АНАЛИЗ СТРУКТУРЫ СЕТЕВЫХ ВЗАИМОДЕЙСТВИЙ: КОНТЕКСТНО-ЗАВИСИМЫЕ МЕРЫ ЦЕНТРАЛЬНОСТИ

Кузнецов Е. Н.[1]
(*ФГБУН Институт проблем управления им. В.А. Трапезникова РАН, Москва*)

Классические индексы центральности для анализа системы взаимосвязей и взаимодействия некоторого множества объектов и распространенные методы выделения ключевых элементов сети при помощи ранжирования узлов по величине общесетевого индекса центральности не всегда учитывают интенсивности внутригрупповых взаимодействий в различных частях системы. В данной работе предлагается использовать контекстно-зависимые меры центральности – на основе внутригруппового взаимодействия элементов сети. То есть предлагаемая мера центральности каждого элемента зависит от того подмножества элементов, для которого рассматривается в данный момент. Для общей интегральной характеристики важности, влиятельности и т.д. некоторой группы вершин предлагается использовать минимальное или максимальное, а не среднее, значение индекса центральности вершин этой группы. 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. В качестве примера применения предлагаемого подхода рассмотрена сеть экспортных связей стран – членов Евросоюза. Выявлена особая – пограничная роль Великобритании в ядре Евросоюза.

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

1. Введение

Для представления и анализа системы взаимосвязей и взаимодействия некоторого множества объектов часто используется граф («социальная сеть») [18]. Вершины в этом графе представляют объекты (или субъекты, «акторы»), а ребра – связи между ними[2]. При этом наличие связи может характеризовать близость

[1] *Евгений Николаевич Кузнецов, к.т.н., ст.н.с.* (enkuznetsov@mail.ru).

[2] *Разумеется, не все пары вершин связаны ребрами. Граф, в котором все пары вершин связаны ребрами, называется полным.*

(похожесть) между объектами, различие (расстояние) между ними или факт (интенсивность) их взаимодействия. В первых двух случаях, как правило, речь идет о симметричных связях, и этому схематично соответствует неориентированный граф. В последнем случае обычно рассматривают направленные связи между объектами, что представляется ориентированным графом. 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, объём (величину) 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 её 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. [19].

Так, в анализе социальных сетей рассматривают тесно связанные понятия центральности и индивидуальной власти, лидерства, положения коммуникативного посредника, периферийного или почти изолированного узла [15, 18]. В анализе свойств сетей на рынке межбанковского кредитования и платежных систем проводится количественная оценка значимости элементов финансовых взаимодействий и накопленного системного риска [7]. Индексы влияния и меры центральности используются также для определения ключевых игроков на международном рынке капитала [13]. Разумеется, используется этот аппарат и для исследования структуры и процессов в политических институтах, таких как парламент [1], исследования структуры потоков международной миграции [12] и во многих других областях .

Исследователи социальных сетей разработали множество количественных мер (индексов) центральности. Наиболее известными и распространенными являются следующие четыре группы мер [15, 16]:

– центральность по степени связности (degree centrality);
– центральность по близости (closeness centrality);
– центральность по посредничеству (betweenness centrality);
– центральность по влиятельности (eigenvector centrality).

В то же время, как неоднократно отмечалось в литературе (см., например, [14, 17]), классические индексы не всегда учитывают структуру сети и индивидуальные характеристики отдельных акторов, а также интенсивности прямых и опосредованных, в частности, *внутригрупповых* взаимодействий в системе. Поэтому были предложены и иные меры центральности, в частности, индексы центральности по ближним (short-range interactions centrality, SRIC) и дальним (long-range interactions centrality, LRIC) взаимодействиям, которые позволяют учитывать данные особенности [3, 13, 14]. Предложенный в [13] индекс центральности заёмщика учитывает характер и количественную составляющую связей между субъектами финансовой системы для прямых контрагентов и косвенных контрагентов первого уровня.

Тем не менее ни одна из имеющихся в распоряжении исследователя мер центральности, ни все они в совокупности не позволяют полностью удовлетворить потребности анализа сетевой структуры близости (различия) между элементами, сетевой структуры сложной системы материальных, финансовых, информационных и других потоков или системы взаимодействия между акторами. Поэтому развитие новых подходов к анализу сетевой структуры является актуальным.

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 на основе полученных значений проводится ранжирование всех узлов по каждому индексу. Это позволяет выделить в сети группы узлов с наибольшими или наименьшими значениями этих индексов (ТОП-3, ТОП-5 и т.д.). Или шире – выделить группы узлов, отвечающих тем или иным априорно налагаемым требованиям. Например, группу узлов, индексы центральности которых превышают (или не превышают), некоторое пороговое значение или принимают наибольшие значения. Так, в [2] определяются наиболее важные международные и российские экономические журналы в сетях в результате анализа перекрестного цитирования научных публикаций этих журналов. По каждому используемому индексу определяется группа ТОП-3 наиболее влиятельных журналов по экономике. Разумеется, перечень этих «наиболее влиятельных журналов» по разным индексам сильно отличается.

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 (т.е. функция одного аргумента – узла). Кроме того, для общей характеристики важности, влиятельности и т.д. некоторой группы вершин используются такие стандартные интегральные характеристики этой группы как среднее (реже, суммарное или медианное) значение используемого индекса центральности [19].

В данной работе для сетевого анализа центральности предлагается три нововведения.

