

صلى الله عليه وسلم



دانشگاه تربیت مدرس

دانشکده مهندسی برق و کامپیوتر

گروه کامپیوتر

پایان نامه کارشناسی ارشد رشته فناوری اطلاعات گرایش امنیت اطلاعات

عنوان پایان نامه

روشی برای گمنام سازی درجه با حفظ نیازمندی های حریم خصوصی در گراف های شبکه اجتماعی

مشارش

علی شینی

استاد راهنما

دکتر مهدی آبادی

استاد شاور

دکتر بهزاد اکبری

بهمن ماه ۹۳



دانشگاه تربیت مدرس  
دانشکده مهندسی برق و کامپیوتر

بسمه تعالی

### تاییدیه اعضای هیات داوران حاضر در جلسه دفاع از پایان نامه کارشناسی ارشد

آقای علی شیخی پایان نامه ۶ واحدی خود را با عنوان روشی برای گمنام سازی درجه با حفظ نیاز مندی های حریم خصوصی در گراف های شبکه اجتماعی در تاریخ ۱۳۹۳/۱۱/۲۱ ارائه کردند.

اعضای هیات داوران نسخه نهایی این پایان نامه را از نظر فرم و محتوا تایید کرده، پذیرش آنرا برای اخذ درجه کارشناسی ارشد مهندسی فناوری اطلاعات-امنیت اطلاعات پیشنهاد می کنند.

عضو هیات داوران	نام و نام خانوادگی	رتبه علمی	امضا
استاد راهنما	دکتر مهدی آبادی	استادیار	
استاد مشاور	دکتر بهزاد اکبری	استادیار	
استاد ناظر	دکتر فواد قادری	استادیار	
استاد ناظر	دکتر مسعود اسدیپور	استادیار	
مدیر گروه (یا نماینده گروه تخصصی)	دکتر فواد قادری	استادیار	



## دستورالعمل حق مالکیت مادی و معنوی در مورد نتایج پژوهشهای علمی دانشگاه تربیت مدرس

**مقدمه:** با عنایت به سیاست‌های پژوهشی دانشگاه در راستای تحقق عدالت و کرامت انسانها که لازمه شکوفایی علمی و فنی است و رعایت حقوق مادی و معنوی دانشگاه و پژوهشگران، لازم است اعضای هیات علمی، دانشجویان، دانش‌آموختگان و دیگر همکاران طرح، در مورد نتایج پژوهشهای علمی که تحت عنوان پایان‌نامه، رساله و طرحهای تحقیقاتی که با هماهنگی دانشگاه انجام شده است، موارد ذیل را رعایت نمایند:

**ماده ۱-** حقوق مادی و معنوی پایان‌نامه‌ها / رساله‌های مصوب دانشگاه متعلق به دانشگاه است و هرگونه بهره‌برداری از آن باید با ذکر نام دانشگاه و رعایت آیین‌نامه‌ها و دستورالعمل‌های مصوب دانشگاه باشد.

**ماده ۲-** انتشار مقاله یا مقالات مستخرج از پایان‌نامه / رساله به صورت چاپ در نشریات علمی و یا ارائه در مجامع علمی باید به نام دانشگاه بوده و استاد راهنما مسئول مکاتبات مقاله باشد.

تبصره: در مقالاتی که پس از دانش‌آموختگی بصورت ترکیبی از اطلاعات جدید و نتایج حاصل از پایان‌نامه / رساله نیز منتشر می‌شود نیز باید نام دانشگاه درج شود.

**ماده ۳-** انتشار کتاب حاصل از نتایج پایان‌نامه / رساله و تمامی طرحهای تحقیقاتی دانشگاه باید با مجوز کتبی صادره از طریق حوزه پژوهشی دانشگاه و بر اساس آئین‌نامه‌های مصوب انجام می‌شود.

**ماده ۴-** ثبت اختراع و تدوین دانش فنی و یا ارائه در جشنواره‌های ملی، منطقه‌ای و بین‌المللی که حاصل نتایج مستخرج از پایان‌نامه / رساله و تمامی طرحهای تحقیقاتی دانشگاه باید با هماهنگی استاد راهنما یا مجری طرح از طریق حوزه پژوهشی دانشگاه انجام گیرد.

