

# Введение

---

В наше время научная деятельность играет важную роль в развитии общества и науки. Однако, для достижения наилучших результатов, необходимо эффективное взаимодействие между научными группами и командами. В связи с этим, вопрос выявления научных групп и команд, а также факторов, влияющих на формирование и разрушение научного взаимодействия, становится все более актуальным.

Целью данной курсовой работы является выявление факторов влияющих на формирование научных групп и команд Томского Государственного Университета на основе библиометрического анализа, а также определение факторов, влияющих на формирование и разрушение научного взаимодействия. Для достижения поставленной цели используется стохастическое акторно-ориентированное моделирование (CAOM).

Стохастическое акторно-ориентированное моделирование используется для изучения динамики социальных сетей. Разработал CAOM Tom A.B. Snijders. Класс моделей CAOM объединяет теорию полезности и марковские процессы. CAOM применялся для анализа динамики социальных сетей в исследованиях образования [1], для анализа экономических взаимодействий между географическими акторами [2] и многих других исследованиях социальных сетей.

Я построю несколько временных снимков сети соавторства Томского Государственного Университета используя данные с информационного портала ТГУ.Сотрудники. Исследования с использованием CAOM проводилось для динамики научного сотрудничества в национальной научной системе Словении [3].Результаты исследования могут быть использованы для оптимизации научного взаимодействия в рамках Томского Государственного Университета и дальнейшего развития научной деятельности в целом.

## Глава 1 Теоретическая часть

---

### Научные группы и команды: определение, характеристики, особенности.

---

Научные группы и команды - это организованные сообщества ученых, сотрудничающих для достижения определенных научных целей. Научные группы и команды образуются на базе научных институтов, университетов, научно-исследовательских центров или лабораторий.

#### Характеристики научных групп и команд:

- Члены научной группы или команды имеют общие научные интересы и цели;
- Члены научной группы или команды имеют определенные роли и задачи;
- Члены научной группы или команды сотрудничают для выполнения научных исследований, в том числе и научных публикаций;
- Научные группы и команды могут иметь различные размеры и структуры;
- Научные группы и команды могут работать в рамках одной научной области или пересекаться с другими научными областями.

## Особенности научных групп и команд:

- Результаты работы научных групп и команд могут быть использованы для разработки новых технологий и продуктов;
- Научные группы и команды могут привлекать внешних экспертов для решения конкретных задач;
- Научные группы и команды могут иметь различные формы финансирования, включая государственное и частное финансирование.

## 1.2 Библиометрический анализ: методы и подходы

---

Библиометрический анализ является методом исследования научной литературы и состоит в количественном анализе научных публикаций, цитирований и авторских коллективов с целью выявления закономерностей и тенденций в различных научных областях.

В рамках данной курсовой работы я буду рассматривать такие методы и подходы библиометрического анализа:

- Анализ сетевых структур. Этот подход базируется на анализе сетевых структур, создаваемых научными сообществами в процессе обмена информацией и знаниями. Он позволяет выявить ключевых игроков и их роли в сообществе, а также группы, которые наиболее тесно связаны между собой.
- Кластерный анализ. Этот подход основан на группировке научных работ по подразделению авторов. Он позволяет выявить кластеры научных работ, которые имеют схожую тематику, и определить их особенности и тенденции.

## 1.3 Факторы, влияющие на формирование и разрушение научного взаимодействия

---

Научное взаимодействие представляет собой процесс сотрудничества между учеными с целью достижения общей научной цели. Однако, как и любой другой процесс, научное взаимодействие может изменяться под влиянием различных факторов, которые могут способствовать или, наоборот, препятствовать его формированию и развитию.

Основными факторами, влияющими на формирование и разрушение научного взаимодействия, которые можно выделить в рамках взаимодействия в Томском Государственном Университете:

- Общая научная тематика: Ученые, работающие в одной области науки, имеют больше возможностей для сотрудничества, чем те, кто занимаются разными научными направлениями.
- Финансирование: Наличие достаточных финансовых ресурсов может способствовать формированию научных групп и команд. Однако, в случае недостатка финансирования, ученые могут ощутить трудности в формировании и поддержании научных контактов.
- Личностные факторы: Личностные характеристики ученых, такие как общительность, коммуникабельность и т.д., могут оказывать существенное влияние на формирование и поддержание научных связей.
- Технологический прогресс: Технологический прогресс может предоставить новые возможности для научного сотрудничества.