1. 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 т.д.) элемента и группы элементов. *То есть* ***мера*** *(степень)* ***центральности*** *каждого элемента определяется не единожды, по всей сети, а* ***зависит от контекста*** *– от того подмножества элементов, для которого рассматривается в данный момент*. Таким образом, значение индекса центральности каждой вершины явно зависит от подмножества вершин, для которого оно вычисляется. Актор, занимающий лидирующие позиции в одной группе вершин, может быть вполне периферийным узлом в составе другой группы. Этот

подход как бы реализует идею известной поговорки «Молодец – среди овец, а на молодца – и сам овца».

2. 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). То есть группа характеризуется по самому влиятельному своему члену или, наоборот, итоговое время марш-броска взвода – по наиболее отставшему солдату.

3. 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 статьи вводятся все необходимые формальные понятия. С математической точки зрения данная работа основана на работе [6]. В третьей части описываются алгоритмы решения поставленной задачи. В четвертой части рассматриваются свойства построенных структурных объектов. Наконец, в пятой части работы кратко описывается практический пример применения предлагаемого подхода к сетевому анализу системы торгово-экономических потоков между странами-членами Евросоюза.

Автор благодарит Ф.Т. Алескерова и Ю.Э. Гурову за ценные замечания и помощь в работе над статьей.

2. Контекстно-зависимая мера сетевой центральности

Пусть $V = \{1, 2, \ldots, n\}$ – конечное множество связанных узлов (вершин, элементов, акторов и т.д.). Характер связей между ними может быть самый произвольный. Широко известны по крайней мере три типа таких связей:

1. Мера близости (похожести) или расстояния (различия) между двумя узлами.

2. Величина (объем, масса) потока между ними (материального, финансового, информационного и др.).

3. Интенсивность взаимодействия (влияния) одного элемента на другой.

Рассмотрим взвешенный граф $G = \langle V, E, W \rangle$, где V, $|V| = N$ – множество вершин, $E \subseteq V \times V$ – множество ребер, а $W = \{w_{ij}\}$ – множество весов ребер, где w_{ij} – неотрицательное действительное число, соответствующее ребру $(i, j) \in E$. Граф G называется ориентированным, если ребро $(i, j) \in E$ характеризует направленную связь между узлом i и узлом j (например, величину некоторого потока от узла i к узлу j, меру влияния узла i на узел j и т.п.), и неориентированным, если ребро $(i, j) \in E$ характеризует ненаправленную связь между узлами i и j (например, меру сходства или различия между ними).

Невзвешенный граф G может быть представлен с помощью матрицы смежности $A= [a_{ij}]_{N\times N}$, где $a_{ij} = 1$, если $(i, j) \in E$, и $a_{ij} = 0$ в противном случае, а взвешенный – с помощью матрицы весов $W = [w_{ij}]_{N\times N}$. При этом предполагается, что все вершины графа пронумерованы и, тем самым, различимы. Граф G называется полным, если каждая вершина соединена с каждой вершиной, т.е. $E = V \times V$. Однако часто под полным графом понимают полный граф без петель. То есть граф, в котором каждая вершина соединена только с каждой другой вершиной.

Исходя из смысла весов ребер, любой взвешенный граф можно представить в виде полного графа, соответствующим образом дополнив матрицу весов (парных связей). Например, если вес w_{ij} характеризует близость (похожесть) вершин i и j, и при этом максимальной близости соответствует значение веса, например, равное 1, то естественно заполнить диагональ матрицы весов единицами (каждая вершина в наибольшей степени похожа сама на себя). Если вес – это расстояние между вершинами (например, расстояние, вычисляемое по какой-либо формуле в пространстве признаков), то вес отсутствующих ребер в исходном графе можно определить как кратно увеличенное максимальное возможное расстояние между вершинами (заменяя, таким образом, для практических нужд последующего анализа бесконечную величину). Диагональ в этом случае – расстояния от каждой вершины до самой себя – естественно заполнить нулями. Если вес w_{ij} характеризует величину какого-либо потока между вершинами (материального, финансового или информационного), то, наоборот, значения весов отсутствующих ребер в исходном графе естественно заполнить нулями. Аналогичным образом поступают и в том случае, если вес ребер характеризует интенсивность взаимодействия (прямого влияния) одного элемента на другой.

В любом случае, мы имеем конечное множество V связанных узлов (вершин, элементов, акторов и т.д.) и матрицу $W = [w_{ij}]_{N\times N}$, количественно характеризующую парные связи между ними. (Сначала мы рассматриваем только симметричные матрицы связей.) Для определенности будем предполагать, что $w_{ij} \geq 0$, $\forall i \neq j$, $w_{ij} = 0$.

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 $\pi(i, H)$, приписываемое этой функцией элементу i на подмножестве H, назовем весом элемента i на H. Таким образом, вес каждого элемента подмножества H равен сумме связей этого элемента со всеми элементами того же подмножества. 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 или, можно сказать, измеренной на подграфе исходного графа, определяемом подмножеством узлов H.

Система функций $\pi(i, H)$, так определенных на множестве всех подмножеств множества V, очевидно, удовлетворяет условию монотонности (напомним, что веса ребер неотрицательны, $w_{ij} \geq 0$). Действительно, для любых двух вложенных подмножеств $H_1 \subseteq H_2$ и для любого элемента $i \in H_1$ выполняется неравенство $\pi(i, H_1) \leq \pi(i, H_2)$, т.е.

(2) $\pi(i, H_1) \leq \pi(i, H_2), \quad \forall i, H_1, H_2 : i \in H_1 \subseteq H_2 \subseteq V.$

Свойство монотонности можно записать и иначе (можно показать, что эти две записи эквивалентны):

(3) $\pi(i, H \setminus k) \leq \pi(i, H), \quad \forall i \neq k,\ i, k \in H,\ \forall H \subseteq V.$

Такого рода системы названы в [9] монотонными. Существенный вклад в развитие теории таких систем внесли работы [5, 8, 10] и др.

