

## **МОДЕЛИ И АЛГОРИТМЫ В ТОПОЛОГИЧЕСКИ ГЕТЕРОГЕННЫХ СОЦИАЛЬНЫХ СЕТЯХ, ФУНКЦИОНИРУЮЩИХ В УСЛОВИЯХ ДЕСТАБИЛИЗИРУЮЩИХ ИНФОРМАЦИОННЫХ ВОЗДЕЙСТВИЙ И ФАКТОРОВ**

Владимирский государственный университет им. А.Г. Н.Г. Столетовых

На сегодняшний день основная цель моего исследования это определение модели топологии участка реальной социальной сети. Для достижения поставленной цели были поставлены следующие задачи:

- составление набора моделей топологий сетей, предположительно применимых в рамках текущей проблемы;
- получение топологии участка реальной социальной сети;
- выявление классификационных признаков, достаточных для определения типа топологии социальной сети.

В своих работах исследователи приводят различные классификации сетевых топологий и характерные для них свойства. По одной из классификаций, существует три класса сетей «тесного мира»: широкомасштабные сети, одномасштабные сети и безмасштабные сети. Первый тип характеризуется функцией распределения степеней связности, которая представляет собой смесь степенного закона и экспоненциального распределения. Для графов второго типа характерно резкое распределение, которое означает, что узлы с очень высокой степенью связности отсутствуют. Одним из таких примеров является модель Уоттса-Строгатца. Наконец, безмасштабные сети (SF) строго описываются степенным законом распределения; подразумевается, что вероятность наличия узлов с большой степенью связности достаточно высока. Наиболее популярный пример SF сети - модель Барабаши и Альберт (BA).

Ученые сходятся во мнении, что социальные сети представляют собой SF сети или же small world. Поэтому в нашей работе мы в основном ориентируемся на эти два типа топологии.

В рамках исследования были получены топологические данные об участке социальной сети Facebook размером 6927 узлов. На данном этапе работы имеется также большой участок Facebook с количеством узлов приблизительно в 1.5 млн. Ведется анализ и реализация обработки такого большого количества информации.

Ниже приведены признаки, по которым характеризуют топологию сети:

– *число связей в сети из N узлов* – число всех возможных связей между N узлами.

– *средняя длина пути l* – средняя длина пути между всевозможными парами вершин.

– *распределение степеней вершин* – зависимость для узла  $k_i$  вероятности того, что узел будет иметь степень  $k$ , от числа вершин, с которыми соединен  $i$ -й узел.

– *коэффициент кластеризации* определяется как вероятность того, что два узла, связанных с данным узлом, также связаны между собой.

$$C_i = \frac{E_i}{\frac{1}{2}k_i(k_i - 1)} \quad (1),$$

где  $k_i$  – число вершин, с которыми соединен данный узел,  $E_i$  – общее число связей, между  $k_i$  узлами. Общий коэффициент кластеризации для сети из N узлов будет равен  $C = \frac{1}{N} \sum_i C_i$ .

В рамках нашего исследования для классификации мы брали две характеристики: распределение степеней вершин и коэффициент кластеризации.

Топологии и соответствующие значения характеристик приведены в таблице.

Таблица 1

Параметр	WS Small World	Scale free BA	Random Graphs
Число связей	$kN/2$	$m(N-1)$	$\langle k \rangle N/2$
Средняя длина пути	$\sim \frac{\ln N}{\ln k}$	$m = 1: l \sim \ln N;$ $m \geq 2: l_{BA}^{\alpha > 3} \approx \ln N;$ $l_{BA}^{\alpha = 3} \approx \ln N / \ln \ln N;$ $l_{BA}^{2 < \alpha < 3} \approx \ln \ln N.$	$\sim \frac{\ln N}{\ln \langle k \rangle}$
Кластерный коэффициент	$C_{p \rightarrow 1} \sim k/N,$ $C_{p > 0} \gg C_{p \rightarrow 1}$	$C \sim 5 \frac{\langle k \rangle}{N}$ [1]	$C \sim \frac{\langle k \rangle}{N}$
Распределение степеней вершин	Закон Пуассона: $p(k) \sim e^{-pN} \frac{(pN)^k}{k!}$	Степенной закон: $p(k) \sim k^{-a}$	Закон Пуассона: $p(k) \sim e^{-pN} \frac{(pN)^k}{k!}$

В рамках проведенных исследований были получены следующие результаты:

- 1) распределение узлов по степени связности носит ярко выраженный степенной характер.