- Наличие научной инфраструктуры: Наличие необходимой инфраструктуры, такой как лаборатории, библиотеки, научные журналы, тоже является важным фактором для формирования научных групп и команд.

## Глава 2 Методология исследования

---

### 2.1 Объект исследования

---

Объектом исследования данной курсовой работы являются научные группы и команды Томского Государственного Университета, а конкретно данные, собранные о сотрудниках ТГУ, их научных публикациях, подразделениях в котором они работают и сетевых связях. Используя библиометрический анализ и стохастическое акторно-ориентированное моделирование, можно выявить факторы, влияющие на формирование и разрушение научного взаимодействия внутри этих групп и команд.

### 2.2 Методика библиометрического анализа

---

Для исследования использованы сетевые структуры и метрики. Данные выгружены с веб сайта ТГУ.Сотрудники. Используемая информация о авторах: Опыт работы, подразделение в ТГУ и список публикаций автора, образующий сеть соавторства. Исходный код парсера данных (программы по выгрузке и систематизации данных), а так же данные, которые были использованы для исследования доступны по ссылке.

### 2.3 Стохастическое акторно-ориентированное моделирование: суть метода, его применимость в исследовании научных групп и команд

---

Стохастическое акторно-ориентированное моделирование [4] (Stochastic Actor-Oriented Modeling, SAOM) - это метод социально- сетевого анализа, который позволяет моделировать и предсказывать эволюцию социальных сетей на основе поведения и взаимодействий индивидуальных акторов в них. SAOM используется для изучения процессов формирования и разрушения социальных связей в группах и командах, а также для оценки влияния различных факторов на эти процессы.

Применительно к исследованию научных групп и команд, SAOM может быть использован для моделирования процессов формирования научных связей между учеными в рамках группы или команды.

Используя SAOM возможно выявить такие факторы, влияющие на формирование или разрушение связей, как: влияние стажа на то, с какой частотой актор будет принимать решения относительно изменения сети, с какой частотой изменяется сеть, как влияет научное подразделение актора на динамику сети, и т.д.

Преимущества SAOM включают возможность моделирования процессов на уровне индивидуальных акторов, учет динамической природы социальных связей, а также возможность оценки влияния различных факторов на эти процессы. Однако SAOM также имеет некоторые ограничения, такие как необходимость большого количества данных для моделирования и ограниченная возможность предсказания долгосрочных эффектов.

Таким образом, SAOM является мощным инструментом для исследования научных групп и команд, который может помочь в выявлении факторов, влияющих на формирование научных связей и определении эффективных мер по их сохранению.

### 2.3.1 Допущения исследования

1. Необходимо учесть факт, что сеть соавторства является ненаправленной:

$X_{i,j} = X_{j,i}$ , где  $X_{i,j} \in \{0, 1\}$  является элементом графа и обозначает связь  $i \leftrightarrow j$ . В силу ненаправленности сети формирование или разрушение связи между акторами происходит по одному из двух сценариев:

1.1 Односторонняя инициатива: Выбирается один актор  $i$ , который получает возможность произвести изменение.

1.2 Двусторонняя инициатива: Выбирается упорядоченная пара акторов  $(i, j)$ , при и получает возможность принять новое решение о связи  $i \leftrightarrow j$ .

Процесс выбора моделируется по одному из двух сценариев:

D. Диктаторский(Dictatorial): Один актор может навязывать решение о связи другому актору.

M. Взаимный(Mutal): Оба актора дают согласие на существование связи между ними.

2. Между временными наблюдениями происходят изменения, которые не наблюдаются.
3. Вероятность изменения состояния сети зависит от её текущего состояния, и не зависит от истории изменений.
4. В любой момент изменений сети происходит не более одного изменения. Этими изменениями могут быть: создание или уничтожение связи.

