

**Значимость основных международных
и российских экономических журналов:
сетевой анализ**

**Ф.Т. Алескеров
НИУ ВШЭ и ИПУ РАН**

Авторы

- **Ф.Т. Алескеров** НИУ ВШЭ, ИПУ РАН, Москва
- **Д.Н. Бадгаева** НИУ ВШЭ, Москва
- **В.В. Писляков** НИУ ВШЭ, Москва
- **И.А. Стерлигов** НИУ ВШЭ, Москва
- **С.В. Швыдун** НИУ ВШЭ, ИПУ РАН, Москва

Рассматривается сеть, вершинами которой являются журналы, а на ребрах находится информация о числе цитирований одного журнала на другой.

В ряде работ (Leydesdorff, 2006, 2007, 2008) был предложен метод анализа перекрестного цитирования (далее CrC), основанный на построении матрицы журналов с последующим выявлением сети коммуникаций между ними. Цитирования работ одних журналов работами других журналов образуют сеть, связывающую все издания в единый комплекс (Price, 1980, p. 7).

Сеть может быть определена так же по совместному цитированию (co-citation), в основе которого лежит принцип выделения взаимосвязи между двумя публикациями на основе одновременного цитирования этих публикаций в третьей (Small, 1973; McCain, 1991).

При построении матрицы цитирований для сетевого анализа мы использовали окно публикации 2011–2015 гг., окно цитирования 2011–10 апреля 2016 г. Учитывались все типы публикаций.

Основные понятия

Мы работаем с графом $G = (V, E)$, где $V = \{1, \dots, n\}$ набор вершин, $E \subseteq V \times V$ набор взвешенных дуг (дуга $(i, j) \in E$). Мы рассматриваем направленные графы, т.е. графы, где существование дуги (i, j) не подразумевает существование дуги (j, i) . Однако, наш подход к оценке мощности также может быть применен к простым сетям (ненаправленным, невзвешенным).

Классические меры центральности

Центральность по степени: кол-во инциденций

Центральность по собственному вектору: λ_{\max} : $Ax = \lambda_{\max}x$

Не только соседние, но также длинные связи

Центральность по посредничеству: $B_i = \sum_{uv} \sigma_{uv}(i) / \sigma_{uv}$

Число кратчайших путей, проходящих через узел

Центральность по близости: $C_i = 1 / \sum_j d_{ij}$

Чем больше близких вершин, тем лучше

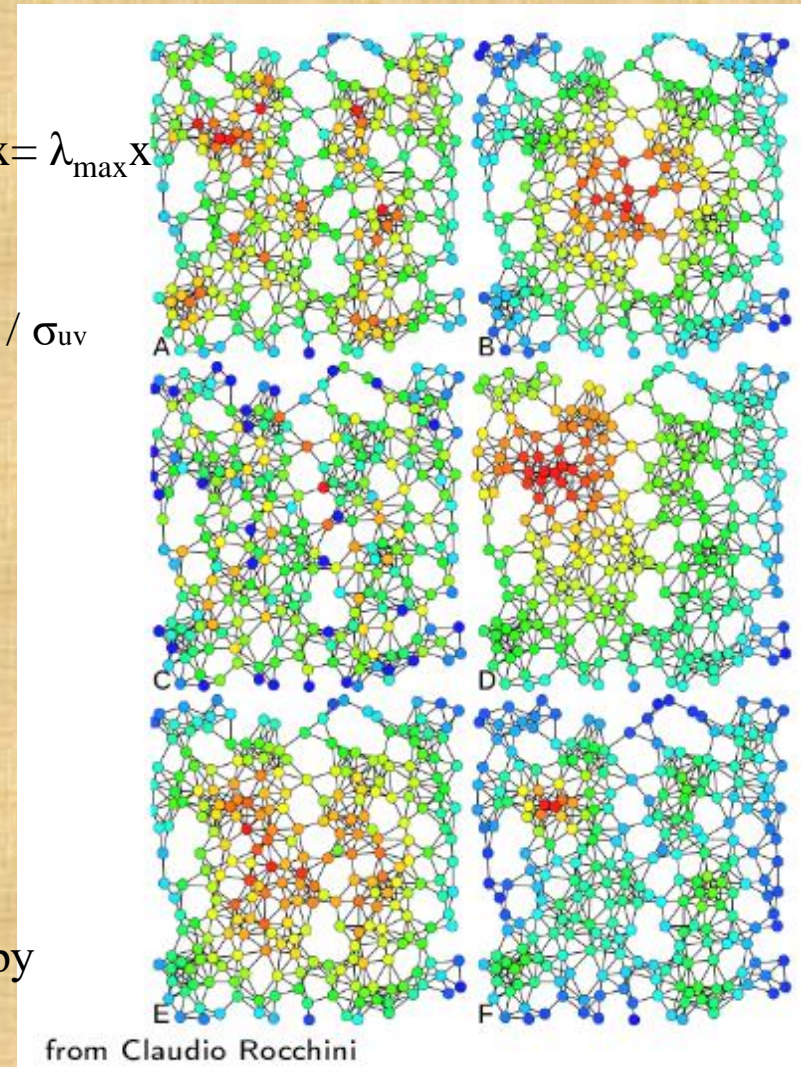
Центральность PageRank: $x_i = \alpha \sum_j A_{ij} x_j / n_j + \beta$