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

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

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, т.е. 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

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

Подмножества K, $K \subseteq V$, на которых функция $F(H)$ принимает максимальное значение, названы в [9] 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.

Иначе говоря, 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.

Предположим, что вес w_{ij} характеризует величину какого-либо потока между вершинами (материального, финансового или информационного). Пусть узлы – это субъекты некоторой экономической системы (например, страны Евросоюза), а вес w_{ij} парной связи характеризует объем торговли (либо только экспорт из i-й страны в j-ю, либо весь товарооборот между ними). Тогда ядро K включает страны с наибольшим внутригрупповым товарооборотом в смысле критерия (5) в составе экономической системы.

Ядра монотонных систем обладают многими замечательными свойствами [4–6, 9, 10], некоторые из которых описаны ниже в разделе 4. Но сначала рассмотрим алгоритм выделения ядра монотонной системы.

3. Алгоритм выделения ядра – центрального подмножества элементов взаимосвязанной системы

За годы, прошедшие с появления теории монотонных систем [9], было предложено несколько различных алгоритмов выделения ядер монотонной системы (см., например, [5] и др.). Экстремальные свойства ядра монотонной системы, разумеется, не зависят от алгоритма, применяемого для его выделения. Здесь дадим краткое описание простейшего алгоритма, следуя работе [6].

В основе алгоритма решения задачи (5) используется процедура построения так называемой определяющей последовательности элементов и выделения в ней наибольшего ядра – определимого множества [9].

Алгоритм описывается рекуррентным образом по шагам.

Первый шаг. Положим $H_1 = V$. Находим такой элемент $\alpha_1 \in V$, что

$$\pi(\alpha_1, V) = \min_{i \in V} \pi(i, V).$$

Элемент α_1 первым «удаляется» из множества V и объявляется первым элементом выстраиваемой определяющей последовательности[1]. Обозначим его вес $\pi(\alpha_1, V)$ через δ_1. Это число будет использоваться как начальное значение порога для сравнения с весами других элементов, но рассматриваемых уже на множестве $H_2 = H_1 \setminus \alpha_1$, на следующем шаге. Сравнение осуществляется для «удаления» других элементов. В то же время значение порога сравнения в дальнейшем будет изменяться.

k-й шаг. Пусть после выполнения $(k - 1)$ шагов алгоритма выстроена последовательность элементов $\langle \alpha_1, \alpha_2, \ldots, \alpha_{k-1} \rangle$ и в качестве текущего значения порога для сравнения получено некоторое число δ_{k-1}. Обозначим через $H_k = H_{k-1} \setminus \alpha_{k-1}$ подмножество всех оставшихся к k-му шагу, т.е. еще не включенных в определяющую последовательность, элементов множества V:

$$H_k = V \setminus \langle \alpha_1, \alpha_2, ..., \alpha_{k-1} \rangle.$$

Находим такой элемент $\alpha_k \in H_k$, что

$$\pi(\alpha_k, H_k) = \min_{i \in H_k} \pi(i, H_k),$$

и производим сравнение его веса с текущей величиной порога δ_{k-1}. Если выполняется условие

$$\pi(\alpha_k, H_k) \le \delta_{k-1} \tag{6}$$

и $k < N = |V|$, то переходим к $(k + 1)$-му шагу, причем величина порога сохраняется:

$$\delta_k = \delta_{k-1}.$$

В противном случае, т.е. если не выполняется неравенство (6), перед переходом к следующему шагу величина порога изменяется:

$$\delta_k = \pi(\alpha_k, H_k).$$

---

[1] *Если таких элементов несколько, то в определяющую последовательность вставляется любой из них. Как показано в работах* [5, 9], *это не влияет не только на состав выделяемого в последующем наибольшего ядра, но и на состав всех квазиядер – см. далее. Аналогичное свойство имеет место на каждом шаге алгоритма.*

В этом случае, очевидно, $\delta_k > \delta_{k-1}$ (т.е. это неравенство выполняется только для некоторых k). Алгоритм заканчивает работу, когда полностью исчерпано исходное множество V, т.е. все его элементы выстроены в последовательность $\langle \alpha_1, \alpha_2, \ldots, \alpha_N \rangle$, которая и названа определяющей. При этом получена также сопутствующая последовательность подмножеств $\bar{H}$:

$$\bar{H} = \left\langle H_1, H_2, ..., H_N \right\rangle,$$

где $H_1 = V$, $H_k = H_{k-1} \setminus \alpha_{k-1}$, $(H_N = \alpha_N)$ а α_k – k-й элемент определяющей последовательности. Последовательность элементов $\langle \alpha_1, \alpha_2, \ldots, \alpha_N \rangle$ и сопутствующая ей последовательность подмножеств $\bar{H}$ могут быть выстроены не единственным образом, если на некотором шаге оказалось более одного элемента с минимальным весом. Однако, как показано далее, это не помешает решению поставленной задачи.

