

Министерство науки и высшего образования Российской Федерации  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)  
Институт прикладной математики и компьютерных наук

ОТЧЕТ ПО ПРАКТИКЕ

Вид практики: производственная практика

Тип практики: Научно-исследовательская работа

Энгельке Сергей Алексеевич

Направление подготовки 01.03.02 Прикладная математика и информатика  
Направленность (профиль) «Математические методы в экономике»

Руководитель практики от НИ ТГУ  
д-р физ.-мат. наук, доцент

\_\_\_\_\_ А.Н. Моисеев  
*подпись*  
« \_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

Руководитель практики от  
профильной организации  
ученая степень, звание, место работы,  
должность

\_\_\_\_\_ М.П. *подпись*  
« \_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

Составитель отчета  
студент группы № 932023

\_\_\_\_\_ С.А. Энгельке  
*подпись*  
« \_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ г.

Томск –2023

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	3
Описание модели .....	4
Описание имитационной модели .....	8
Допущения реализуемой модели.....	10
Описание работы алгоритма сетевой динамики.....	11
ЗАКЛЮЧЕНИЕ .....	12
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ..	13

## ВВЕДЕНИЕ

В современном мире изучение социальных сетей приобретает всё большую значимость. Это связано с необходимостью анализа и оптимизации различных социальных процессов, будь то в науке, экономике или политике. С развитием сети интернет появилась возможность собирать и анализировать объективные данные о социальных сетях, и их структурах. Это вызвало развитие такой дисциплины, как сетевой анализ. В рамках сетевого анализа существует класс акторно-ориентированных моделей. Стохастическое акторно-ориентированное моделирование является мощным инструментом для анализа и предсказания сетевой динамики.

Объектом данного исследования является класс стохастических акторно-ориентированных моделей, а именно построение этих моделей, и построение имитационной модели. Целью данной работы является построение и анализ стохастических акторно-ориентированных моделей и симуляция сетевой динамики. Для достижения поставленных целей можно выделить такие задачи:

1. Описание модели САОМ для которой будет реализована имитационная модель;
2. Описать допущения реализуемой модели;
3. Разработать алгоритм работы имитационной модели.

Полученная в ходе данной работы имитационная модель будет использоваться для разработки и реализации собственных компонент САОМ для пакета «RSiena».

## Описание модели

Обозначим  $X(t)$  как состояние сети в момент времени  $t$ . В любой момент времени в сети происходит не более одного изменения. Этими изменениями могут быть: создание или уничтожение связи.

Вероятность изменения состояния сети зависит от её текущего состояния, и не зависит от истории изменений. Таким образом  $X(t)$  – непрерывная цепь Маркова. В этом контексте непрерывная цепь Маркова означает, что в каждый момент времени вероятность перехода сети в новое состояние определяется исключительно текущими связями и атрибутами акторов, без учета последовательности предыдущих изменений. На рисунке 1 можно увидеть пример изменения процесса  $X(t)$  за один шаг. В данном случае изменением является создание новой связи между акторами 1 и 3. Где структура сети на рисунке 1 представлена на рисунке 2, новое ребро выделено кругом.

