

محمد مهدی دلیری خمامی^۱، علیرضا رضوانیان^۲ و محمد رضا میبدی^۳

^۱ دانشگاه آزاد اسلامی، واحد قزوین، گروه مهندسی نرم افزار، قزوین، ایران، m.daliri@qiau.ac.ir

^۲ دانشگاه صنعتی امیرکبیر، آزمایشگاه محاسبات نرم، گروه مهندسی نرم افزار، تهران، ایران، a.rezvanian@aut.ac.ir

^۳ دانشگاه صنعتی امیرکبیر، آزمایشگاه محاسبات نرم، گروه مهندسی نرم افزار، تهران، ایران، mmeybodi@aut.ac.ir

چکیده - با توجه به ارتباطات میان کاربران در شبکه های اجتماعی آنلاین، انتشار اقلام مختلف اطلاعاتی نیز بسیار مورد توجه قرار گرفته است. به طوریکه توجه بسیاری را در بازاریابی مدرن در این شبکه ها به این مسئله جلب نموده است. تا کنون مدلها مخالفی برای مدلسازی انتشار در شبکه های اجتماعی ارائه شده است که در اکثر این مدلها از حالت ساده ای به صورت کاربران فعال و غیرفعال استفاده شده است. با توجه به این نکته که در دنیای واقعی، نحوه تاثیرپذیری افراد در انتشار اطلاعات به صورت مرحله ای صورت میگیرد، در این مقاله از این ایده برای پیشنهاد یک مدل جدید برای انتشار به صورت مرحله ای استفاده شده است. برای مدل پیشنهادی از اتماتای سلوی نامنظم بهره برده شده است. که در این مدل حالت های مختلفی از شروع انتشار تا پذیرش توسط کاربر به عنوان وضعیتی از سلوهای اتماتای سلوی نامنظم در نظر گرفته شده است. به منظور ارزیابی مدل پیشنهادی آزمایشای مختلفی بر روی دادگان استاندارد شبکه های اجتماعی معروف صورت گرفته است و براساس پارامترهای مختلفی نیز مورد مقایسه قرار گرفته است. نتایج شبیه سازی حاکی از برتری نسبی مدل پیشنهادی در مقایسه با مدل انتشار آستانه خطی می باشد.

کلید واژه - شبکه های اجتماعی، گسترش انتشار، مدل انتشار، اتماتای سلوی نامنظم

مقدمه

استفاده کردند. آنها از الگوریتم ژنتیک برای جستجو و انتخاب ترکیب

بهینه از مشخصه فیزیکی شبکه استفاده نموده اند. بن و همکاران [۶]

یک الگوریتم فرا ابتکاری بر پایه الگوریتم ژنتیک برای انتخاب نودهای اولیه پیشنهاد داده اند. اکثر الگوریتم های اشاره شده به مدل انتشار بستگی شدیدی دارند. به عبارت دیگر کاهش زمان اجرا در مدل انتشار، منجر به اکتساب نتایج در زمان بهتری می شود. بدین گونه کارایی فرایند انتشار از مقیاس پذیری بهتری بهره می گیرد.

علاوه بر الگوریتم های انتخاب نودهای اولیه در فرآیند انتشار، بخشی از تحقیقات بر روی مدلها انتشار تمرکز یافته است. یکی از چالش ها در مدل انتشار، اینست که چگونه می توان فرایند انتشار را مناسب تر شبیه سازی نمود. یکی از اولین مدل های انتشار مدل باس می باشد که برای شبیه سازی فرایند تاثیرپذیری یک محصول بکار رفته است. اما این مدل صرفاً بر روی روند انتشار تاکید داشته و نحوه انتخاب کاربران اولیه را در نظر نمی گیرد همچنین این مدل فاقد شبیه سازی دنیای واقعی می باشد. کیمپی و همکاران دو مدل پایه ای از فرایند انتشار را معرفی نمودند. مدل آستانه خطی^۱ و مدل آبشاری مستقل که بعدها بصورت گسترده ای مورد استفاده قرار گرفته است. این دو مدل شبیه سازی بهتری در مقایسه با مدل باس ارائه می کنند.

