УДК 519. 87 + 51-77

КОНТЕКСТНО-ЗАВИСИМЫЕ МЕРЫ ЦЕНТРАЛЬНОСТИ

Е.Н. Кузнецов
Институт проблем управления им. В.А. Трапезникова РАН
Россия, 117997, Москва, Профсоюзная ул., 65
E-mail: enkuznetsov@mail.ru

Ключевые слова: меры центральности, ключевые элементы системы взаимосвязанных объектов, монотонные системы.

Аннотация: Для анализа системы взаимосвязей и взаимодействия некоторого множества объектов определяются контекстно-зависимые меры центральности – предлагаемая мера центральности каждого элемента зависит от подмножества элементов, для которого рассматривается в данный момент. В качестве общей характеристики важности, влиятельности и т.д. некоторой группы вершин предлагается использовать минимальное или максимальное, а не среднее, значение индекса центральности вершин этой группы. Для выделения подмножества ключевых узлов сети предлагается использовать алгоритмы специального кластер анализа – алгоритмы выделения ядра монотонной системы. Это обеспечивает определение глобального экстремума функционала в соответствующей оптимизационной задаче, а также позволяет более подробно проанализировать структуру сети. Для примера предлагаемый подход применен для анализа структуры сети экспортных связей стран – членов Евросоюза.

1. Введение

Для представления и анализа системы взаимосвязей и взаимодействия некоторого множества объектов часто используется граф («социальная сеть») [1]. Вершины в этом графе представляют объекты (или субъекты, «акторы»), а ребра – связи между ними. Наличие связи может характеризовать близость (похожесть) между объектами, различие (расстояние) между ними или факт (интенсивность) их взаимодействия. При этом рассматривают как симметричные, так и направленные связи между объектами. Наконец, часто используются взвешенные графы, в которых каждому ребру приписано некоторое число – вес, характеризующее меру близости или расстояния (степень похожести или различия) между объектами, или интенсивность воздействия (влияния) одного объекта на другого, объем материального, информационного или финансового потока между ними и тому подобное.

Для анализа той роли, которую играет каждый объект во взаимосвязанной системе (каждый субъект в социальной сети) рассматривается сама концепция центральности и вводятся различные ее числовые характеристики – меры центральности (centrality indices). Концепция центральности широко используется в социальных и поведенческих науках, а также в области политических наук, управления, социальных сетей в интернете, экономики, биологии и т. д. [2].

Исследователи социальных сетей разработали множество количественных мер (индексов) центральности. Наиболее известными и распространенными являются следующие четыре группы мер [3, 4]: центральность по степени связности (degree centrality);

центральность по близости (closeness centrality); центральность по посредничеству (betweenness centrality); центральности по влиятельности (eigenvector centrality).

В то же время, как неоднократно отмечалось в литературе (см., например, [5, 6]), классические индексы не всегда учитывают структуру сети, а также интенсивности прямых и опосредованных, в частности, внутригрупповых взаимодействий в системе. Поэтому развитие новых подходов к анализу сетевой структуры является актуальным.

Многие прикладные работы по сетевому анализу, как правило, состоят из двух основных этапов. Сначала для всех узлов имеющейся сети (акторов) рассчитываются те или иные индексы (меры центральности). Затем производится ранжирование всех узлов по каждому индексу и выделяются группы узлов с наибольшими или наименьшими значениями этих индексов (ТОП-3, ТОП-5 и т.д.) [7].

Общей чертой всех известных автору работ по анализу сетевой структуры на основе мер центральности является то, что значения индексов вычисляются для каждой вершины раз и навсегда – по всем вершинам сети. Иначе говоря, индекс центральности рассматривается как свойство или функция узла на всей сети (то есть функция одного аргумента – узла). Кроме того, для общей характеристики важности, влиятельности и т.д. некоторой группы вершин используются среднее (реже, суммарное или медианное) значение соответствующего индекса центральности [2].

Для сетевого анализа центральности мы предлагаем **три нововведения**.

Первое. Предлагается рассмотреть контекстно-зависимые меры центральности. То есть, с самого начала мера центральности вводится как функция двух аргументов: некоторого узла сети и некоторого подмножества узлов сети. Иначе говоря, мера центральности рассматривается как мера отношения (связи, роли, взаимодействия и так далее) элемента и группы элементов. То есть мера (степень) центральности каждого элемента явно зависит от контекста – от того подмножества элементов, для которого рассматривается в данный момент. Актор, занимающий лидирующие позиции в одной группе вершин, может быть вполне периферийным узлом в составе другой группы.

