بیشینه‌سازی تأثیر در شبکه‌های اجتماعی با مدل رقابتی مبتنی بر موضوع

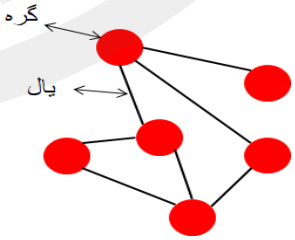
فاطمه طهماسبی[۱]، محمدعلی منوچهری[۲] محمدصادق هل فروش[۳] حبیب اله دانیالی[۴]
[۱]دانشگاه صنعتی شیراز، f.tahmasebi@sutech.ac.ir
[۲] دانشگاه صنعتی شیراز، m.manouchehri@sutech.ac.ir
[۳]دانشگاه صنعتی شیراز، ms_helfroush@sutech.ac.ir
[۴]دانشگاه صنعتی شیراز، danyali@sutech.ac.ir

شبکه‌های اجتماعی در دنیای امروز نقش اساسی در گسترش اطلاعات دارند. پژوهشگران بسیاری با بررسی ساختار شبکه‌های اجتماعی، سعی در شناخت و چگونگی کارکرد آن داشته‌اند. یکی از موضوعات مهم یافتن بهترین گره‌ها برای بیشتر کردن تأثیر در شبکه است. در این مسأله، انتشار اطلاعات درگراف شبکه‌های اجتماعی تحت یک مدل انتشار اتفاق می‌افتد. امروزه تبلیغات فراوانی از طریق شبکه‌های اجتماعی صورت می‌گیرد. توجه زیاد به بحث رقابت در شبکه‌های اجتماعی، منجر به شکل‌گیری مدل رقابتی شده‌است. همچنین، با توجه به تأثیرگذاری موضوعات مختلف روی کاربران، مدل مبتنی بر موضوع ارائه گردید. حال به منظور در نظرگرفتن تأثیر موضوعات مختلف در مدل رقابتی، برای اولین بار در این مقاله، مدل رقابتی مبتنی بر موضوع ارائه شده است. در روش پیشنهادی با استفاده از الگوریتم نمونه‌برداری، با انتخاب مناسب تعداد نمونه‌ها می‌توان، تأثیرگذارترین مجموعه از گره‌ها را بدست آورد. به لحاظ تئوری الگوریتم ارائه شده، دقتی معادل الگوریتم حریصانه دارد. اما پیچیدگی محاسباتی آن بسیار کمتر است. نتایج پیاده سازی بر روی داده‌های استاندارد، مؤثر بودن روش پیشنهادی را نشان می‌دهد.

کلید واژه-الگوریتم نمونه برداری،بیشینه‌سازی تأثیر،شبکه‌های اجتماعی،مدل رقابتی مبتنی بر موضوع

۱- مقدمه

شبکه‌های اجتماعی[۱] بر ابعاد مختلف زندگی افراد تاثیر می‌گذارند و باعث افزایش تعاملات اجتماعی می‌شوند. پژوهشگران بسیاری با بررسی ساختار شبکه‌های اجتماعی، سعی در شناخت و چگونگی کارکرد آن داشته‌اند. این شبکه‌ها به صورت یک گراف $G = (V, E)$ مدل سازی می‌شوند. گره‌های آن V به عنوان اعضای شبکه و یال‌های آن E ارتباطات بین آنها می‌باشند. در شکل ۱ نمونه‌ای از آن، نشان داده شده‌است.



شکل ۱: ساختار گراف اجتماعی

یکی از موضوعات اساسی در این شبکه‌ها انتشار اطلاعات[۲] و بیشینه سازی تأثیر[۳] روی کاربران است. فرایند انتشار برای مدل‌سازی پدیده‌های مختلفی مانند ویروس کامپیوتری، اتخاذ بازاریابی ویروسی[۴] برای محصول، شبکه‌های اجتماعی و غیره به کار می رود.