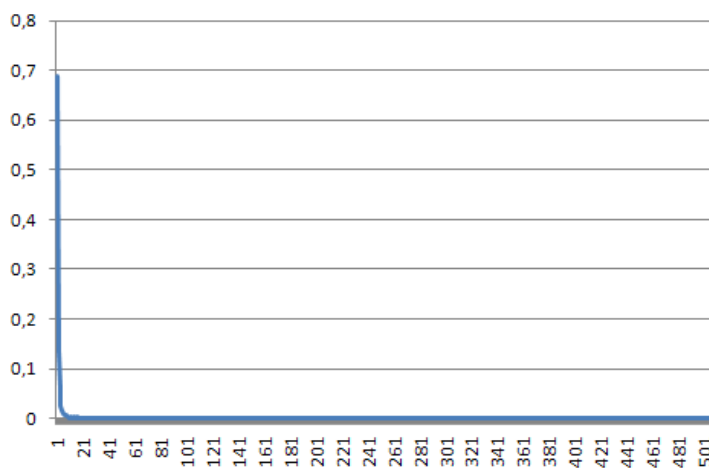


Рис. 1а. Вероятность указанного числа смежных узлов

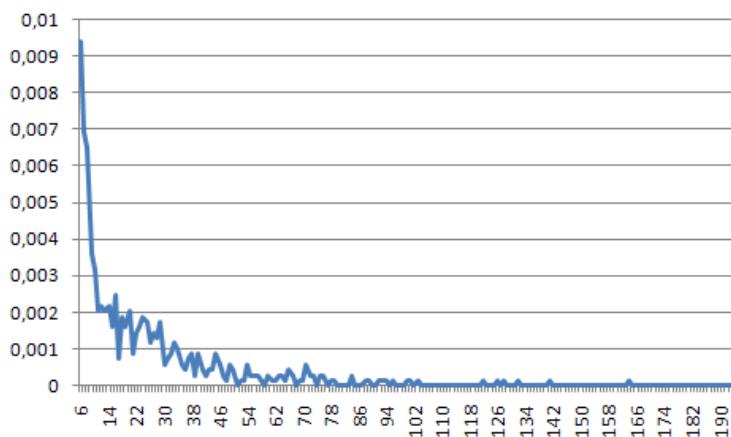


Рис. 1б. Вероятность указанного числа смежных узлов на отрезке [6;195]

- 2) значение кластерного коэффициента отличается на два порядка. Данный результат анализируется. Имеются предположения о не совсем удачном выбранном участке сети.

Другим важным направлением моего исследования является анализ перколяционного кластера.

Одним из важнейших понятий для SF-сетей является перколяционный кластер (большая связная компонента). В понимании авторов перколяционный кластер представляет собой совокупность связанных между собой узлов с высоким кластерным коэффициентом. Иммунизация узлов кластера коренным образом влияет на распространение нежелательной информации в сети. Алгоритм поиска кластера состоит из следующих шагов:

- 1) для всех узлов в сети высчитывается кластерный коэффициент по формуле (1);
- 2) задается пороговый коэффициент, значение которого должно находиться между максимальным и минимальным значениями коэффициентов в заданной сети;
- 3) берется узел с наибольшим значением кластерного коэффициента и от этого узла через его соседей рекурсивно формируется кластер. Критерий вхождения узла в кластер - значение его кластерного коэффициента должно быть выше порогового.

При подсчете кластерных коэффициентов была построена функция распределения этих значений. Результат неоднозначен, ведется его анализ.

