

## TOSI: یک روش ارزیابی اثر شبکه اجتماعی در شبکه های اجتماعی زمینه ای

چکیده:

شبکه های اجتماعی آنلاین (OSN) به عنوان ابزاری برای طیف وسیعی از کاربرد ها استفاده شده است. برای مثال، بستر شبکه اجتماعی در سیستم اشتغال، تجارت الکترونیک و سیستم CRM برای بهبود کیفیت توصیه ها با کمک شبکه های اجتماعی استفاده شده است. با این حال، روش های ارزیابی اثر اجتماعی موجود به طور کامل شرایط اجتماعی را در نظر می گیرد یعنی روابط اجتماعی و اعتماد اجتماعی بین شرکت کننده ها و سلايق شرکت کننده ها را در نظر نمی گیرد که همگی اثر معنی داری بر روی ارزیابی اثر در OSN دارد. از اینرو، این روش های موجود قادر به ارایه نتایج ارزیابی اثر اجتماعی صحیح نمی باشند. در این مقاله، یک روش ارزیابی اثر اجتماعی اعتماد محور موسوم به TOSI با در نظر گرفتن شرایط اجتماعی پیشنهاد می شود. آزمایشاتی بر روی دو مجموعه شبکه اجتماعی انجام شد یعنی اپینیون و dblp. نتایج آزمایشی نشان می دهد که روش TOSI عملکرد بهتری بر روی روش پیشرفته SOCAP از حیث اثر بخشی، کارایی و استواری دارد.

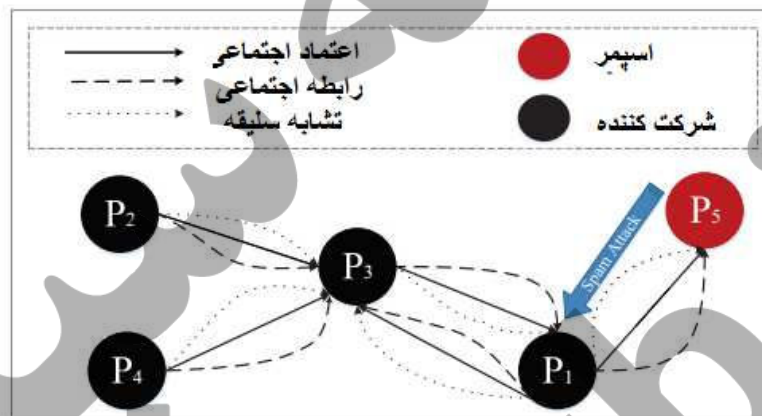
کلمات کلیدی: شبکه اجتماعی، اثر اجتماعی و اعتماد

### 1- مقدمه

#### 1-1 پیش زمینه

شبکه های اجتماعی آنلاین (OSN) روز به روز محبوب تر شده و به عنوان ابزاری در شرایط مختلف نظیر اشتغال، CRM و تجارت الکترونیکی استفاده شده است. در این شرایط، اثر اجتماعی شرگت کننده ها بر تصمیم گیری سایرین اثر دارد. برای مثال، اپینیون که یک بستر تجارت الکترونیکی مبتنی بر OSN می باشد، خریدار قادر به نوشتن نظر خود در مورد محصول برای طبقه بندی محصولات و فروشنده متناظر است. این بازنگری و امتیاز دهی را می توان توسط سایر خریداران در نظر گرفت که همگی بر تصمیم گیری در خرید روش های یکسان اثر دارد. همان طور که در مطالعات توسط سایر خریداران نشان داده شده است و می تواند بر روی تصمیم گیری در خرید محصولات یکسان اثر داشته باشد. همان طور که در مطالعات روان شناسی اجتماعی و علوم کامپیوتری (1-2-3) نشان داده شده است، یک فرد توصیه های شرکت کننده های با اثر اجتماعی بالاتر را در حوزه خاص می پذیرد. از اینرو، ارزیابی صحیح اثر اجتماعی شرکت کننده ها و شناسایی عوامل موثر در شبکه های اجتماعی مهم است.

در منابع، بسیاری از روش های ارزیابی اثر اجتماعی پیشنهاد شده است [5, 6, 7, 8, 9, 10, 11, 12, 13, 14] که در آن مدل آبخاری مستقل (IC) یک مدل برای یافتن گروه های K برتر می باشد و دارای اثر اجتماعی بیشینه در یک شبکه است. متعاقباً، برخی از کار های مهم (9-11)، برای بهبود مقیاس پذیری IC پیشنهاد شده است. به علاوه در سال های اخیر، روش بیشینه سازی اثر محلی (15) برای ارزیابی اثر اجتماعی شرکت کننده خاص در OSN پیشنهاد شده است. به علاوه چون برخی OSN در حال تبدیل شدن به یک ژنراتور زمان واقعی بزرگ می باشد، برخی از روش های استریمینگ برای ارزیابی اثر اجتماعی شرکت کننده ها در OSN پیشنهاد شده است.



شکل 1: شبکه اجتماعی اپینیون

## 1-2 مسئله و انگیزه

همان طور که در روان شناسی اجتماعی نشان داده شده است (18-19-20)، اعتماد اجتماعی بین شرکت کننده ها (برای مثال، دانشجویان به استاد خود در یک زمینه خاص اعتماد می کنند)، رابطه اجتماعی بین شرکت کنندگان (به عنوان مثال، رابطه بین پدر و پسر)، و شباهت اولویت بین شرکت کنندگان (به عنوان مثال، همه آنها علاقه مند به بازی بسکتبال هستند) اثر معنی داری بر روی تصمیم گیری شرکت کننده ها و اثرات و نفوذ اجتماعی دارند. با این حال، این شرایط اجتماعی توسط روش های ارزیابی اثر اجتماعی عمومی در نظر گرفته نشده اند. از این روی این روش ها نمی توانند منجر به نتایج ارزیابی اثر اجتماعی شوند. به علاوه با رشد مقیاس و پیچیدگی شبکه،

شبکه های اجتماعی به انواع مختلف حملات ناخواسته سازگار هستند. یک راهبرد رتبه بندی اعتماد محور برای دفاع در برابر حمله پیشنهاد می شود.

مثال 1: شکل 1 شبکه اجتماعی اپینیتون را نشان می دهد که شامل 5 شرکت کننده از P1 تا P5 می باشد که همگی خریدار هستند. رابطه اعتماد (خطوط پر رنگ) P1 و P2، P3 و P4 و P3 بر اساس کیفیت ارزیابی محصول P3 است. رابطه اجتماعی و سلايق را می توان تز تاریخچه خرید و پروفیل آن ها کاوش کرد (22). فرض کنید که P1 دارای روابط اجتماعی نزدیک باشد و دارای سلايق مشابه با P3 نسبت به P2 است و P3 می تواند بر رفتار خرید P1 نسبت به P2 اثر داشته باشد. با این حال، در روش های ارزیابی اثر اجتماعی (5-10)، احتمال اثر بین دو گره در مدل LT تصادفی است و منعکس کننده اثر واقعی شرکت کننده ها است. به علاوه اگر P5 یک اسپر باشد، می تواند از طیف وسیعی از همسایه ها برای اثر گذاری اجتماعی قوی بر روی تصمیم گیری P1 داشته باشد

مسائل فوق امکان توسعه یک روش ارزیابی اثر را برای ارزیابی اثر اجتماعی شرکت کننده ها در OSN می دهد. در این صورت، با در نظر گرفتن شرایط اجتماعی مهم، روش ارزیابی اثر اجتماعی اعتماد محور موسوم به TOSI با استفاده از روش تکراری پیشنهاد می شود. چون روش ما همگرا است، می توان به نتایج ارزیابی اجتماعی صحیح دست یافت

