

حداکثر سازی تاثیر مقیاس بندی با تجرید شبکه

چکیده

تحت محدودیت بودجه‌ی تبلیغاتی، بیشینه‌سازی انتخاب محصول در یک شبکه‌ی اجتماعی مشتری، یک مورد خاص و مهم از مساله‌ی عمومی بیشینه‌سازی تاثیر می‌باشد. تکنیک‌های بهینه‌سازی خاصی که همبستگی‌ها و تاثیرات جامعه‌ی محلی را در نظر می‌گیرند، می‌توانند عملکرد بهتری نسبت تکنیک‌های شبکه‌بنیان داشته باشند که باعث تعاملاتی می‌شوند که منبعت از بازاریابی محصولات متعدد برای یک گروه مشتری می‌باشد. با اینحال، این انجام‌پذیر است که از روشهای بهینه‌سازی دقیق استفاده کنیم که از عملیات ماتریس پرهزینه روی شبکه‌های بزرگ، بدون تکنیک‌های محاسباتی موازی استفاده کند. در این فصل، یک رویکرد بیشینه‌سازی تاثیر سلسله‌مراتبی را برای بازاریابی محصول ارائه می‌دهیم که یک سلسله‌مراتب تجرید را برای مقیاس‌بندی تکنیک‌های بهینه‌سازی برای شبکه‌های بزرگ، می‌سازد. یک راه‌حل دقیق روی پارتیشن‌های کوچکتر شبکه اعمال می‌شود و مجموعه‌ای کاندید از گروه‌های تاثیرگذار، به سمت بالا و به باز نمود مجرد شبکه‌ی اریجنال منتشر می‌شود که اطلاعات مسافت را حفظ می‌کند. این فرایند تجرید، راه‌حل و انتشار، تا زمانی تکرار می‌شود که شبکه‌ی مجرد حاصله، آنقدر کوچک شود که بتواند دقیقاً حل گردد.

واژگان کلیدی: بیشینه‌سازی؛ بازاریابی؛ شبیه‌سازی اجتماعی چندعاملی؛ بهینه‌سازی.

در بازار امروز، تبلیغ صرفاً موضوعی که مشتریان را به خرید محصول متقاعد کند نیست، بلکه متقاعد کردن شبکه‌ی اجتماعی آنها به انتخاب یک سبک زندگی، مدنظر است. اکنون کاملاً مشخص شده است که پیوندهای اجتماعی بین کاربران نقش مهمی در شکل‌دهی به رفتار آنها ایفا می‌کند. یکی از راه‌های تحقق این امر، از طریق تاثیر اجتماعی است که طی آن، یک رفتار یا ایده می‌تواند بین دوستان فراگیر شود. با مدنظر قرار دادن فاکتورهایی نظیر نوع دوستی و متغیرهای مداخله‌گر مشاهده‌نشده، می‌توان به بررسی دقیق و آماری همبستگی این رفتارها در شبکه‌ی اجتماعی پرداخت. هدف از راهبردهای بازاریابی ویروسی، ارتقای همبستگی این رفتارها جهت خلق زنجیره‌ای از اطلاعات است که تعداد زیادی از مشتریان در آن، از مجموعه‌ای بسیار کوچکتر از افراد آگاه تقلید می‌کنند، که قبل از سایرین توسط طرح‌های بازاریابی اقناع شده‌اند.

بازاریابی با بودجه‌ی محدود را می‌توان نسخه‌ای تخصصی از مساله‌ی بیشینه‌سازی تاثیر تلقی کرد که هدف از آن، تبلیغ گره‌های دانه‌ای جهت اصلاح دیدگاه‌های درون شبکه - بر مبنای مدل انتشار تاثیر- می‌باشد. مدل‌های انتشار پرکاربرد نظیر مدل آستانه‌ی خطی (LTM) و مدل مستقل آبخاری (ICM) فرض را بر این می‌گذارند که احتمال پذیرش یک گره، مشروط به دیدگاه‌های همسایگی شبکه‌ی محلی می‌باشد. قسمت اعظم وظیفه‌ی پیشین بیشینه‌سازی تاثیر، از این دو مدل تفاکتور استفاده می‌کند. از زمان مدل LT و مدل ارجینال، سایر مدل‌های تعمیم‌یافته، برای حیطة‌های مختلف و مصارف تخصصی مطرح شده‌اند. مثلاً مدل آبخاری کاهنده، مدل‌هایی را که در جامعه‌شناسی و انجمن‌های اقتصادی بکار می‌رود را تعمیم می‌دهد، چراکه در این جوامع، یک رفتار بصورت آبخاری و بر مبنای یک قاعده‌ی احتمالاتی پراکنده می‌شود و با مجموعه‌ای از گره‌هایی که آن رفتار را می‌پذیرند، شروع می‌شود. در مقابل مدل ارجینال IC، در مدل آبخاری کاهنده، احتمال انتشار تاثیر از یک گره فعال، ثابت نیست. همینطور، نسخه‌های تعمیم‌یافته‌ی مدل آستانه‌ی خطی نیز معرفی شده‌اند. سادگی این مدل‌های انتشار، تحلیل نظری را تسهیل می‌کند اما مدلی واقع‌گرایانه از ملاحظات بازاریابی به دست می‌دهد: نظیر تفاکتور بین تبلیغات محصولات متعدد و اثرات عضویت در جامعه بر انتخاب محصول.

جهت پرداختن به این مسائل، در کار قبلی (رفرنس 21)، مدلی از انتخاب محصول در شبکه‌های اجتماعی را مطرح کردیم که این فاکتورها را در نظر می‌گیرد و از یک فرمولاسیون بهینه‌سازی جهت محاسبه‌ی بهترین راهبرد بازاریابی -با فرض بودجه‌ی محدود- استفاده می‌کند. این فاکتورهای اجتماعی می‌توانند از متغیرهای مستقل مختلفی استفاده کنند: نظیر پیوندهای بین دوستان و همسایگان، جایگاه اجتماعی و وضعیت اقتصادی عامل‌ها¹. مشخصات مشابهی نیز بر مردم در سایر حیطه‌ها تاثیر می‌گذارند؛ مثلاً، آرال و والکر² تاثیر جایگاه اجتماعی را بر فاکتور «تاثیر» مردم در فیسبوک اثبات کرده‌اند. باورمان بر این است که در بازاریابی، تمام این فاکتورها بر احتمال تاثیرپذیری مشتریان و تانایی آنها در تاثیرگذاری بر سایرین، موثرند.

برخورداری از یک مدل واقع‌گراتر، جهت غلبه بر اثرات تبلیغات منفی مفید واقع می‌شود؛ این تبلیغات منفی که از جانب شرکتهای رقیب به سوی مصرف‌کننده‌ها سرازیر می‌شوند، قصد دارند مصرف‌کننده را به خرید کالاهای خود ترغیب و از خرید کالاهای سایر شرکتهای منصرف سازند. ضروری است که انتشار تاثیر منفی را نیز مدلسازی کنیم، چراکه انتشار می‌یابد و می‌تواند تاثیری قوی‌تر و مسری‌تر از تاثیر مثبت، بر تصمیم‌گیری‌های افراد داشته باشد.

محدودیت اصلی این نوع رویکرد و رویکردهای مشابه، این است که شامل معکوس ماتریس می‌شود که اندکی کمتر $O(N^3)$ است و یک فاکتور محدودساز است که مانع تبدیل این الگوریتم‌ها به شبکه‌های بزرگتر می‌شود. در این فصل، یک رویکرد پیشینه‌سازیِ تاثیر سلسله‌مراتبی را مطرح می‌کنیم که طرفدار «تقسیم و غلبه» است -یعنی شبکه به چند شبکه‌ی کوچکتر تقسیم می‌شود که می‌توان آنها دقیقاً با تکنیک‌های بهینه‌سازی و با در نظر گرفتن یک مدل IC ی تعمیم‌یافته جهت تعیین مجموعه‌ای از گره‌های دانه، حل نمود. از گره‌های کاندید جهت خلق یک نسخه‌ی خلاصه‌ی حافظ مسافت از شبکه، استفاده می‌شود که مدل تاثیر متراکم را بین پارتیشن‌ها حفظ می‌کند. در اینجا نشان می‌دهیم که این تکنیک خلاصه‌سازی را چطور می‌تواند الگوریتم‌های پیشینه‌سازیِ تاثیر را جهت سناریوهای بزرگتر انتخاب محصول، مقیاس‌بندی کند. بعلاوه قضیه‌ای را مطرح می‌سازیم که نشان می‌دهد مدل سیستم‌های اجتماعی واقع‌گرا، دارای نقطه‌ای ثابت است، که راهبرد بهینه‌سازیِ انتخاب محصول را معتبر می‌سازد.

¹ Agents

² Aral and Walker

این فصل به ترتیب زیر سازماندهی شده است. بخش 2 مروری بر کارهای پژوهشی دارد که در زمینه‌ی بیشینه‌سازیِ تاثیر انجام گرفته‌اند. بخش 3 روش مطروحه‌ی ما، بیشینه‌سازیِ تاثیر سلسله‌مراتبی (HIM) را معرفی می‌کند، و عملیات مدل واقع‌گرای انتخاب محصول را که در رفرنس 21 ارائه شده است، خلاصه می‌کند. در بخش 4، ما روش خودمان را در مقابل سایر رویکردهای بیشینه‌سازیِ تاثیر، هم در شبکه‌های واقعی و هم ساختگی، ارزیابی می‌کنیم. این فصل، ضمیمه‌ای بر کار قبلی ما (رفرنس 22) است که تکنیک‌های پیش‌پردازش را برای شبکه‌های بزرگ معرفی می‌کند و ارزیابی جامع‌تری از چارچوب ما، در خصوص سه دیتابیس دنیای واقعی بزرگتر ارائه می‌دهد. این فصل را با مبحثی در خصوص پژوهش‌های آتی به پایان می‌بریم.

2. پژوهش‌های مرتبط

بیشینه‌سازیِ تاثیر را می‌توان اینگونه توصیف کرد: تعیین مجموعه‌ای کوچک از گره‌ها که قادر به راه‌اندازی آبشارهای بزرگی از رفتار هستند که در تمام شبکه پخش می‌شود. این مجموعه از گره‌ها را می‌توان با استفاده از رویکردهای احتمالاتی (مثلاً رفرنس‌های 2 و 17) یا تکنیک‌های مبتنی بر بهینه‌سازی کشف نمود. رفرنس‌های 12 و 21، با بهینه‌سازی تاثیر، بصورت یک مساله‌ی بهینه‌سازی محدب رفتار می‌کنند؛ این مساله برای تاثیرگذاری بر جوامع کوچک انجام‌پذیر است اما برای مسائل بزرگتر خیر/ به دلیل ضوابط محاسباتی ماتریس، این رویکردها زمانی ناکام می‌مانند که تعداد عامل‌های سیستم افزایش یابند. الگوریتم HIM با استفاده از یک رویکرد سلسله‌مراتبی جهت تبدیل سیستم به ماتریس‌های کوچکتر، بر این ناکارآمدی غلبه می‌کند.