Второе. Для общей интегральной характеристики важности, влиятельности и т.д. некоторой группы вершин предлагается использовать крайние значения индекса центральности вершин этой группы (минимальное или максимальное, а не среднее).

Третье. Для определения подмножества ключевых узлов сети предлагается использовать алгоритмы специального кластер анализа – алгоритмы выделения ядра монотонной системы. Это не только обеспечивает определение глобального экстремума функционала в соответствующей оптимизационной задаче, но и позволяет более подробно проанализировать структуру сети.

Во второй части доклада вводятся основные формальные понятия. С математической точки зрения данная работа основана на работе [8]. В ней и в работе [9] приведены все доказательства. В третьей части работы кратко описывается практический пример применения предлагаемого подхода к сетевому анализу системы торгово-экономических потоков между странами-членами Евросоюза.

2. Контекстно-зависимая мера сетевой центральности

Пусть $V=\{1,2,...,n\}$ конечное множество связанных узлов (элементов, акторов и т.д.). Кроме того пусть есть матрица $W=\left[w_{ij}\right]_{N\times N}$, количественно характеризующая парные связи между узлами. Для определенности будем предполагать, что $w_{ij}\geq 0,\ \forall i\neq j,\ w_{ii}=0$.

Рассмотрим произвольное подмножество $H \subseteq V$. Определим на каждом таком подмножестве скалярную числовую функцию его элементов

(1) $$\pi(i,H)=\sum_{j\in H} w_{ij},\ \forall i\in H$$

Число $\pi(i,H)$ назовем весом элемента i на подмножестве H. Таким образом, вес каждого элемента подмножества H равен сумме связей этого элемента со всеми элементами того же подмножества (это простейший способ определения веса элемента). В соответствии с содержательным смыслом весов ребер можно интерпретировать вес $\pi(i,H)$ как близость элемента и подмножества узлов исходного графа, или как расстояние элемента от подмножества элементов, или как величину суммарного потока от одного узла в группу узлов, или как меру влияния одного узла на подмножество узлов и тому подобное. То есть в данном случае число $\pi(i,H)$ – это прямой аналог меры степенной центральности (взвешенной), только измеренной не по всей сети, а на некотором подмножестве узлов H (внутригрупповая мера центральности).

Система функций $\pi(i,H)$, так определенных на множестве всех подмножеств множества V, очевидно, удовлетворяет условию монотонности. Действительно, для любых двух вложенных подмножеств $H_1 \subseteq H_2$ и для любого элемента $i \in H_1$ выполняется неравенство $\pi(i,H_1) \le \pi(i,H_2)$, то есть

$$\pi(i,H_1) \le \pi(i,H_2),\quad \forall i,H_1,H_2 : i\in H_1 \subseteq H_2 \subseteq V$$

Такого рода системы названы в [10] монотонными. Существенный вклад в развитие теории таких систем внесли работы [9, 11, 12] и другие.

Определим теперь на множестве всех подмножеств множества узлов одну скалярную функцию по следующему правилу

(2) $$F(H)=\min_{i\in H}\pi(i,H),\quad \forall H\subseteq V$$

Таким образом, функция $F(H)$ есть величина наименьшей меры центральности подмножества узлов H. То есть каждое подмножество узлов будем характеризовать не средней, а наименьшей мерой центральности его элементов. Напомним, что мера центральности каждого узла здесь контекстно-зависима, то есть характеризует связи этого узла с другими узлами только в пределах рассматриваемого подмножества.

Теперь поставим задачу выделить такое наибольшее подмножество K, $K \subseteq V$ узлов исходного графа, на котором функция $F(H)$ принимает максимальное значение

(3) $$F(K)=\max_{H\subseteq V}F(H)=\max_{H\subseteq V}\min_{i\in H}\pi(i,H)$$

Подмножества K, $K \subseteq V$, на которых функция $F(H)$ принимает максимальное значение, названы в [10] ядрами монотонной системы. В этой же работе показано, что множество всех ядер монотонной системы замкнуто относительно операции объединения множеств, и, тем самым, в любой монотонной системе существует одно наибольшее по вложению ядро.