در این مقاله یک مدل جدید برای انتشار بر مبنای اتماتای سلوی نامنظم ارائه شده است که در مدل پیشنهادی برای هر کاربر مقادیر مختلفی از آستانه تاثیر در نظر گرفته می شود تا بتوان به حالت نزدیکتری از واقعیت دست یافت. در واقع هدف از این مقاله، در ابتدا ارائه یک مدل حقیقی از فرایند انتشار که بتواند فرایند انتشار را در دنیای واقعی شبیه سازی کند دوم تسریع فرایند انتشار بطوریکه برای

اخیرا با گسترش خدمات آنلاین بر روی اینترنت مانند وب سایتهاي شبکه های اجتماعی، مطالعه بر روی شبکه های اجتماعی از جنبه های مختلفی بیش از گذشته مورد توجه قرار گرفته است. تقریباً انواع وب سایت های شبکه های اجتماعی با داشتن یک ساختار اجتماعی از روابط میان کاربران، سعی بر تاثیر به برقراری ارتباط، اشتراک تصاویر و انتشار اطلاعات دارند. ویژگی هایی چون رغبت کاربران به اشتراک گذاری اطلاعات در وب سایت های شبکه های اجتماعی، نظر مدیران بازاریابی را برای ترویج محصول بخود بیش از پیش جلب می نماید. [۱] بازاریابی اشاعه ای، اولین بار توسط دومینگوس^۲ و همکاران از جنبه داده کاوی معرفی شد. بازاریابی ویروسی به طور فرمال در گراف به صورت مسئله مسئله را فرموله نموده و ثابت کردند که مسئله و همکاران [۲] این مسئله را فرموله نموده و ثابت کردند که مسئله دارای پیچیدگی سخت بوده و یک الگوریتم حریصانه با ضربت تقریب $\frac{1}{e}$ - ۱ پیشنهاد نمودند. ایده اصلی الگوریتم آنها روش حریصانه بوده سعی در انتخاب نودهایی با بیشترین تاثیرپذیری در هر تکرار را دارد. این روش دارای بار محاسباتی بالایی می باشد. برای حل این مسئله لسکوک و همکاران [۳] الگوریتم CELF^۴ را با استفاده از ویژگی زیرپیمانه ای^۵ پیشنهاد کردند. الگوریتم CELF تعداد دفعات اجرای الگوریتم مدل انتشار را کاهش می دهد. با اینحال چن وی و همکاران [۴] نشان دادند که انتخاب مجموعه کوچکی از نود ها ساعت ها بطول می انجامد. حال چنانچه زمان در فرایند انتشار بهبود یابد این مدل می تواند در بازاریابی اشاعه ای مورد مطالعه قرار گیرد. فورست استوندهال [۵] از معیار فیزیکی شبکه برای چینش نودهای اولیه

در هر سلول فقط بستگی به مقدار همسایگان اطراف آن دارد. و مقدار جدید هر سلول فقط بستگی به مقادیر تعداد محدود(معمولاً یک مرحله) از مراحل قبل دارد. قوانین به صورت تابعی از وضعیت همسایگی‌ها می‌باشد؛ در گروهی از قوانین مقدار یک سلول در مرحله بعدی به مقادیر همسایه‌ها و در گروهی دیگر، تنها به مجموع همسایه‌ها بستگی دارد [٧].

آtomاتای سلولی نامنظم در اواسط دهه ۸۰ در مرجع [٨] معرفی شده است بصورت غیر رسمی، آtomاتای سلولی نامنظم دارای یک پیکربندی از نقاط در فضا بدون هیچ محدودیت از پیش تعیین شده می‌باشد. هر نقطه در این فضا شامل تعداد نقاطی تحت عنوان همسایه می‌باشد. چند ضلعی ورونی برای تعیین ارتباط و مثلث بنده دلانی برای تقسیم فضا و تعیین همسایگان هر نود نمونه مثال هایی از آtomاتای سلولی نامنظم می‌باشد. [٩] دیاگرام ورنوی روشنی برای تقسیم فضا به تعدادی ناحیه می‌باشد. در این دیاگرام به هر مجموعه‌ای از نقاط (که دامنه‌ها، سایتها و یا ژناتورها نامیده می‌شوند) ناحیه‌ای اختصاص داده می‌شود. این نواحی سلول‌های ورنوی نامیده می‌شود. برای یک مجموعه از نقاط دیاگرام ورنوی سطح را به مناطقی تقسیم بنده می‌کند که برای هر نقطه از مجموعه نقاط یک منطقه تعیین می‌شود. به طوری که تمام نقاط این منطقه به نقطه تولید کننده آن منطقه نزدیکتر می‌باشد. در حالی که مثلث بنده دلانی مثلث بنده نقاط در نومدار ورنوی بوده بطوری که دایره محاط از مثلث بنده نقاط تپه می‌باشد. در این مقاله، ما از آtomاتای سلولی نامنظم عنوان یک گراف جهت دار استفاده نموده این بطوریکه هر راس در این گراف شامل یک سلول از آtomاتای سلولی نامنظم می‌باشد.