مدل اچ آی ام (HIM = بیشینه‌سازیِ تاثیر سلسله‌مراتبی) جهت کار روی سیستم اجتماعی پیچیده‌ای طراحی شده است که در آن، فاکتورهای چندگانه بر انتشارِ تاثیر، اثر می‌گذارند. حالتی ساده‌تر نیز توسط گروه‌های پژوهشی متعدد مورد بررسی قرار گرفته است که هدف از آن بهبود پژوهش‌های اولیه‌ی کمپ³ روی رویکردهای حریصانه جهت بیشینه‌سازیِ تاثیر می‌باشد. مثالهایی از تسریع احتمالی، شامل نوآوری‌هایی نظیر استفاده از یک مدل آبشاریِ تاثیر

³ Kempe

مبتنی بر کوتاه‌ترین مسیر یا یک الگوریتم بهینه‌سازی *lazy-forward* می‌باشند که جهت کاهش تعداد ارزیابی‌های صورت‌گرفته روی گستره‌ی تاثیر گره‌ها می‌باشند. روش‌های اکتشافی هوشمند با موفقیت بکار رفته‌اند تا سرعت محاسبات را هم در مدل LT و هم در مدل IC بالا ببرند. در این فصل، به جای استفاده از مدل‌های آبخاری اریجنال که توسط کمپ و همکارانش تهیه شد، یک مدل آبخاری معرفی می‌کنیم که تعاملات محصول و تفاوت‌های جامعه را در انتشار تاثیر، مدنظر قرار می‌دهد.

مدل‌های پیشنهادی جهت بررسی نحوه‌ی انتشار و تاثیر ایده‌ها در یک شبکه، در بسیاری از حیطه‌ها بکاررفته‌اند، از جمله دیفیوژن فناوری، انتخاب راهبرد در ستینگ‌های *game-theoretic* و پذیرش محصولات جدید در بازار. برای بازاریابی ویروسی، گره‌های نافذ را می‌توان یا بواسطه‌ی دیتای تعاملی یا راهبردهای احتمالاتی زیر شناسایی کرد. مثلاً هارت‌لاین⁴ یک مساله‌ی بیشینه‌سازی درآمد را جهت واکاوی راهبردهای بازاریابی اثربخش، حل نموده است. رفرنس 26 یک روش بازاریابی هدفمند را بر مبنای تعامل زیرگروه‌ها در شبکه‌ی اجتماعی، ارائه کرده است. مشابه این کار، باقرجیران و پارک⁵ هموفیلی خرید در شبکه‌های اجتماعی را ارتقا می‌دهند. اما به جای یافتن گره‌های تاثیرگذار، آنها راهبرد تبلیغاتی خود را بر اطلاعات پروفایل کاربران می‌گذارند. دستیابی به تاثیر عمیق در بازار، یک جنبه‌ی مهم از بازاریابی می‌تواند باشد؛ شاکاریان و داموند⁶ یک راهبرد بازاریابی ویروسی جهت انتخاب گره‌های دانه ارائه می‌دهند که گسترش کلام به سراسر شبکه را ضمانت می‌کند. کار ما از این لحاظ با کارهای مرتبط متفاوت است که مدل ماهم فاکتورهای اجتماعی را مدنظر قرار می‌دهد و هم تاثیر منفی تبلیغات محصول رقابتی و همبستگی بین تقاضا برای محصولات مختلف را مدنظر قرار می‌دهد. رویکرد بهینه‌سازی ما، عمدتاً تحت تاثیر پیچیدگی اضافی قرار ندارد، چراکه این فاکتورها فقط بر ارزش بلندمدت مورد نظر تاثیر می‌گذارند و نه بر روش پاسخ‌یابی واقعی.

برخی محققان بر جنبه‌ی نامطلوب رقابت در مقابل سایر تبلیغات تمرکز می‌کنند. در این حالت، فرض این است که تبلیغ‌گر قادر نیست گره‌ها را بصورت یکجانبه انتخاب نماید. در رفرنس 5، یک مدل طبیعی که بلحاظ ریاضیاتی قابل

⁴ Hartline

⁵ Bagherjeiran and Parekh

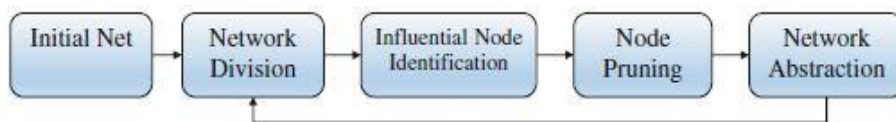
⁶ Shakarian and Damon

ردگیری است، جهت دیفیوژن نوآوری‌های چندگانه در شبکه، ارائه می‌شود. پژوهش ما فرض را بر این می‌گذارد که گره‌های تاثیرگذار به صورت مرکزی بین تبلیغگران درون یک فرایند آفلاین خصمانه، انتخاب می‌شوند.

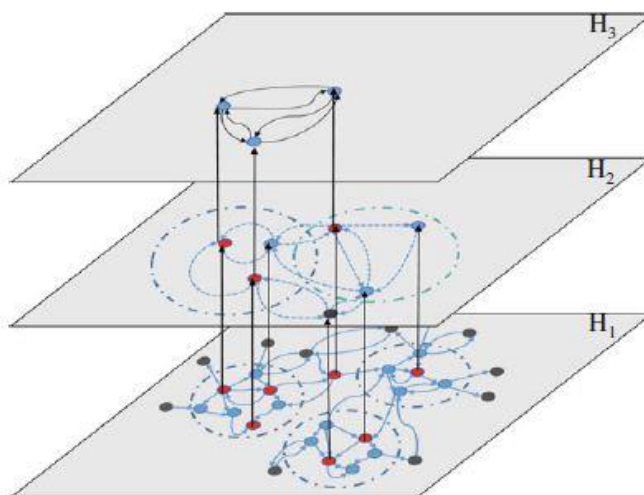
3. روش

رویکرد سلسله‌مراتبی پیشنهادی ما، به‌قرار زیر است:

1. یک شبکه‌ی محلی برای هر گره تهیه کنید که از همسایگان و همسایگان همسایگان تشکیل شده باشد؛
 2. با تخصیص گره مجازی به هر گره مرزی، تاثیر شبکه‌ی بیرونی را مدلسازی نمایید تا فعالیت را بیرون پارتیشن محلی، تلخیص کنید.
 3. بر مبنای مدل و کانکشن‌های شبکه، پارامترهای تعامل را برای گره مجازی روزآمد سازید؛
 4. با استفاده از بهینه‌سازی محدب جهت بهینه‌سازی انتخاب همیشگی محصول، مجموعه‌ای کاندید از گره‌های تاثیرگذار برای هر شبکه‌ی محلی تهیه کنید؛
 5. مجموعه‌ی کاندید را به سطح بالاتری از تلخیص انتشار دهید و گره‌های مجرد را بر مبنای کوتاه‌ترین مسیرهایشان در شبکه‌ی قبلی، پیوند زنید.
 6. فرایند تلخیص را تا زمانی تکرار کنید که شبکه‌ی حاصله به اندازه‌ی کافی برای بهینه‌سازی بصورت یک پارتیشن واحد، کوچک باشد؛ سپس مجموعه‌ی حاصله از گره‌های کاندید، برای تبلیغ در نظر گرفته می‌شود. تصویر 1 فلوچارت این الگوریتم را نشان می‌دهد.
- تصویر 2 فرایند الگوریتم را با سه سلسله‌مراتب نشان می‌دهد. گره‌های منتخب در هر همسایگی محلی، که با رنگ قرمز نشان داده‌ایم، به سلسله‌مراتب بالاتر انتقال می‌یابد و بر مبنای کوتاه‌ترین مسافتهای مسیر از سطح پایین‌تر، مجدداً کانتکت می‌شوند.



تصویر 1. فلوجارت الگوریتم ما، بیشینه‌سازی تاثیر سلسله‌مراتبی (HIM).



تصویر 2. در هر سطح سلسله‌مراتبی (H_i)، همسایگی‌های محلی و گره‌های مجازی (مشکی)، تولید می‌شوند. با استفاده از یک تکنیک بهینه‌سازی، گره‌های تاثیرگذار (قرمز) انتخاب می‌شوند. گره‌هایی که دست‌کم یکبار بعنوان گره تاثیرگذار انتخاب شده‌اند، به سطح بعدی سلسله‌مراتب منتقل می‌شوند. در سطوح بالاتر، ارتباط بین گره‌های منتخب، با استفاده از کوتاه‌ترین مسیر در شبکه‌ی اریجنال تعریف می‌شود. این فرایند تکرار می‌شود تا زمانی که آخرین مجموعه از گره‌های تاثیرگذار، کوچکتر از کل بودجه‌ی تبلیغاتی گردند.

همین فرایند در سلسله‌مراتب بعدی، جهت انتخاب گره‌های تاثیرگذارتر، تکرار می‌شود. این فرایند در آخرین سلسله‌مراتب متوقف می‌شود، یعنی وقتی که تعداد گره‌های تاثیرگذار، نهایتاً کوچکتر از بودجه‌ی تبلیغاتی می‌شود.

3.1 مدل بازار

جهت بررسی کارایی روش بیشینه‌سازی تاثیر سلسله‌مراتب (HIM) در بازاریابی تجاری، ما از مدل سیستم چندعامله که در رفرنس 21 آمده است، استفاده کرده‌ایم تا یک سیستم اجتماعی از مشتریان بالقوه را شبیه‌سازی کنیم. ما تعریف برخی پارامترهای مدل را اندکی تغییر دادیم تا به یک مدل مفهوم‌تر با قابلیت‌های تعمیم‌یافته دست یابیم.

در این مدل، جمعیت N عامل که با $A = \{a_1, \dots, a_N\}$ نشان می‌دهیم، از دو نوع عامل تشکیل شده است $(A = A_R \cup A_P)$ ، که عامل‌های (R) Regular و (P) Product نامیده می‌شوند. عامل‌های R مشتریان بالقوه در بازار هستند که گهگاه نگرش‌شان در خصوص خرید محصولات، بر مبنای تاثیری که یا از همسایه‌ها یا از عامل‌های Product می‌گیرند (کسانی که یک کالای خاص را در نقش فروشنده، پیشنهاد می‌دهند)، تغییر می‌کند. عامل‌های R به یک شبکه‌ی اجتماعی متصل تعلق دارند که در آن لینک‌های موزون جهت‌مند در این شبکه، تاریخچه‌ای از تعاملات گذشته‌ی میان عامل‌ها را در اختیار دارد. این شبکه‌ی اجتماعی با استفاده از ماتریس همسایگی (E) مدل‌سازی می‌شود که $e_{ij} = 1$ وزن لبه‌ی جهت‌مند از عامل a_i به عامل a_j است و در درجه‌ی این-نُذ و آوت-نُذ⁷ عامل a_i حاصل‌جمع اوزان این-نُذ و آوت-نُذ می‌باشند.

در این مدل، یک بردار \bar{x}_i به هر عامل تخصیص می‌یابد (هم عامل R و هم عامل P)، که بیانگر نگرش یا تمایل عامل بسوی تمام محصولات در بازار است. هر المان از این بردار x_{ip} ، یک متغیر رانندوم در بازه‌ی $[1 -1]$ است که بیانگر تمایل عامل a_i به خرید یک آیت‌م یا مصرف یک محصول خاص (P) است.

در شبیه‌سازی اجتماعی، هر عامل با یک عامل دیگر به شیوه‌ی دوبه‌دو تعامل می‌کنند که بصورت فرایند پواسن با نرخ 1، فارغ از تمام عامل‌های دیگر، مدل‌سازی می‌شود. با در نظر گرفتن فرایند پواسن تعامل، ادعا می‌کنیم که حداکثر یک تعامل در هر زمان وجود دارد. در اینجا، احتمال تعامل لین عامل‌های a_i و a_j با استفاده از P_{ij} نشان داده می‌شود که بصورت جزئی از وزن اتصال بین این عامل‌ها، بر کل کانکشن‌هایی تعریف می‌شود که عامل i با سایر عامل‌ها شکل می‌دهد. بنابراین،

$$P_{ij} = \begin{cases} \frac{e_{ij}}{d_{out}^i} & i, j \in A_R \\ \frac{u_{ji}}{Threshold} & i \in A_R, j \in A_P \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

⁷ in-node and out-node

در اینجا، پارامتر «آستانه»، کل تعداد لینک‌هایی است که عامل P می‌تواند با عامل‌های R ایجاد کند. کران‌های روی «آستانه»، تبعات طبیعی بودجه‌ی محدود شرکتها در تبلیغ محصولاتشان است. پارامتر u_{ij} یک شاخص است که نشان می‌دهد که آیا عامل P با عامل R مرتبط است یا خیر.