Выделим из определяющей последовательности $\langle \alpha_1, \alpha_2, \ldots, \alpha_N \rangle$ специальную подпоследовательность $\left\langle \alpha_{j_1}, \alpha_{j_2}, ..., \alpha_{j_p} \right\rangle$, элементы которой определяют шаги алгоритма со сменой значений порога сравнения (положим $\alpha_{j_1} = \alpha_1$). Обозначим общее число таких различных значений порога через $\boldsymbol{p}$. Для упрощения дальнейшего рассмотрения введем специальные обозначения для этой последовательности элементов и связанной с ней последовательности подмножеств:

$$(7) \quad \bar{\gamma} = \left\langle \gamma_1 = \alpha_{j_1}, \gamma_2 = \alpha_{j_2}, ..., \gamma_p = \alpha_{j_p} \right\rangle,$$

$$(8) \quad \bar{\Gamma} = \left\langle \Gamma_1 = H_{j_1}, \Gamma_2 = H_{j_2}, ..., \Gamma_p = H_{j_p} \right\rangle.$$

Аналогичным образом переобозначим и числовую последовательность возникающих порогов:

$$(9) \quad \bar{\varepsilon} = \left\langle \varepsilon_1 = \delta_{j_1}, \varepsilon_2 = \delta_{j_2}, ..., \varepsilon_p = \delta_{j_p} \right\rangle,$$

где $\varepsilon_{j-1} < \varepsilon_j$, $j = 1, \ldots, p$.

Обозначим через K соответствующее подмножество $\Gamma_p = H_{j_p}$, т.е. подмножество элементов, включенных в определяющую последовательность после последнего изменения значения порога на j_p-м шаге, начиная с самого элемента γ_p до конца определяющей последовательности. В [9] доказано, что на подмножестве $K = \Gamma_p$ функция $F(H)$ достигает глобального максимума и такое подмножество $K = \Gamma_p$ единственно[1]. Тем самым множество узлов K есть решение поставленной задачи о выделении подмножества «наиболее взаимосвязанных (взаимовлияющих)» элементов.

Алгоритм является достаточно простым в вычислительном отношении и поэтому пригоден для обработки больших систем взаимосвязанных элементов.

Для того чтобы наглядно продемонстрировать работу алгоритма, рассмотрим следующий простой пример системы взаимосвязанных элементов, которой соответствует взвешенный граф на рис.1.

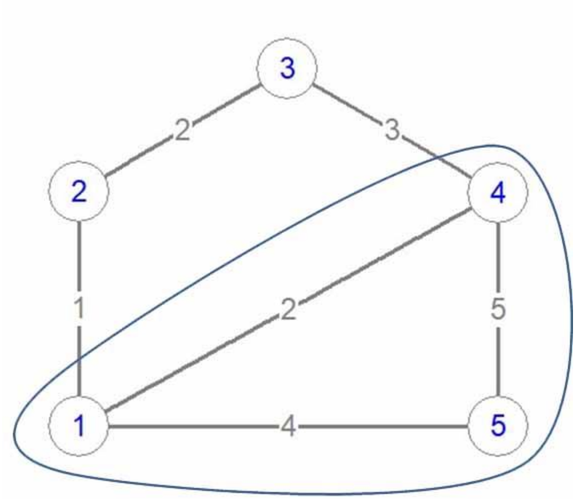

Рис. 1. Пример простейшей системы взаимосвязанных элементов. Ядро выделено плавной кривой

[1] *В [5] доказано, что, несмотря на произвольный выбор элемента с минимальным весом на каждом шаге алгоритма, в случае, если таких элементов несколько, последовательность квазиядер* $\overline{\Gamma}$ *также определяется единственным образом.*

Описанный алгоритм выделяет на этом взвешенном графе ядро, состоящее из вершин 1, 4 и 5 (т.е. ядро $K = \Gamma_p = \{1, 4, 5\}$). При этом строится определяющая последовательность $\langle 2, 3, 1, 4, 5\rangle$, последовательность множеств $\bar{\Gamma} = \left\langle \Gamma_1 = \{1, 2, 3, 4, 5\},\ \Gamma_p = \{1, 4, 5\} \right\rangle$ и последовательность возрастающих порогов $\bar{\varepsilon} = \left\langle \varepsilon_1 = 3, \varepsilon_2 = 6 \right\rangle$.

4. Свойства монотонной системы на матрице связей как основа для выделения подмножества ключевых элементов сети

Важная особенность алгоритма состоит в том, что вместе с определением подмножества «наиболее взаимосвязанных» элементов в ходе решения задачи отыскивается система вложенных друг в друга подмножеств $\bar{\Gamma} = \left\langle \Gamma_1, \Gamma_2, ..., \Gamma_p \right\rangle$ последнее из которых, Γ_p, является искомым ядром. Эти подмножества обладают особыми свойствами [5, 6, 9].

Прежде всего сделаем два общих замечания, вытекающих непосредственно из самого построения алгоритма решения задачи. Во-первых, последовательность значений порога, выстраиваемая в ходе выполнения алгоритма, строго упорядочена по величине

$$\varepsilon_1 < \varepsilon_2 < ... < \varepsilon_j < \varepsilon_{j+1} < ... < \varepsilon_p,$$

где

$$\varepsilon_j = \min_{i \in \Gamma_j} \pi(i, \Gamma_j) = \min_{i \in \Gamma_j} \sum_{k \in \Gamma_j} w_{ik}.$$

Во-вторых, множества последовательности $\bar{\Gamma}$ представляют собой точки локальных экстремумов (максимумов) функции F в последовательности подмножеств $\bar{H}$:

$$F(\Gamma_j) = \max_{\Gamma_j \supset H \supset \Gamma_{j+1}} F(H),\ \forall H \in \bar{H},\ j = 1, 2, ..., p-1,$$

или

$$\left.\begin{array}{l} F(H) \le F(\Gamma_j) = \varepsilon_j \\ F(H) < F(\Gamma_{j+1}) = \varepsilon_{j+1} \end{array}\right| \begin{array}{c} \forall H \in \bar{H},\ \Gamma_j \supset H \supset \Gamma_{j+1}, \\ j = 1, 2, ..., p-1. \end{array} \quad (10)$$

Более того, как это отмечается в [5], соотношения (10) верны для любого подмножества $H, \Gamma_j \supset H \supset \Gamma_{j-1}$, не обязательно являющегося членом последовательности подмножеств $\bar{H}$.

