

Робастное, адаптивное и сетевое управление

© 2024 г. Н.Г. МЕЩЕРЯКОВА (natamesc@gmail.com),
С.В. ШВЫДУН, канд. физ.-мат. наук (shvydun@hse.ru)
(Национальный исследовательский университет
«Высшая школа экономики», Москва)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕР ЦЕНТРАЛЬНОСТИ В СЕТЕВЫХ СТРУКТУРАХ¹

Рассматривается задача нахождения центральных элементов в сетевых структурах. Поскольку концепция центральности не имеет четкого определения, в литературе известно большое число моделей определения наиболее важных элементов сети. Приведен обзор существующих подходов для сравнения мер центральности, осуществлен сравнительный анализ чувствительности известных индексов центральности к различным изменениям в структуре сети.

Ключевые слова: сетевой анализ, индексы центральности, аксиомы центральности, устойчивость.

DOI: 10.31857/S0005231024080039, **EDN:** WPLVJA

1. Введение

Методы сетевого анализа активно используются при моделировании и изучении сложных систем. В частности, с помощью сетевого анализа можно описать большинство существующих технологических, биологических, социальных и других систем, а также смоделировать различные социально-экономические, транспортные, эпидемические, производственные и прочие процессы.

Одной из важных задач в области анализа сложных сетевых структур является выявление ключевых или наиболее важных участников [1]. К сожалению, концепция важности в сети не имеет четкого определения и зависит от поставленной перед исследователем задачи, а также протекающих в сети процессов. В литературе известно большое число различных моделей определения центральных элементов, которые основаны на подсчете числа взаимосвязей у каждого элемента, количестве путей между различными элементами, учете индивидуальных атрибутов элементов сети и возможности группового влияния вершин. Поскольку общее число моделей центральности

¹ Данная работа осуществлена в рамках гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – кандидатов наук (грант МК-3867.2022.1.6. “Исследование свойств индексов влияния в сетевых структурах”).

исчисляется сотнями [2], а результаты от их применения могут быть совершенно различными, возникает необходимость в оценке качества полученных результатов и сравнении данных моделей.

Сравнение индексов центральности между собой представляет собой сложную задачу из-за субъективного характера определения важности элементов сети. Так, в ряде исследований полученные результаты применения индексов центральности сопоставляются с заведомо известной информацией о важности вершин или некоторой экспертной информацией [3, 4]. В рамках конкретной задачи и при наличии заранее известного распределения важности среди элементов сети можно выбрать наиболее близкий индекс центральности для его дальнейшего использования в данной сети и ее модификациях (например, для выявления ключевых элементов в динамике, когда истинное распределение важности уже неизвестно или известно частично). К сожалению, эталонные данные об истинной важности вершин крайне редко встречаются в реальных задачах.

Существует также другой подход к оценке качества мер центральности в сетевых структурах. Ряд исследований предлагает сравнивать классические меры центральности с точки зрения выполнения или невыполнения различных свойств, известных в литературе как *аксиомы центральности* [5–13]. Данные свойства дают понимание о работе рассматриваемых индексов и о том, будет ли выявлен тот или иной ключевой элемент, если граф имеет определенную структуру. Данный подход помогает оценить, насколько хорошо та или иная мера центральности удовлетворяет ряду логичных свойств. Тем не менее текущий список свойств не является единым и исчерпывающим, что приводит к тому, что разные исследования посвящены изучению того или иного набора свойств.

Наконец, одним из способов сравнения моделей центральности является анализ их устойчивости к изменению структуры сети. Поскольку множество существующих реальных сетей зачастую содержат неполную или некорректную информацию о связях между элементами сети (например, ошибки или пропуски в исходных данных), результат применения различных индексов центральности может быть искажен. Таким образом, возникает вопрос, насколько устойчиво и репрезентативно текущее ранжирование элементов сети, которое было получено в результате применения той или иной модели центральности. На данный момент в литературе известен ряд работ, в которых исследуется чувствительность различных моделей центральности [14–19]. Тем не менее большинство данных исследований ограничивается анализом только классических индексов центральности на малом наборе сетей.

В рамках данного исследования проводится обзор существующей аксиоматики, характеризующей понятие центральности в сетевых структурах. Кроме того, в работе осуществлен анализ чувствительности некоторых классических и других мер центральности к малым изменениям в сети. Результаты данного

направления позволят получить более глубокое понимание того, как сильно изменения во взаимосвязях между элементами сети влияют на результат применения известных моделей выявления наиболее важных элементов.

Работа поделена на четыре части. В первой части приводится описание известных классических моделей центральности. Во второй части приведен обзор существующих исследований по изучению аксиоматических свойств моделей центральности. Третья часть посвящена исследованию устойчивости мер центральности к изменениям в сети на примере классических и реальных сетевых структур. В четвертой части приводится заключение по полученным результатам.

2. Модели центральности в сетевых структурах

Рассмотрим граф $G = (V, E)$, где $V = \{1, \dots, n\}$ – множество вершин, $|V| = n$, $E \subseteq V \times V$ – множество ребер между вершинами графа. Граф G называется *ненаправленным* (неориентированным), если $\forall i, j \in V: (i, j) \in E \implies (j, i) \in E$, и *направленным* (ориентированным) в противном случае. Граф G может быть также представлен с помощью матрицы смежности $A = [a_{ij}]$, где $a_{ij} = 1$, если $(i, j) \in E$, и $a_{ij} = 0$ в ином случае. Граф называется *взвешенным*, если он описывается с помощью матрицы весов $W = [w_{ij}]$, где значение w_{ij} отражает вес ребра $(i, j) \in E$. Каждая вершина может иметь набор атрибутов $\{w_i^k\}$, где $i \in V$ – вершина графа, k – номер атрибута, $k \in K$, а также пороговое значение q_i , при достижении которого вершина i подвержена влиянию со стороны других вершин. Под *центральностью* вершины i подразумевается некоторое числовое значение $c_i \in \mathbb{R}^+$, характеризующее степень ее важности в графе G .