در هر تعامل، فرصتی برای عامل‌ها وجود دارد تا بر یکدیگر تاثیر بگذارند و بردار تمایل خود را برای خرید یا مصرف یک محصول تغییر دهند. طی این تعاملات، عامل‌های P هیچگاه نگرش خود را تغییر نمی‌دهند و یک بردار تمایل ثابت 1 را به سمت خودشان و یک بردار 1- را به سمت سایر شرکت‌های تبلیغاتی، حفظ می‌کنند. احتمال اینکه عامل i پذیرنده‌ی عامل j باشد را با a_{ij} نشان می‌دهیم و اینگونه محاسبه می‌کنیم:

$$\alpha_{ij} = \begin{cases} \frac{e_{ji}}{d_{in}^i} & i, j \in A_R \\ cte & i \in A_R, j \in A_P \end{cases} \quad (2)$$

یک پارامترهای مهم دیگر در فرایند تاثیر عامل، ϵ_{ij} است که مشخص می‌کند عامل j تا چه مقدار بر عامل i تاثیرگذار است. این پارامتر نقش فاکتورهای اجتماعی را در تصمیم‌گیری عامل‌ها نشان می‌دهد. بر خلاف کار قبلی، ما این پارامتر را به یک توزیع خاص محدود نمی‌کنیم تا بتوانیم مدل منعطف‌تری داشته باشیم. علاوه بر این، در زندگی واقعی هم یک همبستگی بین تقاضای کاربر برای محصولات مختلف در بازار وجود دارد. خواسته‌ی مشتریان برای یک محصول خاص، به میل وی در خصوص سایر محصولات مشابه بستگی دارد. ماتریس M این همبستگی را مدلسازی می‌کند و ما تاثیر آنرا در فرمولاسیون خودمان در نظر می‌گیریم. هدف نهایی مساله‌ی بازاریابی ما، این است که عامل‌های تاثیرگذار را در گراف شناسایی کنیم و مجموعه‌ای از ارتباط‌های بین عامل‌های A_P و عامل‌های A_R را تعریف نماییم، به طریقی که بتوان میل بلندمدت جهت خرید آن کالاها را در وجود عامل‌ها پیشینه ساخت. توجه کنید که پیوندهای بین عامل‌های P و عامل‌های R لینک‌های مستقیمی از محصولات به عامل‌ها می‌باشند و نه در جهت عکس.

3.2 مدل مستقل آبخاری تعمیم یافته

ما از یک نسخه‌ی تعمیم یافته‌ی ICM، مشابه [13,21] استفاده می‌کنیم. دینامیک‌های مدل در هر تکرار k به قرار زیر است:

1. عامل i بر اساس یک توزیع احتمال یونیفورم روی تمام عامل‌ها، تعامل را شروع می‌کند. آنگاه عامل i یک عامل دیگر را در میان همسایگانش، با احتمال p_{ij} انتخاب می‌کند. توجه کنید که پویایی تمایل می‌تواند با احتمال $1/N(p_{ij}+p_{ji})$ رخ دهد، چراکه نگرش عامل i می‌تواند تعیین کند که آیا آن تعامل را آغاز می‌کند یا توسط عامل j انتخاب می‌شود.

2. مشروط به تعامل i و j است:

* با انتشارپذیری α_{ij} ، عامل i تمایلش را تغییر می‌دهد:

$$\begin{cases} \vec{X}_i(k+1) = \varepsilon_{ij} M \vec{X}_i(k) + (1 - \varepsilon_{ij}) M \vec{X}_j(k) \\ \vec{X}_j(k+1) = \vec{X}_j(k) \end{cases} \quad (3)$$

توجه کنید که M ماتریس از پیش تعریف شده‌ای است که بیانگر همبستگی بین تقاضا برای محصولات مختلف است. * با احتمال $(1-\alpha_{ij})$ ، عامل i تحت تاثیر عامل دیگر قرار نمی‌گیرد:

$$\begin{cases} \vec{X}_i(k+1) = \vec{X}_i(k) \\ \vec{X}_j(k+1) = \vec{X}_j(k) \end{cases} \quad (4)$$

لازم است متذکر شویم که اگر ما $\varepsilon_{ij} = 0$ ، $M = I$ قرار دهیم و p_{ij} ها را برابر درست بعد از فعالسازی هر گره، برابر با 1 قرار دهیم و در باقی زمان‌ها برابر با 0 قرار دهیم، مدل تعامل فوق را می‌توان به مدل IC تنزل داد. همچنین از آنجاکه مقادیر بردار تمایل، از $[-1 \ 1]$ است، $x_{ips} \in [0 \ 1]$ and $x_{ips} \in [-1 \ 0]$ را می‌توان به 0 و 1 تغییر دهیم تا با باز نمود فعالسازی و عدم فعالسازی مدل IC منطبق گردد.

3.3 الگوریتم HIM

با استفاده از مفروضاتی راجع به پویایی‌های انتخاب محصول توسط مشتری، یک تکنیک بهینه‌سازی مقیاس‌پذیر ابداع کردیم و آنرا بهینه‌سازی تاثیر سلسله‌مراتبی (HIM) نامیدیم. شبه‌گد الگوریتم پیشنهادی ما را در جدول 1 ملاحظه می‌کنید. در اینجا، ماتریس E بیانگر ماتریس کانکشن میان عامل‌های Regular و ماتریس‌های P و A حاوی تمام α_{ij} و π_{ij} های مدل بازار می‌باشند. بعبارت دیگر، تمام تعامل‌ها و احتمالات تاثیر بین دو جفت از عامل‌های R یا (A_R) ، در المان‌های این ماتریس‌ها تعبیه شده‌اند. عامل حاوی تمام اطلاعات مرتبط با مشخصات عامل‌های R و P بود، از جمله بردارهای تمایل $(\bar{X}_i's)$ و بردارهای تاثیر \bar{T}_i با سایز P می‌باشد، که l_{ip} در آن بیانگر تعداد دفعاتی است که عامل i بعنوان یک‌گره تاثیرگذار برای محصول p انتخاب شده است. خدمات الگوریتم، تمام دیتای موجود در خصوص عامل‌ها و مدل را بعنوان ورودی دریافت می‌کند و خروجی الگوریتم، ماتریس U است که حاوی ماموریت‌های U_{ij} ها است و ماتریس کانکشن نهایی بین تمام محصولات و گره‌های دانه‌ی نافذ را نشان می‌دهد.

سطح سلسله‌مراتب توسط پارامتر H نشان داده می‌شود که تا زمانی افزایش می‌یابد که معیارهای توقف، رعایت شوند. در هر سلسله مراتب (H)، روی تمام گره‌ها (آنها) تکرار را در شبکه‌ی سلسله‌مراتب (E^H) انجام می‌دهیم و عامل‌های همسایه‌ی هر گره را فهرست می‌کنیم. شعاع همسایگی که با پارامتر ۲ نشان می‌دهیم، بیانگر دانه‌ی بودن تحلیل است. بر مبنای شعاع ۲، ما شبکه را به زیربخش‌هایی (E_i^H) تقسیم می‌کنیم و ماتریس‌های احتمال A , P را برای آن زیربخش آپدیت می‌کنیم. HIM با استفاده از یک تکنیک بهینه‌سازی، عامل‌های تاثیرگذار را در آن شبکه انتخاب می‌کند و آنها را برای استفاده‌ی آتی تگ می‌کند. فرایند انتخاب گره را به تفصیل در بخش 3.3.2 شرح خواهیم داد. آنگاه این گره‌های تاثیرگذار را به مجموعه‌ای از گره‌های تاثیرگذاری اضافه می‌کنیم که در سایر همسایگی‌ها -در همان سلسله‌مراتب-، شناسایی شده‌اند.

Table 1 HIM Algorithm

```

HIM (Agent, E, P, A, AR, Hmax, r)
H = 0
EH = E
NH = |AR|
While stopCriteria do
  H = H + 1
  infList = NULL
  for i = 1 to NH do
    neighborList = FindNeighborList (i, r, EH)
    EiH = Subgraph (neighborList, EH)
    EiH = AddOutsideWorld (EH, EiH)
    (Pi, Ai) = UpdateMat (EH, P, A, neighborList)
    L = Optimize (Agent, EiH, Pi, Ai)
    infList = infList ∪ L
  Agent = UpdateAgent (infList)
  end for
  NH = |infList|
  U = MakeU (Agent)
  stopCriteria = UpdateCriteria (infList, H)
  EH = UpdateHierarchy (infList)
end while
return U

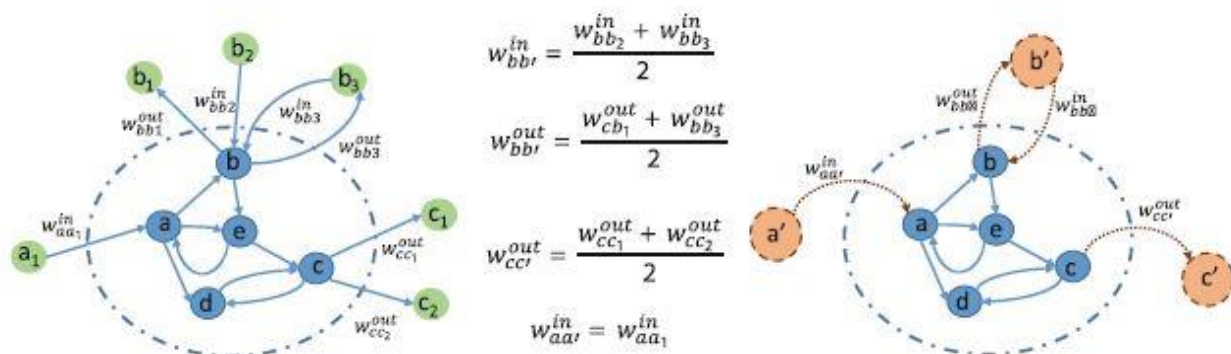
```

3.3.1 تاثیر دنیای بیرون

وقتیکه یک همسایگی محلی از شبکه‌ی کامل جدا می‌شود، گره‌هایی مرزی وجود دارند که به گره‌های بیرون از همسایگی کانتکت می‌شوند. این کانکشن‌ها که به بیرون از همسایگی می‌افتند، بصورت بالقوه می‌توانند بر بردار تمایل عامل‌ها -درون همسایگی-، تاثیر بگذارند. یک رویکرد احتمالی، نادیده گرفتن این تاثیرات و صرفاً مدنظر قرار دادن گره‌های درون پارتیشن است. در این فصل، با تخصیص یک گره مجازی به هر گره مرزی، این تاثیرات را مدنظر قرار می‌دهیم. این گره مجازی، نماینده‌ی تمام گره‌های بیرون از همسایگی است که به گره مرزی متصلند. تصویر 3 تجریدی از تاثیر دنیای بیرون را به تصویر کشیده است و نشان می‌دهد که پارامترهای مدل، به چه نحو بین هر مرز و گره مجازی محاسبه می‌شوند.

3.3.2 انتخاب گره

فرایند انتخاب گره‌های تاثیرگذار، در هر سلسله‌مراتب و در هر همسایگی محلی که i را احاطه کرده است، تکرار می‌شود. با پیروی از پژوهش‌های پیشین، ما پویایی مطلوب تمام عامل‌ها را بعنوان یک زنجیره‌ی مارکوف، مدل‌سازی می‌کنیم که در آن، وضعیت همسایگی محلی، ماتریسی از بردارهای تمایل عامل‌های موجود در هر تکرار خاص k می‌باشد و انتقال‌های وضعیت بصورت احتمالاتی از تعامل دوبه‌دوی بین عامل‌های متصل به شبکه محاسبه می‌شوند.



