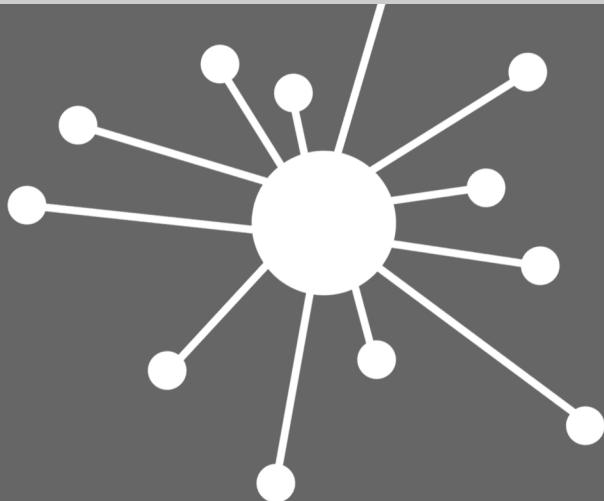


دانشگاه علامه طباطبائی



معلوٰتیانہ

فصلنامه علمی - دانشجویی علوم رایانه، سال اول، شماره اول، بهار ۱۳۹۳

صاحب امتیاز: انجمن علمی دانشجویی علوم رایانه دانشگاه علامه طباطبائی

مدیر مسئول: مجتبی اصلاحی

سردیبیر: مهدی کلهری

صفحه آرایی: محمد حمید عظیمی

ناظر علمی: دکتر محمدرضا اصغری اسکوئی

نویسندها: علیرضا غفاری صدر، سروش عبادی، وحید حیدری پور،

مجتبی اصلاحی، سید سعید حاجی سید جوادی، امیرحسام ملتجي

پست الکترونیکی: css@atu.ac.ir



طراح لوگو: امیر علیزاده طرح جلد: محمد حمید عظیمی - مجتبی اصلاحی

فهرست:

سخن مدیر مسئول ب
سخن دبیر انجمن علمی علوم رایانه پ
سخن سردیبیر ت
تکنیک های داده کاوی و کاربردها: مرور یک دهه از سال ۲۰۰۰ تا سال ۲۰۱۱ ۱
مقایسه عملکرد الگوریتم های ژنتیک، تبرید شبیه سازی شده و پسگرد در حل مسئله چند وزیر ۱۰
مدل سازی انتشار اطلاعات و عضویت در اجتماع با استفاده از بهینه سازی احتمالی ۱۶
شبکه های اجتماعی : تعاریف، ابزارها و روش های تحلیل ۲۴
سخن اساتید - دکتر فرشته آزادی پرند: علوم کامپیوتر و تئوری بازی ۲۷
آشنائی با گروه رایانه دانشگاه علامه طباطبائی ۲۸
همایش ها و کنفرانس های علوم رایانه ۳۰
۳۸ Procedural Modeling and User Interaction

سخن مدیر مسئول

خداآوند منان را شاکریم که این توان و این فرصت را به ما ارزانی داشت تا بتوانیم با به ثمر رساندن اولین شماره از نشریه علمی پژوهشی «علوم رایانه» اولین قدم را در جهت حرکتی علمی پژوهشی در حوزه علوم رایانه در دانشگاه علامه طباطبایی برداریم. هر چند گفته می شود اولین قدم معمولاً سخت ترین مرحله است اما به خوبی به این نکته آگاهیم که زمانی که در مورد یک حوزه علمی پژوهشی صحبت می کنیم تداوم حرکت از اهمیت فراوان برخوردار است.

به طور عام برای تداوم یافتن یک حرکت علمی ما نیازمند «وجود مسئله» و «احساس نیاز به حل آن» هستیم. در شرایطی که ما قرار داریم هر دو مورد در حوزه های مختلف به فراوانی به چشم می خورد. حال برای مواجه با این مسائل ابتدا باید مسائل را تعریف کنیم. «تعریف مسئله» این پاشنه‌ی آشیل بسیاری از کارهای علمی پژوهشی کشورمان را نباید دست کم گرفت. شاید در حوزه های علوم پزشکی ما کمتر با این مشکل مواجه باشیم اما در حوزه های علوم فنی و مهندسی این مشکل جدیست. بسیاری از دانشجویان مقاطع تحصیلات تکمیلی ما نمی دانند مسئله چیست که بخواهند آن را حل کنند! چه بسیارند دانشجویانی که نیمی از دوره مقطع تحصیلات تکمیلی خود را گذرانده اند اما هنوز نمی توانند حوزه ای را بیابند که کسی در آن مسئله ای تعریف کرده باشد و آن ها بتوانند به حل آن اقدام کنند. این موضوع نه فقط آزاردهنده نظام آموزشی ما که آزاردهنده دانشجویان نیز هست. آن ها که با هزاران امید و آرزو پا در این مسیر گذاشته اند اغلب نامید و ناخشنود از نتیجه دوره خود را به پایان می گذرانند.

آیا تعریف مسئله بر عهده دانشجوست؟ قطعاً خیر. دانشجو در اصل وجود مسئله و همچنین تعریف مسئله بی تقصیر و به گمانم بدون دید جامع و گستردگی و تجربه شده است. او اما توان، پتانسیل و تمایل خود را برای حل مسئله نشان داده و اتفاقاً می خواهد که از این ظرفیت ها در جهت حل مسائل استفاده شود. فراموش نکنیم برای یک دانشجو اولین چیزی که مدنظر است این است که چه مسیری زندگی ابتدایی او را تامین خواهد کرد. او نمی تواند و تمایلی هم ندارد که خلاف جهت اقتصادی و اجتماعی جامعه اش حرکت کند. اگر حل مسائل داخل کشور با سیاست های پژوهشی هم سو نباشد او به خارج از کشور تمایل پیدا می کند و اگر در خارج از کشور هم این فضا فراهم نمی بود شاید قید دانشجو بودن را می زد. صرف مدرک گرفتن در حال حاضر یک حرکت سود ده است. پس دانشجویان ما زیاد خواهند شد و این طبیعی است. حل مسائل واقعی جامعه هم باید سود ده باشد تا دانشجو به آن تمایل و گرایش داشته باشد و به آن سمت حرکت کند. سود ده بودن هم نه به این معنا که پول فراوانی در اختیار دانشجو قرار گیرد. حداقل این است که او احساس نکند دارد حرص و جوش می خورد، زندگی سخت دانشجویی را تحمل می کند و ضرر هم می کند.

در حوزه علوم رایانه اگر بنا باشد حل مسائل داخل کشور تمایل دانشجویان را به خود جلب نکند - که متسافانه در چنین وضعیتی قرار داریم - آن گاه دانشجوی ما با یک بحران بزرگ تر نیز روبرو می شود. حالا نه فقط کسی برای او مسئله تعریف نمی کند بلکه فاصله بسیار زیاد علمی باعث می شود او نیازمند آن باشد که مطالعات گستردگی ای را در پیش بگیرد تا بتواند با جنس مسائل علوم رایانه در کشورهای پیشرفته دنیا آشنا شود. سپس باید مشابه مسائل آن ها را برای خود تعریف کند تا تازه بتواند به حل آن ها اقدام کند. بیچاره دانشجویان!

کاری که ما می توانیم انجام دهیم آن است که با گسترش حرکت های جمعی چون انتشار نشریات، برگزاری همایش ها و کارهایی از این دست فرآیند شناخت حوزه های تحقیقی، شناخت مسئله و تعریف آن را برای خود و دیگران هموارتر کند. امیدواریم حرکت هایی این چنین بتواند زمینه ساز تحقق موارد مورد اشاره باشد.

اولین شماره از یک نشریه حتماً دارای کم و کاستی هایی خواهد بود. حتماً موارد مدنظر خود را با ما از طریق ایمیل css@atu.ac.ir در میان بگذارید. کیفیت مطالب ارائه شده در نشریه به عنوان مهم ترین هدف نشریه در شماره های آتی مدنظر گردانندگان آن خواهد بود. در این مورد نگرانی هایی وجود دارد اما امیدواریم دانشجویان و اعضای هیئت علمی با همکاری های خود بتوانند باعث تحقق هدف ما یعنی ارائه مطالب خوب و سازنده باشند.

مجتبی اصلاحی
اردیبهشت ۱۳۹۳

سخن دبیر انجمن علمی علوم رایانه

خرد هر کجا گنجی آرد پدید

زنام خدا سازد آن را کلید

خداآوند منان را شاکریم که بار دیگر توفیقی عطا نمود تا در جهت پیشرفت علم، گام کوچکی برداریم. انجمن علمی علوم رایانه دانشگاه علامه طباطبائی فعالیت خود را به طور رسمی در آبان ماه ۱۳۹۲، با هدف رشد و ارتقاء هرچه بیشتر دانش فردی و گروهی دانشجویان، آغاز نموده است. این فعالیت‌ها شامل شرکت در همایش‌ها، کنفرانس‌ها و مسابقات دانشجوئی مرتبط در زمینه علوم رایانه، برگزاری کلاس‌های آموزشی مفید برای تمامی دانشجویان و همچنین چاپ مقالات دانشجوئی می‌باشد.

نشریه حاضر، نخستین چاپ خود را تجربه می‌کند که با همت جمعی از دانشجویان کارشناسی ارشد علوم کامپیوتر دانشگاه علامه طباطبائی به انجام رسیده و محورهای اصلی آن به شرح زیر است:

- الف) معرفی رشته علوم کامپیوتر و گرایش‌های آن در دانشگاه علامه طباطبائی.
- ب) معرفی استادی رشته علوم کامپیوتر دانشگاه علامه طباطبائی.
- ج) مقالات برگسته‌ای در زمینه کامپیوتر و همچنین فراخوان عمومی برای چاپ آتی.
- د) برنامه‌های پیش رو برای شرکت در همایش‌ها و کنفرانس‌های مرتبط.

در اینجا لازم می‌دانیم تا از استاد مشاور انجمن علمی علوم رایانه، جناب آقای دکتر اسکوئی به واسطه راهنمایی‌های ارزشمندانه، قدردانی نمائیم. در پایان نیز از تمامی استادی و دانشجویان محترم خواهشمندیم تا جهت بهبود کیفی هرچه بیشتر این نشریه، ما را از نظرات و پیشنهادات گرانبهای خود آگاه سازند.

علیرضا غفاری صدر

دبیر انجمن علمی علوم رایانه دانشگاه علامه طباطبائی

css@atu.ac.ir

سخن سردبیر

به نام آنکه جان را فکرت آموخت چراغ دل به نور جان برافروخت

با استعانت از درگاه ایزد دانا، اولین شماره نشریه علمی دانشجویان در انجمن علمی علوم رایانه دانشگاه علامه طباطبایی است، را منتشر می کنیم.

در این نشریه که بصورت فصلنامه منتشر می شود، برآنیم تا با رویکردی علمی به مباحث حوزه علوم رایانه پردازیم. لذا از تمامی اعضای هیات علمی و دانشجویان رشته های مرتبط با علوم رایانه دعوت می کنیم تا با ارسال مقالات و تابیج پژوهش های خود، به غنای علمی این مجله بیافزایند و ما را در طی این مسیر که هنوز در ابتدای آن هستیم، یاری رسانند. بی شک ادامه این راه، تنها با مشارکت و همیاری شما عزیزان میسر می باشد.

سعی ما در این نشریه معرفی موضوعات مطرح روز در زمینه علوم رایانه و ارائه پژوهش های انجام گرفته بر روی آنها می باشد. جهت گیری اصلی ما حرکت به سمت مزد های دانش در این حوزه می باشد، اما می دانیم که این امر به آسانی اتفاق نخواهد افتاد و نیازمند تلاش فراوان و استفاده از تمامی ظرفیت های موجود می باشد. همچنین در این نشریه سعی می کنیم مخاطبان را در جریان اتفاقات علمی مرتبط از جمله کنفرانس ها و همایش ها قرار بدهیم.

امید است که فصلنامه علوم رایانه بتواند جایگاه مناسبی را در بین علاقمندان این حوزه بدست آورد و رضایت خاطر آنها را فراهم آورد. از کلیه عزیزان و صاحب نظران دعوت می کنیم تا با پیشنهادات سازنده خود ما را در این مهم یاری برسانند.

مهدی کلهری

اردیبهشت ۹۳

تکنیک های داده کاوی و کاربردها: مرور یک دهه از سال ۲۰۰۰ تا سال ۲۰۱۱

Data mining techniques and applications – A decade review from 2000 to 2011

Shu-Hsien Liao, Pei-Hui Chu, Pei-Yuan Hsiao

Expert Systems with Applications

ترجمه: مجتبی اصلاحی، دانشجوی کارشناسی ارشد علوم کامپیوتر، دانشگاه علامه طباطبایی

کلیدواژه:

داده کاوی
تکنیک های داده کاوی
کاربردهای داده کاوی
ارزیابی مقاطعات

به منظور تبیین چگونگی توسعه تکنیک های داده کاوی و کاربردهای آن ها در طول دهه گذشته، این مقاله به بررسی تکنیک های داده کاوی و کاربردهای آن ها و همچنین نحوه توسعه آن ها از طریق ارزیابی مقاطعات و دسته بندی آن ها از سال ۲۰۰۰ تا سال ۲۰۱۱ می پردازد. برای این منظور از کلمات کلیدی و چکیده مقالات برای شناسایی ۲۱۶ مقاله مرتب با کاربردهای داده کاوی از ۱۵۹ نشریه معتبر (به دست آمده از ۵ دیتابیس آنلاین) استفاده شده است. این مقاله به ارزیابی و دسته بندی تکنیک های داده کاوی با توجه به ۳ حوزه می پردازد: نوع داشن، نوع تحلیل و نوع معماری به ضمیمه کاربردهای آن ها در حوزه های مختلف تحقیقاتی و عملیاتی است. به علاوه بحثی در مورد مسیر توسعه های آتی حوزه روش شناسی و کاربرد تکنیک های داده کاوی مطرح شده است: (۱) کاربردهای داده کاوی در حال رشد به سمت تخصصی تر شدن پیش می رود و توسعه کاربردهای آن یک حوزه مسئله محور است. (۲) پیشنهاد شده است که روش شناسی های علوم اجتماعی از جمله روانشناسی، علوم شناختی و رفتارشناسی انسان می توانند به جای روش های موجود، براساس تکنیک های داده کاوی توسعه و پیاده سازی شوند. (۳) توانایی تعییر مدام و به دست آوردن درک تازه نقش نیروی پیش برنده تکنیک های داده کاوی را بازی می کنند و اجازه کاربردهای جدید را در آینده به آن می دهند.

۱. مقدمه

مقالات از سال ۲۰۰۰ تا سال ۲۰۱۱ و دسته بندی آن ها می پردازد. کاربردهای گوناگون تکنیک های داده کاوی در این مدت مورد بررسی قرار گرفته است. این دوره از آن جهت حائز اهمیت است که اینترنت این سال ها در اختیار عموم قرار گرفته بود و دسترس پذیری گستردگی اطلاعات و تکنولوژی های ارتباطات نقش مهمی را نه تنها در تکنیک های داده کاوی بلکه در توسعه روش شناسی های جمع آوری و نگهداری داده از پایگاه های داده آنلайн بازی می کرد.

دوره مورد بررسی این مقاله از ژانویه ۲۰۰۰ آغاز شده است. در آگوست ۲۰۱۱ جستجویی در پایگاه های آنلайн Elsevier SCOPUS، Springerlink، IEEE Wiley InterScience (Xplore، EBSCO (Electronic Journal Device چکیده مقاطعات شامل عبارت «تکنیک داده کاوی» انجام شد. برای دوره ۲۰۰۰ تا ۲۰۱۱ تعداد ۲۰۱۱ مقاله یافت شد. با ترکیب نتایج گذشته و محدودیت اعمالی دیگری به صورت «کاربردهای داده کاوی»^۱ این تعداد به ۲۱۶ عدد از ۱۵۹ نشریه کاهاش یافت. این مقاله با استفاده از این ۲۱۶ مقاله در مورد کاربردهای داده کاوی، به ارزیابی و دسته بندی تکنیک های داده کاوی در ۹ حوزه می پردازد: شبکه های عصبی، معماري الگوريتم، رویکرد مبتنی بر پیش بینی پویا^۲، تحلیل معماري سیستم ها، سیستم های کارگزار هوشمند^۳، مدل سازی، سیستم های مبتنی بر داشن، بهینه سازی سیستم و سیستم های اطلاعاتی به علاوه کاربردهای آن ها در حوزه های تحقیقاتی و عملیاتی مختلف.

تکنیک های داده کاوی یک شاخه از هوش مصنوعی کاربردی را از سال ۱۹۶۰ شکل داده است. در طول دهه های حضور آن، نوآوری های مهم سیستم های کامپیوتری باعث معرفی تکنولوژی های جدید برای آموزش مبتنی بر وب (Ha, Bae, & Park, ۲۰۰۰) شده است. داده کاوی اجازه جستجو اطلاعات ارزشمند را در حجم بالایی از داده ها فراهم می کند (& Weiss, ۱۹۹۸). رشد انفجاری پایگاه های داده نیاز به توسعه تکنولوژی هایی برای استفاده از اطلاعات و داشن را به وجود آورده است. بنابراین تکنیک های داده کاوی تبدیل به یک حوزه تحقیقاتی با درجه اهمیت رو به

رشد گردیده است (Fayyad, Djorgovski, & Weir, ۱۹۹۶) از تکنیک های داده که به تازگی توسعه یافته است، چند روش مهم داده کاوی شامل عمومیت یابی^۴، توصیف صفات اختصاصی^۵، دسته بندی، خوش بندی، وابستگی^۶، تکامل، تطابق الگو، نمایش داده، کاوش براساس فرآقوانین^۷ در اینجا بررسی شده است. همچنین تکنیک های کاوش داشن از انواع گوناگون پایگاه های داده شامل پایگاه های داده رابطه ای، تراکنشی^۸، شی گر، پایگاه های داده سه بعدی و فعل^۹ و همچنین سیستم های اطلاعاتی سراسری^{۱۰} مورد بررسی قرار گرفته است.

به عنوان زیر مجموعه ای از تحقیقات حوزه تکنیک های داده کاوی، این مقاله به ارزیابی توسعه این تکنیک ها در طی یک بررسی

1) generalization

6) Spatial and active databases

2) Characterization

7) global information systems

3) Association

8) Data mining technique

4) meta-rule guided mining

9) Data Mining Applications

5) Transactional

10) Dynamic prediction-based approach

11) Intelligent Agent System

چند کاربرد که توسط الگوریتم ها پیاده سازی شده است عبارتند از: تشخیص تعامل خودکار مربع خی، مدل ها و الگوریتم ها، GRASP، پردازش تحلیلی آنالین^{۱۸}، K-means، الگوریتم های خوش بندی، الگوریتم های درختان تصمیم، درختان دسته بندی و رگرسیون، فاصله اقلیدسی، الگوریتم های خوش بندی بسته شده، منطق فازی، قوانین پیوستگی، C&RT، CRISP-DM، غیرطبیعی، الگوریتم های خوش بندی ژنتیک، مدل های تحریک غدد و ماسشین بردار پشتیبانی، معماری های الگوریتم و کاربردهای آن ها در جدول ۳ آورده شده است.

دیگر بخش های این مقاله به این ترتیب است: بخش های ۳ الی ۱۱ نتایج ارزیابی روش شناسی تکنیک های داده کاوی و کاربردهای آن ها براساس دسته بندی های نام برده شده در بالا ارائه می دهد. بخش ۱۲ پیشنهاداتی را برای توسعه های آتی روش شناسی تکنیک های داده کاوی و کاربردهای آن ها ارائه می دهد. درنهایت بخش ۱۳ شامل یک نتیجه گیری کوتاه است.

۲. سیر حرکت تکنیک های داده کاوی

داده های جمع آوری شده شامل مقالات از ۲۰۰۰ تا ۲۰۱۱ هستند. سیر حرکتی کلمات کلیدی مقالات تحقیقاتی: داده کاوی درخت تصمیم، شبکه عصبی مصنوعی، خوش بندی، قوانین وابستگی^{۱۹}، هوش مصنوعی، بیوانفورماتیک، ارتباط مشتری، منطق فازی و کاربردهای آن ها. این موارد در جدول ۱ به نمایش درآمده است.

Table 1
2000-2011 DMT keywords trends.

Keyword	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Total
Data mining	1	0	6	5	8	8	12	16	10	16	16	16	114
Decision tree	0	0	1	0	0	2	0	1	2	4	3	4	17
Artificial neural network	1	1	2	1	2	2	2	0	2	4	2	3	22
Clustering	0	0	1	0	0	3	0	1	0	2	1	1	9
Association rule	0	0	0	1	0	0	0	2	1	3	0	1	8
Artificial intelligence	0	0	0	0	0	0	0	1	0	1	2	1	6
Bioinformatics	0	0	0	0	0	0	0	3	0	1	0	0	4
Customer relationship management	0	0	0	1	0	0	0	0	0	0	3	0	4
Fuzzy logic	0	0	0	1	1	0	0	1	0	1	0	0	4
Total	2	1	10	9	11	15	15	24	16	32	27	26	188

جدول ۱

Table 2
Neural networks and their applications.

Neural networks applications	Authors
Feed forward neural networks	Trafalis et al. (2002)
Radial basis function networks	Srivastava et al. (2005)
Neural classification	Dutta et al. (2004)
Bayesian confidence propagation neural networks	Cesana et al. (2007)
Gene regulatory networks	Ma and Chan (2007)
Fuzzy recurrent neural networks	Aliev et al. (2008)
Neural nets	Tsai and Chen (2008)
Back-propagation artificial neural networks	Lin et al. (2011)
Bayesian networks	Rivas et al. (2011)
General regression neural networks	Tu et al. (2011)
Flow networks	Zhang and Ramirez-Marquez (2011)

جدول ۲

Table 3
Algorithm architecture and its applications.

Algorithm applications	Authors
Gap statistic algorithm	Huang and Lin (2002)
Chi-square automated interaction detection	Rygielski et al. (2002)
Models and algorithms	Lancashire et al. (2005)
GRASP	Ribeiro et al. (2006)
OLAP	Singhal and Jajodia (2006)
K-Means	Adderley et al. (2007)
Clustering algorithms	Balzano and Del Sorbo (2007)
Decision forest algorithms	Hsia et al. (2008)
Classification and regression tree	Qiang et al. (2008)
Euclidean distance	Fan et al. (2009)
Bagged clustering algorithms	Huang et al. (2009)
Fuzzy logic	Jiménez et al. (2009)
Association rule	Shih et al. (2009)
C&RT	Ture et al. (2009)
Apriori algorithms	Chen and Bai (2010)
C5	Marx (2010)
Anomaly-based IDS	Nikolova and Jecheva (2010)
Clustering	Piquer et al. (2010)
Genetic algorithms	Ahn et al. (2011)
CRISP-DM model	Chen and Huang (2011)
Thyroid stimulating hormone	Henderson and Grey (2011)
SVM	Ravishankar et al. (2011)

جدول ۳

۳. شبکه های عصبی و کاربردها آن ها

مرسم بوده است که عبارت «شبکه های عصبی» به شبکه یا حلقه ای از نورون های زیستی اشاره کند. کاربرد جدید این عبارت اغلب به شبکه های عصبی مصنوعی اشاره دارد که شامل نورون ها یا نودهای مصنوعی است. مشابه پیام دهی^{۲۰} الکترونیکی، شبکه های دیگر پیام دهی از پراکنده گی فرستنده های عصبی نشات می گیرد که بر پیام دهی الکترونیکی تاثیر می گذارد. به همین ترتیب، شبکه های عصبی دارای پیچیدگی های بسیار زیادی هستند.

چند کاربرد شبکه های عصبی عبارتند از شبکه های مبتنی برتابع پایه شعاعی^{۲۱}، دسته بندی عصبی، شبکه های عصبی انتشار اطمینان بیزین^{۲۲}، شبکه های نظم دهنده^{۲۳}، شبکه های عصبی بازگشت کننده فازی، شبکه های عصبی، شبکه های عصبی مصنوعی پس انتشار خطأ، شبکه های بیزین، شبکه های عصبی رگرسیون عمومی و شبکه های جریان. تکنیک های شبکه عصبی و کاربردهای آن در جدول ۲و قابل مشاهده است.

۴. معماری الگوریتم و کاربردها

معماری الگوریتم به صورت تعداد محدودی دستور خوش تعریف برای محاسبه یکتابع توصیف شده است. الگوریتم ها برای محاسبه، پردازش داده و استدلال خودکار (automated reasoning) کاربرد دارد. به طور ساده الگوریتم یک فرآیند مرحله به مرحله برای محاسبه است. ارائه یک تعریف رسمی از این مفهوم با تلاش ها برای حل مسئله توقف^{۲۴} که توسط David Hilbert در سال ۱۹۲۸ مطرح شد، آغاز گردید.

12) Association rule

13) Signaling

14) radial basis function networks

15) Bayesian confidence propagation neural networks

16) Gene regulatory networks

17) Entscheidungs problem

18) OLAP

درختان مدل M5، تحلیل فاکتور، تحلیل سبد خرید، پالایش گروهی، تحلیل داده، مدل های مبتنی بر درخت تصمیم، تحلیل مولقه اصلی، گزینش چند مشخصه ای، تشخیص نفوذ و تحلیل دیالیز خون. این معماری های سیستم و کاربردهای آن ها در جدول پنج ذکر گردیده اند.

Table 5
Analysis of systems architecture and its applications.

Analysis of systems architecture applications	Authors
Correlation analysis	Lee et al. (2000)
Semantic analysis	Sui and Meng (2001)
Regression analysis	Li et al. (2002)
Statistical analysis	Lu et al. (2005)
Discriminative analysis	Adachi et al. (2006)
Association analysis	Al-Hamami et al. (2006)
Penalized discriminative analysis	Granitto et al. (2007)
Process parameter analysis	Martínez-de-Pisón et al. (2007)
Cluster analysis	Hoontakul and Sahadev (2008)
Decision making	Ranjan et al. (2008)
Decision support systems	Bára et al. (2009)
Consumer behavior analysis	Hsieh and Chu (2009)
Binary logistic regression analyses	Jilani et al. (2009)
M5 model trees	Küçükşille et al. (2009)
Factor analysis	Parhizi et al. (2009)
Market basket analysis	Shahrobi and Neyestani (2009)
Collaborative filtering	Kim et al. (2010)
Data analysis	Miranda et al. (2010)
Decision tree based models	Bae and Kim (2011)
Principal component analysis	Bhramaramba et al. (2011)
Multi-feature selection	Chang et al. (2011)
Intrusion detection	Thiruvadi and Patel (2011)
Hem dialysis	Yeh et al. (2011)

جدول ۵

۷. سیستم های کارگزار هوشمند

در حوزه هوش مصنوعی، سیستم های کارگزار هوشمند یک موجودیت مستقل است که در یک محیط مشاهده می کند و عکس العمل نشان می دهد. عامل های هوشمند همچنین می توانند یاد بگیرند یا از داشن استفاده کنند تا به هدف خود برسند. آن ها می توانند بسیار ساده یا بسیار پیچیده باشند. یک ترمومترات یک عامل هوشمند است، همان طور که یک انسان یک عامل هوشمند است و همچنین گروهی از انسان ها که برای رسیدن به یک هدف با یکدیگر همکاری می کنند Russell and Norvig رفتار بر مبنای هدف را به عنوان جوهر و ماهیت هوش معرفی کرده اند و بنابراین از عبارت عامل منطقی^{۱۹} برگفته شده از علوم اقتصادی برای این منظور استفاده کرده اند.

چند کاربرد برای سیستم های کارگزار هوشمند عبارتند از: سیستم های چند عاملی، سیستم های پیچیده، طراحی رابط کامپیوتر، سیستم های پایگاه داده چند کاربره، تحلیل هوشمند، هوش صنعتی، سیستم های تدریس خصوصی هوشمند، ماشین های بردار پشتیبانی، سیستم های تشخیص برنامه، سیستم های ناظرتی و تخصصی، هوش محاسباتی، هوش مصنوعی و سیستم های تشخیص و پشتیبانی Mahalanobis Taguchi هوشمند و کاربردهای آن ها در جدول شش آمده است.

۵. رویکردهای مبتنی بر پیش بینی پویا و کاربردهای آن
رویکردهای مبتنی بر پیش بینی پویا یک مدل ریاضی برای دینامیک های تصادفی است که برای مدل کردن مولکول ها استفاده می شود و البته کاربردهایی در بازار بورس نیز دارد. مهم ترین ویژگی دینامیک لائزون^{۲۰} وجود یک نویز تصادفی گوسین است. اصول اولیه مکان یابی موقعت^{۲۱} به طور گسترده توسط Peter در سال ۱۹۶۸ مورد بررسی و کاربرد قرار گرفت. این مطالعات نشان می دهد که یک ریسمان^{۲۲}، دسترسی به یک مجموعه کاملاً تصادفی از آدرس ها را در هنگام تراکنش دارا نیست.

چند کاربرد رویکرد مبتنی بر پیش بینی پویا شناسایی تومورهای چشمی، تشخیص خطای وسایل نقلیه، رایانش شبکی، موج های دودوبی، واکنش اولیه، مدل های پیش بینی ترمیم خط، مدل های پیش بینی های بازفعالی شیمیایی، ردیابی وسایل نقلیه بلاذرگ، پیش بینی، تشخیص وضیعت نامطلوب، پیش بینی سرهم بندی، ژنتیک مقایسه ای، پیش بینی های بالینی، و مدل های پیش بینی کننده. جدول چهار تحقیقات حوزه پیش بینی را به نمایش گذاشته است.

Table 4
Dynamic prediction based approach and its applications.