در مسئله بیشینه سازی تأثیر، زیرمجموعه‌ای از گره‌ها به نحوی انتخاب می‌شوند که بتوانند انتشار در شبکه را بیشینه کنند. بدست آوردن این گره‌ها یک مسأله NP_hard است. برای یافتن گره‌های موثر، یکی از روش‌های ارائه شده الگوریتم حریصانه[۵] است.

به منظور محاسبه تأثیر یگ گره در شبکه‌های اجتماعی از روش

[۱] Social network
[۲] Information Diffusion
[۳] Influence Maximization
[۴] Viral marketing
[۵] greedy

مونت کارلو[۶] استفاده می‌شود. به شرط وجود دو خاصیت سابمادولار[۷] و یکنواختی[۸] خروجی الگوریتم حریصانه، حداقل $1-\frac{1}{e}-\varepsilon$ برابر پاسخ بهینه است (۶۳ درصد پاسخ صحیح را به ما می‌دهد). در انتشار اطلاعات از طریق شبکه‌ها چندین مدل ارائه شده است. دیوید کمپ و همکارانش[۱]، مسأله بیشینه سازی تأثیر در شبکه‌های اجتماعی را به صورت ریاضی مدل کردند.

مدل‌ها چگونگی تأثیرپذیری افراد مختلف بر یکدیگر را نشان می‌دهند. از انواع آن‌ها، مدل‌های آستانه[۹] [۱]، [۲]، [۳] آبشاری[۱۰] [۱] معرفی شده‌اند.

آبشاری: در این مدل هر یال (v,u) دارای یک وزن $p(v,u)$ در بازه $[۰،۱]$ است که نشان دهنده‌ی احتمال فعال شدن گره u توسط v است. رأس v بعد از فعال شدن، یک شانس برای فعال کردن همسایه‌های غیر فعال خود u پیدا می‌کند. فعال شدن هر گره u، مستقل از تلاش‌های قبلی همسایگانش برای فعال نمودن وی می‌باشد. همچنین گره‌ای که قبلا فعال شده است دیگر فعالیتی روی آن انجام نمی‌شود. فرآیند فعال شدن، تا زمانی که هیچ گره دیگری نتواند فعال شود، ادامه می‌یابد.

- آستانه: در این مدل برای فعال شدن هر گره در نتیجه تأثیر پذیری از همسایه‌هایش، آستانه‌ای بین $[۰،۱]$ تعریف شده است. بعضی افراد مقدار آستانه کمتری دارند به این معنی که راحت‌تر تحت تأثیر قرار می‌گیرند. بعضی دیگر آستانه بالایی دارند و سخت‌تر می‌توان روی آن‌ها تأثیر گذاشت. برای هر لینک احتمالی تعریف می‌شود که فعال شدن هر گره بستگی به احتمال لینک‌هایی دارد که به آن متصل هستند. در صورتی که مجموع احتمالات لینک‌های متصل به گره مورد نظر بیشتر از حد آستانه مشخص شده باشد، آن گره فعال می‌شود.

در حال حاضر شرایط فعلی اقتصاد جهانی به گونه‌ای است که اغلب کشورها و شرکت‌ها به دنبال دست‌یابی به جایگاهی رقابتی در سطح بین المللی هستند.

حضور گستردهٔ شرکت‌های اقتصادی در شبکه‌های اجتماعی سبب ایجاد مدل رقابتی[۱۱] [۴]، [۵]، [۶]، [۷] شده‌است.

این مدل، سودمند برای شرکت‌هایی است که تمایل به حضور در بازار رقابتی را دارند.

