

А. В. СЕМЁНОВ, А. В. БУХАНОВСКИЙ

МЕТРОЛОГИЧЕСКИЙ АНАЛИЗ В СОЦИАЛЬНЫХ СЕТЯХ

Рассматриваются особенности экспериментального определения количественных и статистических характеристик виртуальных социальных сообществ в сети Интернет.

Ключевые слова: социальная сеть, виртуальные сообщества, погрешность измерений, статистические измерения.

Современный Интернет в свете парадигмы Web 2.0 характеризуется интенсивным развитием социальных сетей. Социальная сеть представляет собой структурированный web-сайт, включающий в себя профили пользователей (информация, представленная ими самостоятельно), связанные между собою различными способами („друзья“, „поклонники“ и пр.). В настоящее время крупнейшей социальной сетью в мире является сеть Facebook (около 500 млн пользователей). Русскоязычная социальная сеть vkontakte.ru содержит около 100 млн пользователей и является одним из самых популярных web-сайтов России. Помимо этого существует значительное число тематических социальных сетей меньшего размера, включая профессиональные виртуальные сообщества [1].

Социальные сети являются принципиально новым объектом социологических измерений, поскольку представляют собой виртуальные „слепки“ общественных отношений в сети Интернет, воспроизводящие с определенной долей достоверности структуру контактов между людьми в реальном обществе. С этим связано появление новых методов экспериментальных исследований, в частности, метода виртуальных этнографий [2]. Однако их продвижение требует развития соответствующего метрологического обеспечения, гарантирующего проведение различных измерений с заданной точностью и достоверностью в рамках единой системы. Это позволяет на общей основе сопоставлять результаты измерений, выполненных в разное время на примере различных сетей, с использованием разных методов и средств измерений.

Одним из наиболее востребованных приложений анализа социальных сетей является выявление скрытых сообществ: групп пользователей, объединяемых общей *неформальной* идеей и состоящих в непосредственном общении между собой. Объектом измерений в данном случае выступает графовая структура с нетривиальными топологическими свойствами — комплексная сеть [3], ассоциированная виртуальному сообществу. Методы измерений строятся на основе алгоритмов поиска топологических структур на данном графе, а средствами измерений являются реализующие их поисковые роботы (web-crawler), выполняющие процедуру прямых измерений характеристик сообществ: их числа, количества членов, структуры связей. На их основе выполняются косвенные измерения различных топологических характеристик. В частности, внутренняя структура сообщества характеризуется такими величинами, как диаметр, средняя длина пути, коэффициенты связности и централизации, распределение степеней вершин. Напротив, для описания свойств сообщества по отношению ко всей сети используются такие характеристики, как количество связей, направленных наружу, количество внешних участников, имеющих более одной связи с членами сообщества, а также метрики, описывающие „перемычки“, или „мосты“ между отдельными слабосвязанными сообществами.

В связи с высокой вычислительной сложностью алгоритмов определения данных характеристик на всей социальной сети процедура их измерения основывается на методе двухэтапного сканирования структуры сообщества. На первом этапе средствами поисковых роботов самой социальной сети выявляются стартовые профили, формально соответствующие заданной

топологии, на втором выполняется последовательный обход смежных узлов сети с проверкой принадлежности к сообществу. Как следствие, погрешность измерений зависит от двух факторов: количества стартовых профилей на первом этапе процедуры и их релевантности формальному описанию члена сообщества. Степень релевантности определяется характеристиками алгоритма поисковой машины самой социальной сети и порождает неустранимую инструментальную погрешность измерений. Количество стартовых профилей определяет погрешность статистических измерений; ее величина обратно пропорциональна средней степени вершин в социальной сети.

Дополнительным источником погрешности измерений характеристик сообществ являются нестационарность и структурная изменчивость, связанная с процессами самоодерации и самоорганизации в социальной сети. Величина погрешности во многом определяется конъюнктурой тематики, определяющей вхождение новых членов в сообщество (например, при обсуждении каких-либо „громких“ событий) и не может быть определена априори. Для ее учета может быть использован метод последовательных серий, позволяющий итерационно уточнять значения измеренных характеристик, одновременно оценивая тенденцию изменения погрешности.

Существенная зависимость характеристик составляющих погрешности от топологии сети и внешних факторов, обуславливающих ее динамику, приводит к необходимости выполнения метрологического анализа с использованием численных алгоритмов моделирования комплексных сетей, идентифицированных по интегральным данным измерений [4]. Это дает возможность, применяя метод Монте-Карло, строить интервальные оценки выборочных оценок характеристик сообщества и обосновывать выбор числа стартовых профилей и вложенности обходов в измерительной процедуре.

Работа выполнена при частичной поддержке ФЦП „Научные и научно-педагогические кадры инновационной России“ на 2007—2013 гг., проекта „Интеллектуальные технологии распределенных вычислений для моделирования сложных систем“, а также проектов в рамках реализации постановлений Правительства РФ № 218 „О мерах государственной поддержки развития кооперации российских высших учебных заведений и организаций, реализующих комплексные проекты по созданию высокотехнологичного производства“ и № 220 „О мерах по привлечению ведущих ученых в российские образовательные учреждения высшего профессионального образования“.

СПИСОК ЛИТЕРАТУРЫ

1. *Hu D., Kaza S., Chen H.* Identifying Significant Facilitators of DarkNetwork Evolution // *J. American Society for Inf. Sci. and Techn.* 2009. Vol. 60, N 4. P. 655—665.
2. *Ameripour A., Newman M., Nicholson B. A.* Convivial Tool? // *The Case of the Internet in Iran. J. of Inf. Techn.* 2010. Vol. 25. P. 244—257.
3. *Vaccara N.* Modeling Complex Systems. NY: Springer, 2004. 397 p.
4. *Иванов С. В., Колыхматов И. И., Бухановский А. В.* Параллельные алгоритмы моделирования комплексных сетей // *Изв. вузов. Приборостроение.* 2008. Т. 51, № 10. С. 5—12.

Сведения об авторах

- Александр Валерьевич Бухановский** — д-р техн. наук, профессор; НИИ Научно-технологических компьютерных технологий Санкт-Петербургского государственного университета информационных технологий, механики и оптики; директор;
E-mail: avb_mail@mail.ru
- Александр Владимирович Семёнов** — аспирант; Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кафедра информационных систем; E-mail: avsemyonov@gmail.com

Рекомендована кафедрой
информационных систем

Поступила в редакцию
13.01.11 г.