

АЛГОРИТМ ПОИСКА СОЦИАЛЬНЫХ СВЯЗЕЙ НА ОСНОВЕ СОЦИАЛЬНОГО ГРАФА И ГРАФА ИНТЕРЕСОВ

Архипов Глеб Максимович

студент, Самарский университет им. С.П. Королёва, РФ, г. Самара

Паращенко Андрей Владимирович

студент, Самарский университет им. С.П. Королёва, РФ, г. Самара

Додонова Наталья Леонидовна

научный руководитель, канд. физ.- мат. наук, доцент, Самарский университет им. С.П. Королёва, РФ, г. Самара

Аннотация: В настоящее время компьютеры и социальные сети стали неотъемлемой частью практически всех людей. В интернете мы можем находить не только полезную информацию, но также и новых и старых знакомых. Мы разработали актуальный алгоритм поиска одноклассников в сети.

Ключевые слова: интернет, социальный граф, граф интересов, связи, сети.

Цели и Задачи:

- Познакомиться с понятиями «социальный граф» и «граф интересов»;
- Найти связи между интересами человека;
- Реализовать алгоритм поиска одноклассников на основе изученной информации.
- Найти практическое применение для такого вида разработки.

Социальный граф

Социальный граф — это граф, в котором узлы представлены социальными объектами, такими как аккаунты пользователей в социальных сетях, медиа-контентом и т. д., а ребрами называются — связи между ними.

При помощи социальных графов решаются многие задачи, например:

- Социальный поиск;
- Ориентация на контент пользователей по их активности;
- Генерация рекомендаций пользователя по потребляемому контенту;
- Сбор информации в открытом доступе для моделирования графа;
- Идентификация пользователей;

Из-за разнообразия социальных сетей и закрытости данных о пользователях возникают некоторые проблемы с обработкой данных социальных графов.

Идентификация пользователей: Обнаружение профилей в социальных сетях,

принадлежащих одному человеку. С помощью идентификации пользователей получается составить более полный социальный граф, который поможет в решении следующих задач:

1. **Социальный поиск** - поисковая система, осуществляющая поиск по людям, фотокарточкам с помощью социальных связей и предпочтений пользователей.
 2. **Генерация рекомендаций:** Важной задачей является поиск точных алгоритмов генерации рекомендаций и предложений пользователям, который так же используется при создании графа интересов на основе социального графа.
- Рекомендация друзей — пользователи сами не зная того, делят свои контакты на кластеры, через взаимодействия в рамках социальной сети.
 - Рекомендации контента — рекомендации медиа-контента, сообществ, новостей ...

3) Выявление «настоящих» связей

С помощью различных алгоритмов и аналитики профилей пользователей, выявление истинных связей между людьми, т.е. друзей, коллег по работе, родственников и т.п.

Граф интересов

Граф интересов — это представление интересов конкретного пользователя, полученное из анализа его активности в соцсетях. Вершинами такого графа являются интересы пользователя, ребра графа показывают взаимоотношения между вершинами. При помощи графа интересов можно узнать, чем человек увлекается, как проводит досуг, что хочет купить, за каким видом спорта он следит, или куда он хотел бы съездить отдохнуть.

Связи в графе интересов

В графе интересов различают некоторые типы связей, которые позволяют пользователю выходить за рамки обычных социальных сетей. Например, человек хочет найти ответ на интересующий его вопрос, но среди друзей и знакомых никто не осведомлен в этой области. В этом случае выстраивается цепочка из трех типов связей:

- человек-человек (пользователи в социальной сети взаимодействуют напрямую)
- человек-интерес (то с чем пользователь взаимодействует с социальной сети)
- интерес-интерес (схожие интересы могут быть взаимосвязаны)

Граф интересов также может быть взвешенным, в этом случае вес ребра означает силу взаимосвязи между вершинами.