Рассматриваются степени соседей

Центральность по Кацу: $x = \alpha Ax + \beta 1$

Обобщенная центральность по собственному вектору

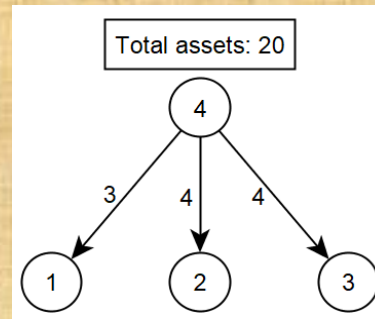
(штрафы, нет ненулевых решений)



Основные понятия

- Каждая вершина определяется вектором параметров, скажем, $\bar{\alpha}^i = (\alpha_1^i, \dots, \alpha_s^i)$ для каждой вершины $i \in V$. Эта возможность была упомянута Ньюманом [Newman 2003]: “И вершины или дуги могут иметь множество свойств, числовых или иных, связанных с ними”.
- Для социальных сетей, где дуги представляют людей, эти параметры могут быть возрастом респондента, ее годовым доходом, уровнем образования, населением поселения, где она живет, политическими взглядами (скажем, на левой-правой шкале), etc.
- В банковской сети, где вершины являются банками, этими параметрами могут быть общие активы, общая сумма кредитов, рискованные кредиты, просроченные кредиты, etc.
- Каждая вершина ij также описывается набором параметров $\bar{w}_{ij} = (w_{ij}^1, \dots, w_{ij}^r)$, которые могут представлять различные детали соединения.

Пример



Малая банковская система

$q_4 = 0.5 * \text{совокупные}$
активы = 10

Пусть совокупные активы банка 4 (то есть кредитор в сети) равны 20. Допустим, кредитор предоставляет кредит трем другим банкам (банки 1, 2 и 3) и страдает от убытка, если заемщики не будут выплачивать 50% своих общих активов (отсюда, $q_4 = 20 \cdot 50\% = 10$). В этом случае каждый отдельный заемщик 1, 2, 3 с кредитами $w_{41} = 3$, $w_{42} = 4$, $w_{43} = 4$ не создают проблему для банка 4, т.к. $w_{4k} < q_4$ для $k = 1, 2, 3$. Тем не менее, $w_{41} + w_{42} + w_{43} > q_4$.

Критическая группа. Ключевые вершины.

Простой путь.

Определение. Критическая группа $\Omega(j)$ для узла j является подмножеством вершин, чье групповое воздействие превышает некоторый порог q_j . Более формально $\Omega(j) \subseteq V$ является критической группой узла j , если

$$\sum_{i \in \Omega(j)} w_{ij} \geq q_j$$

Определение. Вершина является ключевой для некоторой группы, если ее исключение из этой группы делает группу некритической. В одной группе может быть несколько ключевых вершин. Формально $\Omega^p(j) \subseteq \Omega(j)$ является подмножеством ключевых вершин группы $\Omega(j)$, если

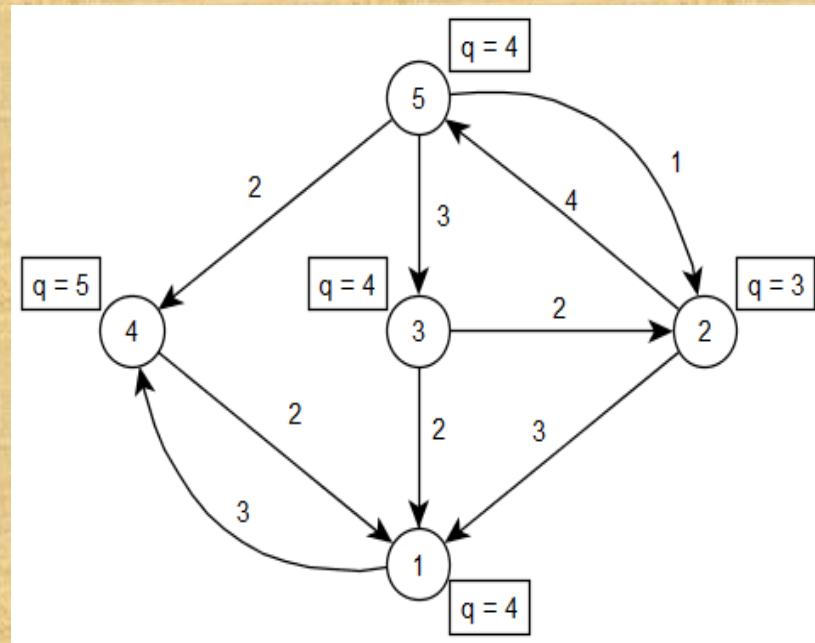
$$\forall k \in \Omega^p(j) \quad \sum_{i \in \Omega(j) \setminus \{k\}} w_{ij} < q_j$$

Определение. Простой путь между вершинами i и j в графе G представляет собой последовательность ребер, которая соединяет их и содержит различные вершины. Точнее,

$$(i, k_1), \quad (k_1, k_2), \quad (k_2, k_3), \quad \dots, \quad (k_{s-1}, j), \\ i \neq k_1 \neq \dots \neq k_{s-1} \neq j$$

- простой путь длины s , обозначаемый как $P_{i-j}^t(s)$, где t - номер пути.

Ближнее взаимодействие. Пример.