Центральность является одним из ключевых показателей, используемых при исследовании сложных сетевых структур [1–3]. В литературе известен ряд мер, которые помогают выявить наиболее центральные элементы в той или иной сетевой структуре. Описание наиболее известных мер приведено ниже.

2.1. Степенные центральности

Простейшей центральностью i -й вершины ненаправленного графа является ее степень (degree centrality), т.е. количество ребер, инцидентных i -й вершине [20]:

$$(1) \quad c_i = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji}.$$

В случае направленного графа существует две отдельные меры центральности по степени: входящая степень и исходящая степень. Входящая степень указывает число связей, направленных к вершине, а исходящая – число связей, направленных от вершины. Степенная центральность может быть также

адаптирована к взвешенным сетям и рассчитана с учетом веса w_{ij} каждого ребра $(i, j) \in E$.

2.2. Спектральные центральности

Степенные центральности являются локальными мерами, которые никак не учитывают важность соседних вершин. Обобщением степенной центральности является так называемая центральность по собственному вектору [20], которая учитывает не только прямые, но и не прямые взаимосвязи между вершинами. Центральность вершин по собственному вектору (eigenvector centrality) \vec{c} присваивает относительную важность каждой вершине следующим образом: более важные соседи вершины дают больший вклад в ее собственную центральность по сравнению с менее важными соседями. Вычисление данной модели центральности сводится к нахождению собственного вектора \vec{c} исходной матрицы смежности A :

$$(2) \quad A\vec{c} = \lambda\vec{c},$$

где λ – наибольшее собственное значение матрицы A . На практике центральность по собственному вектору используется при анализе ненаправленных графов, поскольку вычисление собственного вектора на направленных графах является более сложной задачей, которая может привести к появлению нулевых или комплексных собственных значений.

Другим примером данного класса центральностей является центральность по PageRank [21], которая заключается в определении вероятности посещения каждой вершины в рамках случайных блужданий в графе:

$$(3) \quad c_i = \alpha \sum_{j=1}^n \frac{c_j}{\sum_{k=1}^n a_{jk}} a_{ji} + \frac{1 - \alpha}{n},$$

где α – вероятность продолжения случайного блуждания в графе, $0 \leq \alpha \leq 1$. Отметим, что в отличие от центральности по собственному вектору данный индекс может использоваться также и для направленных графов.

Существуют также и другие модели центральности, основанные на подсчете собственного вектора: центральность Каца (Katz centrality) [22], алгоритм HITS (hubs and authorities) [23], подграфовая центральность (subgraph centrality) [24] и т.д.

2.3. Центральности по нахождению кратчайших путей в графе

Данный класс центральностей основан на подсчете кратчайших путей между вершинами. Наиболее известным примером является центральность по близости (closeness centrality) [20, 25], которая учитывает, насколько близко к вершине i расположена к остальным вершинам сети:

$$(4) \quad c_i = \frac{1}{\sum_{j=1}^n d_{ij}},$$

где d_{ij} – длина кратчайшего пути между вершинами i и j .

Другим примером центральности, основанной на нахождении кратчайших путей между вершинами, является центральность по посредничеству (betweenness centrality) [26, 27]. Центральность по посредничеству характеризует, насколько важную роль играет вершина i как посредник между всевозможными парами других вершин:

$$(5) \quad c_i = \sum_{j \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}},$$

где σ_{jk} показывает число кратчайших путей между вершинами j и k , а $\sigma_{jk}(i)$ – число кратчайших путей между вершинами j и k , проходящих через вершину i .

2.4. Центральности по ближним и дальним взаимодействиям

Большинство приведенных выше классических мер центральности не позволяет учитывать важные характеристики – индивидуальные атрибуты каждой вершины и возможность их группового взаимодействия. Более того, существующие методы не в полной мере оценивают интенсивность дальних взаимодействий и зачастую учитывают незначимые связи между элементами сети. В связи с этим в [3, 28] предложен индекс дальних взаимодействий (LRIC), который позволяет учитывать данные особенности сложных систем.

Обозначим через N_j набор вершин, имеющих прямую связь с вершиной j . Тогда группа вершин $\Omega(j) \subseteq N_j$ является *критической* для вершины j , если $\sum_{i \in \Omega(j)} w_{ij} \geq q_j$, а вершина $l \in \Omega(j)$ – *ключевой*, если ее исключение из данной группы делает группу неключевой. Набор ключевых вершин в группе $\Omega(j)$ обозначим через $\Omega_p(j)$. Отличительной особенностью индекса дальних взаимодействий является определение влияния вершины i на вершину j через поиск такой критической группы $\Omega_p(j)$, в которой вершина i имеет максимальное влияние, т.е.

$$(6) \quad c_{ij} = \begin{cases} \max_{\Omega_p(j): i \in \Omega_p(j)} \frac{w_{ij}}{\sum_{k \in \Omega_p(j)} w_{kj}}, & \text{если } \exists \Omega_p(j): i \in \Omega_p(j), \\ 0 & \text{в ином случае.} \end{cases}$$

Значение c_{ij} показывает, как вершина i влияет на вершину j напрямую. Найденные значения можно представить в виде матрицы прямых влияний $C = [c_{ij}]_{n \times n}$, которую далее можно использовать для расчета непрямых влияний между вершинами. Модель LRIC имеет ряд модификаций. В частности, известны модели ближних взаимодействий (SRIC) [29], а также индекс группового влияния [30]. Стоит отметить, что индекс LRIC показал свою эффективность при анализе сетей международного рынка капитала, международной миграции, общей торговли, торговли оружием, торговли продовольствием, цитирования экономических журналов, а также сети взаимодействий между террористическими группами [3, 31, 32].