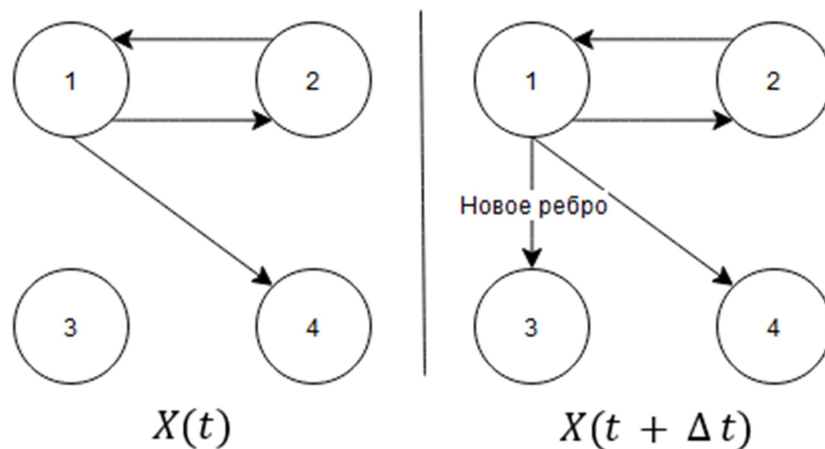


Рисунок 1 – Пример изменения процесса  $X(t)$

	1	2	3	4		1	2	3	4	
1	0	1	0	1		1	0	1	1	
2	1	0	0	0		2	1	0	0	
3	0	0	0	0		3	0	0	0	
4	0	0	0	0		4	0	0	0	
$X(t)$						$X(t + \Delta t)$				

Рисунок 2 – Пример представления процесса  $X(t)$

Частота принятия решений актором  $i$  зависит от параметра  $\lambda_i$ . Обозначим  $\lambda \triangleq \sum \lambda_i$ .

$$P\{\text{Следующая возможность изменения наступит в } t + \Delta t \mid \text{Текущее время} - t\} = \exp(-\lambda_i * \Delta t)$$

Моменты времени распределены в соответствии с экспоненциальным распределением.

$$P\{\text{Следующая возможность изменения предоставлена актору } i\} = \frac{\lambda_i}{\lambda}$$

Первый получивший возможность произвести изменение делает свой выбор и все начинается с начала, но уже в новом состоянии.

У каждого актора есть  $n$  вариантов выбора, и всего  $n - 1$  вариантов изменения сети. Построить\разрушить связь с кем то из других вершин, или ничего не делать. Для выбранного актора строится функция полезности для оценки вероятности выбора каждого из  $n$  вариантов. Функция полезности необходима для определения вероятности того, что при следующей смене связей данный актор перейдет из состояния  $x$  в  $x^{(\pm ij)}$ .

Функция полезности:

$$f_i(x, x^{(\pm ij)}, \beta),$$

где  $i$  – индекс актора принимающего решение о связи с актором  $j$ ;

$x$  – текущее состояние сети, для графа и  $i \neq j$ ,

$x^{(\pm ij)}$  – граф, который идентичен по всем рёбрам, кроме пары  $(i, j)$ , для которой существование связи является переключателем  $x^{(\pm ij)} = 1 - x_{ij}$ , причём возможен случай когда  $i = j$  это значит, что актор не изменит текущее состояние сети.

$f_i(x, x^{(\pm ij)}, \beta)$  определена на множестве всех пар сети  $x^{(\pm ij)}$  и  $x$ . В интерпретации полезности функция полезности может рассматриваться как чистая полезность, которую получает актор  $i$  от перехода из состояния  $x$  в  $x^{(\pm ij)}$ .

Предположим, что в момент времени  $t + \Delta t$ , при текущей сети  $x$  у актора появилась возможность изменения в состояние  $x^{(\pm ij)}$

$$\frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\sum_h \exp(f_i(x, x^{(\pm ih)}, \beta))}.$$

Это распределение вероятностей выборов актора представляет собой модель множественного выбора при  $n$  вариантов решений у вершины. Подробнее про построение функции полезности:

Функция полезности:

$$f_i(x, x^{(\pm ij)}, \beta) = \sum_{k=1}^K \beta_k s_{ki}(x, x^{(\pm ij)}).$$

Функция полезности описывает полезность перехода из состояния  $x$  в  $x^{(\pm ij)}$  для  $i$  актора.

В рассматриваемой модели функция полезности – есть линейная комбинация параметров  $\beta$  и функции сети, где функция сети описывает то, как выглядит изменение сети из текущего состояния в состояние новое с точки зрения актора  $i$ .

Функция сети:

$$s_{ki}(x, x^{(\pm ij)}),$$

где  $k$  – индекс параметра модели. Существует несколько реализаций функции сети [6] [3].

Вероятность изменения связи  $i \leftrightarrow j$ :

$$P\{X(t + \Delta t) = x^{(\pm ij)} | X(t) = x\} = \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\sum_h \exp(f_i(x, x^{(\pm ih)}, \beta))},$$

где  $X(t)$  – матрица смежности с элементами  $x_{ij}$  представляющая сеть в момент времени  $t$ .