این شبکه‌ها به کاربران خود سرویس‌هایی را ارائه می‌دهند که به کمک آن‌ها، اعضا قادر به اشتراک‌گذاری محتوای اجتماعی خودشان با دیگران می‌باشند. اطلاعات افراد در شبکه‌های اجتماعی به صورت متن، تصویر و ویدئو منتشرمی‌شوند. با توجه به علاقه‌مندی کاربران و تأثیراتی که در موضوعات مختلف بر یکدیگر می‌گذارند، محتوای اطلاعات منتشرشده به یکی از فاکتورهای مهم در شبکه تبدیل شده‌است. برهمین اساس مدل مبتنی بر موضوع[۱۲] ایجاد شد.

مدل مبتنی برموضوع [۸]، [۹]، [۱۰] تأثیر کاربران بر یکدیگر در مسئله‌های گوناگون را مورد بررسی قرار می‌دهد.

حال به منظور در نظرگرفتن تأثیر موضوعات مختلف در مدل رقابتی، برای اولین بار در این مقاله، مدل رقابتی مبتنی بر موضوع[۱۳] مطرح شده است.

۲- تعریف مسأله

امروزه رقابت کردن، لازمه ماندگاری فعالیت در بازارهای بین‌المللی و عرصه سیاسی است. از طرف دیگر موضوع و محتوای منتشر شده در شبکه‌های اجتماعی، در چگونه انتخاب کردن و چگونه اندیشیدن کاربران این رسانه‌ها حائز اهمیت است. برهمین اساس برای بهره‌مندی از دو خصوصیت رقابتی و تسلط بر موضوع منتشر شده در شبکه، برای اولین بار در این مقاله مدل رقابتی مبتنی بر موضوع ارائه شده است. حال به بررسی جزئیات بیشتری درحوزه بیشینه کردن تأثیر در شبکه‌های اجتماعی در مدل پیشنهادی پرداخته خواهد شد. روند انتشار نیز به صورت آبشاری مستقل[۱۴] است.

در مدل آبشاری مستقل مبتنی بر موضوع ارتباط کاربران با یکدیگر مستقل بوده و احتمال تأثیرگذاری روی یکدیگر بستگی به موضوع دارد. مدل IC و LT به نسخه‌های مبتنی بر موضوع گسترش یافته‌اند. فرض می شود آیتم i که توزیعی از موضوعات $z \in [1,k]$ است، در شبکه منتشر شود. تابع توزیع آیتم i با γ_i نشان داده می‌شود. براساس خصوصیات تابع توزیع داریم:

[۶] Monte Carlo
[۷] Submodular
[۸] Monotone
[9] Linear threshold model (LTM)
[10] Cascade model (CM)
[۱۱] Competitive Influence Maximization
[۱۲] Topic-aware Influence Maximization (TIM)
[۱۳] Topic-aware competitive Influence Maximization (TCIM)
[۱۴] Independent Cascade (IC)

$$\gamma_i^z = p(Z = z|i)\ with\ \sum_{z=1}^{k} \gamma_i^z = 1 \quad (۲)$$

i آیتم مورد نظر و $z \in [1, k]$موضوع منتشر شده‌است و k تعداد موضوعات است. p احتمال تأثیرکاربران بر یکدیگر با توجه به موضوع است و γ_i تابع توزیع احتمالات است.

احتمال $p_{u,v}^z$ تأثیر گره u بر گره v از طریق موضوع z است. احتمال روی یک آیتم که تمام موضوعات را شامل می‌شود، به صورت امید ریاضی نشان داده می‌شود.

$$p_{u,v}^i = \sum_{z=1}^{k} p_{u,v}^z \gamma_i^z \quad (۳)$$

با توجه به احتمال محاسبه شده براساس موضوعات، گره‌های تأثیرگذار موردنظر یافت می‌شوند.