تصویر 3. شبکه‌ی سمت چپ، مثالی از همسایگی پیرامون گره e است؛ شبکه‌ی سمت راست، شبکه‌ی هم‌تراز با گره‌های مجازی می‌باشد که بیانگر تاثیر دنیای بیرون است. در اینجا، W می‌تواند هرگونه پارامتر تعامل نظیر وزن لینک، α ، ϵ باشد. جهت تعامل با گره مجازی، مبتنی بر نوع لینک‌هایی است که گره مرزی، با گره‌های بیرون از همسایگی دارد. مقدار پارامتر، متوسط تمام انواع تعاملات مشابهی است که با دنیای بیرون انجام می‌گیرد.

وضعیت شبکه‌ی محلی پیرامون عامل i در تکرار k ام، یک بردار متغیرهای راندم است که با

نشان داده می‌شود (که با اتصال بردارهای N_i^H به سایز P حاصل می‌شود) و اینگونه بیان

می‌گردد:

$$\mathbf{X}_i(k) = \begin{pmatrix} [\vec{X}_1(k)] \\ \vdots \\ [\vec{X}_{N_i^H}(k)] \end{pmatrix}$$

ما تمایل بلندمدت مورد انتظار عامل‌ها را در هر شبکه‌ی محلی، پیرامون عامل i محاسبه می‌کنیم و این محاسبه منتج

به فرمول زیر می‌شود:

$$E[X_i(k+1)] = E[X_i(k)] + Q_i E[X_i(k)]. \quad (5)$$

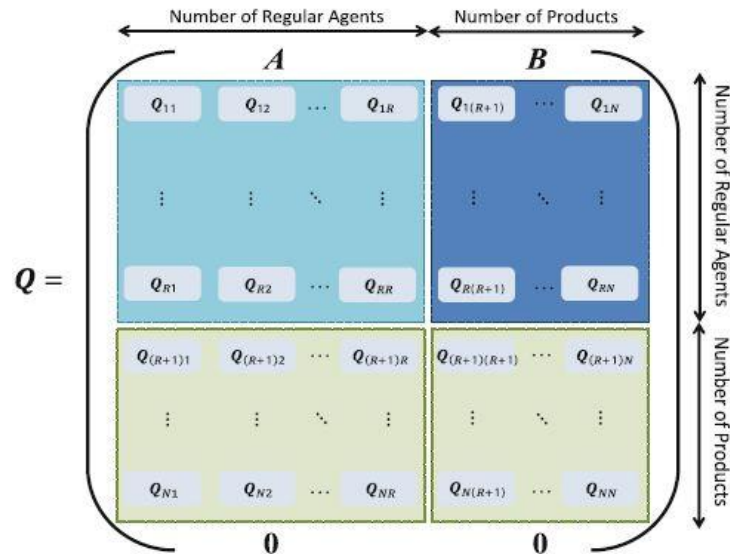
جهت حل کارآمد این سیستم معادلات، ماتریس‌ها را تجزیه می‌کنیم:

$$Q = \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \vec{\mu}_X(\infty) = \begin{pmatrix} \vec{\mu}_R \\ \vec{\mu}_P \end{pmatrix} \quad (6)$$

در اینجا $A \in \mathbb{R}^{RP \times RP}$ ماتریس فرعی است که بیانگر تعاملات مورد انتظاری است که قرار است بین عامل‌های R

صورت گیرد، در حالیکه $B \in \mathbb{R}^{RP \times P^2}$ بیانگر تعاملات بین عامل‌های R و عامل‌های P است. تصویر 4، تجزیه‌ی

ماتریس Q را نشان داده است.



تصویر 4. ماتریس Q یک ماتریس بلوکی با سائز $N \times N$ است که N تعداد کل عامل‌ها $(R+P)$ و سائز هر بلوک نیز

$P \times P$ است. ماتریس‌های A و B بخش غیرصفر این ماتریس هستند که تعاملات میان عامل‌های R و تعاملات بین

عامل‌های R و P را نشان می‌دهند.

علاوه بر این، بردارهایی هستند که بیانگر تمایل بلندمدت عامل‌های R و عامل‌های P در تکرار

می‌باشند. توجه کنید که بردار $\vec{\mu}_P$ معلوم است چراکه عامل‌های P -تبلیغ‌گران-، عامل‌های

تغییرناپذیری هستند که هیچگاه تمایلشان را تغییر نمی‌دهند. حل مساله برای $\vec{\mu}_R$ بردار تمایل بلندمدت مورد

انتظار را برای تمام عامل‌های **regular**، برای مجموعه‌ای مفروض از احتمالات تاثیرگذار روی یک شبکه‌ی اجتماعی قطعی، حاصل می‌کند.

$$\mathbf{A} \vec{\mu}_R + \mathbf{B} \vec{\mu}_P = 0 \Rightarrow \vec{\mu}_R = \mathbf{A}^{-1}(-\mathbf{B} \vec{\mu}_P) \quad (7)$$

از این رو می‌توان گره‌های تاثیرگذار در شبکه را شناسایی کرد و محصولات را به شیوه‌ای به آن عاملها متصل کرد که تمایل بلندمدت عامل‌ها را در سیستم اجتماعی بیشینه سازد. ما تابع هدف را اینگونه تعریف می‌کنیم: بیشینه‌سازی متوسط موزون تمایل بلندمدت مورد انتظار تمام عامل‌های R در شبکه به سوی تمام محصولات، بصورت:

$$\max_u \sum_{1 \leq k \leq P} \sum_{i \in A_R} (\rho_i \cdot \vec{\mu}_{R,i}) \quad (8)$$

$\vec{\mu}_{R,i}$ بخشی از $\vec{\mu}_R$ است که به عامل i تعلق دارد، و پارامتر ρ_i صرفاً وزنی است که می‌توانیم بر مبنای اهمیت عامل‌ها در شبکه، به آنها تخصیص دهیم. در صورتیکه برای تمام عاملها $\rho_i = 1$ باشد، تابع فوق به میانگین ریاضیاتی بردارهای مطلوب بلندمدت مورد انتظار برای تمام عاملها، کاهش می‌یابد.

3.3.3 همگرایی

با استفاده از قضیه‌ی نقطه‌ی ثابتِ براور، اثبات می‌کنیم که هر همسایگی محلی، از یک نقطه‌ی ثابت برخوردار است؛ بنابراین حل معادله 5 در یک وضعیت باثبات، یک انتخاب معتبر است. این قضیه بیان می‌دارد:

قضیه 1: هر تابع پیوسته‌ای از یک گوی بسته‌ی فضای اقلیدسی به خودش، دارای یک نقطه‌ی ثابت است.

طبق محاسبه‌ی معادله‌ی 5، $E[X_i(k+1)]$ یک تابع پیوسته است، چراکه حاصلجمع دو تابع پیوسته می‌باشد.

همچنین از آنجاکه $\vec{X}_i(k+1)$ در معادله 3، یک تابع کراندار در $[-1 \ 1]$ است، امید ریاضی آن

$(E[X_i(k+1)])$ نیز کراندار خواهد بود. در نتیجه، ما از یک تابع پیوسته‌ی کراندار برخورداریم که قضیه‌ی

نقطه-ثابتِ براور، تضمین می‌دهد که یک تابع نقطه-ثابت باشد. این به ما امکان می‌دهد که مساله‌ی خودمان با

الگوریتم بهینه‌سازی پیشنهادی مان حل کنیم تا تخصیص زلاها را به شیوه‌ای پیدا کنیم که بردار بلندمدت عاملها را برای تمام محصولات موجود در بازار، بیشینه سازد.

3.3.4 سلسله‌مراتب آپدیت

وقتیکه از یک سلسله‌مراتب به سلسله‌مراتب بعدی پیش می‌رویم، گره‌های منتخب که به سلسله‌مراتب بالایی انتشار می‌یابند، الزاماً همجوار نیستند. بنابراین، لازم است بر اساس جایگاه آنها در شبکه‌ی واقعی، مدلی تعاملی را بین آنها تعریف کنیم. تابع `UpdateHierarchy` مسئول ایجاد اتصال مناسب شبکه و مدل تعامل برای سلسله‌مراتب بعدی، بر مبنای گره‌های تاثیرگذار منتخب در سلسله‌مراتب فعلی می‌باشد. این گره‌ها با انتخاب شدن بعنوان گره‌های تاثیرگذار در دست‌کم یک همسایگی محلی، به سلسله‌مراتب بالاتر انتشار یافتند. برای یک گره، امکانپذیر است که در پارتیشن‌های متعددی حضور داشته باشد و بیش از یکبار انتخاب شود.

توجه کنید که احتمالش کم است که تمام گره‌های منتخب در شبکه‌ی واقعی (E)، گره‌های همسایه یا مجاور هم باشند. بنابراین لازم است روشی جهت ایجاد اتصال بین آنها ایجاد کنیم تا E^H شکل گیرد. جهت انجام این کار، به نزدیکترین مسیر بین این گره‌ها در شبکه‌ی E نگاه می‌کنیم و از آن جهت محاسبه‌ی وزن لبه‌ها در E^H استفاده می‌کنیم. در شبکه‌ی E^H ، وزن لینک بین دو گره منتخب، محصول اوزان کوتاه‌ترین مسیر بین این دو گره در سلسله‌مراتب پیشین است. همچنین احتمالات تعامل و تاثیر بین گره‌های تاثیرگذار، طور تعیین می‌شود که محصول احتمالات در راستای کوتاه‌ترین مسیر بین آنها باشد.

3.3.5 معیارهای پایان بخشی

جهت پایان دادن به یک لوپ، دو معیار مختلف را در تابع `UpdateCriteria` تعیین می‌کنیم. این تابع، معیارهای توقف را بر اساس سطح سلسله‌مراتب بررسی می‌کند و گره‌های تاثیرگذار را فهرست می‌کند. یک معیار مبتنی بر بیشینه تعداد سطوح سلسله‌مراتب است و دیگری بر مبنای نسبت گره‌های تاثیرگذار منتخب و بودجه‌ی تبلیغات

است. بر اساس خروجی `stopCriteria`، این الگوریتم تصمیم می‌گیرد که آیا به سلسله‌مراتب بالاتر برود یا جستجو را متوقف سازد و به ماتریس U بازگردد تا بعنوان ماموریت تبلیغاتی از آن استفاده شود.

3.3.6 رویه و روال بهینه‌سازی

بهترین تخصیص عاملهای P به عاملهای R از طریق حل مسأله‌ی بهینه‌سازی زیر حاصل می‌شود:

$$\begin{aligned} & \underset{\hat{u}}{\text{maximize}} && \|A^{-1} \text{Vec}(M \hat{\mu}_P \hat{u})\|_1 \\ & \text{subject to} && x_{ip} \in [-1, 1], \quad \forall i \in A_R, \\ & && \sum_{j \in A_R} u_{ij} = \text{cte}. \end{aligned} \quad (9)$$

در اینجا، به دنبال مجموعه‌ای از u ها هستیم که هزینه‌ی ما را بیشینه می‌سازد، یا بعبارت دیگر، مقدار تمایل عاملها را بیشینه می‌سازد. از آنجاکه u ها دلالت بر وجود یا عدم وجود اتصال بین عاملهای P و R دارند، متغیرهای باینری هستند و می‌توان با استفاده از برنامه‌سازی عدد صحیح مختلط، آنها را شناسایی کرد. جهت حل نمودن مسأله‌ی بهینه‌سازی مان، ما از بسته‌ی GLPK (کیت برنامه‌سازی خطی GNU) استفاده کردیم که جهت حل مسائل مربوط به برنامه‌سازی خطی بزرگ‌مقیاس و برنامه‌سازی عدد صحیح مختلط، طراحی شده است. GLPK مجموعه‌ای از روتین‌هاست که بصورت کُد ANSI نوشته شده است و بصورت کتابخانه‌ی قابل فراخوانی سازماندهی شده است که می‌توان آنرا از این آدرس دانلود کرد: www.gnu.org/software/glpk

4. ارزیابی

4.1 ستاپ آزمایشی

ما مجموعه‌ای از آزمایشات شبیه‌سازی را جهت ارزیابی اثربخشی روش انتخابی گره پیشنهادی مان در یک سیستم اجتماعی شبیه‌ساخته با یک شبکه‌ی استاتیک، ارزیابی کردیم. پارامترهای مدل تعاملی برای تمام اجراها، در جدول 2a ملاحظه می‌کنید. تمام نتایج از متوسط حدود 100 بار اجرا بدست آمده‌اند که شبیه‌سازی‌های مختلفی را روی هر ده ساختار شبکه نشان می‌دهند.