### 3-1 اهمیت و هدف

اهمیت این مقاله به شرح زیر است

- بر اساس آخرین گزارشات این اولین کاری است که در ارزیابی اثر اجتماعی در نظر گرفته می شود
- برای دفاع در برابر حمله اسپم، اسپم فارم (23)، در فرایند ارزیابی اثر اجتماعی، ما یک روش ارزیابی اثر اجتماعی اعتماد محور را پیشنهاد می کنیم که از طیف اسپم برای اندازه گیری احتمال مهاجم بودن شرکت کننده استفاده می کند
- یک روش ارزیابی اثر اجتماعی جدید TOSI پیشنهاد می شود که N تعداد گره ها در OSN و  $l$  دفعات تکراری در محاسبه است

- آزمایشاتی بر روی دیتاست های شبکه اجتماعی واقعی انجام شد. اپینیون و DBPL. با مقایسه با روش

ارزیابی اثرات اجتماعی پیشرفته، SOCAP، روش TOSI عملکرد بهتری از SOCAP داشته است

این مقاله به صورت زیر ارزیابی شده است. بخش دوم مربوط به مرور منابع است. بخش سوم به معرفی موارد

مقدماتی، بخش چهارم شامل پیشنهاد روش TOSI است. در بخش 5، اثر بخشی و کارایی روش پیشنهادی با

مقایسه روش پیشرفته SOCAP تعیین می شود. بخش 6 شامل نتیجه گیری است

## 2- مرور منابع

در منابع، رویکرد های ارزیابی اثر اجتماعی به چهار گروه زیر طبقه بندی شده اند: بیشینه سازی اثر جهانی،

بیشینه سازی اثر محلی، یادگیری استریم جریان و ارزیابی اثر فردی. این روش ها در ادامه بررسی شده اند

## 1-2 بیشینه سازی اثر جهانی

بیشینه سازی اثر جهانی به معنی یافتن گروهی از گره هایی است که بر تعداد بیشینه سایر گره ها در OSN اثر

دارد. کمپ و همکاران (5) یک الگوریتم حریص را پیشنهاد کرده اند که  $(1 - 1/e)$  نسبت بیشینه است. با این حال

این الگوریتم دارای کارایی عملی پایین بوده و با اندازه شبکه غیر مقیاس پذیر است. برای بهبود مقیاس پذیری، 9

یک الگوریتمی را پیشنهاد می کند که دارای پارامتر ساده است و گاربران قادر به کنترل توازن بین زمان اجرا و

توزیع اثر الگوریتم است. جانگ و همکاران (7) یک الگوریتم IRIE را ارائه کرده اند که مزیت های رتبه بندی اثر و

روش های برآورد اثر را برای بیشینه سازی اثر کلی ترکیب می کند. 10 یک الگوریتم بیشینه سازی اثر مقیاس

پذیر موسوم به الگوریتم مسیر مستقل برای مدل IC ارائه کرده است. در این مدل، آن ها به بررسی IPA با در

نظر گرفتن مسیر اثر مستقل به عنوان یک واحد ارزیابی پرداخته اند. به علاوه برای توسعه الگوریتم ارزیابی (11)

یک الگوریتم CELF را ارائه کرده و از مدولاریته زیرین برای یافتن مجموعه های موثر تر استفاده می کند. به علاوه،

لاسکوگ و همکاران (24-25) روابط بین دو شرکت کننده را به دو مورد اثر مثبت و منفی تقسیم کرده و آن ها را

از طریق تئوری های تعادل و وضعیت از روان شناسی اجتماعی بررسی می کند. بر اساس این مطالعه، یاهونا و

همکاران (26) به بررسی اثر انتشار اثر و بیشینه سازی اثر در OSN با دو رابطه منفی و مثبت پرداختند

## 2-2 بیشینه سازی اثر محلی

بیشینه سازی اثر اجتماعی، به معنی یافتن گروه هایی از گره ها با اثر بیشینه بر روی شرکت کننده خاص است. یان و همکاران (13) به مطالعه رابطه بین رتبه اعتماد و محصول در سایت ارزیابی مصرف کننده پرداخته اند. به علاوه آن ها روشی را برای ارزیابی نقاط قوت رابطه اعتماد برای برآورد اثر اعتماد در میان شرکت کننده ها ارایه کرده اند. به علاوه، گایو و همکاران (15) مدلی را برای یافتن K مدل ارایه کرده اند. به علاوه، ایتاوا و همکاران (27) مدل احتمالی را برای کشف اثر پنهان بین شرکت کننده در OSN ارایه کردند. این مدل برای یافتن شرکت کنند های موثر و کشف رابطه بین شرکت کننده ها استفاده می شود

### 3-2 یادگیری استریم اثر

در سال های اخیر، OSN در حال تبدیل شدن به یک تولید کننده زمان واقعی دیتا استریم نظیر تویتر است. روش های استریمینگ اثر اجتماعی محبوب تر است. کاتسکوفو همکاران یک روش استریمینگ را ارایه کرده اند که برای محاسبه مقاومت اثر در لینگ OSN استفاده شده است. به علاوه کارتیک (17) رویکردی را برای کاوش الگو های جریان بر اساس محدودیت های اعتبار جریان ارایه کرده است. با محدود سازی بستر های میکرو بلاگینگ، OSN های دیگر می توانند زمینه کافی را برای اجرای کشف الگوی جریان اطلاعات ارایه کنند. از این روی روش های استریمینگ را نمی توان برای یافتن عامل اثر گذار در بستر های تجارت الکترونیکی OSN استفاده کرد

### 4-2 مسئله ارزیابی اثر انفرادی

برای ارزیابی اثر اجتماعی یک شرکت کننده خاص، اسابین و همکاران (12) یک رویکرد موسوم به SOCAP را برای یافتن عوامل مثر در OSN با استفاده از مقادیر سرمایه اجتماعی ارایه کرده اند. مسئله یافتن عامل اثر گذار به عنوان مسئله تخصیص ارزش مدل سازی شد که در آن مقدار تخصیص داده شده بیانگر سرمایه اجتماعی است. به علاوه، فرانگس و همکاران (14)، روشی را برای شناسایی عوامل موثر در سیستم های چند عاملی با استفاده از روش ماتریس فاکتور گیری اندازه گیری گره ها در شبکه ارایه کرده است

### جمع بندی

روش های موجود به طور کامل شرایط اجتماعی نظیر روابط و اعتماد بین شرکت کننده ها را در نظر نمی گیرد. همان طور که در روان شناسی اجتماعی 1-2-3 و علوم گامپیوتر نشان داده شده است (4)، این زمینه های اجتماعی برای ارزیابی اثر اجتماعی مفید هستند. از این روی این روش ها به نتایج صحیح منتهی نمی شوند

## 3- موارد مقدکاتی

هدف روش ما ارزیابی اثر اجتماعی منطقی هر شرکت کننده در شبکه های اجتماعی زمینه ای بر اساس ساختار شبکه اجتماعی و زمینه های اجتماعی است

## 1-3 شبکه اجتماعی زمینه ای

یک شبکه اجتماعی زمینه ای (CSN) (28) یک گراف جهت دار دارای برچسب  $G = (V, E, LV, LE)$  است که

-  $V$  مجموعه رئوس است

-  $E$  مجموعه ای از رئوس است و  $(v_i, v_j) \in E$  نشان دهنده یک راس جهت دار از  $v_i$  تا  $v_j$  است

$LV$  تابع تعریف شده بر روی  $V$  است به طوری که هر راس  $v$  در  $V$ ،  $LV(v)$  مجموعه ای از برچسب های

$v$  است. برچسب های رئوس برای نقش های اجتماعی یا اثر اجتماعی در حوزه خاص است

-  $LE$  تابع تعریف شده  $E$  است که برای هر  $(v_i, v_j) \in E$ ،  $LE(v_i, v_j)$  مجموعه ای از برچسب ها برای