با گسترش تولید انواع کالاهای با کیفیت و برندهای مختلف، شرکت‌ها به دنبال یافتن راهی برای تبلیغ محصول و کالای خود در بازارهای رقابتی هستند. از این رو مدل رقابتی شکل گرفت. این روش گسترش یافته مدل آبشاری مستقل است اما به جای که یک شرکت فعالیت داشته باشد چندین فعال اقتصادی، سیاسی و غیره با یکدیگر در یک محصول خاص و یا نظر و ایده خاصی در شبکه فعالیت می‌کنند. در این مسأله شرکت‌ها به طور جداگانه با داشتن اطلاعات از رقبای خود می‌توانند زیر مجموعه‌ای از کاربران را انتخاب کنند، که بیشترین تأثیر را در شبکه داشته باشند. در این مقاله مجموعه گره‌های $S_A \subseteq V$ به عنوان رقیب، در نظر گرفته می‌شوند. مجموعه S_B باید به گونه‌ای انتخاب شوند که تأثیر مورد انتظار را با توجه به مجموعه S_A به حداکثر برساند.

$$S_B^* = argmax_{S_B=\{S\subseteq V\setminus S_A, |S|=K\}} \sigma(S_B|S_A) \quad (۴)$$

تابع $\sigma(S_B|S_A)$، میزان تأثیر مجموعه B در برابر رقیب خود را نشان می‌دهد.

از آنجایی که محتوای منتشر شده توسط یک کاربر، علایق و تمایلات کاربر را نشان می‌دهد، این داده‌ها و معلومات را می‌توان در مدل رقابتی استفاده کرد. از این رو مؤثرترین و کامل‌ترین روش رقابتی، با توجه به تمایلات کاربران در شبکه در این مقاله مطرح شده‌است. کاربران موضوع و یا محتوایی که شامل اطلاعات مختلف است را در شبکه منتشر می‌کنند.

با توجه به انتشار موضوعات مختلف، و در نظر گرفتن تأثیرات آن‌ها در مدل رقابتی، مدل رقابتی مبتنی بر موضوع در شبکه شکل می‌گیرد. برای یافتن مؤثرترین زیرمجموعه از گره‌ها، روش نمونه برداری [۱۱] استفاده می‌شود.

۳- ارائه راه حل پیشنهادی

راه حل پیشنهادی با هدف پیدا کردن بهترین زیر مجموعه از گره‌ها در مدل رقابتی مبتنی بر موضوع، الگوریتم نمونه برداری است. با توجه به آیتم منتشر شده در شبکه و ارتباطی که کاربران در موضوعات مختلف با یکدیگر دارند، برای هر یال سه احتمال در سه موضوع مختلف در نظر گرفته می‌شود. با در نظر گرفتن معادله ۳ هر یال تبدیل به یک احتمال بین [۰،۱] می‌شود.

برای یافتن گره‌های تأثیرگذار در این مدل، ابتدا تعریف زیر ارائه داده می‌شود.

تعریف۱: مجموعه دستیابی معکوس[۱۵] (RR): در ابتدا به صورت تصادفی گره‌ای انتخاب می‌شود. یال‌های ورودی به آن در نظرگرفته می‌شوند. گره‌های متصل به آن در صورتی که رقیب نباشد، می‌توانند گره موردنظر را فعال کنند. در صورت فعال شدن، گره‌های متصل به مجموعه RR وارد می‌شوند. با استفاده از این مجموعه‌ها گره‌های تأثیرگذار بدست می‌آیند.

برای بدست آوردن مجموعه B که حداکثر تأثیر گذاری مورد انتظار را داشته باشد، تعداد نمونه برداری باید به طور مناسب بدست آید. برای تخمین این تعداد از قضیه چرنوف باند استفاده می‌شود. در این روش نیز به تقریب $1 - \frac{1}{e} - \varepsilon$ خواهیم رسید.

مراحل روش حل مسئله:

۱) با استفاده از الگوریتم STCIM[۱۶] [۱۱] از قضیه چرنوف باند، تعداد نمونه‌ها بدست می‌آیند.