3. Свойства индексов центральности: обзор литературы

Ряд исследований предлагает сравнивать меры центральности с точки зрения выполнения или невыполнения различных свойств, определяющих разумность полученного набора наиболее важных элементов [5–11]. В данном разделе рассматриваются наиболее известные аксиомы центральности, а также приводится список индексов центральности, для которых были исследованы данные аксиомы.

3.1. Аксиома анонимности (*Anonymity axiom*)

Аксиома анонимности была предложена в [5]. Согласно данной аксиоме центральность вершины i зависит только от ее расположения в сети, т.е.

$$(7) \quad c_i(V, E) = c_{f(i)}(V, \{(f(i), f(j)) : (i, j) \in E\}),$$

где $c_i(V, E)$ – центральность вершины в графе $G = (V, E)$, а f – биекция на множестве вершин, $f : V \rightarrow V$. Другими словами, центральность вершины совпадает с центральностью соответствующей вершины в изоморфном графе. Отметим, что все индексы центральности, не использующие атрибуты вершин, удовлетворяют анонимности.

3.2. Аксиома роста конечной точки (*Endpoint increase axiom*)

Аксиома роста конечной точки [5] выполняется, если добавление ребра между двумя вершинами не приводит к уменьшению их центральности, т.е. $\forall G, \forall i, j \in V$

$$(8) \quad \begin{cases} c_i(V, E \cup (i, j)) \geq c_i(V, E), \\ c_j(V, E \cup (i, j)) \geq c_j(V, E), \end{cases}$$

где $c_i(V, E \cup (i, j))$ – центральность вершины i в графе G' , который образуется при добавлении ребра (i, j) к графу $G = (V, E)$. Таким образом, у любой пары несвязных вершин имеется стимул к появлению связи между ними.

3.3. Аксиома монотонности (*Monotonicity axiom*)

Аксиома монотонности [5] выполняется, если добавление в граф ребра (i, j) не уменьшает значение центральности никакой вершины графа, т.е. $\forall G, \forall i, j, k \in V$

$$(9) \quad c_k(V, E \cup (i, j)) \geq c_k(V, E).$$

Если модель центральности предполагает нормирование значений индексов центральности, то данная аксиома практически во всех случаях не будет выполняться, поскольку увеличение центральности одной вершины приводит к уменьшению относительной центральности других вершин. Кроме того, если аксиома роста конечной точки не выполняется, то аксиома монотонности также не выполняется для рассматриваемой модели центральности.

3.4. Аксиома центральной вершины (Top node axiom)

Согласно аксиоме центральной вершины [5] вершина i , имеющая наибольшее значение центральности в исходном графе G , должна иметь также наибольшее значение центральности в графе G' , который образуется при добавлении инцидентного ребра к данной вершине в графе G :

$$(10) \quad \begin{aligned} c_i(V, E) \geq c_k(V, E) \quad \forall k \in V &\implies \\ \implies \forall j, k \in V \quad c_i(V, E \cup (i, j)) \geq c_k(V, E \cup (i, j)). \end{aligned}$$

Аксиома центральной вершины не выполняется, если добавление инцидентных связей уменьшает центральность наиболее влиятельной вершины, что противоречит интуиции понятия центральности в сети.

3.5. Аксиома справедливости (Fairness axiom)

Аксиома справедливости [6] выполняется, если добавление ребра (i, j) изменяет центральность инцидентных вершин на одинаковую величину, т.е. $\forall G, \forall i, j \in V$

$$(11) \quad c_i(V, E \cup (i, j)) - c_i(V, E) = c_j(V, E \cup (i, j)) - c_j(V, E).$$

Таким образом, вклад новой связи должен быть одинаковым для каждой из инцидентных вершин.

3.6. Аксиома сбалансированного вклада (Balanced contributions axiom)

Аксиома сбалансированного вклада приведена в [7] и аксиоматизирована в [9]. Модель центральности удовлетворяет данной аксиоме, если удаление всех ребер E_i , инцидентных вершине i , влияет на центральность вершины j в той же мере, как и удаление всех инцидентных ребер E_j на вершину i , т.е. $\forall i, j \in V$

$$(12) \quad c_i(V, E) - c_i(V, E \setminus E_j) = c_j(V, E) - c_j(V, E \setminus E_i),$$

где E_i – множество всех ребер, инцидентных вершине i .

3.7. Аксиома добавления нового ребра (Add edge distance axiom)

Аксиома добавления нового ребра предложена в [9]. Согласно данной аксиоме добавление нового ребра между парой вершин i и j , находящихся на одинаковом расстоянии от вершины k , не влияет на значение центральности данной вершины, т.е. $\forall i, j, k \in V: d_{ik} = d_{jk}$

$$(13) \quad c_k(V, E) = c_k(V, E \cup (i, j)).$$

Таким образом, добавление ребра между двумя вершинами не должно влиять на центральность вершин, расположенных от них на одинаковом расстоянии.

3.8. Аксиома моста (*Bridge axiom*)

Согласно аксиоме моста [9], если ребро (i, j) является единственной связью между двумя несвязанными компонентами связности (т.е. является мостом), то центральность вершины i должна быть выше центральности вершины j , если компонента связности, имеющая связь с вершиной i , больше второй компоненты связности.