Таким образом Стохастические Акторно-ориентированные модели (САОМ) работают по следующему алгоритму:

1. Зададим  $t = t_m, x = x(t_m)$ .
2. Генерируем  $\Delta t$  в соответствии с экспоненциальным временем с параметром  $\lambda$ .
3. Если  $t + \Delta t > t_{m+1}$  установить  $t = t_{m+1}$  и остановиться.
4. Выбираем случайный  $i \in 1..n$  используя распределение вероятностей  $\frac{\lambda_i}{\lambda}$ .
5. Выбираем случайным образом  $x \in A_i(x)$ , используя вероятности  $p_i(x, x^{(\pm ij)}, \beta)$ .  $A_i(x)$  - множество всех состояний в которое может измениться сеть  $x$  в следствии решения актора  $i$ .
6. Задать  $t = t + \Delta t$ .
7. Задать  $x = x^{(\pm ij)}$ .
8. Вернуться к шагу (2).

На принятие решения может влиять множество эффектов, и параметров при этих эффектах. Дальнейшей задачей является оценка этих параметров.

## Описание имитационной модели

Исследование реальных данных проводилось с использованием пакета RSiena. RSiena предоставляет широкий спектр различных реализаций компонентов модели и распространяется с открытым исходным кодом, а также имеет активную поддержку сообщества для написания пользовательских компонентов.

Пользовательский компонент может представлять собой эффект, например, влияние временного ряда научных конференций на степень кластеризации всей сети. Такой компонент должен включать в себя анализ временного ряда, использование dummy variable и классический компонент из RSiena - связные тройки.

Для проверки корректности пользовательских эффектов необходима имитационная модель CAOM.

Была построена модель для направленной сети, где возможные результаты решения актора включают создание новой связи, разрушение существующей связи или оставление сети в том виде, в котором она есть в текущий момент времени. Разрушение связи происходит в одностороннем порядке, а создание связи также осуществляется в одностороннем порядке.

Вероятность изменения связи  $i \leftrightarrow j$ :

$$P\{X(t + \Delta t) = x^{(\pm ij)} | X(t) = x\} = \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\sum_h \exp(f_i(x, x^{(\pm ih)}, \beta))},$$

Функция полезности – есть линейная комбинация функций сети и их параметров. Для имитационной модели были взяты такие сетевые эффекты как:



1. Влияние количества исходящих связей,
2. Влияние количества общих связей.

В таком случае функция полезности данной модели имеет вид

$$f_i(x, x^{(\pm ij)}, \beta) = \beta_2 * \left( \sum_j x_{ij} x_{ji} \right) + \beta_1 * \left( \sum_j x_{ij} \right) \quad (1)$$

## Допущения реализуемой модели

Из-за отсутствия использования в имитационной модели эффектов, связанных с ковариатами, выбор актора, принимающего решение, можно упростить. Для каждого актора интенсивности рассматриваются одинаковыми, что позволяет применить равномерное распределение вероятностей при выборе актора на этапе принятия решения.

Сеть представлена квадратной матрицей, заполненной нулями и единицами. Начальное состояние сети на момент  $t_0$  генерируется случайно с использованием равномерного распределения.

Интервалы времени между снимками сети одинаковы по длительности.

## Описание работы алгоритма сетевой динамики

В начале генерируется начальное состояние сети, представленное матрицей смежности размерности  $N$ , где  $N$  – количество акторов в сети. Начальная матрица заполняется единицами с заданным пользователем коэффициентом разреженности. Затем стимулируется социальная динамика до момента времени начала следующего снимка сети  $t_{m+1}$ .

Процесс повторяется дважды, используя каждый раз снимок, полученный на предыдущем шаге, таким образом, в результате работы получается три снимка сети.