**ماده ۵-** این دستورالعمل در ۵ ماده و یک تبصره در تاریخ ۱۳۸۴/۴/۲۵ در شورای پژوهشی دانشگاه به تصویب رسیده و از تاریخ تصویب لازم‌الاجرا است و هرگونه تخلف از مفاد این دستورالعمل، از طریق مراجع قانونی قابل پیگیری می‌شود.

نام و نام خانوادگی

امضاء  


### آیین نامه چاپ پایان نامه (رساله) های دانشجویان دانشگاه تربیت مدرس

نظر به اینکه چاپ و انتشار پایان نامه (رساله) های تحصیلی دانشجویان دانشگاه تربیت مدرس، مبین بخشی از فعالیتهای علمی - پژوهشی دانشگاه است بنابراین به منظور آگاهی و رعایت حقوق دانشگاه، دانش آموختگان این دانشگاه نسبت به رعایت موارد ذیل متعهد می شوند:

ماده ۱: در صورت اقدام به چاپ پایان نامه (رساله) ی خود، مراتب را قبلاً به طور کتبی به «دفتر نشر آثار علمی» دانشگاه اطلاع دهد.

ماده ۲: در صفحه سوم کتاب (پس از برگ شناسنامه) عبارت ذیل را چاپ کند:


«کتاب حاضر، حاصل پایان نامه کارشناسی ارشد/ رساله دکتری نگارنده در رشته فناوری اطلاعات است که در سال ۱۳۹۲ در دانشکده برق و کامپیوتر دانشگاه تربیت مدرس به راهنمایی جناب آقای دکتر مهدی آبادی، مشاوره جناب آقای دکتر بهزاد اکبری از آن دفاع شده است.»

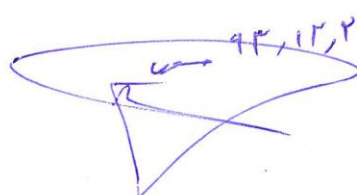
ماده ۳: به منظور جبران بخشی از هزینه های انتشارات دانشگاه، تعداد یک درصد شمارگان کتاب (در هر نوبت چاپ) را به «دفتر نشر آثار علمی» دانشگاه اهدا کند. دانشگاه می تواند مازاد نیاز خود را به نفع مرکز نشر در معرض فروش قرار دهد.

ماده ۴: در صورت عدم رعایت ماده ۳، ۵۰٪ بهای شمارگان چاپ شده رابه عنوان خسارت به دانشگاه تربیت مدرس، تأدیه کند.

ماده ۵: دانشجو تعهد و قبول می کند در صورت خودداری از پرداخت بهای خسارت، دانشگاه می تواند خسارت مذکور را از طریق مراجع قضایی مطالبه و وصول کند؛ به علاوه به دانشگاه حق می دهد به منظور استیفای حقوق خود، از طریق دادگاه، معادل وجه مذکور در ماده ۴ را از محل توقیف کتابهای عرضه شده نگارنده برای فروش، تامین نماید.

ماده ۶: اینجانب علی شیخی دانشجوی رشته فناوری اطلاعات مقطع کارشناسی ارشد تعهد فوق و ضمانت اجرایی آن را قبول کرده، به آن ملتزم می شوم.

نام و نام خانوادگی: 

تاریخ و امضا: 

تقدیم به مقدس ترین واژه مادر لغت نامه دلم،

مادر مهربانم که زندگیم را دیون مهر و عطفوت آن می دانم

پدر، برادر و حامی

برادر، همراه، همسنگی

## قدردانی

بدینوسیله از زحمات جناب آقای دکتر مهدی آبادی که وقت گرانبهای خود را صرف راهنمایی‌های پژوهش حاضر نموده‌اند کمال تشکر و سپاسگزاری را داشته و برای ایشان سلامتی و سرفرازی روز افزون را از خداوند متعال آرزومندم.