Список критических групп и ключевых вершин в этих группах для вершины 1

Критические группы, $\Omega_k(1)$ где k – номер группы	Ключевые вершины, $\Omega_k^p(1)$ где k – номер группы
$\Omega_1(1) = \{2, 3\}$	$\Omega_1^p(1) = \{2, 3\}$
$\Omega_2(1) = \{2, 4\}$	$\Omega_2^p(1) = \{2, 4\}$
$\Omega_3(1) = \{3, 4\}$	$\Omega_3^p(1) = \{3, 4\}$
$\Omega_4(1) = \{2, 3, 4\}$	$\Omega_4^p(1) = \{\emptyset\}$
$\Omega_5(1) = \{2, 3, 5\}$	$\Omega_5^p(1) = \{2, 3\}$
$\Omega_6(1) = \{2, 4, 5\}$	$\Omega_6^p(1) = \{2, 4\}$
$\Omega_7(1) = \{3, 4, 5\}$	$\Omega_7^p(1) = \{3, 4\}$
$\Omega_8(1) = \{2, 3, 4, 5\}$	$\Omega_8^p(1) = \{\emptyset\}$

Прямые и косвенные взаимодействия

Нормализованные интенсивности прямых воздействий и косвенных воздействий через одну промежуточную вершину.

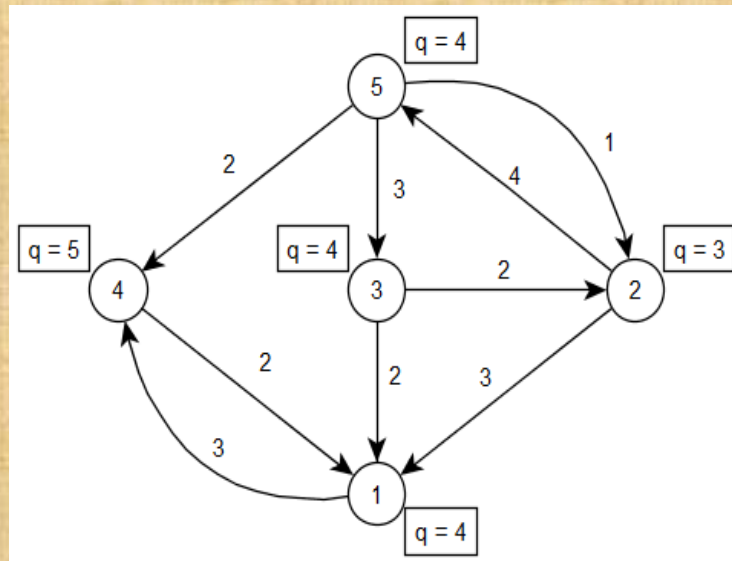
Обратите внимание, что косвенное воздействие рассчитывается только в том случае, если существует прямая связь с вершиной 1.

$p_{ij} = \frac{w_{ij}}{\sum_k w_{kj}}$ – интенсивность прямого воздействия вершины i на вершину j ,

$$p_{ihj} = \begin{cases} \frac{w_{ih}}{\sum_k w_{kj}}, & \text{if } w_{ij} > 0, \quad w_{hj} > 0, \quad w_{hj} \geq w_{ih}, \quad i \neq h \neq j, \\ \frac{w_{hj}}{\sum_k w_{kj}}, & \text{if } w_{ij} > 0, \quad w_{hj} > 0, \quad w_{hj} < w_{ih}, \quad i \neq h \neq j, \\ 0, & \text{иначе.} \end{cases} \quad (4)$$

p_{ihj} – интенсивность косвенного воздействия вершины i на вершину j через вершину h .

Прямое и косвенное взаимодействия



Прямые интенсивности воздействий с вершиной 1 равны $p_{21} = 3/7$; $p_{31} = 2/7$; $p_{41} = 2/7$ и косвенные интенсивности влияний с вершиной 1 равны $p_{321} = 2/7$.

Интенсивности влияний ключевых вершин

Зная критические группы, ключевые вершины и нормализованные интенсивности соединений, можно вычислить интенсивности в пределах каждой критической группы для каждой ключевой вершины узла.

Мы также обращаем внимание на размер группы (чем больше группа, тем меньше вклад, который имеет каждая ключевая вершина).

Следовательно, введем интенсивность влияния ключевой вершины i на j внутри группы $\Omega_k(j)$, обозначаемой $f(i, \Omega_k(j))$

$$f(i, \Omega_k(j)) = \frac{p_{ij} + \sum_{h \in \Omega_k(j)} p_{ihj}}{|\Omega_k(j)|}, i \in \Omega_k^p(j).$$