Dynamic prediction based applications	Authors
Ophthalmic oncology	Jegelevičius et al. (2002)
Vehicle fault diagnosis	Fong et al. (2003)
Grid computing	Sánchez et al. (2004)
Dyadic wavelet	Yu et al. (2006)
Pre-fetching	Jain (2007)
Fault restoration prediction model	Hwang et al. (2008)
Fault prediction model	Bae et al. (2009)
Financial distress prediction model	Chen and Du (2009)
Vlasov–Maxwell equations	Assous and Chaskalovic (2010)
Chemical reactivity prediction	Borghini et al. (2010)
Real time vehicle tracking	Constantinescu et al. (2010)
Forecast	He et al. (2010)
Anomaly detection	Mahesh et al. (2010)
Churn prediction	Tsai and Lu (2010)
Comparative genomics	Bhramaramba et al. (2011)
Clinical predictions	Gregori et al. (2011)
Predictive model	Li et al. (2011)

جدول ۴

۶. تحلیل معماری سیستم و کاربردهای آن

تحلیل معماری سیستم از یک مدل مفهومی برای تعریف ساختار، رفتار و دیگر وجوه یک سیستم استفاده می کند. معماری سیستم از هر دو عنصر نرم افزار و سخت افزار استفاده می کند که اجازه طراحی سیستم های ترکیبی را می دهد. یک معماری خوب می تواند به صورت یک دیاگرام تقسیم کننده یا یک الگوریتم تصور شود که به طور کامل نیازمندی های فعلی و قابل پیش بینی در آینده سیستم را به مجموعه ای از زیرسیستم های کران دار کاربردی تقسیم بندی کند.

چند کاربرد تحلیل معماری سیستم عبارتند از: تحلیل معنایی، تحلیل رگرسیون، تحلیل آماری، تحلیل تفکیک، تحلیل وابستگی، تحلیل تفکیک دارای جریمه، تحلیل پارامتر فرآیند، تحلیل خوش، تصمیم گیری، سیستم های پشتیبانی تصمیم، تحلیل رفتار مصرف کننده، تحلیل رگرسیون لاجیستیکی دودوبی ،

19) Langevin

20) temporal locality

21) thread

22) rational agent

۹. سیستم های متنی، بی دانش، و کاربردهای آن ها

سیستم های مبتنی بر دانش ابزارهای هوش مصنوعی هستند که در یک حوزه محدود کار می کنند و تصمیم های هوشمند فراهم می آورند. معمول ترین تعریف سیستم های مبتنی بر دانش با محوریت انسان است به این دلیل که ریشه های چنین سیستم هایی درون حوزه های هوش مصنوعی قرار دارد. این سیستم ها نشانگر تلاش ها برای شناخت دانش های انسانی و وارد کردن آن ها در سیستمهای کامپیوتی، است (Wiig, 1994).

دانش توسط تکنیک ها، قوانین و چارچوب های گوناگون ارائه داشت، به دست آمده و نمایش داده می شود. مزیت های اولیه ای که توسط این سیستم ها ارائه می شود، مستندسازی دانش، پشتیبانی تصمیم هوشمند، خودآموزی، استدلال و توصیف است. Akerkar و Sajja در سال ۲۰۰۹ شرح دادند که سیستم های مبتنی بر دانش سیستم هایی هستند که از تکنیک ها و روش های هوش مصنوعی استفاده م کنند.

چند کاربرد سیستم های مبتنی بر دانش عبارتند از: تکنیک های یادگیری، تکنیک های کنترل خودکار، اکتشاف دانش در پایگاه های داده، حلقه های دانش، تکنیک های ارتباطی، سنجش دانش، استخراج دانش، فرآگیری دانش، مدیریت دانش، نمایش دانش، کتابخانه های دیجیتال، داده کاوی مبتنی بر تئوری کسب اطلاعات. کاربردهای سیستم های مبتنی بر دانش در جدول هشت نمایش داده شده است.

Table 6
Intelligence agent systems and their applications.

Intelligence agent systems applications	Authors
Inductive learning systems	Kaichang et al. (2000)
Multi-agent systems	Symeonidis et al. (2003)
Complex systems	Trafalis and White (2003)
Computer interface design	Ye et al. (2003)
Multiuser database systems	Feng and Lu (2004)
Intelligent analysis	Mamčenko and Kulvietiene (2005)
Manufacturing intelligence	Hsieh (2007)
Intelligent tutoring systems	Chen and Chen (2009)
Support vector machines	Mucherino et al. (2009)
Program diagnostics systems	Riquelme et al. (2009)
Supervisory and specialist systems	De Andrade et al. (2010)
Computing intelligence	Chou et al. (2011)
Artificial intelligence	Fiol-Roig and Miró-Julià (2011)
Mahalanobis Taguchi system	Su et al. (2005)

جدول ٤

۸. مدل سازی و کاربردهای آن

مدل سازی در مهندسی نرم افزار فرآیند ساخت یک مدل داده به وسیله ترکیب توصیف های مربوط به مدل های داده قراردادی با استفاده از تکنیک های مدل سازی داده است. تکنولوژی مدل سازی می تواند روش های کیفی برای تحلیل داده به منظور نشان دادن یا به دست آوردن دانش خبره را فراهم آورد که بنابراین هوش مصنوعی، علوم شناختی و دیگر حوزه های تحقیقاتی سکه های، وسیع تری، باء، تبعه تکنیک های، داده کاهه، خواهند داشت.

کاربردهای مدل سازی عبارتند از: مدل سازی هزینه، تشخیص مبتنی بر مدل، مدل سازی تکثیر آتش جنگل، مدل سازی خروجی های آماری، مدل سازی لهجه (نحوه تلفظ)، مدل سازی استناد XML، مدل سازی خطرهای نسیی Cox، چند جمله ای ها، موج های دارای اشکال مشابه، سیستم تصمیمی گیری چند خصوصیتی^{۳۳}، کنترل عددی کامپیوتر، فرا آموزش، به کارگیری دارو، کاربردهای تکنیک های مدل سازی در جدول ۷ نمایش داده شده است.

Table 7
Modeling and its applications.

Modeling applications	Authors
Language modeling	Chen (2004)
Cost modeling	Popovic (2004)
Model-based diagnosis	Saitta et al. (2005)
Forest fire spreading modeling	Xiao et al. (2006)
Model output statistics	Besse et al. (2007)
Intonation modeling	Escudero-Mancebo and Cardeñoso-Payo (2007)
XML document modeling	Mei and Zhang (2007)
Cox proportional hazard modeling	Pelletier and Diderrick (2007)
Load Damage Exponents	Chen et al. (2008)
Polynomials	Rezania et al. (2008)
Similar waveforms	Vega et al. (2008)
Simple additive weight	Laosiritaworn and Holimchayachotikul (2010a).
Computer numerical control	Laosiritaworn and Holimchayachotikul (2010b)
Meta learning	Radosavljevic et al. (2010)
Drug utilization	Dakheel et al. (2011)

جدول ٧

۰۱. بهینه سازی سیستم و کاربردهای آن

فرما و لاگرانژ برای شناسایی جواب بهینه از فرمول های حساب دیفرانسیل و انتگرال استفاده کردند در حالی که نیوتون و گاؤس روش های تکرارشونده را به کار برداشتند. به صورت تاریخی عبارت بهینه سازی برای برنامه ریزی خطی به کار می رفته است که توسط Dantzig معرفی شد هر چند بخش های

دیگری از نظری موسسه Kantorovich را سل ۱۹۴۷ معرفی کرد.^{۴۴}

Dantzig simplex را در سال ۱۹۴۷ منتشر کرد و تئوری دوگانه را نیز در همان سال توسعه داد. بهینه سازی سیستم به انتخاب بهترین گزینه از بین چندین مجموعه از جواب های ممکن می پردازد. در ساده ترین حالت، زمانی که حداقل کردن مقدار یکتابع مدنظر است، به طور منظم مقداری طبیعی یا حقیقی از یک مجموعه مجاز انتخاب می گردد.

23) weighted linear combination ↴ simple additive weight

- 23) weight
- 24) Duality

25)

Table 10
Information systems and their applications.

Information systems/applications	Authors
Patient characteristics	Mylod and Kaldenberg (2000)
Catchment characteristics	Hall et al. (2002)
Mobile databases	Saygin and Ulusoy (2002)
Self-organizing feature map	Chen and Mynett (2003)
Insurance claims database	Chen et al. (2003)
Alternating current field measurement	Kang et al. (2004)
Fracture acidizing	Kang et al. (2004)
Latest time sub-series	Wang and Zhang (2006)
Destination choice	Wong et al. (2006)
Attribute relevance study	Bagui et al. (2007)
Fraudulent financial statements	Kirkos et al. (2007)
Sequence similarity	Zhu and Xiong (2007)
Case-based reasoning	Chun et al. (2008)
Anthropometric data	Lin et al. (2008)
Regression splines	Paul (2008)
Economic imbalances	Andronie and Andronie (2009)
Medium-voltage customer fault	Bae et al. (2009)
Maintenance and engineering	Cruz et al. (2009)
Bank lending	Ince and Aktan (2009)
Reinforcement learning	Kheradmandian and Rahmati (2009)
Supervised learning	Oh et al. (2009)
Arousal	Shmiel et al. (2009)
Information visualization	Neto et al. (2010)
Customer retention	Ranjan and Bhatnagar (2010)
Chum management	Su et al. (2010)
Pattern discovery	Tremblay et al. (2010)
Customer relationship management	Wang et al. (2010)
Uniaxial compressive strength	Tinoco et al. (2011)

جدول ۱۰

حوزه داده کاوی مجموعه داده ها برای تطابق الگوهای آتی نگهداری می شد. این مقاله نشان می دهد که کاربردها و توسعه تکنیک های داده کاوی بسیار گسترده است و در کنار آن محققان با پیش زمینه ها، تخصص ها و عالیق مختلف بر روی آن کار می کنند. به همین دلیل بعضی از محققان دارای مقالات با کاربردهای گوناگون و یا استفاده از روش های متفاوت هستند. همچنین بعضی از تکنیک ها دارای مفاهیم و روش های مشابه هستند. برای مثال روش های مبتنی بر پیش بینی پویا، سیستم های مبتنی بر دانش و شبکه های عصبی از این نوع هستند. تعداد محدودی از محققان بر روی مقالات با روش ها و کاربردهای مختلف کار کرده اند. این موضوع نشان از گستردگی و تفاوت توسعه روش ها بر مبنای عالیق و حوزه تحقیقاتی محقق است. به نظر می آید این موضوع نشانگر آن است که توسعه تکنیک های داده کاوی حول تخصص و متأثر از آن است.

گذشته از این، بعضی از مسائل ظرفیت این را دارا هستند که چندین روش را مورد استفاده قرار بدهند. برای مثال دسته بندی عصبی، شبکه های عصبی انتشار اطمینان بیزین، شبکه های نظم دهنده ژن، شبکه های عصبی بازگشت کننده فازی، شبکه های عصبی، درختان دسته بندی و رگرسیون، الگوریتم های استقرایی،²⁵ سیستم تشخیص نفوذ مبتنی بر رفتار غیرطبیعی، الگوریتم های خوش بندی ژنتیک، مدل های CRISP-DM، تحریک گد، معادله های Vlasov-Maxwell، پیش بینی های بازفعالی شبیه سانی، ریابی وسایل نقلیه بلاذرنگ، پیش بینی، تشخیص وضعیت نامطلوب، پیش بینی سرهم بندی، سنجش دانش، استخراج دانش، فرآگیری دانش، مدیریت دانش و نمایش دانش همگی موضوعات دارای روش های مختلف اند. این کاربردها نشان دهنده بخش مهمی از توسعه تکنیک های داده کاوی است اما روش های زیادی هستند که خاص یک مسئله توسعه می یابند. این موضوع نشان دهنده آن است که توسعه های آتی تکنیک های داده کاوی بیشتر مسئله محور خواهند بود.

چند کاربرد بهینه سازی سیستم ها عبارتند از: تحریک عصب الکتریکی، مقادیر بهینه سازی عملیات، تقسیم بندی عمودی، رگرسیون لاجستیکی، فرآیندهای سلسه مراتبی تحلیلی، رگرسیون چندجمله ای، بهینه سازی مبتنی بر زیست جغرافی، بهینه سازی از دحام ذرات، روش های عنصر محدود، روش های گستته حدود بالا و پایین مجموعه²⁶، روش های محاسبات موازی. کاربردهای تکنیک های بهینه سازی سیستم در جدول ۹ قابل مشاهده است.

Table 9
System optimization and its applications.

System optimization applications	Authors
Electrical nerve stimulation	Tam et al. (2004)
R peak detection	Yu et al. (2004)
Latent reference individual extraction method	Tamechika et al. (2006)
Operation optimization value	Li et al. (2007)
Vertical partitioning	Gorla and Betty (2008)
Logistic regression	Setoguchi et al. (2008)
Analytical hierarchy process	Shen and Chuang (2009)
Polynomial regression	Wi et al. (2009)
Biogeography based optimization	Chandrakala et al. (2010)
Particle swarm optimization	Durán et al. (2010)
Finite element method	Fernández et al. (2010)
Discrete rough set methods	Wan et al. (2010)
Asymptotic methods	Assous and Chaskalovic (2011)
Parallel computing	Jian et al. (2011)

جدول ۹

۱۱. سیستم های اطلاعاتی و کاربردهای آن ها

سیستم های اطلاعاتی از نتایج رشته های دانشگاهی است. این سیستم ها در جایی بین جهان تجارت و علوم کامپیوتر جای می گیرد و حوزه تجارت و حوزه خوش تعریف علوم کامپیوتر را به هم متصل می کند که باعث توسعه یک حوزه جدید علمی می شود. یک سیستم اطلاعاتی بر پایه های تئوری اطلاعات و محاسبات متکی است که به محققان اجازه می دهد تا در یک موقعیت منحصر بفرد توانند در مطالعات دانشگاهی مدل های مختلف تجارت و فرآیند های مرتبط الگوریتمی که در حوزه علوم کامپیوتر هستند شرکت داشته باشند. به طور عام سیستم های اطلاعاتی بر روی پردازش اطلاعات درون سازمان به خصوص سازمان های بزرگ تمرکز می کنند (Hoganson,2001) محصولات این فرآیند در ادامه می تواند به جامعه ارائه شود.

چند کاربرد سیستم های اطلاعاتی عبارتند از: خصوصیات بیمار، خصوصیات حوزه آبخیز، خودسازمان دهی، نقشه های خصوصیت، پایگاه داده ادعای خسارت بیمه، سنجش جایگزینی رشته کنونی، زیرسروی های آخرین زمان، انتخاب مقصد، مطالعات ارتباط خصوصیت، صورت وضعیت مالی کلاهبردار، تشابه دنباله، استدلال موردى، داده های مرتبط با اندازه گیری بدن انسان²⁷، عدم تعادل اقتصادی، خطاهای سطح میانی مشتری، نگهداری و مهندسی، وام دهی بانک، یادگیری تقویتی، یادگیری نظارتی، نمایش اطلاعات، نگهداری مشتری، مدیریت سرهم بندی، کشف الگو، مدیریت ارتباط با مشتری، قدرت فشرده تک محوره. کاربردهای سیستم های اطلاعاتی در جدول ۹ آورده شده است.

۱۲. مباحث، محدودیت ها و پیشنهادات

۱۲.۱ مباحث

یک تحلیل سطح بالا از تکنیک های داده کاوی باید بر روی حفاظت و نگهداری از داده ها تمرکز کند. در اولین تلاش های

25) discrete rough set methods

26) anthropometric

و همچنین مقالات دیگر مجالات و زرناقل ها و گزارش های عملی در این مقاله مورد بررسی قرار نگرفته اند. این مقالات می توانستند در رسیدن به اطلاعات کاملتری در مورد چگونگی توسعه تکنیک های داده کاوی به ما کمک کنند. در آخر مقالات غیر انگلیسی از این بررسی دور مانده اند. مطمئناً تحقیقات حول کاربردهای تکنیک های داده کاوی مورد بررسی و انتشار به زبان های دیگر نیز قرار داشته اند.

۳۰۱۲ پیشنهادات

(۱) روش های دیگر علوم اجتماعی. به این دلیل که روش های دیگری مانند روش های علوم اجتماعی در این مقاله مورد بررسی قرار نگرفته اند، تعریف تکنیک های داده کاوی این مقاله کامل نیست، اگرچه پرسشنامه های کیفی و روش های آماری از جمله تکنیک های مورد استفاده در علوم اجتماعی هستند. برای مثال علوم شناختی، روانشناسی و رفتارشناسی انسان هر کدام به طریقی به جستجوی مسائل انسانی می پردازند. بنابراین دیگر روش های علوم انسانی، می توانند در تحقیقات آتی، مورد بررسی، قرار گیرند.

(۲) ترکیب روش ها. تکنیک های داده کاوی یک موضوع تحقیقی بین رشته ای است بنابراین توسعه های آتی باید به وسیله یکپارچه شدن روش های مختلف انجام گیرد. این یکپارچگی روش ها و تحقیق بین رشته ای می تواند افق های جدیدی را برای مسائل موجود در تکنیک های داده کاوی نشان دهد.

(۳) تغییر، یکی از منابع توسعه های آتی سیستم های خبره است. تغییر، با هر دلیل اجتماعی و تکنیکی می تواند باعث توسعه و یا بازدارنگی استفاده از روش های مبتنی بر سیستم های خبره و توسعه کاربردهای آن باشد. قابل تصور است که سکون حاصل از استفاده از روش های معمول برای حل مسائل، منابع دانش ثابت و اتکا بر تجربه های پیشین می تواند مانع یادگیری و خلاقیت افداد و سازمان ها گردد.

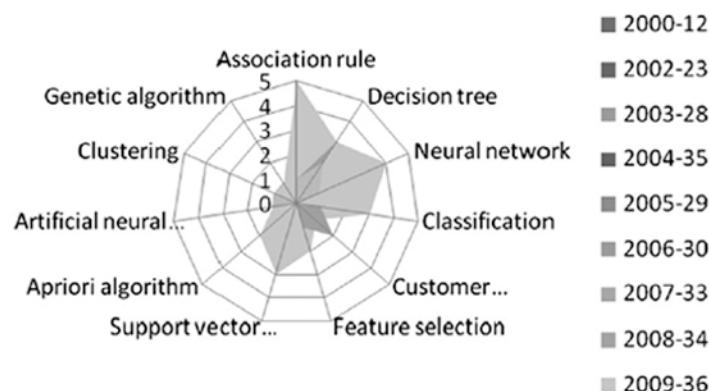
ادامه تولید، به اشتراک گذاری، یادگیری و کسب دانش در مورد روش‌ها و کاربردهای گوناگون نقشی کلیدی در توسعه سیستم‌های خبره بازی می‌کند.

نتیجہ گیری:

این مقاله به بررسی مقالات ارائه شده سال های ۲۰۰۰ تا ۲۰۱۱ در مورد تکنیک های داده کاوی و کاربردهای آن می پردازد. برای این منظور جستجویی در میان کلمات کلیدی و عنوان های مقالات انجام شد. نتایج حاصل بر این نکته تأکید می کنند که گرایش توسعه تکنیک های داده کاوی به سمت تخصصی تر شدن پیش می رود. همچنین کاربردهای تکنیک های داده کاوی به سمت تمرکز بر مسئله و مسئله محوری پیش می روند. در این مقاله پیشنهاد شده است که روش شناسی های تحقیقی علوم اجتماعی مانند روانشناسی، علوم شناختی و رفتارشناسی انسان به سمت استفاده از روش داده کاوی به جاوی روش های پیشین حرکت کنند. یکپارچگی روش های کیفی، کمی و علمی و یکپارچگی مطالعات روش شناسی تکنیک های داده کاوی به فهم بهتر موضوع کمک خواهد کرد. در نهایت توانایی تغییر مدام و فراهم آوردن درک جدید، از اصلی ترین مزایای تکنیک های داده کاوی است و در آنende از حمله عناصر اصلی. تکنیک های داده کاوی، خواهد بود.

در این بررسی، مقالات از رشته های گوناگون شامل علوم کامپیوتر، مهندسی، دارو شناسی، ریاضیات، علوم زمین و سیاره، زیست شیمی، ژنتیک، زیست شناسی مولکولی، مدیریت تجارت و حسابداری، علوم اجتماعی، علوم تصمیمی، علوم بین رشته ای، علوم محیطی، انرژی، علوم کشاورزی و زیستی، پرستاری، علوم مواد، دارو شناسی، سسم شناسی، شیمی، علوم سلامت، فیزیک و نجوم، اقتصاد، اقتصاد سنجی و مالی، روانشناسی، عصب شناسی، مهندسی شیمی و دامپزشکی مورد بررسی قرار گرفتند. این مقالات از پایگاه های داده Elsevier SCOPUS، Springerlink، IEEE Xplore، EBSCO elec آنلاین (-

کاربردهای صنعتی تکنیک های داده کاوی در بازه سال های ۲۰۰۰ تا ۲۰۱۱ افزایش داشته است. شکل ۱ میزان تمایل برای هر کدام از موارد یاد شده را نمایش می دهد.



نمی توان نتیجه گرفت که روش های داده کاوی و کاربردهای آن در دیگر علوم مورد بررسی قرار نگرفته است. با این حال باید مطالعات بیشتری در مورد کاربردهای روش های داده کاوی در رشته های مختلف منتشر شود تا به گسترش حوزه تکنیک های داده کاوی در سطح آکادمیک و عملی منجر شود.

۲۰۱۲ محدودت ها

بررسی های این مقاله دارای محدودیت هایی است. اولاً یک بررسی جامع تکنیک های داده کاوی و کاربردهای آن به دلیل نیاز به پیش زمینه های علمی گسترده در هنگام جمع آوری، مطالعه و دسته بندی مقالات، کاری دشوار است. با وجود پیش زمینه های محدود، این مقاله یک بررسی اجمالی از مقالات مرتبط با تکنیک های داده کاوی از سال ۲۰۰۰ تا ۲۰۱۱ سال انجام داد تا مشخص کند تکنیک های داده کاوی و کاربردهای آن چگونه گسترش و توسعه یافته اند. در حقیقت دسته بندی روش ها و کاربردهای آن ها براساس کلمات کلیدی و چکیده مقالات جمع آوری شده برای این تحقیق است. ممکن است تحقیقات دیگری از روشهای مشابه استفاده کرده باشد اما به طور خاص از عبارت «تکنیک های داده کاوی» استفاده نکرده باشد بنابراین این مقاله یک نمونه مشابه و دارای اشتراک با مقالات دیگر نخواهد بود. بنابراین اولین محدودیت این مقاله محدودیت علم نویسنده به این موضوع است. دوماً با این که ۲۶۱ مقاله از ۱۵۹ ژورنال (پنج پایگاه داده آنلاین) در این مقاله مرجع قرار گرفته اند، مقالات آکادمیک دیگر موجود در Science

References

- Adachi, H., Kikuchi, M., & Watanabe, Y. (2006). Electric switch machine failure detection using data-mining technique. *Quarterly Report of RTRI (Railway Technical Research Institute) (Japan)*, 47(4), 182–186.
- Adderley, R., Townsley, M., & Bond, J. (2007). Use of data mining techniques to model crime scene investigator performance. *Knowledge-Based Systems*, 20(2), 170–176.
- Ahn, H., Ahn, J. J., Oh, K. J., & Kim, D. H. (2011). Facilitating cross-selling in a mobile telecom market to develop customer classification model based on hybrid data mining techniques. *Expert Systems with Applications*, 38(5), 5005–5012.
- Akerkar, R. A., & Sajja Priti Srinivas (2009). *Knowledge-based systems*. Sudbury, MA, USA: Jones & Bartlett Publishers.
- Al-Hamami, A. H., Al-Hamami, M. A., & Hasheem, S. H. (2006). Applying data mining techniques in intrusion detection system on web and analysis of web usage. *Information Technology Journal*, 5(1), 57–63.
- Aliev, R. A., Aliev, R. R., Guirimov, B., & Uyar, K. (2008). Dynamic data mining technique for rules extraction in a process of battery charging. *Applied Soft Computing Journal*, 8(3), 1252–1258.
- Andronie, M., & Andronie, M. (2009). Data mining techniques used in metallurgic industry. *Metalurgia International*, 14(12), 17–22.
- Assous, F., & Chaskalovic, J. (2010). Méthodes de data mining pour l'analyse d'approximations numériques: Le cas de solutions asymptotiques des équations de Vlasov-Maxwell= data mining techniques for numerical approximations analysis: A test case of asymptotic solutions to the Vlasov-Maxwell equations. *Comptes Rendus.Mécanique*, 338(6), 305–310.
- Assous, F., & Chaskalovic, J. (2011). Data mining techniques for scientific computing: Application to asymptotic paraxial approximations to model ultrarelativistic particles. *Journal of Computational Physics*, 230(12), 4811–4827.
- Bae, J. K., & Kim, J. (2011). Product development with data mining techniques: A case on design of digital camera. *Expert Systems with Applications*, 38(8), 9274–9280.
- Bae, S. H., Kim, J., & Lim, H. (2009). A study on constructing the prediction system using data mining techniques to find medium-voltage customers causing distribution line faults. *Transactions of the Korean Institute of Electrical Engineers*, 58(12), 23–54.
- Bagui, S., Mint, D., & Cash, P. (2007). Data mining techniques to study voting patterns in the US. *Data Science Journal*, 6, 46–63.
- Balzano, W., & Del Sorbo, M. R. (2007). Genomic comparison using data mining techniques based on a possibilistic fuzzy sets model. *BioSystems*, 88(3), 343–349.
- Bâra, A., Lungu, I., & Oprea, S. V. (2009). Public institutions' investments with data mining techniques. *WSEAS Transactions on Computers*, 8(4), 589–598.
- Besse, P., Milhem, H., Mestre, O., Dufour, A., & Peuch, V. (2007). A comparison of data mining techniques for the statistical adaptation of ozone forecasts of chemistry-transport MOCAGE model. *Pollution Atmosphérique*, 195(45), 285–292.
- Bhramaramba, R., Allam, A. R., Kumar, V. V., & Sridhar, G. R. (2011). Application of data mining techniques on diabetes related proteins. *International Journal of Diabetes in Developing Countries*, 31(1), 22–25.
- Borghini, A., Crotti, P., Pietra, D., Favero, L., & Bianucci, A. M. (2010). Chemical reactivity predictions: Use of data mining techniques for analyzing regioselective azidolysis of epoxides. *Journal of Computational Chemistry*, 31(14), 2612–2619.
- Cesana, M., Cerutti, R., Grossi, E., Fagioli, E., Stabilini, M., Stella, F., et al. (2007). Bayesian data mining techniques: The evidence provided by signals detected in single-company spontaneous reports databases. *Drug Information Journal*, 41(1), 16–28.
- Chandrakala, D., Sumathi, S., & Saraswathi, D. (2010). Blur identification with image restoration based on application of data mining techniques. *International Journal of Imaging*, 4(10 A), 99–122.
- Chang, C., Wang, C., & Jiang, B. C. (2011). Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors. *Expert Systems with Applications*, 38(5), 5507–5513.
- Chen, C., & Chen, M. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers and Education*, 52(1), 256–273.
- Chen, Q., & Mynett, A. E. (2003). Integration of data mining techniques and heuristic knowledge in fuzzy logic modelling of eutrophication in taihu lake. *Ecological Modelling*, 162(1–2), 55–67.
- Chen, S., & Bai, S. (2010). Using data mining techniques to automatically construct concept maps for adaptive learning systems. *Expert Systems with Applications*, 37(6), 4496–4503.
- Chen, S. C., & Huang, M. Y. (2011). Constructing credit auditing and control & management model with data mining technique. *Expert Systems with Applications*, 38(5), 5359–5365.
- Chen, T., Chang, J., & Chen, D. (2008). Applying data mining technique to compute LDE for rutting through full scale accelerated pavement testing. *Road Materials and Pavement Design*, 9(2), 227–246.
- Chen, T., Chou, L., & Hwang, S. (2003). Application of a data-mining technique to analyze coprescription patterns for antacids in taiwan. *Clinical Therapeutics*, 25(9), 2453–2463.
- Chen, W., & Du, Y. (2009). Using neural networks and data mining techniques for the financial distress prediction model. *Expert Systems with Applications*, 36(2), 4075–4086.
- Chen, Y. (2004). Constructing language model by using data mining technique. *Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University*, 38(9), 1590–1592.
- Chou, J., Chiu, C., Farfoula, M., & Al-Taharwa, I. (2011). Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques. *Journal of Computing in Civil Engineering*, 25(3), 242–253.
- Chun, S., & Park, Y. (2006). A new hybrid data mining technique using a regression case based reasoning: Application to financial forecasting. *Expert Systems with Applications*, 31(2), 329–336.
- Chun, S., Kim, J., Hahn, K., Park, Y., & Chun, S. (2008). Data mining technique for medical informatics: Detecting gastric cancer using case-based reasoning and single nucleotide polymorphisms. *Expert Systems*, 25(2), 163–172.
- Constantinescu, Z., Marinou, C., & Vladou, M. (2010). Driving style analysis using data mining techniques. *International Journal of Computers, Communications and Control*, 5(5), 654–663.
- Cruz, A. M., Aguilera-Huertas, W. A., & Díaz-Mora, D. A. (2009). A comparative study of maintenance services using the data-mining technique. *Colombia's Journal of Public Health broad-casts research*, 11(4), 64–72.
- David, H. (1928). Die Grundlagen der Mathematik. *Abhandlungen aus dem Seminar der Hamburgischen Universität*, 6(65–68), 464–479.
- Dakheel, F. I., Smko, R., Negrat, K., & Almarimi, A. (2011). Using data mining techniques for finding cardiac outlier patients. *Proceedings of World Academy of Science, Engineering and Technology*, 73, 443–447.
- De Andrade, A. A., Pereira, S. L., Dias, E. M., & Fontana, C. F. (2010). Using data mining techniques for development expert systems equipped with learning capabilities for use in automated industrial plants. *WSEAS Transactions on Systems and Control*, 5(6), 21–31.
- Durán, O., Rodriguez, N., & Consalter, L. A. (2010). Collaborative particle swarm optimization with a data mining technique for manufacturing cell design. *Expert Systems with Applications*, 37(2), 1563–1567.
- Dutta, M., Mukhopadhyay, A., & Chakrabarti, S. (2004). Effect of galvanising parameters on spangle size investigated by data mining technique. *ISIJ International*, 44(1), 129–138.
- Escudero-Mancebo, D., & Cardeñoso-Payo, V. (2007). Applying data mining techniques to corpus based prosodic modeling. *Speech Communication*, 49(3), 213–229.
- Fan, Y., Chaovalltwongse, W. A., Liu, C., Sachdeo, R. C., Iasemidis, L. D., & Pardalos, P. M. (2009). Optimisation and data mining techniques for the screening of epileptic patients. *International Journal of Bioinformatics Research and Applications*, 23(2), 121–135.
- Fayyad, U., Djorgovski, S. G., & Weir, N. (1996). Automating the analysis and cataloging of sky surveys. In U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 471–494). Cambridge, MA: MIT Press.
- Feng, L., & Lu, H. (2004). Managing multiuser database buffers using data mining techniques. *Knowledge and Information Systems*, 6(6), 679–709.
- Fernández, J., Pernía, A., Martínez-de-Pisón, F. J., & Lostado, R. (2010). Prediction models for calculating bolted connections using data mining techniques and the finite element method. *Engineering Structures*, 32(10), 3018–3027.
- Fesharaki, M., Shirazi, H., & Bakhshi, A. (2011). Knowledge acquisition from database of information management and documentation softwares by data mining techniques. *Information Sciences and Technology*, 26(2), 259–283.
- Fiol-Roig, G., & Miró-Julià, M. (2011). Stock market analysis using data mining techniques: A practical application. *International Journal of Artificial Intelligence*, 6(11 S), 129–143.
- Fong, A. C. M., Hui, S. C., & Jha, G. (2003). A hybrid data mining technique for vehicle fault diagnosis. *Engineering Intelligent Systems*, 11(3), 113–122.
- Gajzler, M. (2010). Text and data mining techniques in aspect of knowledge acquisition for decision support system in construction industry. *Technological and Economic Development of Economy*, 16(2), 219–232.
- Gorla, N., & Betty, P. W. Y. (2008). Vertical fragmentation in databases using data-mining technique. *International Journal of Data Warehousing and Mining*, 4(3), 35–53.
- Granitto, P. M., Biasioli, F., Aprea, E., Mott, D., Furlanello, C., Märk, T. D., et al. (2007). Rapid and non-destructive identification of strawberry cultivars by direct PTR-MS headspace analysis and data mining techniques. *Sensors and Actuators*, 33(1), 44–56.
- Gregori, D., Petrinco, M., Bo, S., Rosato, R., Pagano, E., Berchialla, P., et al. (2011). Using data mining techniques in monitoring diabetes care. the simpler the better? *Journal of Medical Systems*, 35(2), 277–281.
- Hall, M. J., Minns, A. W., & Ashrafuzzaman, A. K. M. (2002). The application of data mining techniques for the regionalisation of hydrological variables. *Hydrology and Earth System Sciences*, 6(4), 685–694.
- Ha, S., Bae, S., & Park, S. (2000). Web mining for distance education. In *IEEE international conference on management of innovation and technology* (pp. 715–719).
- He, Y., Wang, Y., Luo, T., He, A., & Wang, J. (2010). Urban residential load combined forecast model based on data mining techniques and panel data theory. *Journal of Computational Information Systems*, 6(6), 1801–1808.
- Henderson, M. P. A., & Grey, V. (2011). Establishing and evaluating pediatric thyroid reference intervals on the roche modular analytics E 170 using computational statistics and data-mining techniques. *Clinical Biochemistry*, 44(10–11), 767–770.
- Hoganson, Ken. (2001). Alternative curriculum models for integrating computer science and information systems analysis, recommendations, pitfalls, opportunities, accreditations, and trends. *Journal of Computing Sciences in Colleges*, 17(2), 313–325.
- Hoonakker, F., Lachiche, N., Varnek, A., & Wagner, A. (2011). A representation to apply usual data mining techniques to chemical reactions illustration on the rate constant of SN2 reactions in water. *International Journal on Artificial Intelligence Tools*, 41(1), 31–45.
- Hoontarakul, P., & Sahadev, S. (2008). Application of data mining techniques in the on-line travel industry: A case study from thailand. *Marketing Intelligence and Planning*, 26(1), 60–76.
- Hsia, T., Shie, A., & Chen, L. (2008). Course planning of extension education to meet market demand by using data mining techniques - an example of chinkuo technology university in Taiwan. *Expert Systems with Applications*, 34(1), 596–602.
- Hsieh, K. (2007). Employing data mining technique to achieve the parameter optimization based on manufacturing intelligence. *Journal of the Chinese Institute of Industrial Engineers*, 24(4), 309–318.

