتهای دنباله الحاق کرده مجدداً دنباله‌ی جدید را به مدل ورودی دهیم. اما همان‌طور که در بخش ۴-۱.۲ گفته شد، این کار باعث انتشار خطا می‌شود زیرا در صورت اشتباه در یک مرحله تمامی انتخاب‌های بعدی نیز دچار خطا شده‌اند. برای کاهش این انتشار خطا از روش جستجوی پرتو استفاده شده است، به این صورت که در هر مرحله m دنباله‌ی پیش‌بینی شده نگه داشته می‌شوند و با تولید s دنباله جدید از بین هر کدام از این دنباله‌ها، نهایتاً از بین $m \times s$ دنباله m دنباله برحسب $P(\{u_{it+1}, \dots, u_{i\tau}\} | \{u_{i1}, u_{i2}, \dots, u_{it-1}\})$ انتخاب می‌شوند. شمای کلی این فرآیند در تصویر ۴-۳ نمایش داده شده است. همان‌طور که مشخص شده در هر مرحله واحد تشخیص کاربر بعدی که همان مدل ارائه شده است، برای هر دنباله تعدادی کاربر پیشنهاد می‌شود. سپس واحد انتخاب دنباله‌های محتمل‌تر از بین همه‌ی این دنباله‌ها تعدادی را برای ادامه انتخاب می‌کند و روند تشخیص کاربران بعدی نیز به همین روال پیش می‌رود.



شکل ۴-۳: نحوه تولید دنباله‌های محتمل‌تر با استفاده از جستجوی پرتو

۴-۵ جمع‌بندی

در این فصل پس از بیان مفاهیم مورد نیاز، مسأله‌ی مقابله را به صورت رسمی تعریف کردیم. سپس با شرح رویکرد انتخابی برای مقابله به بیان جزییات مدل، ورودی و نحوه‌ی آموزش آن پرداختیم.