با انتشار داده‌های شبکه اجتماعی فرصت‌های بسیاری برای تحقیقات فراهم می‌شود. در اثر این انتشار ممکن است حریم خصوصی کاربران حاضر در شبکه اجتماعی نقض شود. بنابراین داده‌ها قبل از انتشار باید گمنام شوند. یکی از مهم‌ترین موضوعات در انتشار داده‌های شبکه اجتماعی برای اهداف تحقیقاتی و تحلیل‌های تجاری، ایجاد یک مصالحه بین حریم خصوصی افراد حاضر در شبکه و سودمندی داده‌ها می‌باشد. تا کنون محققان روش‌های گمنام‌سازی و مدل‌های حریم خصوصی بسیاری را برای جلوگیری از بازشناسایی افراد با توجه به خصوصیات ساختاری شبکه اجتماعی ارائه داده‌اند. اما این روش‌ها به سطح نیازمندی حریم خصوصی کاربران توجهی ندارند. این امر واضح است که کاربران مختلف در شبکه اجتماعی سطح حریم خصوصی متفاوتی دارند. همچنین اغلب روش‌های ارائه شده ویژگی‌های ساختاری شبکه را حفظ نمی‌کنند. این در حالی است که بسیاری از تحقیقات انجام شده بر روی شبکه اجتماعی تنها بر روی ویژگی‌های ساختاری متمرکز هستند.

در این پایان‌نامه روشی با نام SocialPDA ارائه شده است که شبکه اجتماعی را با توجه به نیازمندی‌های سطح حریم خصوصی افراد موجود در شبکه گمنام می‌کند به طوری که مصالحه خوبی بین حریم خصوصی و سودمندی داده‌ها برقرار می‌شود. روش پیشنهادی شامل سه مرحله می‌باشد که در ابتدا گره‌ها خوشه‌بندی می‌شوند. سپس در مرحله دوم مقادیر گمنام‌سازی برای هر گره مشخص می‌شود و در مرحله پایانی با تغییراتی بر روی گراف اصلی، گراف گمنام‌شده تولید می‌شود. نتایج حاصل از ارزیابی روش پیشنهادی با روش‌های پیشین بر روی مجموعه داده Polbook نشان می‌دهد که روش پیشنهادی با فراهم نمودن حریم خصوصی یکسان ویژگی‌های ساختاری شبکه اجتماعی را بهتر حفظ می‌کند.

**کلمات کلیدی** - حریم خصوصی در داده‌های شبکه اجتماعی،  $k$ -درجه گمنامی، ریز تجمیع، سطوح مختلف حریم خصوصی، حفظ ویژگی‌های گراف.



# فهرست مطالب

۱ فصل اول کلیات .....	۱
۱,۱ مقدمه .....	۲
۲,۱ بیان مساله و فرضیات .....	۳
۳,۱ اهداف و نتایج حاصل از پژوهش .....	۴
۴,۱ جنبه‌های نوآوری .....	۴
۵,۱ مروری بر فصل‌ها .....	۵
۲ فصل دوم انتشار داده‌ها با حفظ حریم خصوصی .....	۶
۱,۲ مقدمه .....	۷
۲,۲ $k$ -گمنامی در انتشار داده‌های جدولی .....	۹
۳,۲ تحلیل شبکه‌های اجتماعی .....	۱۴
۱,۳,۲ انواع شبکه‌ها در دنیای واقعی .....	۱۴
۱,۱,۳,۲ شبکه اجتماعی .....	۱۵
۲,۱,۳,۲ شبکه اطلاعات .....	۱۵
۳,۱,۳,۲ شبکه تکنولوژیکی .....	۱۵
۴,۱,۳,۲ شبکه زیستی .....	۱۵
۲,۳,۲ تاریخچه شبکه‌های پیچیده .....	۱۶
۳,۳,۲ هدف مطالعه‌ی شبکه‌های پیچیده .....	۱۶
۴,۳,۲ تعاریف مرتبط با شبکه پیچیده .....	۱۷
۱,۴,۳,۲ گراف .....	۱۷
۲,۴,۳,۲ چگالی .....	۱۸
۳,۴,۳,۲ ضریب خوشه‌بندی .....	۱۸
۵,۳,۲ معیارهای مرکزیت شبکه .....	۱۹
۱,۵,۳,۲ درجه .....	۱۹

