

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ СОЦИОДИНАМИКИ НА КОМПЛЕКСНЫХ СЕТЯХ В РАСПРЕДЕЛЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ СРЕДЕ

В.В. Каширин

СПбГУ ИТМО

E-mail: kashirin.victor@gmail.com

Введение

Социальные сети являются основой для быстрого распространения информации, правдоподобность которой не поддается адекватной проверке – слухов, в том числе, распространяемых умышленно. Моделирование распространения слухов [1,2] обычно учитывает топологию сети, однако не отражает специфические неоднородности и гетерогенную природу связей внутри сообществ и групп, входящих в нее. При этом моделируемые сети включают в себя большое число вершин и связей между ними, что повышает вычислительную сложность моделирования, как структуры сети, так и процессов на ней. В работе рассматривается моделирование процесса распространения слухов в сети с неоднородной структурой неформальных сообществ с использованием многофункциональной инструментально-технологической платформы CLAVIRE (CLoud Applications VIRtual Environment).

Сервисы моделирования сценариев

Моделирование распространения слухов условно делится на два этапа: моделирование структуры специфической контактной сети и моделирование динамики распространения информации (слуха) в ней. Для моделирования структуры сети предлагается расширенная конфигурационная модель, которая позволяет построить сеть из вершин нескольких классов, где каждому классу соответствует определенный закон распределения степеней среди вершин этого класса. Для моделирования процесса распространения слухов используется модель Далея-Кендалла [3]. Она оперирует тремя множествами вершин: неинформированные (I), распространяющие (S) и неактивные (R). На каждом шаге при взаимодействии вершины типа I с вершиной типа S, она с заданной вероятностью переходит во множество S, а при взаимодействии с вершиной из S или R – во множество R. Значения вероятностей можно задавать для всей сети, для каждого класса вершин и так же для каждой пары классов, вершины которых вступают во взаимодействие. За шаг алгоритма можно взять сутки (что связано с цикличностью просмотра страниц пользователями социальных сетей). Взаимодействие между двумя вершинами определяется наличием связи между ними.

Реализация процесса моделирования в распределенной среде

Для повышения производительности процедура построения модели сети распараллеливается на несколько вычислителей на этапе связывания узлов сети друг с другом. Так, при связывании узлов внутри одного класса может быть использовано от 2 до Q вычислителей, где Q – число классов вершин сети, а при связывании вершин разных классов может быть использовано от 1 до Q/2 вычислителей. Процесс распространения слухов не поддается распараллеливанию, поскольку требует активного обмена боль-

шим количеством информации между вычислителями, вследствие чего распределение процесса моделирования между несколькими процессорами является неэффективным.

Для моделирования явлений социодинамики стало необходимым применение распределенной системы, поскольку исследования проводятся на больших наборах модельных графов с различными топологическими свойствами и в широком диапазоне изменений параметров динамических процессов. С учетом этого методы были реализованы с использованием многофункциональной инструментально-технологической платформы CLAVIRE (CLOUD Applications VIRTUAL Environment). Задачи моделирования сетей различных типов, моделирования динамики распространения слухов и анализа результатов разделены между собой и внедрены в распределенную систему, а обмен данными производится через централизованное хранилище. Таким образом, расчеты проводятся независимо друг от друга и в порядке поступления новых данных. Данный подход показал свою эффективность, поскольку на одной модельной сети могут параллельно исследоваться несколько процессов распространения слухов с различными характеристиками, которые существенно более затратные по вычислениям, нежели генерация сети.

Сценарии использования

Предлагаемый подход может быть использован для моделирования ситуаций, различных по параметрам и характеристикам, с целью изучения ключевых факторов, влияющих на процесс распространения слухов в социальной среде. При этом могут численно проверяться гипотезы о структуре сети, а также строиться интервальные оценки прогнозируемых процессов. В качестве примера на рис. 1 приведены результаты моделирования сети, состоящей из двух классов вершин. Распределение степеней первого класса характеризуется степенным законом с показателем степени 4, а второго – законом Пуассона с показателем 7. Вершины первого класса (80%) определяют традиционную структуру социальной сети, тогда как второго (20%) соответствуют сплоченному сообществу с большим количеством связей между участниками.