$(v_i, v_j)$  نظیر روابط اجتماعی، اعتماد اجتماعی و سلاقی در دامنه خاص است

## 2-3 زمینه های اجتماعی

فرض کنید که  $P$  مجموعه ای از شرکت کننده ها و  $R$  بردار زمینه های اجتماعی

که در آن  $\vec{R}_{i,j}(p)$  و  $\vec{R}_{i,j}(t)$ ،  $\vec{R}_{i,j}(s)$  بیانگر اعتماد، رابطه

وسلیقه بین  $P_i$  و  $P_j$  است و علاوه،  $IN_i$  نشان دهنده همسایه های ورودی  $P_i$  و  $ON_i$  نشان دهنده

همسایه های خروجی است

- اعتماد اجتماعی: این بر اساس اثر متقابل افراد است (28). فرض کنید که  $t$  نشان دهنده اعتماد بین

شرکت کننده ها است.  $\vec{R}_{i,j}(t) = 1$  نشان می دهد که  $P_i$  و  $P_j$  و  $\vec{R}_{i,j}(t) = 0$  نشان می دهد که  $P_i$  به

طور کامل به  $P_j$  بی اعتماد است. مقدار  $ST$  بر اساس تجربه اثر متقابل است

- رابطه اجتماعی: این به معنی درجه اجتماعی بین دو شرکت کننده (19) است. فرض کنید که  $S$  نشان

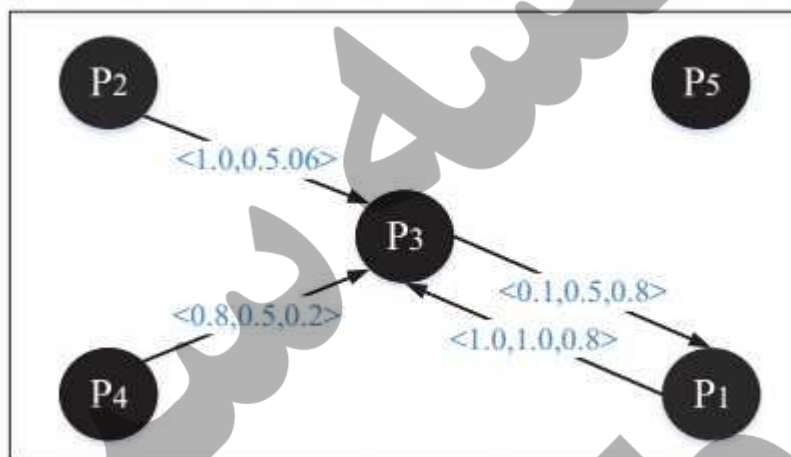
دهنده نزدیکی بین روابط اجتماعی است و از این روی  $\vec{R}_{i,j}(s) = 0$  نشان می دهد که  $P_i$  با  $P_j$  در

تماس نبوده است. نزدیکی رابطه بین دو شرکت کننده بر اساس اطلاعات شخصی آن ها یا تعامل آنلاین اندازه گیری می شود

- تشابه سلیقه: تشابه سلیقه منعکس کننده رابطه بین سلايق دو شرکت کننده است. فرض کنید  $p$ ، مقدار

تشابه سلیقه بین دو شرکت کننده باشد.  $\vec{R}_{i,j}(p) = 1$  نشان می دهد که سلايق  $p_i$  و  $p_j$  دقیقا یکسان

هستند و  $\vec{R}_{i,j}(p) = 0$  نشان می دهد که چیز مشترک بین  $p_i$  و  $p_j$  وجود ندارد



شکل 2: یک شبکه اجتماعی زمینه ای

ایجاد روابط اجتماعی، اعتماد و تشابه سلیقه در همه ابعاد سخت است و از این روی آن ها در جوامع اجتماعی

با استفاده از روش های داده کاوی امکان پذیر بوده است. کاوش مقادیر زمینه ای چالش بر انگیز است

مثال 2: شکل 2 شبکه اجتماعی زمینه ای را نشان می دهد که شامل زمینه های

$$\vec{R}_{2,3} = \langle 1.0, 0.5, 0.6 \rangle,$$

$$\vec{R}_{4,3} = \langle 0.8, 0.5, 0.2 \rangle, \vec{R}_{1,3} = \langle 1.0, 1.0, 0.8 \rangle \text{ and } \vec{R}_{3,1} = \langle 0.1, 0.5, 0.8 \rangle.$$

است.

### 3-3 مسئله ارزیابی اثر اجتماعی

با توجه به شبکه اجتماعی زمینه ای  $G = (V, E, LV, LE)$  و مجموعه ای از عوامل،  $r$ ، اثر اجتماعی شرکت

کننده و مجموعه اثرات SI ارزیابی می شود

4- روش ارزیابی اثر مبتنی بر اعتماد

در این بخش، روش ارزیابی اثر اجتماعی مبتنی بر اعتماد TOSI با روش تکراری ارزیابی اثر اجتماعی پیشنهاد می شود. چون روش ما همگرا است، می توان به نتایج ارزیابی اجتماعی صحیح دست یافت

#### 1-4 توصیف الگوریتم

در روش TOSI، اثر اجتماعی شرکت کننده ها محاسبه شده و تا زمان رسیدن به همگرایی با روش تکراری جایگزین می شود. سپس، فرایند تکرار معرفی می شود

اثرات اجتماعی در زمان T+1 بر اساس عوامل اثر گذاری است که در زمان تکرار T قرار دارد. در فرایند ارزیابی

اثرات اجتماعی، زمینه های اجتماعی مطلوب است.  $SI_i^*$  بیانگر اثر اجتماعی شرکت کننده PI می باشد

$$SI_i^{t+1} = \sum_{P_k \in IN_i} SI_k^t \cdot \rho_{k,i} \quad (1)$$

که  $\sum_{P_i \in ON_k} \rho_{k,i} = 1$  و

$$\rho_{k,i} = \frac{\vec{R}_{k,i}(t)}{3 \cdot TTTR_k} + \frac{\vec{R}_{k,i}(s)}{3 \cdot TTSR_k} + \frac{\vec{R}_{k,i}(p)}{3 \cdot TTPS_k} \quad (2)$$

$$TTTR_k = \sum_{P_j \in ON_k} \vec{R}_{k,j}(t), \quad TTSR_k = \sum_{P_j \in ON_k} \vec{R}_{k,j}(s) \text{ and}$$

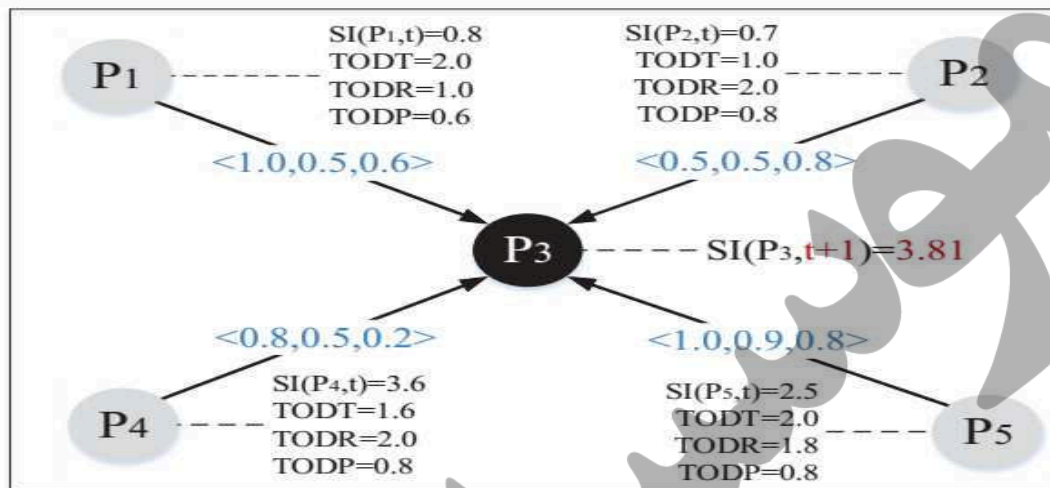
$$TTPS_k = \sum_{P_j \in ON_k} \vec{R}_{k,j}(p). \text{ میباید.}$$