Допущения 2 и 3 позволяют рассматривать модель как непрерывную цепь Маркова.

### 2.3.2 Спецификация Акторно ориентированной модели:

Модель является акторной в том смысле, что изменение связи моделируется, как результат выбора сделанного актором. Модель изменения связи состоит из двух компонент: время и выбор.

Время изменения определяется в терминах возможности изменения, а не факта изменения. В случае если актора устраивает его текущее состояние он не будет изменять свои связи.

Рассмотрим текущий момент времени и обозначим текущее состояние сети  $X(t)$ . Частота принятия решений актором  $i$  зависит от параметра  $\lambda_i$ . Обозначим  $\lambda \triangleq \sum \lambda_i$ .

$$P\{\text{Следующая возможность изменения наступит в } t + \Delta t | \text{текущее время} - t\} = 1 - \exp(-\lambda_i * \Delta t)$$

Моменты времени распределены в соответствии с экспоненциальным распределением.

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

Эта формула соответствует модели «first past the post». Все участники имеют стохастическое время ожидания. Первый получивший возможность произвести изменение делает свой выбор и все начинается с начала, но уже в новом состоянии.

Совокупность выборов  $i$  актора образуют собой поток Пуассона  $\delta_i$  с параметром  $\lambda_i$ .

Пример ( вставить рисунки в офисном контейнере они сохранены и корректно подписаны)

$\delta_1$  с параметром  $\lambda_1$

$\delta_2$  с параметром  $\lambda_2$

$\delta_3$  с параметром  $\lambda_3$

$\delta_4$  с параметром  $\lambda_4$

Образуют собой совокупный поток  $N(t)$  с интенсивностью  $\lambda$ .

Сумма всех потоков – есть так же поток Пуассона с параметром  $\lambda$ , где времена между моментами скачков независимы и распределены экспоненциально.

Вероятность того, что актер примет решение о изменении сети в выбранный промежуток времени:

$$P(X_{t+\Delta t} - X_t) = \lambda_i * \Delta t + o(\Delta t),$$

где  $o(\Delta t)$  остаточный член высшего порядка.

Вероятность того, что какой то актер в выбранный промежуток времени совокупного потока  $N(t_1)$ , где  $t \leq t_1 \leq t + \Delta t$  примет решение о изменении сети :

$$P(X_{t+\Delta t} - X_t) = \lambda * \Delta t + o(\Delta t).$$

У каждого актора есть  $n$  вариантов изменения сети, и всего  $n - 1$  вариантов выбора. 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 определяет вероятности того, что при следующей смене связей данный актер перейдет из состояния  $x^{(0)}$  в  $x$ .

Функция полезности:

$$f_i(x^{(0)}, x^{(\pm ij)}, \beta),$$

- где  $i$  – итератор актора,
- $x^{(0)}$  – текущее состояние сети, для графа и  $i \neq j$
- $x^{(\pm ij)}$  – граф, который идентичен по всем ребрам, кроме пары  $(i, j)$ , для которой существование связи является переключателем  $x^{(\pm ij)} = 1 - x_{ij}$ , причем возможен случай когда  $j = i$  это значит, что актер не изменит текущее состояние сети.

$f_i(x^{(0)}, x^{(\pm ij)}, \beta)$  определена на множестве всех пар сети  $x^{(0)}$  и  $x$ , таких, что  $x^{(0)} \neq x$ . В интерпретации полезности функция полезности может рассматриваться как чистая полезность, которую получает актер от перехода из состояния  $x^{(0)}$  в  $x$ .

Для определения этой вероятности используем следующие обозначения: Предположим, что в момент времени  $t + \Delta t$ , при текущей сети у актора появилась возможность изменения состояния сети в состояние задана

$$\frac{\exp(f_i(x^{(0)}, x^{(\pm ij)}, \beta))}{\sum_h \exp(f_i(x^{(0)}, x^{(\pm ih)}, \beta))}$$

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

$s_{ki}(x^{(0)}, x)$ , где  $k$  – индекс параметра модели. Существует несколько реализаций функции сети [6] [4]. В качестве примера приведу некоторые:

- Базовый компонент исходящих степеней:  $s_{1i} = \sum_j x_{ij}$  Он подобен постоянному члену в регрессионных моделях и всегда включается. Этот компонент балансирует между созданием и прерыванием связей. Учитывая предыдущее состояние  $x^{(0)}$ , следующее состояние либо имеет на одну связь больше, либо на одну связь меньше, или они идентичны.