Теорема 1[1]. *Для любого собственного подмножества* $H, \Gamma_j \supset H \supset \Gamma_{j-1}$ *справедливо*

$$\left.\begin{array}{r} F(H) \le F(\Gamma_j) = \varepsilon_j, \\ F(H) < F(\Gamma_{j+1}) = \varepsilon_{j+1}, \end{array}\right| \begin{array}{l} \Gamma_j \supset H \supset \Gamma_{j+1}, \\ j = 1, 2, \ldots, p-1. \end{array}$$

Из теоремы 1 следует, что последовательность множеств $\bar{\Gamma}$ есть последовательность точек строго возрастающих локальных максимумов функции F на множестве всех подмножеств множества узлов V. Иначе говоря, в множестве всех подмножеств множества узлов V как на решетке выделяется последовательность вложенных интервалов подмножеств $H, \Gamma_j \supset H \supset \Gamma_{j-1}$, $j = 1, 2,\ldots, p-1$, границы которых служат локальными максимумами функции F. Последний из этих максимумов является в то же время глобальным. По этой причине подмножества последовательности $\bar{\Gamma}$ называют также квазиядрами.

Таким образом, функция $F(H)$ имеет весьма простую структуру: у нее имеется всего $p \le N = |V|$локальных максимумов. Именно эта простота, являющаяся следствием монотонности системы, определяет возможность построения быстрого алгоритма решения задачи (5): алгоритм можно понимать как простую процедуру последовательного перебора всех локальных максимумов функции $F(H)$, т.е. квазиядер монотонной системы.

Рассмотренная особенность алгоритма выделения ядра монотонной системы позволяет решать исходную задачу с дополнительными ограничениями. Так, в частности, в условиях задачи может быть априорно указан перечень особых ключевых элементов, которые должны обязательно войти в результирующее центральное подмножество; в другом случае могут быть заданы допустимые размеры искомого подмножества (желательный диапазон). В обоих случаях решением будет наименьшее множество Γ_j, удовлетворяющее соответствующему усло-

[1] *Доказательства теорем приведены в* [6].

вию, т.е., например, наименьшее квазиядро Γ_j, включающее априорно заданные ключевые элементы.

С точки зрения содержательной интерпретации результатов, весьма желательно для оценки качества получаемого решения иметь общие количественные характеристики близости (удаленности, связности, интенсивности взаимодействия и т.д.) элементов ядра K и однородности их взаимного расположения (взаимосвязей). В качестве таких характеристик предлагается использовать соответственно величины $F(K)$ и $Q(K)$, которые определяются следующим образом:

(11) $$F(K) = \varepsilon_p = \min_{i \in K} \pi(i, K) = \min_{i \in K} \sum_{k \in K} w_{ik}.$$

(12) $$Q(K) = \max_{H \subset K} \max_{i \in K \setminus H} \sum_{k \in H} w_{ik},\ H \in \bar{H},\ (H \subset K,\ K \setminus H \neq \phi).$$

Величина $F(K)$, как уже отмечалось, определяет особый, крайний элемент ядра – γ_p (центр или антицентр ядра, в зависимости от содержательного смысла весов ребер между узлами), а характеристика $Q(K)$ выделяет некоторую собственную часть H' ядра $K = \Gamma_p$, причем из свойства монотонности следует, что элемент $\alpha_k \in K \setminus H'$, такой, что $\sum_{k \in H'} w_{\alpha k} = \max_{i \in K \setminus H'} \sum_{k \in H'} w_{ik}$, стоит непосредственно перед первым элементом H' в определяющей последовательности. Тем самым, величина $Q(K)$определяет некий максимальный разрез ядра K (в определяющей последовательности).

Целесообразность использования величин $F(K)$ и $Q(K)$ в качестве мер общей близости элементов ядра и однородности их взаимосвязей обосновывается тем, что они удовлетворяют определенным соотношениям.

Теорема 2. *Мера однородности ядра не превосходит величину порога* $\varepsilon_p = F(K)$:

$$Q(K) \le \min_{i \in K} \sum_{k \in K} w_{ik}\ .$$

Иными словами, нельзя ядро, расположенное у правого конца определяющей последовательности, разрезать на две части таким образом, чтобы максимальная суммарная связь элемента из одной части («левой») со всеми элементами другой части («правой») превысила минимальную связь первого элемента ядра со всем ядром $\varepsilon_p = F(K)$.

Теорема 3. *Минимальная сумма связей какого-либо элемента ядра K со всеми остальными его элементами, т.е. величина порога* $\varepsilon_p = F(K)$ *не меньше максимальной величины парной связи между любыми двумя элементами всего множества V*:

$$F(K) \geq \max_{i,k \in V} w_{ik} \cdot$$

В отличие от теоремы 2, гарантирующей в определенном смысле «плотность» ядра K, теорема 3 устанавливает существенное ограничение снизу на величину минимальной связи, т.е. «минимальной взаимосвязанности» элементов в этом множестве.