- Hsieh, N., & Chu, K. (2009). Enhancing consumer behavior analysis by data mining techniques. *International Journal of Information and Management Sciences*, 20(1), 39–53.
- Huang, S., & Lin, J. (2002). Enhancement of power system data debugging using GSA-based data-mining technique. *IEEE Transactions on Power Systems*, 17(4), 1022–1029.
- Huang, S., Chang, E., & Wu, H. (2009). A case study of applying data mining techniques in an outfitter's customer value analysis. *Expert Systems with Applications*, 36(3 PART 2), 5909–5915.
- Hwang, W., Kim, J., Jang, W., Hong, J., & Han, D. (2008). Fault pattern analysis and restoration prediction model construction of pole transformer using data mining techniques. *Transactions of the Korean Institute of Electrical Engineers*, 57(9), 1123–1154.
- Ince, H., & Aktan, B. (2009). A comparison of data mining techniques for credit scoring in banking: A managerial perspective. *Journal of Business Economics and Management*, 10(3), 233–240.
- Jain, A. (2007). Optimizing web server performance using data mining techniques. *European Journal of Scientific Research*, 17(2), 222–231.
- Jegelevičius, D., Lukoševičius, A., Paunksnis, A., & Barzdžiukas, V. (2002). Application of data mining technique for diagnosis of posterior uveal melanoma. *Informatica*, 13(4), 455–464.
- Jian, L., Wang, C., Liu, Y., Liang, S., Yi, W., Shi, Y., et al. (2011). Data mining techniques on graphics processing unit with compute unified device architecture (CUDA). *The Journal of Supercomputing*, 15(1), 1–26.
- Jilani, T. A., Yasin, H., Yasin, M., Ardit, C. (2009). Acute coronary syndrome prediction using data mining techniques—an application. In *Proceedings of world academy of science, engineering and technology* (Vol. 59, pp. 474–478).
- Jiménez, A., Aroba, J., de la Torre, M. L., Andujar, J. M., & Grande, J. A. (2009). Model of behaviour of conductivity versus pH in acid mine drainage water, based on fuzzy logic and data mining techniques. *Journal of Hydroinformatics*, 11(2), 147–153.
- Kaichang, D., Deren, L., & Deyi, L. (2000). Remote sensing image classification with GIS data based on spatial data mining techniques. *Geo-Spatial Information Science*, 3(4), 30–35.
- Kang, Z., Luo, F., Pan, M., & Chen, D. (2004). Networked non-destructive testing system based on data mining technique. *Journal of Test and Measurement Technology*, 18(3), 263–268.
- Kheradmandian, G., & Rahmati, M. (2009). Automatic abstraction in reinforcement learning using data mining techniques. *Robotics and Autonomous Systems*, 57(11), 1119–1128.
- Kim, K., Ahn, H., Jeong, S. (2010). Context-aware recommender systems using data mining techniques. In *Proceedings of world academy of science, engineering and technology* (Vol. 64, pp. 357–362).
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32(4), 995–1003.
- Ko, M., & Osei-Bryson, K. (2006). Analyzing the impact of information technology investments using regression and data mining techniques. *Journal of Enterprise Information Management*, 19(4), 403–417.
- Küçükşille, E. U., Selbaş, R., & Şençan, A. (2009). Data mining techniques for thermophysical properties of refrigerants. *Energy Conversion and Management*, 50(2), 399–412.
- Lancashire, L. J., Mian, S., Ellis, I. O., Rees, R. C., & Ball, G. R. (2005). Current developments in the analysis of proteomic data: Artificial neural network data mining techniques for the identification of proteomic biomarkers related to breast cancer. *Expert Systems with Applications*, 32(5), 1434–1445.
- Laosiritaworn, W., & Holimchayachotikul, P. (2010a). Metal frame for actuator manufacturing process improvement using data mining techniques. *Chiang Mai Journal of Science*, 37(3), 421–428.
- Laosiritaworn, W. S., & Holimchayachotikul, P. (2010b). Machine scoring model using data mining techniques. *Proceedings of World Academy of Science, Engineering and Technology*, 64, 566–570.
- Lee, I., Liao, S., & Embrechts, M. (2000). Data mining techniques applied to medical information. *Medical Informatics and the Internet in Medicine*, 25(2), 81–102.
- Li, J., Niu, C., & Liu, J. (2006). Application of data mining technique in optimizing the operation of power plants. *Dongli Gongcheng/Power Engineering*, 26(6), 830–835.
- Li, J., Liu, J., Gu, J., & Niu, C. (2007). Study of operation optimization based on data mining technique in power plants. *Frontiers of Energy and Power Engineering in China*, 1(4), 457–462.
- Li, R., Yu, Z., Zhang, J., & Hu, R. (2002). Data warehouse and data mining technique applied in switch hitch management system. *Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems, CMS*, 8(11), 919–924.
- Li, X., Li, H., Zhou, J., Tan, F., He, X., & Chen, Y. (2011). Identifying water flooded layers based on the domain-driven data mining technique. *Shiyu Kantan Yu Kaifa/Petroleum Exploration and Development*, 38(3), 345–351.
- Lin, H., Hsu, C., Wang, M. J., & Lin, Y. (2008). An application of data mining techniques in developing sizing systems for army soldiers in Taiwan. *WEAS Transactions on Computers*, 7(4), 245–252.
- Lin, W. T., Wang, S. J., Wu, Y. C., & Ye, T. C. (2011). An empirical analysis on auto corporation training program planning by data mining techniques. *Expert Systems with Applications*, 38(5), 5841–5850.
- Lu, X., Dong, Z. Y., & Li, X. (2005). Electricity market price spike forecast with data mining techniques. *Electric Power Systems Research*, 73(1), 19–29.
- Ma, P. C. H., & Chan, K. C. C. (2007). An effective data mining technique for reconstructing gene regulatory networks from time series expression data. *Journal of Bioinformatics and Computational Biology*, 5(3), 651–668.
- Mahesh, S., Mahesh, T. R., & Vinayababu, M. (2010). Using data mining techniques for detecting terror-related activities on the web. *Journal of Theoretical and Applied Information Technology*, 16(2), 99–104.
- Mamčenko, J., & Kulvietiene, R. (2005). Data mining technique for collaborative server activity analysis. *WSEAS Transactions on Information Science and Applications*, 2(5), 530–533.
- Martínez-de-Pisón, F. J., Ordieres, J., Perría, A., Alba, F., & Torre, V. (2007). Reduce of adherence problems in galvanised processes through data mining techniques. *Reducción de problemas de adherencia en procesos de galvanizado mediante técnicas de mi*.
- Marx, A. (2010). Detection and classification of bark beetle infestation in pure norway spruce stands with multi-temporal RapidEye imagery and data mining techniques. In *Erkennung von borkenkäferbefall in fichtenreinbeständen mit multi-temporalen RapidEye-S*.
- Mei, D., & Zhang, X. (2007). Data mining techniques for structure of single XML document. *Shiyu Huagong Gaodeng Xuexiao Xuebao/Journal of Petrochemical Universities*, 20(1), 94–98.
- Miranda, T., Correia, A. G., Santos, M., Ribeiro e Sousa, L., & Cortez, P. (2010). New models for strength and deformability parameter calculation in rock masses using data-mining techniques. *International Journal of Geomechanics*, 11(1), 44–58.
- Mucherino, A., Papajorgji, P., & Pardalos, P. (2009). A survey of data mining techniques applied to agriculture. *Operational Research*, 9(2), 121–140.
- Mylod, D. E., & Kaldenberg, D. O. (2000). Data mining techniques for patient satisfaction data in home care settings. *Home Health Care Management and Practice*, 12(6), 18–29.
- Neto, M. A. S., Villwock, R., Scheer, S., Steiner, M. T. A., & Dyminski, A. S. (2010). Visual data mining techniques applied for the analysis of data collected at itaipu power plant. *Técnicas de Mineração Visual de Dados aplicadas aos dados de instrumento*.
- Nikolova, E., & Jecheva, V. (2010). Some similarity coefficients and application of data mining techniques to the anomaly-based IDS. *Telecommunication Systems*, 33(2), 64–75.
- Oh, H., Doh, I., & Chae, K. (2009). Attack classification based on data mining technique and its application for reliable medical sensor communication. *International Journal of Computer Science and Applications*, 6(3), 20–32.
- Parhizi, S., Shahrabi, J., & Pariazar, M. (2009). A new accident investigation approach based on data mining techniques. *Journal of Applied Sciences*, 9(4), 731–737.
- Paul, S. K. (2008). Predictability of price of tea from sensory assessments and biochemical information using data-mining techniques. *Journal of the Science of Food and Agriculture*, 88(8), 1354–1362.
- Peter, J. (1968). The working set model for program behavior. *Communications of the ACM*, 11(5), 323–333.
- Pelletier, R. P., & Diderrich, G. T. (2007). A note on breiman's random forest data mining technique and conventional cox modeling of survival statistics. *Communications in Statistics – Theory and Methods*, 36(10), 1953–1964.
- Piquer, A. G., Albert, F. H., Ribé, E. G., & Cugota Florejachs, L. (2010). Assessment of competences in university degrees using data mining techniques. *Validación de Competencias en Titulaciones Universitarias Usando Minería de Datos Revista Iberoamerica*.
- Popovic, Z. (2004). Implementation of data mining techniques in construction estimating. *Neural, Parallel and Scientific Computations*, 12(1), 37–52.
- Qiang, Y., Guo, Y., Zhang, S., Wang, Q., & Niu, G. (2008). Using data mining techniques to establish solitary pulmonary nodules diagnosis model. *Chinese Journal of Medical Imaging Technology*, 24(3), 438–442.
- Radosavljević, V., Vučetić, S., & Obradović, Z. (2010). A data-mining technique for aerosol retrieval across multiple accuracy measures. *IEEE Geoscience and Remote Sensing Letters*, 7(2), 411–415.
- Ranjan, J., & Bhatnagar, V. (2010). Application of data mining techniques in the financial sector for profitable customer relationship management. *International Journal of Information and Communication Technology*, 2(4), 342–354.
- Ranjan, J., Goyal, D. P., & Ahson, S. I. (2008). Data mining techniques for better decisions in human resource management systems. *International Journal of Business Information Systems*, 3(5), 464–481.
- Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, 50(2), 491–500.
- Rezania, M., Javadi, A. A., & Giustolisi, O. (2008). An evolutionary-based data mining technique for assessment of civil engineering systems. *Engineering Computations*, 25(6), 500–517.
- Ribeiro, M., Plastino, A., & Martins, S. (2006). Hybridization of GRASP metaheuristic with data mining techniques. *Journal of Mathematical Modelling and Algorithms*, 5(1), 23–41.
- Riquelme, J. C., Ruiz, R., Rodriguez, D., & Aguilar-Ruiz, J. S. (2009). JISBD 04: Finding defective software modules by means of data mining techniques. *Latin America Transactions, IEEE (Revista IEEE America Latina)*, 7(3), 377–382.
- Rivas, T., Paz, M., Martín, J. E., Matías, J. M., García, J. F., & Taboada, J. (2011). Explaining and predicting workplace accidents using data-mining techniques. *Reliability Engineering and System Safety*, 96(7), 739–747.
- Russell, Stuart J., & Norvig, Peter. (2003). *Artificial Intelligence: A Modern Approach* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall. ISBN 0-13-790395-2, Chapter 2.
- Rygielski, C., Wang, J., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, 24(4), 483–502.
- Saitta, S., Raphael, B., & Smith, I. F. C. (2005). Data mining techniques for improving the reliability of system identification. *Advanced Engineering Informatics*, 19(4), 289–298.
- Sánchez, A., Peña, J. M., Pérez, M. S., Robles, V., & Herrero, P. (2004). Improving distributed data mining techniques by means of a grid infrastructure. *Expert Systems with Applications*, 26(1), 21–33.
- Saygin, Y., & Ulusoy, Ö. (2002). Exploiting data mining techniques for broadcasting data in mobile computing environments. *IEEE Transactions on Knowledge and Data Engineering*, 14(6), 1387–1399.
- Segura, A., Vidal-Castro, C., Menéndez-Domínguez, V., Campos, P. G., & Prieto, M. (2011). Using data mining techniques for exploring learning object repositories. *Electronic Library*, 29(2), 162–180.



- Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R. J., & Cook, E. F. (2008). Evaluating uses of data mining techniques in propensity score estimation: A simulation study. *Pharmacoepidemiology and Drug Safety*, 17(6), 546–555.
- Shahabi, J., & Neyestani, R. S. (2009). Discovering iranians' shopping culture by considering virtual items using data mining techniques. *Journal of Applied Sciences*, 9(13), 2351–2361.
- Shapira, P., & Youtie, J. (2006). Measures for knowledge-based economic development: Introducing data mining techniques to economic developers in the state of georgia and the US south. *Technological Forecasting and Social Change*, 73(8), 950–965.
- Shen, C., & Chuang, H. (2009). A study on the applications of data mining techniques to enhance customer lifetime value. *WSEAS Transactions on Information Science and Applications*, 6(2), 319–328.
- Shih, C., Chiang, D., Lai, S., & Hu, Y. (2009). Applying hybrid data mining techniques to web-based self-assessment system of study and learning strategies inventory. *Expert Systems with Applications*, 36(3 PART 1), 5523–5532.
- Shmueli, O., Shmueli, T., Dagan, Y., & Teicher, M. (2009). Data mining techniques for detection of sleep arousals. *Journal of Neuroscience Methods*, 179(2), 331–337.
- Singhal, A., & Jajodia, S. (2006). Data warehousing and data mining techniques for intrusion detection systems. *Distributed and Parallel Databases*, 20(2), 149–166.
- Srivastava, A. N., Oza, N. C., & Stroeve, J. (2005). Virtual sensors: Using data mining techniques to efficiently estimate remote sensing spectra. *Geoscience and Remote Sensing, IEEE Transactions on Engineering*, 43(3), 590–600.
- Su, C., Wang, P., Chen, Y., & Chen, L. (2005). Data mining techniques for assisting the diagnosis of pressure ulcer development in surgical patients. *Journal of Medical Systems*, 3(1), 1–13.
- Su, Q., Shao, P., & Zou, T. (2010). CRBT customer churn prediction: Can data mining techniques work? *International Journal of Networking and Virtual Organisations*, 7(4), 353–365.
- Sudheep, E. M., Idikkula, S. M., & Alexander, J. (2011). Design and performance analysis of data mining techniques based on decision trees and naive bayes classifier for employment chance prediction. *Journal of Convergence Information Technology*, 6(5), 89.
- Sugihara, T. (2006). An approach of electric power demand forecasting using data-mining method: A case study of application of data-mining technique to improve decision making. *International Journal of Management and Decision Making*, 7(1), 88–104.
- Sui, L., & Meng, L. (2001). Adding the value of NAVTECH road database: An implementation of spatial data mining techniques. *Journal of Geographical Sciences*, 11(0), 69–73.
- Symeonidis, A. L., Kehagias, D. D., & Mitkas, P. A. (2003). Intelligent policy recommendations on enterprise resource planning by the use of agent technology and data mining techniques. *Expert Systems with Applications*, 25(4), 589–602.
- Tam, S., Cheung, G. L. Y., & Hui-Chan, C. W. Y. (2004). Predicting osteoarthritic knee rehabilitation outcome by using a prediction model developed by data mining techniques. *International Journal of Rehabilitation Research*, 27(1), 65–69.
- Tamechika, Y., Iwatani, Y., Tohyama, K., & Ichihara, K. (2006). Insufficient filling of vacuum tubes as a cause of microhemolysis and elevated serum lactate dehydrogenase levels: use of a data-mining technique in evaluation of questionable laboratory test. *Expert Systems with Applications*, 29(2), 323–335.
- Thiruvadi, S., & Patel, S. C. (2011). Survey of data-mining techniques used in fraud detection and prevention. *Information Technology Journal*, 10(4), 710–716.
- Tinoco, J., Gomes Correia, A., & Cortez, P. (2011). Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time. *Construction and Building Materials*, 25(3), 1257–1262.
- Trafalis, T. B., & White, A. (2003). Data mining techniques for pattern recognition: Tornado signatures in doppler weather radar data. *International Journal of Smart Engineering System Design*, 5(4), 347–359.
- Trafalis, T. B., Richman, M. B., White, A., & Santosa, B. (2002). Data mining techniques for improved WSR-88D rainfall estimation. *Computers and Industrial Engineering*, 43(4), 775–786.
- Tremblay, M. C., Dutta, K., & Vandermeer, D. (2010). Using data mining techniques to discover bias patterns in missing data. *Journal of Data and Information Quality*, 2(1), 23–55.
- Tsai, C., & Chen, M. (2008). Using adaptive resonance theory and data-mining techniques for materials recommendation based on the e-library environment. *Electronic Library*, 26(3), 287–302.
- Tsai, C., & Lu, Y. (2010). Data mining techniques in customer churn prediction. *Recent Patents on Computer Science*, 3(1), 28–32.
- Tu, C., Chang, C., Chen, K., & Lu, H. (2011). Application of data mining technique in the performance analysis of shipping and freight enterprise and the construction of stock forecast model. *Journal of Convergence Information Technology*, 6(3), 331–342.
- Ture, M., Tokatlı, F., & Kurt Omurlu, I. (2009). The comparisons of prognostic indexes using data mining techniques and cox regression analysis in the breast cancer data. *Expert Systems with Applications*, 36(4), 8247–8254.
- Vega, J., Pereira, A., Portas, A., Dormido-Canto, S., Farias, G., Dormido, R., et al. (2008). Data mining technique for fast retrieval of similar waveforms in fusion massive databases. *Fusion Engineering and Design*, 83(1), 132–139.
- Wan, S., Lei, T. C., & Chou, T. Y. (2010). A novel data mining technique of analysis and classification for landslide problems. *Natural Hazards*, 52(1), 211–230.
- Wang, Y., Chiang, D., Lai, S., & Lin, C. (2010). Applying data mining techniques to WIFLY in customer relationship management. *Information Technology Journal*, 9(3), 488–493.
- Wang, Y., & Zhang, X. (2006). Time series data mining technique based on varying series. *Nanjing Hangkong Daxue Xuebao/Journal of Nanjing University of Aeronautics and Astronautics*, 38(15), 154–157. SUPPL.
- Wasan, S. K., Bhatnagar, V., & Kaur, H. (2006). The impact of data mining techniques on medical diagnostics. *Data Science Journal*, 5, 119–126.
- Weiss, S. H., & Indurkha, N. (1998). *Predictive Data Mining: A Practical Guide*. San Francisco, CA: Morgan Kaufmann Publishers.
- Wi, Y., Song, K., & Joo, S. (2009). Data mining technique using the coefficient of determination in holiday load forecasting. *Transactions of the Korean Institute of Electrical Engineers*, 58(1), 18–22.
- Wiig, K. M. (1994). Knowledge management, the central management. *Expert Systems with Applications*, 10(1), 32–45.
- Wong, J., Chen, H., Chung, P., & Kao, N. (2006). Identifying valuable travelers and their next foreign destination by the application of data mining techniques. *Asia Pacific Journal of Tourism Research*, 11(4), 355–373.
- Xiao, H., Zhang, G., Liu, D., & Cai, X. (2006). Selecting forest fire spreading models based on the fuzzy data mining technique. *Beijing Linye Daxue Xuebao/Journal of Beijing Forestry University*, 28(6), 93–97.
- Ye, N., Li, X., & Farley, T. (2003). A data mining technique for discovering distinct patterns of hand signs: Implications in user training and computer interface design. *Ergonomics*, 46(1–3), 188–196.
- Yeh, J., Wu, T., & Tsao, C. (2011). Using data mining techniques to predict hospitalization of hemodialysis patients. *Decision Support Systems*, 50(2), 439–448.
- Yu, H., Xie, Y., Zhou, Z., & Lu, Y. (2004). Application of holter ECG signal analysis based on wavelet and data mining technique. *Transactions of Tianjin University, Remote Sensing*, 10(2), 126–129.
- Yu, H., Zhang, L., Liu, W., Huang, Z., & Ding, M. (2006). Application of holter ECG signal analysis based on wavelet and data mining technique. *Tianjin Daxue Xuebao (Ziran Kexue Yu Gongcheng Jishu Ban)/Journal of Tianjin University Science*, 12(3), 243–254.
- Zhang, C., & Ramirez-Marquez, J. E. (2011). Approximation of minimal cut sets for a flow network via evolutionary optimization and data mining techniques. *International Journal of Performability Engineering*, 7(1), 21–31.
- Zhu, Y., & Xiong, Y. (2007). DNA sequence data mining technique. *Ruan Jian Xue Bao/Journal of Software*, 18(11), 2766–2781.

مقایسه عملکرد الگوریتم‌های ژنتیک، تبرید شبیه سازی شده و پسگرد در حل مسئله چند وزیر

علیرضا غفاری صدر^۱، سروش عبادی^۲، محمدرضا اصغری اسکوئی^۳

۱. دانشکدهی علوم ریاضی و کامپیوتر، دانشگاه علامه طباطبائی، تهران
a.ghsadr@computation.ir

۲. دانشکدهی علوم ریاضی و کامپیوتر، دانشگاه علامه طباطبائی، تهران
soroush.e89@gmail.com

۳. استادیار دانشکدهی علوم ریاضی و کامپیوتر، دانشگاه علامه طباطبائی، تهران
oskoei@atu.ac.ir

کلیدواژه:

مسئله چند وزیر
جستجوی محلی
پسگرد
الگوریتم ژنتیک
تبرید شبیه سازی شده

چکیده:

مسئله‌ی چند وزیر، جزء مسائل پایه‌ای و در عین حال، پیچیده‌ی علوم کامپیوتر به حساب می‌آید که در این مقاله، به مقایسه‌ی عملکرد روش‌های حل آن توسط الگوریتم پسگرد و دو روش جستجوی محلی، شامل الگوریتم ژنتیک و تبرید شبیه سازی شده پرداخته‌ایم و پارامترهای مختلف و همچنین فضای حلالات گوناگون را مقایسه نموده‌ایم. این مقایسه‌ها نسبت به تعداد دوره‌های تکرار، تعداد گره‌های مشاهده شده در هر تکرار و همچنین زمان رسیدن به باسن صورت گرفته است. نتایج بدست آمده نشان می‌دهد که روش پسگرد، زمان و فضای بسیار زیادی را به هدر داده و کارآمد نمی‌باشد. در مقابل، روش جستجوی محلی، کارایی قابل قبولی را نسبت به زمان و فضای ایجاد می‌نماید. در میان دو روش جستجوی محلی ارائه شده نیز، الگوریتم ژنتیک نتایج مناسب‌تری را نسبت به تبرید شبیه سازی شده، نشان می‌دهد.

۱. مقدمه

مسئله‌ی چند وزیر، یکی از مسائل کلاس NP-hard می‌باشد. این نوع مسائل، از مرتبه‌ی اجرایی غیر خطی (نمایی) برخوردار بوده و با افزایش ابعاد مسئله، در زمانی قابل قبول، توسط تکنیک‌های قطعی، قابل حل نیستند. بنابراین از الگوریتم‌های جستجوی محلی (غیر قطعی) جهت حل آن‌ها استفاده می‌گردد که زمان مناسبی را جهت رسیدن به یک پاسخ، سپری می‌نمایند.

تحقیقات گوناگونی در زمینه‌ی حل مسئله‌ی چند وزیر، انجام گرفته است که از جمله‌ی آنها می‌توان به مارتینجاک^[۲]، اشاره نمود. وی از روش‌های تبرید شبیه سازی شده^۱، جستجوی ممنوعه^۲ و الگوریتم ژنتیک برای حل مسئله‌ی چند وزیر استفاده نمود و پارامترهای مختلف فضای حلال در آنها مورد بررسی قرار داد. یکی از نتایج بررسی شده، مسئله‌ی ۱۰ وزیر می‌باشد که در آن، میانگین پاسخ‌های هر سه روش، مقایسه گردیده‌اند. این مسئله، در روش تبرید شبیه سازی شده با میانگین ۹۴۷.۸، در روش جستجوی ممنوعه با میانگین ۴۷۲.۵ و در روش الگوریتم ژنتیک با میانگین ۴۹۱۰ تعداد تکرار، به پاسخ رسیده است.

بازیکوویچ^[۳]، از الگوریتم ژنتیک موازی برای حل مسئله‌ی چند وزیر استفاده نمود که طبق آن، برای مسئله‌ی ۱۰۰ وزیر با ۵۳۷ دوره‌ی تکرار در زمان ۱۵ میلی ثانیه، مسئله‌ی ۲۰۰ وزیر با ۱۳۴۶ دوره‌ی تکرار در زمان ۵۲ میلی ثانیه، مسئله‌ی ۵۰۰ وزیر با ۲۰۷۳ دوره‌ی تکرار در زمان ۵۵۳ میلی ثانیه، مسئله‌ی ۱۰۰۰ وزیر با ۱۱۳۹۵ دوره‌ی تکرار در زمان ۱۹۶ ثانیه و برای ۲۰۰۰ وزیر با ۲۶۱۳۲ دوره‌ی تکرار در زمان ۸.۷ ثانیه به پاسخ رسیده است.

از خر^[۴]، نیز روش‌های بهینه‌سازی اجتماع ذرات^۳ و الگوریتم ژنتیک را در حل

مسئله‌ی چند وزیر به کار برد که طبق آن در مسئله‌ی ۱۶ وزیر، الگوریتم ژنتیک توسط جمعیت اولیه‌ی 10^0 ، احتمال جهش 0.5 ، روش انتخاب تورنومنت و ترکیب نیمه تطبیقی^۴، با 60 دوره‌ی تکرار و در 19999 ثانیه به پاسخ رسیده است. همچنین الگوریتم بهینه‌سازی اجتماع ذرات توسط $(T_{max} - G) \times (0.9 - 0.4)$ و $w = 0.1$ تا 0.9 برابر عددی تصادفی بین $(0.1, 0.9)$ و اندازه‌ی اجتماع 10^0 ، با 2^{3547} دوره‌ی تکرار در 61.9999 ثانیه، پاسخ مسئله را یافت.