قضیه چرنوف باند: فرض کنید که X، مجموع θ تا متغیر تصادفی iid در بازه [۰،۱] باشد، با میانگین μ سپس برای هر $\delta > 0$ داریم:

$$P_r[X > (1+\delta)\theta\mu] \le exp\left(-\frac{\delta^2}{2+\delta} \cdot \theta\mu\right) \quad (۵)$$

$$P_r[X < (1+\delta)\theta\mu] \le exp(-\frac{\epsilon^2}{2} \cdot \theta\mu) \quad (۶)$$

با توجه به مرجع [۱۱] و قضیه چرنوف باند تعداد نمونه برداری‌ها به صورت زیر بدست می‌آیند:

$$\theta \ge (8+2\epsilon)n.\frac{l\,ln\,n+ln\binom{n}{k}+ln\,2}{OPT.\epsilon^2} \quad (۷)$$

θ تعداد نمونه‌برداری، n تعدادگره‌های گراف، L یک در نظر گرفته می‌شود،ϵ عددی بین صفر و یک است که هر چقدر بزرگتر باشد دقت کمتری دارد.

مدل رقابتی [۱۱]، [۱۲] دو خصوصیت سابمادولار و یکنواختی

[۱۵] Reverse Reachable Set (RR)

[۱۶] Sampling Topic-aware competitive Influence Maximization (STCIM)

را دارد. در مدل مبتنی بر موضوع بعد از مشخص شدن وزن‌های یال‌ها با توجه به آیتم مورد نظر [۱۰]، به مدل رقابتی تبدیل می‌شود. برهمین اساس دو خصوصیت سابمادولار و یکنواختی را دارد و مدل ارائه شده نیز هر دو مشخصه را خواهد داشت. در مدل TCIM نیز می‌توان تعداد K گره با بیشترین تأثیر را از طریق الگوریتم حریصانه (الگوریتم ۱) بدست آورد و به تقریب $1-\frac{1}{e}-\varepsilon$ رسید.

اگر شرایط معادله ۷ برای θ برقرار باشد با احتمال $1-n^{-l}$ به تقریب $1-\frac{1}{e}-\varepsilon$ خواهیم رسید.

Algorithm 1 NodeSelection (G,$\boldsymbol{S_A}$,k,$\boldsymbol{\theta}$)

1: Initialize a set R=∅.
2: Generate $\boldsymbol{\theta}$ random RR sets and insert them into R.
3: Initialize a node set $\boldsymbol{S_K^*} = \emptyset$.
4: while $\boldsymbol{S_K^*} \leq \boldsymbol{K}$.
5: Let arg $\boldsymbol{max_u(\sigma(u) = f_R(\{u\}|S_A) for\ all\ u \in V \backslash S_A)}$.
6: Identify the node $\boldsymbol{u}$ that covers the most RR sets in R.
7: Add $\boldsymbol{u}$ into $\boldsymbol{S_K^*}$.
8: Remove from R all sets that are coverd by $\boldsymbol{u}$.
9: return $\boldsymbol{S_K^*}$.

۲) در معادله ۷ مقدار OPT ناشناخته است. باند پایینی برای آن تخمین زده می‌شود، تا مقدار قابل قبولی برای تعداد نمونه برداری بدست‌آید. الگوریتم یافتن باند پایین به صورت زیر است:

در الگوریتم ۲، $\alpha(R)$ به صورت زیر است:

$$\alpha(R) = 1 - (1 - \frac{W(R)}{m})^{k} \quad (۸)$$

V'_R: گره‌هایی مجموعه RR به جز مجموعه A.

W(R): تعداد لینک‌هایی که به گره‌های V'_R متصل هستند.

m′: لینک‌های کل گراف به جز لینک‌هایی که به گره‌های مجموعه A ختم می‌شوند.

پس از یافتن مجموعه‌های RR بهترین گره‌ها از طریق الگوریتم ۱ برای مجموعه B بدست می‌آید.