از این روی،  $\rho_{k,i}$  نشان دهنده احتمال اثر کل از PK تا pi است

مثال 3: در شکل 3، برای شرکت کننده p1، زمینه های اجتماعی، p2=p5 وجود دارد. جزییات شرایط اجتماعی

و اثر اجتماعی در زمان t در شکل 3 نشان داده شده است





شکل 3: محاسبه اثر اجتماعی در فرایند تکراری

اثر اجتماعی  $SI_i(t+1) = 0.8 \times (1.0/2.0 + 0.5/1.0 + 0.6/0.6)/3 + 0.7 \times (0.5/1.0 + 0.5/2.0 + 0.8/0.8)/3 + 3.6 \times (0.8/1.6 + 0.5/2.0 + 0.2/0.8)/3 + 2.5 \times (1.0/2.0 + 0.9/1.8 + 0.8/0.8)/3 = 3.81$  است.

#### 2-4 همگرایی تکرار

برای روش تکراری، یک کار اساسی، تضمین همگرایی است. هاولیا 20 اقدام به تحلیل همگرایی PageRank

بر اساس ایده های روش تکراری کرده است و ما از همگرایی روش TOSI بهره می بریم. اثر اجتماعی واقعی

شرکت کننده PI است

$$Error(t) = \sum_{i=1}^N |SI_i^t - SI_i^*| \quad (3)$$

قضیه 1: TOSI همگرا است زیرا،  $Error(t) < Error(t - 1)$ .

اثبات 1: چون  $SI_i^*$  یک راه حل واقعی است. بر طبق معادله 1، حل معادله زیر مطلوب است

$$SI_i^* = \sum_{P_k \in IN_i} SI_k^* \cdot \rho_{k,i} \quad (4)$$

برای شرکت کننده PI، خطای تکرار به صورت زیر است

$$SI_i^t - SI_i^* = \sum_{P_k \in IN_i} (SI_k^{t-1} - SI_k^*) \cdot \rho_{k,i} \quad (5)$$

با نامساوی مثلث، می توان معادله زیر را نوشت

$$|SI_i^t - SI_i^*| \leq \sum_{P_k \in IN_i} |SI_k^{t-1} - SI_k^*| \cdot \rho_{k,i} \quad (6)$$

سپس، جمع کل خطا های شرکت کننده ها برای بدست آوردن خطای کل محاسبه می شود. توجه کنید که

$$\sum_{P_i \in ON_k} \rho_{k,i} = 1: \text{ است.}$$

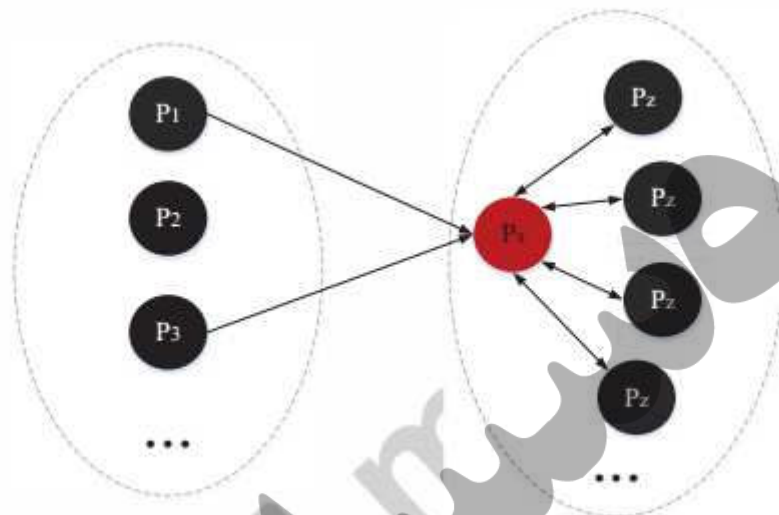
$$\begin{aligned} Error(t) &= \sum_{i=1}^N |SI_i^t - SI_i^*| \\ &\leq \sum_{i=1}^N \sum_{P_k \in IN_i} |SI_k^{t-1} - SI_k^*| \cdot \rho_{k,i} \\ &= \sum_{\vec{R}_{k,i} \in R} |SI_k^{t-1} - SI_k^*| \cdot \rho_{k,i} \\ &= \sum_{k=1}^N |SI_k^{t-1} - SI(P_k)^*| \cdot \sum_{P_i \in ON_k} \rho_{k,i} \\ &= Error(t-1) \end{aligned} \quad (7)$$

با توجه به معادله 6 میتوان دریافت که  $Error(t) = Error(t-1)$  اگر تنها اگر

باشد. با این حال روش تکراری موجب  $\forall P_k \in P, SI_k^{t-1} - SI_k^* > 0 \text{ or } \forall P_k \in P, SI_k^{t-1} - SI_k^* < 0$

کاهش اثر اجتماعی به 1 می شود یعنی  $\sum_{k=1}^N SI_k^{t-1} = \sum_{k=1}^N SI_k^* = 1$  است. از این روی

است.  $Error(t) < Error(t-1)$ . لذا قضیه 1 اثبات می شود.



شکل 4: حمله اسپم

## 3-4 حمله اسپم

- در این روش ما در برابر جمله اسپم دفاع می کنیم یعنی فارم اسپم برای دفاع در برابر حمله اسپم، اسپم فارم (23)، در فرایند ارزیابی اثر اجتماعی، ما یک رو ارزیابی اثر اجتماعی اعتماد محور را پیشنهاد می کنیم که از طیف اسپم برای اندازه گیری احتمال مهاجم بودن شرکت کننده استفاده می کند برای مثال در فیسبوک طرفدار زامبی معمولاً نقش شرکت کننده را ایفای کند. این ساختار شبکه نظیر شکل 4، موجب می شود تا اسپم دارای اثر اجتماعی غیر واقعی باشد بر اساس تراست رنک (30)، از راهبرد رتبه بندی اعتماد محور برای دفاع استفاده می کنیم و  $TR_i$  برای نشان دادن اعتماد استفاده می کنیم در ابتدا، تراست رنک  $TR^{t+1}$  در زمان  $t+1$  بر اساس رتبه های اعتماد  $TR^t$  است. هم چنین، فرایند اثر اجتماعی، زمینه های اجتماعی در نظر گرفته می شود. معادله به صورت زیر است

$$TR_i^{t+1} = \sum_{P_k \in IN_i} TR_k^t \cdot \rho_{ki} \quad (8)$$

که  $\rho_{ki}$  در معادله 2 تعریف شده است

سپس برای دفاع در برابر حمله، یک سری گره های قابلاعتماد انتخاب می شوند. برای مثال ویبو می تواند در زمانی انتخاب شود که TN بر اساس مجموعه الگوریتم استفاده شود. در غیر این صورت، ضریب تجزیه  $\alpha$  برای تراست رنک استفاده می شود معادله 8 به صورت زیر بازنویسی می شود

$$TR_i^{t+1} = \begin{cases} \alpha \cdot (\sum_{P_k \in IN_i} TR_k^t \cdot \rho_{k,i}) + (1 - \alpha)/J \\ \alpha \cdot (\sum_{P_k \in IN_i} TR_k^t \cdot \rho_{k,i}) \end{cases}$$