Нужно определить такое подмножество узлов K, в котором даже узел с наименьшей мерой центральности (в пределах этого подмножества) имеет наибольшую меру центральности среди всех подмножеств $H \subseteq V$. Можно трактовать это и таким образом, что речь идет о выделении подмножества «наиболее взаимосвязанных (взаимовлияющих)», центральных элементов исходной системы, образующих ее ядро.

Ядра монотонных систем обладают многими замечательными свойствами (см., например, [8-10, 12, 13]). Было предложено несколько различных алгоритмов выделения ядер монотонной системы (см., например, [9] и др.). Экстремальные свойства ядра монотонной системы, разумеется, не зависят от конкретного алгоритма, применяемого для его выделения. Простейший алгоритм описан в работе [8].

В основе алгоритма решения задачи **Ошибка! Источник ссылки не найден.** используется процедура построения так называемой определяющей последовательности элементов [10]. На каждом шаге алгоритма определяется элемент с минимальным весом на подмножестве элементов, еще не включенных в определяющую последовательность к этому шагу, и фиксируется пороговое — максимальное значение веса включаемого элемента. Элементы, включенные в определяющую последовательность после этого, образуют наибольшее ядро. Алгоритм является достаточно простым в вычислительном отношении и поэтому пригоден для обработки больших систем взаимосвязанных элементов.

3. Ядро Евросоюза по данным взаимного экспорта стран – членов (2016 год)

В качестве примера предлагаемый подход был применен для выявления особенностей торговых отношений государств Евросоюза внутри ЕС по данным об экспорте (внешней торговле товарами) за 2016 год (www.eurostat.eu).

Евросоюз – самое известное и наиболее развитое экономическое и социально-политическое объединение стран в мире. На долю ЕС в последние годы приходилось около четверти мирового валового внутреннего продукта, и это при том, что доля стран ЕС в мировом населении составляет не более десяти процентов [14]. Совокупный размер ВВП стран ЕС превышает размер ВВП США [15]. Внешнеторговые связи стран – членов Евросоюза не только характеризуют степень вовлеченности их во внутрирегиональный торговый оборот, но и выступают как существенный фактор их регионального единства и развития. Велика роль Евросоюза и в мировой торговле: доля ЕС (с учетом внутрирегиональной торговли) составляет около 40% международной торговли товарами [14, 15]. Однако, несмотря на постоянно развивающийся процесс интеграции, Евросоюз не стал пока полностью однородным образованием. И хотя объем и доля внешней торговли внутри объединения у стран – членов Евросоюза последние десятилетия постоянно росли, они развивались по-разному у разных стран ЕС.

В этом случае V – это множество стран – членов Евросоюза, $|V| = N = 28$, а w_{ij} – объем экспорта из страны i в страну j. Если взять произвольное подмножество $H \subseteq V$, то функция **Ошибка! Источник ссылки не найден.** будет определять суммарный экспорт из страны i во все страны подмножества H и может рассматриваться как внутригрупповая мера центральности.

Нужно определить такое подмножество узлов K, в котором даже узел с наименьшей мерой центральности (в пределах этого подмножества) имеет наибольшую внутригрупповую меру центральности среди всех подмножеств $H \subseteq V$. Иными словами, нужно определить подмножество K «наиболее взаимосвязанных (взаимовлияющих)» центральных элементов исходной системы внутренних экспортных связей стран Евросоюза. Таким образом, ядро K включает страны с наибольшим внутригрупповым товарооборотом в смысле критерия **Ошибка! Источник ссылки не найден.** в составе экономической системы Евросоюза.

По данным об экспорте (внешней торговле товарами) стран Евросоюза за 2016 год получены следующие результаты.

В качестве ядра (доставляющего глобальный максимум функционалу **Ошибка! Источник ссылки не найден.**) алгоритм выделил следующие страны: **Великобритания,**

Испания, Италия, Франция, Бельгия, Германия, Нидерланды. Подчеркнем: это выделение сделано не человеком – экспертом, возможно учитывающим всю полноту информации о странах Евросоюза, а предлагаемым формальным алгоритмом на основе только данных об объемах экспорта товаров из каждой страны в каждую другую страну.

Первым элементом ядра Евросоюза в определяющей последовательности (по данным внутригруппового экспорта 2016 года) стала Великобритания. То есть это как раз первый (особый, крайний) из ключевых элементов европейского сообщества.

