

Бондаренко В.В., к.э.н. Юхименко Б.И
МОДИФИКАЦИЯ АЛГОРИТМА ПОИСКА ЗНАКОМЫХ В
ЛОКАЛЬНЫХ СОЦИАЛЬНЫХ СЕТЯХ НА ОСНОВЕ
ГРУППИРОВКИ ДАННЫХ И ВНЕДРЕНИЯ КЛЮЧЕВЫХ
ПАРАМЕТРОВ

Bondarenko V.V., Ph.D. Yuhimenko B.I.
MODIFICATION OF THE ALGORITHM OF SEARCHING FOR
KNOWN IN LOCAL SOCIAL NETWORKS BASED ON GROUPING OF
DATA AND INTRODUCTION OF KEY PARAMETERS

АННОТАЦИЯ. В работе предложена модификация алгоритма поиска знакомых среди членов частной социальной сети или группы пользователей одного социального круга. Использована дополнительная группировка данных и введены дополнительные ключевые параметры.

Вступление. Не смотря на то, что алгоритм поиска знакомых на основе выделения значимых признаков в локальных социальных сетях[3] имеет хорошие результаты, он так же содержит и ряд проблем, которые необходимо было решить. В частности одной из таких проблем являлось достаточно длительное время обработки и выдачи результата. В связи, с чем возникла необходимость его доработки и дополнительного анализа данных пользователей.

Цель данной работы: Улучшение алгоритма поиска знакомых с помощью использования групп значимых признаков и введения ключевых параметров.

Основная часть работы. В ходе исследования было установлено, что благодаря сравнению ряда параметров из данных пользователей, мы можем с высокой точностью определить, что эти пользователи не знакомы. Такие параметры, демонстрирующие явное логическое противоречие исключающее факт знакомства пользователей друг с другом, были названы ключевыми и были использованы для увеличения скорости алгоритма. Стоит отметить, что набор ключевых параметров не является статичным и изменяется в зависимости от ряда условий, уникальных для каждой социальной группы и ряда обстоятельств.

Использование значимых признаков позволяет выяснить факт знакомства [3], однако существует достаточно весомая проблема,

**Матеріали VIII Міжнародної науково-практичної конференції
«Інформаційні управляючі системи та технології»
23 - 25 вересня 2019, Одеса**

связанная с тем, что для разных социальных групп кол-во и набор ключевых параметров, а так же значимых признаков отличались.

Это создает необходимость переопределять приоритет признаков и ключевых параметров для каждой новой пары пользователей. Для решения этой проблемы было решено собрать признаки в группы и присвоить им коэффициент значимости [4] от 0 до 1 для дальнейшего использования их в качестве дополнительных модификаторов анализа в нейронной сети прямого распространения [1].

На основании сходства свойств данных были получены следующие группы:

1. Подтверждающие данные – позволяют с высокой точностью определять знакомства людей. Компоненты этой группы могут или однозначно подтвердить, или же опровергнуть факт знакомства двух человек. Из этой группы чаще всего берутся ключевые параметры. К признакам этой группы относится: информация о месте проживания, времени, школе и классе котором учился человек и т.д.

2. Данные по интересам – повышают вероятность наличия знакомства людей друг с другом. Особенность группы заключается в том, что знакомые люди почти всегда имеют какие-либо точки соприкосновения, общие интересы и антипатии.

3. Персонализированные данные - определяют шанс знакомства людей и являются узкоспециализированными. Указывают на какую-то конкретную черту или особенность. Важно отметить, что для положительного результата эти черты не обязаны совпадать.

4. Общие данные – группа данных, которые практически не влияют на вероятность знакомства. В частном случае к ним могут относиться: рост, вес и прочие общие параметры человека. Несмотря на то, что такие данные скорее помеха, их полное отсутствие в представлении объекта приводит к значительному ухудшению результата.

5. Непостоянные или неполные данные – личные данные, которые могут быть внесенными не до конца или вовсе не предоставляемыми пользователем. Такие данные весьма интересны и часто весьма противоречивы. Они имеют особенность разрастаться до больших размеров и терять актуальность, что может привести к замедлению процедуры поиска знакомых. Эта группа требует наиболее тщательного анализа, проводимого в совокупности с данными других типов.