Если имеет параметр то для создания связи имеет вклад в функцию полезности  $\beta_1$ , для разрушения связи  $-\beta_1$ . Таким образом, роль эффекта степени исходящих связей в модели заключается в вкладе в пользу создания связей по сравнению с их прерыванием. Обычно сети разрежены, поэтому возможностей для создания связей гораздо больше, чем для их прерывания.

- Взаимный выбор:  $s_{1i} = \sum_j x_{ij}x_{ji}$  Фундаментальный аспект почти всех направленных социальных сетей, потому что почти всегда существует некий обмен или другая взаимная зависимость. Это отражается в степени взаимности.

Функция полезности:

$$f_i(x^{(0)}, x, \beta) = \sum_{k=1}^K \beta_k s_{ki}(x^{(0)}, x).$$

Функция полезности описывает полезность для  $i$  актора перехода из состояния  $x^{(0)}$  в  $x$ .

Вероятность изменения связи  $i \leftrightarrow j$ :

$$P\{X(t + \Delta t) = x^{(\pm ij)} | X(t) = x\} = \frac{\exp(f_i(x^{(0)}, x^{(\pm ij)}, \beta))}{\sum_{h=1}^n \exp(f_i(x^{(0)}, x^{(\pm ih)}, \beta))}, \text{ где}$$

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

Учитывая факт того, что в рамках этой работы рассматривается ненаправленная сеть то создание связи может быть как односторонней инициативой, так и совместной, в отличии от разрушения связи. Разрушение связи происходит в одностороннем порядке. Рассмотрим разные варианты моделирования возникновения связи.

$p_{ij}(x, \beta)$  – есть вероятность возникновения связи в разных сценариях.

- Диктаторский сценарий:
  - Односторонняя инициатива:  $p_{ij}(x, \beta) = \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\sum_{h=1}^n \exp(f_i(x, x^{(\pm ih)}, \beta))}$
  - Двусторонняя инициатива:  $p_{ij}(x, \beta) = \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\exp(f_i(x, x, \beta)) + \exp(f_i(x, x^{(\pm ij)}, \beta))}$
- Взаимный сценарий:  $p_{ij} = \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\sum_{h=1}^n \exp(f_i(x, x^{(\pm ih)}, \beta))} \left( \frac{\exp(f_i(x, x^{(\pm ij)}, \beta))}{\exp(f_i(x, x, \beta)) + \exp(f_i(x, x^{(\pm ij)}, \beta))} \right)^{1-x_{ij}}$ 
  - $1 - x_{ij}$  Позволяет учесть факт существования или отсутствия связи в рассматриваемый момент времени.

$$P\{j \text{ примет предложение связи от } i\} = \frac{\exp(f_j(x, x^{(\pm ij)}, \beta))}{\exp(f_j(x, x, \beta)) + \exp(f_j(x, x^{(\pm ij)}, \beta))}$$

Скорость изменения модели с односторонней инициативой:

$$q_{ij} = \lambda_i p_{ij}(x, \beta) + \lambda_j p_{ij}(x, \beta)$$

Скорость изменения модели с двусторонней инициативой:

$$q_{ij} = \lambda_{ij} p_{ij}(x, \beta) + \lambda_{ji} p_{ij}(x, \beta)$$

+ где  $q_{ij}$  – скорость изменения пары.

Таким образом Акторно-ориентированная модель (CAOM) работает 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 вероятности  $\frac{\lambda_i(\alpha, x)}{\lambda(\alpha, x)}$ .
5. Выбираем случайным образом  $x \in A_i(x)$ , используя вероятности  $p_i(x, x^{(0)}, \beta)$ .  $A_i(x)$  - множество всех состояний в которое может измениться сеть  $x$  в следствии решения актора  $i$ .
6. Задать  $t = t + \Delta T$ .
7. Задать  $x^{(0)} = x$ .
8. Вернуться к шагу (2).