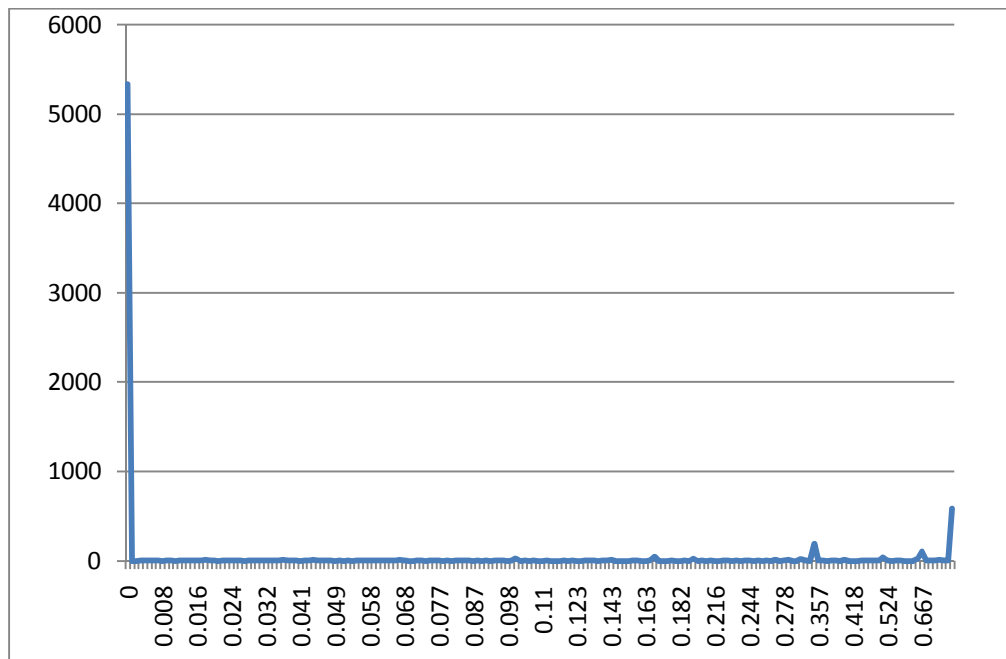


Рис. 2 Распределение кластерных коэффициентов

Авторами была выдвинута следующая гипотеза. В ходе процесса распространения нежелательной информации в социальной сети скорость распространения будет коррелировать с количеством узлов в перколяционном кластере.

Для подтверждения или опровержения данной гипотезы было проведено имитационное моделирование. Для того, чтобы проследить динамику изменения узлов в перколяционном кластере, использовался следующий алгоритм:

- 1) если в ходе процесса в перколяционном кластере иммунизируется узел, то этот узел «выбрасывается» из сети;
- 2) пересчитываются кластерные коэффициенты для оставшихся узлов в перколяционном кластере;
- 3) если у узла из перколяционного кластера коэффициент падает ниже порогового, то данный узел «выбрасывается» из сети.

По результатам моделирования были сделаны следующие выводы. В конце процесса распространения нежелательной информации в сети в кластере остаются подверженные узлы, которые были отсечены в ходе моделирования, по причине иммунизации или снижения кластерного

коэффициента. Если говорить о выдвинутой гипотезе, то значение корреляционного коэффициента оказалось равным 0.4, что говорит о слабой корреляции.

Были проведены исследования по изучению зависимости количества узлов в кластере от порогового значения. При моделировании пороговый коэффициент вычислялся по следующей формуле

$$lev=(max\_val+min\_val)/k$$

Значение  $k$  изменялось в диапазоне [2, 50]. Полученные результаты представлены на рис.

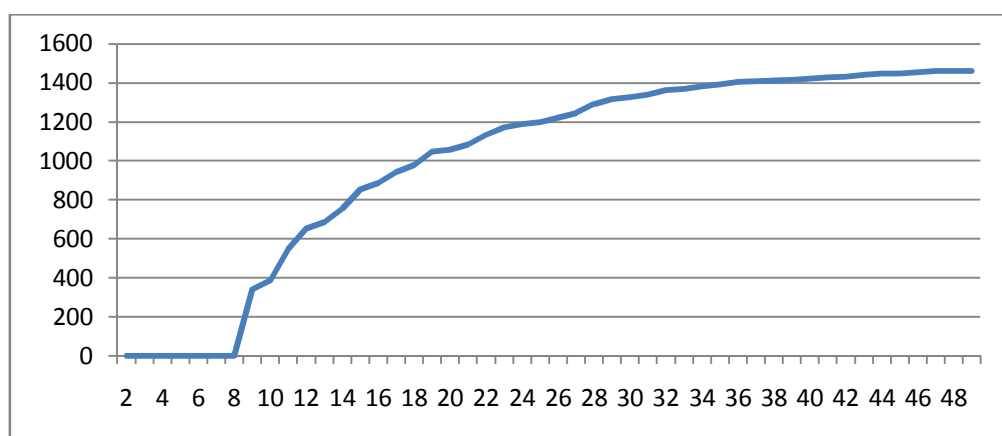


Рис. 3 Зависимость числа узлов в кластере от  $k$

По результатам эксперимента видно, что при значении  $k=8.9$  происходит фазовый скачок.

## Литература

1. Груздева Л.М., Монахов М.Ю. Алгоритм оптимизации функционирования распределенной системы защиты // Вестник Костромского государственного университета им. Н.А.Некрасова. - №2, 2008. - С. 80-82.
2. Жаринов И.В., Крылов В.В. Конструирование графов с минимальной средней длиной пути // Вестник Ижевского государственного технического университета -№4, 2008. –С. 164-169.
3. Монахов М.Ю., Груздева Л.М. Алгоритм раннего обнаружения атак на информационные ресурсы АСУП // Автоматизация в промышленности. - №3, 2008. С. 12-14.
4. Leveille J. Epidemic Spreading in Technological Networks // M. Sc. Thesis, HP Labs Bristol, 2002 – 100 pp.

### **Сведения об авторах**

Монахов Юрий Михайлович, Владимирский государственный университет, доцент кафедры ИЗИ, к.т.н., unklefck@gmail.com

Абрамов Константин Германович, Владимирский государственный университет, инженер кафедры ИЗИ, аспирант, abramovkostya@mail.ru