

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра социальной информатики

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ГРАФА
ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНОЙ СЕТИ**

(автореферат бакалаврской работы)

студентки 4 курса 451 группы
направления 09.03.03 - Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Яковлевой Анжелины Дмитриевны

Научный руководитель
кандидат социологических наук, доцент

_____ К.В. Мохнаткина
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г. Малинский
подпись, дата

Саратов 2022

ВВЕДЕНИЕ

Актуальность исследования. Социальная сеть - это социальная структура, состоящая из набора социальных акторов (физических лиц или организаций), наборов диадических связей и других социальных взаимодействий между участниками.

Яркими примерами социальных сетей являются Facebook, Twitter, ВКонтакте и др.

Рост информации в социальных сетях обусловил необходимость в создании более совершенных методах их исследования. Современная социальная сеть - это междисциплинарная область, которая объединяет математиков, ИТ-специалистов, экономистов и социологов.

SNA (Socialnetworkanalysis, анализ социальных сетей) - направление современной компьютерной социологии, которое занимается описанием и анализом возникающих в ходе социального взаимодействия и коммуникации связей (сетей) различной плотности и интенсивности.

Исследование социальных сетей основано на методах анализа, использующих локальные и глобальные паттерны данных структур для определения местоположения ключевых акторов и изучения динамики сети.

Одними из таких методов являются методы визуализация графов.

Исследование социальных сетей методами визуализация графов является актуальным и представляет научно-практический интерес.

Объект исследования: социальная сеть.

Предмет исследования: методы визуализации графов.

Цель данной работы: исследование социальных сетей методами визуализации графов.

Для достижения данной цели необходимо выполнить следующие **задачи:**

— проанализировать методы анализа социальных сетей на основе теории графов;

— проанализировать модели и алгоритмы визуализации социальной

сети;

— выполнить экспериментальный анализ социальной сети с помощью методов визуализации графов.

Методы исследования – методы теории графов, методы анализа социальных сетей.

Практическая значимость дипломной работы заключается в исследовании возможностей существующего программного обеспечения для анализа социальных сетей методами визуализации графов.

Степень научной разработанности данной проблемы. Широкий пласт научных статей и исследований посвящаются активно развивающейся тенденции использования социальных сетей в разных целях. Социальные сети исследуются в социологии как способ взаимодействия людей друг с другом, например, в трудах А. В. Щекотурова «Гендерные самопрезентации подростков на страницах социальной сети «ВКонтакте»¹. Родоначальником теории графов считается выдающийся математик, член Петербургской академии наук Леонард Эйлер.

В 1736 году в одном из своих писем он формулирует и предлагает решение задачи о семи кёнигсбергских мостах, ставшей впоследствии одной из классических задач теории графов.

Структура работы. Бакалаврская работа состоит из введения, четырех глав, заключения и списка использованных источников.

¹Щекотуров, А. В. Гендерные самопрезентации подростков на страницах социальной сети «ВКОНТАКТЕ» // Мониторинг общественного мнения. 2012. № 4 С. 80–88.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе «Теоретико-методологические подходы к анализу социальных сетей.» рассматривается история термина «социальная сеть», самые популярные крупнейшие сети в России, причины их популярности и использования, а также положительные и отрицательные их стороны, методология анализа социальных сетей, а именно анализ центральности, более подробно рассмотрены все меры центральности.

Понятие «социальная сеть» приобрело большую популярность среди западных исследователей общества во второй половине 20 века. Различные исследователи предполагают, что самым первым ресурсом социальной сети является сеть Classmates.com, созданная Рэнди Конрадсом в 1995 году. Classmates переводится как одноклассники. Сайт предоставил пользователям возможность восстановить связь с бывшими одноклассниками, однокурсниками, коллегами, друзьями. Сеть работает по сей день и насчитывает более 50 миллионов пользователей в США и Канаде. Кроме того, его услуги доступны жителям Швеции, Германии, Австрии и Франции. Но довольно долгое время этот социальный портал не поддерживал функции создания личных профилей и добавления друзей. То есть пользователь мог только подключиться к своему учебному заведению и предоставить ему список студентов, обучавшихся в этом учреждении.

Например, ВКонтакте считается одной из самых популярных социальных сетей в России, ВКонтакте — современный молодежный интернет-мессенджер для быстрого и удобного общения, а также поиска людей по всему миру. Самый популярный в России и странах ближнего зарубежья. Основана петербуржцем Павлом Дуровым в 2006 году. Изначально задумывалась как сайт поиска однокурсников и одноклассников, но с момента первого запуска претерпела множество изменений и была упрощена. Одна из возможностей ресурса — смотреть видео и слушать музыку онлайн. Регистрация доступна каждому, достаточно ввести номер мобильного телефона и на телефон придет логин и пароль для входа.