گره‌های مجموعه B با ۳ روش حریصانه (الگوریتم۱)، max_degree که نشان دهنده گره‌هایی هستند که بیشترین یال‌های خروجی را دارند و الگوریتم تصادفی (random)، نیزبدست می‌آیند و با یکدیگر مقایسه می‌شوند.

Algorithm 2 OPTEstimation (G, k)

1: for $\boldsymbol{i = 1\ to\ log_2 n - 1}$ do
2: Let$\boldsymbol{N_i = (6l\ log\ n + 6\ log\ (log_2 n - 1)).2^i}$.
3: Let $\boldsymbol{ave} = 0$.
4: for j = 1 to $\boldsymbol{N_i}$ do
5: Generate a random RR set R.
6: $\boldsymbol{\alpha(R) = 1 - (1 - \frac{W(R)}{m})^{k}}$
7: $\boldsymbol{ave = ave + \alpha(R)}$.
8 if $\boldsymbol{ave/N_i > \frac{1}{2^i}\ then}$
9: return $\boldsymbol{OPT = n.ave/2 \cdot N_i}$
10: return OPT =1.

۴- نتایج تجربی

مشخصات آماری داده‌های مورد مطالعه در جدول ۱ آمده است.

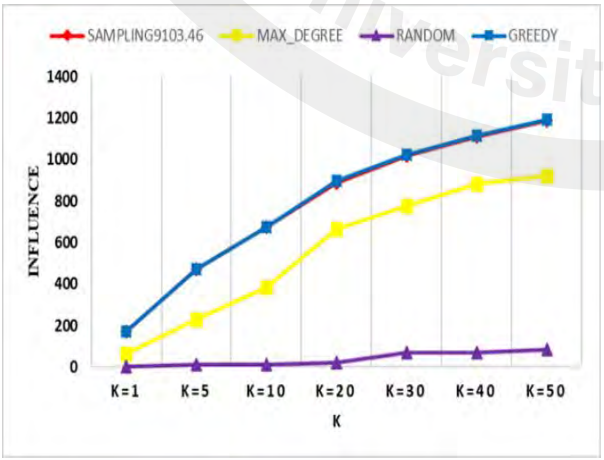
در این بخش مقایسه کارایی الگوریتم‌های مطرح شده با یکدیگر از نظر میزان تأثیر و زمان اجرای الگوریتم‌ها مورد بررسی قرار می‌گیرند. نتایج از اعمال ۴ الگوریتم متفاوت بر روی داده‌ها بدست‌آمده‌اند. برای همه آن‌ها مدل رقابتی مبتنی بر موضوع در نظر گرفته می‌شود.

در شکل ۲ و ۳ میزان تأثیر مجموعه بدست آمده B، با الگوریتم‌های مختلف در حضور رقیب با Kهای متفاوت مقایسه می‌شوند. مجموعه رقیب A، برای تمام نتایج‌ها ۵۰ گره یکسان در نظر گرفته می‌شود.

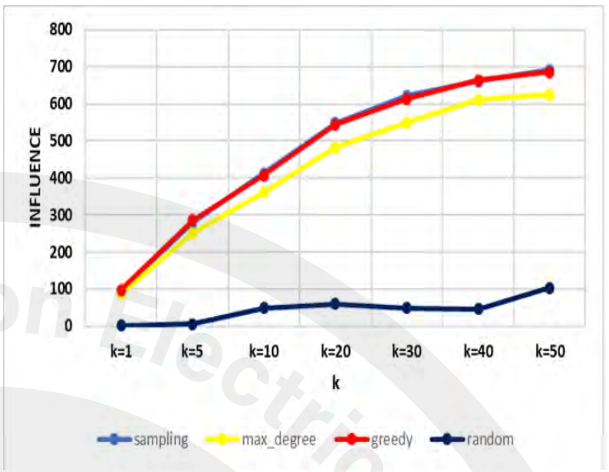
تأثیر داده ۷۵۸۷۹ گره‌ای همانطور که در شکل ۳ نشان داده شده است با ۳ الگوریتم مقایسه می‌شود.