تارکی^[۵]، از الگوریتم ژنتیک جهت حل مسئله‌ی چند وزیر استفاده نمود. این الگوریتم با احتمال ترکیب 0.8 ، احتمال جهش 0.2 ، روش انتخاب تورنومنت، ترکیب دو نقطه‌ای، جهش تک بیتی و جمعیت 200 ، مسئله‌ی ۱۰۰ وزیر را در 3.5537 ثانیه حل نموده است.

این مقاله، مقایسه‌ی روش پسگرد و الگوریتم‌های جستجوی محلی شامل الگوریتم ژنتیک و تبرید شبیه سازی شده را در مسئله‌ی چند وزیر، ارائه می‌دهد که ضمن بررسی فضای حلالات مختلف، پارامترهای هر روش بصورت بهینه تنظیم شده است. این الگوریتم‌ها نسبت به دوره تکرار، تعداد مشاهده‌ی گره‌ها در هر تکرار و همچنین زمان رسیدن به پاسخ، با یکدیگر مقایسه شده اند. بخش ۲، شامل شرح مسئله‌ی چند وزیر می‌باشد. در بخش ۳، به توصیف روش پسگرد در حل مسئله‌ی چند وزیر و همچنین بدست آوردن نتایج حاصل از آن پرداخته‌ایم. بخش ۴ نیز شامل شرح روش‌های جستجوی محلی تبرید شبیه سازی شده و الگوریتم ژنتیک، در حل این مسئله است. نتایج بدست آمده از این بخش را نیز در بخش ۵، با یکدیگر مقایسه نموده‌ایم.

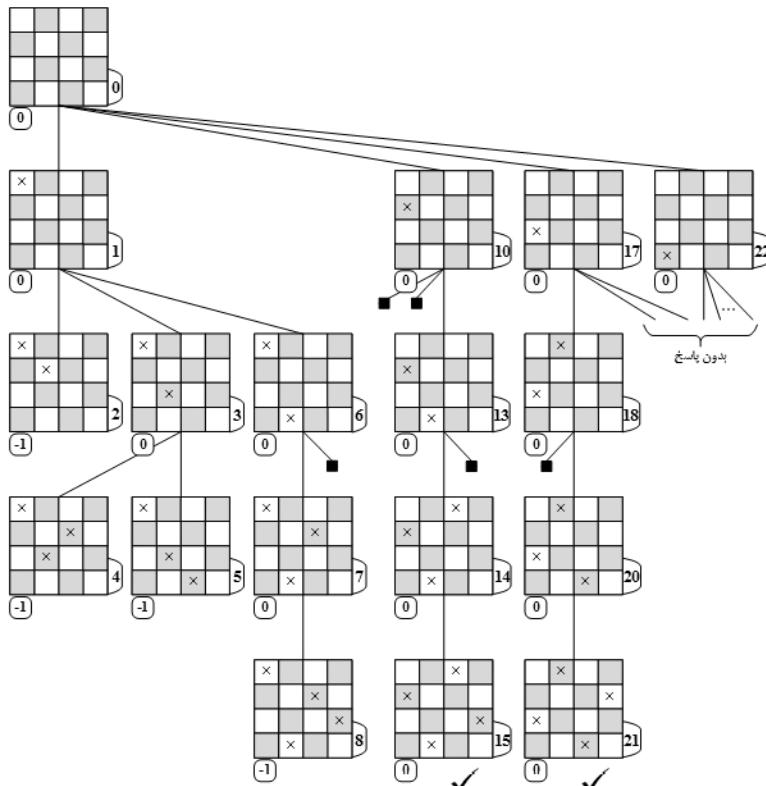
1) Simulated annealing
2) Tabu search

3) Particle Swarm Optimization
4) Partial Matched Crossover

گرفتن محدودیت‌های موجود، حل نماید. هدف اصلی، رسیدن به پاسخی برای مسئله‌ی چند وزیر، در کوتاه‌ترین زمان ممکن می‌باشد. در نتیجه استفاده از روشی مانند پسگرد، ناکارآمد خواهد بود. زیرا زمان بسیار زیادی را سپری می‌نماید. روش‌های جستجوی محلی مانند تبرید شبیه سازی شده و الگوریتم زنگنه، قادرند تا در مدت زمان کوتاهی، پاسخ را بدست آورند.

۳. روش پسگرد

راهبردهای جستجو در فضای حالت، به دو دسته‌ی ناآگاهانه و آگاهانه تقسیم‌بندی می‌شوند. در راهبرد نوع اول، هیچ اطلاعاتی به جز تعريف مسئله، در اختیار الگوریتم نیست. جستجوی اول عمق، یکی از این راهبردها می‌باشد. راهبردهای جستجوی آگاهانه، می‌دانند که از بین گره‌های غیر هدف، کدامیک با احتمال بیشتری ما را به پاسخ مسئله، تزدیک می‌کنند. تکنیک پسگرد، یکی از روش‌های شناخته شده در مسائل اراضی محدودیت و همچنین یک روش جستجوی ناآگاهانه در فضای حالت می‌باشد. این الگوریتم همانند جستجوی اول عمق است، با این تفاوت که فرزندان یک گره تنها هنگامی ملاقات می‌شوند که آن گره امید بخش باشد و در آن، حلقه‌ای وجود نداشته باشد. با توجه به اینکه هیچ دو وزیری نباید یکدیگر را تهدید کنند و همچنین طبق محدودیت‌های اعمال شده در مسئله که نمی‌توانند در یک سطر و ستون قرار گیرند، تعداد حالات کل، برابر با 4^n خواهد بود. در این مسئله، هر وزیر می‌تواند مهره‌هایی که در خانه‌های قطری آن قرار دارند را تهدید کند [۱]. برای سادگی تشریح این مسئله با استفاده از روش پسگرد، مسئله‌ی ۴ وزیر را در نظر می‌گیریم. مراحل جستجو برای یافتن پاسخ را بدین صورت دنبال می‌کنیم که:

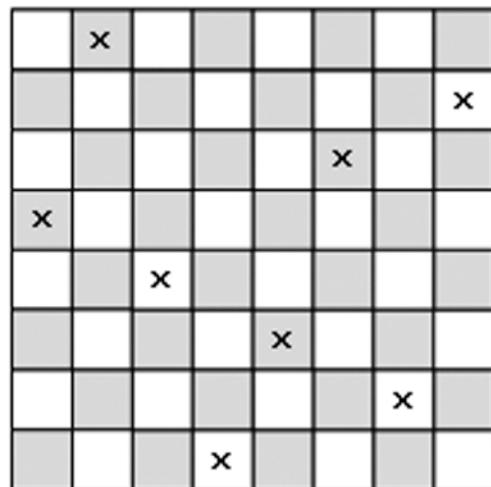


شکل ۲: حل مسئله‌ی ۴ وزیر با روش پسگرد؛ در این روش از هرس درخت جستجو جهت رسیدن به پاسخ استفاده شده است

۲. مسئله‌ی چند وزیر

یکی از مسائل مورد توجه در طراحی الگوریتم‌ها، حل معماهی چند وزیر می‌باشد. صورت مسئله بدين شکل است که n وزیر را می‌بایست در یک صفحه $n \times n$ به گونه‌ای قرار دهیم که هیچ دو وزیری یکدیگر را تهدید نکنند. تهدید به صورت سطحی، ستونی و قطری اتفاق می‌افتد. در نتیجه برای چیدمان مهره‌ها باید تدبیری اندیشید که تهدید رخ ندهد. شکل(۱)، یک حالت چیدمان صحیح از مهره‌ها را روی صفحه 8×8 نمایش می‌دهد. مسئله‌ی ۸ وزیر دارای ۹۲ پاسخ است که از این تعداد، تنها ۱۲ پاسخ، منحصر به فرد می‌باشند. مابقی از تقارن این ۱۲ پاسخ بدست می‌آیند [۲].

برای ساده‌تر نمودن حل یک مسئله، می‌توان از اعمال برخی محدودیت‌ها استفاده نمود. زیرا زمان رسیدن به پاسخ را سرعت می‌بخشد. در مسئله‌ی چند وزیر نیز می‌توان با اعمال محدودیت، بدون در نظر گرفتن روش حل آن، فضای جستجو را کاهش داد تا سرعت دستیابی به پاسخ، افزایش یابد [۱]. بدین ترتیب، اعمال محدودیت، امکان استفاده از روش‌های کلاسیک اراضی محدودیت از جمله روش پسگرد را فراهم می‌آورد. به عنوان مثال می‌توان روند چیدمان را این گونه پیش برد که هر وزیر باید تنها در یک سطر و ستون قرار گیرد. با توجه به این مطلب، هر چیزی توسعه یک n -تایی مرتب، قبل نمایش خواهد بود. بنابراین شکل(۱) را می‌توان به صورت (۴, ۱, ۵, ۸, ۶, ۳, ۷, ۲) نمایش داد. این نحوه نمایش، تضمین می‌نماید که پیچیدگی این الگوریتم، $(n!)^n$ خواهد بود [۳].



شکل ۱: یک پاسخ برای مسئله‌ی ۸ وزیر

اعمال این محدودیت، برخوردهای سطحی و ستونی را حذف می‌نماید. اما مشکل برخوردهای قطری همچنان پابرجا است. برای بر طرف ساختن برخوردهای قطری نیز باید یک محدودیت اعمال شود که طی آن، همیشه رابطه‌ی (۱) برقرار باشد.

$$|q_i - q_j| \neq i - j; i \neq j$$

که در آن q_i و q_j وزیرهایی هستند که در ستون‌های i و j قرار دارند. به عبارت دیگر، اختلاف عددی هیچ دو وزیری (غیر تکراری)، نباید با اختلاف اندیس آنها در n -تایی مرتب، برابر باشد. هر n -تایی مرتب، حاوی یک چیدمان از مهره‌ها و یک تابع برازنده‌ی متناظر با آن می‌باشد که تعداد تهدیدهای آن را مشخص می‌نماید. در نتیجه هرچه مقدار این تابع کمتر باشد، به یکی از پاسخ‌ها نزدیکتر خواهیم شد و زمانی که این مقدار به ۰ برسد، به یک راه حل رسیده‌ایم. اکنون باید روشی ارائه داد تا این مسئله را با در نظر

1) Fitness function
2) Uninformed search

3) Informed search
4) Depth first search

n	فضای کل ($n!$)	تعداد پاسخ‌ها ممکن	تعداد پاسخ‌های منحصر به فرد	زمان اجرای جستجوی پاسخ‌ها (h:m:s.ms)
۱	۱	۱	۱	۰:۰۰:۰۰
۲	۲	۶	۱	۰:۰۰:۰۰
۳	۶			
۴	۲۴	۲	۱	۰:۰۰:۰۰
۵	۱۲۰	۱۰	۲	۰:۰۰:۰۰
۶	۷۲۰	۴	۱	۰:۰۰:۰۰
۷	۵۰۴۰	۴۰	۶	۰:۰۰:۰۰
۸	۴۰۳۲۰	۹۲	۱۲	۰:۰۰:۰۰
۹	۳۶۲۸۸۰	۳۵۲	۴۶	۰:۰۰:۰۰
۱۰	۳۶۲۸۸۰۰	۷۲۴	۹۲	۰:۰۰:۰۰
۱۱	۳۹۹۱۶۸۰۰	۲۶۸۰	۳۴۱	۰:۰۰:۰۰
۱۲	۴۷۹۰۰۱۶۰۰	۱۴۲۰۰	۱۷۸۷	۰:۰۰:۰۰
۱۳	۶۲۲۷۰۰۸۰۰	۷۳۷۱۲	۹۲۳۳	۰:۰۰:۰۰.۳۰
۱۴	۸۷۱۷۸۲۹۱۲۰۰	۳۶۵۵۹۶	۴۵۷۵۲	۰:۰۰:۰۰.۱۷
۱۵	۱,۳۰E+۱۲	۲۲۷۹۱۸۴	۲۸۵۰۵۳	۰:۰۰:۱۰.۸
۱۶	۲,۰۹E+۱۳	۱۴۷۷۲۵۱۲	۱۸۴۶۹۵۵	۰:۰۰:۴۶.۹۷
۱۷	۳,۵۵E+۱۴	۹۵۸۱۵۱۰۴	۱۱۹۷۷۹۳۹	۰:۰۰:۴۸.۳۹
۱۸	۶,۴۰E+۱۵	۶۶۶۰۹۰۶۲۴	۸۳۲۶۳۵۹۱	۰:۰۵:۵۱.۸۲۸
۱۹	۱,۲۱E+۱۷	۴۹۶۸۰۵۷۸۴۸	۶۲۱۰۱۲۷۵۴	۰:۰۴۷:۵۰.۴۰۶
۲۰	۲,۴۳E+۱۸	۳۹۰۰۲۹۱۸۸۸۴	۴۸۷۸۶۶۸۰۸	۰:۰۵۷:۳۳.۱۸

میزان فضایی را که این الگوریتم جهت اجرا اشغال می نماید، ممکن است حافظه را پر کرده و خطای سربیز اتفاق بیفتد. این امر موجب توقف الگوریتم خواهد شد. در نتیجه برای حل این گونه مسائل، تکنیک‌هایی نیاز است که بتوانند یک پاسخ را در زمانی معقول، بدست آورند. تفاوت میان روش پسگرد و روش جستجوی محلی، این است که در روش اول از جستجوی نا‌آگاهانه و در مورد دوم، از جستجوی آگاهانه استفاده می گردد.

۴. روش جستجوی محلی

در علم کامپیوتر، جستجوی محلی یک روش فرا اکتشافی جهت حل مسائل بهینه سازی ساخت، بصورت محاسباتی می باشد. جستجوی محلی، بیشتر در مسائلی مورد استفاده قرار می گیرد که بتواند به عنوان راه حلی برای بالا بردن یک معیار در میان تعدادی از راه حل‌های پیش رو، مطرح شود. الگوریتم‌های جستجوی محلی در فضایی از راه حل‌های پیش رو (فضایی جستجو)، از یک راه حل به راه حل دیگر با استفاده از تغییرات محدود حرکت می کنند تا یک پاسخ به نظر مطلوب یافت شود و یا زمانی خاص سپری گردد و در این راه، تنها بخشی از فضایی حالت را مورد جستجو قرار می دهند. لذا در این روش‌ها رسیدن به پاسخ قطعی نبوده و به حالت شروع و همچنین نحوه جستجو در محدوده‌ی فضایی حالت بستگی دارد. با توجه به ماهیت تصادفی این روش‌ها، ارزیابی آنها با تکرار چندین مشاهده و بدست آوردن میانگین نتایج، قابل استناد خواهد بود.

۱.۴ تبرید شبیه سازی شده

تبرید شبیه سازی شده، یکی از روش‌های جستجوی محلی است که جهت یافتن نقطه‌ی بهینه ساخته شده، به کار می رود. این روش از صنعت ذوب فلزات برگرفته شده است و می تواند برای جستجو در یک

وزیر اول را در ستون و سطر اول قرار می دهیم. در ستون دوم، از دومین سطر تا سطر چهارم به دنبال خانه‌ای می گردیم که مورد تهدید وزیر اول نباشد و وزیر دوم را در آن خانه قرار می دهیم. به همین ترتیب این روند را ادامه می دهیم تا جایی که به پاسخ صحیح برسیم. در این میان اگر یک تهدید اتفاق بیفتد، به مرحله‌ی گذشته رجوع کرده و مهره‌ی موجود در آن ستون را به یک خانه‌ی دیگر که مورد تهدید مهره‌های قبلی نیست، انتقال می دهیم. همچنین اگر امکان انتقال مهره در یک مرحله قبل وجود نداشته باشد، می بایست به اولین ستونی بازگردیم که بتوان این عمل را انجام داد. پس از انتقال مهره، بطور مجدد روند اصلی را از آن قسمت به بعد ادامه می دهیم تا پاسخ حاصل شود.

شکل (۲) یک جستجوی پسگرد برای این مسئله را نمایش می دهد. در درختواره‌ی حاصل از الگوریتم پسگرد برای مسئله‌ی چند وزیر، یک گره، بیان گر حالتی از چیدمان مسئله می باشد. گرهی والد، گره‌ای است که در ابتدا بصورت فضای خالی ایجاد می گردد. گره‌های فرزند، با افزودن یک وزیر به هر گره، توسط گره‌های والد بوجود می آیند. گرهی امید بخش، گره‌ای است که بطور احتمالی، می تواند یک پاسخ باشد. یک گره نیز، غیر امید بخش است هرگاه هنگام ملاقات آن، مشخص شود که بطور احتمالی منجر به حل نمی گردد. پسگرد روندی است که توسط آن، پس از رسیدن به گره غیر امید بخش، به گره والد بازمی گردد و جستجوی اول عمق را روی فرزند دیگری ادامه می دهد. این فرایند در جستجو، هرس نام دارد [۱]. شکل (۳)، شبه کد این روش را نمایش می دهد.

```
function Backtracking(Q [1 .. n], r)
inputs: Q[1 .. n] , array for n queen
       r, recursive counter
local variables: legal, checking solution
if r = n+1
    Return Q
else
    for j ← 1 to n
        legal ← TRUE
        for i ← 1 to r - 1
            if (Q[i] = j) or (Q[i] = r - j + i) or (Q[i] = r + i - j)
                legal ← FALSE
        if legal
            Q[r] ← j
            Backtracking (Q[1 .. n], r + 1)
```

شکل ۳: شبه کد روش پسگرد

۱-۳ بررسی نتایج روش پسگرد

بررسی نتایج طبق جدول (۱)، روی ۲۰ فضایی حالت از مسئله، حاصل گردید که رشد نمائی این روش را به وضوح نشان می دهد. در این روش، اگر دو نمونه با مقدار n یکسان داشته باشیم، الگوریتم برای یافتن یک پاسخ، ممکن است برای یکی از آن‌ها تنها چند گره و برای دیگری، تمام درخت فضایی حالت را جستجو نماید [۱]. این روش همچنین، تمامی گره‌های فضایی مسئله را جستجو خواهد نمود که دارای فضایی برابر با $n!$ می باشد. با توجه به این نتایج، برای فضایی حالات بسیار بزرگ نمی توان از این روش استفاده نمود. زیرا ممکن است گاه تا چندین ماه یا سال به طول انجامد. همچنین با افزایش ابعاد مسئله،

شبیه کد این روش برای مسئله‌ی چند وزیر در قالب شکل(۵)، نمایش داده شده است.

function Simulated-Annealing (Problem , T_0 , α , I_{\max})

Inputs: problem , a problem

T_0 , a "temperature" for first step

α , Cooling ratio for mapping "temperature"

I_{\max} , maximum iteration

Local variables: current, a node

next, a node

T , a "temperature" controlling prob. of downward steps

current \leftarrow make-node(Initial-state[problem])

while ($i \leq I_{\max}$)

$T = \alpha T_0$

 next \leftarrow a randomly selected successor of current

$\Delta \leftarrow \text{fitness}[\text{next}] - \text{fitness}[\text{current}]$

 if $\Delta > 0$ then current \leftarrow next

 else current \leftarrow next only with probability $e^{\Delta/T}$

end

return current

شکل ۵: شبیه کد روش جستجوی محلی تبرید شبیه سازی شده

۱.۱.۴ برورسی نتایج روش تبرید شبیه سازی شده

طبق جدول(۳)، روی ۲۰ فضایی حالت از مسئله و همچنین ۱۰۰ بار مشاهده در هریک (جز حالات ۱، ۲، ۳ و -۳ وزیر) از آنها، حاصل گردید که با $T=100$ و $\alpha=0.9$ نجام گرفته است. این روش در هر دوره‌ی تکرار، تنها یک گره را مشاهده می‌نماید. در نتیجه زمان اندرکی صرف پردازش روی یک دوره خواهد شد.

فضایی حالت، استفاده گردد. در روش مذکور، ابتدا فلز در دمای بسیار بالا ذوب می‌شود، سپس دما به آرامی کاهش یافته تا به صفر برسد. هرچه درجه حرارت بالاتر باشد، جنبش مولکولی و در پی آن، انعطاف پذیری بیشتری برای فلز ایجاد می‌گردد که با کاهش دما، از قابلیت انعطاف آن کاسته شده و در نهایت هنگامی که دما به صفر برسد، فرم فلز دیگر تغییر نخواهد کرد. این الگوریتم، یک نقطه‌ی تصادفی را به عنوان وضعیت شروع در نظر گرفته، سپس مراحلی را بدین شرح تکرار می‌کند:

یکی از همسایه‌های وضعیت فعلی را بطور تصادفی انتخاب می‌نماید. اگر تابع برآزنده‌ی آن، از نقطه‌ی فعلی مناسب‌تر بود، آنگاه به همسایه‌ی انتخابی تغییر وضعیت می‌دهد. در غیر این صورت، با احتمال $e^{-\Delta/T}$ ، به آن تغییر وضعیت خواهد داد. تنشان دهنده‌ی دما و اختلاف سطح انرژی است که توسط رابطه‌ی (۲) محاسبه می‌گردد.

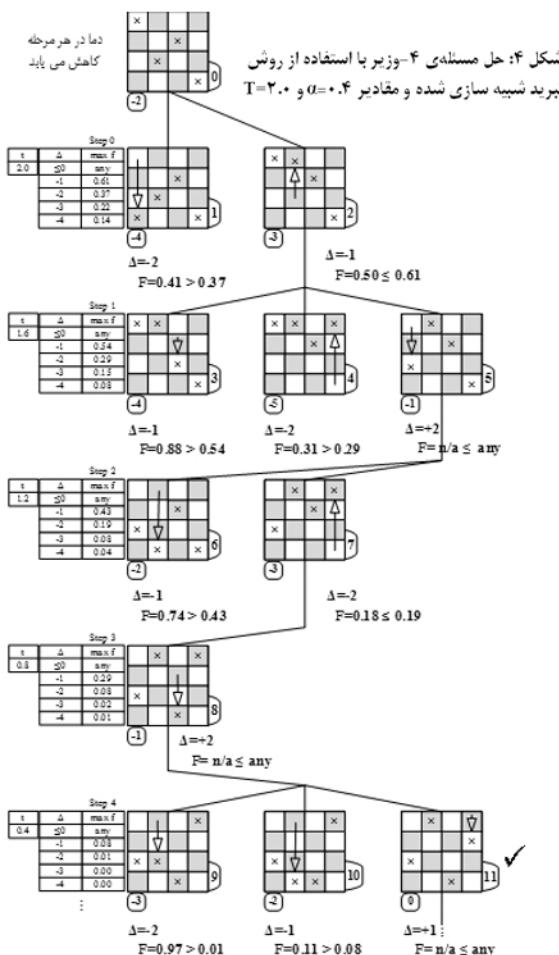
$$\Delta = \text{Value}_{\text{همسایه}}[\text{نقطه فعلی}] - \text{Value}_{\text{همسایه}}[\text{نقطه فعلی}] \quad (2)$$

همچنین مقدار T در زمان اجرای الگوریتم، توسط نرخ سرد شدن^{۱۰} که با نامگذاری کرده‌ایم، کاهش می‌یابد. هنگامی که دما بالا است، مقدار نیز $e^{-\Delta/T}$ بوده و الگوریتم با احتمال بیشتری به نقاط بدتر از نقطه‌ی فعلی « $\Delta < 0$ » بود. اما با پائین آمدن دما، این مقدار کاهش می‌یابد و در نهایت زمانی که T برابر با صفر شود، الگوریتم متوقف شده و نقطه فعلی را به عنوان پاسخ، باز می‌گرداند. انعطاف پذیری این روش، بطور معمول به α بستگی دارد. نحوی استفاده از این الگوریتم برای مسئله‌ی ۴ وزیر، در قالب شکل(۴)، نمایش داده شده است که طبق آن، دما در هر مرتبه α ، کاهش یافته و با احتمال $e^{-\Delta/T}$ یکی از همسایه‌ها انتخاب می‌گردد. این روند تا هنگامی که پاسخی بدست $e^{-\Delta/T}$ نیامده و همچنین مقدار T ، غیر صفر باشد، ادامه خواهد یافت.

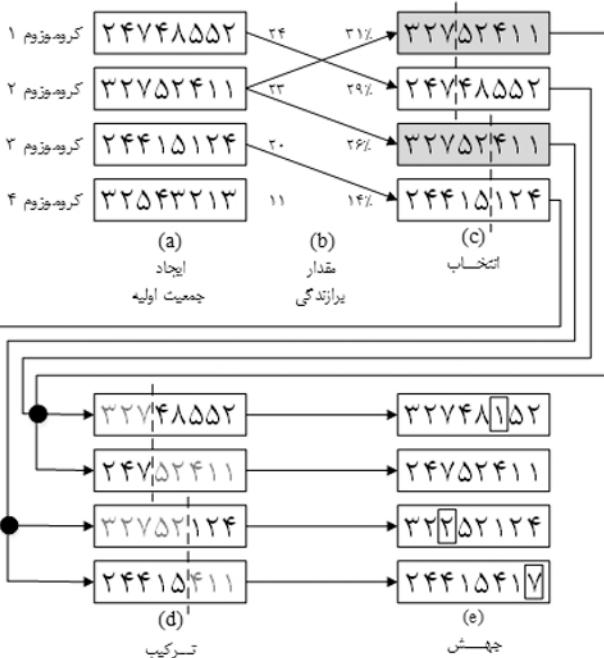
n	زمان اجرای نمونه (m:s.ms)	میانگین تعداد دوره‌ها	کمترین دوره‌ها برای ۱۰۰ نمونه	بیشترین دوره‌ها برای ۱۰۰ نمونه
۱	۰۰:۰۰:۰۰	۱	۱	۱
۲				
۳				
۴	۰:۰۰:۰۰	۱۴	۱	۳۱
۵	۰:۰۰:۰۰	۱۱	۱	۳۷
۶	۰:۰۰:۰۰	۲۰۸	۱	۹۷۰
۷	۰:۰۰:۰۰	۱۵۵	۸	۱۶۵۵
۸	۰:۰۰:۰۰	۵۱۱	۷	۲۷۸۰
۹	۰:۰۰:۰۰	۱۰۲۸	۱۲	۷۰۴۸
۱۰	۰:۰۰:۰۱	۵۶۶۶	۶	۲۹۷۲۹
۱۱	۰:۰۰:۰۳۴	۱۴۹۴۳	۲۰۷	۷۵۵۶۱
۱۲	۰:۰۰:۰۹۷	۳۶۷۱۹	۴۴۵	۱۴۶۰۹۹
۱۳	۰:۰۰:۰۳۱۵	۱۰۵۲۶۸	۴۰۸	۵۷۵۰۷۹
۱۴	۰:۰۰:۰۸۴	۲۶۹۸۲۸	۱۱۶۱	۱۲۳۷۶۷۷
۱۵	۰:۰۲:۱۱	۵۵۹۱۱۹	۷۶۹	۳۳۵۹۰۴۱
۱۶	۰:۰۵:۸۸۶	۱۴۹۱۳۵۶	۲۶۷۰	۶۱۱۰۶۹۵
۱۷	۰:۰۱:۵۲۲۰	۳۵۸۶۹۵۰	۵۳۱۱	۲۱۲۵۶۵۲۸
۱۸	۰:۰۵:۲۳	۱۱۰۱۵۳۹۹	۱۷۴۵۴	۴۵۸۵۷۷۸
۱۹	۲:۱۹:۱۴۱	۲۵۵۴۹۵۳۰	۵۲۶۴۰۲	۹۶۹۹۶۵۷۳
۲۰	۶:۵۵:۵۴	۷۵۶۳۷۸۵۶	۵۶۱۸۳۷۱	۳۰۲۲۶۶۳۱۷

جدول ۲: نتایج بدست آمده برای مسئله‌ی چند وزیر در روش تبرید شبیه سازی شده با فضایی حالات متفاوت؛ این نمونه‌ها با $T=100$ و $\alpha=0.9$ بدست آمده است

10) Cooling ratio



(الف) برازنده‌گی تمام اعضای جمعیت را جمع کنید و نتیجه را برازنده‌گی کل بنامید. ب) عدد n را به صورت تصادفی تولید کنید. بطوری که عددی بین صفر و برازنده‌گی کل باشد. ج) اولین عضو از جمعیت را که برازنده‌گی آن به اضافه‌ی برازنده‌گی اعضای پیشین جمعیت، بزرگتر یا مساوی n باشد را باز گردانید[۶].



شکل ۷: روند ایجاد جمیعت جدید در مسئله‌ی آوزیر توسط الگوریتم زنگیک با جمیعت اولیه‌ی ۴

٢٠٢٤ء، نتائج الگو، نتیجہ ڈنستک

بررسی نتایج طبق جدول(۳)، روی ۲۰ فضای حالت از مسئله و همچنین ۱۰۰ بار مشاهده در هریک (جز حالات ۱، ۲ و ۳ وزیر) از آن‌ها، حاصل گردید که با جمعیتی برابر با ۱۰۰ کروموزوم، $p_c = 0.1$ و $p_m = 0.1$ روش در هر دوره تکرار، حداقل به تعداد $n \times 100$ گره را مشاهده می‌نماید.