Введенные с помощью теорем 2 и 3 количественные характеристики взаимосвязанности и однородности (плотности) элементов ядра, т.е. подмножества узлов, на котором функция $F(H)$ имеет глобальный максимум, можно обобщить для оценки всех локальных экстремумов этой функции. Определим их аналогичным образом:

$$F(\Gamma_j) = \min_{i \in \Gamma_j} \pi(i, \Gamma_j) = \min_{i \in \Gamma_j} \sum_{k \in \Gamma_j} w_{ik},\ j = 1, 2, ..., p-1,$$

$$Q(\Gamma_j) = \max_{H} \max_{i \in \Gamma_j \setminus H} \sum_{k \in H} w_{ik},\ \Gamma_j \supset H \supset \Gamma_{j+1},\ H \in \bar{H},\ j = 1, 2, ..., p-1.$$

Теорема 4. *Для любого множества* Γ_j *справедливо следующее неравенство*:

$$Q(\Gamma_j) \leq \min_{i \in \Gamma_j} \sum_{k \in \Gamma_j} w_{ik},\ j = 1, 2, ..., p-1 \cdot$$

Теорема 5. *Для любого множества* Γ_j *минимальная сумма связей его элемента со всеми другими его элементами, т.е. величина порога* $\varepsilon_j = F(\Gamma_j)$, *не меньше максимальной величины связи между двумя элементами i и k, не принадлежащими множеству* Γ_{j+1}, *т.е.* $i, k \in V \setminus \Gamma_{j+1}$,,

$$F(\Gamma_j) \geq \max_{i,k \in V \setminus \Gamma_{j+1}} w_{ik},\ j = 1, 2, ..., p-1.$$

Теоремы 4 и 5 служат обоснованием возможности использования локальных экстремумов функции F, т.е. множеств $\langle \Gamma_j \rangle$, j =1, 2, …, p–1, в качестве приемлемого решения задачи выделения подмножества «наиболее взаимосвязанных» элементов для случая, когда на это подмножество накладываются дополнительные априорные ограничения (см. выше). Введенные в этих теоремах меры «взаимосвязанности» и однородности элементов подмножеств, доставляющих локальные максимумы функции F, и сравнение их с соответствующими характеристиками для глобального максимума указывает, насколько искажается идеальное решение введением дополнительных ограничений.

В заключение рассмотрим теорему 6, устанавливающую, что множества $\langle\Gamma_j\rangle$, $j = 1, 2, \ldots, p-1$, являются в определенном смысле точками локальных экстремумов также и функционала, равного средней попарной связи между элементами подмножества.

Теорема 6. *Пусть* $\forall H \subseteq V$

(13) $$f(H) = \frac{1}{|H|(|H|-1)} \sum_{i,k \in H} w_{ik}, \ \forall H \subseteq V .$$

Тогда справедливы утверждения:

А. *Для любого подмножества* $H \subset \Gamma_j$, $j=1, 2, \ldots, p-1$, *с числом элементов* $|H|$, *равным* $|\Gamma_{j+1}|$, *отличающегося от множества* Γ_{j+1} *не более чем одним элементом, выполняется неравенство*

(14) $$f(H) < f(\Gamma_{j+1}) .$$

Б. *Для любого подмножества* $H \subset \Gamma_j$, $j=1, 2, \ldots, p-1$, *с числом элементов* $|H|$, *равным* $(|\Gamma_{j+1}| + 1)$, *содержащего множество* Γ_{j+1}, *т.е.* $\mathrm{H} \supset \Gamma_{j+1}$, *также выполняется неравенство* (14).

Рассмотренные свойства монотонной системы, построенной на матрице связей, предоставляют широкие возможности для содержательных интерпретаций.

Замечание. Формальная постановка задачи о выделении подмножества «наиболее взаимосвязанных» элементов и алгоритм ее решения полностью сохраняются и тогда, когда элемент w_{ij} матрицы связей акторов носит смысл не близости, меры сходства, интенсивности взаимодействия между ними, а расстояния. Задача в этом случае состоит в выделении подмножества «наиболее изолированных», в смысле (5) элементов множества *V*. Более того, весь предложенный подход полностью применим и к анализу сети с несимметричной матрицей связей узлов, не удовлетворяющей неравенству треугольника, а именно, к анализу потоковой матрицы связей.

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

В качестве примера применения предлагаемого подхода рассмотрим задачу выделения ядра Евросоюза 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 (по паритету покупательной способности), и это при том, что доля стран ЕС в мировом населении составляет не более десяти процентов [20]. Совокупный размер ВВП стран ЕС превышает размер ВВП США [11]. Как известно, важнейшей частью экономики является внешняя торговля. 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 [11, 20]. Однако, несмотря на постоянно развивающийся процесс интеграции, Евросоюз не стал пока полностью однородным образованием. Все чаще идут разговоры о «Европе двух скоростей». 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.

Для выявления особенностей торговых отношений государств Евросоюза внутри ЕС применим рассмотренный выше подход.

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 $H \subseteq V$, функция (1)

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

будет определять суммарный экспорт из страны i во все страны подмножества H.

Таким образом, вес $\pi(i, H)$ каждого элемента подмножества H равен сумме объемов экспортных связей этого элемента со всеми элементами того же подмножества, т.е. величине суммарного потока от одного узла в группу узлов, или мере влияния одного узла на подмножество узлов. То есть в данном случае, как уже отмечалось выше, число $\pi(i, H)$ – это прямой аналог меры степенной центральности (взвешенной), измеренной не по всей сети, а на некотором подмножестве узлов H (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 «наиболее взаимосвязанных (взаимовлияющих)» центральных элементов исходной системы внутренних экспортных связей стран Евросоюза, образующих ее ядро. Таким образом, ядро K включает страны с наибольшим внутригрупповым товарооборотом в смысле критерия (5) 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 (4)) алгоритм выделил следующие страны: *Великобритания, Испания, Италия, Франция, Бельгия, Германия, Нидерланды* (или GB, ES, IT, FR, BE, DE, NL – в двухбуквенной кодировке). 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.