3.9. Аксиома аддитивности центральности общей вершины (*Cut-vertex additivity*)

Согласно аксиоме аддитивности центральности общей вершины [10], если объединить два графа G и G' , имеющих только одну совпадающую вершину i , то центральность этой вершины в объединенном графе определяется как сумма ее центральностей в графах G и G' , т.е. $\forall G = (V, E), G' = (V', E') : V \cap V' = \{i\}$

$$(14) \quad c_i(V \cup V', E \cup E') = C_i(V, E) + C_i(V', E').$$

Отметим, что данная аксиома имеет альтернативную формулировку:

$$(15) \quad \frac{1}{c_i(V \cup V', E \cup E')} = \frac{1}{c_i(V, E)} + \frac{1}{c_i(V', E')}.$$

Ряд аксиом центральности сформулирован в контексте рассмотрения конкретных графовых структур.

3.10. Аксиома плотности (*Density axiom*)

Аксиома плотности предложена в [8]. Рассмотрим граф $G_{k,p}$, который представляет собой полный граф из k вершин, соединенный одним ребром (i, j) с циклом длины p ($p \geq 3$), где вершина i – вершина полного графа, а j – вершина цикла. Аксиома плотности выполняется, если при $k = p$ центральность вершины i будет выше центральности вершины j .

3.11. Аксиома размера (*Size axiom*)

Аксиома размера рассмотрена в [8]. Рассмотрим граф $G_{k,p}$, который представляет собой полный граф из k вершин, соединенный одним ребром (i, j) с циклом длины p ($p \geq 3$). Аксиома размера выполняется, если при любом k найдется такое значение $p \geq k$, что центральность вершин в цикле будет выше центральности вершин в полном графе, и, наоборот, для любого k найдется такое значение $p < k$, что центральность вершин в цикле будет ниже центральности вершин в полном графе.

Ряд аксиом уже был исследован для некоторых моделей центральности (см. табл. 1) [33]. Известные в литературе результаты по анализу свойств индексов центральности, упомянутых в разделе 2, приведены в [8, 10, 11, 34].

Таблица 1. Свойства классических индексов центральности («+» – аксиома выполняется, «-» – аксиома не выполняется, «?» – аксиома не была исследована в литературе)

№	Модели центральности	Анонимность	Рост конечной точки	Монотонность	Центральная вершина	Справедливость	Сбалансированный вклад	Добавление нового ребра	Мост	Аудитивность	Плотность	Размер
1	Степенные центральности	+	+	+	+	+	+	+	-	+	+	-
2	Центральность по собственному вектору	+	-	-	-	-	-	-	-	-	+	-
3	PageRank	+	?	-	?	-	-	-	-	-	+	-
4	Центральность Каца	+	+	+	?	-	-	-	-	-	+	-
5	HITS	+	?	?	?	-	-	?	?	?	?	?
6	Подграфовая центральность	+	?	?	?	-	-	?	?	?	?	?
7	Центральность по близости	+	+	+	-	-	-	+	+	-	+	-
8	Центральность по посредничеству	+	-	-	-	-	-	-	-	-	-	-
9	LRIC	-	?	-	?	-	-	?	?	?	?	?

Заметим, что на данный момент даже классические индексы центральности исследованы лишь частично.

Согласно табл. 1 большинству приведенных аксиом удовлетворяет только степенная центральность. Данную особенность можно объяснить тем фактом, что степенные центральности являются локальными и основаны на расчете только прямых связей между вершинами. Тем не менее большинство аксиом сформулированы таким образом, чтобы учитывать глобальную структуру сети и оценивать эффект от непрямых связей между вершинами. В связи с этим, поскольку степенные центральности не изменяют свою центральность от изменения структуры взаимосвязей других вершин, то данная мера удовлетворяет большинству аксиоматических свойств. Что касается сравнения других моделей центральности, то большинство мер не удовлетворяет приведенным аксиомам. Все рассматриваемые модели центральности, не учитывающие индивидуальные атрибуты вершин, удовлетворяют аксиоме анонимности. Аксиомы роста конечной точки и монотонности выполняются только для центральности по близости и центральности Катца. Аксиома плотности выполняется для центральности по собственному вектору, центральности Катца, PageRank и центральности по близости.

Рассматриваемый список аксиом не является исчерпывающим, а в ряде других исследований можно встретить другую аксиоматику [35–38]. Кроме того, можно предложить аксиомы из других смежных научных областей. На-

пример, известна *аксиома согласия* для функций выбора [39], согласно которой альтернатива, выбираемая на двух подмножествах, должна быть также выбрана на их объединении. В теории выбора данное условие характеризует рациональный выбор на множестве альтернатив. В связи с этим можно расширить аксиому аддитивности центральности общей вершины и предположить, что если у двух графов G и G' есть общая вершина i , которая имеет наибольшую центральность в обоих графах, то данная вершина должна также иметь наибольшую центральность и на их объединении.

Стоит отметить, что невыполнение для конкретной модели каких-либо свойств не ведет к тому, что мера является некачественной, а лишь только позволяет выявить и лучше понять основные особенности модели центральности.