Интенсивность воздействия на вершину 1 внутри группы $\Omega_k(j)$

Критические группы, $\Omega_k(1)$ где k - число групп	Ключевые вершины, $\Omega_k^p(1)$ где k - число групп	Интенсивность влияния ключевого узла i на узел j в критической группе $\Omega_k(j) \rightarrow f(i, \Omega_k(j))$
$\Omega_1(1) = \{2, 3\}$	$\Omega_1^p(1) = \{2, 3\}$	<ul style="list-style-type: none"> • $f(2, \Omega_1(1)) = \frac{3/7}{2} = 3/14$ • $f(3, \Omega_1(1)) = \frac{2/7+2/7}{2} = 4/14$
$\Omega_2(1) = \{2, 4\}$	$\Omega_2^p(1) = \{2, 4\}$	<ul style="list-style-type: none"> • $f(2, \Omega_2(1)) = \frac{3/7}{2} = 3/14$ • $f(4, \Omega_2(1)) = \frac{2/7}{2} = 2/14$
$\Omega_3(1) = \{3, 4\}$	$\Omega_3^p(1) = \{3, 4\}$	<ul style="list-style-type: none"> • $f(3, \Omega_3(1)) = \frac{2/7}{2} = 2/14$ • $f(4, \Omega_3(1)) = \frac{2/7}{2} = 2/14$
$\Omega_4(1) = \{2, 3, 4\}$	$\Omega_4^p(1) = \{\emptyset\}$	-
$\Omega_5(1) = \{2, 3, 5\}$	$\Omega_5^p(1) = \{2, 3\}$	<ul style="list-style-type: none"> • $f(2, \Omega_5(1)) = \frac{3/7}{3} = 3/21$ • $f(3, \Omega_5(1)) = \frac{2/7+2/7}{3} = 4/21$
$\Omega_6(1) = \{2, 4, 5\}$	$\Omega_6^p(1) = \{2, 4\}$	<ul style="list-style-type: none"> • $f(2, \Omega_6(1)) = \frac{3/7}{3} = 3/21$ • $f(4, \Omega_6(1)) = \frac{2/7}{3} = 2/21$
$\Omega_7(1) = \{3, 4, 5\}$	$\Omega_7^p(1) = \{3, 4\}$	<ul style="list-style-type: none"> • $f(3, \Omega_7(1)) = \frac{2/7}{3} = 2/21$ • $f(4, \Omega_7(1)) = \frac{2/7}{3} = 2/21$
$\Omega_8(1) = \{2, 3, 4, 5\}$	$\Omega_8^p(1) = \{\emptyset\}$	-

После того, как мы получим интенсивности воздействий вершины i на вершину j во всех критических группах вершины j , где i является ключевой, нам необходимо агрегировать их в одно значение или, другими словами, в воздействие между вершинами. Модель агрегирования, представленная в [Aleskerov et al., 2014], выглядит следующим образом

$$\chi_i(j) = \sum_k f(i, \Omega_k(j)).$$

Наконец, эти значения нормируются к их сумме

$$\hat{\chi}_i(j) = \frac{\chi_i(j)}{\sum_k \chi_k(j)}.$$

Для нашего примера

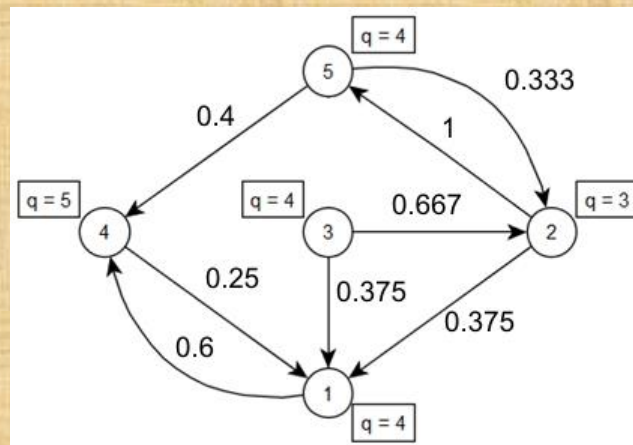
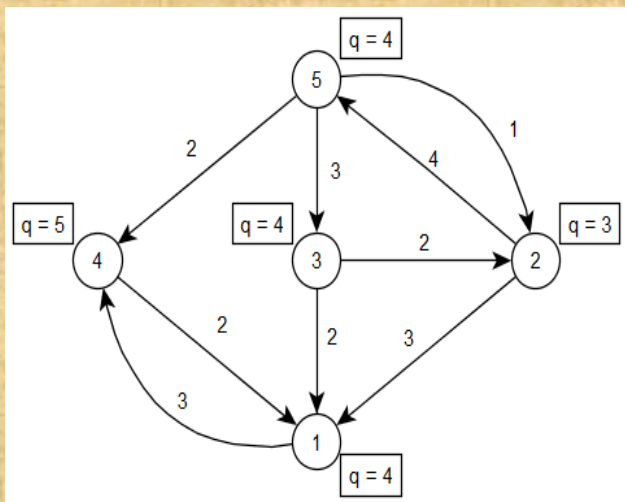
- $\chi_2(1) = \frac{3}{14} + \frac{3}{14} + \frac{3}{21} + \frac{3}{21} = \frac{15}{21}$; $\chi_3(1) = \frac{4}{14} + \frac{2}{14} + \frac{4}{21} + \frac{2}{21} = \frac{15}{21}$;
- $\chi_4(1) = \frac{2}{14} + \frac{2}{14} + \frac{2}{21} + \frac{2}{21} = \frac{10}{21}$; $\chi_5(1) = 0$.

Нормированные значения равны

- $\hat{\chi}_2(1) = \frac{15/21}{15/21+15/21+10/21+0} = \frac{15}{40}$; $\hat{\chi}_3(1) = \frac{15/21}{15/21+15/21+10/21+0} = \frac{15}{40}$;
- $\hat{\chi}_4(1) = \frac{10/21}{15/21+15/21+10/21+0} = \frac{10}{40}$; $\hat{\chi}_5(1) = 0$.