۳. مدل پیشنهادی

مدل پیشنهادی از چهار وضعیت تشکیل شده است که بیانگر سطح تاثیر نودها در شبکه می‌باشد. در واقع رویکرد اصلی این مدل تاثیر پذیری متفاوت افراد از یکدیگر می‌باشد. همانگونه که در شکل ۱ مشاهده می‌کنید در این مدل چهار وضعیت که بترتیب عبارتند از غیرفعال، فعل، پیروی کننده، و تطبیق یافته تعريف شده است. ایده اصلی در این مدل بدین صورت می‌باشد که یک فرد می‌توان در مواجه با افرادی که قبلاً تاثیر گرفته اند سطح تاثیر مختلفی داشته باشد. مثلاً شرکتی را در نظر بگیرید که برای فرایند بازاریابی به عده ای محدود محصولات خود را رایگان ارائه می‌کند و انتظار دارد که این افراد در شبکه با توجه به فرایند تاثیر محصول مورد نظر را اشاعه دهند، در این مورد مسئله مهم سطح تاثیر پذیری افراد برای تطبیق با محصول می‌باشد. در واقع دسته‌ای از افراد هستند که از محصول آگاهی لازم را پیدا می‌کنند اما آگاهی و رغبت آنها به اندازه‌ای نیست که ترغیب به خرید محصول شوند که از این دسته در مدل عنوان اعضای فعل یاد می‌شود، اما دسته‌ای دیگری نیز وجود دارند که دارای آگاهی بوده ولی نیاز به تحریک بیشتری برای خرید محصول دارند در حالتی هم می‌توان افرادی را در نظر گرفت که این افراد به

بازاریابی اشاعه‌ای مناسب باشند. در این مقاله یک مدل انتشار چندگانه با استفاده از اتماتاتی سلولی نامنظم پیشنهاد شده است که در این حالت، کلیه‌ی نودهای شبکه به صورت یک سلول در نظر گرفته می‌شود. همچنین کلیه سلول‌ها وضعیت خود را با توجه به قوانین طراحی شده در حالات تعیین شده در زمان‌های گسسته تغییر می‌دهند. مدل پیشنهادی با اقتباس از فرآیند انتشار در واقعیت عمل می‌کند. در واقع در ابتداء ممکن است یک کاربر فعل شده باشد (یکاید محصول به وی معرفی شده باشد یا یک خبر به وی برسد) اما لزوماً بدین معنی نیست که این کاربر آن را دنبال کند، بلکه حتی ممکن است عنوان ترویج دهنده (تبليغ کننده) محصول باشد. همچنین ممکن است که ترویج دهنده محصول این محصول را بوسیله اشتراک گذاری به دیگر کاربران در جهت خرید یا عدم خرید اشاعه دهد. به هر حال با گذشت زمان، کاربر فعل شده یا دنبال کننده ممکن است محصول مورد نظر را خریداری کند. مدل پیشنهادی بر روی دادگان استاندارد و معروف شبکه‌های اجتماعی مورد ارزیابی قرار گرفته است و در مقایسه با مدل‌های دیگر مورد بحث و تحلیل قرار گرفته است. این مقاله در ادامه به این صورت سازماندهی شده است: در بخش دوم آtomاتای سلولی نامتقاضی معرفی می‌شود. در بخش سوم مدل پیشنهادی انتشار چندگانه مبتنی بر آtomاتای سلولی همراه با جزئیات شرح داده شده است در بخش چهارم الگوریتم پیشنهادی معرفی گردید در بخش پنجم به معرفی دادگان آزمایشی و شبهی سازی اختصاص یافته است. بخش ششم نتیجه‌گیری و کارهای آینده به این مقاله پایان میدهد.

۲- آtomاتای سلولی نامنظم

در این مقاله کاربرد مدل جدیدی تحت عنوان آtomاتای سلولی (CA) می‌باشد در مدل‌سازی انتشار مورد مطالعه و ارزیابی قرار می‌گیرد و نشان داده می‌شود که این مدل می‌تواند مدل مناسب تری نسبت به مدل‌های ذکر شده فوق باشد.

آtomاتای سلولی شبکه‌ای از سلول‌ها است که هر کدام می‌تواند وضعیت داشته باشد در حالت یک بعدی، اگر سلول i دارای دو همسایه $i-1$ و $i+1$ باشد در این صورت وضعیت سلول i در زمان $t+1$ یعنی $a_i^{(t+1)}$ بر اساس تابع φ زیر بدست می‌آید.

$$(1) \quad a_i^{(t+1)} = \varphi(a_{i-1}^{(t)}, a_i^{(t)}, a_{i+1}^{(t)})$$

تابع $a_i^{(t+1)}$ را قانون آtomاتای سلولی می‌نامیم. ایده همسایگی در آtomاتای سلولی را می‌توان بگونه‌ای بسط داد که دو همسایه دیگر و یا بیشتر را شامل شود. اگر چهار سلول در طرفین یک سلول همسایه محسوب شوند (دو سلول در هر طرف) آtomاتای سلولی دارای شعاع همسایگی (۲) مساوی ۲ می‌باشد. سلولها در آtomاتای سلولی می‌توانند با هر ابعادی قرار گیرند