در تعاملات عامل‌های R و P ، پارامترهای α و ϵ برای یک تعامل مفروض، ثابت هستند و در جدول 2a ارائه شده‌اند. فرض را بر این می‌گذاریم که این پارامترها را بتوان توسط شرکتهای تبلیغاتی و بر مبنای مدلسازی کاربر محاسبه کرد. مقادیر P_{ij} برای این نوع تعامل را با استفاده از معادله 1 محاسبه می‌کنیم و پارامتری هستند. جدول 2b پارامترهای الگوریتم HIM را ارائه کرده است (شعاع همسایگی و بیشینه سطح سلسله‌مراتب). بخش باقیمانده‌ی ستاپ سیستم اجتماعی را با ماتریس M نشان می‌دهیم که همبستگی بین تقاضا برای محصولات مختلف را مدلسازی می‌کند. این ماتریس بصورت یکپارچه و با اعداد رانوم بین $[0 \ 1]$ تولید می‌شود، چراکه دارای یک تفسیر احتمالاتی می‌باشد، که حاصلجمع مقادیر در هر ردیف، نشان می‌دهد که کل تقاضا برای یک آیتم برابر با یک است.

جدول 2. تنظیمات پارامتری		
پارامتر	مقدار	شرح
الف) پارامترهای مدل بازار		
آستانه	2	تعداد لینک‌های بین عامل‌های R و P
ϵ	0.4	فاکتور تاثیر بین عاملهای P و R
α	0.8	احتمال تاثیرگذاری بین عاملهای P و R
R	متغیر	تعداد عاملهای Regular
P	10	تعداد عاملهای Product
$N_{Iterations}$	60000	تعداد تکرارها
N_{Run}	10	تعداد اجراها
N_{Net}	10	تعداد شبکه‌های مختلف
ب) پارامترهای HIM		
r	3	شعاع همسایگی
H_{max}	5	بیشینه سطح سلسله‌مراتب

4.2 معیارها

ما الگوریتم سلسله‌مراتبی خودمان را با ورژن غیرسلسله‌مراتبی⁸ و مجموعه‌ای از شاخص‌های مبتنی بر مرکزیت، که عموماً در آنالیز شبکه‌ی اجتماعی جهت شناسایی گره‌های تاثیرگذار بر مبنای ساختار شبکه، بکار می‌روند، مقایسه کردیم.

* OIM: روش بهینه‌سازی‌شده‌ی بیشینه‌سازی تاثیر، گره‌های تاثیرگذار را با استفاده از روش بهینه‌سازی ما، روی شبکه‌ی اریجنال پیدا می‌کند.

* درجه: با فرض اینکه گره‌های درجه‌بالا، گره‌های تاثیرگذار درون شبکه هستند، ما احتمال تبلیغ برای یک عامل R را بر مبنای درجه‌ی بیرونی عاملها محاسبه کردیم و عاملهای P را بر اساس یک مدل الحاقی ترجیحی، لینک کردیم. بنابراین، گره‌هایی با درجه‌ی بالاتر، از شانس بیشتری برای انتخاب شدن بعنوان هدف تبلیغاتی برخوردار بودند.

* بینیت: این متریک مرکزیت، تعداد دفعاتی را اندازه می‌گیرد که یک گره روی ژئودزیک ظاهر می‌شود و تمام دیگر گره‌های شبکه را به یکدیگر متصل می‌کند. گره‌هایی که دارای بالاترین مقدار بینیت بودند، بالاترین شانس انتخاب شدن بعنوان یک گره تاثیرگذار را داشتند.

* PageRank: بر مبنای این فرض که گره‌هایی که دارای بالاترین امتیاز PageRank هستند، شانس بالاتری در تاثیرگذاری بر سایر گره‌ها دارند، احتمال انتخاب گره را بر مقدار PageRank آن قرار دادیم.

* رانوم: در این بیسلاین، صرفاً گره‌ها را بصورت یونیفورم انتخاب می‌کنیم.

جهت ارزیابی این روشها، شبیه‌سازی را با بردار تمایل اولیه که برای تمام عاملها برابر با 0 تنظیم شده است، شروع کردیم و 60000 تکرار از تعاملات عامل را شبیه‌سازی نمودیم. کل فرایند تعامل و تاثیر، در کنترل معادله‌ی 3 و 4 (بخش 3.2) قرار داشت. در هر تکرار، متوسط مقدار تمایل مورد انتظار عاملها را نسبت به تمام محصولات محاسبه کردیم. این مقدار متوسط، حاصل متوسط‌گیری از 100 اجرا (10 شبیه‌سازی روی 10 ساختار شبکه‌ای متفاوت) برای دیتاست ساختگی و 100 اجرا روی دیتاست دنیای واقعی بود. توجه کنید که بردار تمایل عاملهای P، برای

⁸ روش بهینه‌سازی‌شده‌ی بیشینه‌سازی تاثیر (OIM) که در رفرنس 21 به آن پرداخته شده است.

تمام محصولات ثابت باقی می‌ماند؛ در شبیه‌سازی ما، این مقدار برای خودِ محصول برابر با 1 و برای سایر محصولات،

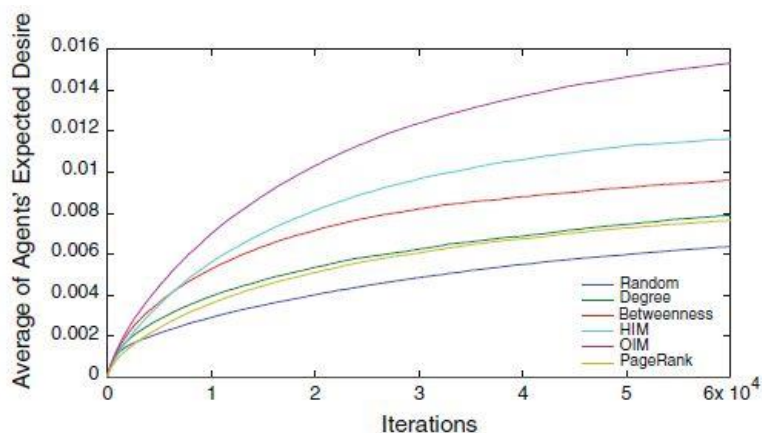
برابر با 0.1- در نظر گرفته شد. (e.g., $\mu_1 = [1 - 0.1 - 0.1 \dots - 0.1]$).

4.3 دیتاستِ ساختگی

برای این دیتاستِ ساختگی، از همان تکنیک تولید شبکه استفاده کردیم که در رفرنس 21 برای تولید شبکه‌های مشتری، شرح دادیم. جهت مقایسه‌ی عملکرد این روشها، متوسط مقدار تمایلِ عامل‌ها در یک شبکه‌ی 150 عاملی، در طول زمان در تصویر 5 نشان داده شده است. در اینجا 150 عامل را بعنوان تعداد بهینه‌ی عامل‌ها انتخاب کردیم تا تمام الگوریتم‌ها را در کنار هم مقایسه نماییم. با عامل‌های کمتر، برخورداری از ده محصول بطور همزمان، شبکه را اشباع می‌کند در حالیکه با تعداد بیشتری از عامل‌ها (OIM)، دچار مسائل مقیاس‌پذیری می‌شود.

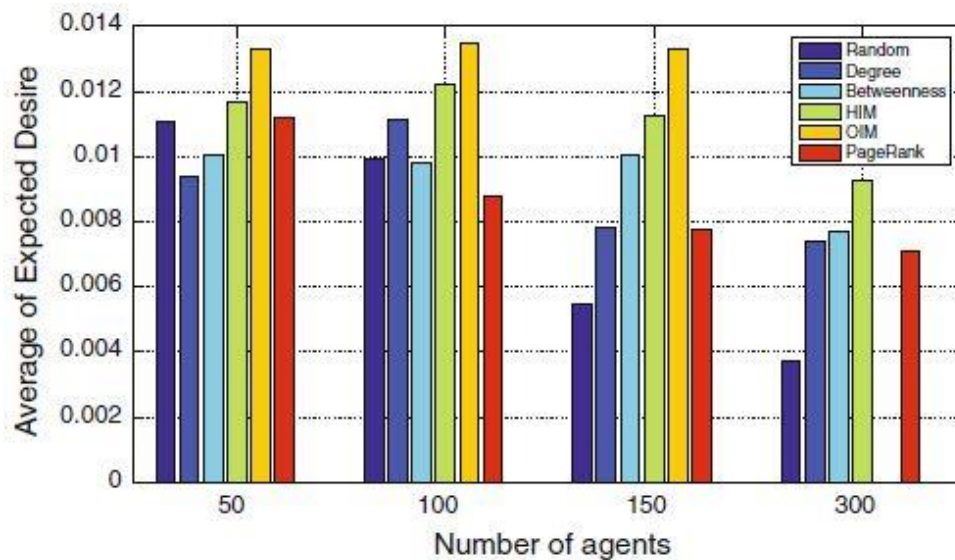
4.3.1 اثربخشی بازاریابی

در تصویر 5، با استفاده از روشهای بهینه‌سازی مختص بازاریابی برای تخصیص بودجه‌ی تبلیغات، مقدار تمایلِ عامل‌ها به تمام محصولات، بیشترین افزایش را می‌یابد، که منتج به بیشترین فروش می‌شود. هرچند که HIM مقداری از عملکرد را فدای مقیاس‌پذیری کرده است، اما بوضوح عملکردش از روشهای سنجش مرکزیت بهتر است. حاصل رویکرد انتخاب بهینگی محلی HIM، عملکرد اندکی پایین‌تر از OIM بهینه‌ی فراگیر است.



تصویر 5. متوسط تمایل مورد انتظار عاملها در مقابل تعداد تکرارها، که روی تمام محصولات و طی 100 اجرا محاسبه شده است (10 اجرای مختلف روی 10 شبکه‌ی مختلف). روشهای بهینه‌سازی، از بالاتری متوسط در مقایسه با روش‌های اکتشافی سنجش مرکزیت برخوردارند. از آنجاکه HIM یک روش فروبهن است، تعجبی ندارد که عملکرد آن بدتر از روش بهینه‌سازی فراگیر (OIM) است.

تصویر 6 مقدار متوسط نهایی تمایل مورد انتظار عاملها را در آخرین تکرار برای تعداد مختلف عاملهای R، نشان می‌دهد. هرچند که OIM با روش بهینه‌سازی فراگیر، عملکرد بهتری نسبت به HIM و سایر روشهای سنجش مرکزیت دارد، اما به دلیل ماتریس تعامل نزدیک به واحد، قادر نیست به 300 عامل و بیشتر در شبکه دست یابد. HIM با توانایی خود در افزایش خطی، یک راه‌حل فروبهن و در عین حال عملی را در انتخاب گره‌های تاثیرگذار در شبکه‌های بزرگ، ارائه می‌دهد.