در نهایت، بر طبق امتیازات تراست رنک، TOSI مقدار اسپم (23) را محاسبه می کند و این نشان دهنده احتمال مهاجم بودن فرد است. هر چه وزناسپم بالا باشد، احتمال مهاجم بودن بالا است. فرض کنید که  $SM_i$  نشان دهنده وزن کل اسپم PI است. معادله به صورت زیر است

$$SM_i = (SI_i - TR_i)/SI_i \quad (10)$$

از این روی روش TASI عوامل موثر اجتماعی را توصیه می کند که اثر اجتماعی قوی دارد

الگوریتم: مجموعه اثر اجتماعی همه شرکت کننده های SI، مجموعه وزنی همه SM به شکل زیر است

```

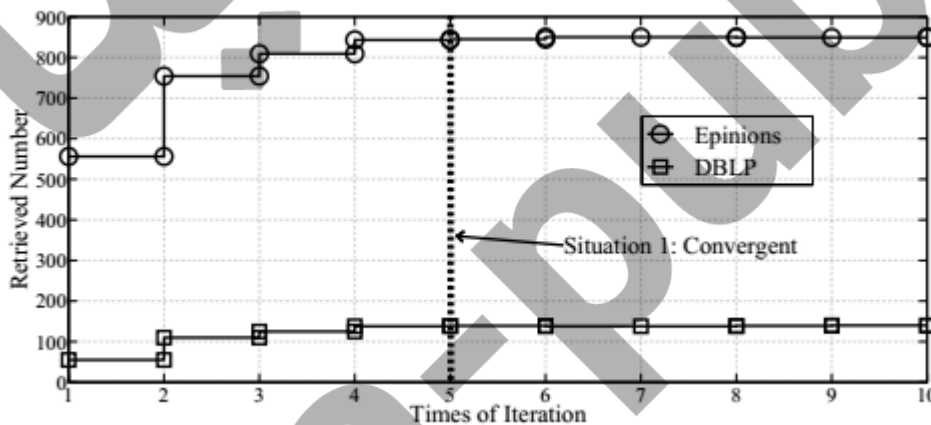
1:  $SI, TR, SM \leftarrow \{rand(1)\};$ 
2:  $NewSI, NewTR \leftarrow \emptyset;$ 
3:  $TotalSI, TotalTR, i \leftarrow 0;$ 
4: while  $i < \lambda$  do /* Iterative evaluate the social influences which are based on
last social influences */
5:  $i \leftarrow i + 1;$ 
6:  $TotalSI \leftarrow 0;$ 
7: for each  $P_j$  in  $P$  do
8: if  $P_j$  is not isolated node then
9:  $v, tr \leftarrow 0;$ 
10: for each node  $P_k$  in the incoming neighbors of  $P_j$  do
11:  $\rho_{k,j} \leftarrow (- \rightarrow R_{k,j}(t) TTTR_k + - \rightarrow R_{k,j}(r) / TTS R_k + - \rightarrow R_{k,j}(p) / TTPS_k) / 3;$  /* Calculate the influence
probability from  $P_k$  to  $P_j$  */
12:  $v \leftarrow v + SI[k] * \rho_{k,j};$  /* Add up the social influence propagating from  $P_k$  to
 $P_j$  */
13: if  $P_j$  is trustworthy participant then /* Add up the TrustRank propagating
from  $P_k$  to  $P_j$  */
14:  $tr \leftarrow tr + \alpha * TR[k] * \rho_{k,j} + (1 - \alpha)/J;$ 

```

```

15: else
16:  $tr \leftarrow tr + \alpha * TR[k] * \rho_{k,j}$ ;
17: end if
18: end for
19:  $NewS I[j] \leftarrow v, NewTR[j] \leftarrow tr$ ;
20:  $TotalS I \leftarrow TotalS I + v, TotalTR \leftarrow TotalTR + tr$ ;
21: end if
22: end for
23:  $NewS I \leftarrow NewS I / TotalS I$ ; /* Reduce the sum of social influences to 1 */
24:  $NewTR \leftarrow NewTR / TotalTR$ ; /* Reduce the sum of TrustRanks to 1 */
25:  $S M \leftarrow (NewS I - NewTR) / NewS I$ ;
26: Replace  $S I$  with  $NewS I$ ,  $TR$  with  $NewTR$ ;
27: end while
28: Return  $S I$  and  $S M$ 

```



شکل 5: همگرایی روش طوسی. محور X زمان تکرار و محور Y عدد بازیابی شده است

#### 4-4 الگوریتم

شبه کد الگوریتم در الگوریتم 1 نشان داده شده است چون TOSI همگرا است و زمان تکرار ثابت است و پیچیدگی زمانی روش TOSI،  $\mathcal{O}(\lambda N^2)$  است که N تعداد شرکت کننده و  $\lambda$  زمان تکرار است

#### 5- آزمایشات

در آزمایشات ما، ما روش TOSI با روش پیشرفته مقایسه می کنیم. SCAP برای بررسی کارایی در منابع، بسیاری از روش های ارزیابی اثر اجتماعی پیشنهاد شده است که در آن مدل آبخاری مستقل (IC) یک مدل برای یافتن گروه های K برتر می باشد و دارای اثر اجتماعی بیشینه در یک شبکه است. متعاقباً، برخی از کار های مهم (11-9)، برای بهبود مقیاس پذیری IC پیشنهاد شده است. به علاوه در سال های اخیر، روش بیشینه سازی اثر محلی (15) برای ارزیابی اثر اجتماعی شرکت کننده خاص در OSN پیشنهاد شده است. به علاوه چون برخی OSN در حال تبدیل شدن به یک ژنراتور زمان واقعی بزرگ می باشد، برخی از روش های استریمینگ برای ارزیابی اثر اجتماعی شرکت کننده ها در OSN پیشنهاد شده است.

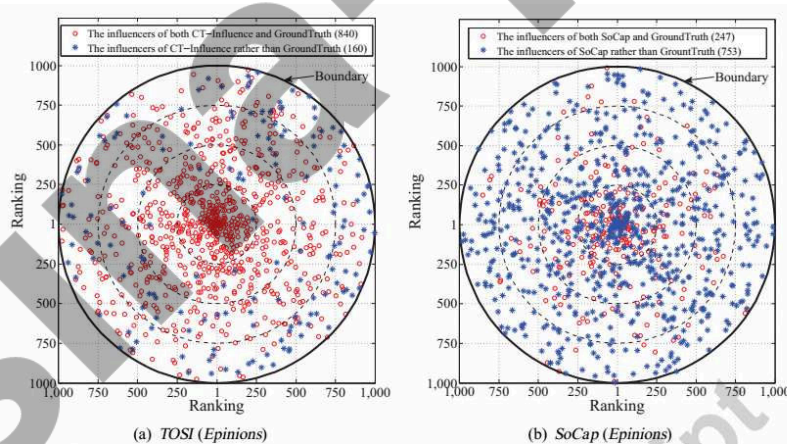
## 1-5 شرایط آزمایشی

جدول 1: دیتاست آزمایشی

Dataset	Epinions	DBLP
Nodes	75,879	317,080
Links	508,837	1,049,866
Average Indegree	6.706	3.311
High Indegree Nodes (Indegree $\geq 50$ )	2032	170
The Ratio of High Indegree Nodes	2.679 %	0.054 %

جدول 2: عملکرد TOSI و SOCAP با تاپ 1000

Method	DataSet	Retrieved Number	Precision	Average Execution Time
TOSI	Epinions	840	0.84	444 ms
SoCap	Epinions	247	0.247	6436 ms
TOSI	DBLP	138	0.138	1244 ms
SoCap	DBLP	44	0.044	8364 ms



شکل 6: 1000 عامل اثر گذار هر روش بر روی اپینیون. محور ایکس و ایکرک رتبه بندی های اثر اجتماعی هستند و حلقه برگ مرزی حاوی 1000 عامل اثر گذار است. دایره های قرمز کوچک شامل هر دو اثر گذاران واقعی و کاذب است

