DOI: 10.17117/na.2015.10.03.244

<http://ucom.ru/doc/na.2015.10.03.244.pdf>

Поступила (Received): 13.10.2015

Соловьев В.А., Шемяхина А.К.
Моделирование графа социальной сети с помощью
булевозначного представления для построения
сценария поиска информации из открытых
источников в сети интернет

Solovev V.A., Shemyahina A.K.
Simulation of a social network graph using the Boolean valued
representation to build the scenario of searching
information from open sources on the Internet

В работе рассматривается один из открытых источников для разведки – социальные сети в сети Интернет. Социальные отношения между объектами представляются в виде графика социальной сети. Социальная сеть график моделируется как сети на основе булевой алгебры, в котором вершины являются объектами социальная сеть и края связь между ними. Края отмечены элементов булевой алгебры. Атомы булевой алгебры соответствуют теме сообщений, передаваемых между объектами

Ключевые слова: разведка по открытым источникам, анализ социальных сетей, булевозначные сети

Соловьев Вадим Анатольевич
 Кандидат технических наук, старший преподаватель
 Омский государственный технический университет
 г. Омск, пр. Мира, 11

Шемяхина Анастасия Константиновна
 Студент
 Омский государственный технический университет
 г. Омск, пр. Мира, 11

In the paper, one of the methods of competitive intelligence called web intelligence on the open sources is reviewed. Usage of social networks as a source of information for the analysis is proposed because its users generate large amounts of data. Social relationships between objects are represented as a graph social network. Social network graph is modeled as network based on Boolean algebra in which the vertices are objects social network, and the edges are the connection between them. The edges are marked by elements of Boolean algebra. Atoms of Boolean algebra are correspond to the subject of messages transmitted between objects
Key words: open source intelligence, social network analysis, boolean valued networks

Solovev Vadim Anatolevich
 Candidate of Engineering Sciences, Senior Lecturer
 Omsk state technical university
 Omsk, Mira ave., 11

Shemyahina Anastasiya Konstantinovna
 Student
 Omsk state technical university
 Omsk, Mira ave., 11

Наряду с военными разведками, которые сразу представляются при осмыслении самого термина «разведка» существует и другие ее виды, например, коммерческие, экономические, политические, инженерные и другие разведки. С другой стороны, термин «разведка» у большинства людей попадает под понятия «шпионаж», «кража информации», «агентура», что по своей сущности представляет собой нелегальный метод сбора информации. Однако 80% информации получается легальным способом, например, прослушиванием радиоканалов, наблюдение при помощи разведывательных спутников, пресса и др. [1]. При этом следует уточнить следующие два термина. Открытый источник – информация без требования сохранения ее конфиденциальности. Открытые источники относятся к среде общедоступной информации, и не имеют ограничения в доступе для физических лиц. Общедоступная информация – данные, факты, инструкций или другие материалы, опубликованные или размещенные для широкого использования; доступные для общественности; законно увиденные или услышанные случайными наблюдателями; или представленные на открытых встречах для общественности. Огромную роль в разведке из открытых источников играет сеть Интернет. С тех пор как всемирная «паутина» стала неотъемлемой частью современного общества большую часть информации о человеке можно узнать из его блога, социальной сети и т.д. Люди открыто делятся своей контактной информацией, высказывают свои взгляды и мнения, делятся чувствами, переживаниями и убеждениями.

Интернет-разведка – это мониторинг данных, которые находятся в свободном доступе, без рассекречивания своего внимания к тому или иному объекту. Интернет стал неотъемлемой частью современного общества, поэтому из всех методов конкурентной разведки именно метод разведки по открытым источникам в сети Интернет наиболее оптимален, так как обладает множеством преимуществ, таких как оперативность, анонимность и общедоступность в финансовом плане.

Большой объем данных можно обнаружить в социальных сетях, которые на сегодняшний день обрели преимущественную популярность. Пользователи социальных сетей сами передают свои контактные данные уже в готовом для анализа виде и распространяют интересующую их информацию, что намного облегчает сбор информации. В ходе развития аппарата анализа социальных сетей появились такие термины как Social Network Analysis (SNA), теория социальных сетей и т.п.

Любую социальную сеть можно математически моделировать графом, где вершины представляют собой объекты сети (люди, организации и т.д.), а ребра графа – связи между объектами. Анализ графов позволяет рассчитать целый ряд параметров (диаметр графа, плотность графа, степень вершины, среднее расстояние, степень промежуточности и др.) и найти решение задач [2]: структурный анализ и анализ поведения связей, статистический анализ социальных сетей, определение сообществ в социальных сетях, анализ содержания социальных сетей, анализ медиаданных, анализ текстовой информации в социальных сетях, анализ мультимедийной информации сети, классификация вершин в социальных сетях и др.

В данной работе в качестве модели графа предлагается использовать булевозначную сеть [3]. Булевозначная сеть представляет из себя ориентированный мультиграф, в котором каждое ребро помечается элементом из конечной булевой алгебры B . Объединение элементов, соответствующих ребрам выбранного пути, называют оценкой пути, а число атомов булевой алгебры B , содержащихся в оценке, – разнообразием пути. Маршрут между двумя вершинами является оптимальным по определению, если его разнообразие является наименьшим среди всех путей, а при равенстве разнообразий двух путей, наименьшим будет тот, в который входит меньшее число ребер.

Будем считать, что информация в социальных сетях сгруппирована по тематикам. Каждой тематике поставим в соответствие атом из булевой алгебры B . Тогда сообщение между двумя объектами на произвольную тему может быть разделено на набор «атомарных» тем, а процесс его передачи представлен на графе как ребро с соответствующей меткой. Такая интерпретация модели социальной сети позволяет применять весь спектр алгоритмов оптимизации в булевозначных сетях, что расширяет возможности анализа социальных сетей как вид открытых источников информации. При построении модели предлагается использовать алгоритм, разработанный Е.В. Щерба и В.А. Соловьевым [4], позволяющий находить и строить максимальный поток в графе, соответствующем булевозначной сети.

Таким образом, показана актуальность предложенной задачи и подход к ее решению с применением современной теории булевозначных сетей. В целом, применение моделей булевозначных сетей открывает новые возможности для решения важных проблем, связанных с моделированием сетей (социальных, вычислительных и других), а разработка новых алгоритмов и технологий на базе этого подхода представляет перспективную область исследований.

Список используемых источников:

1. Ющук Е.Л. Конкурентная разведка: источники информации и организация работы с ними // Известия Уральского государственного экономического университета. 2010. Т. 31. № 5. С. 139–146.
2. Charu C. Aggarwal. *Social Network Data Analytics*. 2011. 520 p.
3. Салий В.Н. Оптимизация в булевозначных сетях // Дискретная математика. 2005. Т. 17. Вып. 1. С. 141–146.
4. Щерба Е.В., Соловьев В.А. Задача о максимальном потоке в булевозначной сети и ее приложения // Аппроксимация логических моделей, алгоритмов и задач – АЛМАЗ'2. Омск: Изд-во ОмГТУ, 2015. С. 70–76.

© 2015, Соловьев В.А., Шемяхина А.К.
Моделирование графа социальной сети с помощью
булевозначного представления для построения
сценария поиска информации из открытых
источников в сети интернет

© 2015, Solovlev V.A., Shemyahina A.K.
*Simulation of a social network graph using the
Boolean valued representation to build the scenario
of searching information from open sources on the
Internet*