ویژگی اساسی آtomاتای سلولی عبارتند از: فضا و زمان بصورت گسسته می‌باشد، هر سلول تعداد محدودی وضعیت ممکن را اختیار می‌کند، تمام سلول‌ها یکسان می‌باشند، بروز در آوردن سلول‌ها بشکل همگام می‌باشد، قوانین بصورت قطعی اعمال می‌شوند، قانون

قوانين بروزرسانی بالا نشان دهنده نحوی قرار گیری نودها در وضعیت های تعریف شده آنها می باشد . در قوانین بالا (S_t) معروف وضعیت نود v در زمان t می باشد، همچنین تابع f_{ij} مقدار تاثیر نود i بر روی نود j را نشان می دهد. فرض ما در این مقاله این است که وقتی یک نود در وضعیت شامل محصول قرار می گیرد حالت خود را حفظ می کند. تابع $T(\Delta t)$ بعنوان تابع کاهشی محسوب میشود که یک تابع غیر صعودی می باشد . در واقع تابع کاهشی $T(\Delta t)$ برای شبیه سازی این واقعیت که تاثیر در طول زمان بطور تدریجی کاهش می یابد . به عبارت دیگر یک کاربر با احتمال بیشتری پیشنهادی که از دوستان خود در زمان اولیه به او داده می شود را قبول می کند. در غیر اینصورت احتمال اینکه به نوبه خود یک محصول را بپذیرد در طول زمان کاهش پیدا می کند . ما تابع کاهش تاثیر را برای شبیه سازی و ارزیابی رویه انتشار بصورت زیر معرفی می کنیم :

$$T(\Delta t) = F\left(\frac{1}{1 + e^{-\Delta t}}\right) \quad (\text{¶})$$

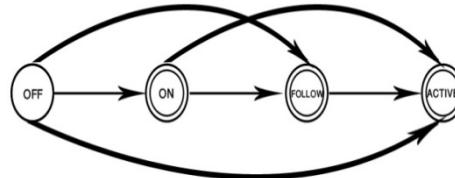
در فرمول (۳) مقدار F تاثیر بر روی یال ها می باشد .

٤. الگوریتم پیشنهادی

براساس مدل انتشار چندگانه پیشنهادی با استفاده از آناتاماتی سلولی نامنظم ، الگوریتمی مبتنی بر آناتاماتی سلولی ارائه می گردد. شبه کد الگوریتم بشرح زیر می باشد . همانطور که در الگوریتم شماره ۱ مشاهده می شود وروی الگوریتم یک گراف جهت دار وزن دار و مجموعه نودهای فعال اولیه که با نماد S نشان داده شده می باشد همچنین K تعداد رؤوس در کل شبکه است. خروجی الگوریتم تعداد نودهای فعال شده در اتمام الگوریتم می باشد. در مرحله ارزش دهی آغازین متغیر زمان را t برابر صفر و مجموعه A را عنوان مجموعه وضعیت نودها در طول اجرای الگوریتم تعریف می کنیم در واقع در این مجموعه تغییر وضعیت لحظه می شود در ابتدا وضعیت نودها off بوده یعنی هیچ نودی توسط الگوریتم فعال یا بعبارت دیگر تغییر وضعیت نداده است . خطوط ۱ تا ۶ الگوریتم برای جستجوی نودهای وروودی به هر نود بکار رفته است . پس از اتمام این مرحله در خطوط ۸ تا ۱۰ وضعیت تمام نودهای اولیه فعال را بروز رسانی می نماییم. در گام بعدی الگوریتم تا زمانی ادامه می یابد که هیچ نودی تغییر وضعیت ندهد زیرا عدم تغییر وضعیت به معنی پایداری و تعادل شبکه می باشد در این مرحله برای هر نود بررسی می شود که آیا نود در حال بررسی می تواند وضعیت خود تغییر دهد یا خیر اگر این نود وضعیت خود را تغییر دهد. وضعیت آن در مجموعه A و در زمان تغییر مثلاً A

بروز می شود شرح مراحل شامل خطوط ۱۱ تا ۱۹ الگوریتم می باشد.

محض ترغیب شدن به محصول آن را خریداری می نمایند که از این دسته بعنوان افراد تطبیق یافته یاد می شود.