Полученные значения могут интерпретироваться как новые веса на дугах между i -ой вершиной и j -ой вершиной: это окончательное модифицированное воздействие вершины i на вершину j , которое включает как прямое, так и косвенное воздействие индивидуально или через группы с другими вершинами с учетом индивидуальных параметров вершин.

Влияние вершины на вершину для численного примера



Это влияние вершины на вершину может быть преобразовано в одно значение, соответствующее каждой вершине (мера влияния на графике) - индекс SRIC

α_{SRIC} , который введен в [Aleskerov et al., 2014]. Идея заключается в том, что воздействие на конкретную вершину важнее, если эта вершина управляет многими другими вершинами.

α_{SRIC} индекс в соответствии с нормированной сверткой по степеням исхода

$$\alpha_{SRIC}(1) = 0.19; \alpha_{SRIC}(2) = 0.34; \alpha_{SRIC}(3) = 0.22; \\ \alpha_{SRIC}(4) = 0.06; \alpha_{SRIC}(5) = 0.19.$$

Другое определение $\alpha_{SRIC}(i)$

Другим вариантом является среднее значение для каждой вершины, т.е.

$$\alpha_{SRIC}(i) = \frac{\sum_{j \in V} \hat{\chi}_i(j)}{|V|}.$$

Для нашего примера

$$\alpha_{SRIC}(1) = 0.12; \alpha_{SRIC}(2) = 0.28; \alpha_{SRIC}(3) = 0.21;$$
$$\alpha_{SRIC}(4) = 0.05; \alpha_{SRIC}(5) = 0.15.$$

Недостатки SRIC

- требует наличия прямой связи для оценки косвенного воздействия, которое обычно приводит к недооценке реального воздействия: в нашем примере вершина 5 может воздействовать на вершину 1 по вершине 4, но SRIC не чувствителен к ней;
- учитывает только соединения по дальним воздействиям и не учитывает воздействие более чем одного промежуточной дуги;
- процедуры формирования групп включают косвенно связанные вершины, и это выглядит нецелесообразным, поскольку эти вершины не будут воздействовать на рассматриваемую.

Важную роль в данной задаче играет выбор пороговых значений для индексов ближних и дальних взаимодействий. При анализе сети цитирований были выбраны пороговые значения, равные 10, 50 и 80% общего числа цитирований конкретного журнала.

Низкие пороговые значения повышают влияние тех журналов, которые цитируют большое число других журналов с учетом их важности.

Высокие пороговые значения позволяют выявить те журналы, общее число цитирований которых составляют подавляющую долю цитирований конкретного журнала. Логично предположить, что высокие значения порогов могут выявить некоторые группы журналов, которые по большей части ссылаются друг на друга.

Мы выбрали первые 100 из 333 журналов экономической направленности, которые по версии JCR являются наиболее авторитетными по IF. Прежде всего, стоит упомянуть "Quarterly Journal of Economics" (двухлетний импакт-фактор которого равен 6,654, а пятилетний — 9,794), который занимает первое место среди журналов в категории. Последний в нашей выборке — "Journal of Agricultural Economics" (двухлетний импакт-фактор которого равен 1,258, а пятилетний — 1,898). Использовались данные JCR за 2011–2014 гг

Ранжирование при $q=10\%$

- AER
- Journal of Banking and Finance
- J of Economic Behavior and Organization
- Energy Economics
- Ecological Economics
- World Development

По всем шести версиям LRIC при $q = 50\%$ в ТОП-3 вошел только один журнал — "Energy Economics", причем в пяти случаях заняв первую позицию (кроме индекса Simul — второе место); по четырем индексам — "Ecological Economics" (за исключением SumPath — четвертое место и Simul — 16-е место) и "Journal of Banking & Finance" (за исключением MaxMin — четвертое место и Simul — 12-е место).