### 1-1-5 مجموعه داده ها

ما از مجموعه داده های شبکه اجتماعی، برای اپینیون (32) و DBLP استفاده می کنیم. مجموعه داده اپینیون دارای 75897 و 508837 لیتک است و هر یک نشان دهنده یک خریدار بوده و هر لینک بین خریداران متناظر است. مجموعه داده DBLP دارای 317980 گره و 10949 گره است. جزییات در جدول 1 نشان داده شده است

### 2-1-5 واقعیت زمینی

بر اساس روش شناسی اجتماعی اگر شرکت کننده بر تعداد زیادی از افراد با نفوذ زیاد اثر داشته باشد، این شرکت کننده نیز اثر بالایی داشته ما اثر گذاران را بر اساس تعداد شرکت به صورت واقعیت زمینی در  $\exp-1$  و  $\exp-3$  است

### 3-1-5 مدل های انتشار

در  $\exp-2$ ، ما از دو روش انتشار کلاسیک استفاده می کنیم مدل استانه خطی و مدل ابشار مستقل. این مدل ها برای بررسی اثر بخشی ارزیابی اثر در 35-36-37 با مقایسه گره ها نشان داده شده است

- مدل استانه خطی: مدل LT اولین مدل برای شبیه سازی فرایند انتشار اطلاعات است. این رویکرد بر

اساس روش های استانه خاص است. در این مدل، در مرحله T، همه گره ها تحت تاثیر قرار می گیرند. یک

شرکت کننده Pi بر اساس تابع  $f(In(i, t)) \in [0, 1]$  تحت تاثیر قرار می گیرد و  $\theta_i \in [0, 1]$  می باشد

$$f(In(i, t)) = \sum_{P_j \in In(i, t)} b_{i,j} \quad (11)$$

که  $In(i, t)$  همسایه های pi در زمان t است

$$b_{i,j} = \frac{\vec{R}_{i,j}(t) + \vec{R}_{i,j}(s) + \vec{R}_{i,j}(p)}{\sum_{P_k \in On_i} (\vec{R}_{i,k}(t) + \vec{R}_{i,k}(s) + \vec{R}_{i,k}(p))} \quad (12)$$

جدول 3: مقایسه TOSI و SOCAP با واقعیت زمینی

Nodes' ID	Ground Truth Ranking	TOSI Ranking	SoCap Ranking
18	1	1	108
737	2	2	269
401	3	3	308
40	4	4	631
118	5	6	669
34	6	7	1184 (missing)
550	7	8	6226 (missing)
136	8	12	6442 (missing)
143	9	23	6448 (missing)
1719	10	32	23842 (missing)

بر اساس استانه ها به ازای هر PI داریم  $\theta_i \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  و  $\sum_{P_i \in On_i} b_{i,j} \leq 1$  همسایه های خروجی هستند. در این آزمایشات برای بررسی اثر بخشی روش

- آبخار مستقل: مدل IC یک مدل ابشار دینامیک برای فرایند انتشار است. این مدل بر اساس سیستم ذرات - تعاملی از نظریه احتمال است. در هر زمان T، هر شرکت کننده یا تحت تاثیر قرار گرفته یا حساس است. یک PJ شرکت کننده در زمان t-1، شانس اثر گذاری بر روی همه همسایه ها وجود دارد. اثرات بر احتمال  $p_{i,j}$  است. احتمال اثر گذاری  $p_i$  بر روی t به شکل زیر است

$$f(i, t) = 1 - \prod_{P_j \in In(i, t-1)} (1 - p_{i,j}) \quad (13)$$

که  $In(i, t-1)$  همسایه های ورودی  $p_i$  در زمان t-1 است.  $p_{i,j} = (\vec{R}_{i,j}(t) + \vec{R}_{i,j}(s) + \vec{R}_{i,j}(p))/3$

در آزمایشات، عامل اثر گذار k برتر توسط TOSI و SOCAP در مدل های انتشاری به صورت  $K \in \{1, 5, 10, 20, 50, 100\}$  قرار می گیرد. بر اساس ویژگی های مدل انتشاری، تعداد گره های متاثر از مدل می

توانند نشان داده شوند

#### 4-1-5 محیط آزمایش

همه آزمایشات بر روی کامپیوتر با پردازنده Intel Core i5-3470 CPU 3.20 GHz با حافظه 8 گیگ

انجام شد. این کد با استفاده از Visual C++ 2012 اجرا شده داده ها با MySQL مدیریت شدند

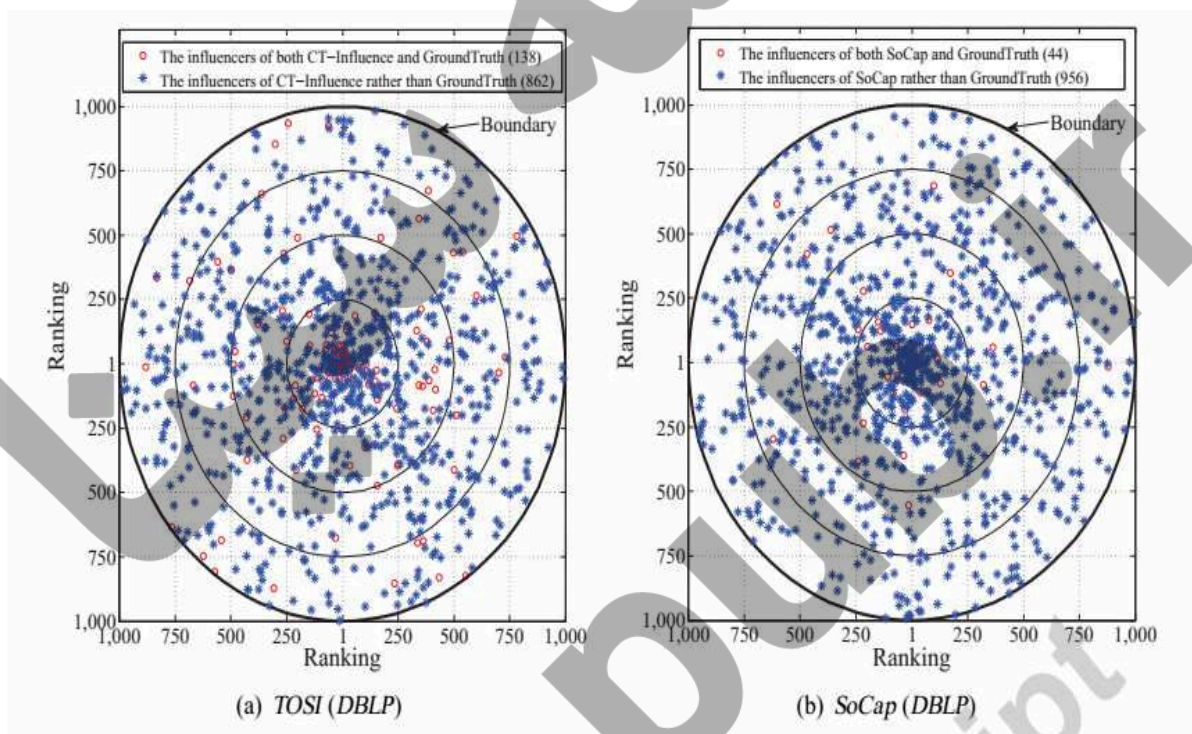


## 2-5 نتایج آزمایشی و تحلیل

## EXP-1 1-2-5، اثر بخشی

دقت با 1000 عامل اثر گذار مقایسه شد:

- در ابتدا، روند تغییرات عدد بازیابی شده با افزایش زمان تکرار برای بررسی همگرایی روش TOSI بررسی شد. عدد بازیابی شده تعداد عوامل اثر گذار است. نتایج آزمایشی بر اساس مجموعه داده های اپینیون و مجموعه داده های DBLP در شکل 5 نشان داده شده است و از این روی این داده ها بر مبنای روش TOSI هستند و در زیر تعداد تکرار  $\lambda$  مطلوب است



شکل 7: 1000 اثر گذار بر تر هر روش بر روی DBLP. محور ایکس و ایگرک بر اساس رتبه بندی اثرات

اجتماعی بوده و دایره بزرگ مرز حاوی 1000 عامل اثر گذار بر اساس هر روش است. دایره کوچک به صورت

1000 عامل اثر گذار و ده عامل اثر گذار می باشد. ستاره های ابی 1000 عامل موثر تر است.