4. Исследование устойчивости индексов центральности

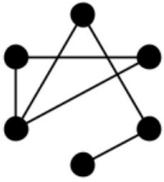
При анализе реальных систем одним из важнейших критериев выбора модели центральности является степень устойчивости полученных результатов. К сожалению, ряд существующих наборов данных является неполным или может содержать некорректную информацию. Например, в [40, 41] рассматриваются криминальные сети, которые являются неполными (в силу специфики предметной области), могут содержать ошибки (ошибки при сборе данных либо намеренно искаженная информация) и противоречащую информацию (сбор данных осуществлен из разных источников). В [31] произведен анализ сетевых взаимодействий на международном рынке капитала, при этом исходные данные содержат порядка 94% от общего числа информации, поскольку ряд стран не предоставляют отчетность. В [42] рассматривалась проблема асимметрии в данных международной торговли, связанной с различными классификациями номенклатуры, оценке объема торговли (транспортные и страховые расходы), а также задержкой по времени. Таким образом, возникает вопрос, насколько сильно результаты применения моделей центральности подвержены небольшим изменениям в структуре сети.

4.1. Устойчивость индексов центральности на классических структурах

В данном разделе проводится оценка устойчивости известных моделей центральности к добавлению или удалению ребер в ненаправленных невзвешенных графах. Для оценки чувствительности известных моделей центральности произведем серию вычислительных экспериментов на классических графовых структурах, состоящих из $n = 100$ вершин (см. рис. 1):

1. Случайный граф Эрдёша–Реньи (Erdos–Renyi graph): случайный граф с n вершинами, в котором связи образуются с вероятностью p .

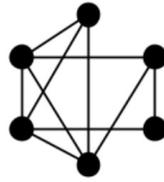
2. Безмасштабный граф (Barabasi–Albert graph): случайный граф с n вершинами, который образуется последовательным присоединением новых вершин, при этом на каждом этапе образуется m связей с вероятностью, пропорциональной текущей степени вершин графа.



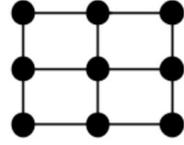
(1) Случайный граф



(2) Безмасштабный граф



(3) Сеть малого мира



(4) Граф квадратной решетки

Рис. 1. Классические сетевые структуры.

3. Сеть малого мира (Small-world Watts–Strogatz graph model): случайный граф, в рамках которого создается регулярный граф с n вершинами, а затем каждое ребро переписывается с вероятностью p на другой произвольный узел.

4. Граф квадратной решетки (Square lattice) с n вершинами, где вершины соответствуют точкам на координатной плоскости, которые впоследствии соединены ребром, если вершины находятся на расстоянии 1.

Вычислительный эксперимент осуществлен следующим образом. На первом этапе происходит построение $N = 100$ графов заданной структуры (для графа квадратной решетки $N = 1$). В рамках исследования был установлен параметр $p = 0,05$ для случайных графов и сетей малого мира, а также $m = 4$ для безмасштабных графов². На втором этапе происходит изменение структуры каждого рассматриваемого графа $K = 100$ способами (для графа квадратной решетки $K = 10\,000$). Далее производится оценка центральности вершин в исходном и измененных графах, после чего между индексами центральности осуществляется расчет коэффициента корреляции Кендалла, учитывающего относительную важность вершин. Итоговое значение устойчивости индекса центральности определяется как среднее арифметическое среди рассчитанных характеристик по $N \times K = 10\,000$ экспериментам. Отметим, что увеличение числа экспериментов не приводит к значительному изменению результатов.

Вычислительный эксперимент осуществлен для следующих моделей центральности: степенная центральность (degree), центральность по собственному вектору (eigenvector), PageRank, центральность по близости (closeness), центральность по посредничеству (betweenness), подграфовая центральность (subgraph), а также индекс дальних взаимодействий (LRIC). Полученные результаты приведены на рис. 2.

Как показано на рис. 2, наибольшую устойчивость индексы центральности демонстрируют на случайных и безмасштабных графах: после удаления/добавления 20 ребер в графе корреляция индексов центральности остается высокой (выше 0,8). Наоборот, небольшие изменения в структуре сети малого мира и графа квадратной решетки приводят к значительным измене-

² Авторы также рассматривали другие значения параметров p (от 0,01 до 0,1 с шагом 0,01) и m (от 2 до 5). Было выявлено, что относительное расположение индексов центральности с точки зрения их устойчивости при других значениях параметров меняется незначительно.