۲۰	..... مرکزیت میانی	۲,۵,۳,۲
۲۰	..... مرکزیت نزدیکی	۳,۵,۳,۲
۲۰	..... مرکزیت بردار ویژه	۴,۵,۳,۲
۲۱	..... معیار مرکزی رتبه‌بندی صفحه	۵,۵,۳,۲
۲۱	..... مدل‌های شبکه‌ی تصادفی	۶,۳,۲
۲۲	..... گراف تصادفی ER	۱,۶,۳,۲
۲۲	..... شبکه‌های دنیای کوچک	۲,۶,۳,۲
۲۳	..... شبکه‌های مقیاس آزاد	۳,۶,۳,۲
۲۴	..... شبکه‌های پودمانی	۴,۶,۳,۲
۲۵	..... انواع نقض حریم خصوصی در شبکه‌های اجتماعی	۴,۲
۲۶	..... افشای هویت	۱,۴,۲
۲۷	..... افشای خصیصه	۲,۴,۲
۲۹	..... افشای ارتباط اجتماعی	۳,۴,۲
۳۰	..... افشای روابط وابستگی	۴,۴,۲
۳۳	..... ریزتجمیع داده‌های تک متغیره	۵,۲
۳۳۵	..... جمع بندی	۶,۲
۳۶	..... فصل سوم حفظ حریم خصوصی در انتشار داده‌های شبکه اجتماعی	۳
۳۷	..... مقدمه	۱,۳
۳۸	..... روش‌های مبتنی بر دستکاری گراف	۲,۳
۳۹	..... روش‌های مبتنی بر $k$ -گمنامی در شبکه اجتماعی	۳,۳
۴۲	..... روش‌های مبتنی بر $k$ -درجه گمنامی	۴,۳
۴۳	..... الگوریتم برنامه نویسی پویا	۱,۴,۳
۴۵	..... روش اضافه کردن گره	۲,۴,۳
۴۶	..... روش مبتنی بر الگوریتم تکاملی	۳,۴,۳
۴۲۸	..... جمع بندی	۴,۳

#### ۴ فصل چهارم روش پیشنهادی برای گمنام‌سازی شبکه اجتماعی ..... ۴۹

۵۰	.....	۱,۴	مقدمه
۵۱	.....	۲,۴	گمنام‌سازی دنباله درجه‌های شخصی سازی شده
۵۳	.....	۳,۴	روش پیشنهادی
۵۴	.....	۱,۳,۴	افراز شخصی‌سازی شده
۵۹	.....	۲,۳,۴	گمنام‌سازی درجات
۶۲	.....	۱,۲,۳,۴	بازنمایی ذرات
۶۳	.....	۲,۲,۳,۴	بازسازی حریصانه
۶۴	.....	۳,۲,۳,۴	تابع برازش
۶۵	.....	۳,۳,۴	دستکاری گراف
۷۱	.....	۱,۲,۳,۴	بازنمایی ذرات
۷۲	.....	۲,۳,۳,۴	تابع برازش
۷۲	.....	۳,۳,۳,۴	ساخت گراف گمنام‌شده

#### ۵ فصل پنجم ارزیابی روش پیشنهادی ..... ۷۴

۷۵	.....	۱,۵	مقدمه
۷۵	.....	۲,۵	مدل ارزیابی
۷۶	.....	۳,۵	معیارهای ارزیابی
۷۸	.....	۴,۵	ارزیابی روش پیشنهادی
۷۸	.....	۱,۴,۵	ارزیابی MicroPDA
۷۹	.....	۲,۴,۵	ارزیابی ParticlePDA
۸۰	.....	۳,۴,۵	ارزیابی ModPDA
۸۲	.....	۵,۵	مقایسه روش پیشنهادی بدون شخصی سازی با سایر روش‌ها
۸۳	.....	۱,۵,۵	ارزیابی میانگین همساز در کوتاهترین فاصله
۸۴	.....	۲,۵,۵	ارزیابی تریایی
۸۵	.....	۳,۵,۵	ارزیابی بیشترین مقدار ویژه در ماتریس مجاورت A
۸۶	.....	۴,۵,۵	ارزیابی چگالی

۶,۵	مقایسه روش پیشنهادی با سایر روش‌ها	۸۷
۷,۵	جمع‌بندی	۸۹
۶	فصل ششم نتیجه‌گیری و پژوهش‌های آینده	۹۰
۱,۶	نتیجه‌گیری	۹۱
۲,۶	پژوهش‌های آینده	۹۲
۷	مراجع	۹۳