При построении графа интересов изначально вводится точка зрения о том, что взаимосвязи между людьми имеют одинаковую силу. Например, интерес к мотоциклам и к политике неизвестен, и взаимосвязь между этими двумя интересами представляется в виде бесконечного числа. Затем, если в процессе анализа обнаружится что люди, увлекающиеся мотоциклами, ведут себя похожим образом, что и люди, увлекающиеся политикой, значение веса ребра взаимосвязи между вершинами, примет конкретное значение.

Отношения между графом интересов и социальным графом

Граф интересов и социальный граф похожи по своему принципу действия, но между ними существуют важные отличия.

Главным отличием является то, что графы интересов служат для создания сети интересов людей. В то время как различные социальные сети строятся вокруг друзей человека, то есть вокруг социального графа, сети интересов созданы вокруг увлечений пользователей, их графа интересов.

Таким образом, увлечения человека, представленные в виде графа интересов, обеспечивают средства для дальнейшей персонализации веб-пространства, основанной на пересечении графа интересов с веб-контентом. Граф интересов или сеть интересов в некоторых случаях

могут быть получены из социального графа или социальной сети и могут поддерживать и обновлять связи между вершинами на основе данной социальной сети.

Использование графа интересов

Граф интересов, может быть, полезен как для обычных пользователей, так и для предпринимателей.

Для пользователей граф интересов может быть применён для установления связей между пользователями социальной сети, поиска людей по общим увлечениям и т. д.

Граф интересов широко применяется в маркетинге. С помощью него происходит анализ интересов аудитории, которой в дальнейшем предлагается интересующая их продукция. Также такие графы используются для таргетированной рекламы, которая сейчас сильно распространена и эффективна.

Алгоритм поиска по социальным сетям

Ознакомившись с темой «граф интересов» и «социальный граф», мы поставили себе задачу реализовать алгоритм по поиску возможных одноклассников в соц.сети на языке с++. Программа реализована по определенным шагам.

1) Создаем или парсим(переносим) базу данных студентов. Если создается вручную, то заполняются поля для каждого студента, такие как: Имя, Фамилия, Отчество, Дата Рождения, Дата поступления на специальность, Университет и факультет обучения, статус страницы (открыта или закрыта для общего обозрения) и основная сфера интересов. Из полученной информации составляется граф, где вершинами являются студенты, а ребрами связь между ними как одноклассники).

Далее проверяется у каждого студента тип его страницы в соц.сетях. Если она закрыта, то удаляются все ребра, связанные с вершиной этого студента.

```
[2]    Список студентов:

(1)    Архипов Глеб Максимович
        Дата рождения: 04.09.2000
        Страницка в соц. сетях открыта
        Учится в СНИУ С.П.Королева до 2023 на специальности 100503
        Основная сфера интересов - Программирование.

(2)    Богданчиков Артем Сергеевич
        Дата рождения: 19.09.2000
        Страницка в соц. сетях открыта
        Учится в СамГУ до 2025 на специальности 310867
        Основная сфера интересов - Хирургия.

(3)    Соколов Александр Валерьевич
        Дата рождения: 28.07.1996
        Страницка в соц. сетях открыта
        Учится в СамГУ до 2018 на специальности 090304
        Основная сфера интересов - Программирование.
```

Рисунок 1. Список студентов

```

(4) Горин Иван Дмитриевич
Дата рождения: 11.10.2001
Страничка в соц. сетях открыта
Учится в СНИУ С.П.Королёва до 2023 на специальности 100501
Основная сфера интересов - Программирование.

(5) Паращенко Андрей Владимирович
Дата рождения: 23.02.2001
Страничка в соц. сетях открыта
Учится в СНИУ С.П.Королёва до 2023 на специальности 100503
Основная сфера интересов - Программирование.

(6) Кузнецов Вадим Олегович
Дата рождения: 06.10.2000
Страничка в соц. сетях ЗАКРЫТА
Учится в - до - на специальности -
Основная сфера интересов -.

```

Рисунок 2. Список студентов

2) Создание матрицы смежности путем проверки совпадений такого поля данных, как **основная сфера интересов**.