ن	میانگین زمان اجرای ۱۰۰ نمونه (m:s.ms)	میانگین تعداد دوره ها	کمترین دوره ها برای ۱۰۰ نمونه	بیشترین دوره ها برای ۱۰۰ نمونه
۱	۰:۰۰:۰۰۰	۱	۱	۱
۲				
۳				
			بدون پاسخ	
۴	۰:۰۰:۰۳۰	۱	۱	۱
۵	۰:۰۰:۰۸۰	۱	۱	۱
۶	۰:۰۰:۰۱۶۰	۱	۱	۵
۷	۰:۰۰:۰۵۰	۲	۱	۴
۸	۰:۰۰:۰۹۰	۳	۱	۱۳
۹	۰:۰۰:۰۷۰	۷	۱	۳۲
۱۰	۰:۰۰:۷۸۴۰	۸	۱	۴۹
۱۱	۰:۰۱۲:۰۳۰	۱۳	۳	۸۰
۱۲	۰:۰۱۹:۸۴۰	۱۷	۵	۱۱۵
۱۳	۰:۰۲۲:۳۸۰	۲۱	۱۳	۱۲۸
۱۴	۰:۰۴۷:۳۵۰	۵۳	۱۴	۱۷۴
۱۵	۰:۰۵:۰۸۲۰	۹۲	۱۸	۱۹۰
۱۶	۱:۰۰:۰۶۴۰	۱۱۳	۲۸	۲۰۳
۱۷	۱:۰۷:۰۳۲۰	۱۴۲	۴۹	۲۲۸
۱۸	۲:۰۳:۰۲۹۰	۱۹۶	۵۵	۳۰۶
۱۹	۳:۰۸:۰۳۶۰	۲۵۳	۶۴	۴۷۵
۲۰	۴:۰۷:۰۱۲۰	۳۰۷	۸۶	۵۴۰

جدول ۳ : نتایج بدست آمده برای مسئله‌ی چند وزیر در روش الگوریتم زنگنه با فضای حالات متفاوت؛ این نمونه‌ها با جمعیتی برابر با 10^6 کروموزوم، $p_m = 0.1$ و $p_c = 0.8$ مده است

٤-٢ الگو، یتم و نتیک

الگوريتم ژنتيک، يك الگوريتم جستجو و بهينه‌سازی است که بر اساس اينde
ى تکامل طبیعی بنا نهاده شده است. طبق اين اиде، در هر جامعه، بطور
معمول افراد قوی تر، می‌توانند بهره‌ی بیشتری از منابع طبیعی پيروان خود
دریافت کنند و با احتمال بیشتری زنده می‌مانند. اما افراد ضعیفتر نیز با
احتمال کمتری می‌توانند زنده بمانند. با استفاده از اين شيوه، می‌توان برخی از
مسائل در حوزه‌ی بهينه‌سازی که دارای يچيدگی زيادي هستند را پاسخ
گفت. فضای مسئله در اين روش، جمعیت آن می‌باشد که از کروموزوم‌ها
تشکيل شده است. اين کروموزوم‌ها نيز از تعداد مشخصی ژن بوجود آمداند
که در هر دوره، نسل جديدي توسيط آنها ايجاد مي‌گردد. بدويهي است که
هر چه تابع برازنده‌ی کرموزوم بهتر باشد، احتمال انتخاب آنها نيز بيشتر خواهد
شد. کرموزوم‌های انتخاب شده در هر دوره، بطور احتمالي با يكديگر ترکيب
شده و همچنین جهش می‌يابند تا در نهايت جمیت جديد را ايجاد نمایند.
شكل(۶) روند انجام الگوريتم ژنتيک را توصيف مي‌کند[۶].

بکارگیری این شیوه در مسئله‌ی چند وزیر، بدین صورت خواهد بود که ابتدا جمعیتی متشکل از تعدادی کروموزوم که هریک حاوی n ژن می‌باشد، پدید می‌آید. سپس تابع برازنده‌ی هریک محاسبه و احتمال انتخاب آنها بدست می‌آید. در نهایت پس از انتخاب کروموزوم‌ها برای ایجاد نسل جدید، عمل ترکیب و چesh بر روی آنها و بصورت احتمالی انجام خواهد پذیرفت. شکل(۷)، نحوه‌ی اجرای الگوریتم زتیک روی مسئله‌ی -وزیر را نمایش می‌دهد.

۱.۲.۴ روش انتخاب کروموزوم ها

روش‌های متفاوتی در انتخاب کروموزوم‌ها استفاده می‌شود که در این مقاله، از چرخ گردان بهره‌گرفته شده است. این روش شامل مراحلی بدین شرح است:



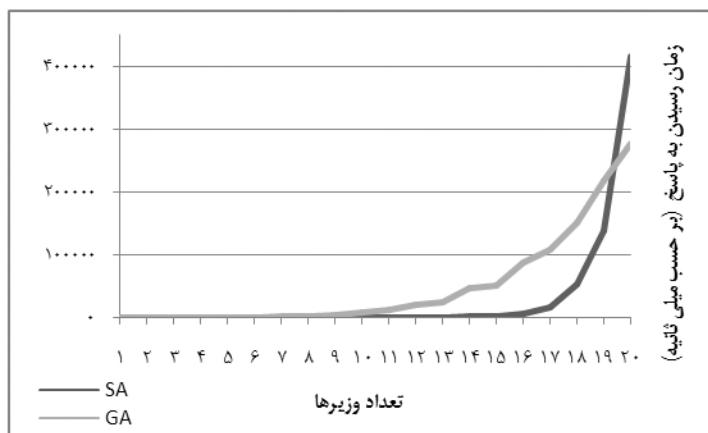
شکل ۶: فلوچارت روند اجرای الگوریتم ژنتیک

۷. منابع

- [۱] نیویلیتان، ریچارد، نعیمی پور، کیومرث، طراحی الگوریتم‌ها، ترجمه‌ی جعفرنژاد قمی، عین‌الله، ویراسته‌ی جعفرنژاد قمی، عین‌الله، ویرایش سوم، بابل، نشر علوم رایانه، تابستان ۱۳۸۶.
- [۲] Martinjak, I.; Golub, M., "Comparison of Heuristic Algorithms for the N-Queen Problem", Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on , vol., no., pp.759,764, 25-28 June 2007.
- [۳] Bozikovic, M.; Golub, M.; Budin, L., "Solving n-Queen problem using global parallel genetic algorithm", EUROCON 2003. Computer as a Tool. The IEEE Region 8 , vol.2, no., pp.104,107 vol.2, 22-24 Sept. 2003.
- [۴] Azhar W. Hammad, Dr. Ban N. Thanoon. "Genetic algorithm versus particle swarm optimization in n-Queen problem", Journal of Al-Nahrain University, Vol.10(2), pp.172-177, 2007.
- [۵] Turky, A.M.; Ahmad, A. "Using genetic algorithm for solving N-Queens problem", Information Technology (ITSim), 2010 International Symposium in, On page(s): 745 - 747 Volume: 2, 15-17 June 2010.
- [۶] Michael Negnevitsky, Artificial Intelligence A Guide to Intelligent Systems, Second edition, Person Education publishing, 2005.

۵. مقایسه روش‌های جستجوی محلی ارائه شده

نتایج بدست آمده در دو روش تبرید شبیه سازی شده و الگوریتم ژنتیک، نسبت به زمان، در قالب نمودار شکل (۸) قابل نمایش می‌باشد که تنها با مدل‌های ارائه شده، بدین صورت حاصل گردیده است. همچنین از مقایسه‌ی تعداد دوره‌ها در این دو روش، به دلیل اختلاف بسیار زیاد بین میانگین تعداد دوره‌ها، صرف نظر شده است.



شکل ۸: نمودار مقایسه‌ی روش‌های تبرید شبیه سازی شده و الگوریتم ژنتیک در حل مسئله‌ی چند وزیر

۶. نتیجه

طبق نتایج بدست آمده، روش پسگرد در تعداد وزیرهای بالا، به دلیل پیمایش تمامی گره‌های درخت فضای جستجو، زمان بسیار زیادی را سپری نموده و همچنین حجم بزرگی از حافظه را اشغال می‌نماید. بنابراین استفاده از آن، در حل این مسئله، و همچنین مسائلی از این دسته، کارآمد نمی‌باشد. در نتیجه می‌توان از روش‌های جستجوی محلی که پاسخی تقریبی اما با کارایی قابل قبول را بازمی‌گردانند، بهره گرفت. استفاده از واژه‌ی تقریبی در این قسمت به دلیل آن است که این نوع الگوریتم‌ها بطور غیرقطعی عمل می‌کنند، اما در موارد زیادی پاسخ را تولید می‌نمایند. این نوع روش‌ها، فضای بسیار کوچکی از حافظه را نیز اشغال می‌نمایند. در بررسی دو روش جستجوی محلی تبرید شبیه سازی شده و الگوریتم ژنتیک نیز الگوریتم ژنتیک نتایج مناسب‌تری را نسبت به روش دیگر، نمایش می‌دهد که این قیاس، تنها در مدل‌های ارائه شده‌ی این مقاله قابل استناد خواهد بود. زیرا مدل‌های بسیار زیادی را می‌توان توسط پارامترهای متغیر موجود در این دو روش ارائه داد. اما بهترین مدل‌های انتخابی در این مقاله، الگوریتم ژنتیک را در حل مسئله‌ی چند وزیر، پیشنهاد می‌نماید که با جمعیت ۱۰۰ کروموزوم، احتمال ترکیب ۰.۸ و احتمال جهشی برابر با ۰.۰۵ بدست آمده است.

مدل سازی انتشار اطلاعات و عضویت در اجتماع با استفاده از بهینه سازی احتمالی

Modeling Information Diffusion and Community Membership using Stochastic Optimization

Alireza Hajibagheri, Ali Hamzeh, Gita Sukthankar

IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2013

ترجمه: وحید حیدری پور، دانشجوی کارشناسی ارشد علوم کامپیوتر، دانشکده علوم اقتصادی

چکیدہ:

اجتمعات را به عنوان چرخهای کارا در انتشار اخبار، شایعات و عقاید در شبکه‌های اجتماعی انسانی می‌شناسیم. مدلسازی انتشار اطلاعات از طریق یک شبکه مارا قادر خواهد ساخت که به فهم برتر تابعی از تأثیرات ساختار شبکه مانند اجتماعات بر انتشار اطلاعات دست یابیم. فرض اولیه در اینجا این خواهد بود که فرم که از تابع پیروی می‌کند-بازیکن‌های عقلائی ساز و کارهای انتخاب اجتماعی را سیک و سنجین می‌کنند و بسته به میزان اطلاعاتی که در اختیار آن‌ها در اجتماعات مختلف قرار می‌گیرد به اجتماعات مختلف می‌پیوندند. بهینه‌سازی از دحام ذرات^۱ در اصل برای شبیه‌سازی مجموع رفتار اجتماع طراحی شده بود؛ در مدل انتشار مطرح شده‌ی ما،^۲ جریان اطلاعات در یک شبکه را با ساخت اجتماعات ذرات برای همسایگان محلی شبکه که یک مدل پیوسته از توابع هدف^۳ را بهینه می‌کنند مدلسازی می‌کنند. در این مقاله، نشان خواهیم داد که چطور راهکار ما با دیگر کارهای مدلسازی پیشگام در این زمینه تفاوت دارد و خواهیم دید که روش ما در چندین مجموعه داده از شبکه‌های اجتماعی از مدل‌های برپایه مدلسازی تشخیص اجتماع موجود بهرتر است.

۱. مقدمه

انتشار اطلاعات، به دلیل اهمیتش در زندگی روزانه به یک مسئله کلیدی پژوهش در اطلاعات اجتماعی بدل شده است. فیزیک انتشار اطلاعات با اتخاذ ایده از اینترنت و جهان پنهان و ب تغییر کرده است. تا چند سال قبل، مشکل اصلی کسی که میخواست اطلاعاتی را در یک اجتماع پخش کند هزینه ای بود که باید برای سیستم پایه ای فنی مورد نیاز به منظور دستیابی به تعداد زیادی از افراد پرداخت می کرد. امروزه، با دسترسی وسیع به اینترنت، این گلوگاه تا حد زیادی مرتقق شده است. اخیراً علاقه و توجه نه تنها به مشاهده جریان اطلاعات و اکتشافات، بلکه در ساخت، تأثیرگذاری و شبیهسازی آن وجود دارد.

مدلسازی انتشار اطلاعات یک بخش کلیدی برای آدرس دهی و پاسخ دادن به سؤال پژوهشی زیر است:

۱. بیشینه‌سازی گستره تأثیرگذاری از طریق عقاید، ایده‌ها و پیشنهادات؛
 ۲. تحت کنترل داشتن سریع غذا و بیماری.

در این مقاله، ما از انتشار اطلاعات برای فهمیدن درمورد مسئله سومی استفاده کردیم: مشخص کردن عضویت در اجتماعات در

1) Particle Swarm Optimization

- 2) Particle Swarm Optimization Diffusion Model

3) Holland's hyperplane-defined

4) Agent

۲. پیشینه

مجموع رفتارهای اجتماعی مورد استفاده قرار گرفت) روش جدیدی برای مدلسازی سازوکارهای انتشار اطلاعات در شبکه‌های اجتماعی ارائه دادیم. استفاده از PSO به ما این اجازه را داد که به جای استفاده از رشته دودویی برای ذخیره سازی اطلاعات از بردارهای با مقادیر واقعی نزدیک تر به واقعیت استفاده کنیم، هر گره شبکه اجتماعی موردنظر به شکل یک ذره در مجموعه ذرات دیده می‌شود؛ حالت اطلاعات کنونی هر گره با موقعیت ذره که به شکل یک بردار با مقادیر حقیقی کدشده ارائه می‌شود. اطلاعات تبادلی با محاسبه مقدار مرتبط از اطلاعات یک گره در مقایسه با همسایگان محلی با استفاده از یک HDF تغییر یافته مدل می‌شوند. این تغییر در مدلسازی انتشار اطلاعات الگوریتم تشخیص اجتماع ما را قادر خواهد ساخت که از روش‌های تشخیص اجتماع برپایه مدل موجود، شامل استفاده از GADM عملکرد بهتری داشته باشد.

۳. کارهای مرتبط

تشخیص اجتماعات در اطلاعات جهان واقعی مانند شبکه‌های اجتماعی، شبکه جهانی وب و شبکه‌های بیولوژیکی یک مسئله مهم است که در سالی‌های اخیر توجه زیادی را به خود جلب کرده است. یک اجتماع شبکه، که گاماها یک خوش^{۱۳} نامیده می‌شود، به عنوان یک گروه از گره‌ها شناخته می‌شوند که درین خودشان به نسبت با دیگر گره‌های موجود در شبکه تعاملات بیشتری دارند.

تشخیص اجتماع یک ابزار ارزشمند برای درک بهتر سازوکار شبکه‌های پیچیده برایمان دارد. گراف‌های سیاری در جهان واقعی به صورت طبیعی به اجتماعاتی که در آن گره‌ها با چگالی بالایی به هم متصل هستند شکسته تجزیه می‌شوند که این اجتماعات با یکدیگر اتصالات سیار تنکتری دارند. اجتماعات در شبکه‌های اجتماعی بزرگ ارزش علمی و عملی بالایی دارند چرا که معمولاً به بخش‌های کارای شبکه مرتبط می‌شوند، مانند گروه‌های اجتماعی یا واحدهای پروتئین.

تشخیص اجتماع برپایه مدل، دسته‌بندی گسترده‌ای از روش‌ها که عمل پردازش پویا با استفاده از شبکه را بررسی می‌کنند یا مدل سازنده از چیدمان شبکه، را بیان می‌کند. مثال‌هایی از پردازش‌های پویا، انتشار برچسب، جریان اطلاعات که به الگوریتم خوش‌بندی مارکف^{۱۴} شناخته می‌شود و مدل چرخشی محبوب است.

۴. روش

در این بخش از مقاله، توضیحات و فرمول ریاضی برای مدل انتشار اطلاعات برپایه PSO پیشنهادی آمده است. جریان‌های اطلاعاتی که با این روش انتشار نمایان می‌شوند به عنوان ورودی به الگوریتم نظریه بازی‌ای برای تشخیص اجتماع داده می‌شود^{۱۵}. GPSODM. پیاده‌سازی این الگوریتم‌ها برای استفاده عموم در سایت <http://www.eecs.ucf.edu/~alireza/psodm> قرار دارد.

انتشار اطلاعات عموماً به عنوان پیش‌نیاز برای بهینه‌سازی تأثیرگذاری مطالعه می‌شود. مسئله به این شکل تعریف می‌شود، انتخاب مجموعه k عضوی از گره‌ها به هدف فعالیت‌های اولیه که بزرگترین گستره مورد انتظار اطلاعات را پیسازد، که α با بودجه‌ای محدود می‌شود. برای مطالعه پیشینه سازی تأثیرگذاری به یک فهم عمیق درمورد تأثیر سطح ریز و سطح درشت^۱ سازوکارهای انتشار اطلاعات نیازمندیم. که به نوبه خود بروی مدلسازی و پیش‌بینی انتشار اطلاعات در شبکه‌های اجتماعی تمرکز دارد. این مدل‌های انتشار اغلب از تحقیقات در زمینه‌های مختلف شامل همه‌گیر شدن، جامعه شناسی، بازاریابی و فیزیک برمی‌آیند.

دو مدل بسیار پر استفاده در انتشار، مدل آستانه خطی^۲، که فرض می‌کند که یک گره فعال می‌شود اگر تعداد گره‌های فعال در همسایگی آن از آستانه شخصی بالاتر باشد، و مدل به جریان افتادن مستقل^۳، که فرض می‌کند زمانی که یک گره فعال می‌شود هر کدام از همسایگانش احتمالی برای فعال شدن دارند، که سبب جریان فعالیت برای حرکت در شبکه می‌شود. علاوه بر ارائه مدل‌ها برای پردازش انتشار، این تکنیک می‌تواند برای کارهای مانند شناسایی عقیده پیشرو استفاده شود. اخیراً Lahiriri و Cerbian نشان دادند که مدل‌های محبوب می‌توانند به چشم مدل‌های خاصی از مدل انتشار الگوریتم ژنتیک^۴ دیده شوند.

فرض اصلی در GADM این است که اشخاص در شبکه به هدف افزایش ارزش اطلاعات شخصی خودشان که با یک تابع هدف عمومی مشخص می‌شود با هم ارتباط دارند. نکته ای که نگارنده به عنوان یک تقاضه کلیدی با مدل‌های قبلي مورد توجه قرار داده است این است که آنها می‌توانند چندین بخش از اطلاعات را با استفاده از HDF^۵ متشکل از چندین شما مدلسازی کنند. هر گره در شبکه با یک رشته دودویی که اطلاعات شخصی را مشخص می‌کند مقداردهی اولیه می‌شود. یک HDF برای پارس کردن رشته‌های دودویی و دادن امتیازات از چندین شما تشکیل شده است فیتنس^۶ کلی HDF از مجموع امتیازات شماها محاسبه می‌شود. تعاملات اجتماعی از طریق یک تقاطع جابه‌جاوی انتهای^۷ که گره‌ها زیررشته‌ها را تبادل می‌کنند شبیه‌سازی شده است. رشته‌های تولید شده به شرطی که امتیاز HDF بالاتری نسبت به رشته‌های کنونی داشته باشند نگه داشته می‌شوند. Lahiriri و Cerbian نشان داده‌اند که هدف میانگین نرمال شده نهایی^۸ از هر گره شبکه که با استفاده از GADM محاسبه می‌شود به راحتی با استفاده از خصایص شبکه مانند درجه قابل پیش‌بینی نیست.

اگرچه که GADM یه مدل بسیار عمومی است، با این حال در انتشار اطلاعات کافی نیست. عمل تقاطع که بخش جدایی ناپذیر آن است موجب از دست رفتن اطلاعات می‌شود و تعاملات اجتماعی همیشه به شکل جفتی فرض می‌شوند در حالی که باید برپایه همسایگی محلی باشند. با توجه به همه خوبی‌های GADM، ما با استفاده از تکنیک بهینه‌سازی احتمالی(بهینه‌سازی حرکت ذرات، که با موفقیت برای مدلسازی تأثیر

5) Micro and macro-level

6) Linear Threshold Model (LTM)

7) Independent Cascade Model

8) Genetic Algorithm Diffusion Model (GADM)

9) Hyperplane-defined function

10) Fitness

11) Tail-swap cross-over

12) Average normalized objective (ANO)

13) Cluster

14) Markov Cluster Algorithm

15) Game + Particle Swarm Optimization Diffusion Model

تأثیر کلی الگوریتم در حرکت هر گره به سمت برداری با ارزشمندترین اطلاعات در شبکه محلی است. بردار اطلاعات شخصی گره v رشته S_v است که شامل مقادیر $[0, 1]$ است و به شکل زیر مقداردهی می‌شود:

$$S_v[i] = \text{random}(0,1)$$

برای مثال، جدول زیر بردار حالت گره با طول $\beta = 10$ را نشان میدهد:

1.0	0.8	0.5	0.0	0.2	0.7	0.2	0.6	0.6	0.9
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

درطی شبیه سازی سازوکار انتشار اطلاعات، گره ها تعامل دارند و سعی میکنند تا سطح اطلاعاتی خودشان را بیشینه کنند. روند تعامل و به روزرسانی اطلاعات برپایه الگوریتم PSO است. در بک الگوریتم PSO نرمال، سرعت و موقعیت ذرات مکرراً به روز میشود. اگرچه، در الگوریتم بیان شده، موقعیت ذرات تنها زمانی به روز میشود که یک یال جدید در زمان t داشته باشیم، براساس داده رسانه اجتماعی، که مشاهده مشبود.

مشابه با GADM ها، HDF برای تولید شمای اطلاعاتی استفاده میشوند. یک HDF با تعریف شماها (زیرنویس کوتاه با کلمات که از موقعیت مشخصی شروع میشود) به شکل تصادفی و سلسه مراتبی، که با شمای نسبتاً کوتاه از درجه یک که نقطه شروعش به شکل تصادفی در رشته اتفاق میافتد ساخته میشود. جفت ها از یک شمای با هم الحق میشوند تا شما از درجه دو را بسازند، و هر شما یک امتیاز مثبت یا منفی دریافت میکنند که به شکل تصادفی از یک بازه خواهد بود. مشابهای شماها از درجات بالاتر با همین ساز و کار قابل ساختن هستند. HDF ها یک رشته دو دو یی به عنوان ورودی دریافت میکنند و یک امتیاز هدف باز میگردانند که این امتیاز مجموع همه امتیازات شماها موجود در داخل آن است. به وضوح ما فقط میتوانیم از حالت اصلی HDF ها برای تولید شماهای اطلاعاتی استفاده کنیم نمایش بردار حالت پیوسته به یکتابع هدف جدید نیاز دارد. به منظور یافتن شباهت بین بردار حالت هر گره و شمای اطلاعاتی کلی گام های زیر را اجرا کردیم:
 ۱. تفاوت بین هر عنصر از بردار حالت با معادلش در الگوی تولید شده با HDF را محاسبه کنید. به طور خاص، این تفاوت برای بیت X به شکل $|X - 0|$ تعریف میشود اگر بیت مرتبط در الگو صفر باشد و در غیر این صورت به سادگی مقدار X خواهد بود.

۲. بعد از محاسبه همه تفاوت ها بدون توجه به حالت صفر (وقتی که دو بیت مرتبه برابر هستند)، مقدار امتیاز را برای یک رشته وروdi W.R.T شمای داده شده محاسبه میکنیم. برای اینکار، از حاصلضرب تفاوت های محاسبه شده به عنوان وزن ω برای آن شما استفاده میکنیم. با ضرب ω و امتیاز الگوی داده شده، مقدار امتیاز S برای رشته وروdi مرتبه اینگونه خواهد بود:

$$S_j = \omega S_j$$

۱۶. الف. مدل انتشار حرکت بهینه ذرات

همانطور که پیشتر بحث شد، GADM به دلیل از دست دادن اطلاعات در عملگر تقطیع تک نقطه از تعاملات اجتماعی ناکامل است. برای همین، ما یک مدل از انتقال اطلاعات که تأثیر همه همسایگی‌های محلی بر به روز شدن اطلاعات شخصی یک مأمور را درنظر می‌گیرد معرفی می‌کیم. علاوه بر این، اطلاعات شخصی به شکل یک بردار با مقادیر حقیقی از دانسته‌ها برای مدلسازی، یک زنجیره از تخصص، نشان، داده خواهد شد.

یک شبکه اجتماعی یک گراف جهت دار یا بدون جهت است ($G=(V,E)$ ، که در آن یک رأس $v \in V$ به عنوان یک بازیگر در یک شبکه اجتماعی است و یک عامل بین این دو فرد را نشان می‌دهد. از طرف دیگر، یک یک شبکه اجتماعی پویا یک گراف چندگانه ($G=(V,E_t)$) است که در آن یک مجموعه یال‌ها و هر یال $(u,v) \in E_t$ به عنوان یک عامل بین دو فرد در زمان t است. یک مدل انتشار یک گراف ($G=(V,E)$) را قبول می‌کند، و به عنوان ورودی یک بردار حالت S_v برای هر رأس $v \in V$ می‌پذیرد. برایه حالت همه رئوس افزایشی، برای زمان $t+1$ خواهیم داشت S_v^{t+1} برای یک شبکه اجتماعی پویا، گراف در هر واحد زمانی به شکل ($G_t=(V,E_t)$) نشان داده می‌شود. که در آن $E_t = \{(u,v) : (u,v) \in E\}$ مجموعه یال‌ها در واحد زمانی t است.

یک شبکه اجتماعی پویا شامل مجموعه رئوس $V = \{v_1, \dots, v_n\}$ است که در دوره زمانی T با واحد زمانی گستته با یکدیگر در تعامل هستند. هر کدام ازین رئوس به عنوان یک ذره در الگوریتم PSO هستند که برای انتشار اطلاعات مورد استفاده قرار گرفته است. الگوریتم PSO استاندارد یک نسل از جوابهای کاندیدا را به عنوان ذرات نگه میدارد و ذرات را حوالی فضای جستجو حرکت میدهد تا یک تابع هدف را ارضاء کند. حرکت هریک از ذرات برایه موقعیت و سرعت ذره است. حرکت ذرات متأثر از دو پارامتر است: شناخته شده ترین موقعیت برای ذره مشخص شده($pbest$) و بهترین موقعیت شناخته شده به وسیله ازدحام ذرات($gbest$). سرعت و موقعیت ذرات با توجه به وابط زیر به دو میشوند:

$$v[i+1] = v[i] + (C_1 R_1 \times (pbest[i] - position[i])) + (C_2 R_2 \times (gbest[i] - position[i]))$$

و همچنین

$$position[i + 1] = position[i] + v[i + 1]$$

که در این روابط مقادیر C و R پارامترهای الگوریتم PSO هستند که برای کارایی، بهتر میتوان آنها را میزان کرد.

برای مدلسازی افراد در یک شبکه اجتماعی به شکل ذره در گروه، الگوریتم PSODM ما به این شکل خواهد بود:

۱. هر گره $v \in V$ در گراف $G = (V, E)$ را به عنوان یک ذره PSO در نظر بگیر.

۲. یک بردار حالت با طول β به هر گره به عنوان موقعیتش در PSO متناظر می‌کنیم.

۳. ذرات خودخواهانه سعی دارند تا حالت اطلاعاتی خودشان را بیشینه کنند. ازین رو برای هر ذره، p_{best} موقعیت فعلی ذره خواهد بود، که از بین برنده اولین ترم در رابطه به روز رسانی است. یک ذره در شبکه اجتماعی تنها اجازه دارد که با همسایگانش تعامل داشته باشد. نتیجتاً g_{best} برای هر ذره یکی از همسایگانش، خواهد بود.

۴. ب. تشخیص اجتماع: GPSODM کارهای گذشته نشان میدهد که ترکیب GADM با نظریه بازی میتواند یک چارچوب قوی برای شناسایی ساختار اجتماع یک شبکه باشد. هر گره از گراف به عنوان یک مأمور خودخواه با یک تابع مطلوبیت مدل میشود که با اطلاعاتی که با همسایگانش مبادله میکند محاسبه میشود. مأمورین انتخاب میکنند که به یک اجتماع ملحق شوند یا نه تا زمانی که به حالت تعادل برسند. تعادل نش^{۱۷} این بازی به ساختار اجتماع در شبکه مرتبط است. در این چارچوب، GADM برای بدست آوردن مقدار اطلاعات تبادل شده بین گره ها استفاده شده است که به نوبه خود به عنوان ورودی برای الگوریتم تشخیص اجتماع مورد استفاده قرار میگیرد.

برای ارزیابی کارایی مدل انتشار ما در تشخیص اجتماع، یک چارچوب مشابه ساختیم که از PSODM برای ساخت ورودی مورد نیاز الگوریتم تشخیص اجتماع GPSODM استفاده میکند(که در الگوریتم شماره ۲ نشان داده شده است). هر رأس از شبکه ما به عنوان یک مأمور خودخواه که مطلوبیت خودش را به شکل محلی بیشینه میکند مدل میشود. در ابتدا مقدار مطلوبیت برای همه مأمورین با مقدار صفر مقداردهی میشود. عضویت در اجتماع جریان اطلاعات یک مأمور را تعدیل میکند. هرگاه یک مأمور از نسل نمونه برداری شد، این شناس را خواهد داشت که به یک اجتماع جدید پیوندد، یکی از اجتماعات را ترک کند و یا از یک اجتماع به اجتماع دیگر نقل مکان کند، و همه اینها برایه مطلوبیت فعلی U است. هیچ یک از مأمورین دیگر حالت خودشان را در این دوره زمانی تغییر نمیدهد. اگر افزایش مطلوبیت امکان پذیر نباشد، حالت پیش فرض این است که مأمورین بی قرار هستند و پیوستن به اجتماع جدید را انتخاب میکنند. مطلوبیت هر مأمور، U به شکل زیر محاسبه میشود:

$$U_i(S) = \frac{1}{m} \sum_{j=1, j \neq i}^n I_{ij} \delta_{ij}$$

که $\sum_{j=1}^n I_{ij}$ مقدار اطلاعاتی است که مأمور i از مأمور j دریافت میکند، $\sum_{j=1}^n m$ برابر 1 است اگر مأمور i و مأمور j در یک اجتماع هستند. و همچنین n و m به ترتیب تعداد گره ها و یال ها را مشخص میکنند. در اینجا S مجموعه استراتژی ها برای همه مأمورین است. در این چارچوب، استراتژی بهترین پاسخ برای مأمور i با توجه به استراتژی (S_{-i}) برای ماقی مأمورین به شکل زیر محاسبه میشود:

$$\arg \max_{s'_i \subseteq [k]} U_i(S_{-i}, s'_i)$$

مجموعه همه اجتماعات شدنی شبکه با $[k] = 1, 2, \dots, n$ نشان داده میشود که k چندجمله ای در n است، اگرچه که تعداد اجتماع های نهایی ما ممکن است بسیار کوچک تر از k باشد. استراتژی پروفایل S یک تعادل نش اجتماع از چینش بازی زمانی که همه مأمورین بهترین بازی ممکن را انجام دهنده ارائه میکند. در تعادل نش هیچ مأموری نمیتواند مطلوبیت خودش را افزایش دهد؛ و به مطلوبیت کوتني قانع شده است:

و همین رویه را با توجه به الگوهای تولید شده از HDF تکرار میکنیم تا به امتیاز کل زیر بررسیم:

$$total-score = \sum_{i=1}^n S_i$$

کارایی PSODM خیلی به انتخاب تابع حساس نیست؛ به طور بالقوه دیگر دیگر معیارهای فاصله میتوانند برای محاسبه تابع هدف به کار گرفته شوند. به روز رسانی اطلاعات زمانی اتفاق می افتد که یک یال جدید در زمان t بررسد و به شکل زیر انجام میشود. ابتدا، سرعت ذرات را به روز میکنیم:

$$\begin{aligned} velocity_v^{t+1}[i] &= velocity_v^t[i] \\ &+ (C_1 R_1(pbest[i] - position_v^t[i])) \\ &+ (C_2 R_2(gbest[i] - position_v^t[i])) \end{aligned}$$

یادآوری میکنیم که pbest همیشه موقعیت فعلی هر گره را نشان میدهد، پس داریم:

$$velocity_v^{t+1}[i] = velocity_v^t[i] + (CR \times (gbest[i] - position_v^t[i]))$$

سپس برایه سرعت موقعیت هر گره را به روز میکنیم که سطح اطلاعاتی آن را نشان میدهد:

$$position_v^{t+1}[i] = position_v^t[i] + velocity_v^{t+1}[i]$$

هربار که یک جفت از افراد تعامل دارند، اطلاعات تبادل شده با الگوریتم گفته شده مدل‌سازی میشود. این رویه میتواند موجب افزایش در سطح اطلاعاتی یک گره شود. استفاده از PSO ما را قادر به مدل‌سازی دانش جزئی و اجتناب از از دست دادن اطلاعات میسازد که راست نمایی مدل را افزایش میدهد. الگوریتم ۱ PSODM را خلاصه میکند.