## فهرست شکل‌ها

- شکل ۱-۲ گمنام سازی ساده در یک شبکه اجتماعی کوچک. ۸
- شکل ۲-۲ مثال‌هایی از رابطه‌های اجتماعی حساس. ۳۰
- شکل ۱-۳ عمل جابجایی برای چهار گره. ۴۴
- شکل ۱-۴ الف) گراف شبکه اجتماعی ب) گراف گمنام شده شکل الف. ۵۱
- شکل ۲-۴ سه مرحله اصلی SocialPDA. ۵۴
- شکل ۳-۴ الگوریتم افراز شخصی سازی شده. ۵۶
- شکل ۴-۴ گراف  $G\Phi$  مربوط به گراف اجتماعی شکل ۱-۴ الف. ۵۷
- شکل ۵-۴ الگوریتم گمنام سازی درجه‌ها. ۶۱
- شکل ۶-۴ سه عمل دستکاری گراف. ۶۶
- شکل ۷-۴ مثالی از دستکاری گراف با سه عمل اصلی. ۶۸
- شکل ۸-۴ الگوریتم تولید مجموعه جابجایی. ۷۰
- شکل ۱-۵ نتایج بدست آمده از ارزیابی الگوریتم ModPDA. ۸۲
- شکل ۲-۵ ارزیابی معیار H. ۸۴
- شکل ۳-۵ ارزیابی معیار T. ۸۵
- شکل ۴-۵ ارزیابی معیار  $\lambda$ . ۸۶
- شکل ۵-۵ ارزیابی معیار D. ۸۷
- شکل ۶-۵ نتایج حاصل از ارزیابی SocialPDA و SocialDA. ۸۸

## فهرست جدول‌ها

- جدول ۱-۲ نمایش انواع خصیصه‌ها در جدول بیماران بیمارستان ..... ۱۰
- جدول ۲-۲ حذف خصیصه‌ی نام به منظور گمنام‌سازی ..... ۱۱
- جدول ۳-۲ اطلاعات جدول رای‌دهندگان و امکان بروز حمله‌ی پیوند ..... ۱۲
- جدول ۴-۲ مشخصات بیماران یک بیمارستان ..... ۱۳
- جدول ۵-۲ ۲-گمنامی جدول ۴-۲ ..... ۱۴
- جدول ۶-۲ نمونه اطلاعات منتشر شده پرس‌وجو AOL در سال ۲۰۰۶ ..... ۳۲
- جدول ۴-۱ میزان سطح حریم خصوصی، مقدار گمنام‌سازی و درجه افراد موجود در گراف اجتماعی ..... ۵۷
- جدول ۵-۱ مجموعه داده‌های مورد استفاده در ارزیابی ..... ۷۶
- جدول ۵-۲ اندازه مجموعه کاندیدا در مجموعه دادگان اصلی ..... ۷۸
- جدول ۵-۳ زمان اجرا و تعداد  $w$ -خوشه‌ها در الگوریتم MicroPDA ..... ۷۹
- جدول ۵-۴ نتایج حاصل از اجرای الگوریتم ParticlePDA ..... ۸۰
- جدول ۵-۵ نتایج حاصل از اجرای الگوریتم ModPDA ..... ۸۱
- جدول ۵-۶ مقایسه روش پیشنهادی با سایر روش‌ها ..... ۸۳
- جدول ۵-۷ خطر افشا برای روش SocialPDA و SocialIDA ..... ۸۹

## اصطلاحات و علائم اختصاری

DPSO	Distinct Particle Swarm Optimization
EAGA	Evolutionary algorithm for graph anonymization
SSN	Social Security Number
MicroPDA	Microaggregation for Personalized Degree Anonymity
ModPDA	Graph Modification for Personalized Degree Anonymity
ParticlePDA	PSO for Personalized Degree Anonymity
BPSO	Binary Particle Swarm Optimization
SocialDA	Social Graph Degree Anonymity
SocialPDA	Social Graph Personalized Degree Anonymity