Заметим, что механическое удаление Великобритании из состава ядра приводит к его разрушению – никакая другая группа стран Евросоюза без Великобритании не дает такое же или большее значение функции **Ошибка! Источник ссылки не найден.**, характеризующей, как отмечалось выше, минимальную взаимосвязанность группы стран. Напомним, что ядро – это подмножество стран с максимальной среди всех возможных групп стран минимальной внутригрупповой взаимосвязанностью.

Чтобы оценить в целом внутригрупповые и внешние экспортные связи стран ядра рассмотрим суммарный экспорт стран – ядра Евросоюза внутри ядра, внутри всего Евросоюза и во все страны мира кроме стран ЕС.

Оказалось, что среди группы стран, составляющих ядро Евросоюза, Великобритания имеет наименьший суммарный экспорт в страны ядра, а Германия, как и можно было ожидать, – наибольший. Кроме того именно и только Великобритания продает в остальные страны мира кроме ЕС больше (194.459 млрд евро), чем внутри Евросоюза (174.980 млрд евро). Остальные страны ядра гораздо в большей степени ориентированы на внутригрупповую внешнюю торговлю. Экспорт некоторых из них (Испания, Бельгия, Нидерланды) внутри Евросоюза превышает экспорт в остальной мир более чем вдвое. Трудно представить, чтобы кто-то из них захотел выйти из Евросоюза.

Разумеется, здесь представлены только первые, самые поверхностные результаты, в качестве примера применения предлагаемого подхода к сетевому анализу центральности системы взаимосвязанных элементов.

Список литературы

1. Wasserman S., Faust K. Social Network Analysis: Methods and Applications. Cambridge: Cambridge Univercity Press, 1994. 828 p.
2. Models and Methods in Social Network Analysis / Ed. by. P.J. Carrington, J. Scott, S. Wasserman. Cambridge: Cambridge Univercity Press, 2005. 334 p.
3. Bonacich P. Power and Centrality: A family of Measures // American Journal of Sociology. 1987. Vol. 92, No. 5. P. 1170-1182.
4. Bonacich P., Lloyd P. Eigenvector-like measures of centrality for asymmetric relations // Social Networks. 2001. Vol. 23, No. 3. P. 191-201.
5. Landherr A.A., Friedl B., Heidemann J. Critical Review of Centrality Measures in Social Networks // Business & Information Systems Engineering. 2010. Vol. 2, No. 6. P. 371-385.
6. Aleskerov F. Meshcheryakova N., Shvydun S. Centrality measures in networks based on nodes attributes, long-range interactions and group influence // Mathematical methods for decision making in economics, business and politics 2016. P. 1-44.
7. Алескеров Ф.Т., Бадгаева Д.Н., Писляков В.В., Стерлигов И.А. Значимость основных российских и международных экономических журналов: сетевой анализ // Журнал Новой экономической ассоциации. 2016. Т. 30, № 2. С. 193-205.
8. Кузнецов Е.Н. Анализ структуры матрицы связей с помощью построения на ней монотонной системы // Автоматика и телемеханика. 1980. № 7. С. 128-136.
9. Кузнецов Е.Н., Мучник И.Б., Шварцер Л.В. Монотонные системы и их свойства // Анализ нечисловой информации в социологических исследованиях. 1985. С. 29-57.
10. Муллат И.Э. Экстремальные подсистемы монотонных систем. I // Автоматика и телемеханика. 1976. № 5. С. 130-139.

11. Малишевский А.В. Свойства порядковых функций множеств // Качественные модели в теории сложных систем. 1998. С. 169-173.
12. Мучник, И.Б., Шварцер Л.В. Субмодулярные функции множеств и монотонные системы в задачах агрегирования. I // Автоматика и телемеханика. 1987. № 5. С. 135-148.
13. Закс, Ю.М., Мучник И.Б. Монотонные системы для построения неполных классификаций конечного множества объектов // Автоматика и телемеханика. 1989. № 4. С. 155-164.
14. World Trade Statistical Review 2016. WTO, 2017. 165 p.
15. Родионова, И. А.; Умерова И.А. Позиции стран-членов ЕС в мировой и внутрирегиональной торговле // Зарубежный опыт 2009. Т. 41. № 134. С. 103-108.