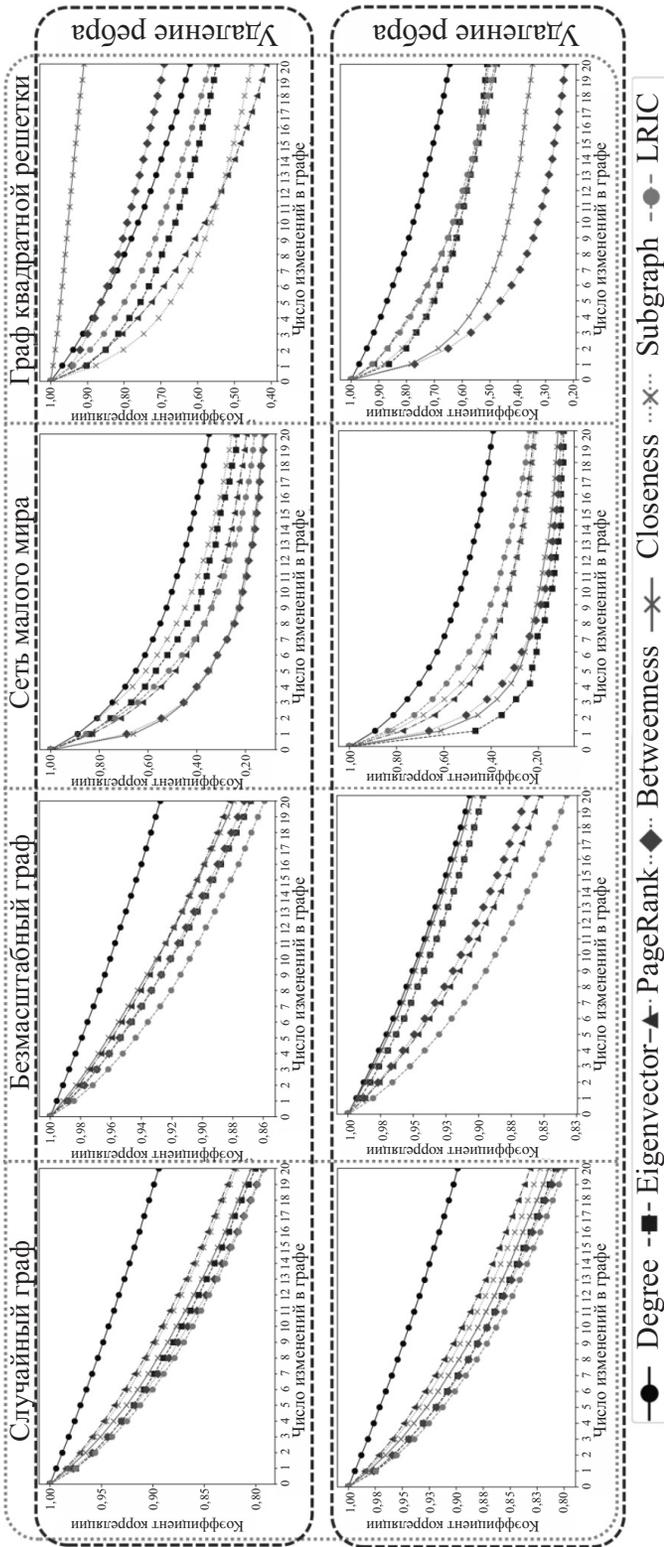


Рис. 2. Устойчивость индексов центральности на классических структурах.

ниям центральности вершин. Таким образом, интерпретировать индексы центральности на этих сетевых структурах нужно с осторожностью, поскольку большинство моделей центральности крайне неустойчивы (исключение составляют степенная центральность при добавлении ребра, а также центральность по близости при удалении ребра в графе квадратной решетки). Среди индексов центральности, учитывающих глобальную структуру сети и имеющих низкую чувствительность к малым возмущениям, стоит выделить индекс PageRank и подграфовую центральность (случайные и безмасштабные графы, а также сети малого мира), а также LRIC (граф квадратной решетки). Стоит отметить, что собственный вектор и центральности, основанные на кратчайших путях, наибольшим образом подвержены изменениям в большинстве структур сети. Таким образом, выбор наиболее подходящей модели центральности по критерию устойчивости сильно зависит от рассматриваемой структуры сети.

4.2. Устойчивость индексов центральности на реальных сетевых структурах

Поскольку в ряде прикладных задач исследователь не обладает полной информацией о реальной структуре сети, возникает вопрос, насколько устойчив и репрезентативен полученный набор элементов сети. В рамках данного раздела анализ чувствительности моделей центральности производится на основе анализа двух реальных сетевых структур (криминальная сеть, сеть международной торговли), информация о структуре которых доступна лишь частично.

Анализ криминальной сети

Исследуемая сеть содержит информацию, собранную итальянской полицией в рамках масштабной операции в 2007 г., о взаимодействиях между членами сицилийской мафии. Рассматриваемая сеть была построена на основе зафиксированных встреч между участниками мафии и представляет собой ненаправленный граф, состоящий из 101 вершины и 256 ребер [41]. Чувствительность индексов центральности в криминальной сети осуществлена из предположения о том, что в сети отсутствует порядка 10% от общего числа зафиксированных связей. Были рассмотрены две основные стратегии добавления новых ребер в граф: случайным образом (сценарий 1) и пропорционально расстоянию между вершинами (сценарий 2). Оценка чувствительности оценивалась на основе сравнения среднего значения индекса корреляции Кендалла между исходной и модифицированной структурой сети и среднего коэффициента Жаккарда между множествами наиболее центральных элементов (ТОП-5). Полученные результаты приведены в табл. 2.

Согласно табл. 2 случайные изменения в структуре графа (сценарий 1) оказывают большее влияние на общее ранжирование всех вершин по сравнению с изменениями структуры графа на основе расстояния между вершинами (сценарий 2). Степенная центральность показывает наиболее устойчивые

Таблица 2. Устойчивость индексов корреляции в криминальной сети

Модель центральности	Сценарий 1		Сценарий 2	
	Корреляция	ТОП-5	Корреляция	ТОП-5
Степенные центральности	0,91	0,98	0,92	0,98
Собственный вектор	0,89	0,85	0,89	0,84
PageRank	0,78	0,94	0,85	0,93
Центральность Каца	0,91	0,95	0,9	0,94
Подграфовая центральность	0,92	0,91	0,92	0,88
Центральность по близости	0,81	0,88	0,83	0,87
Центральность по посредничеству	0,66	0,84	0,71	0,85
LRIC	0,74	0,8	0,79	0,81

результаты, что согласуется с результатами раздела 4.1. Среди других мер стоит также отметить центральность Каца и подграфовую центральность, которые показывают достаточно устойчивые результаты и учитывают глобальную структуру сети. Напротив, центральности, основанные на кратчайших путях, а также индекс LRIC имеют наибольшую чувствительность на рассматриваемом наборе данных.