جدول۱: مشخصات اماری مجموعه داده‌های مورد استفاده

نام داده‌ها	تعدادگره‌ها	تعداد یال‌ها
Epinions	۷۵۸۷۹	۵۰۸۸۳۷
Email_Eu_core	۱۰۰۵	۲۵۵۷۱
ca-CSphd	۱۸۸۲	۱۷۴۰



شکل۲: تأثیر داده ca-CSphd



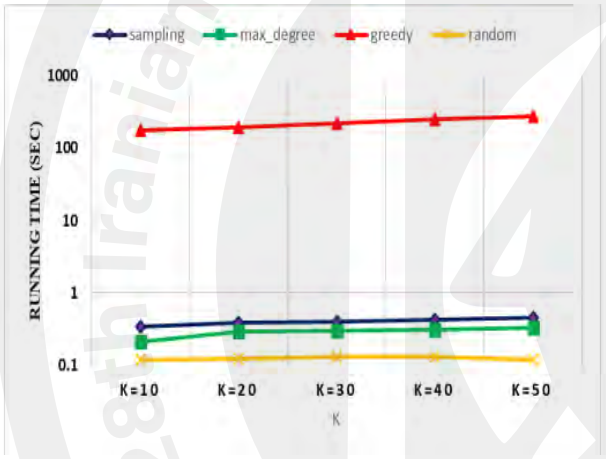
شکل۳: تأثیر داده Email_Eu_Core

شکل۴: تأثیر داده Epinions

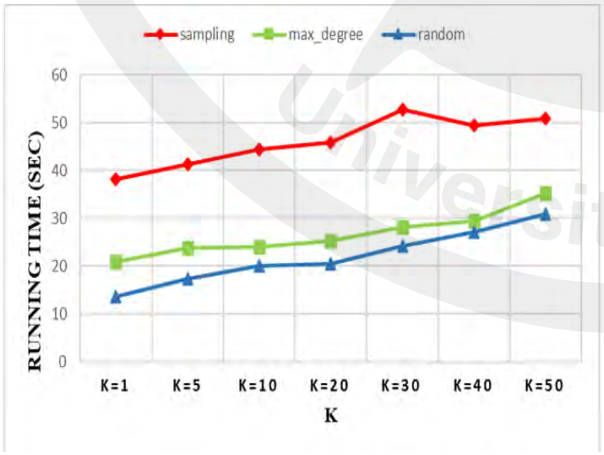
با مقایسه عملکرد روش پیشنهادی (نمونه‌برداری) و الگوریتم حریصانه، نتایج بیانگر برتری هر دوی این روش‌ها از نظر بیشینه سازی تأثیر نسبت به سایر روش‌ها می‌باشند.

زمان اجرای الگوریتم‌ها متفاوت است. مقایسه زمانی اجرای الگوریتم‌ها از این نظر حائز اهمیت است که مشخص شود، کدام روش در زمان بهینه‌تری پاسخ بهتری دارد. در شکل ۴ و ۵ زمان اجرای سه الگوریتم با یکدیگر برای هر کدام از داده‌ها بررسی می‌شوند. برای تمام داده‌ها در محاسبه معادله ۷ مقدار $\epsilon = 0.1$ است. در نمودارهای زیر محور افقی تعدادگره‌های گراف و محور عمودی زمان اجرای الگوریتم را بر حسب ثانیه مشخص می‌کند.

شکل۵: زمان اجرای الگوریتم‌ها Email_Eu_Core



شکل۶: زمان اجرای الگوریتم‌ها داده ca-CSphd



شکل۸: زمان اجرای داده Epinions با الگوریتم‌های متفاوت

از جنبه زمانی اجرای برنامه، روش حریصانه مدت زمان زیادی برای یافتن تأثیرگذارترین گره‌ها نیاز دارد. به لحاظ تئوری الگوریتم

ارائه شده، دقتی معادل الگوریتم حریصانه دارد. اما روش پیشنهادی(نمونه برداری) زمان اجرای قابل قبول و پیچیدگی محاسباتی کمتری در مقایسه با دیگر روش‌ها دارد.