6. Ситуативные данные – группа данных, которые описывают ситуацию, в которой будет проводиться анализ. Может включать в себя сведения о погоде, а так же другие признаки, которые так или иначе могут повлиять на поведение субъектов, но не являются данными,

**Матеріали VIII Міжнародної науково-практичної конференції
«Інформаційні управляючі системи та технології»
23 - 25 вересня 2019, Одеса**

описывающими самого субъекта. Как и «общие данные» эта группа практически не влияет на результат напрямую, но в совокупности с другими способна значительно улучшить результат анализа. При некорректном сборе данных для этой группы могут возникать дополнительные ошибки второго рода

7. Специализированные данные – данные, характерные для каких-то социальных кругов или мероприятий, не относящиеся к вышеизложенному списку.

В табл.1 приведены коэффициенты значимости для каждой из групп. Важно отметить, что текущие значения коэффициентов подобраны эмпирическим путем, на основе небольших выборок данных. И в реальности могут колебаться в зависимости от количества данных пользователя, а так же дополнительных модификаций алгоритма.

Таблица 1 - Коэффициенты значимости групп для нейронной сети

Подтверждающие данные	Данные по интересам	Персональные данные	Общие данные	Неполные данные	Ситуативные данные	Специализированные данные
0,7360	0,4426	0,8012	0,318	0,7942	0,7917	0,4411-0,8965

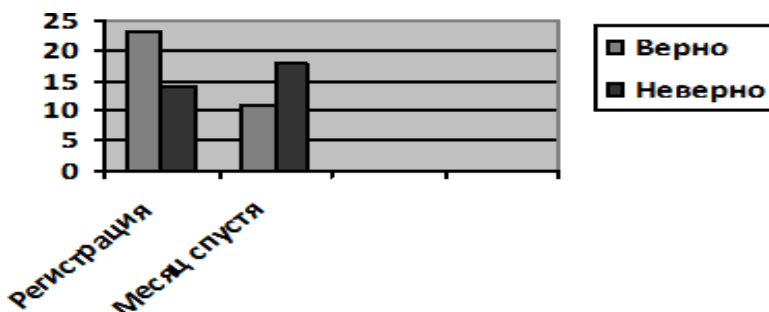


Рис. 1. Среднее кол-во уведомлений о возможных знакомых

Вывод: В работе приведена модификация алгоритма поиска знакомых среди членов частной социальной сети [3]. Группировка ключевых параметров уменьшает количество пересматриваемой информации при поиске, быстрее констатирует факт оценки знакомых, увеличивая

**Матеріали VIII Міжнародної науково-практичної конференції
«Інформаційні управляючі системи та технології»
23 - 25 вересня 2019, Одеса**

скорость работы. На рис.1 приведены результаты модифицированного алгоритма. Следует так же отметить, что группировка данных в позициях 1-5 значительно увеличивает распознавание. Введение же ключевых параметров – повышает скорость работы алгоритма

Литература

1. I.Sutskever, J.Martens, G.Dahl, G.Hinton. On the importance of initialization and momentum in deep learning. J. of Machine Learning Research, 2013. – Vol.28. – P. 1139 – 1147.
2. W.S.McCulloch, W.Pitts. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 1993. – Vol.5. – P. 115 – 133.
3. Бондаренко В. Алгоритм поиска знакомых на основе выделения значимых признаков пользователя в локальных социальных сетях
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2004. – 1072 с.

УДК 004.055

Information Control Systems and Technologies, pp. 140-143

**Ухіна Г.В., Тепличук А.М., Кір'як Ю.Ф., Богатова О.О.,
д.т.н. Ситніков В.С.
УПРАВЛІННЯ СМУГОВИМЧАСТОТНО-ЗАЛЕЖНИМ
КОМПОНЕТОМ ІНФОРМАЦІЙНО-УПРАВЛЯЮЧОЮ
СИСТЕМОЮ ДЛЯ УСУНЕННЯ ДЕТОНАЦІЇ ДВИГУНА
ВНУТРИШНЬОГО ЗГОРАННЯ**

**Ukhina H.V., Teplichuk A.M., Kiryak U.F., Bogatova O.O.,
Dr.Sci. Sytnikov V.S.
MANAGEMENT BANDPASS FREQUENCY-DEPENDENT
COMPONENTS OF INFORMATION MANAGEMENT SYSTEM FOR
ELIMINATING DETONATION COMBUSTION ENGINE INTERIOR**

Сучасний стан розвитку гібридних та електричних автомобілів не знімає проблеми підвищення економічності бензинового двигуна внутрішнього згорання і зниження токсичних газів у вихлопах автомобіля є однією із важливих задач при розробці та експлуатації автомобіля. Несумісність цих показників приводить до ускладнення системи керування двигуном.