Подробнее про симуляцию сетевой динамики. Положим  $t_m = 0, t_{m+1}$  задаётся пользователем, вектор параметров  $\beta$  при эффектах задаётся пользователем. Генерируем  $\Delta t$  экспоненциально, с параметром  $\lambda$ . Пока  $t_m + \Delta t$  меньше  $t_{m+1}$  используя распределение акторов выбирается актор принимающий решение о изменении сети в момент  $t_m + \Delta t$ . Если  $t_m + \Delta t$  меньше  $t_{m+1}$  то генерируются все возможные состояния сети, в которые можно перейти из текущего состояния. Выбранный актор (актор  $i$ ) используя распределение вероятностей переходов во все возможные варианты новой сети  $x^{(\pm ij)}$  принимает решение о изменении сети. Процесс заканчивается когда  $t_m + \Delta t > t_{m+1}$ .

Для симуляции принятия решения строится вектор полезностей для каждого актора при зафиксированном  $i$  акторе. Каждый элемент этого вектора отражает полезность изменения сети из текущего значения в следующее (минишаг). Считается функция полезности для каждого возможного изменения, где конкретный вид функции полезности задан формулой (1). Используя полученные функции полезности, актор  $i$  принимает решение о том, какое изменение ему внести в сеть (либо не вносить изменение, в случае если большая полезность у не изменения сети). Измененное состояние сети записывается как текущее, и начинается новая итерация цикла с генерации  $\Delta t$  и следующего актора, которому будет предоставлено право изменения сети.

## **ЗАКЛЮЧЕНИЕ**

В рамках данной работы была разработана и реализована имитационная модель сетевой динамики для стохастической акторно-ориентированной модели.

Класс моделей САОМ показывает высокую эффективность для анализа сетевой динамик, однако является сложным инструментом для анализа, в силу высокой вычислительной сложности, а так же требования социологических гипотез. Не смотря на то, что существует пакет RSiena, позволяющий проводить анализ социальных сетей, имитационная модель необходима для углубленного анализа конкретных сетей. При необходимости написания собственных компонент пакета RSiena, будь то пользовательские эффекты сети, или варианты принятия решения, необходима проверка корректности работы данных эффектов. Что позволить провести динамику социальной сети и проверить на согласованность этих компонент в пакете RSiena. Собственные компоненты могут быть необходимы для более точного описания социологической модели взаимодействия акторов в сети, в отличие от стандартных эффектов.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ И ЛИТЕРАТУРЫ

1. Докука С.В., Валеева Д.Р. Статистические модели для анализа динамики социальных сетей в исследованиях образования // Вопросы образования. - 2015. - №1. - С. 201-213.
2. Anuška Ferligoj, Luka Kronegger, Franc Mali, Tom A B Snijders, Patrick Doreian Scientific collaboration dynamics in a national scientific system // Scientometrics. - 2015. - №104. - С. 985–1012.
3. Snijders, T.A.B., & Pickup, M. (2016). Stochastic Actor-Oriented Models for Network Dynamics. Retrieved June 10, 2016.
4. Manual for RSiena / M. R. Ruth. — Текст : электронный // University of Oxford: Department of Statistics; Nuffield College University of Groningen: Department of Sociology : [сайт]. — URL: [https://www.stats.ox.ac.uk/~snijders/siena/RSiena\\_Manual.pdf](https://www.stats.ox.ac.uk/~snijders/siena/RSiena_Manual.pdf) (дата обращения: 14.01.2024).
5. Snijders The statistical evaluation of social network dynamics / Snijders, A T. — Текст : непосредственный // Sociological Methodology. — 2001. — № 31. — С. 361-395.
6. Siena algorithms // Department of Statistics – University of Oxford URL: [https://www.stats.ox.ac.uk/~snijders/siena/Siena\\_algorithms.pdf](https://www.stats.ox.ac.uk/~snijders/siena/Siena_algorithms.pdf) (дата обращения: 06.01.2024).
7. Snijders, T. A. B. The Statistical Evaluation of Social Network Dynamics // Sociological Methodology. - 2001. - №31. - С. 361-395.
8. Tom Broekel, Pierre-Alexandre Balland, Martijn Burger, Frank Oort Modeling knowledge networks in economic geography: a discussion of four methods // The Annals of Regional Science. - 2014. - №53. - С. 423-452.