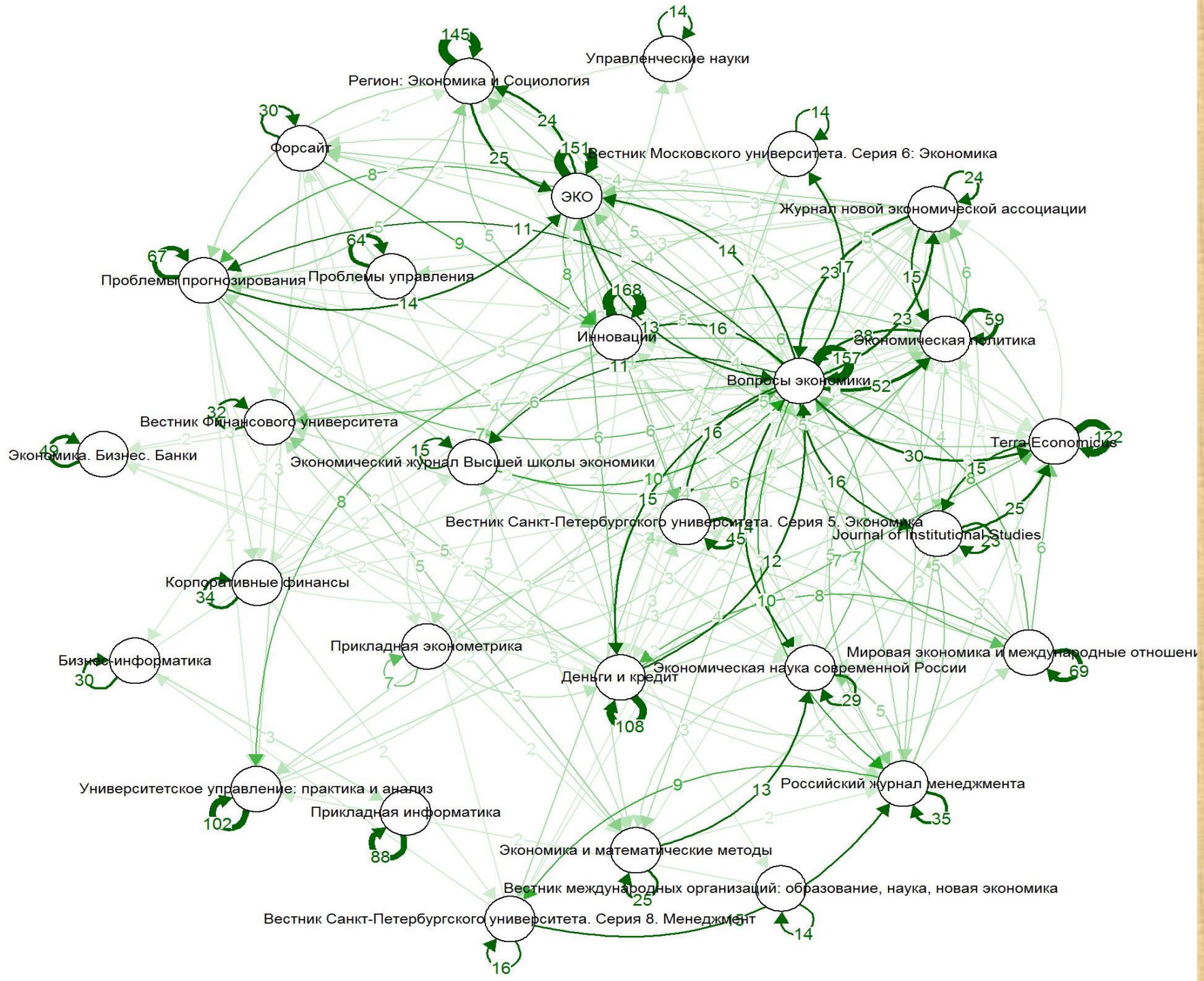
В число лучших при высоком пороговом значении $q = 80\%$ вошли "Ecological Economics" (первое место по всем индексам, кроме SumPaths и Simul), "Transportation Research Part A-policy and Practice", "Transport Policy", "Journal of Transport Geography", "Cambridge Journal of Economics", "Pharmacoeconomics".

Ранги журналов по индексам центральности

№ п/п	Название журнала	Классический индекс			SRIC			LRIC																	
		Weighted In-degree	PageRank	Eigenvector	10%	50%	80%	SumPaths			MaxPath			MaxMin			MaxT			MultT			Simul		
								10%	50%	80%	10%	50%	80%	10%	50%	80%	10%	50%	80%	10%	50%	80%	10%	50%	80%
1	American Economic Review	1	1	4	1	32	39	25	22	14	9	12	12	12	20	13	12	18	13	10	12	12	4	6	-
2	Journal of Financial Economics	2	3	3	9	9	46	57	20	40	46	23	40	40	33	42	42	27	42	45	23	40	22	-	-
3	Journal of Finance	3	4	6	30	43	52	63	30	-	54	28	-	59	38	-	60	37	-	55	37	-	24	-	-
4	Energy Economics	4	20	2	4	2	5	4	1	15	1	1	15	1	1	16	1	1	16	1	1	15	5	2	-
5	Transportation Research Part A-policy and Practice	11	30	18	18	11	12	1	5	1	17	6	2	20	5	3	17	5	3	17	6	5	32	-	-
6	Journal of Banking & Finance	10	34	1	2	1	10	18	2	6	4	2	4	4	4	6	4	3	6	4	2	4	10	12	-
7	Ecological Economics	12	26	7	5	3	3	8	4	4	3	3	1	3	2	1	3	2	1	3	3	1	3	-	-
8	Journal of Economic Behavior & Organization	27	44	9	3	13	60	9	15	34	6	18	33	6	15	36	6	23	35	5	20	33	8	-	-
9	World Development	15	32	26	6	5	7	5	3	7	5	4	6	5	3	5	5	4	5	6	4	3	2	3	-
10	American Journal of Agricultural Economics	40	45	45	59	52	45	2	34	39	22	33	38	21	23	35	23	35	34	22	32	38	53	-	-
11	Journal of Economic Surveys	78	91	78	43	47	31	3	14	16	16	10	16	14	11	11	15	14	11	16	11	16	51	14	-
12	Transport Policy	25	56	41	52	61	29	12	9	2	38	9	3	57	10	2	54	10	2	47	9	8	23	-	-
13	Journal of Transport Geography	6	24	12	14	7	2	28	17	3	44	11	7	63	13	4	68	12	4	43	10	6	33	4	2
14	International Review of Economics & Finance	20	75	8	38	29	19	29	25	45	2	25	45	2	25	46	2	16	46	2	25	45	50	-	-
15	Technological and Economic Development of Economy	69	50	81	51	37	24	73	-	-	91	-	-	97	-	-	95	-	-	92	-	-	1	15	-
16	Cambridge Journal of Economics	65	39	84	49	40	25	69	91	28	70	91	28	80	-	29	71	91	27	66	91	28	88	-	1
17	Econometrica	7	2	10	23	59	40	61	57	29	52	54	29	47	59	32	46	54	31	52	54	29	40	7	-
18	Pharmacoeconomics	43	71	95	11	4	1	79	11	5	53	17	5	50	27	7	50	20	7	51	16	2	16	-	-
19	Journal of Regional Science	60	81	74	17	8	11	24	27	17	16	22	17	34	18	18	33	15	18	25	22	17	25	1	-