(После данного шага мы получаем сформированную, но не до конца отфильтрованную матрицу смежности, так как студенты могут интересоваться одним и тем же, но обучаться в разных университетах и на разных направлениях. В данном случае изменяется значение 1 на 0. Тем самым в графе между вершинами, которыми являются студенты убирается ребро)

```

Матрица смежности:

   [1] [2] [3] [4] [5] [6]
[1] 0  1  1  1  1  0
[2] 1  0  1  1  1  0
[3] 1  1  0  1  1  0
[4] 1  1  1  0  1  0
[5] 1  1  1  1  0  0
[6] 0  0  0  0  0  0

1 - одnogруппник; 0 - не одnogруппник

Для получения более корректного результата нужно отфильтровать по определенным критериям:
1. По основной сфере интересов
2. По университету
3. По специальности

Выберите метод фильтрации -> 1

```

Рисунок 3. Результат после фильтрации

- Проверка получившегося результата по критерию «**Университет**»

(После данного шага получаем сформированную, но не до конца отфильтрованную матрицу смежности, так как студенты могут или могли обучаться в одном университете, но на разных направлениях, что не делает их одnogруппниками. Так же, как и при предыдущей фильтрации 1 меняются на 0 в случае несовпадения полей данных при сравнении).

```
Матрица смежности после фильтрации по основной сфере интересов:

  [1] [2] [3] [4] [5] [6]
[1] 0  0  1  1  1  0
[2] 0  0  0  0  0  0
[3] 1  0  0  1  1  0
[4] 1  0  1  0  1  0
[5] 1  0  1  1  0  0
[6] 0  0  0  0  0  0

1 - одnogруппник; 0 - не одnogруппник

Для получения более корректного результата нужно отфильтровать по определенным критериям:
2. По университету
3. По специальности

Выберите метод фильтрации -> 2
```

Рисунок 4. Результат после фильтрации

- Проверка получившегося результата по критерию «**Специальность**»

После данного шага получаем сформированную и отфильтрованную матрицу смежности, что позволяет нам узнать кто является одnogруппником для изначального студента.

```
Матрица смежности после фильтрации по специальности:

  [1] [2] [3] [4] [5] [6]
[1] 0  0  0  0  1  0
[2] 0  0  0  0  0  0
[3] 0  0  0  0  0  0
[4] 0  0  0  0  0  0
[5] 1  0  0  0  0  0
[6] 0  0  0  0  0  0

1 - одnogруппник; 0 - не одnogруппник

Для получения более корректного результата нужно отфильтровать по определенным критериям:
Произведена фильтрация по всем доступным критериям.
```

Рисунок 5. Результат после фильтрации

Алгоритм сравнения полей данных реализован так, что каждый студент сравнивается с каждым по всем критериям, поэтому из общего списка студентов могут быть выявлены одnogруппники из разных вузов.

3) Вывод полученного результата.

- Преобразование матрицы смежности в список смежности для более удобного вывода в дальнейшем полученной информации.

```
[4]                Одногоруппники:

После всех фильтратий одногоруппниками являются (1) Архипов и (5) Паращенко.

Нажмите ESC для возвращения в меню
```

Рисунок 6. Результат алгоритма

Вывод:

Я считаю, что с практической точки зрения данный алгоритм можно использовать на сайтах личных кабинетов для составления списков студентов, обучающихся в одной группе. Так же если добавить другие поля данные и иные поля фильтрации, можно получить группы людей, сформированные по новым данным.

Список литературы:

1. Анализ социального графа [Электронный ресурс]: <https://habr.com/ru/post/81225/>
2. Додонова Л.Н. Конспект лекций по дисциплине «Теория конечных графов и ее применения». Самара, 2019. — стр. 52.
3. Теория графов [Электронный ресурс]: Карпов Д.В. – Режим доступа: https://logic.pdmi.ras.ru/~dvk/graphs_dk.pdf (дата обращения: 03.12.2019).