Напомним также, что в предлагаемом подходе очень важен порядок элементов. Более того, важнейшую информацию несет сама определяющая последовательность всех элементов и сопровождающая ее последовательность значений весовой функции $\pi(i, H)$, характеризующей сумму внутригруппового экспорта каждой страны i в страны, расположенные в этой последовательности справа от нее[1] (см. рис. 2):

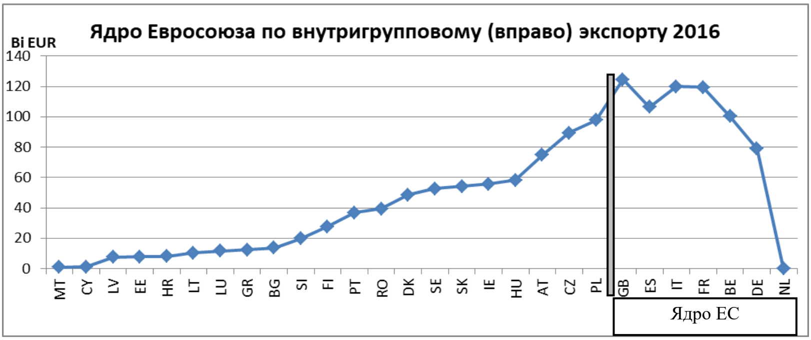

Рис. 2. Ядро Евросоюза по внутригрупповому экспорту 2016

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

[1] *Пусть не смущает значение веса, равное нулю, у Голландии: здесь показан суммарный экспорт каждой страны в группу стран, расположенных в последовательности справа от нее. Голландия была включена в последовательность последней, после Германии: это означает, что когда остались только две страны (Германия и Голландия), то экспорт из Германии в Голландию оказался меньше, чем экспорт из Голландии в Германию, и поэтому она (Германия) была включена в определяющую последовательность раньше Голландии, оставшейся последней.*

группа стран Евросоюза без Великобритании не дает такое же или большее значение функции (4), 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 следующую таблицу[1].

Таблица 1. Суммарный экспорт стран – ядра Евросоюза (млрд евро), 2016 год

Страна	**Суммарный экспорт внутри ядра Евросоюза**	**Суммарный экспорт внутри всего Евросоюза**	**Суммарный экспорт во все страны мира кроме стран ЕС**
Великобритания	124,354	174,980	194,459
Испания	126,087	171,666	86,999
Италия	163,312	232,525	183,856
Франция	218,065	268,807	183,767
Бельгия	216,134	260,131	100,681
Германия	410,200	707,545	499,720
Нидерланды	313,074	391,170	124,015
Итого	**1571,226**	**2206,824**	**1373,497**

Легко видеть, что действительно среди группы стран, составляющих ядро Евросоюза, Великобритания имеет наименьший суммарный экспорт в страны ядра, а Германия, как и можно было ожидать, – наибольший. Видим также, что именно и только Великобритания продает в остальные страны мира кроме ЕС больше (194,459 млрд евро), чем внутри Евросоюза

[1] *Данные о торговле со странами – не членами Евросоюза в этой работе не участвуют в исследуемой системе взаимосвязанных элементов, а используются для интерпретации состава выделенного ядра.*

(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.

Мера однородности ядра $Q(K)$, как оказалось, соответствует разрезу ядра по Италии:

$$Q(K) = 120\,.$$

Ее значение, близкое по величине к максимальному порогу $F(K) = 124{,}4$, как раз и свидетельствует о том, что крайний элемент ядра – Великобритания – по данным внутригруппового экспорта весьма близок к центру ядра.

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

6. Заключение

Для анализа той роли, которую играет каждый объект во взаимосвязанной системе («сети»), в социальных и поведенческих науках, в экономике и других областях широко используется концепция центральности. Сеть при этом представляется взвешенным графом, в котором веса связей определяются мерой близости или расстояния (степенью похожести или различия) между объектами или интенсивностью воздействия (влияния) одного объекта на другого, объёмом материального, информационного или финансового потока между ними и т.п.

Однако классические индексы центральности и распространенные методы выделения ключевых элементов при помощи ранжирования узлов по величине общесетевого индекса центральности не всегда учитывают структуру сети и, прежде всего, интенсивности прямых и *внутригрупповых* взаимодействий в системе.

Индекс центральности при этом рассматривается как функция одного аргумента – узла, то есть вычисляются для каждой вершины *раз и навсегда* – по всем вершинам сети.

Подход к анализу структуры сети взаимосвязанных элементов, *предлагаемый в данной работе*, имеет *три основные особенности*.

Первая. Предлагается использовать контекстно-зависимые меры центральности – на основе внутригруппового взаимодействия элементов сети. *То есть мера (степень) центральности каждого элемента определяется не единожды, по все сети, а зависит от контекста – от того подмножества элементов, для которого рассматривается в данный момент*.

Вторая. 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 значение).

Третья. 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.

В качестве примера применения предлагаемого подхода рассмотрена сеть экспортных связей стран – членов Евросоюза. Выделенное алгоритмом ядро по данным экспорта за 2016 год включает семь стран: Великобритания, Испания, Италия, Франция, Бельгия, Германия, Нидерланды. Первым (особым, крайним) элементом из подмножества ключевых элементов европейского сообщества оказалась Великобритания, единственная из них, которая продает в остальные страны мира кроме ЕС больше, чем внутри Евросоюза.