На этапе моделирования принятия решения получается множество параметров. Дальнейшей задачей является оценка этих параметров.

### 2.3.3 Оценка параметров:

Для оценки параметров используется имитационное моделирование.

Реальная система заменяется моделью, и производится множество повторений алгоритма. Когда в среднем модель будет хорошо согласовываться с данными процесс оценки будет остановлен, и мы будем считать, что параметры оценены.

Оценка параметров может быть произведена несколькими способами. В рамках пакета RSiena, как следствие, и в моей курсовой работе используется метод моментов [5].

Обозначим оцениваемый параметр как  $\theta = (\rho, \alpha, \beta)$ . Для каждого существует статистика, чувствительная к этому параметру. Для  $\rho_m$ , влияющего на общее количество изменений используется статистика Hamming distance,

$$D(X(t_{m+1}), X(t_m)) = \sum_{ij} X_{i,j}(t_m + 1) - X_{i,j}(t_m)$$

Для параметра  $\alpha_k$ , обозначающего насколько сильно скорость изменения актора  $i$  зависит от  $u_{i,k}(X)$ , существует статистика – ковариата или позиционная характеристика актора, такая как исходящая степень  $\sum_j x_{ij}$ . Имеется в виду что скорость изменения может зависеть от пола, возраста, и тд.

$$A(X(t_{m+1}), X(t_m)) = \sum_{ij} u_{i,k}(X(t_m)) + X_{i,j}(t_m + 1) - X_{i,j}(t_m).$$

Для функции полезности, где  $s_{i,k}(x^{(0)}, x)$  не зависит от  $x^{(0)}$ , большие значения  $\beta_k$  будут стремиться к сетям  $x$  для которых значение  $s_{ik}(x)$  больше для любого актора  $i$ . Для оценки статистика имеет вид

$$s_k(X(t_m)) = \sum_i s_{ik}(X(t_m)).$$

Комбинируя статистики и используя предположение о марковской цепи для наблюдаемых данных оценочные уравнения имеют вид.

$$D(X(t_{m+1}), X(t_m)) = E_{\theta}\{D(X(t_{m+1}), x(t_m) \vee X(t_m))\}$$

$$(m = 1, \dots, M - 1)$$

$$\sum_{m=1}^{M-1} A_k$$

$$(k = 1, \dots, K_{\alpha})$$

$$\sum_{m=1}^{M-1} s_k$$

$(k = 1, \dots, K_\beta)$

где  $K_\alpha$  и  $K_\beta$  номера элементов  $\alpha$  и  $\beta$ .

Для решения этих уравнений используется стохастическая оптимизация на основе алгоритма Роббинса-Монро. Алгоритм применяет многомерную версию алгоритма Роббинса-Монро с улучшениями, предложенными Поляком и Руппертом. Техника "двойного усреднения" также применяется для улучшения результатов. Алгоритм реализован в пакете RSiena на языке R.

Он состоит из трех фаз:

1. Определение чувствительности ожидаемых статистик к параметрам.
2. Обновляет параметры с использованием симуляций динамики сети.
3. Используется для оценки приближения полученных уравнений и вычисления стандартных ошибок.

Для вычисления производных ожидаемых значений по отношению к параметрам используется метод функции оценки. Этот алгоритм является надежным, но затратным по времени.

## Глава 3 Анализ результатов

Формирование начальных гипотез, а так же формирование социологических выводов – это задача прикладных специалистов.

В рамках моей курсовой работы я использую некоторые предположения выдвинутые в подобном исследовании для национальной научной системы Словении [3].