Algorithm 1 PSODM

- 1: **Input:** Initial state vector $S_v^{(t)}$ for node v
- 2: **Output:** New state vector $S_v^{(t+1)}$
- 3: $position_v^{(t)} = S_v^{(t)}$
- 4: $gBest = \emptyset$
- 5: Neighbors = all neighbors of v
- 6: **for** $u \in \text{Neighbors}$ **do**
- 7: **if** ($position_u^{(t)} > position_{gBest}$) **then**
- 8: $gBest \leftarrow u$.
- 9: **end if**
- 10: **end for**
- 11: Update $velocity_v^{(t+1)}$ based on (3)
- 12: Update $position_v^{(t+1)}$ based on (4)

این مجموعه اطلاعاتی به منظور شبیه سازی روند انتشار اطلاعات اعمال کردیم.

برایی با تولید بردار حالت تصادفی برای مقادیر اولیه دادن به هر گره شروع کردیم، با یک HDF تصادفی با طول دلخواه ادامه دادیم. بعد از اجرای PSODM بروی یالهای زماندار، مقدار نهایی هدف (سطح اطلاعات) همه گره ها را به نسبت بیشینه مقادار هدف در جمعیت ذرات نرمال کردیم. یک گره میتواند به مقدار هدف بالاتری دست یابد، اگر و تنها اگر ترکیب ارزشمندتری از واحدهای اطلاعاتی را در مقایسه با سایرین جمع کند. همانطور که در بحث شد، برای محاسبه بایاس آخر، این الگوریتم را با HDF های متفاوت و همینطور بردار حالت های تصادفی جدید برای هر گره اجرا کردیم. و مقدار متوسط نرمال شده هدف را برای هر گره محاسبه کردیم. این مقدار نشان دهنده ارتباط ارزش اطلاعات یک فرد، و یک ذره با مقدار متوسط نرمال شده هدف بالاتر دارای واحدهای اطلاعاتی با ارزش تر است. نتایج شکل ۱ توزیع تجمعی مقدار متوسط نرمال شده هدف همه افراد در شبکه را برای PSODM (بالایی) و GADM (پایینی) برای زمان مقدار دهی و نهایی نشان میدهد. مقادیر اولیه توزیع تجمعی مقدار متوسط نرمال شده هدف به شدت خوش بندی شده اند و پراکندگی کمی برای هر دو مدل دارند. همانطور که در ذکر شد، یک توزیع مقدار متوسط نرمال شده هدف نهایی مشابه انتظار میرود اگر همه افراد در ترم موقعیتشان در شبکه قابل مقایسه بودند، به عنوان مثال، هیچ گره ای نباید به طور مداوم با شروع از حالت های مختلف به مقدار بالای نرمال شده هدف به پایان برسد. اگرچه، از هر دو توزیع نهایی آشکار است که این مورد نیست. ما میتوانیم توزیع نهایی مقدار متوسط نرمال شده هدف ها را برای هر دو مدل بیینیم که شامل تعداد کمی از رئوس اند که به طور مداوم با بالاترین امتیاز شما، بدون توجه به حالت ابتدایی شان به پایان میرسند. این آینده شبکه، نخبه گرایی اطلاعاتی نامیده میشود.

همانطوری که Cerbian و Lahiri بحث کرده اند، نخبه گرایی اطلاعات قابل توضیح با مشخصات بدیهی گراف مانند درجه نیست. آنها نشان دادند که هیچ ارتباطی بین مقادیر متوسط نرمال شده هدف نهایی از ۲ درصد بالایی گره ها و چندین خصیصه ساده از هر گره وجود ندارد، مانند درجه ورودی (تعداد شرکای ایمیلی ورودی)، تعداد کل ایمیل های دریافت شده از همسایگان، و زمانی که گره برای اولین بار در مجموعه داده ها دیده شده است. این پدیده کاهش یافته زمانی که PSODM را اعمال کردیم، گره همانطور که در شکل ۱(a) نشان داده شده است. در مقابل GADM، گره های PSODM اطلاعاتشان را با بهترین همسایه تبادل میکنند (gbest).

به وسیله پهینه سازی محلی رویه انتشار اطلاعات، نخبه گرایی اطلاعاتی در شبکه کاهش قابل ملاحظه ای دارد.

ب. الگوریتم تشخیص اجتماع (GPSODM)

اگرچه نتایج انتشار اطلاعات دیدگاه جالبی را حاصل کرد، هدف تحقیق ما بهمود در رویه تشخیص اجتماع نبود. اینجا، نتایج بدست آمده برای چارچوب GPSODM تشخیص اجتماعی را نشان میدهیم، GPSODM با GGADM (Game+GADM) مجموعه داده جهان واقعی و دو مجموعه داده ساختگی مقایسه شده است. دو محک مجموعه رسانه اجتماعی از [۲۶] برای نشان دادن مقیاس پذیری

$$\forall i, s'_i \neq s_i, U_i(S_{-i}, s'_i) \leq U_i(S_{-i}, s_i)$$

از آنجایی که رسیدن به تعادل نش عمومی در این بازی شدنی نیست، تعادل نش محلی را محاسبه میکنیم. استراتژی پروفایل S یک تعادل محلی است اگر همه مأمورین بهینه محلی خود را بازی کنند. در اینجا N_{agents} به فضای استراتژی محلی مأمور i اشاره دارد:

$$\forall i, s'_i \in ls(s_i), U_i(S_{-i}, s'_i) \leq U_i(S_{-i}, s_i)$$

Algorithm 2 GPSODM

```

1: Input: underlying network graph  $G$ 
2: Output: communities as a final division of  $G$ 
3: communities = {}
4: Agents = {agent1, agent2, ..., agentn}
5: repeat
6:   agenti = Random_Select(Agents)
7:   actioni = Best_Operation (join, leave, switch)
8:   u'_i = Utility_Calculate(agenti, actioni)
9:   if (ui < u'_i) then
10:    ui ← u'_i.
11:    Update si
12:    Update communities
13:   else actioni = no Operation
14:   end if
15: until (local equilibrium is reached)

```

۵. نتایج تجربی

همانطور که گفته شد، PSODM میتواند برای مدلسازی انتشار اطلاعات در شبکه های اجتماعی استفاده شود. برای نشان دادن این موضوع، ما PSODM را برای شبیه سازی انتشار اطلاعات بر روی یک شبکه اجتماعی پویا اعمال کردیم. این شبکه از مجموعه داده های ایمیل Enron که توسط Lahiri و Cerbian برای ارزیابی شان در [۹] آمده است استخراج شده بود. نتایج نهایی از هر دو مدل (PSODM و GADM) بر روی این مجموعه اطلاعات تفاوت بین این دو الگوریتم را نشان میدهد.

الف. انتشار اطلاعات با PSODM

مجموعه اطلاعات پست الکترونیک Enron یک مخزن از ایمیل های رد و بدل شده بین مدیران اجرایی Enron سابق است. که یکی از بزرگترین مجموعه داده های محیط ایمیل یک شرکت است، و طبیعتاً به عنوان یک شبکه اجتماعی پویا شناخته میشود. هر ایمیل ارسال شده را میشود به عنوان یک یال جهت دار و زمان دار در یک شبکه اجتماعی پویا بررسی کرد. همچنین هر آدرس ایمیل در این مجموعه اطلاعاتی به عنوان یک رأس در گراف متناظر شناخته میشود. Cerbian و Lahiri را تجزیه کردن و مشخص شد که ۱۳۲۶۷۱ یال مرتبط با ایمیل های را تجزیه کرده اند و مشخص شد که ۲۱۵۸۴۱ زمان ثبت یکتا بود که به شکل غیریکنواخت در دوره حدوداً ۴ ساله بوده است. ما الگوریتم ها را بر روی

جريان اطلاعات به نتيجه ماندگاری اطلاعات غالب در هر گره خواهد انجامید، که میشود برای مشخص کردن ساختار اجتماع از آن استفاده کرد.

[LPA]^[۱۸]: در این الگوریتم، هر گره با یک برجسب یکتا مقدار دهی اولیه میشود و در هر مرحله هر گره برجسبی را اتخاذ میکند که اکثر همسایگانش در آن لحظه دارند. در طی این رویه تکراری گروه های چگال متصل از گره ها یک اجماع بر روی یک برجسب یکتا به منظور ایجاد اجتماع ایجاد میکند.

[HA]^[۱۹]: فرمی گسترده از LPA است. بالقوه های ابتکاری پیش نهاد شده، در این کار ارائه شده اند، که میتوانند به LPA افزوده شوند و میانگین کارایی تشخیص و سازگاری آن را افزایش دهند. همچنین، به وسیله افزایش سرعت پارامترها الگوریتم جدید میتواند با انواع مختلفی از شبکه ها سازگار باشد.

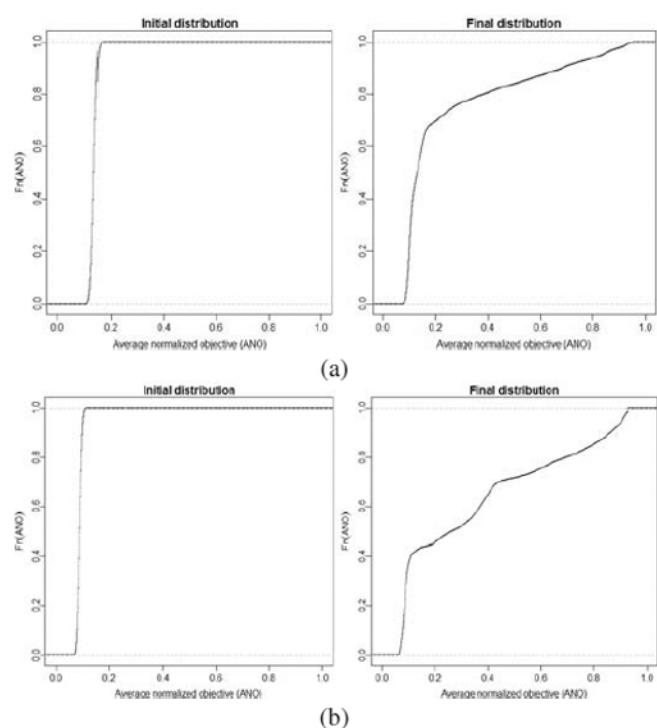
[InfoMap]^[۳۰]: یک روش مرکز بربایه حداقل طول است. از احتمال جريان گام های تصادفی بر روی شبکه مانند یک پروفیل برای جريان اطلاعات در سیستم واقعی استفاده میکند و شبکه را به مازول هایی با فشرده سازی یک توضیح از احتمال جريان تجزیه میکند. نتيجه نقشه ای خواهد شد که نظم و قاعده را در ساختار و روابطشان هم ساده و هم هایلایت کرده است.

۱. ارزشیابی: به منظور اندازه گیری شباهت بین پارتبیشن اجتماع شناسایی شده و پارتبیشن حقیقی میتوانیم از NMI استفاده کنیم. اگرچه، زمانی که حقیقت ناشناخته است، مازول بندی کردن میتواند برای ارزشیابی کیفیت تشخیص اجتماع مورد استفاده قرار گیرد. پیمانه ای بودن محبوب ترین اندازه کیفی برای تشخیص اجتماعات در شبکه های اجتماعی است، اگرچه این اندازه اشکالاتی دارد و غیرقابل اعتماد میشود زمانی که شبکه بسیار تنک میشود.

۲. مجموعه اطلاعات: آزمایشات بر روی سه شبکه واقعی و دو شبکه ساختگی انجام گرفت. نتایج بدست آمده نشان میدهد که روش ما بر روی همه مجموعه داده ها عملکرد خوبی دارد. همچنین ما نتایج بدست آمده را از نظر پیمانه ای بودن و پیچیدگی زمانی بر روی مجموعه اطلاعات اجتماعی رسانه هم بررسی کردیم. ابتدا، نتایج بر روی شبکه های ساختگی و را به منظور نشان دادن قدرت روشنان بر روی مجموعه اطلاعات وسیع نشان میدهیم.

در پایان از سه مجموعه اطلاعات واقعی استفاده کردیم، که برای هریک ground truth داریم، و کارایی الگوریتم خان LFR^[۱۷] را نشان میدهیم. از آنجایی که PSODM از مدل انتشار اطلاعات واقعی تری استفاده میکند، میتواند بر روی این مجموعه اطلاعاتی به نتایج بهتری برسد.

GPSODM استفاده شده است. از دو استاندارد متريک برای نشان دادن کارایی الگوریتم خود استفاده کردیم: نرمال سازی اطلاعات متقابل^[۱۸] [۲۷] و پیمانه ای بودن^[۱۹] [۲۸]. جدول ۱ پارامترهای آزمایش را نشان میدهد.



شکل ۱: توزیع تجمعی مقادیر ANO مجموعه ایمیل های Enron در ابتداء (چپ) و پایان (راست) (a) توزیع تجمعی (b) PSODM توزیع تجمعی GADM.PSODM به سبب مدل انتشار کارآمدش نخبه گرایی اطلاعاتی کمتری به نمایش می گذارد.

جدول ۱: پارامترهای استفاده شده در آزمایش

Algorithms	Parameters
GPSODM	$C = 2.0$
GGADM [36]	None
HA [40]	$m = 0.1, \delta = 0.05$
MMC [41]	$\alpha = 2.65, \beta = 2, \rho = 0.9, \mu = 1.08, \eta = 0.7, \gamma = 0.1$
LPA [42]	None
InfoMap [43]	None

همه الگوریتم ها به زبان جاوا بیاده سازی شده اند و بر روی یک سیستم با ۴ گیگابایت حافظه رم و پردازنده ۲.۵۳ گیگاهرتزی اینتل اجرا شده است. روش تشخیص اجتماع زیر به عنوان محک استفاده شده اند:

[MMC]^[29]: این الگوریتم یک روش انتشار برجسب گذاری است که از برجسب ها برای نشان دادن انتشار اطلاعات در شبکه های اجتماعی استفاده میکند. هر گره قادر به تبادل اطلاعات با همسایگان موجود است. با به کارگیری عملیات مشابه تورم در الگوریتم خوش بندی مارکف برای به روز رسانی احتمالات انتقال،

18) Normalized Mutual Information (NMI)
19) Modularity (Q)

جدول ۲: نتایج تشخیص اجتماع بروی داده‌های واقعی (میانگین از ۱۰۰ بار اجرا)

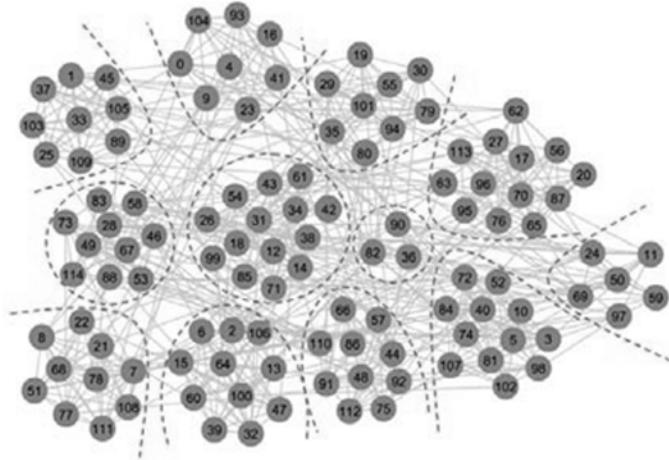
	NMI			Modularity		
	Football	Dolphin	Karate	Football	Dolphin	Karate
GPSODM	1.000	0.723	1.000	0.583	0.580	0.412
GGADM	0.910	0.736	1.000	0.598	0.544	0.381
HA	0.907	0.707	0.754	0.566	0.449	0.300
MMC	0.885	0.579	1.000	0.595	0.526	0.371
LPA	0.927	0.710	0.751	0.597	0.450	0.362
InfoMap	0.899	0.695	0.643	0.575	0.514	0.354

کارایی الگوریتم های متفاوت بروی این شبکه ها برای متريک های پيمانه ای بودن و NMI در جدول دو نشان داده است.

American College Football Net- work- Bottlenose Dolphin Network- Zachary Karate Club Network

و محک های مجموعه اطلاعات شبکه های رسانه ای:

Flicker-Blobcatalog



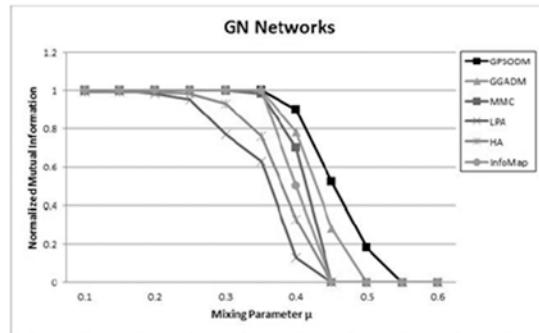
شکل ۴: بهترین ساختار شبکه فوتبال کالج آمریکایی یافته شده با روش ما و $NMI = 1$

جدول ۳: خصایص آماری مجموعه اطلاعات Flickr و BlogCatalog

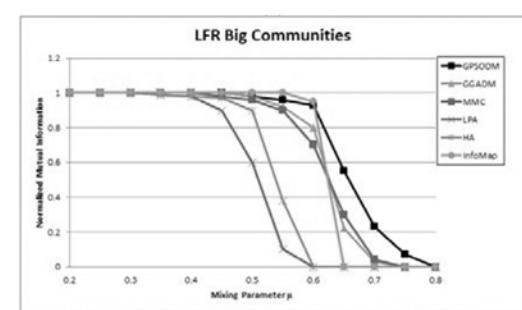
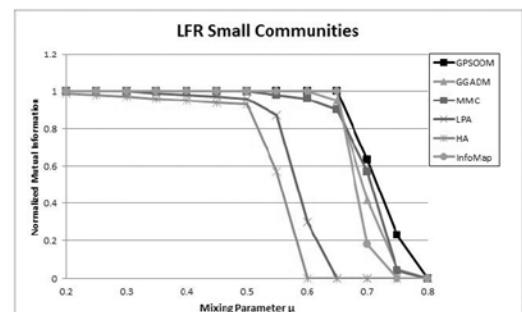
Dataset	# Categories	# Nodes	# Edges	Max Degree
BlogCatalog	39	10,312	333,983	3,992
Flickr	195	80,513	5,899,882	5,706

جدول ۴: زمان محاسباتی از دو روش و تعداد اجتماعات کشف شده بروی BlogCatalog مجموعه اطلاعات

Method	Time (Seconds)	#Communities
GPSODM	8,834	15,340
GGADM	5,430	15,645



شکل ۲: تشخیص اجتماع (NMI) بروی شبکه های ساختگی GN. دربرابر روش های دیگر در موارد چالش برانگیز با پارامتر پیچیده بیشتر بهتر عمل می کند.



شکل ۳: تشخیص اجتماع (NMI) بروی شبکه های ساختگی LFR

(a) اجتماعات کوچک (۵۰-۲۰) گره

(b) اجتماعات بزرگ (۱۰۰-۲۰) گره

GPSODM دربرابر روش های دیگر در موارد چالش برانگیز با پارامتر پیچیده بیشتر بهتر عمل می کند.

- [10] J. H. Holland, "Building blocks, cohort genetic algorithms, and hyperplane-defined functions," *Evolutionary Computation*, vol. 8, pp. 373–391, 2000.
- [11] A. Hajibagheri, H. Alvari, A. Hamzeh, and S. Hashemi, "Community detection in social networks using information diffusion," in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2012, pp. 702–703.
- [12] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, no. 6, Dec. 2004.
- [13] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, p. 75, 2010.
- [14] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [15] A. Lancichinetti and S. Fortunato, "Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities," *Physical Review E*, vol. 80, no. 1, p. 016118, 2009.
- [16] S. E. Schaeffer, "Survey: Graph clustering," *Computer Science Review*, vol. 1, no. 1, pp. 27–64, Aug. 2007.
- [17] I. A. Kovacs, R. Palotai, M. S. Szalay, and P. Csermely, "Community landscapes: An integrative approach to determine overlapping network module hierarchy, identify key nodes and predict network dynamics," *PLoS ONE*, vol. 5, no. 9, 2010.
- [18] U. Raghavan, R. Albert, and S. Kumara, "Near linear time algorithm to detect community structures in large-scale networks," *Physical Review E*, vol. 76, no. 3, 2007.
- [19] I. Leung, P. Hui, P. Lio, and J. Crowcroft, "Towards real-time community detection in large networks," *Physical Review E*, vol. 79, no. 6, Jun. 2009.
- [20] S. Gregory, "Finding overlapping communities in networks by label propagation," 2009.
- [21] S. M. van Dongen, "Graph Clustering by Flow Simulation," Ph.D. dissertation, University of Utrecht, The Netherlands, 2000.
- [22] J. Reichardt and S. Bornholdt, "Statistical mechanics of community detection," *Phys. Rev. E*, vol. 74, Jul 2006.
- [23] M. B. Hastings, "Community detection as an inference problem," *Phys. Rev. E*, vol. 74, Sep 2006.
- [24] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, 1995, pp. 1942–1948 vol.4.
- [25] C. Alos-Ferrer and A. Ania, "Local equilibria in economic games," *Economics Letters*, vol. 70, no. 2, pp. 165–173, 2001.
- [26] L. Tang and H. Liu, "Relational learning via latent social dimensions," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2009, pp. 817–826.
- [27] L. Danon, A. Daz-guilera, and J. Duch, "Comparing community structure identification," *Journal of Statistical Mechanics: Theory and Experiment*, 2005.
- [28] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, Feb. 2004.
- [29] W. Chen, "Discovering communities by information diffusion," in *Proceedings of the International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, vol. 2, 2011, pp. 1123–1132.

جدول 5: زمان محاسباتی از دو روش و تعداد اجتماعات کشفشده بروی
مجموعه اطلاعات Flickr

Method	Time (Seconds)	#Communities
GPSODM	53,570	128,591
GGADM	36,029	202,171

۶. نتیجه گیری و کارهای آینده

اطلاعات میتواند عقیده عمومی را شکل دهد، دلیل وحشت و شورش باشد، خروجی انتخابات را تغییر دهد، یا مشتریان را جذب به اتخاذ یک محصول کند.

در این مطالعه، یک مدل انتشار اطلاعات برپایه بهینه سازی حرکت ذرات ارائه شد که میتواند برای یافتن ساختار اجتماع در شبکه مورد استفاده باشد. در این چارچوب هر گره از گراف به عنوان یک مأمور خودخواه در نظر گرفته میشود که میخواهدتابع مطلوبیت خودش را به شکل محلی بهینه کند و پخش اطلاعات خودش را بیشینه کند. نتایج بدست آمده نشان میدهد که روش ما در مجموعه اطلاعات واقعی و ساختگی از روش های موجود عملکرد بهتری دارد. برای کارهای آینده، طرح داریم تا مفهوم تداخل را به وسیله تابع کنترلی از پیش تعیین شده پشتیبانی کنیم.

مراجع:

- P. Domingos and M. Richardson, "Mining the network value of customers," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2001, pp. 57–66.
- M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002, pp. 61–70.
- D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, pp. 137–146.
- M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi, "Measuring user influence in Twitter: the million follower fallacy," in *Proceedings of International AAAI Conference on Weblogs and Social Media*, 2010.
- M. Goetz, J. Leskovec, M. McGlohon, and C. Faloutsos, "Modeling blog dynamics," in *Proceedings of the International Conference on Weblogs and Social Media*, May 2009.
- J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2009, pp. 497–506.
- D. Liben-Nowell and J. Kleinberg, "Tracing information flow on a global scale using Internet chain-letter data," *Proceedings of the National Academy of Sciences*, vol. 105, no. 12, pp. 4633–4638, 2008.
- M. Granovetter, "Threshold Models of Collective Behavior," *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.
- M. Lahiri and M. Cebrian, "The genetic algorithm as a general diffusion model for social networks," in *AAAI Conference on Artificial Intelligence*, 2010. [Online]. Available: <http://www.aaai.org/ocs/index.php/AAAI/AAAI10/paper/view/1785>

شبکه‌های اجتماعی: تعاریف، ابزارها و روش‌های تحلیل

سید سعید حاجی سید جوادی

دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر (پلی تکنیک تهران)

کلیدواژه:

تحلیل شبکه های اجتماعی
شبکه های پیچیده
شبکه های اجتماعی آنلاین

در دو دهه گذشته، شبکه‌های اجتماعی در زمینه‌های گوناگون تحلیل ارتباطات افراد و تعیین اهمیت ساختارها و الگوهای این ارتباطات به شدت مورد توجه قرار گرفته است. در این مقاله تلاش شده است ابتدا به تعاریف مختلف شبکه‌های اجتماعی و شبکه‌های اجتماعی آنلاین پرداخته شود. پس از ارائه خلاصه‌ای از مطالعات و تحقیقات صورت گرفته، انواع تحلیل‌های مرسوم شبکه‌های اجتماعی را دسته‌بندی و معرفی می‌کنیم. سپس بعد از معرفی و بررسی اجمالی این تحلیل‌ها به سراغ کاربردها و مسایل مطرح شده در این زمینه می‌پردازم.

چکیده:

(نه صرفاً اجتماعی) امکان پذیر شود، دسترسی مناسب به شاخص‌های فردی و گروهی نهفته در آن هاست بطوریکه در بعضی موارد استخراج یک شاخص مهم در شبکه به تنها بیان می‌تواند اطلاعات بسیار مفیدی را در اختیار داده کاوی تحلیلگر قرار دهد. به دلیل اجتماعی بودن شبکه‌ها مورد بحث، بعد جدیدی از شاخص‌های اجتماعی (مانند ضریب نفوذ، شاخص قدرت، سرعت انتشار اخبار...) تعریف می‌شود. همچنین تحلیل‌های ثانویه مشخصه‌های گوناگون فردی و اجتماعی بدست آمده از شبکه، می‌تواند اطلاعات نهفته بازارشی را آشکار سازد که برداشت‌های صورت گرفته از آن‌ها مختص تعاملات اجتماعی افراد با یکدیگر است و در شبکه‌های مشابه دیگر (مانند شبکه‌های ترافیکی، شبکه‌های پروتکلی و...) معنای دیگری خواهد داشت.

۱. شبکه‌های اجتماعی

شبکه اجتماعی، ساختار اجتماعی است که توسط مجموعه‌ای از افراد (یا سازمان یا موجودیت‌های اجتماعی دیگر) شکل گرفته است. این مجموعه از رابطه‌های اجتماعی مانند مبادله اطلاعات، همکاری، دوستی، خویشاوندی، تبادلات مالی، تجارت و... مرتبط شده‌اند. واسمن و فاست می‌گویند: شبکه اجتماعی می‌تواند بعنوان سیستم رابطه‌ی اجتماعی که توسط مجموعه بازیگران و ارتباط اجتماعی آنها توصیف می‌شود، تعریف شود. تحقیق در تعدادی از زمینه‌های آکادمیک نشان داده است که شبکه‌های اجتماعی در بسیاری از سطوح به کار گرفته می‌شوند، از خانواده‌ها گرفته تا ملت‌ها و نقش مهمی در تعیین راه حل مسائل، اداره کردن تشکیلات و میزان موفقیت افراد در رسیدن به اهدافشان ایفا می‌کند. یک شبکه بعنوان مجموعه‌ای از اشیاء (یا گره‌ها) و مجموعه‌ای از روابط مشخص می‌شود و می‌توان شبکه‌های روابط اجتماعی را با اصطلاحات راس و یال نگریست که در آن رأس‌ها بازیگران فردی درون شبکه‌ها هستند و یال‌ها روابط میان این بازیگران. در ساده‌ترین شکل یک شبکه اجتماعی نگاشتی از تمام یال‌های مربوط، میان راس‌های مورد مطالعه است.

۱. مقدمه
انسان‌ها اجتماعی ترین موجودات، در میان همه حیوانات هستند که با هم ارتباطات وسیعی دارند. این روابط، جامعه جهانی پیچیده ما را شکل می‌دهد. اکنون می‌توان پیچیده‌ترین جامعه‌ها را که به کمک ابزارهای نوین ساخت افزاری و نرم افزاری از سوی تحلیل‌گران شبکه اجتماعی به جهان کوچک تبدیل کرد و روی میز قرار داد و سپس به تحلیل آن پرداخت. بنابراین، ورود جامعه شناسی به علم به نسبت نوپای تحلیل شبکه اجتماعی نه تنها جهان امروز را به جهانی کوچک برای تحلیل تبدیل کرده بلکه تبیین‌های مربوط را نیز تا حد بسیاری کمی نموده و نزدیکی علوم اجتماعی به علوم دقیق را میسر کرده است.