شکل ۱: مدل پیشنهاد شده چند گانه همراه با وضعیت‌های قرار گرفتن هر نودیک گراف جهت دار $G(v, E)$ که v معرف مجموعه تمام نودها در شبکه اجتماعی و E مجموعه تمام یال‌ها نشان داده می‌شود. برای هر یال (u, v) مقدار احتمال وجود دارد، که نشانگر تاثیر کاربر u بر کاربر v می‌باشد. مدل انتشار پیشنهادی به عنوان ورودی تعدادی از کاربران را به عنوان کاربران اولیه دریافت می‌کند و خروجی مدل کاربران فعال شده در پایان فرایند انتشار بر روی شبکه می‌باشد. در مدل پیشنهادی از آناتاماتی سلولی نامنظم استفاده شده است که توسط $I = (S, A, N, R, Sum)$ توصیف شده است. بطوریکه مجموعه S مجموعه تمام حالت‌های سلولی که در شکل ۲-۲ $s = \{0 - 0.5, 0.5 - 1\}$ راene شده می‌باشد. درینجا مجموعه $0.6 - 0.8, 0.6 - 0.8$ در نظر گرفته شده است. بدین معنی که اگر سلولی مقدار آستانه آن در اجرای الگوریتم بین ۰.۵ و ۰.۶ باشد این نod غیر فعال بود و اگر ۰.۵ و کمتر از ۰.۶ باشد این نod فعال شده و اگر بین ۰.۶ تا ۰.۸ باشد این نod بعنوان نod پیروی کننده محصول^۷ و اگر ۰.۸ تا ۱ بعنوان تطبیق یافته محصول^۸ در نظر گرفته می‌شود. در سایر حالات نودها بعنوان نودهای غیر فعال در نظر گرفته می‌شوند. همانگونه که زمان پیش می‌رود نودها بیشتر و بیشتر پیشنهادات همسایگان خود را قبول می‌کنند و فرایند تاثیر پذیری ادامه می‌یابد. اجتماع تمام حالات نودها در شبکه می‌باشد $A = S^n$ که n تعداد نودها می‌باشد. در هنگام شروع همه همچنین $A_0 = \{0, 0, 0, 0, 0\}$ در نودها غیر فعال می‌باشد بنابراین A_t در زمان t وقتی که $A_{t-1} = A_t$ مدل اجرای خود را متوقف می‌کند زیرا هیچ تغییری در وضعیت نودها ظاهر نمی‌شود. برای هر سلول مجموعه نودهای ورودی تعریف می‌کنیم. بعنوان مثال مجموعه نودهای ورودی برای نod v_1 بصورت $\{v_2, v_3, v_9, \dots\}$ مولفه $\Omega_{v_1} = \{v_2, v_3, v_9, \dots\}$ مجموعه کل نودهای ورودی می‌باشد^۹. مولفه $N = \{\Omega_{v_1}, \Omega_{v_2}, \dots\}$ مجموعه کل نودهای فعال شده می‌باشد. R قوانین اثamatی سلولی، نامنظم برای انتشار می‌باشد که بصورت (۲) تعریف می‌شود:

$$S_{t+1}(v_i) = \begin{cases} off & S_t(v_i) = off, \sum_{j \in \Omega(i)} (S_t(j)F_t(v_{i,j}) - T(t)) < 0.5 \\ on & S_t(v_i) = off, 0.5 \leq \sum_{j \in \Omega(i)} (S_t(j)F_t(v_{i,j}) - T(t)) < 0.6 \\ follow & S_t(v_i) = off \text{ or } S_t(v_i) = on, 0.6 \leq \sum_{j \in \Omega(i)} (S_t(j)F_t(v_{i,j}) - T(t)) < 0.8 \\ & S_t(v_i) = off \text{ or } S_t(v_i) = on \text{ or } S_t(v_i) = follow \\ active & 0.8 \leq \sum_{j \in \Omega(i)} (S_t(j)F_t(v_{i,j}) - T(t)) = 1 \end{cases},$$

(¶)

عنوان مدیران شبکه مدیریت می‌گردد. این مجموعه دادگان در تاریخ ۲۰۰۸ جمع آوری شده است. مجموعه داده چهارم Gnutella که یک شبکه اشتراک گذاری فایل می‌باشد که بخشی از این شبکه در اگوست ۲۰۰۲ جمع آوری شده است. جزئیات مربوط به هریک از این دادگان در جدول (۱) ارائه شده است.