- دوما پس از 5 تکرار، نتایج آزمایشی در جدول 2 نشان داده شده است. برای اپینیون، روش TOSI، 840

اثر گذار دارد در حالی که SOCAP تنها 247 اثر گذار را یافته است. بر اساس تابع دقت در معادله 8،

دقت روش TOSI 84 درصد است. بر عکس، تنها 24.7 درصد روش SOCAP است. از این روی مقایسه

با SOCAP نشان می دهد که روش ما موجب افزایش دقت ارزیابی تا 240 درصد می شود. برای DBLP، روش TOSI 138 مورد را از واقعیت های زمینی تعیین کرده است.

دقت = مرتبط / ابازیابی شده / ابازیابی شده

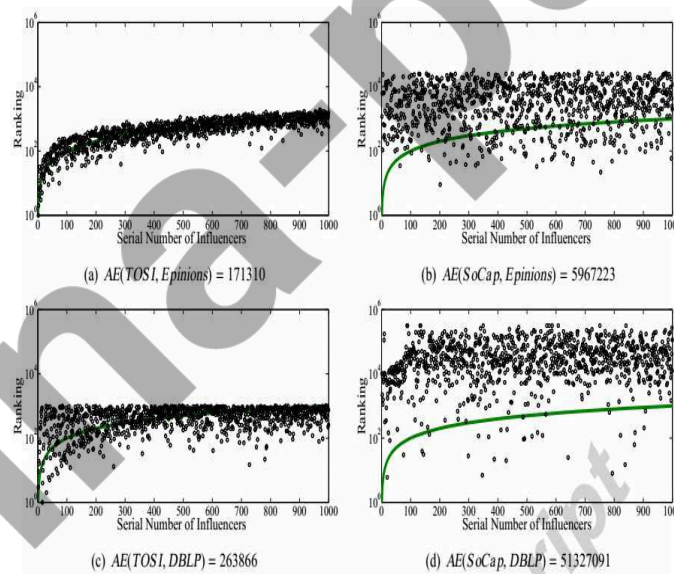
- سپس فهرست نتایج 10 اثر گذار برتر ابازیابی شده توسط روش در جدول 3 نشان داده می شود. بر اساس جدول 3، روش TOSI قادر به یافتن 10 اثر گذار برتر است و رتبه بندی TOSI نزدیک به رتبه بندی ده واقعیت زمینی است. ولی اثر گذاران دارای رتبه بندی 10 اثر گذار بودند

- نتایج آزمایشی اپینیون و DBLP در شکل 6 و 7 نشان داده شده است که در آن تعداد اثر گذاران بر اساس روش طوسی بیش از SOCAP با رتبه بندی بالاتر است. از این روی TOSI نتایج ارزیابی بهتری نسبت به SOCAP دارد

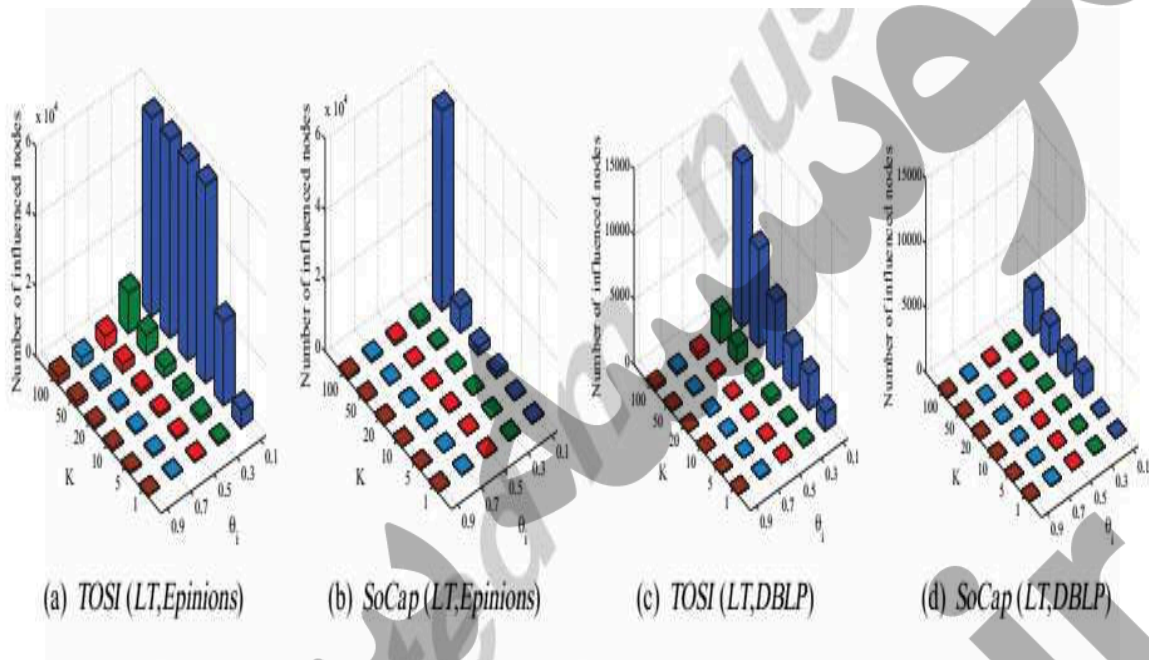
- ما از تابع خطای مطلق برای اندازه گیری خطای هر روش استفاده می کنیم. خطا، مقدار مطلق بین رتبه بندی اثر با هر روش و رتبه بندی واقعیت زمینی است

$$AE(\text{method}, \text{dataset}) = \sum_{P_i \in GT} |RA(P_i) - GTRA(P_i)| \quad (15)$$

که GT مجموعه ای از 1000 اثر گذار واقعی،  $RA(P_i)$  رتبه بندی اثر PI است که توسط هر روش ارزیابی می شود و  $GTRA(P_i)$  رتبه بندی واقعیت زمینی PI است.