در دو دهه گذشته، شبکه‌های اجتماعی در زمینه تحلیل ارتباطات افراد و تعیین اهمیت الگوهای این ارتباطات به شدت مورد توجه قرار گرفته است. در سال‌های اخیر نیز شبکه‌های اجتماعی آنلاین به دلیل افزایش ناگهانی وسائل برقراری ارتباط با اینترنت، از جمله کامپیوتراهای شخصی، تلفن‌های همراه و ابداعیات ساخت افزاری جدیدی همچون تبلت‌ها، بسیار محبوب شده است. همچنین شبکه‌های اجتماعی آنلاین قادر محدودیت‌های جغرافیایی و اجتماعی شبکه‌های اجتماعی مرسوم هستند و ارتباطات افراد به شکل ساده تری نسبت به ملاقات‌های چهره به چهره صورت می‌گیرد. از این‌رو بیشتر تحلیل‌ها و مطالعات سال‌های اخیر در زمینه تحلیل چنین شبکه‌هایی متتمرکز بوده است که نمونه‌های شاخص آن‌ها سایت‌های LinkedIn و Twitter، Facebook هستند.

ساختار اصلی بنا شده حول شبکه‌های اجتماعی می‌تواند امکان تحلیل گوناگون داده‌ای مانند جستجو، تحلیل متن، تحلیل تصاویر و... را در اختیار تحلیل‌گیران قرار دهد. همچنین تحلیل رشد و پیشرفت ساختار یک شبکه اجتماعی، کشف انجمن‌های موجود در آن، پیش‌بینی ارتباطات افراد، تحلیل میزان شهرت و نفوذ افراد، تحلیل نحوه پخش جریان اطلاعات و شایعات در آن و... از جمله مسائل جالب و مورد توجه مطرح شده در این زمینه است. آنچه که موجب می‌شود تا به کمک آن امر تحلیل شبکه‌های گوناگون

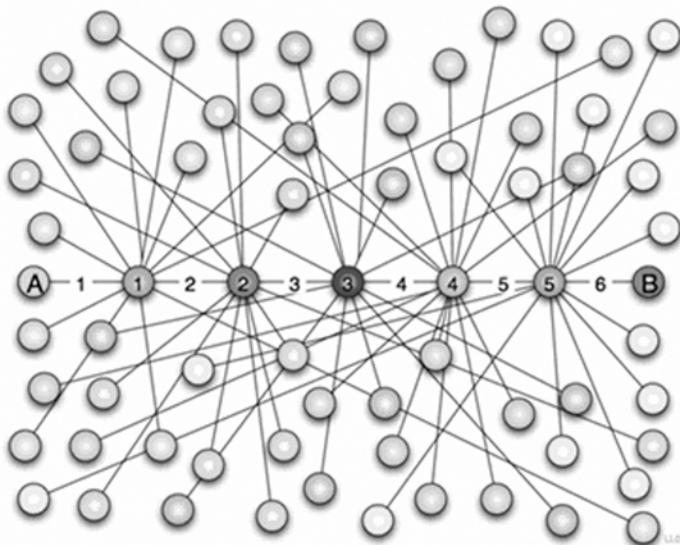
و صحت سنجی این نظریه‌ها به صورت بسیار ساده فراهم گشته است.

۴. نظریه ۶ درجه جدایی

یکی از قدیمی‌ترین و مشهورترین قضیه‌های مطرح شده در زمینه تحلیل شبکه‌های اجتماعی نظریه ۶ درجه جدایی یا پدیده دنیای کوچک نام دارد. در دهه ۶۰ میلادی یک جامعه شناس آمریکایی به نام استنلی میل‌گرام نظریه ای به نام ۶ درجه جدایی Six degrees of separation مطرح کرد که طی آن هر دو شخص دلخواه روی کره زمین با ۶ واسطه به هم مربوط می‌شوند.^[۳]

(شکل ۱)

استنلی میل‌گرام به دنبال پاسخی برای چگونگی ارتباط افراد با یکدیگر بود. او می‌خواست بداند که ارتباط انسان‌ها یک ارتباط یک به یک است یا در شبکه‌ای محدود روابط شکل می‌گیرد. او برای رسیدن به مکانیزم روابط انسانی دست به یک آزمایش زد. ۱۶۰ پاکت آماده کرد و برای ۱۶۰ نفر که در ایالت نیبراسکا زندگی می‌کردند فرستاد و از آنان خواست که به هر طریق که صلاح می‌دانند پاکت را بدهند یک کارگزار بورس که در ایالت ماساچوست زندگی می‌کرد برسانند. نتیجه نهائی مطالعه میل‌گرام این بود که اکثر نامه‌ها پس از طی ۵ یا ۶ مرحله به دست فرد مورد نظر رسیده بود.



۵. زمینه‌های مختلف شبکه‌های اجتماعی

اصلی ترین بخش تعریف شبکه‌های اجتماعی بر اساس حضور و تعاملات انسان‌ها بنا شده است. از این رو همواره به عنوان یک فیلد تحقیقاتی مناسب برای رشته جامعه شناسی مطرح بوده است. پیش از ظهور اینترنت و محبویت شبکه‌های اجتماعی میان مردم، داده‌های واقعی را بصورت دستی از میان تعاملات افراد با یکدیگر جمع آوری می‌کردند و مورد تحلیل قرار می‌دادند. برای مثال همانطور که اشاره شد میل‌گرام برای آزمایش نظریه خود (شیش درجه جدایی) از سیستم پستی استفاده نمود. اما ممکن است در این روش به دلیل عدم دریافت بعضی از نامه‌های ارسالی یا دیر ارسال نمودن نامه‌های دریافتی نتیجه آزمایش به کلی تحت تاثیر قرار گیرد. از آنجایی که بررسی اینگونه آزمایشات تجربی در اجتماعات بسیار بزرگ عالم‌گیر ممکن بود، لذا در آن زمان نظریه مذکور تنها از منظر کیفی مورد قبول واقع شد.

شبکه‌های اجتماعی برای بررسی چگونگی تأثیرات متقابل میان تشکیلات، توصیف بسیاری از اتصالات غیررسمی که مجریان را به یکدیگر متصل می‌کند، نیز مورد استفاده قرار گرفته است و در این زمینه‌ها نیز به خوبی برقراری ارتباطات فردی میان کارمندان در سازمان‌های مختلف عمل می‌کند. شبکه‌های اجتماعی نقش کلیدی در موفقیت‌های تجاری و پیشرفت‌های کاری ایفا می‌کنند. شبکه‌ها راه‌هایی را برای شرکت‌ها فراهم می‌کند که اطلاعات جمیع اوری کنند، از رقابت پرهیزند و حتی برای تنظیم قیمت‌ها و سیاست‌ها با هم تبادل کنند.

۲. تحلیل شبکه‌های اجتماعی

تحلیل شبکه‌های اجتماعی رویکردی برای مطالعه فعل و انفعال شبکه‌های بشمری است و برای بررسی الگوهای، ساختار ارتباطی یا سازماندهی شبکه‌های اجتماعی استفاده شود. در علوم اجتماعی، شبکه شامل مجموعه‌ای از بازیگران (یا گره‌ها، یا عامل‌ها، یا نقاط، یا رئوس) که ممکن است دارای روابطی (یا پیوندها، یا یالها) با یکدیگر باشند تحلیل شبکه‌های اجتماعی عبارتست از نگاشت و اندازه گیری روابط و همکاری‌های بین افراد، گروه‌ها، سازمان‌ها و هر موجودیتی که قابلیت پردازش اطلاعات و دانش داشته باشد. گره‌ها در شبکه افراد و گروه‌ها را نشان می‌دهد در حالی که پیوندها، ارتباطات یا جریان بین گره‌ها را نشان میدهد.

۳. شبکه‌های اجتماعی آنلاین

واضح است که تعریف شبکه‌های اجتماعی منحصر به موارد خاصی همچون شبکه‌های اجتماعی تحت وب مانند فیس بوک و توییتر نمی‌باشد. مسائل شبکه‌های اجتماعی غالباً در رشته جامعه شناسی با رویکرد تحلیل تعاملات بین هر گروهی از بازیگران مورد مطالعه قرار می‌گرفته است. این تعاملات می‌توانسته به صورت مرسوم یا غیر مرسوم مانند ارتباط چهره به چهره، ارتباط از طریق ایمیل، ارتباط از طریق پست... باشد. مطالعات معمول در زمینه تحلیل شبکه‌های اجتماعی تنها بر تعاملات آنلاین تمرکز ندارد و از لحاظ تاریخی قبل از ظهور و محبویت کامپیوترها یا اینترنت بوده است.

توجه داشته باشیم که تعریف شبکه اجتماعی آنلاین می‌تواند شامل محدوده وسیع تری از سایت‌هایی مانند Facebook و Twitter که به عنوان شبکه‌های اجتماعی تبلیغ می‌کنند، باشد. هر وب سایت یا نرم افزار تحت وبی که تجربیات اجتماعی را به فرم تعاملات کاربران فراهم می‌کند می‌توان به عنوان گونه‌ای از شبکه اجتماعی آنلاین مورد قبول قرار داد. برای مثال سایت‌های به اشتراک گذاری رسانه‌ای مانند Flickr و YouTube به طور رسمی جزء شبکه‌های اجتماعی قلمداد نمی‌شوند، در حالیکه آنها اجزای تعاملات اجتماعی را به فرم تبادلات اطلاعات درمورد محتوایی که به اشتراک گذاشته می‌شود را به کاربران می‌دهند.

به همین منوال در شبکه ارتباط تلفنی، شبکه سراسری وب، نرم افزارهای چت آنلاین نیز جنبه خاص اجتماعی وجود دارد. فیلد تحلیل شبکه‌های اجتماعی آنلاین در سالهای اخیر به شدت مورد توجه قرار گرفته است. یکی از دلایل عمده آن وجود داده‌های بسیار غنی در این سایتها برای تحلیل و داده کاوی است. با گسترش و فراگیر شدن شبکه‌های اجتماعی آنلاین امکان آزمایش

قدرت پردازش زبان طبیعی و پردازش تصویر خود از قصد آن شخص خبردار شوند و به او گزینه‌های اقامت در هتل‌ها و استفاده از تورهای تحت قرارداد با گوگل را پیشنهاد دهند.

به تازگی تحلیل گران به این نتیجه رسیده اند که ترکیب این دو نوع تحلیل نتایج موثرتری را برای کاربردهای مختلف در اختیارشان قرار خواهد داد. برای مثال پیدا کردن یک انجمان متخصص در زمینه خاص نیازمند به کارگیری تراکنشات انتقامی و حتی انتقام گیری باشد.

نقاوت اصلی دیگری که در زمینه الگوریتم‌های شبکه‌های اجتماعی مطرح می‌شود، اختلاف بین تحلیل پویا و تحلیل ایستا است.

٣) تحلیل اپستا

در این تحلیل فرض بر این است که شبکه اجتماعی به کندی در حال تغییر و تحول است و ارتباطات و شاکله اصلی شبکه در مدت زمان کوتاهی به کلی تغییر نخواهد کرد.

٤) تحلیل بہا

این نوع تحلیل بر روی شبکه‌های پویا مورد استفاده قرار می‌گیرد. شبکه‌های پویا شبکه‌های هستند که در آن‌ها به طور مداوم تعاملات بسیار زیادی بصورت پیوسته در طول زمان جایه‌جا می‌شود که این انتقال اطلاعات به جریان شبکه^۲ منجر خواهد شد. تحلیل چنین شبکه‌هایی بسیار پیچیده‌تر از تحلیل شبکه‌های ایستا است و از جمله موضوعات پرطرفدار در پژوهش‌های اخیر در زمینه تحلیل شبکه‌های اجتماعی بوده است. از جمله این شبکه‌ها می‌توان به شبکه مخابراتی تلفن همراه یا نرم افزارهای چت و گفتگوی آنلاین اشاره کرد، که ارتباطات افراد با یکدیگر در شبکه در طول زمان بطور مداوم در حال تغییر و جایه‌جایی هستند.

اینگونه شبکه‌ها را می‌توان با گراف‌های پویا مدل نمود که در آن یال‌ها دائماً در حال تغییر هستند، به دلیل رهگیری هم‌زمان ارتباطات فوق العاده زیاد میان گره‌ها، این گونه گراف‌ها در پردازش دچار چالش بزرگی خواهد بود. در چنین مواردی نرم افزارهای جریان کاوی گراف^۴ جهت تحلیل موثر، مورد نیاز است.

جمع بندی

در این مقاله به معرفی اجمالی انواع شبکه‌های اجتماعی و اهمیت و جایگاه آن در میان پژوهش‌های اخیر پرداختیم. همچنین در ادامه زمینه‌های مختلف تحلیل این گونه شبکه‌ها را مورد بررسی قرار دادیم و در نهایت تحلیل‌های مسimum در این موضع معاف ننمودیم.

در مقاله بعدی به مباحث تحقیقاتی به روز موجود در زمینه تحلیل شبکه‌های اجتماعی، مبادراتی و پرایو هر کدام مثالی، خواهیم آورد.

منابع

1. C. Aggarwal, Charu. *Social Network Data Analytics*. New York : Springer Press, 2011, pp. 1-14
 2. Wasserman, Stanley; Faust, Katherine, *Social Network Analysis. Methods and Applications*, Cambridge: Cambridge University Press, 1994
 3. Stanley Milgram, "The Small World Problem", *Psychology Today*, 1967, Vol. 2, pp. 60-67.
 4. C. Aggarwal, Charu ; Wang. *Managing and Mining Graph Data*. New York : Springer Press, 2010

سایت‌های اشتراک گذاری میدیا مانند YouTube، Flicker و آن جهت که اجازه انتقال حجم وسیعی از ارتباطات را بین کاربران می‌دهند، می‌توانند به عنوان اشکال غیر مستقیم شبکه‌های اجتماعی در نظر گرفته شوند. در چنین مواردی تعاملات حول یک سرویس خاص مانند به اشتراک گذاری محتوا تمرکز یافته است. در نتیجه تعدادی از شبکه‌های اجتماعی می‌تواند حول یک نوع ارتباط خاص در انجمن‌های گوناگون، پدید آید. یک مثال مشهور آن انجمن‌های علمی است که شبکه‌های Bibliography آن می‌تواند با محور قراردادن مولف و یا مقالات شکل گیرد. هر وبسایت یا نرم افزار تحت وبی که تجربیات اجتماعی را به فرم تعاملات کاربران فراهم می‌کند می‌توان به عنوان گونه‌ای از شبکه اجتماعی آنلاین مورد قبول قرار داد. برای مثال سایت‌های Flicker و YouTube به اشتراک گذاری رسانه‌های مانند Flickr و YouTube بطور رسمی جزو شبکه‌های اجتماعی قلمداد نمی‌شوند، در حالیکه آن‌ها اجازه تعاملات اجتماعی را در فرم تبادلات اطلاعات درمورد محتوایی که به اشتراک گذاشته شده، به کاربران می‌دهند.

به همین منوال در شبکه ارتباط تلفنی، شبکه سراسری و ب، نرم افزارهای چت آنلاین نیز جنبه خاص ارتباطات اجتماعی وجود دارد. ظهور نرم افزارهای گفتگوهای آنلاین مانند messenger، MSN messenger، Skype و ... - که می‌توان آن‌ها را به عنوان اشکال غیرمستقیم شبکه‌های اجتماعی در نظر گرفت- به دلیل شفاف بودن ارتباطات افراد با یکدیگر و امکان رهگیری^۱ این روابط، شرایط بررسی نظریه‌های مطرح شده در زمینه شبکه‌های اجتماعی مانند نظریه ۶ درجه جایی را به وجود آورند.

۶. دسته بندی انواع تحلیل شبکه های اجتماعی

در این بخش به معرفی و بررسی انواع تحلیل‌های مرسوم در زمینه تحلیل شبکه‌های اجتماعی می‌پردازیم.

۱) تحلیل ساختاری، با ارتبا

در این نوع تحلیل به دنبال درک درست از ساختار شبکه و روابط افراد با یکدیگر در آن هستیم. از جمله مسائلی که در اینگونه تحلیل مطرح می‌شود عبارتند از: پیدا کردن گرههای مهم و تاثیرگذار، کشف انجمن‌ها و گروههای، حدس و پیش‌بینی ارتباطات آینده گره‌های شبکه، تحلیل ناحیه‌های در حال رشد شبکه.

این تحلیل‌ها، دیدخوبی در مورد نحوه پیشرفت کلی شبکه برای تحلیل گر فراهم می‌کنند.

٢) تحلیل محتوایی

بسیاری از شبکه های اجتماعی از جمله Flicker، YouTube، Facebook حاوی مقدار عظیمی از محتوا هستند که زمینه خوبی را برای داده کاوی و تحلیل دادهای بر روی آنها پرداخت می آورد. به طور مشابه شبکه بلاگ ها، شبکه پست های الکترونیک و انجمن ها، حاوی اطلاعات متنی غنی برای تحلیل و بررسی هستند. از جمله کاربردهای چنین تحلیلی می توان به سرویس پیشنهاد دهنده تبلیغاتی سرویس گوگل اشاره نمود که بصورت خودکار با خواندن و یادگیری از متن ارسالی و دریافتی اشخاص به نمایش تبلیغات مرتبط با آن ها می پردازد برای مثال شخصی به دوستش از سفر قریب الوقوعش به شهر خاصی سخن می گوید و یا در مرحله بعدی عکس هایی از مناظر دینی آن شهر برای دوستش ارسال می کند. سرویس های جستجوی متن و تصویر گهگا، مه توانند با

1) Trace

- 1) Trace
- 2) Network Stream

3) Graph stream mining applications



علوم کامپیوتر و تئوری بازی

خانم دکتر فرشته آزادی پرند

چه میزان ارتباط به منظور محاسبه مقدار یکتابع لازم است. برای درک ارتباط این مساله و اقتصاد، به این مساله توجه نمایید که سازوکارهای اقتصادی با این هدف طراحی می‌شوند که عامل‌ها انگیزش لازم برای بازگو نمودن الوبت‌های خود را با صداقت داشته باشند. لیکن میزان ارتباطات لازم برای تحقق این هدف عاملی بازدارنده در اجرای ان‌ها برای محیط‌های واقعی است.

۲. هزینه بی نظمی

در سیستم‌های کامپیوتری، موقعیت‌هایی است که باید میان به کارگیری راه حل مرکزی و راه حل توزیع شده انتخاب نمود. در راهکارهای مرکزی به هر عامل دیکته می‌شود که چه کارهایی را باید انجام دهد. در راه حل توزیع شده، هر عامل سعی در بهینه کردن منفعت شخصی دارد. بدیهی است که در راه حل‌های مرکزی، بیشینه کردن سود اجتماعی امکان پذیر است، لیکن در مقابل خرابی آسیب پذیرتر بوده و با موقعیت‌های واقعی که عامل‌ها مستقل اند سازگاری ندارد. حال با توجه به محاسبن سیستم‌های توزیع شده این سوال مطرح می‌شود که با انتخاب راه حل توزیع شده چه میزان از کارایی بهینه که در سیستم‌های مرکزی قابل حصول است دور می‌شویم؟ برای اندازه‌گیری این موضوع از تناسبی به نام هزینه بی نظمی که نسبت سود اجتماعی در حالت مرکزی به سود اجتماعی در حالت توزیع شده را اندازه‌گیری می‌کند، استفاده می‌شود. این معیار در کابرد‌های مانند تصمیم‌گیری در مورد کنترل ترافیک، اشتراک طیف امواج و ... کاربرد دارد.

۳. نظریه بازی و محاسبات توزیعی

حوزه کاربرد نظریه بازی و محاسبات توزیع شده هر دو یکسان است: محیطی که عامل‌های زیاد، عدم قطعیت و امکان وجود اهداف متفاوت، از خصوصیات آن است. لیکن تفاوت‌های بسیاری در تاکیدات دو حوزه وجود دارد: در محاسبات توزیع شده تاکید به روی مسائلی مانند تحمل پذیری در مقابل خطا، هم گام سازی، مقیاس پذیری و اثبات صحت الگوریتم است. در حالی که در نظریه بازی تاکید بر روی رفتار استراتژیک می‌باشد.

جمع بندی

دانشمندان حوزه نظریه بازی تمرکز بالایی بر روی روش‌های حل مساله بدون توجه به محدودیت‌های عملی آن داشته‌اند. با ورود محققان علوم کامپیوتر جنبه‌های پیاده سازی این روش‌ها در محیط‌های واقعی مورد بررسی قرار گرفته است. از سویی دیگر بسیاری از این روش‌ها در محیط‌های توزیع شده مانند اینترنت به کار گرفته شده‌اند. بنابراین می‌توان حلقه اشتراک علم کامپیوتر و نظریه بازی را در دو حوزه خلاصه نمود: بررسی امکان پذیری روش‌ها و سازوکارهای تئوری بازی در محیط‌های واقعی به کار گیری روش‌های تئوری بازی در زیرساخت‌های توزیع شده مطرح مانند شبکه‌های اجتماعی، رایانش ابری و اینترنت.

کارهای میان رشته‌ای بسیاری در علوم کامپیوتر و تئوری بازی در دهه گذشته انجام شده است. تئوری بازی یک جزء مهم بسیاری از کنفرانس‌های علوم کامپیوتر بوده و دانشمندان علم کامپیوتر در کنفرانس‌های مهم نظریه بازی دعوت شده‌اند. در این نوشتار به طور خلاصه چند کاربرد علم کامپیوتر در نظریه بازی شرح داده می‌شود.

۱. ملاحظات محاسباتی

نفوذ علوم کامپیوتر در تئوری بازی بیشتر به واسطه مبحث نظریه پیچیدگی می‌باشد. در ادامه به مطالعه چند مورد می‌پردازیم:

محاسبه تعادل نش

نش نشان داد که هر بازی در تعداد متناهی بار، دارای یک تعادل نش با فرض انتخاب راهبرد مخلوط می‌باشد. اما مساله آن است که رسیدن به نقطه تعادل چه میزان دشواری زمانی دارد؟ برای کلاس‌های خاصی از بازی‌ها، الگوریتم هایی با مرتبه چندجمله ای وجود دارند. لیکن پیداکردن تعادل نش در بسیاری از بازی‌ها دارای پیچیدگی NP-hard می‌باشد. برای نمونه نشان داده شده است که در یک بازی در فرم نرمال در حالی که هر بازیکن پرداختی با میزان حداقل دارد، پیچیدگی پیداکردن تعادل نش در هر قالب از بازی‌های مشخص نمودن پیچیدگی پیداکردن تعادل نش در هر قالب از بازی‌های مطرح، دارای ارزش بالایی است، زیرا که علاوه بر اثبات وجود تعادل نش، زمان رسیدن به تعادل دارای اهمیت کاربردی است.

طراحی سازوکارهای الگوریتمیک

مساله طراحی سازوکار، طراحی یک بازی است به گونه‌ای که هر عامل شرکت کننده در بازی تنها انگیزشی برای منافع شخصی دارد. لیکن در عین برآورده سازی اهداف شخصی، اهداف طراح بازی نیز برآورده می‌شود. به زبان دیگر با طراحی درست قوانین بازی، برآورده سازی اهداف طراح با برآورده سازی منافع شخصی بازیگران هم سو می‌شود. این مساله اشتراکات بسیاری با طراحی پروتکل‌های کامپیوتری دارد. در طراحی سازوکار، ملزومات پیچیدگی نادیده گرفته شده است. لیکن توسط محققان حوزه علوم کامپیوتر نشان داده شده است که حتی در ترتیمات ساده، برای بهینه سازی رفاه اجتماعی که معمولاً جزئی از اهداف طراح سازوکار است، پیچیدگی مساله از درجه NP-hard است. بنابراین سازوکارهای مطرح برای بهینه سازی رفاه اجتماعی در محیط‌های بزرگ قابلیت اجرای شدن ندارند. با دانستن پیچیدگی ذاتی مساله می‌توان به روش‌های تقریبی تکیه نمود. لیکن مشکل اساسی دیگری به وجود می‌آید: اثبات شده که خاصیت خودانگیزشی برای راستگویی عامل‌ها در زمانی که پاسخ تقریبی مدنظر طراح بازی باشد از بین می‌رود.

پیچیدگی ارتباطات

پیچیدگی ارتباط به بررسی این موضوع می‌پردازد که برای n عامل

آشنائی با گروه رایانه دانشگاه علامه طباطبائی

گروه رایانه دانشگاه علامه طباطبائی (ره) در سال ۱۳۹۰ با تفکیک گروه درسی "آمار، ریاضی و رایانه" (که نزدیک به ۲۵ سال فعالیت داشت) به سه گروه مجازء و تشکیل دانشکده علوم ریاضی و رایانه، تاسیس گردید. اولین دوره کارشناسی ارشد علوم رایانه در مهر ماه ۱۳۹۰ از طریق آزمون کشوری پذیرش شد. تفکیک این سه گروه و ارتقاء به یک دانشکده علوم پایه درون دانشگاه تخصصی علوم انسانی، حاصل تلاش و پیگیری اعضاء اولیه گروه و مسئولینی بود که به اهمیت و نقش علوم پایه در کنار علوم انسانی ایمان و اعتقاد داشتند.

دانشکده علوم ریاضی و رایانه متشکل از سه گروه آمار، ریاضی و رایانه در حال حاضر پذیرای یک رشته در مقطع کارشناسی، پنج رشته در مقطع کارشناسی ارشد و یک رشته در مقطع دکتری با مجموع ۴۰۰ دانشجو و ۲۰ عضو هیات علمی در خود است. به لحاظ محدودیت فیزیکی، دانشکده علوم ریاضی و رایانه فعلاً در دانشکده اقتصاد فعالیت می نماید و مسئولین دانشگاه در صدد تامین فضای فیزیکی مستقل برای این دانشکده می باشند.

گروه رایانه در حال حاضر سه عضو هیئت علمی تمام وقت و سه عضو مدعو دارد و از بدوم تاسیس هر سال در دو گرایش (زمینه) سیستم های رایانه ای و سیستم های هوشمند دانشجو پذیرفته است. این گروه با اتکاء به تجربیات علمی، پژوهشی و اجرائی اعضاء و حمایت های مجموعه مدیریت دانشگاه، افق توسعه هم به لحاظ آموزشی و پژوهشی و هم به لحاظ نیروی انسانی، تجهیزات و فضای فیزیکی برای دو سال آینده ترسیم نموده و به فضل الهی به سوی آن قدم بر می دارد.