Литература

1. АЛЕСКЕРОВ Ф. Т., БЛАГОВЕЩЕНСКИЙ Н.Ю. [и др.]. *Влияние и структурная устойчивость в российском парламенте* (1905–1917 *и* 1993–2005 *гг*.) / Н.Ю. Алескеров, Ф. Т. Благовещенский, Г.А. Сатаров, А.В. Соколова, В.И. Якуба, 2-е изд., Москва: Физматлит, 2009. 312 с.
2. АЛЕСКЕРОВ Ф.Т. [и др.]. *Значимость основных российских и международных экономических журналов: сетевой анализ* // Журнал Новой экономической ассоциации. 2016. № 2 (30). С. 193–205.
3. АЛЕСКЕРОВ Ф.Т., МЕЩЕРЯКОВА Н.Г., ШВЫДУН С.В. *Влияние в сетевых структурах с использованием индексов дальних взаимодействий* // XIII Всероссийская школа – конференция молодых ученых «Управление большими системами». 2016. С. 13–22.
4. ЗАКС Ю.М., МУЧНИК И.Б. *Монотонные системы для построения неполных классификаций конечного множества объектов* // Автоматика и телемеханика. 1989. № 4. С. 155–164.
5. КУЗНЕЦОВ Е.Н., МУЧНИК И.Б., ШВАРЦЕР Л.В. *Монотонные системы и их свойства* // Анализ нечисловой информации в социологических исследованиях. 1985. С. 29–57.
6. КУЗНЕЦОВ Е.Н. *Анализ структуры матрицы связей с помощью построения на ней монотонной системы* // Автоматика и телемеханика. 1980. № 7. С. 128–136.
7. ЛЕОНИДОВ А.В., РУМЯНЦЕВ Е.Л. *Оценка системных рисков межбанковского рынка России на основе сетевой топологии* // Журнал Новой экономической ассоциации. 2013. № 3 (19). С. 65–80.
8. МАЛИШЕВСКИЙ А.В. *Свойства порядковых функций множеств* // Качественные модели в теории сложных систем. 1998. С. 169–173.
9. МУЛЛАТ И.Э. *Экстремальные подсистемы монотонных систем. I* // Автоматика и телемеханика. 1976. № 5. С. 130–139.

10. МУЧНИК И.Б., ШВАРЦЕР Л.В. *Субмодулярные функции множеств и монотонные системы в задачах агрегирования. I* // Автоматика и телемеханика. 1987. № 5. С. 135–148.
11. РОДИОНОВА И. А.; УМЕРОВА И.А. *Позиции стран-членов ЕС в мировой и внутрирегиональной торговле* // Зарубежный опыт. 2009. № 134 (41). С. 103–108.
12. ALESKEROV F. [и др.]. *Network analysis of international migration* // Математические методы анализа решений в экономике, бизнесе и политике. 2016. (WP7/2016/0). С. 1–55.
13. ALESKEROV F., ANDRIEVSKAYA I., PERMJAKOVA E. *Key borrowers detected by the intensities of their short-range interactions* // Springer Proceedings in Mathematics and Statistics. 2016. (156). С. 267–280.
14. 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. С. 1–44.
15. BONACICH P. *Power and Centrality: A family of Measures* // American Journal of Sociology. 1987. № 5 (92). С. 1170–1182.
16. BONACICH P., LLOYD P. *Eigenvector-like measures of centrality for asymmetric relations* // Social Networks. 2001. № 3 (23). С. 191–201.
17. LANDHERR A., FRIEDL B., HEIDEMANN J. *A Critical Review of Centrality Measures in Social Networks* // Business & Information Systems Engineering. 2010. № 6 (2). С. 371–385.
18. WASSERMAN S., FAUST K. *Social Network Analysis: Methods and Applications* / S. Wasserman, K. Faust, Cambridge: Cambridge Univercity Press, 1994. 828 с.
19. *Models and Methods in Social Network Analysis*. под ред. P.J.Carrington, J. Scott, S. Wasserman, Cambridge: Cambridge Univercity Press, 2005. 334 с.
20. *World Trade Statistical Review 2016*. WTO, 2017. 165 с.

NETWORK INTERACTIONS STRUCTURE ANALYSIS: CONTEXT-SENSITIVE CENTRALITY MEASURES

Eugene Kuznetsov, Institute of Control Sciences of RAS, Moscow, cand.sc., professor (enkuznetsov@mail.ru).

Abstract: Classical centrality indexes for the analysis of system of interrelating and interacting objects and widespread methods of key elements detection by ranging nodes on the value of common network centrality index not always consider intensity of intra-group interactions in the system. We offer context-dependent measures of centrality – based on intra-group interaction of elements in the network. Namely the proposed centrality measure of each element depends on the subset of elements for which is considered. For general integrated characteristic of importance, influence, etc. of some group of nodes we offer to use minimum or maximum, but not an average, value of the centrality index of nodes in this group. For definition of subset of key nodes in the network, we offer to use algorithms of special cluster analysis – algorithms for monotone system kernel detection. It not only provides determination of global extremum of functional in the corresponding optimization task, but also allows to analyze structure of network in more detail. As an example of application of the offered approach, the network of export links of the member countries of the European Union is considered. The special front boundary role of United Kingdom in the core of the European Union is revealed.

Keywords: centrality measures, key elements of the interconnected objects system, monotone systems.

УДК 519. 87 + 51-77
ББК 22.176

Статья представлена к публикации членом редакционной коллегии Ф.Т. Алескеровым.

Поступила в редакцию 17.01.2019.
Опубликована 31.07.2019.