### 3.1 Формирование начальных социологических гипотез.

Согласно исследованию Anuška Ferligoj, Luka Kronegger, Franc Mali, Tom A B Snijders, Patrick Doreian в журнале Scientometrics (2015, №104, с. 989–990): «В целом, мы предлагаем следующие механизмы влияния на научное соавторство: встраиваемость в сеть: соавторы соавторов стремятся стать соавторами соавторов; преимущественная привязанность: авторы предпочитают ищут соавторов, у которых уже много соавторов; институциональная встроенность (принадлежность к одной исследовательской группе и одной научной дисциплине, возрастное сходство также может подпадать под эту категорию, поскольку означает принадлежность к общей когорте ученых, которые взаимодействуют друг с другом больше, чем представители разных когорт) и контрольные переменные, в частности стаж.»

### 3.2 Оценка сформированных гипотез.

Коэффициент сходимости модели достаточно низок ( Overall maximum convergence ratio: 0.1611 ) , для того, чтобы считать модель сходящейся[7].

Estimates, standard errors and convergence t-ratios

|                  | Estimate | Standard |    |
|------------------|----------|----------|----|
| Convergence      |          | Error    | t- |
| ratio            |          |          |    |
| Rate parameters: |          |          |    |



|   |  |         |            |
|---|--|---------|------------|
| 0.1                                       | Rate parameter period 1                        | 0.2317  | ( 0.0053 ) |
| 0.2                                       | Rate parameter period 2                        | 0.1678  | ( 0.0091 ) |
| Other parameters:                         |  |         |            |
| 1.  | rate effect ln(degree+1) on rate co_authorship | 1.6070  | ( 0.0224 ) |
| 0.0157                                    |  |         |            |
| 2.  | eval degree (density)                          | -3.2147 | ( 0.0123 ) |
| -0.0174                                   |  |         |            |
| 3.  | eval transitive triads                         | 0.4268  | ( 0.0140 ) |
| -0.0005                                   |  |         |            |
| 4.  | eval trans. triads intermed. exp_covar         | -0.0018 | ( 0.0006 ) |
| 0.0112                                    |  |         |            |
| Overall maximum convergence ratio: 0.0508 |  |         |            |

### 3.2.1 Интерпретация анализов

Необходимо учитывать, что функция эволюции сети задана формулой:

$$f_i^{net}(x) = \sum \beta_k^{net} s_{ik}^{net}(x) k + \epsilon, \text{ где } \epsilon \sim N(0, 1.28)$$

$s_{1k}^{beh}(x, z) = z_i s_{ik}^0(x, z)$ , где  $s_{ik}^0(x, z)$  не зависит от  $z_i$  причем может зависеть от  $z_j$  для другого актора  $j$ .

Она отражает полезность для  $i$  актора. Используется для сравнения полезности различных изменений. С другой стороны когда  $i$  актер делает изменение из  $x^{(0)}$  в  $x$ , и  $x_a x_b$  два возможных результата шага то отношение вероятностей этих вариантов есть:  $\exp(f_i^{net}(x_a) - f_i^{net}(x_b))$ . Следовательно сетевые параметры стоит расценивать как логарифмические.

Для оценки зависимых параметров необходимо учитывать, что возможные изменения в переменной  $+1, 0, -1$ , и оценки не должны выходить за пределы допустимого диапазона. Параметры целевой функции это вклады в логарифмические вероятности увеличения зависимой переменной на

1 единицу при увеличении эффекта на 1 единицу. Где зависимая переменная либо  $X_{ij}$ , либо  $s_{ik}$ .

Пример:

Рассчитан эффект average alter он равен 1, 1561. Это означает, что при сравнении двух акторов равных во всех отношениях, кроме того, что связи первого в среднем на 1 выше по шкале рассматриваемой ковариаты (Индекс Хирша, и тд), чем у второго, шанс увеличения значения ковариаты по сравнению с отсутствием изменений ( в рамках одного шага относительно ковариаты) выше в  $\exp(1, 1561) = 3.17751676428$  раза, чем для второго.

Не любой эффект можно трактовать как в приведенном примере.