جدول ۱: شرح جزئیات مربوط به مجموعه دادگان آزمایشی

متوجه درجه	چگالی	قطر	ضریب خوشگی	یالها	رتوس	دادگان آزمایشی
۰.۴۴۰	۰.۰۱۳	۶	۰.۴۴	۵۰۹	۶۹۷	Squeakfo undation
۵.۸۵۰	۰.۰۰۱	۱۱	۰.۲۲	۳.۴۹۲	۱۶۹۸	Robots.net
۲۵.۰	۰.۰۰۲	۷	۰.۱۴۰۹	۱۰۳۶۸۹	۷۱۱۵	Wiki-vote
۴.۷۲	۰.۰۰۰۴	۱۰	۰.۰۰۶۷	۳۱۵۲۵	۸۷۱۷	Gnutella

در آزمایش اول، هدف بررسی تاثیر تعداد نودهای اولیه در انتشار در یک شبکه اجتماعی می‌باشد. بدین منظور از دادگان معرفی شده برای آزمایش استفاده شده است. تعداد نودهای اولیه از ۵٪ تا ۲۵٪ از نودهای هر شبکه به عنوان نودهای اولیه به طور تصادفی در نظر گرفته شده است. در مدل انتشار پیشنهادی که جزئیات آن در بخش ۲ توضیح داده شد با گذشت زمان نودهای غیرفعال موجود در شبکه، از نودهای اولیه همسایگی تاثیر پذیرفته و انتشار در شبکه گسترش پیدا می‌کند. این گسترش تا زمانی که دیگر تغییری در وضعیت شبکه حاصل نشود ادامه می‌یابد. نتایج انتشار برای تعداد نودهای اولیه با شرایط مذکور برای مدل آستانه خطی با عنوان LTM و برای مدل پیشنهادی با عنوان MVICA در جداول (۲)، (۳)، (۴)، (۵) نمایش داده شده است. در نتایج گزارش شده میانگین و انحراف معیار زمان لازم برای انتشار و نرخ انتشار (نرخ پوشش) با توجه به تعداد متغیر نودهای اولیه ذکر شده است.

جدول ۲: نتایج مقایسات مدل‌های انتشار LTM و MVICA برای شبکه Squeakfoundation

LTM		MVICA		تعداد نودهای اولیه
نرخ پوشش	زمان انتشار	نرخ پوشش	زمان انتشار	
۵۶.۳±۷.۹	۲.۷±۰.۶	۵۸.۳±۹.۷	۰.۴±۲.۷	۳۵
۱۱۲.۲±۱۹.۲	۲.۹±۰.۴	۱۲۸.۱±۳۴.۶	۲.۰۳ ۰.۴±	۷۰
۲۰۲.۱±۲۵.۵	۳.۲±۰.۵	۲۲۰.۸±۷۰.۰	۳.۱±۰.۳	۱۰۵
۲۷۰.۷۰±۷۹.۱	۲.۲±۰.۴	۳۱۴.۴±۶۹.۵	۳.۵±۰.۶	۱۴۰
۳۴۵.۲±۸۴.۷	۳.۴±۰.۶	۳۴۹.۲±۷۶.۸	۳.۴±۰.۵	۱۷۵

Algorithm 1: proposed cellular automata based algorithm

Input:

$G=(V,E)$ is a Directed graph

$N=|V|$ // number of vertices

$K=|S|$ //number of seed set.

Output:

Number of active node's according to proposed model

Pseudo-code:

Initialize $t=1$; $A=\{off, off, \dots, off\}$

01:**Begin Algorithm**

02:**For** $i=1$ to n

03: **For** $j=1$ to n

04: **If** $((V_i, V_j) \in E)$ **then**

05: $\Omega_{V_j} \leftarrow V_i$

06: **End for**

07: **End for**

08:**End For**

09:**For** each V in S

10: $V.state = active$

11:**End for**

12:**Repeat**

13:**If** $V_j.state$ Is Equal Off Or $V_j.state$ Is Equal On or $V_j.state$ Is Equal Follow) **Then**