تصویر 6. متوسط بردارهای نهایی مورد انتظار برای تعداد مختلفی از عاملهای R (Regular) و 10 عامل P (Product). روشهای مبتنی بر بهینه‌سازی (OIM و HIM) نسبت به سایر روشها، در انتخاب گره‌های دانه‌ای، عملکرد بهتری دارند. در حالیکه OIM موفقیت بیشتری نسبت به HIM در انتخاب گره‌های تاثیرگذار دارد، اما قادر نیست برای شبکه‌هایی با 300 عامل یا بیشتر، مقیاس خود را بزرگ کند.

Number of agents	OIM (s)	HIM (s)
50	10.67	74.09
100	94.76	160.80
150	290.67	208.97
200	897.51	354.35

جدول 3. مقایسه‌ی روتین بین OIM و HIM.

4.3.2 زمان اجرا

جدول 3 یک مقایسه‌ی زمان اجرا را بین دو روش بهینه‌سازی HIM- (روش پیشنهادی ما) و OIM (روش اریجنال) نشان می‌دهد. در شبکه‌های کوچک، زمان اجرای روش بهینه‌سازی فراگیر، کمتر از زمان اجرای روش سلسله‌مراتبی است اما با افزایش سایز شبکه، زمان اجرای آن بصورت نمایی افزایش می‌یابد، در حالیکه زمان اجرای HIM با نرخ

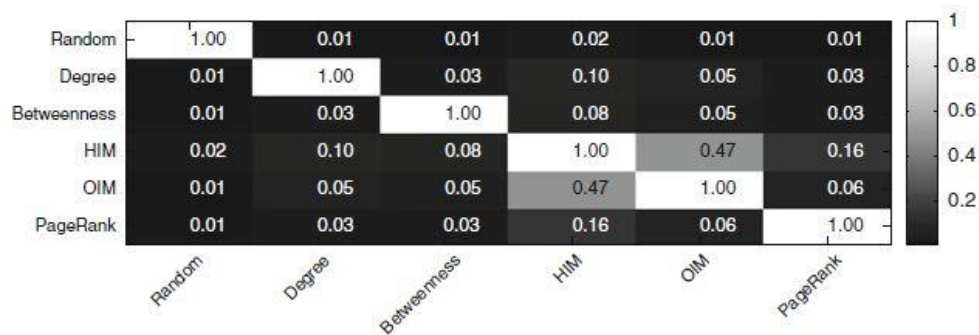
کمتری بالا می‌رود. زمان اجرای طولانی OIM برای شبکه‌هایی که بیش از 200 گره دارند، الگوریتم را برای یافتن گره‌های تاثیرگذار در شبکه‌های بسیار بزرگ، غیرعملی می‌سازد.

4.3.3 تشابه ژاکارد

جهت تحلیل تفاوت‌های بین انتخاب الگوریتم‌های گره‌ها، از سنجش تشابه ژاکارد استفاده می‌کنیم. این سنجش با استفاده از تقسیم تعداد اشتراک دو مجموعه‌ی منتخب بر تعداد اتحاد این دو مجموعه حاصل می‌شود. تصویر 7 این سنجش را برای تمام جفت الگوریتم‌ها نشان می‌دهد. الگوریتم‌های OIM و HIM دارای بالاترین تشابه (مقدار تشابه: 0.47) در مقایسه با سایر روش‌هاست. سایر جفت روش‌ها دارای تشابهات بسیار کمی هستند، که باعث تشکیل مربع‌های تیره‌ری تصویر می‌شوند. تعجبی ندارد که «راندوم» دارای کمترین تشابه در انتخاب گره‌ها نسبت به سایر روش‌هاست. این مساله نشان می‌دهد که HIM بسیاری از همین گره‌ها را بصورت الگوریتم اریجنال OIM - با هزینه‌ی اجرای بسیار پایین‌تر - پیدا می‌کند.

4.4 دیتاست (مجموعه‌ی داده‌ها) دنیای واقعی

ما عملکرد و مقیاس‌پذیری الگوریتم پیشنهادی‌مان را روی شبکه‌های دنیای واقعی در پروژه‌ی تحلیل شبکه‌ی استانفورد، ارزیابی نمودیم. آدرس اینترنتی: snap.stanford.edu



تصویر 7. متوسط سنجش‌های تشابه ژاکارد، بین روش‌های مختلف که از 100 اجرا بدست آمده اتس (10 اجرا روی 10 شبکه‌ی مختلف). مربع‌های روشن‌تر بیانگر تشابه بیشتر بین یک جفت الگوریتم هستند. توجه کنید که انتخاب HIM گره‌ها، تقریباً به انتخاب بهینه‌ی OIM نزدیک است.

* WikiVote: شبکه‌ای که حاوی دیتای رای‌گیری ویکی‌پدیا، از زمان شروع به کار ویکی‌پدیا در ژانویه سال 2008 است. گره‌های درون شبکه، کاربران ویکی‌پدیا را نشان می‌دهند و لبه‌ی جهت‌مند از گره A تا گره B بیانگر این است که کاربرد A به کاربر B رای داده است.

* SlashDots: یک وبسایت اخبار مرتبط با فناوری است که به خاطر جامعه‌ی محلی کاربرانش معروف است. این وبسایت شامل اخبار مرتبط با فناوری است که کاربران روی وبسایت قرار می‌دهند. این وبسایت در سال 2002 به معرفی ویژگی Slashdot Zoo پرداخت که به کاربران اجازه می‌دهد همدیگر را بصورت دوستان یا مخالفان، تگ نمایند. این شبکه شامل لینک‌های دوست/مخالف بین کاربران Slashdot است که در فوریه 2009 بدست آمده است.

* Epinions: این شبکه‌ای است که از سایت Epinions.com استخراج می‌شود. گره‌ها اعضای از سایت هستند که محصولات را مورد بررسی قرار داده‌اند. یک لبه‌ی جهت‌مند از A به B بیانگر این است که B به بررسی‌های A مرتبط است (و از این رو A بر B تاثیرگذار است).

در تمام آزمایشاتی که روی رسانه‌های اجتماعی دنیای واقعی انجام گرفته‌اند، ما شبکه‌ها را پیش‌پردازش کردیم، تا گره‌های ایزوله و گره‌های مرزی را حذف نماییم (گره‌هایی با درجه‌ی یک). جدول 4a,b آمار این شبکه‌های دنیای واقعی را قبل و بعد از مراحل پیش‌پردازش نشان می‌دهند. ما از پارامترهای آزمایشی مشابهی استفاده کردیم (که در

بخش 4.1 خواهید خواند). تنها تفاوت‌هایی که وجود دارند، تعداد محصولات و بودجه‌ی تبلیغاتی است که به ترتیب برابر با 10 و 50 می‌باشند.

ما روش‌های بهینه‌سازی‌مان را در مقابل دو روش برجسته‌ی بیشینه‌سازیِ تاثیر، آزمودیم: یعنی PMIA و DegreeDiscount، در کنار سنجش‌های مرکزیت.

* PMIA: الگوریتم جستجوی کاشف یا هیوریستیک (رفرنس 25) که همسایگی محلی هر گره را جهت یافتن الگوی تاثیر در هر درخت‌سانی محلی مورد بررسی قرار دادیم تا انتشار تاثیر را روی شبکه برآورد نماییم. تا آنجا که می‌دانیم، الگوریتم PMIA مقیاس‌پذیرترین راه‌حل در قبال مسأله‌ی بیشینه‌سازیِ تاثیر تحت مدل مستقل آبشاری است.

* DegreeDiscount: این الگوریتم هیوریستیک که چن⁹ و آنرا ارائه کرده است، با تنزیل درجه‌ی گره‌ها، در جایی که یک همسایه قبلاً بعنوان یک گره تاثیرگذار شناخته شده است، روش درجه را اصلاح می‌کند.

Dataset	WikiVote	SlashDot	Epinion
(a) Before pre-processing			
<i>#Nodes</i>	7 K	82 K	76 K
<i>#Edges</i>	100 K	950 K	509 K
<i>Average Degree</i>	14.6	13.4	6.7
<i>Maximal Degree</i>	1,167	3,079	3,079
<i>Diameter</i>	7	11	14
(b) After pre-processing			
<i>#Nodes</i>	2 K	72 K	20 K
<i>#Edges</i>	38 K	840 K	3700
<i>Average Degree</i>	31.1	10.5	28.9
<i>Maximal Degree</i>	714	5,059	256
<i>Diameter</i>	7	13	12

جدول 4. آمار شبکه‌های دنیای واقعی.

هرچند که استفاده از رویکرد سلسله‌مراتبی باعث کاهش مسأله‌ی رسیدگی به ماتریس‌های تعاملی بزرگ می‌شود، اما همچنان برای پارتیشن‌های شبکه امکان‌پذیر است که به اندازه‌ی کافی بزرگ باشند، در صورتیکه روی یک گره درجه بالا تمرکز داشته باشند که به بخش بزرگی از شبکه متصل شده باشد. این گره‌ها علاوه بر خلق ماتریس‌های تعاملی

⁹ Chen

بزرگ، زیرگراف‌های ستاره‌ای شکل را خلق می‌کنند، که منتج به یک راه‌حل انجام‌پذیر برای فرایند بهینه‌سازی می‌شود. یکی دو راه‌حل برای پرداختن به این گره‌های بسیار درجه‌بالا وجود دارد: 1) وقتی که شبکه را پارتیشن می‌کنیم، آنها را نادیده بگیریم و فرض کنیم که هم‌پیوندی بالای آنها تضمین می‌دهد که آنها درون همسایگی شبکه‌ی سایر گره‌ها ظاهر خواهند شد یا 2) برخی از همسایه‌های درجه‌پایین گره را نادیده بگیریم. در آزمایشات زیر، نخستین رویکرد را در پرداختن به این پارتیشن‌های بزرگ اتخاذ کردیم. بنابراین، در تمام شبکه‌ها فقط پارتیشن‌ها را حول گره‌هایی با درجه‌ی کمتر 100 تمرکز دادیم. بررسی متوسط درجه‌ی گره‌ها در تمام دیتاست‌هایی که در جدول 4b ملاحظه می‌کنید، نشان می‌دهد که این انتخاب هم از ماتریس‌های بزرگ و زیرگراف‌های ستاره‌ای شکل اجتناب می‌کند و هم به ما درصد بالایی از گره‌ها را جهت پردازش ارائه می‌دهد. نتایج زیر، برای دیتاست‌های WikiVote و Epinion تولید شده‌اند.

4.4.1 اثربخشی بازاریابی

تصویر 8 متوسط مقدار تمایل مورد انتظار را برای تمام عاملها در طول زمان و برای 300K تکرار بازار شبیه‌سازی شده، نشان می‌دهد. در این نتیجه، الگوریتم OIM بالاترین مقدار را دارد، در حالیکه الگوریتم HIM از آن دقیقاً پیروی می‌کند. عملکرد الگوریتم HIM، به روش بهینه‌سازی فراگیر (OIM) نزدیک می‌شود. عملکرد هیوریستیک DegreeDiscount (PMIA)، و الگوریتم‌های PageRank – بدون اختلافات قابل توجه – بسیار به یکدیگر نزدیک است.

در حالیکه الگوریتم‌های ما عملکرد بهتری نسبت به سایر الگوریتم‌ها روی دیتاست WikiVote دارد، اما روی دیتاست Epinion، الگوریتم‌های درجه‌بنیان (مبتنی بر درجه)، عملکرد بهتری دارند. تصویر 9 نتایج تمام بنچمارک‌ها و الگوریتم HIM را نشان می‌دهد. هرچند که عملکرد HIM بهتر از PMIA و PageRank است، اما بر الگوریتم‌های درجه‌بنیان – نظیر Degree و DegreeDiscount غلبه نمی‌کند.