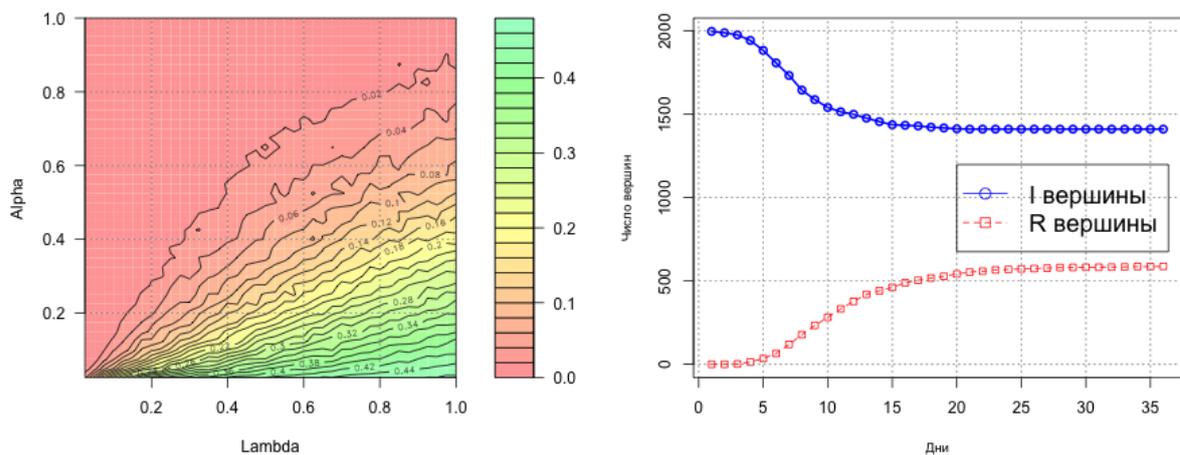


Рис. 1. а)

б)

На рис. 1а приведен график, демонстрирующий степень покрытия вершин сети слухом в зависимости от параметров процесса распространения. Видно, что наилучшее покрытие обеспечивается при высокой вероятности передачи слуха и низкой вероятности его забывания. На рис. 1б приведена динамика распространения слуха, выраженная через число вершин во множествах I и R на определенном шаге работы алгоритма. Предлагаемый подход можно развивать при помощи различных моделей структуры

контактных сетей и процессов распространения информации, а так же их модификаций, с целью формирования универсального набора методов моделирования для различных специфических случаев распространения информации. В общем, типовыми сценариями исследования посредством разработанной технологии являются:

- построение социограммы неформального сообщества, анализ скорости и каналов распространения информации;
- анализ и прогноз индексов общественных настроений;
- выявление групп влияний в социальной сети и определение «лидеров мнений»;
- мониторинг манипуляций мнениями;
- выявление призывов к офф-лайн общественным/экстремистским акциям;
- обнаружение источников умышленной дезинформации.

Заключение

В целом разработанная технология моделирования представляет собой эффективный инструмент изучения предмета социодинамики, как в рамках исследовательских экспериментов, так и в условиях экстренных вычислений, когда осуществить прогнозирование определенного процесса необходимо в сжатые сроки и с определенной точностью. Распределенная среда позволила разделять задачи и сценарии моделирования структуры, моделирования процессов и анализа результатов, что повлияло на скорость и эффективность обработки данных при исследованиях.

Литература

1. Nekovee M., Moreno Y., Bianconi G., Marsili M. Theory of rumour spreading in complex social networks // *Physica A: Statistical Mechanics and its Applications*. 2007. 374 (1). P. 457-470.
2. Pittel B. On Spreading a Rumor // *SIAM Journal on Applied Mathematics*. 1987. Vol. 47, No. 1. P. 213-223.
3. Daley D., Kendall D. Epidemics and rumours // *Nature*. 1964.