# فصل اول

## کلیات



# کلیات

## ۱،۱ مقدمه

امروزه با وجود وب ۲ تعداد شبکه‌های اجتماعی رشد بسیار زیادی داشته‌اند و حاوی مقدار قابل توجهی از اطلاعات ارزشمند برای اهداف تحقیقاتی و تجاری هستند. با این حال اطلاعات موجود، معمولاً حاوی اطلاعاتی از حریم خصوصی افراد و رابطه حساس بین آن‌ها در شبکه اجتماعی هستند. حذف اطلاعات کاربران موجود در شبکه اجتماعی مثل نام و شماره امنیتی اجتماعی (SSN) نمی‌تواند حریم خصوصی را برای کاربران فراهم کند و مهاجم با حملات موسوم به حملات ساختاری<sup>۱</sup> می‌تواند به اطلاعاتی دست یابد [۱]. بنابراین با انتشار این داده‌ها احتمال افشای حریم خصوصی بسیار بالا می‌رود. روش‌های بسیاری برای گمنام‌سازی شبکه‌های اجتماعی مطرح شده است. این روش‌ها باید همواره دو مفهوم اصلی در حوزه داده‌کاوی با حفظ حریم خصوصی را مد نظر قرار دهند. اول مفهوم اتلاف اطلاعات<sup>۲</sup> که اشاره دارد به میزان دقت داده‌ها به منظور استفاده آنها در کاربردهای تحقیقاتی و داده‌کاوی. دوم مفهوم از دست رفتن حریم خصوصی که مشخص می‌کند تا چه میزان احتمال دارد داده‌های اولیه را بتوان از داده‌های گمنام شده استنتاج نمود و حریم خصوصی موجودیت‌های صاحب داده‌ها را به مخاطره انداخت. در کل هرچه حریم خصوصی افراد حاضر در شبکه اجتماعی بیشتر باشد، نتایج به دست آمده از داده‌کاوی کم‌اثرتر خواهد بود و برعکس. در نتیجه هدف اصلی روش‌های ارائه شده در گمنام‌سازی شبکه اجتماعی فراهم

---

<sup>۱</sup> Structural attack

<sup>۲</sup> Information lost

کردن یک نوع مصالحه<sup>۱</sup> بین حریم خصوصی افراد حاضر در شبکه اجتماعی و دقت داده‌ها برای اهداف تحقیقاتی می‌باشد [۲].

## ۲,۱ بیان مساله و فرضیات

تحقیقات بسیاری در زمینه گمنام‌سازی داده‌های رابطه‌ای<sup>۲</sup> انجام شده‌است که به صورت بهینه داده‌ها را گمنام می‌کند [۳, ۴]. اما با توجه به خصوصیات ساختاری شبکه اجتماعی نمی‌توان این روش‌ها را به طور مستقیم بر روی این شبکه‌ها اعمال کرد. همان‌طور اشاره شد، با انتشار داده‌های شبکه اجتماعی احتمال نقض حریم خصوصی کاربران وجود دارد که برای حل این مشکل باید یک روش نظام‌مند<sup>۳</sup> قبل از انتشار داده‌ها اعمال گردد. با این وجود گمنام‌سازی داده‌های شبکه اجتماعی چالش برانگیزتر از گمنام‌سازی داده‌های رابطه‌ای است. اولین چالش موجود مربوط به دانش پس زمینه مهاجم است. در داده‌های رابطه‌ای مهاجم معمولاً از خصیصه‌های شبه‌شناسه آگاه است و با توجه به آن‌ها اطلاعات مورد نظر خود را در بین تعدادی از چندتایی‌های<sup>۴</sup> جدول بدست می‌آورد. اما در شبکه‌های اجتماعی تعداد بسیاری از مشخصات می‌تواند منجر به متمایز شدن یک کاربر در تمام شبکه شود. به عنوان مثال برچسب یک گره، یال‌های متصل به یک گره، برچسب یال‌ها، زیرگراف همسایگان و ترکیب آن‌ها. دومین چالش موجود، پیچیدگی محاسبات اطلاعات از دست رفته بعد از گمنام‌سازی شبکه اجتماعی نسبت به داده‌های رابطه‌ای است. در داده‌های رابطه‌ای اطلاعات از دست رفته با جمع اطلاعات از دست رفته هر چندتایی به دست می‌آید. این در حالی است که شبکه‌های اجتماعی شامل مجموعه‌ای از گره‌ها و یال‌هاست. مقایسه دو شبکه اجتماعی با توجه به گره‌ها و یال‌ها کار بسیار پیچیده‌ای می‌باشد. به عنوان مثال دو شبکه اجتماعی با تعداد برابری گره و یال می‌تواند دارای خصوصیات شبکه‌ای متفاوتی همچون اتصال<sup>۵</sup>

---

<sup>۱</sup> Trade-off

<sup>۲</sup> Relational data

<sup>۳</sup> Systemic

<sup>۴</sup> Tuple

<sup>۵</sup> Connectivity

بینابینی<sup>۱</sup> و قطر<sup>۲</sup> باشد. بنابراین تعریف از دست‌رفتن اطلاعات در شبکه‌های اجتماعی می‌تواند بسیار متنوع باشد.