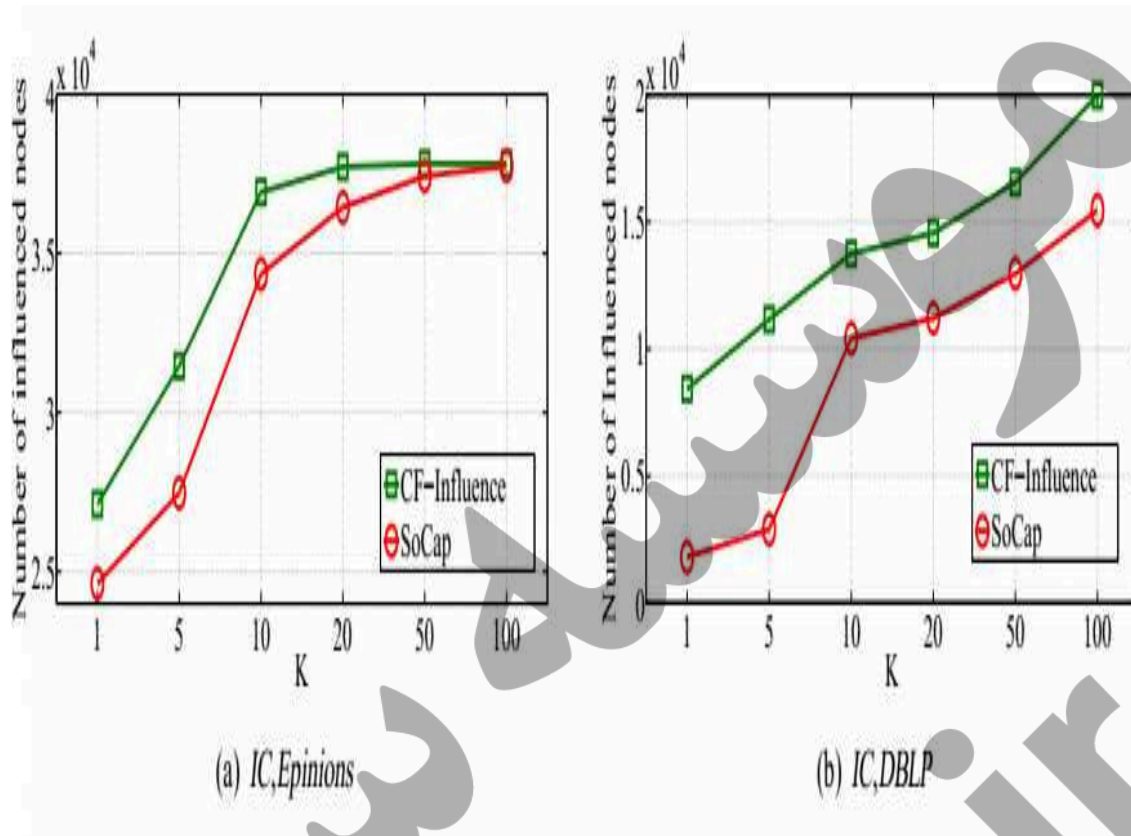
شکل 8: خطای هر روش بر روی هر مجموعه داده. محور X تعداد سریال اثر گذاران است. محور Y رتبه بندی اثر اجتماعی است. منحنی سبز، رتبه بندی واقعیت زمینی و دایره سیاه، رتبه بندی عوامل اثر گذار است



- در نهایت، ما به مطالعه خطای مطلق و صحت روش TOSI با مقایسه 1000 اثر گذار بر تر بر اساس هر روش و واقعیات زمینی می پردازیم. چون SOCAP رابطه اجتماعی را در نظر گرفته است، قادر به ارزیابی نتایج نیست. به این ترتیب روش TOSI عملکرد بهتری از SOCAP در ارزیابی بر اساس نتایج واقعی دارد

#### EXP-2 2-2-5: اثر بخشی بر اساس مدل های انتشار

شکل 9، نتایج آزمایشی مدل LT را نشان می دهد و تعداد گره های تحت تاثیر توسط TOSI با K متفاوت شناسایی شده است و  $\theta_i$  بیش از SO CAP است. تعداد متوسط گره ها با TOSI برابر با 560918 می باشد و از این رو با SO CAP برابر با 137102 و کم تر از TOSI به میزان 75.56 درصد است. به علاوه تعداد گره های متاثر شناسایی شده با دو روش با افزایش K افزایش می یابد. زیرا با افزایش K، تعداد منابع برای اطلاعات افزایش می یابد. به علاوه، تعداد گره های متاثر با دو روش با افزایش  $\theta_i$  کاهش می یابد.



شکل 10: تعداد گره های متاثر بر روی مدل ابشار مستقل

جدول 4: وزن اسپم بر روی حمله کننده ها

Tag	Dataset	TOSI Ranking, $SM(\text{Attacker Probability})$	SoCap Ranking
Honest Influencer	Epinions	1, -1.13(low)	1199
SF-Attacker-50	Epinions	53, 0.94(high)	2127
SF-Attacker-100	Epinions	48, 0.93(high)	809
SF-Attacker-200	Epinions	51, 0.96(high)	1170
SF-Attacker-500	Epinions	42, 1.06(high)	423
SF-Attacker-1000	Epinions	58, 1.97(very high)	230
Honest Influencer	DBLP	1, -2.24(very low)	65256
SF-Attacker-50	DBLP	20, 0.86(high)	19754
SF-Attacker-100	DBLP	17, 0.95(high)	5165
SF-Attacker-200	DBLP	18, 1.21(high)	30
SF-Attacker-500	DBLP	9, 1.93(very high)	856
SF-Attacker-1000	DBLP	19, 2.23(very high)	82

شکل 10 تعداد گره های متاثر را با TOSI و SOCAP نشان می دهد که با افزایش  $K$  در مدل IC، می توان دید که تعداد گره های متاثر TOSI بیش از SOCAP در همه 6 مورد در دو مجموعه داده است. تعداد متوسط گره های متاثر با TOSI برابر با 244192 است و Socap برابر با 210695 است و 13.8 درصد کم تر از قبلی است این بر

اساس خواص مدل  $ic$  در مدل های انتشاری است. به علاوه با افزایش  $k$ ، تعداد گره های متاثر از گره های  $k$  با هر دو روش افزایش می یابد

از نتایج آزمایشی در مدل های انتشار کلاسیک  $ic$  -  $it$ ، عوامل موثر بر  $k$  با  $TOSI$  موجب بهبود کارایی تا 90 درصد می شود. از این روی روش  $TOSI$  عملکرد بیشتری از  $So\ cap$  در اثر بخشی بر اساس دو مدل انتشاری دارد

**5-2-3-3-exp، استواری**

در ابتدا 5 حمله کننده اسپکرم فارم با 500 تکرار تصادفی تنظیم می شوند. از این ری  $M \in \{50, 100, 200\}$  در  $\{500, 1000\}$  در زیر تعیین شده است. سپس عوامل اثر گذار واقعیت زمینی [23] 0.85 استفاده می شود. در

نهایت شرکت کننده های اسپم مس بر اساس معادله 10 با اثر گذار صادق مقایسه می شود. از این روی  $TOSI$  قادر به شناسایی عوامل اثر گذار برای بهبود استواری اجتماعی است.  $SOCAP$  قادر به یافتن این عوامل نیست

#### 5-2-4-XP کارایی

جدول 2 دفعات اجرای ارزیابی اثر دو روش را نشان می دهد. در مجموعه داده اپینیون، 444 برای  $TOSI$  است و از این روی این روش قادر به صرفه جویی 93.1 درصد در زمان اجرا است بر روی مجموعه  $DBLP$ ، 1.244 میلی ثانیه برای  $TOSI$  و 8.364 برای  $SOCAP$  است. به طور متوسط روش ما قادر به صرفه جویی 85.1 است. از این روی بر اساس قضیه 1، همگرایی  $TOSI$  سریع است. روش  $TOSI$  عملکرد بهتری از نظر کارایی نسبت به  $SO$  دارد

#### 6- نتیجه گیری و کار های آینده

در این مقاله، ما یک روش ارزیابی اثر اجتماعی مبتنی بر اعتماد را پیشنهاد کرده ایم که بر اساس سه زمینه اجتماعی هستند. یک روش ارزیابی اثر اجتماعی اعتماد محور موسوم به  $TOSI$  با در نظر گرفتن شرایط اجتماعی پیشنهاد شد. آزمایشاتی بر روی دو مجموعه شبکه اجتماعی انجام شد یعنی اپینیون و  $dblp$ . نتایج آزمایشی نشان می دهد که روش  $TOSI$  عملکرد بهتری بر روی روش پیشرفته  $SOCAP$  از حیث اثر بخشی، کارایی و استواری دارد.

در مطالعات آینده هدف ما استفاده از روش  $TOSI$  به  $OSN$  برای بررسی نتایج ارزیابی اثر و توصیه پارانتر های مطمئن تر است.