Российские журналы

- Источник: Научная электронная библиотека eLIBRARY.RU (<http://elibrary.ru/>)
- Вершины: 29 журналов, отобранных экспертами РИНЦ в рамках проекта "1000 российских журналов на платформе Web of Science»
- Связи между вершинами: общее число цитирований за 2011-2016 годы на статьи 2011-2015 года



Результаты: классические индексы

Индексы центральности	In-degree	Weighted In-degree	Weighted Out-degree	Weighted Degree	PageRank	Eigenvector
Вопросы экономики	1	1	1	1	4	1
ЭКО	2	2	2	2	5	2
Инновации	4	3	3	3	2	3
Terra Economicus	4	4	6	5	1	6
Экономическая политика	3	6	7	7	11	5
Деньги и кредит	11	7	5	6	20	7
Журнал Новой экономической ассоциации	7	17	11	12	26	8
Экономический журнал Высшей школы экономики	16	23	22	23	27	14
Вестник Санкт-Петербургского университета. Серия 8 Менеджмент	16	25	24	24	18	23
Бизнес-информатика	23	26	24	25	17	27

Результаты: наши индексы (q=10%)

Индексы центральности	SRIC	LRIC (Simul)	LRIC (Sum)	LRIC (Max)	LRIC (MaxMin)	LRIC (MaxT)	LRIC (MultT)
Вопросы экономики	1	15	18	3	3	3	3
ЭКО	4	18	22	8	8	8	8
Инновации	13	19	23	20	20	20	19
Terra Economicus	18	24	24	21	21	21	21
Экономическая политика	7	9	9	1	1	1	1
Деньги и кредит	2	17	20	6	6	6	6
Журнал Новой экономической ассоциации	3	13	13	2	2	2	2
Экономический журнал Высшей школы экономики	16	1	1	9	9	9	9
Вестник Санкт-Петербургского университета. Серия 8 Менеджмент	17	2	2	12	12	12	12
Бизнес-информатика	10	23	25	26	25	24	26

Результаты: наши индексы (q=50%)

Индексы центральности	SRIC	LRIC (Simul)	LRIC (Sum)	LRIC (Max)	LRIC (MaxMin)	LRIC (MaxT)	LRIC (MultT)
Вопросы экономики	1	1	1	1	1	1	1
ЭКО	3	8	4	4	5	5	4
Инновации	12	11	19	17	20	21	17
Terra Economicus	20	21	18	21	21	24	21
Экономическая политика	6	15	3	3	3	2	3
Деньги и кредит	5	6	5	5	7	6	5
Журнал Новой экономической ассоциации	2	2	2	2	2	3	2
Экономический журнал Высшей школы экономики	17	17	8	10	11	10	10
Вестник Санкт-Петербургского университета. Серия 8 Менеджмент	15	3	14	11	9	9	11
Бизнес-информатика	7	7	24	23	23	20	22

Результаты: наши индексы (q=80%)

Индексы центральности	SRIC	LRIC (Simul)	LRIC (Sum)	LRIC (Max)	LRIC (MaxMin)	LRIC (MaxT)	LRIC (MultT)
Вопросы экономики	1	2	1	1	1	1	1
ЭКО	9	11	11	10	15	14	10
Инновации	13	15	16	15	19	19	15
Terra Economicus	19	19	21	21	21	21	21
Экономическая политика	5	10	6	8	5	5	8
Деньги и кредит	2	4	4	4	6	6	4
Журнал Новой экономической ассоциации	3	9	2	2	2	2	2
Экономический журнал Высшей школы экономики	16	20	17	17	10	10	17
Вестник Санкт-Петербургского университета. Серия 8 Менеджмент	14	6	13	11	16	15	11
Бизнес-информатика	4	1	7	6	14	13	6