14: Automaton V_i Compute

$(\sum_{j \in \Omega(i)} (S_t(j)F_t(v_{i,j}) - T(t)))$

15: **If** ($V_j.state$ is changed) **Then**

16: Update V_j State

17: Update A_t

18: **End if**

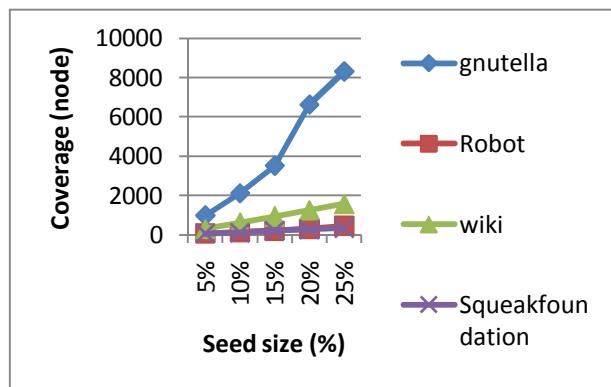
19:**End if**

20:**Until** A_{t-1} Equal to A_t

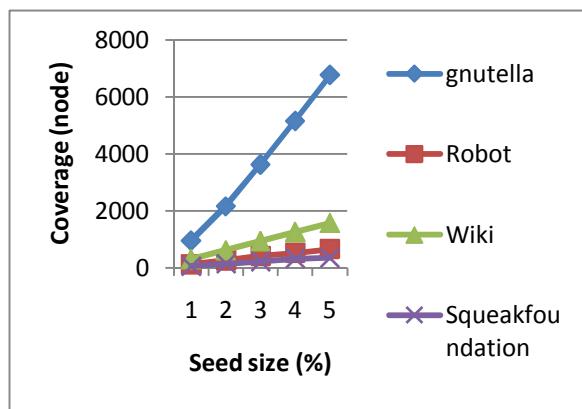
21:**End Algorithm**

۵. آزمایشات

برای شبیه‌سازی از چهار نوع مجموعه داده استاندارد مربوط به شبکه‌های اجتماعی استفاده شده است. اولین مجموعه داده squeakfoundation یک وب سایت شبکه اجتماعی آنلاین است که جزئیات مربوط به این شبکه اجتماعی در آدرس <http://people.squeakfoundation.org>. مجموعه داده‌های مورد استفاده در این شبکه در تاریخ ۱۴ دسامبر سال ۲۰۰۷ به صورت ایستا جمع آوری شده است. مجموعه داده دوم یک شبکه اجتماعی تخصصی مربوط به رباتیک می‌باشد که از طریق سایت <http://robots.net> قابل مشاهده است، که این شبکه در ۱۰ می ۲۰۰۸ جمع آوری شده است. مجموعه داده سوم Wiki می‌باشد که یک سایت دایره المعارف رایگان می‌باشد که توسعه کاربران خاصی



شکل ۶: مقایسه نرخ پوشش انتشار بر حسب درصد نودهای فعال شده اولیه ۵ تا ۲۵ درصد برای LTM



شکل ۷: مقایسه نرخ پوشش انتشار بر حسب درصد نودهای فعال شده اولیه برای MVICA

با توجه به نتایج ارائه شده در جداول ۲ تا ۵ و شکل های ۷ و ۸ برای شبکه های مختلف با ویژگی های مختلفی چون تعداد نود ها، چگالی نودها نتایج متفاوتی بدست آمده است. در نتایج حاصله در مدل پیشنهادی در کلیه حالات به جز زمانیکه از تعداد نودهای اولیه بالای استفاده می شود و شبکه خیلی خلوت است، شاهد نرخ پوشش بالاتری نسبت به مدل استانه خطی می باشیم. چنانچه در جدول شماره ۲ تا ۵ مشاهده می شود با افزایش تعداد نودهای اولیه میزان انتشار به طور کلی نیز افزایش می یابد، در حالیکه برای زمان انتشار در مقایسه با مدل استانه خطی از میزان کمتری برخورد است که از دلایل آن می توان به وجود ساختار ریز پیمانه ای اشاره کرد. در واقع از آنجائیکه انتشار در روش پیشنهادی نسبت به مدل های آستانه خطی، شبیه سازی نزدیکتری به پدیده انتشار در دنیای واقعی دارد، این مورد در مدل پیشنهادی زمان بیشتری را برای انتقال از حالتی به حالت دیگر را صرف خواهد کرد که این مورد در مدل آستانه خطی صرفاً برای ۲ حالت وجود دارد. اختلاف زمان انتشار در مدل انتشار چندگانه بسیار نزدیک به مدل آستانه خطی و قابل چشم پوشی می باشد. با توجه به نتایج بدست آمده ارتباط معکوسی بین دو بین چگالی گراف و زمان انتشار وجود دارد، به طوریکه با کاهش چگالی گراف شاهد افزایش زمان انتشار بوده ایم. همچنین با افزایش تعداد رئوس در شبکه های مختلف شاهد افزایش نرخ انتشار در شبکه نیز

جدول ۳: نتایج مقایسات مدل های انتشار LTM و MVICA برای Robot شبکه

LTM	MVICA	
نرخ پوشش	نرخ پوشش	زمان انتشار
زمان انتشار		تعداد نودهای اولیه
$10.8.5 \pm 6.9$	2.8 ± 0.6	0.6 ± 2.07
269.3 ± 65.3	2.3 ± 0.4	0.4 ± 2.3
351 ± 51.7	3.1 ± 0.7	2.2 ± 0.4
$50.7.9 \pm 58.5$	2.5 ± 0.5	2.4 ± 0.5
621.4 ± 56.4	3.7 ± 0.4	2.5 ± 0.5
		42.7