Сейчас социальные сети дают людям доступ к различным навыкам. Личный успех человека теперь все больше зависит от его умения использовать то, что знают и умеют другие. Поэтому разнообразные связи помогут вам объективнее и шире смотреть на вещи, полнее раскрыть свой творческий потенциал. И когда вы делитесь навыками или информацией с людьми с разным опытом, вы делитесь уникальными и очень ценными ресурсами друг с другом.

Причины использования социальных сетей зрелыми людьми могут быть разными: личная выгода, управление подчиненными (иногда гораздо проще написать сотрудникам в Одноклассники или ВКонтакте, чем звонить им всем), обмен данными, простое общение. Для молодежи социальные сети являются способом удовлетворения своих потребностей: объединение в социальные группы, общение, дружба. Но есть и те, кто пытается реализовать себя, завоевать уважение с помощью социальных сетей.

Сейчас благодаря интернет-ресурсам количество участников социальных сетей превышает миллионы, ведь соц. сети они объединяют людей любых социальных групп, обеспечивают им общение и доступ к различным развлечениям, таким как музыка, игры, видео и т.д., и это то, что привлекает людей - они удобны и удовлетворяют потребности человека, и именно поэтому они составляют основу современного общества.

Граф представляет собой важную сложную сетевую модель, описывающую отношения между различными объектами в реальных приложениях, включая граф знаний, социальную сеть и сеть трафика.

Соответственно, анализ социальных сетей (SNA) рассматривает социальные отношения с точки зрения теории сетей, состоящей из узлов и связей (также называемых ребрами или связями)

Узлы - это отдельные субъекты внутри сетей, а связи - это отношения между акторами. Результирующие структуры на основе графа часто очень сложны. Между узлами может быть много видов связей.

В рамках теории графов и сетевого анализа существуют различные меры центральности вершины в графе, которые определяют относительную важность вершины в графе.

Анализ центральности по степени (Degreecentrality) является самым простым. Центральность по степени определяется как число связей, приходящихся на узел (т.е. количество связей, которые имеет узел).

Центральность по близости (Closenesscentrality) является показателем, насколько быстро распространяется информация в сети от одного участника к остальным. В топологии и смежных областях математики близость является одним из основных понятий в топологическом пространстве. В теории графов близость является мерой центральности вершины в графе. Вершины, которые «неглубокие» по отношению к другим вершинам (то есть те, которые имеют тенденцию иметь короткие геодезические расстояния до других вершин в графе), имеют более высокую близость.

Посредническая центральность (Betweennesscentrality) является мерой центральности вершины в графе. Вершины, которые встречаются на многих кратчайших путях между другими вершинами, имеют более высокую промежуточность чем те, которые этого не делают.

Центральность по собственному вектору (Eigenvectorcentrality) является мерой важности узла в сети. Она назначает относительные оценки всем узлам в сети на основе принципа, согласно которому подключения к узлам с высокой оценкой вносят больший вклад в оценку рассматриваемого узла, чем равные подключения к узлам с низкой оценкой. GooglePageRank является вариантом меры центральности по собственному вектору.

В теории графов центральность Каца является обобщенной мерой центральности в сети. В отличие от типичных мер центральности, которые рассматривают только кратчайший путь (геодезический) между парой действующих лиц, центральность Каца измеряет влияние, принимая во внимание общее количество прогулок между парой действующих лиц.

Во второй главе «Модели и алгоритмы визуализации социальной сети в виде графов. Программное обеспечение для анализа методами визуализации графов.» речь идет о модели визуализации графов и принципах её укладки графов, алгоритмах, представлена постановка задачи визуализации соц. сети в виде графа, а также программы для визуализации графов, непосредственно, программа Gephi для анализа социальных сетей, предоставлены примеры кода визуализации.

Визуализация графа представляет собой процедуру указания координат его вершин и изображения ребер графа в виде отрезков между вершинами.

В настоящее время существует множество алгоритмов, основанных на силовой модели укладки графа: прямая линия, метод прямоугольника, метод сетки, круговой метод, метод дисконтирования основаны на математических методах, в соответствии с математической характеристикой графики они получают графические координаты вершин по уравнению. Силовые модели укладки графов используются для преобразования математической информации графа в двумерную или трехмерную геометрическую информацию, основанную на относительном положении узлов и ребер, автоматического вытягивания графа и, таким образом, реализации его визуализации.