آقای دکتر حسن روشنیدی

دانشیار، دکتری علوم رایانه از دانشگاه اسکس انگلستان،

زمینه پژوهشی: مهندسی نرم افزار، سیستم های پشتیبانی تصمیم گیری و زبان های برنامه سازی

Email: hrashi@atu.ac.ir



آقای دکتر محمد رضا اصغری اسکوئی

استادیار، دکتری علوم رایانه از دانشگاه اسکس انگلستان، مدیر گروه

زمینه پژوهشی: سیستم های هوشمند، برنامه ریزی منطق، پردازش تصویر و بینائی ماشین، روباتیک

Email: oskoei@atu.ac.ir



خانم دکتر فرشته آزادی پرند

استادیار، دکتری مهندسی رایانه از دانشگاه علم و صنعت ایران،

زمینه پژوهشی: نظریه علوم رایانه، پایگاه داده توزیعی،

Email: parand@atu.ac.ir

دروس گروه علوم رایانه – کارشناسی ارشد

سیستم های هوشمند	سیستم های کامپیوتری	دروس
۱۰۱ - نظریه علوم کامپیوتر	۱۰۱ - نظریه علوم کامپیوتر	الزامی - مشترک
۵۰۱ - هوش مصنوعی پیشرفته ۵۰۲ - برنامه سازی منطق	۴۰۱ - نظریه طراحی سیستم ها ۴۰۲ - طراحی نرم افزار پیشرفته	الزامی - تخصصی
۵۰۴ - پردازش تصویر ۵۰۵ - منطق محاسباتی ۵۰۶ - بینائی ماشین ۵۰۸ - یادگیری ماشین ۵۱۰ - شبکه های عصبی	۴۰۳ - سیستم های عامل پیشرفته ۴۰۷ - سیستم های تصمیم یار ۴۰۴ - پایگاه داده پیشرفته ۴۰۵ - پایگاه داده توزیعی ۴۱۵ - الگوریتم های زیستی	اختیاری - تخصصی
۱۰۲ - سمینار ۱۰۳ - پایان نامه	۱۰۲ - سمینار ۱۰۳ - پایان نامه	پژوهشی
۹۱۰ - ریاضیات مهندسی	۹۱۰ - ریاضیات مهندسی	پیش نیاز

همایش ها و کنفرانس های علوم رایانه

۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۰۱	دانشگاه شهید بهشتی و کمیته دائمی کنفرانس برق	بیست و دومین کنفرانس مهندسی برق ایران
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۰۱	انجمن بیوانفورماتیک ایران و مرکز تحقیقات بیوشیمی	پنجمین همایش بیوانفورماتیک ایران
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۰۶	بیوفیزیک (IBB)	ششمین کنفرانس فناوری اطلاعات و دانش
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۰۷	دانشگاه صنعتی شاهroud	اولین کنفرانس ملی ریاضیات صنعتی
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۰۸	دانشگاه تبریز	همایش ملی مهندسی رایانه و مدیریت فناوری اطلاعات
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۱۰	شرکت علم و طلوع فرزین	اولین همایش منطقه ای فناوری اطلاعات برق پالایش
۱۳۹۳/۰۲/۳۰ - ۱۳۹۳/۰۳/۲۲	مرکز علمی کاربردی گچساران ۱	اولین همایش ملی مهندسی برق و کامپیوتر در شمال کشور
۱۳۹۳/۰۴/۰۶	موسسه آموزش عالی موج	چهارمین کنفرانس بین المللی مدیریت فناوری اطلاعات، ارتباطات و کامپیوتر
۱۳۹۳/۰۵/۰۹	موسسه آموزشی تحقیقاتی نورباران اندیشه	همایش ملی الکترونیکی دستاوردهای نوین در علوم مهندسی و پایه
۱۳۹۳/۰۵/۲۸ - ۱۳۹۳/۰۵/۳۰	مرکز پژوهشی زمین کاو	ششمین کنفرانس مهندسی برق و الکترونیک ایران
۱۳۹۳/۰۵/۲۸ - ۱۳۹۳/۰۵/۳۰	دانشگاه آزاد اسلامی گناباد	چهاردهمین کنفرانس سیستم های فازی ایران
۱۳۹۳/۰۶/۱۲ - ۱۳۹۳/۰۶/۱۳	دانشگاه صنعتی سهند تبریز	یازدهمین کنفرانس بین المللی انجمن رمز ایران
۱۳۹۳/۰۸/۱۸	انجمن رمز ایران	دویمین کنفرانس دستاوردهای نوین در مهندسی برق و کامپیوتر
۱۳۹۳/۱۲/۱۲ - ۱۳۹۳/۱۲/۱۴	دانشگاه آزاد اسلامی واحد جوبیار	بیست و سیمین کنفرانس ملی سالانه انجمن کامپیوتر ایران
	انجمن کامپیوتر ایران، دانشگاه فردوسی مشهد	سیزدهمین کنفرانس سیستم های هوشمند ایران
	انجمن ماشین بینایی و پردازش تصویر ایران	نهمین کنفرانس ماشین بینایی و پردازش تصویر ایران

References

1. Smelik, R.M., Tutenel, T., Bidarra, R., Benes, B.: A survey on procedural modelling for virtual worlds. *Computer Graphics Forum* (2014) n/a–n/a
2. Zhou, H., Sun, J., Turk, G., Rehg, J.M.: Terrain synthesis from digital elevation models. *IEEE Transactions on Visualization and Computer Graphics* 13(4) (July 2007) 834–848
3. Gain, J., Marais, P., Stra, W.: Terrain sketching. In: *Proceedings of the 2009 Symposium on Interactive 3D Graphics and Games. I3D '09*, New York, NY, USA, ACM (2009) 31–38
4. Hnaidi, H., Gurin, E., Akkouche, S., Peytavie, A., Galin, E.: Feature based terrain generation using diffusion equation. *Computer Graphics Forum* 29(7) (2010) 2179– 2186
5. Longay, S., Runions, A., Boudon, F., Prusinkiewicz, P.: Treesketch: Interactive procedural modeling of trees on a tablet. In: *Proceedings of the International Symposium on Sketch-Based Interfaces and Modeling. SBIM '12*, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2012) 107–120
6. Anastacio, E., Prusinkiewicz, P., Sousa, M.C.: Sketch-based parameterization of l-systems using illustration-inspired construction lines and depth modulation. *Computers and Graphics* 33(4) (2009) 440 – 451
7. Lieberman, H.: User interface goals, ai opportunities. *AI MAGAZINE* 30(4) (2009)
8. SAUNDERS, R.L.: *Terrainosaurus realistic terrain synthesis using genetic algorithms*. Master's thesis, Texas AM University (2006)
9. Burt, T.: *Interactive evolutionary computation by duplication and diversification of l-systems*. Master's thesis, University of Calgary (2013)

and shown to the user. The user chooses the plant that is closest to what he wants. The program uses evolutionary programming and slightly change the L-System of the chosen plant and generates a new set of plants. The new plants are similar to what the user has chosen. This happens because the chosen plant serves as a parent and main program for generating all the new child plants. Again the user chooses one plant and the process is repeated until the user gets what he wants. Diversity of the generated plants are controllable by the user. To use this program no knowledge about the L-System and its complicated grammars is required. The user just selects the plants that he thinks are more similar to what he wants. This program completely demonstrates the power of AI for making simple user interfaces. Figure 4 shows two generation of plants and their relation.

6- Conclusion

User interaction is important and applications must be designed such that it would be easy to learn and use them. Moreover, applications have to provide sufficient tools for the user to achieve his goal. Designing user interface/interaction for PM applications is a challenging task. This is mostly because of complex and compressed structure of PM. Many researchers have tried to address this problem by utilizing sketch based interface. They try to translate sketches into rules and parameters of PM. However, there are some problems. For terrain generation, the current tools do not provide enough local control for the user. They also do not let the user to remove or erase a portion of the terrain to create holes and caves. One possible solution which help the designers to create better interfaces is to use AI methods. AI can help by amplifying the input data and also.

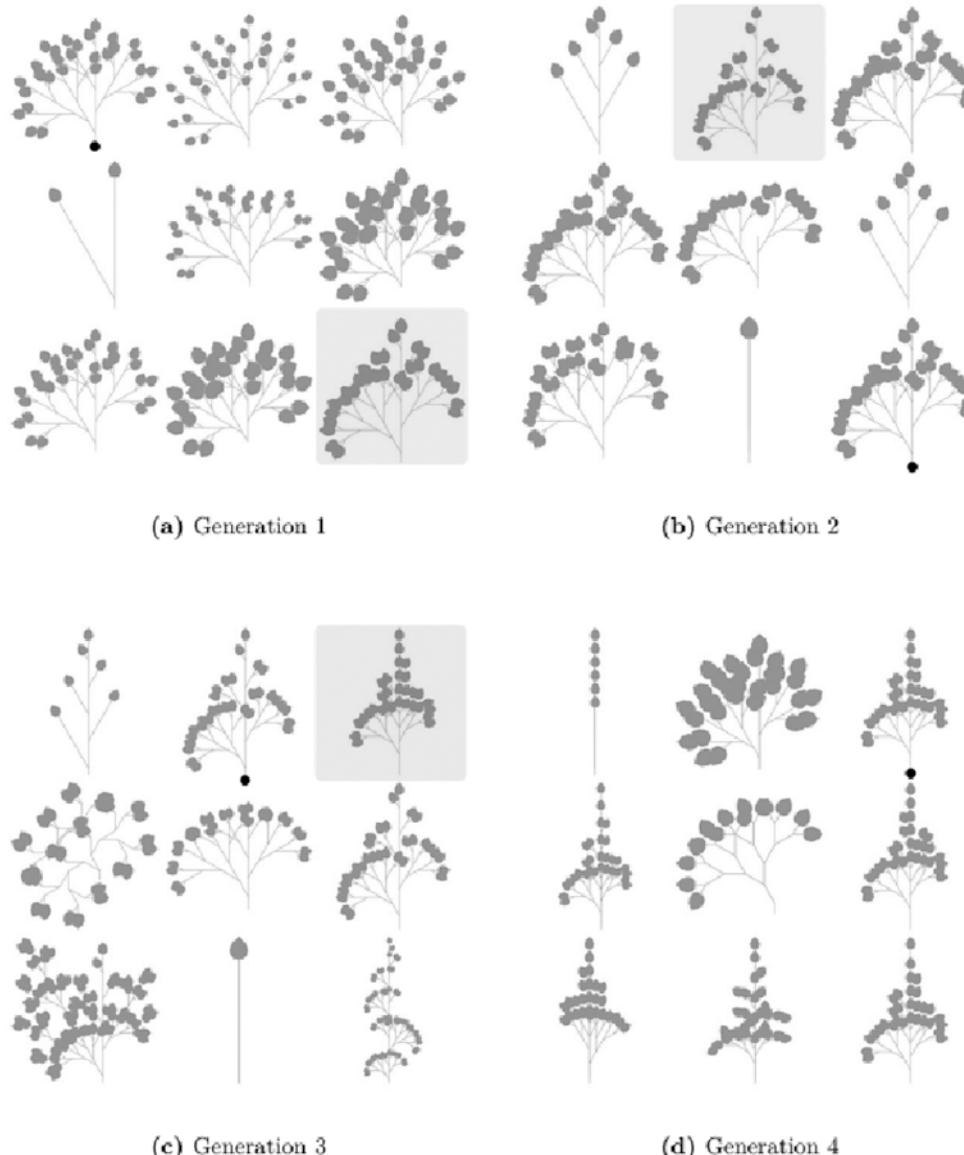


Fig. 4: Different generation of plants created by evolutionary programming with user-based selection.[9]

compressing the information. While the literature is filled with many different PM algorithm, it is still an open question how to control the complex behavior of procedural models. The UI for PM cannot be simple because the ability of user to create complex objects is reduced. Hence, even in sketch based systems the user is either limited or has many options. It seems that AI is the promising way of handling this problem as it can amplify the inputs. In conclusion, despite many recent advancements, there is no general solution for creating good UI for PM and it is still an open question to create interactive PM software programs that allow the user to have control over all the features, local and global parameters of PM. For future the researchers must also consider the problem of integrating different PM tools and techniques and try to find a general way of handling the user interaction with PM

define the overall shape of the plant and its branches. Another way for interactively modifying the plant is to translate the user input into some parameters of the growth. For example, in some algorithms for generating trees the amounts of light that the tree receives affects its growth and hence, the user can interactively change the light position or its intensity in the environment to control the results. There are also some types of algorithm for development of trees which work based on competition and amount of resources available. One possible way for controlling the tree growth in these applications is to translate the sketch into resources like water.

In general for controlling PM of trees there are four important factors which all can be controlled interactively if the generating algorithm supports them. 1- limb shape 2- branch distribution and proliferation 3- topiary shape of tree 4- gravimorphism. TreeSketch is an example application that provides interactive edition of all the factors. The problem is that sometimes controlling all these factors simultaneously is hard and confusing. And some simplification is required to increase the usability of these kind of applications.

5- Role of Artificial Intelligence

Can artificial intelligence help us to improve user interaction with PM? A good user interface make the program easier to use and learn for the user. If we look at the activity cycle, the ideal AI can help users to form their intention and action plans and even to execute them. However, the current AI system are unreliable. Their behavior are undesirable sometimes. This happens mostly because many AI algorithms are based on heuristic search methods. Lieberman argues in his paper [7] this is not something that prevents designers from using AI methods for the design of UI. Although the performance of many AI algorithms is not satisfactory, using AI methods with delicate care can improve the user interface.

Many AI algorithms act like data amplifier. They can add details and more information to the input. In PM these AI methods can receive the high level information from the user and translate it to many detailed parameters and procedures of PM. This can help us to create interfaces that are more intuitive and easy to use. Furthermore, by using AI we can develop PM applications for users

who has a little or no knowledge about PM. On the other hand AI methods can be used for compressing the complex information of PM rules, grammars and parameters in a sensible way. This can be used in the gulf of evaluation to provide sensible feedback for the user.

Saunders et al. [8] proposed to use genetic algorithm for generating terrains. GA is a common meta-heuristic search algorithm. They use GA to create a height map that matches the users requests. In fact the user sketches some region on the terrain and assigns a type of elevation profile to each of them. Then the program has to search for the right elevation map (in a library of terrain data) that is close to the input as much as possible. It seems that his method of terrain generation is not a PM method but the library that he uses can be generated using PM methods.

Burt et al.[9] uses the evolutionary programming algorithm with L-System to produce different types of plants. His programs interface is quite simple. First a set of diverse plants are generated and shown to the user. The user chooses the plant that is closest to what he wants. The program uses evolutionary programming and slightly change the L-System of the chosen plant and generates a new set of plants. The new plants are similar to what the user has chosen. This happens because the chosen plant serves as a parent and main program for who has a little or no knowledge about PM. On the other hand AI methods can be used for compressing the complex information of PM rules, grammars and parameters in a sensible way. This can be used in the gulf of evaluation to provide sensible feedback for the user.

Saunders et al. [8] proposed to use genetic algorithm for generating terrains. GA is a common meta-heuristic search algorithm. They use GA to create a height map that matches the users requests. In fact the user sketches some region on the terrain and assigns a type of elevation profile to each of them. Then the program has to search for the right elevation map (in a library of terrain data) that is close to the input as much as possible. It seems that his method of terrain generation is not a PM method but the library that he uses can be generated using PM methods.

Burt et al.[9] uses the evolutionary programming algorithm with L-System to produce different types of plants. His programs interface is quite simple. First a set of diverse plants are generated

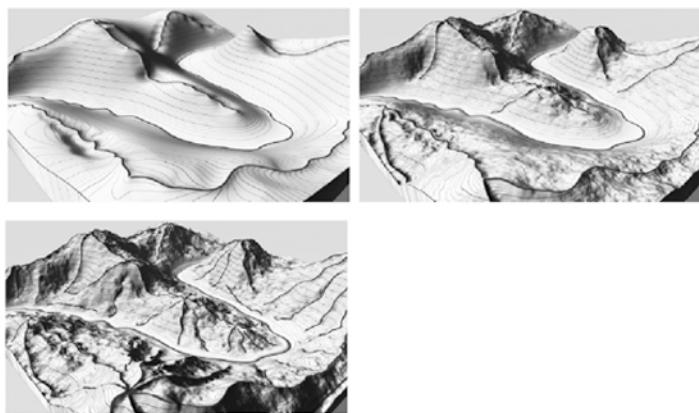


Fig. 3: Step mode user interaction for generating terrain.⁴

for generating terrains have unpredictable and non-intuitive control parameters (especially because they have stochastic behaviors). The other applications provide better control for users with sketch based interfaces. However their interface doesn't provide enough control for the user to design what he intends. In these applications the sketch based interfaces are limited to what the program offers and the user can apply a limited number of options. Therefore sometimes translation of the user intends to the right terrain is impossible or hard to achieve. Especially because most of them do not support local editing and erasing operations.

4- User Interaction with Procedural Modeling of Plants

PM has been used for modeling of plants and trees since the development of L- System. L-System is a framework developed by biologists to model the growth of trees and plants. It is the most common and convenient way of modeling plants. L-System works based on a set of linear grammars that generate the object.

L-Studio is an application that provides an editor and some tools to work with open L-systems. The program is designed for the first group of users who know how PM works and how to program L-Systems. Working with this application is not easy even for professional users. Most of the times a small change in one parameter or grammar can devastate the result and there is no convenient way for debugging it. While it is hard to learn and use it, it is quite flexible and powerful as it allows the user to explore all the possible features of L-Systems and produce any kind of objects. For developing plants by L-Studio the user defines the

appropriate grammars and creates basic objects like leaves and runs the simulator. This activity process is not interactive and intuitive. However, it is possible to use a special set of grammars that allows the user to intuitively control the plant's growth by sketching, brushes and mouse interaction. In order to do this the programmer has to use environmental queries. Environmental queries are a set of query parameters that are set during simulation by the program. The user can use environmental queries to use the mouse positions, etc. to modify the model interactively.

For plant modeling the sketch based interface can be used for setting different types of parameters and constraints. First, the system can provide a brush. The user moves the brush and the plant grows toward the new brush position. This tool acts like an attractor for the plant growth. As another interactive tool the simulation can start by a simple internode and apex and then the user interactively selects the position of new branches to grow. In more advanced editing mode the user can sketch the shape of the new branch in addition to its position.

The only problem of using this interface is that the user sketch is a 2D curve while the new branch grows in a 3D environment. Hence the program has to handle this such that the resulting branch matches the user sketch and also provides a botanically plausible and nice looking result. The sketch can also be used for defining the overall shape of a plant. In TreeSketch [5] the user can sketch the exterior boundary of a tree. This is in fact a virtual topiary. The most challenging problem of creating this interface is to find a way for converting the 2D sketch into a 3D bounding surfaces. There are two common ways for doing this. First, to use extrusion. In extrusion a copy of the sketched curve is translated along the normal of projection surface and then the two curves are sewn up to form a 3D bounding mesh. Another way is to use inflation. In inflation the sketched curve forms the silhouette and boundary of the 3D object and the silhouette is inflated such that its center has the maximum depth and the boundary has no depth at all. The inflation method is more natural for using in PM. Anastacio et al.[6] proposed an interactive sketching method for creating plants. In their system (implemented by L-System) the user sketched three strokes to

the user to help him satisfy his requirements. Failure takes place if the application doesn't provide the proper tools or has a complicated control mechanism. For terrain synthesizing applications, the most important tool is the ability to control the shape of the terrain. This control must be scale independent which means that the user must be able to perform small changes in small regions as well as defining large scale and global changes. On the other hand this control must be intuitive and easy to learn and use.

One of the best intuitive interfaces is sketch based interfaces and brushes. There are some technical challenges when a system uses this type of interfaces. When the user sketches something, the application has to interactively interpret the sketch and translates it into the proper changes. If the translation is not done appropriately then the behavior of the application becomes unpredictable. Another important fact is that a sketch is done on a 2D plane while the terrain is defined in the 3D domain. Therefore we need to find a proper mapping between the sketch and our 3D environment. To find this mapping, for each point we should cast a ray from the camera to 3D environment and find the intersect-

ing object (there are other ways to find the mapping like using zed buffer or rendering with special color assigned to each object etc.). Finding this mapping (un-projection mapping) is sometimes hard in PM. If the PM App uses view dependent generating algorithms then finding the right mapping is impossible. This is because that based on the camera position, it is possible that a portion of our sketch points to some far objects which are not created with enough details (as they are far from the camera). Therefore the right correspondence between the sketch and terrain couldn't be found such that the user modification takes place accordingly.

Sketch based interaction can be done in two different ways. First it can be a continuous interaction. This means that as the user sketches the program interprets the sketch and apply changes interactively and smoothly. Brushes provide continuous interaction as their strokes are interpreted individually regardless of the previous changes and brush instances. In the other way sketching can be implemented as a step mode interaction. In this way, first the user sketches and then he notifies the program to interpret and

apply the appropriate changes. Interpretation of sketches in the second way is harder as they can get quite complex. In the continuous mode the interpretation is restricted to the limited local area of the brush stroke and hence it is much simpler. In literature both methods have been used.

Zhou et al [2] proposed procedural brushes for editing terrains. They allow the user to use a continuous brush with arbitrary brush stroke to add features to the terrain. The user can also control the size of brush strokes. The main problem of their application is that there is no way for erasing and removing a feature or portion the terrain to create

intuitive for the user to define a brush stroke and therefore he will be limited to the pre-defined strokes of the application. There also other terrain generator applications that use sketch based interfaces. Gain et al. [3] introduced a terrain generating application that allows the user to sketch the silhouette of the mountains. Their program interprets the sketch and generates the mountain. They also allow the user to modify the mountain using a curtain based sketching system [Fig 2]. The main problem of this algorithm is that it doesn't provide any control tools for adding local and small features and also adding caves and holes. Hnaidi et al. [4] proposed a terrain generating algorithm that uses a step mode sketch interface (Figure 3). The user sketches some feature curves on the current terrain. Then the program interprets the sketches, produces some vector fields and generates a terrain based on the vector field. The main problem of this method is that the user has no control over the local changes and again the system doesn't support holes and caves. In conclusion, based on what I have found many of the PM applications

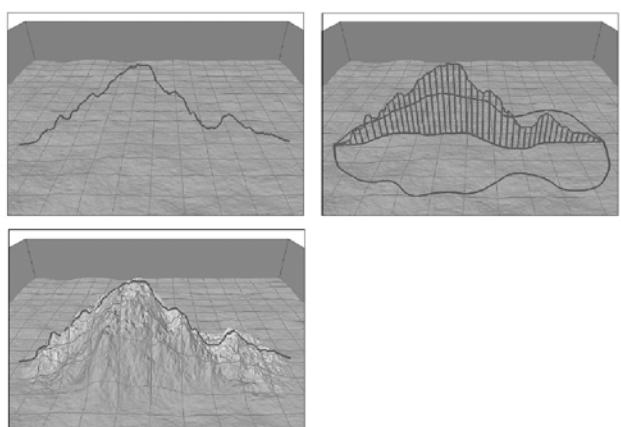
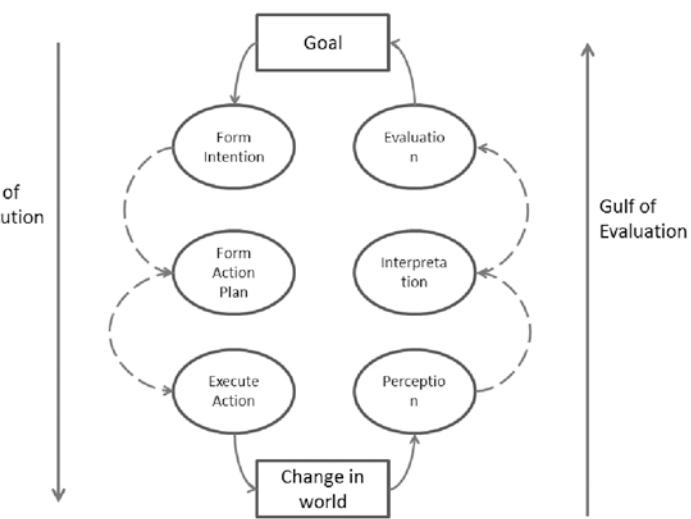


Fig. 2: Sketch based generation of mountains.3

interprets the information and realizes that something is wrong in the code. The last step is evaluations. The human evaluates what has happened in the world against what he intended. For our example the evaluation is done immediately and the user realizes that there is problem in the code. However, sometimes it may take a while for the user to evaluate the results. This may occur due to the limitations of what the world presents or due to the lack of knowledge of the user for comparing the results with what the goal is. After the evaluation the user may form new intentions and the action cycle is repeated. Regarding the action



cycle there are some design concerns which should be taken into account during the design process of UI. How much the user is knowledgeable is the key concept for designing UI. A good UI in the gulf of execution helps the users to form the goals, tasks and action plans. It should be designed such that the user can easily learn to use it and the actions provided by the system match the users requirements. In the gulf of evaluation, the UI must provide sufficient informative feedback for the user such that the user can perceive the systems state and interpret it. In addition UI must not provide too much information as it is confusing and misleading for the user. This again depends on the users level of skill and knowledge.

3- User Interaction with Procedural Modeling of Terrains

Terrain modeling is one of the most common application of PM. The simplest procedural method for generating a terrain is the triangle subdivision algorithm.

In this algorithm we start by a mesh of triangles (usually a 2D grid) and at each iteration we subdivide the triangles and displace the new vertices by random numbers. In the midpoint displacement algorithm, for each triangle only one new vertex is generated which its elevation is set to the average of its neighbors plus a random number [1]. The range of random numbers decreases after each iteration based on a parameter. For this algorithm, the user can only change the number of subdivision steps and the range of random numbers. The user has no control over the shape of the terrain and cannot modify the terrain locally. And besides because the model is based on random numbers, the user gets different results by undoing and redoing a subdivision step (unless the program saves the previous generated random numbers). All the stochastic methods have this problem. Another problem of these algorithms is that the meaning of parameters is not intuitive. The user has no idea of changes happening by increasing or decreasing the parameters. Particularly because of high sensitivity of these algorithms to the initial parameters, even a small change in the parameters can produce something completely different. Setting these parameters is even hard for the first group of users (knowledgeable users) because the results are almost unpredictable. Based on the activity cycle the main problem of this algorithm is that the system doesn't provide sufficient tools to meet the users requirements and therefore the user cannot plan the right actions. On the other hand it is hard for the user to interpret the results to tune the parameters after the evaluation step.

In some more advanced applications, some physical phenomenon like erosion and weathering is simulated over the terrain to produce more realistic results. However they still suffer from the same interaction problems. Another problem of these methods is that they may not be interactive due to the high computation cost of simulation. Having non-interactive applications significantly reduces the usability of that application.

Before we continue to analyze other PM tools, let's discuss what a user requires from the terrain generating applications and how we can address it. The users goal is to model a specific terrain. To this end the user forms an action plan. The application must provide an intuitive user interface for

users of our product. In this case we have two groups of people who may want to use the software.

The first group consists of all the people who want to explore the foundation of procedural modeling and want to experience and play with the low level structure of the procedural models. Scientists, students and innovators are members of this group. For this group of users the designer should propose a system that allows the users to conveniently edit and modify all the parameters and components of a procedural model.

Members of the second group are artists. They want to use the product for making beautiful and nice looking objects. They do not want to work with the low level parameters and details of procedural models. They do not even know anything about how procedural modeling works. So the product should be designed such that all the related information and algorithm is concealed and only a nice and simple interface is provided for the user. Obviously the goals, skills and knowledge of the groups are quite different, opposing and non-homogenous. Therefore the UI design is completely application and user oriented. The UI design is even harder for PM software products as there is a tension between the control of user over the object and the complexity of objects being modeled. Our focus in this essay is on the second group of users while we consider the first group too.

Before we start analyzing the current procedural modeling methods, we need to establish criteria for measuring how successful a procedural model is for interacting with users and how should it be designed. In the context of human computer interaction there is common concept called Action cycle. This concept was proposed by Donald Norman (1988) in his book *The Design of Everyday Things*. Norman identified two gulfs existing between the human and world. The human has a goal and wants to achieve that goal. The world is the one that must be changed in order to possibly satisfy the humans goal. In our case is the PM application is the world and the users goal is to create a specific object (of course the users goal is not limited to only this one and it can be defined more accurately). The first gulf is the gulf of execution. The human knows what he wants but he doesn't necessarily know how to achieve the goal. This problem is captured by the gulf of execution.

The second gulf is the gulf of evaluation. When something is changed in the world, how well the human can interpret, understand and evaluate that change in the world. This question is captured by the gulf of evaluation. Each gulf has three activities. Figure [1] shows these activities. For the gulf of execution the first activity is forming intention. This is just a cognitive effort for achieving the goal. For example in the case of procedural modeling this means that the user decides to use a computer to make the 3D object (which is the goal).

The next activity is forming an action plan. The action plan describes all the steps that the human would take to change the world and satisfy the intentions. For example for procedural modeling the user decides on the program he wants to run (L-Studio for example) and the grammars that he wants to use, enters the grammars and executes them (by CPFG for instance). This activity is not easy for the user. The human requires having sufficient domain and task knowledge to be able to formulate an action sequence. For example if the user intends to model a tree then he should know how to use L-Studio and what the right grammars are for making a tree.

The last step in the gulf of execution is executing actions. This is a physical step. For example the user clicks on the program icon and enter the grammars using a keyboard and then he runs CPFG. The product should be designed such that the user can easily learn it and executes his actions through it. The execution doesn't necessarily end up with what the user has intended. For example, a syntactical mistake in the grammars prevents CPFG to run. If everything works based on the action plan then something changes in the world and we enter into the next gulf. The first step in the gulf of evaluation is perception. The perception state involves all the things that the human notices after execution. The next step is interpretation. The human interprets what he has perceived. The interpretation depends on the humans domain knowledge and the information that he has perceived. It is possible that the changes in the world do not make sense for the human and therefore he can't interpret them properly. For example, the user runs CPFG but because there is a syntactical error a black screen is shown with a message that there is error. The user sees these changes (perception) then .

Survey

Procedural Modeling And User Interaction

Amirhessam Moltaji

amoltaji@ucalgary.ca

University of Calgary

Abstract:

Procedural Modeling is a generative approach for creating 3D models and objects in the world of Computer Graphics. It is known for its strength in data amplification and reducing modeling' time and cost. However, it is not convenient and easy to use PM techniques directly and 3D artists are not expert enough to use it. There are lots of researches done for creating easy to use interfaces for PM.

In this essay, we analyze the allure, conceptual problems and technical challenges of user interaction with procedural modeling of plants and terrains. In addition, we survey some of the recent papers and achievements in PM. Moreover, we analyze the effectiveness of AI methods for creating UI for PM.

Procedural modeling (PM) is a paradigm for making 3D object in the world of computer graphics. Instead of modeling objects individually with all the details, we can define some rules that generate the same object. PM is a wide area. It has been applied for modeling various types of objects such as plants, buildings and cities, terrain and road net-

cedures and rules to automatically create an object. The main feature of PM is its data amplification capability. You can produce complex and various objects by a set of simple rules. PM reduces the amount of time and effort required for modeling many objects. However, because of the compact and complicated representation, it is quite hard to understand for many users how to use PM and how to modify its rules and parameters to get the right results. Most of the time the results of a change in PM parameters and rules are unpredictable. And it is frustrating for a user to find the right parameters. To alleviate this problem there are many approaches proposed in the literature.

In this essay first I talk about the importance of UI and how it is evaluated. Then I analyze the key concepts and challenges of user control/interaction with procedural model of terrains and plants. Following that discussion I talk about the role of Artificial intelligence in designing UI for PM. In the end I conclude the essay with

a summary of what has discussed.

Importance of User Interaction

Why user interaction is important? Is it even important? The questions sound simple. We are making products for people to buy and people buy the stuff that they can use. If you make something that nobody can use then nobody would buy it. Just imagine you have a desktop computer without mouse and keyboard. How would you use it? Is it even worth to buy it?

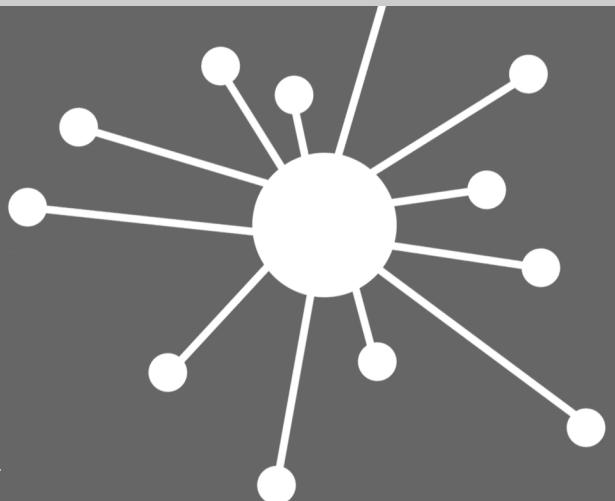
Users are the most important aspect of product development and designers always prioritize the users goals and intention during the design process. However it is not possible always to make a product that anyone in the world can use. People are different. They have different skills, ages, languages, etc. And most importantly not everyone in the world is interested to use your product. So the first step toward the design process is to recognize the different groups of users that possibly will use it. After finding the right group of users then you need find out what the best and easiest way is for them to use your product and then try to make it. Now consider that we want to create a software product based on procedural modeling. As mentioned earlier the first step is to find the possible

مکانیزم

معلمات

ALLAMEH TABATABA'I UNIVERSITY
MATH AND COMPUTER SCIENCE SCIENTIFIC ASSOCIATION

Computer Science



COMPUTER SCIENCE JOURNAL, VOLUME 1, NO 1, SUMMER 2014