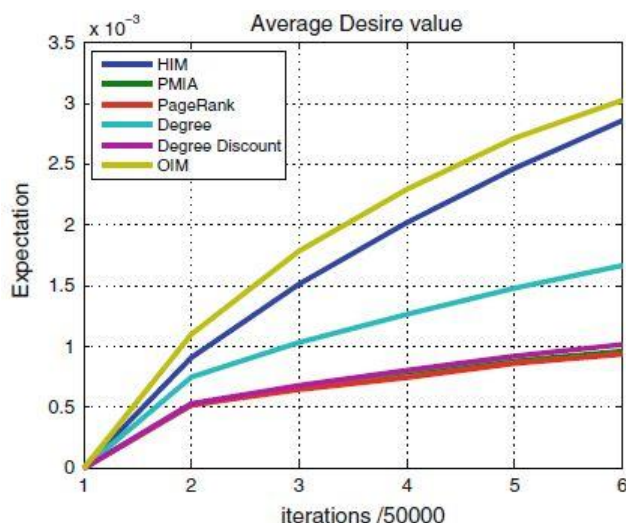
تصویر 10 مقدار نهایی تمایل مورد انتظارِ عاملها را برای الگوریتم‌های مختلف و برای دیتاست‌های مختلف نشان می‌دهد. مقدار پایین بردار تمایل، نتیجه‌ی برخورداری از تعداد پایین تبلیغات درون شبکه‌ی بزرگ است؛ هنگام انتشار تاثیر، بردارهای تمایل عامل، مکرراً در ϵ و α ضرب می‌شوند.

4.4.2 تحلیل توزیع‌های درجه‌ی دیتاست

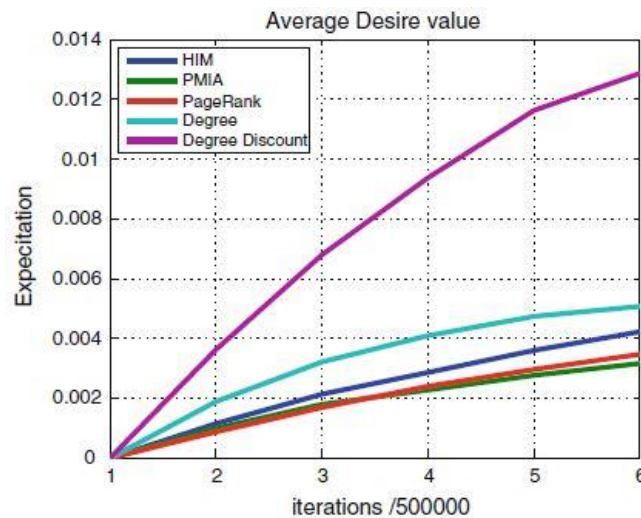
جهت درک عملکرد ضعیف HIM روی دیتاست Epinion، ما ساختار شبکه را بررسی کردیم تا دریابیم که شبکه‌ها تا چه اندازه با هم متفاوتند. جدول 5 آنالیز چندک درجه‌ی گره را برای دیتاست پیش‌پردازش‌شده نشان می‌دهد. بر مبنای این تحلیل، ملاحظه می‌کنیم که شبکه‌ی WikiVote یک شبکه‌ی بسیار کوچک در مقایسه با دو دیتاست دیگر است، با این وجود ماکسیمم درجه‌ی چارک‌های پایین‌تر، بالاتر از سایر شبکه‌هاست. این نشان می‌دهد که شبکه‌ی WikiVote دارای یک توزیع درجه‌ی یونیفورم‌تر است، که در آن درجه‌ی گره، یک ویژگی بسیار متمایز از پتانسیل انتشار تاثیر نمی‌باشد.

Dataset	0 %	25 %	50 %	75 %	100 %
WikiVote	3	25	44	79.25	714
Epinion	0	6	11	33	2,684
SlashDot	3	4	7	17	5,061

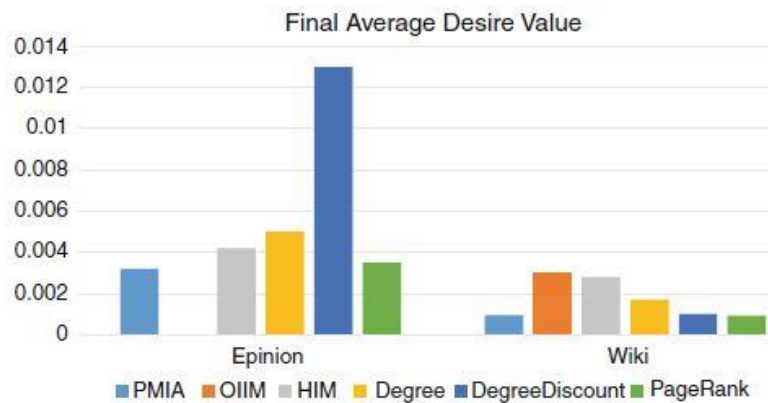
جدول 5. آنالیز چندک درجه‌ی گره در دیتاست پیش‌پردازش شده.



تصویر 8. متوسط تمایل مورد انتظار از عامل‌ها در مقابل تعداد تکرارها برای دیتاست WikiVote، که از تمام محصولات و با 100 اجرا محاسبه شده است. با حذف گره‌های ایزوله‌شده و مرزی، که 2000 گره بدست داد، دیتا پیش‌پردازش شد و شبیه‌سازی برای 300k تکرار، اجرا گردید. روشهای بهینه‌سازی دارای بالاترین متوسط در مقایسه با بقیه‌ی بنچمارک‌ها است. از آنجاکه الگوریتم HIM یک روش فروبهبین است، عملکرد آن کمتر از روش بهینه‌سازی فراگیر است.

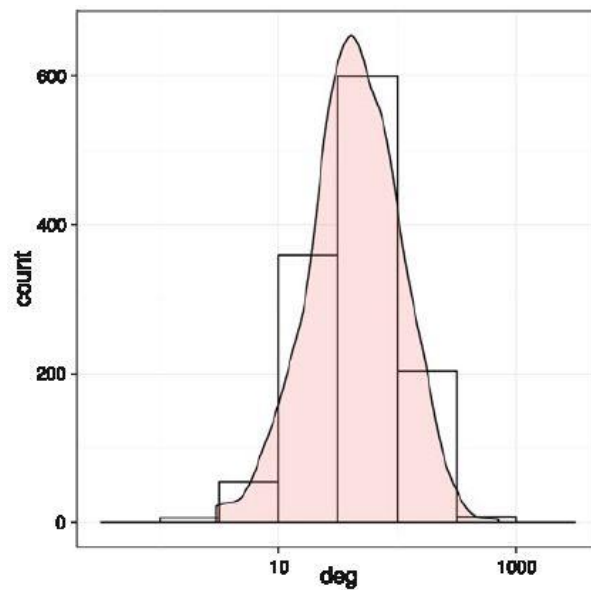


تصویر 9. متوسط تمایل مورد انتظارِ عاملها در مقابل تعداد تکرارها برای دیتاست Epinion، که طی 100 اجرا برای تمام محصولات محاسبه شده است. این دیتاست با حذف گره‌های ایزوله و مرزی، که نتیجه‌ی آن 20k گره بود، پیش‌پردازش گردید و شبیه‌سازی برای 300k تکرار اجرا گردید. HIM عملکرد بهتری نسبت به PMIA و PageRank دارد اما مغلوب الگوریتم‌های درجه‌بنیان Degree و DegreeDiscount- گشت. الگوریتم OIM را نمی‌شد به دلیل سائز شبکه، روی این دیتاست اجرا کرد.



تصویر 10. مقدار نهایی تمایل مورد انتظارِ عاملها در انتهای شبیه‌سازی، برای روشها و دیتاست‌های مختلف. به دلیل

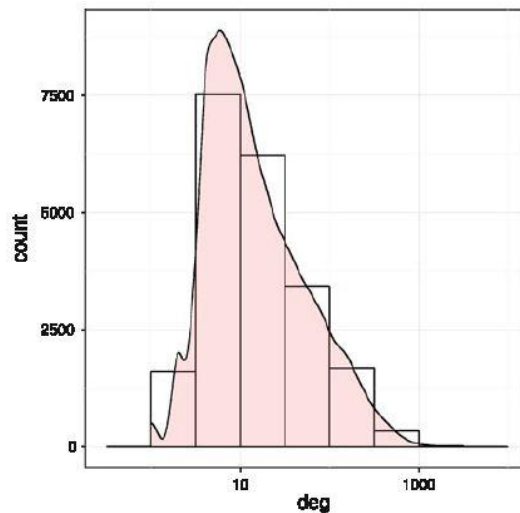
سایز شبکه، الگوریتم OIIM را نمی‌شد روی دیتاست Epinion اجرا کرد.



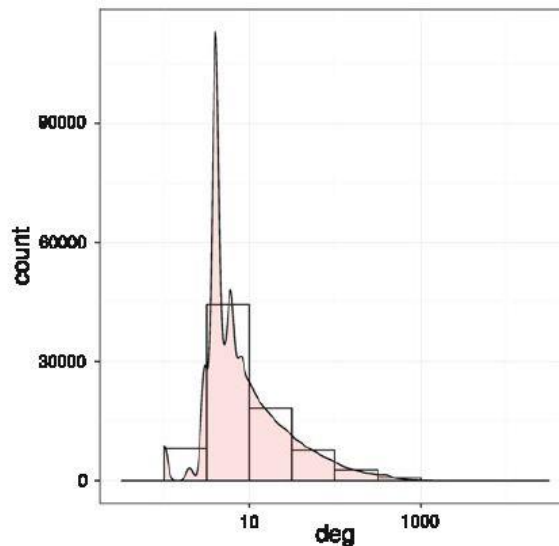
تصویر 11. هیستوگرام درجه‌ی دیتاست WikiVote. محور x مقیاس لگاریتمی درجه را نشان می‌دهد و منحنی نیز

برآورد چگالی هسته‌ها نشان می‌دهد. در این دیتاست، اکثر گره‌ها در رنج میانه قرار دارند و درجه‌ای بین 50 و 100

دارند.



تصویر 12. هیستوگرام درجه‌ی دیتانست Epinion. محور X مقیاس درجه‌ی لگاریتمی و منحنی، برآورد چگالی تراکم را نشان می‌دهد. در این دیتانست، شبکه از یک ساختار پراکنده برخوردار است که اکثریت گره‌ها دارای درجه‌ای کمتر از 10 هستند.



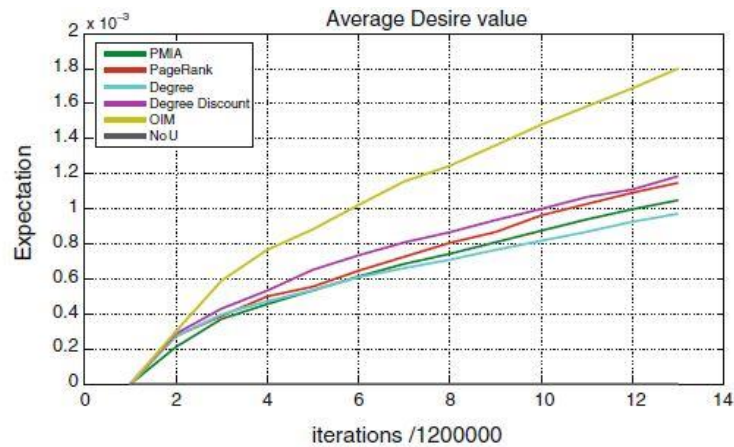
تصویر 13. هیستوگرام درجه‌ی دیتانست SlashDot. محور X مقیاس لگاریتمی درجه و منحنی، برآورد تراکم هسته‌ها را نشان می‌دهند. در این دیتانست، مانند دیتانست Epinion، شبکه از یک ساختار پراکنده برخوردار است، که اکثریت گره‌های آن از یک درجه کمتر از 10 برخوردارند.