۵- نتیجه گیری

شبکه‌های اجتماعی به عنوان عامل ایجاد تعامل میان انسان‌ها درفضای مجازی اهمیت خاصی دارند. یکی از موضوعات مورد توجه در این حوزه انتخاب کاربران با نفوذ، برای تأثیرگذاری حداکثر در این شبکه‌ها است. شرکت‌ها و سازمان‌های مختلف به دنبال یافتن بهترین مجموعه برای تأثیرگذاری موضوع موردنظر خود بوده‌اند. بر اساس اهمیت اطلاعات و رقابتی شدن فضای کسب و کار مدل رقابتی و مبتنی بر موضوع شکل گرفت. در مدل پیشنهادی این دو فاکتور با یکدیگر ادغام شده‌اند و تأثیر موضوع در مدل رقابتی مورد ارزیابی قرار گرفته است. در این مقاله با استفاده از الگوریتم نمونه‌برداری، تأثیرگذارترین مجموعه از گره‌ها بدست می‌آید. از نتایج مشخص است، عملکرد الگوریتم نمونه برداری و حریصانه در بیشینه سازی تأثیر نسبت به دیگر روش‌ها بهتر است. اما در روش پیشنهادی پیچیدگی محاسباتی آن بسیار کمتر است.

۶- مراجع

[1] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining,*.

[2] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010: ACM, pp. 1029-1038.

[3] Y. Tang, X. Xiao, and Y. Shi, "Influence maximization: Near-optimal time complexity meets practical efficiency," in *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, 2014: ACM, pp. 75-86.

[4] S. Kochemazov, "Comparative Study of Combinatorial Algorithms for Solving the Influence Maximization Problem in Networks under a Deterministic Linear Threshold Model," *Procedia Computer Science,* vol. 136, pp. 190-199, 2018.

[5] S. S. Singh, K. Singh, A. Kumar, H. K. Shakya, and B. Biswas, "A survey on information diffusion models in social networks," in *International Conference on Advanced Informatics for Computing Research*, 2018: Springer, pp. 426-439.

[6] C. Budak, D. Agrawal, and A. El Abbadi, "Limiting the spread of misinformation in social networks," in *Proceedings of the 20th international conference on World wide web*, 2011: ACM, pp. 665-674.

[7] M. A. M. A. Kermani, S. F. F. Ardestani, A. Aliahmadi, and F. Barzinpour, "A novel game theoretic approach for modeling competitive information diffusion in social networks with heterogeneous nodes," *Physica A: statistical mechanics and its applications,* vol. 466, pp. 570-582, 2017.

[8] W. Chen, T. Lin, and C. Yang, "Real-time topic-aware influence maximization using preprocessing," *Computational social networks,* vol. 3, no. 1, p. 8, 2016.

[9] N. Barbieri, F. Bonchi, and G. Manco, "Topic-aware social influence propagation models," *Knowledge and information systems,* vol. 37, no. 3, pp. 555-584, 2013.

[10] S. Chen, J. Fan, G. Li, J. Feng, K.-l. Tan, and J. Tang, "Online topic-aware influence maximization," *Proceedings of the VLDB Endowment,* vol. 8, no. 6, pp. 666-677, 2015.

[11] Y. Lin and J. C. Lui, "Analyzing competitive influence maximization problems with partial information: An approximation algorithmic framework," Performance Evaluation, vol. 91, pp. 187–204, 2015.

[12] H. Li, L. Pan, and P. Wu, "Dominated competitive influence maximization with time-critical and time-delayed diffusion in social networks," *Journal of computational science,* vol. 28, pp. 318-327, 2018.