Пример:

Для эффекта out-degree (количество связей) для актора  $i$   $x_{i+}$ . Пусть число связей, чьи значения  $z_j$  меньше, равны или больше значения  $z_i$  для  $i$ , обозначим через  $a$ ,  $b$  и  $c$ . Обозначим диапазон (максимальное минус минимальное значение) через  $r$ . Тогда, в случае минимального шага в отношении поведения, вклады суммарного эффекта сходства в лог-вероятности изменений  $-1, 0, +1$  определяются следующим образом  $\beta_k^{beh}(a - b - c)/r$  и

$\beta_k^{beh}(c - a - b)/r$ . Вклады для среднего

эффекта сходства составляют  $\beta_k^{beh}(a - b - c)/(rx_{i+})$  и  $\beta_k^{beh}(c - a - b)/(rx_{i+})$ . Это показывает, что

влияние связей на эффекты сходства зависит только от того, имеют ли они

большие или меньшие значения, чем у рассматриваемого актора, а не от того, насколько эти значения больше. Этот также показывает, что для эффектов сходства важна дисперсия ценностей связей, а не

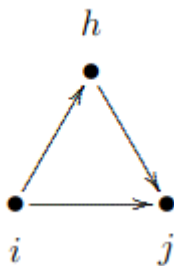
а не только их среднее значение, в то время как для эффекта среднего изменения имеет значение только среднее значение.

Не все эффекты могут быть описаны как изменение некоторой функции полезности. В таких случаях используются элементарные эффекты [8].

Элементарный эффект - это вклад в создание или поддержание связи, определенный напрямую, то есть без выражения его на основе изменения какой-либо функции оценки. Это означает, что элементарные эффекты более общие, чем эффекты оценки, и все эффекты могут быть представлены как элементарные эффекты. Однако для удобства интерпретации предпочтительно использовать формулировку функции оценки.

Элементарные эффекты могут применяться аналогично к созданию и поддержанию связи; или они могут применяться исключительно к созданию связи или исключительно к поддержанию связи.

Основной пример ситуации, когда невозможно описать как изменение функции полезности - тенденция к созданию связные троек (транзитивные замыкания). Если актор  $i$  является центральным, то связи, приводящие к замыканию, представлены как  $i \rightarrow j$  и  $i \rightarrow h$ . Однако первая связь подразумевает создание пути  $i \rightarrow h \rightarrow j$ , в то время как вторая связь означает установление связи с актором  $h$ . Актор  $h$ , в свою очередь, делает тот же исходящий выбор для третьего актора  $j$ , что указывает на структурную эквивалентность; таким образом, это различные процессы.



### density

Оценка параметра этого эффекта указывает на то, что с увеличением исходящих степеней снижается вероятность появления новой связи.

### outRateLog

Логарифмическая зависимость скорости (параметра  $\lambda_i$  для актора  $i$ ) от количества исходящих степеней. С увеличением числа связей актор увеличивает частоту принятия решений.

## transitive triads

Оценка этого эффекта показывает количество транзитивных паттернов относительно актора  $i$ . Для двух акторов, где каждый из них является соавтором с  $i$  индивидуально, вероятность стать соавторами увеличивается в  $\exp(0.4268) = 1.532346162$  раза.

## transTripX exp\_covar

Значение оценки параметра этого эффекта означает, что с увеличением значения эффекта transitive triads этот эффект показывает, что с увеличением стажа ослабевает эффект transitive triads.

# Заключение

---

В рамках курсовой работы я ознакомился с дисциплиной сетевой анализ и ознакомился с современным методом анализа динамики социальных сетей - Стохастическое акторно ориентированное моделирование (Stochastic Actor-Oriented Models for Network Dynamics). Полученные результаты исследования могут оказать практическую значимость и служить основой для дальнейших исследований в данной области.

# Список 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. 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 .
3. 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.
4. Snijders, T.A.B., & Pickup, M. (2016). Stochastic Actor-Oriented Models for Network Dynamics. Retrieved June 10, 2016.
5. Snijders, T. A. B. The Statistical Evaluation of Social Network Dynamics // 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. 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).
8. Snijders The statistical evaluation of social network dynamics / Snijders, A T. — Текст : непосредственный // Sociological Methodology. — 2001. — № 31. — С. 361-395.