با نگاهی به توزیع درجه‌ی دیتانست‌ها (تصویر 11 و 12 و 13)، این موضوع را می‌توان تایید کرد. در دیتانست Epinion و SlashDot، ما کوچکترین تعداد گره‌ها را با درجات بسیار بالا داریم، در حالیکه اکثر گره‌ها در شبکه، از

درجه‌ی پایین‌تر از 10 برخوردار هستند. در این شبکه‌ها، تعداد کمی از گره‌ها بعنوان هاب (hub) عمل می‌کنند و اتصال محکمی دارند، در حالیکه سایر گره‌ها اتصالات اندکی دارند و در بدترین حالت، اصلاً به گره درجه‌بالا متصل نیستند. بنابراین روش اکتشافی ما در عدم تمرکز پارتیشن‌ها پیرامون گره‌های درجه‌بالا، عملکرد فرایند بهینه‌سازی HIM را تخریب می‌کند. از سوی دیگر، الگوریتم‌های درجه‌بنیان، نظیر WikiVote یا شبکه‌های ساختگی که در آنها درجه‌ی گره یونیفورم‌تر است، HIM کارآمد است چراکه گره‌ها در bin‌های میانی، پرشمارتر هستند و اتصال بهتری با کل شبکه دارند. در این حالت و از آنجاکه «درجه» یک تمایزگر محسوب نمی‌شود، الگوریتم‌های درجه‌بنیان عملکرد ضعیفتری دارد.

4.4.3 بهینه‌سازی با روش اکتشافی درجه‌بنیان

بر مبنای این نتیجه، ما فرایند پیش‌پردازش‌مان را اصلاح کردیم تا از یک هیوریستیک درجه‌بنیان جهت انتخاب گره‌هایی که توسط تکنیک بهینه‌سازی‌مان مدنظر قرار گرفتند، استفاده کنیم. در اینجا، 5 درصد از درجه‌بالاترین گره‌ها را در دیتاست Epinion انتخاب کردیم و یک شبکه‌ی مجرد تک‌سطحی را بر مبنای کوتاه‌ترین مسیر، میان این گره‌ها خلق کردیم. آنگاه تکنیک بهینه‌سازی‌مان (OIM) را روی شبکه‌ی واحد اجرا کردیم. تصویر 14 نتیجه‌ی OIM و سایر بنچمارک‌ها را روی این شبکه‌ی پیش‌پردازش‌شده نشان می‌دهد. این نتیجه نشان می‌دهد که بکارگیری بهینه‌سازی در شبکه‌ی مجرد، قطعاً از سایر بنچمارک‌ها عملکرد بهتری دارد.



تصویر 14. متوسط تمایل مورد انتظارِ عامل‌ها در مقابل تعداد تکرارها برای دیتاست Epinion، که روی تمام محصولات و روی ده اجرای مختلف، برای 300k تکرار محاسبه گردید. این دیتاست با انتخاب 1٪ از گره‌های درجه‌بالا و ایجاد یک زیرگراف بر مبنای کوتاهترین مسیر بین این گره‌ها، پیش‌پردازش گردید و گراف را آنقدر کوچک ساخت که بتواند مستقیماً با OIM پردازش شود. OIM عملکرد بهتری نسبت به روشهای درجه‌بنیان دارد.

5. نتیجه و کارهای آتی

در این فصل، به مسأله‌ی پیشینه‌سازی تأثیر در شبکه‌های اجتماعی، به‌منظور تبلیغات پرداختیم. در حیطه‌ی تبلیغات، هدف ما شناسایی گره‌های تأثیرگذار در یک شبکه‌ی اجتماعی بعنوان اهداف تبلیغگر بر مبنای ساختار شبکه، تعاملات میان عاملها در شبکه و بودجه‌ی محدود تبلیغات می‌باشد. ما مدلسازی عامل‌بنیان را اختیار کردیم تا چنان سیستم اجتماعی را مدلسازی کنیم، چراکه ابزاری توانمند جهت بررسی پدیده‌هایی است که بررسی آنها درون آزمایشگاه دشوار است. همچنین تلاش کردیم بکارگیری فاکتورهایی نظیر همبستگی محصول و عضویت گروهی از عاملها، بازار، تعاملات و انتشار تأثیر و انتخاب کالا را واقع‌گرایانه‌تر مدلسازی نماییم.

در اینجا یک رویکرد عمومی سلسله‌مراتبی را برای بکارگیری تکنیک‌های بهینه‌سازی جهت پیشینه‌سازی تأثیر ارائه می‌دهیم. مزیت روش ما بر تکنیک‌های انتخاب دانه این است این روش می‌تواند همبستگی‌ها و تأثیرات جامعه را بر نرخ انتخاب محصول مدنظر قرار دهد. روش ما به انتخاب بهینه‌ی گره — با هزینه‌های روتین بسیار پایین‌تر —، نزدیک

می‌شود. اما تحلیل پیشینی از توزیع درجه‌ی شبکه از شبکه، جهت تشخیص پیش‌پردازش صحیح و فرایند تجزید، ضروری است. الگوریتم HIM را می‌توان جهت بهبود مقیاس‌پذیری بیشینه‌سازی تاثیر روی شبکه‌ها با یک توزیع درجه‌ی شبه‌یونیفورم، بکار برد. در شبکه‌هایی که دارای مرکزگرایی بالایی هستند، بکارگیری تکنیک بهینه‌سازی‌مان برای یک نسخه‌ی مجرد از شبکه که از گره‌های درجه‌بالا تشکیل شده است، توصیه می‌شود. در این فصل، ما به معرفی یک رویکرد جهت پارتیشن‌بندی شبکه به بخش‌های همپوشان و اجرای بیشینه‌سازی تاثیر بر پارتیشن‌ها پرداختیم. یک راه دیگر این است که تقسیمات فعلی شبکه را که با الگوریتم‌های انتخاب جامعه محاسبه شده‌اند، برای نخستین سطح سلسله‌مراتب، بیشتر کنیم. بعلاوه، کار با شبکه‌های پویایی که در آن عاملها می‌توانند وارد شبکه شوند یا از آن خارج شوند، برای مصارف عملی که در آن مجموعه‌ی مشتریان مدام در حال تغییر است، می‌تواند مفید واقع شود.

یک ضمیمه‌ی بالقوه مهم برای این اثر می‌تواند تعمیم شبیه‌سازی بازار، جهت مدلسازی صریح از اثرات نامساعد بین تبلیغ‌گران رقابتجو بعنوان رقابت Stackelberg باشد، که در آن یک تبلیغ‌گر یک آگهی قرار می‌دهد و رقبای بعدی، از وجود آن آگهی آگاهند. در این فصل، فرض را بر این گذاشتیم که احتمال تعامل و تاثیر بین دو عامل، کم است؛ در مقایسه با سایز شبکه که باعث می‌شود عاملها در آن به مدت قابل قبولی به یک تصمیم بچسبند. اما اگر شبکه کوچکتر باشد یا احتمال تعامل افزایش یابد، نوسانات زیادی در بردار تمایل عاملها بوجود می‌آید. بکارگیری یک پارامتر در مدلی که عاملها را مجبور می‌کند تصمیماتشان را به مدت کمی حفظ کنند، فارغ از تعاملات بیرونی، این مساله را اصلاح خواهد کرد. یک چارچوب عمومی‌تر برای مدلسازی و شبیه‌سازی انتخاب محصول توسط مشتری درون شبکه‌های اجتماعی، از اهمیت زیادی برخوردار است؛ مدل ما گامهای اولیه‌ای در راستای این هدف بلندپروازانه برداشته است.

References

1. Anagnostopoulos A, Kumar R, Mahdian M (2008) Influence and correlation in social networks. In: *Proceeding of the ACM SIGKDD international conference on knowledge discovery and data mining*, pp 7–15
2. Apolloni A, Channakeshava K, Durbeck L, Khan M, Kuhlman C, Lewis B, Swarup S (2009) A study of information diffusion over a realistic social network model. In: *Proceedings of the international conference on computational science and engineering*, pp 675–682
3. Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341
4. Bagherjeiran A, Parekh R (2008) Combining behavioral and social network data for online advertising. In: *IEEE international conference on data mining workshops (ICDMW)*, pp 837–846
5. Bharathi S, Kempe D, Salek M (2007) Competitive influence maximization in social networks. In: Deng X, Graham FC (eds) *Internet and network economics*. Springer, Berlin, pp 306–311
6. Borodin A, Filmus Y, Oren J (2010) Threshold models for competitive influence in social networks. In: Saberi A (ed) *Internet and network economics*. Springer, Berlin, pp 539–550
7. Chen W, Collins A, Cummings R, Ke T et al (2011) Influence maximization in social networks when negative opinions may emerge and propagate. In: *Proceedings of the SIAM international conference on data mining*
8. Chen W, Wang C, Wang Y (2010) Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, pp 1029–1038
9. Chen W, Wang Y, Yang S (2009) Efficient influence maximization in social networks. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, pp 199–208
10. Chen W, Yuan Y, Zhang L (2010) Scalable influence maximization in social networks under the linear threshold model. In: *Proceedings of the IEEE international conference on data mining (ICDM)*, pp 88–97
11. Hartline J, Mirrokni V, Sundararajan M (2008) Optimal marketing strategies over social networks. In: *Proceeding of the international conference on world wide web*. ACM, pp 189–198
12. Hung B (2010) Optimization-based selection of influential agents in a rural Afghan social network. Master's thesis, Massachusetts Institute of Technology
13. Hung B, Koltiz S, Ozdaglar A (2011) Optimization-based influencing of village social networks in a counterinsurgency. In: *Proceedings of the international conference on social computing, behavioral-cultural modeling and prediction*, pp 10–17
14. Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, pp 137–146
15. Kempe D, Kleinberg J, Tardos É (2005) Influential nodes in a diffusion model for social networks. In: *Automata, Languages and Programming*, pp 1127–1138
16. Kimura M, Saito K (2006) Tractable models for information diffusion in social networks. In: *Knowledge discovery in databases (PKDD)*, pp 259–271
17. Kimura M, Saito K, Nakano R, Motoda H (2009) Finding influential nodes in a social network from information diffusion data. *Social computing and behavioral modeling*. Springer, New York, pp 1–8
18. Leborgne D (1982) *Calcul différentiel et géométrie*. Presses universitaires de France
19. Leskovec J, Krause A, Guestrin C, Faloutsos C, VanBriesen J, Glance N (2007) Cost-effective outbreak detection in networks. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, pp 420–429
20. Liow L, Cheng S, Lau H (2012) Niche-seeking in influence maximization with adversary. In: *Proceedings of the annual international conference on electronic commerce*. ACM, pp 107–112
21. Maghami M, Sukthankar G (2010) Identifying influential agents for advertising in multi-agent markets. In: *Proceedings of the international conference on autonomous agents and multiagent systems*, pp 687–694
22. Maghami M, Sukthankar G (2013) Hierarchical influence maximization for advertising in multi-agent markets. In: *Proceedings of the IEEE/ACM international conference on advances in social networks analysis and mining*. Niagara Falls, Canada, pp 21–27
23. Pathak N, Banerjee A, Srivastava J (2010) A generalized linear threshold model for multiple cascades. In: *International conference on data mining (ICDM)*, pp 965–970
24. Shakarian P, Paulo D (2012) Large social networks can be targeted for viral marketing with small seed sets. In: *Proceedings of the IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, pp 1–8
25. Wang C, Chen W, Wang Y (2012) Scalable influence maximization for independent cascade model in large-scale social networks. *Data Min Knowl Discov* 1–32
26. Yang W, Dia J, Cheng H, Lin H (2006) Mining social networks for targeted advertising. In: *Proceedings of the annual Hawaii international conference on system sciences*. IEEE Computer Society