Анализ сети международной торговли продовольствием

Исследуемая сеть содержит информацию об объеме торговли продовольствием за 2020 г. между 222 странами мира [43]. Поскольку рассматриваемый граф является направленным и взвешенным (наибольший объем торговли зафиксирован между США и Канадой – около 5,5 млрд. долл.), то анализ чувствительности был проведен только для тех индексов центральности, которые использовались ранее для анализа торговых сетей. В рамках проведенного эксперимента было осуществлено несколько сценариев по изменению структуры сети: изменение текущего объема торговли между странами в пределах 5% (учет возможных ошибок в оценке торгового потока, сценарий 1), а также добавление новых потоков между странами суммарной стоимостью в 1% от общего объема мировой торговли (для учета отсутствующих потоков из-за отсутствия отчетности некоторых стран, сценарий 2). Оценка чувствительности мер центральности проводилось на основе коэффициента корреляции Кендалла и коэффициента Жаккарда между множествами наиболее центральных элементов (ТОП-10). Полученные результаты приведены в табл. 3.

Согласно табл. 3 погрешность в оценке торговых потоков в пределах 5% (сценарий 1) не оказывает практически никакого влияния на значение индексов центральности, при этом отсутствие информации о торговле между некоторыми странами (сценарий 2) может значительно повлиять на ранжирование наиболее центральных элементов сети. В частности, наибольшая корреляция ($> 0,71$) наблюдается для исходящей степенной центральности, а также для моделей hubs (один из индексов алгоритма HITS) и LRIC. Довольно низкое значение корреляции может быть объяснено значительным изменени-

Таблица 3. Устойчивость индексов корреляции в торговой сети продовольствием

Модель центральности	Сценарий 1		Сценарий 2	
	Корреляция	ТОП-10	Корреляция	ТОП-10
Входящая степень (импорт)	0,99	0,99	0,38	0,22
Исходящая степень (экспорт)	0,99	1	0,79	0,92
Разность степеней (чистый экспорт)	0,99	1	0,14	0,81
Hubs	0,99	1	0,78	0,72
Authorities	0,99	1	0,26	0,22
PageRank	0,99	1	0,43	0,26
LRIC	0,99	0,98	0,71	0,5

ем в ранжировании тех стран, которые не входят в список наиболее важных в исходной сети. Набор наиболее влиятельных элементов (ТОП-10) является достаточно устойчивым для ряда степенных центральностей, а также для модели hubs.

5. Заключение

Концепция влияния в сетевых структурах не имеет четкого определения, что привело к возникновению большого числа различных моделей центральности. В связи с этим одним из основных инструментов для выбора наиболее подходящей меры центральности является использование аксиоматического подхода, а также анализ чувствительности моделей к изменению структуры сети.

В рамках данного исследования представлен обзор существующих аксиом центральности в сетевых структурах и показаны текущие результаты по проверке свойств классических индексов. Результаты направления являются основой для дальнейшего исследования известных индексов центральности, поскольку многие аксиомы для существующих моделей до сих пор не были проверены.

В работе проведена оценка устойчивости индексов центральности на различных графовых структурах. Поскольку большинство существующих сетей изменяются во времени, а исследователю доступна лишь частичная информация о структуре сети, возникает вопрос, насколько неслучайным является полученный набор центральных элементов. В рамках текущего направления была проведена оценка чувствительности классических моделей центральности к различным изменениям в структуре искусственных и реальных сетей. Заметим, что модели, имеющие высокую чувствительность к изменениям структуры сети, требуют более осторожного использования при анализе наиболее влиятельных элементов сети.

Полученные результаты расширяют теоретические знания о понятии центральности в сетевых структурах. Результаты исследования могут быть ис-

пользованы при выборе наиболее подходящего индекса центральности, в том числе в условиях неполной информации о реальной структуре графа. Кроме того, результаты исследования могут быть использованы для разработки новых методов по оценке центральности.

СПИСОК ЛИТЕРАТУРЫ

1. *Newman M.E.J.* Networks: Second Edition. Oxford: Oxford University Press, 2018.
2. Centiserver: The most comprehensive centrality resource and web application for centrality measures calculation. [онлайн] доступно по ссылке: <https://www.centiserver.org/centrality/list/> [дата обращения – 7.12.2023].
3. *Aleskerov F.T., Shvydun S., Meshcheryakova N.* New Centrality Measures in Networks: How to Take into Account the Parameters of the Nodes and Group Influence of Nodes to Nodes. Boca Raton: CRC Press, 2022.
4. *Shvydun S.* The Impact of COVID-19 on the Air Transportation Network // Advances in Intelligent Systems and Computing. Springer, Cham. 2023. V. 1435. P. 94–107.
5. *Sabidussi G.* The centrality index of a graph // Psychometrika. 1966. V. 31. No. 4. P. 581–603.
6. *Myerson R.B.* Graphs and cooperation in games // Math. Oper. Res. 1977. V. 2. P. 225–229.
7. *Myerson R.B.* Conference structures and fair allocation rules // Int. J. Game Theory. 1980. V. 9. P. 169–182.
8. *Boldi P., Vigna S.* Axioms for centrality // Internet Math. V. 10. 2014. P. 222–262.
9. *Skibski O., Michalak T.P., Rahwan T.* Axiomatic characterization of game-theoretic centrality // J. Artif. Int. Res. 2018. V. 62. No. 1. P. 33–68.
10. *Skibski O., Sosnowska J.* Axioms for Distance-Based Centralities // Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI). 2018. P. 1218–1225.
11. *Chebotarev P.* Selection of centrality measures using Self-consistency and Bridge axioms // J. Complex Networks. 2023. V. 11. No. 5. P. 1–23.
12. *Хитрая В.А., Мазалов В.В.* Теоретико-игровая центральность вершин ориентированного графа // МТИП. 2023. Т. 3. № 15. С. 64–87.
13. *Щербакоева Н.Г.* Аксиоматика центральности в комплексных сетях // Проблемы информатики. 2015. Т. 3. № 28. С. 3–14.
14. *Borgatti S., Carley K., Krackhardt D.* On the robustness of centrality measures under conditions of imperfect data // Soc. Networks. 2006. V. 28. No. 2. P. 124–136.
15. *Frantz T.L., Cataldo M., Carley K.M.* Robustness of centrality measures under uncertainty: Examining the role of network topology // Comput. Math. Organ. Theor. 2009. V. 15. No. 4. P. 303–328.
16. *Segarra S., Ribeiro A.* Stability and Continuity of Centrality Measures in Weighted Graphs // IEEE Transactions on Signal Processing. 2016. V. 64. No. 3. P. 543–555.
17. *Martin C., Niemeyer P.* Influence of measurement errors on networks: Estimating the robustness of centrality measures // Network Sci. 2019. V. 7. No. 2. P. 180–195.
18. *Murai S., Yoshida Y.* Sensitivity Analysis of Centralities on Unweighted Networks // The World Wide Web Conference on – WWW'19. 2019. P. 1332–1342.