جدول ۴: نتایج مربوط به انتشار برای شبکه wiki –vote

LTM	MVICA	
نرخ پوشش	نرخ پوشش	زمان انتشار
زمان انتشار		تعداد نودهای اولیه
31.0 ± 2.0	2.1 ± 0.3	0.2 ± 2.07
621.8 ± 3.0	2.17 ± 0.3	0.2 ± 2.07
938.2 ± 14.3	2.27 ± 0.4	2.13 ± 0.3
1254 ± 4.3	2.4 ± 0.5	2.23 ± 0.4
1572.3 ± 8.4	2.47 ± 0.5	2.38 ± 0.4
		153.0

جدول ۵: نتایج مربوط به انتشار برای شبکه Gnutella

LTM	MVICA	
نرخ پوشش	نرخ پوشش	زمان انتشار
زمان انتشار		تعداد نودهای اولیه
979.4 ± 28.6	6.97 ± 2.2	1.3 ± 7.2
$210.8.4 \pm 98.2$	9.97 ± 2.0	1.2 ± 7.8
350.56 ± 145.3	12.7 ± 2.6	8.5 ± 0.9
$66200.0 \pm 90.4.2$	34.57 ± 14.1	8.4 ± 0.8
$830.5.3 \pm 10.1.6$	18.7 ± 3.06	8.1 ± 0.9
		2126

نمودار نودهای فعال شده بر حسب تعداد نودهای اولیه برای شبکه های مختلف به ترتیب در شکل ۷ و ۸ برای نودهای فعال شده تحت مدل انتشار آستانه خطی و مدل انتشار چندگانه پیشنهادی مبتنی بر اتماتای سلولی نامنظم نشان داده شده است.

- [9] D. Stevens, "Integration of an Irregular Cellular Automata Approach and Geographic Information Systems for High-Resolution Modeling of Urban Growth", Msc. Thesis, Department of Geography, University of Toronto, 2003.

هستیم. برای معیارهای دیگری مانند ضریب خوشگی و قطر شبکه نمی‌توان ارتباطی با نحوه انتشار ارائه نمود.

۶: نتیجه گیری

انتشار اطلاعات در شبکه اجتماعی یکی از مسائل مهم در بازاریابی اشاعه ای می‌باشد . ارائه یه مدل نزدیک به واقعیت در این زمینه امری ضروری برای تحلیل شبکه های اجتماعی محسوب می شود . در این مقاله یک مدل انتشار مدل انتشار چند گانه بر پایه اتماماتی سلولی نامنظم ارائه شده است. در مدل پیشنهادی حالتی ای تاثیرپذیری کاربران در هر سلول و قواعدی برای تغییر حالت هر سلول طراحی شده است. آزمایشات متنوعی برای مدل پیشنهادی بر روی شبکه های اجتماعی استاندارد صورت گرفت که نشاندهنده ی بهبود نسبی مدل پیشنهادی بود. در آزمایشات صورت گرفته اهمیت تعداد نودهای اولیه در نتایج حاصله مشهود می‌باشد. هم چنین از ویژگی های مدل پیشنهادی میتوان به نزدیک بودن به دنیای واقعی، مقیاس پذیری و سرعت بالای الگوریتم اشاره کرد. برای مدل پیشنهادی در آینده می‌توان توسعه هایی نیز درباره ماهیت نظرات در نظر گرفت.

۷. مراجع

Domingos Et al¹
Cost Effective Lazy Forward²
Submodular³
Linear threshold modele⁴
Independent Cascade Model⁵
On⁶
Follow⁷
Active⁸

- [1] M.Richardson, P. Domingos, "Mining knowledge-sharing sites for viral marketing", Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, July 23-26, 2002
- [2] D. Kempe, J.Kleinberg , É.Tardos, "Maximizing the spread of influence through a social network", Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, August 24-27, 2003
- [3] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. S. Glance. "Cost-effective outbreak detection in networks". In Proceedings of the 13th ACMSIGKDD Conference on Knowledge Discovery and Data Mining, pages 420–429, 2007.
- [4] W.Chen, C.Wang, Y.Wang," Scalable influence maximization for prevalent viral marketing in large-scale social networks", Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, July 25-28, 2010
- [5] F.Stonedahl, W. Rand, U.Wilensky," Evolving viral marketing strategies", Proceedings of GECCO'10 Proceedings of the 12th annual conference on Genetic and evolutionary
- [6] Y.Gui-sheng, W.Ji-jie, D. Hong bin, L.Jia, , "Intelligent Viral Marketing Algorithm over Online Social Network," Networking and Distributed Computing (ICNDC, pp.319-323, 21-24 Sept. 2011
- [7]L. Gordillo and V. Luna, "Parallel Sort on a Linear Array of Cellular Automata", IEEE Trans. Comput, Vol 2, pp. 1904-1910, 1994.
- [8] H. Couclelis, "Cellular Worlds - a framework for modeling micro-macro dynamics", Environment and Planning, pp. 585-596, 1985.