Пружинная модель. Идея пружинной модели построена на предположении, что каждая вершина графа представляет собой шарик. Так как оптимальные координаты шарика не известны, шарик случайно размещен в двумерных плоскостях. Между шарами имеется пружина. Идеальная длина пружин пропорциональна расстоянию между шариками. Если расстояние между шариками не равно идеальной длине пружин, пружина в системе будет растянута или сжата, и система будет накапливать энергию.

Пружинная электрическая модель. Идея электрической модели пружины такова: предположим, что вершины это шары, которые имеют одинаковый заряд и отталкиваются друг от друга, причем сила отталкивания обратно пропорциональна геометрическому расстоянию между шарами.

Программы для визуализации графов можно разделить на две категории:

1) редакторы — приложения, которые позволяют автоматизировать ручное построение графа и обеспечивают возможность автоматического изменения граней при перемещении вершин графа;

2) программы, которые позволяют отказаться от ручного рисования графа. При этом пользователь задает некоторую базу данных, по которой программа рассчитывает положение вершин и граней и строит граф. Для построения социограмм могут применяться программы обоих типов, поскольку для конкретных случаев та или иная программа может оказаться наиболее подходящей.

Gephi — ведущее программное обеспечение для визуализации и исследования всех видов графиков и сетей. Gephi распространяется с открытым исходным кодом и бесплатно. Программа работает на ОС Windows, Mac OS X и Linux.

Анализ социальных сетей обеспечивается простым созданием коннекторов социальных данных для сопоставления общественных организаций и небольших сетей.

Поддерживает вычисление всех показателей центральности.

Используемые технологии:

- эргономичный интерфейс: не требуются навыки программирования;
- высокая производительность: встроенный движок рендеринга;
- собственные форматы файлов: GDF (GUESS), GraphML (NodeXL), GML, NET (Pajek), GEXF и другие;
- настраивается с помощью плагинов: макеты, метрики, источники данных, инструменты манипуляции, пресеты рендеринга и многое другое.

Выполним экспериментальный анализ групп социальных сетей с помощью программы Gephi.

ЗАКЛЮЧЕНИЕ

Одними из основных методов анализа социальных сетей являются методы визуализация графов.

Исследование социальных сетей методами визуализация графов является актуальным и представляют научно-практический интерес.

Целью бакалаврской работы является исследование социальных сетей методами визуализации графов.

Для достижения поставленной цели были решены следующие задачи:

1. Проанализированы методы анализа социальных сетей на основе теории графов. Для анализа социальных сетей используются различные показатели центральности. В рамках теории графов и сетевого анализа существуют различные меры центральности вершины в графе, которые определяют относительную важность вершины в графе: центральность по степени, центральность по близости, посредническая центральность, центральность по собственному вектору является мерой важности узла в сети. Обобщенной мерой центральности узла в сети в теории графов является центральность Каца. Анализ центральности по степени является самым простым. Недостатком других показателей центральности является их вычислительная сложность.

2. Описаны и проанализированы модели и алгоритмы визуализации социальной сети в виде графа. Как показал анализ, силовые модели укладки графов используются для преобразования математической информации графа в двумерную или трехмерную геометрическую информацию, основанную на относительном положении узлов и ребер, автоматического вытягивания графа и реализации его визуализации. Алгоритмы на основе пружинной электрической модели более эффективны, чем алгоритмы, основанные на пружинной модели. Общим недостатком указанных алгоритмов является недостаточно разработанные критерии их сходимости.

3. Как показал анализ, программы для визуализации графов можно условно разделить на две категории: редакторы — приложения, которые

позволяют автоматизировать ручное построение графа программы, которые позволяют отказаться от ручного рисования графа. Ко второй категории программ относится программа Gephi.

4. С помощью программы Gephi выполнен экспериментальный анализ социальной сети методами визуализации графов. Анализ социальных сетей в Gephi обеспечивается простым созданием коннекторов социальных данных для сопоставления общественных организаций и небольших сетей. Как показало экспериментальное тестирование, программа Gephi позволяет успешно решать задачи анализа социальных сетей с помощью методов визуализации графов.

Результаты дипломной работы представляют научно-практический интерес и могут быть рекомендованы для практического решения задач анализа социальных сетей методами визуализации графов.