

ОПРЕДЕЛЕНИЕ НАИБОЛЕЕ ВЛИЯТЕЛЬНЫХ ОБЪЕКТОВ ВИРТУАЛЬНОЙ СОЦИАЛЬНОЙ СЕТИ

Борискин Александр Владимирович

Аспирант

Факультет ПММ ВГУ, Воронеж, Россия

E-mail: boriskinpost@gmail.com

Маркетологи давно заметили, что профиль пользователя виртуальной социальной сети является бесценным хранилищем информации. Существует отдельное направление — Social Mining, которое занимается исследованием методов обработки «сырых» данных виртуальных социальных сетей для получения практически полезной информации. Таковой является в том числе информация о наиболее влиятельных объектах [1].

Рассмотрим одну из задач, где будут полезными знания о влиятельных объектах. Пусть необходимо распространить информацию среди определенной группы пользователей. Например, информацию о выходе нового объектива, позволяющего запечатлеть самые сложные моменты спортивных событий. Эти сведения должны дойти до некоторой группы фотографов, которая интересуется съемкой именно спортивных событий. В большинстве случаев в подобных группах существуют люди, которые пользуются авторитетом, ко мнению которых прислушиваются. Они и являются влиятельными объектами.

Начинают решение данной задачи с моделированием социальной сети в виде графа, узлами которого являются люди. Отношения (в широком смысле) между людьми обозначаются в виде дуги. Граф может быть взвешенным: например, если задана функция $\mu(x, y) \in [0, 1]$, указывающая, в какой степени объект x **напрямую** влияет на объект y . В одной из первых работ [3], посвященной поиску влиятельных объектов, автор ставит во главу угла понятие центральности вершины графа, которое численно выражает понятие влиятельности. Существенный недостаток подобного подхода состоит в том, что он требует задания функции $\mu(x, y)$ еще до применения метода. Целью данной работы является постановка задачи в виде, не требующем задания $\mu(x, y)$ и выработка ее решения.

Дело в том, что современные социальные сети хранят информацию о множестве аспектов взаимодействия людей. Например, отметил ли объект x объекта y как друга, какое количество фотографий объекта y объект x отметил как «мне нравится» и т.д. Обозначим количество аспектов через m . Тогда функцию непосредственного вли-

яния можно определить следующим образом:

$$\mu(x, y) = \sum_{i=1}^m c_i \cdot F_i(x, y), \quad (1)$$

где $F_i(x, y) \in [0, 1]$ — функции, характеризующие i -ый аспект влияния объекта x на y , а c_i — вещественные коэффициенты, определяющие вклад этого влияния в общую сумму. На них накладываются следующие ограничения:

$$0 \leq c_i \leq 1, \quad 0 \leq \sum_{i=1}^m c_i \leq 1. \quad (2)$$

Возвращаясь к предыдущему примеру с объективом, на момент постановки задачи могут быть уже известны люди, пользующиеся значительным влиянием в своей среде. Таким образом, задачу можно сформулировать в следующем виде: необходимо подобрать такие коэффициенты c_i , чтобы максимизировать суммарную влияние известных нам объектов. В основе алгоритма расчета влияния стоит понятие итерированной силы [4], опирающееся, в свою очередь, на функцию непосредственного влияния μ . Приходим к задаче условной оптимизации, которая решается с помощью метода внутренней штрафной функции [2].

Итогом работы алгоритма является задание функции $\mu(x, y)$ и получение списка объектов, отсортированного по убыванию влияния. Эксперимент проводился на группе пользователей социальной сети «VKontakte». Анализ результатов показал, что 5 человек, выбранные как известные влиятельные объекты, оказались в первой двадцатке полученного списка.

Литература

1. Губанов Д. А. Социальные сети. Модели информационного влияния, управления и противоборства. М.: Физматлит, 2010.
2. Васильев Ф. П. Методы оптимизации. М.:Факториал Пресс, 2002.
3. Freeman L. C. Centrality in social networks. Conceptual clarification. // Social Networks. 1978. №1. P. 215–239.
4. Интернет–портал «Web и Social Mining»: http://www.basegroup.ru/library/web_mining/