روش‌های بسیاری برای گمنام‌سازی داده‌های شبکه اجتماعی ارائه شده‌اند، ولی روش‌های موجود به مقدار نیازمندی هر فرد به محافظت از حریم خصوصی توجهی ندارند و کلیه افراد را به یک میزان گمنام می‌کنند. این امر باعث افزایش اتلاف اطلاعات در داده‌های گمنام شده و کاهش دقت اطلاعات می‌شود. همچنین این روش‌ها به خصوصیات شبکه اجتماعی توجهی ندارند و گراف گمنام شده نسبت به گراف اصلی بسیار متفاوت است.

### ۳,۱ اهداف و نتایج حاصل از پژوهش

در این پایان‌نامه روشی برای شخصی‌سازی حفظ حریم خصوصی افراد موجود در شبکه اجتماعی مبتنی بر عمومی‌سازی ارائه شده است. در این روش با توجه به سطوح متفاوت حریم خصوصی افراد، مقداری از گمنام‌سازی فراهم می‌شود. در ابتدا گره‌ها بوسیله روشی مبتنی بر ریزتجمیع<sup>۳</sup> به صورت بهینه گروه‌بندی می‌شوند. سپس گره‌های گروه‌بندی شده با روشی مبتنی بر الگوریتم بهینه‌سازی ازدحام ذرات گسسته دودویی<sup>۴</sup> گمنام شده و در انتها با حفظ خصوصیات ساختاری شبکه اجتماعی و با سه عملگر تعریف شده گراف دستکاری و گمنام می‌شود. هدف روش ارائه شده، حفظ حریم خصوصی کاربران در شبکه اجتماعی با کمترین میزان اتلاف اطلاعات می‌باشد. روش ارائه شده با برخی از روش‌های موجود برای حفظ حریم خصوصی در شبکه‌های اجتماعی مقایسه شده است.

### ۴,۱ جنبه‌های نوآوری

نوآوری‌های روش پیشنهادی را می‌توان به صورت زیر بیان کرد:

---

<sup>۱</sup> Betweenness

<sup>۲</sup> Diameter

<sup>۳</sup> Microaggregation

<sup>۴</sup> Discrete binary particle swarm optimization

- ارائه روشی نوین برای شخصی‌سازی حریم خصوصی افراد موجود در شبکه اجتماعی با توجه به درخواست کاربران و کمترین میزان اتلاف اطلاعات.
- در نظر گرفتن خصوصیات شبکه اجتماعی اصلی در روش گمنام سازی درجه.

## ۵،۱ مروری بر فصل‌ها

در این فصل به کلیات پژوهش شامل انگیزش، بیان مساله و فرضیات، اهداف و نتایج حاصل از پژوهش و جنبه‌های نوآوری آن پرداخته شد.

در ادامه در فصل دوم نگاهی مختصر بر شبکه‌های اجتماعی و ویژگی‌های ساختاری آنها، تعاریف و مفاهیم اولیه حریم خصوصی و انتشار داده‌ها با حفظ حریم خصوصی خواهیم داشت. در انتهای این فصل روش ریزتجمیع برای گمنام‌سازی داده‌ها شرح داده شده است.

فصل سوم به گمنام‌سازی شبکه‌های اجتماعی و مهم‌ترین کارهای انجام شده در این حوزه پرداخته شده است. این روش‌ها به چهار دسته اصلی تقسیم شده‌اند و در انتها با هم مقایسه شده‌اند.

در فصل چهارم روشی جدید مبتنی بر ریزتجمیع و الگوریتم بهینه سازی ازدحام ذرات گسسته دودویی برای گمنام‌سازی داده‌های شبکه اجتماعی با توجه به نیازمندی‌های حریم خصوصی کاربران موجود در شبکه و کمترین میزان اتلاف اطلاعات ارائه شده است. این روش شامل سه مرحله می باشد که ابتدا گره‌ها به صورت بهینه گروه‌بندی می‌شوند. سپس اعضای هر گروه گمنام شده و در مرحله آخر با توجه به ساختار شبکه اجتماعی و سه عملگر تعریف شده برای دستکاری گراف، گراف شبکه اجتماعی گمنام می‌شود. فصل پنجم به ارزیابی و بررسی کارایی روش پیشنهادی و مقایسه آن با سایر روش‌های مشابه می‌پردازد. در فصل ششم نیز نتیجه گیری و ارائه پیشنهادها و تحقیقاتی آورده شده است.