19. *Meshcheryakova N., Shvydun S.* Perturbation Analysis of Centrality Measures // 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). 2023. P. 407–414.
20. *Freeman L.C.* Centrality in social networks: conceptual clarification // Soc. Networks. 1979. V. 1. No. 3 P. 215–239.
21. *Brin S., Page L.* The anatomy of a large-scale hypertextual Web search engine// Comput. Netw. 1998. V. 30. P. 107–117.
22. *Katz L.* A New Status Index Derived from Sociometric Index // Psychometrika. 1953. V. 18 P. 39–43.
23. *Kleinberg J.M.* Authoritative sources in a hyperlinked environment // J. ACM. 1999. V. 46. No. 5. P. 604–632.
24. *Estrada E., Rodriguez-Velazquez J.A.* Subgraph centrality in complex networks // Phys. Rev. E. 2005. V. 71. No. 5. P. 1–9.
25. *Bavelas A.* Communication Patterns in Task-Oriented Groups // J. Acoust. Soc. Amer. 1950. V. 22. No. 6. P. 725–730.
26. *Anthonisse J.M.* The rush in a directed graph // Technical Report BN 9/71, Stichting Mathematisch Centrum. Amsterdam. 1971.
27. *Freeman L.C.* A set of measures of centrality based upon betweenness // Sociometry. 1977. V. 40. No. 1. P. 35–41.
28. *Aleskerov F.T., Meshcheryakova N.G., Shvydun S.V.* Power in Network Structures // Springer Proceedings in Mathematics & Statistics. 2017. V. 197. P. 79–85.
29. *Aleskerov F.T., Andrievskaya I.K., Permjakova E.E.* Key borrowers detected by the intensities of their short-range interactions // Working papers by NRU Higher School of Economics. Series FE “Financial Economics”. 2014. No. WP BRP 33/FE/2014.
30. *Aleskerov F., Yakuba V.* Matrix-vector approach to construct generalized centrality indices in networks // NRU Higher School of Economics. Series WP7 “Mathematical methods for decision making in economics, business and politics”. 2020. No. 2323.
31. *Aleskerov F., Andrievskaya I., Nikitina A., Shvydun S.* Key Borrowers Detected by the Intensities of Their Interactions // Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning (In 4 Volumes). World Scientific. 2020. P. 355–389.
32. *Aleskerov F., Gavrilenkova I., Shvydun S., Yakuba V.* Power Distribution in the Networks of Terrorist Groups: 2001–2018 // Group. Decis. Negot. 2020. V. 29. P. 399–424.
33. Centrality Measures: from Theory to Applications. [онлайн] доступно по ссылке: <https://centrality.mimuw.edu.pl/> [дата обращения – 7.12.2023].
34. *Skibski O.* Vitality Indices are Equivalent to Induced Game-Theoretic Centralities // IJCAI-21. 2021. P. 398–404.
35. *Nieminen U.J.* On the centrality in a directed graph // Soc. Sci. Res. 1973. V. 2. No. 4. P. 371–378.
36. *Garg M.* Axiomatic Foundations of Centrality in Networks // Soc. Sci. Res. Network. 2009. No. 1372441.
37. *Landherr A., Friedl B., Heidemann J.* A Critical Review of Centrality Measures in Social Networks // Bus. Inf. Syst. Eng. 2010. V. 2. P. 371–385.
38. *Bloch F., Jackson M.O., Tebaldi P.* Centrality measures in networks // Soc. Choice Welf. 2023. V. 61. P. 413–453.

39. *Aizerman M.A., Aleskerov F.T.* Theory of choice. Amsterdam: North-Holland. 1995.
40. *Ficara A., Cavallaro L., Curreri F., Fiumara G., De Meo P., Bagdasar O., Song W., Liotta A.* Criminal networks analysis in missing data scenarios through graph distances // PLOS ONE. 2021. V. 16. No. 8. e0255067.
41. *Ficara A., Cavallaro L., De Meo P., Fiumara G., Catanese S., Bagdasar O., Liotta A.* Social Network Analysis of Sicilian Mafia Interconnections // Studies Comput. Int. 2020. V. 882. P. 440–451.
42. *Meshcheryakova N.* Network Analysis of Bilateral Trade Data Under Asymmetry // 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). 2020. P. 379–383.
43. The World Integrated Trade Solution. [онлайн] доступно по ссылке: <https://wits.worldbank.org/> [дата обращения – 7.12.2023].

Статья представлена к публикации членом редколлегии П.Ю. Чеботаревым.

Поступила в редакцию 11.12.2023

После доработки 19.06.2024

Принята к публикации 10.07.2024