Индексы центральности для сети цитирования

Индексы центральности		№ п/п, название журнала										
		1	2	3	4	5	6	7	8	9	10	
		Вопросы экономики	ЭКО	Инновации	Terra Economicus	Экономическая политика	Деньги и кредит	Журнал Новой экономической ассоциации	Экономический журнал Высшей школы экономики	Вестник Санкт-Петербургского университета. Серия 8 Менеджмент	Бизнес-информатика	
Классические индексы	In-degree	1	2	4	4	3	11	7	16	16	23	
	Weighted In-degree	1	2	3	4	6	7	17	23	25	26	
	Weighted Out-degree	1	2	3	6	7	5	11	22	24	24	
	Weighted Degree	1	2	3	5	7	6	12	23	24	25	
	PageRank	4	5	2	1	11	20	26	27	18	17	
	Eigenvector	1	2	3	6	5	7	8	14	23	27	
Индекс ближних взаимодействий, SRI	10%	1	4	13	18	7	2	3	16	17	10	
	50%	1	3	12	20	6	5	2	17	15	7	
	80%	1	9	13	19	5	2	3	16	14	4	
Индексы дальних взаимодействий, LRI	Simul	10%	15	18	19	24	9	17	13	1	2	23
		50%	1	8	11	21	15	6	2	17	3	7
		80%	2	11	15	19	10	4	9	20	6	1
	Sum	10%	18	22	23	24	9	20	13	1	2	25
		50%	1	4	19	18	3	5	2	8	14	24
		80%	1	11	16	21	6	4	2	17	13	7
	Max	10%	3	8	20	21	1	6	2	9	12	26
		50%	1	4	17	21	3	5	2	10	11	23
		80%	1	10	15	21	8	4	2	17	11	6
	MaxMin	10%	3	8	20	21	1	6	2	9	12	25
		50%	1	5	20	21	3	7	2	11	9	23
		80%	1	15	19	21	5	6	2	10	16	14
	MaxT	10%	3	8	20	21	1	6	2	9	12	24
		50%	1	5	21	24	2	6	3	10	9	20
		80%	1	14	19	21	5	6	2	10	15	13
	MultT	10%	3	8	19	21	1	6	2	9	12	26
		50%	1	4	17	21	3	5	2	10	11	22
		80%	1	10	15	21	8	4	2	17	11	6

Заключение

Использование SRIC и LRIC позволяет определить, как смещаются акценты в распределении рангов журналов. Предложенный подход может быть применен к различным категориям научных журналов для выявления системно-значимых изданий.

References

Aleskerov F.T., Andrievskaya I.K., Permjakova E.E. (2014). Key Borrowers Detected by the Intensities of their Short-Range Interactions. Working papers by NRU Higher School of Economics. Series FE “Financial Economics”. No. WP BRP 33/FE/2014. Available at: <http://ssrn.com/abstract=2479272> (accessed: May 2016).

Aleskerov F., Meshcheryakova N., Shvydun S. (2017) Power in Network Structures. In: Kalyagin V., Nikolaev A., Pardalos P., Prokopyev O. (eds) Models, Algorithms, and Technologies for Network Analysis. NET 2016. Springer Proceedings in Mathematics & Statistics, vol 197. Springer, Cham. First Online: 24 June 2017. DOI: 10.1007/978-3-319-56829-4_7

Leydesdorff L. (2006). Can Scientific Journals Be Classified in Terms of Aggregated Journal–Journal Citation Relations Using the Journal Citation Reports? [Электронный ресурс] // Journal of the American Society for Information Science & Technology. Vol. 57(5). P. 601–613. Режим доступа: <http://dx.doi.org/10.1002/asi.20322>, свободный. Загл. с экрана. Яз. англ. (дата обращения: май 2016 г.).

Leydesdorff L. (2007). Visualization of the Citation Impact Environments of Scientific Journals: An Online Mapping. [Электронный ресурс] // Journal of the American Society for Information Science and Technology (JASIST). Vol. 58(1). P. 25–38. Режим доступа: <http://dx.doi.org/10.1002/asi.20406>, свободный. Загл. с экрана. Яз. англ. (дата обращения: май 2016 г.).

Leydesdorff L. (2008). Caveats for the Use of Citation Indicators in Research and Journal Evaluations. [Электронный ресурс] // Journal of the American Society for Information Science and Technology (JASIST). Vol. 59(2). P. 278–287. Режим доступа: <http://dx.doi.org/10.1002/asi.20743>, свободный. Загл. с экрана. Яз. англ. (дата обращения: май 2016 г.).

Leydesdorff L., Bornmann L. (2011). Integrated Impact Indicators Compared with Impact Factors: An Alternative Research Design with Policy Implications // Journal of the American Society for Information Science and Technology (JASIST). Vol. 62(11). P. 2133–2146. doi:10.1002/asi.21609.

McCain K.W. (1991). Mapping Economics Through the Journal Literature: An Experiment in Journal Cocitation Analysis // Journal of the American Society for Information Science. Vol. 42. P. 290–296.

Price D. de Solla (1980). Foreword. Essays of an Information Scientist // Philadelphia. Vol. 3. P. 5–9.

Newman M.E.J. (2010). Networks: An Introduction. Oxford: Oxford University Press.

Small H. (1973). Co-Citation In Scientific Literature – New Measure of Relationship between 2 Documents // Journal of the American Society for Information Science. Vol. 24. Issue 4. P. 265–269.

Waltman L. (2016). A Review of the Literature on Citation Impact Indicators // Journal of In-formetrics. Vol. 10(2). P. 365–391. doi:10.1016/j.joi.2016.02.007.

Спасибо!

20, Myasnitskaya str., Moscow, Russia, 101000

Tel.: +7 (495) 628-8829, Fax: +7 (495) 628-7931

www.hse.ru

Связанная литература: Классические центральные меры

Центральность по степени: кол-во инциденций

Центральность по собственному вектору: λ_{\max} : $Ax = \lambda_{\max}x$

Не только соседние, но также длинные связи

Центральность по посредничеству: $B_i = \sum_{uv} \sigma_{uv}(i) / \sigma_{uv}$

Число кратчайших путей, проходящих через узел

Центральность по близости: $C_i = 1 / \sum_j d_{ij}$

Чем больше близких вершин, тем лучше

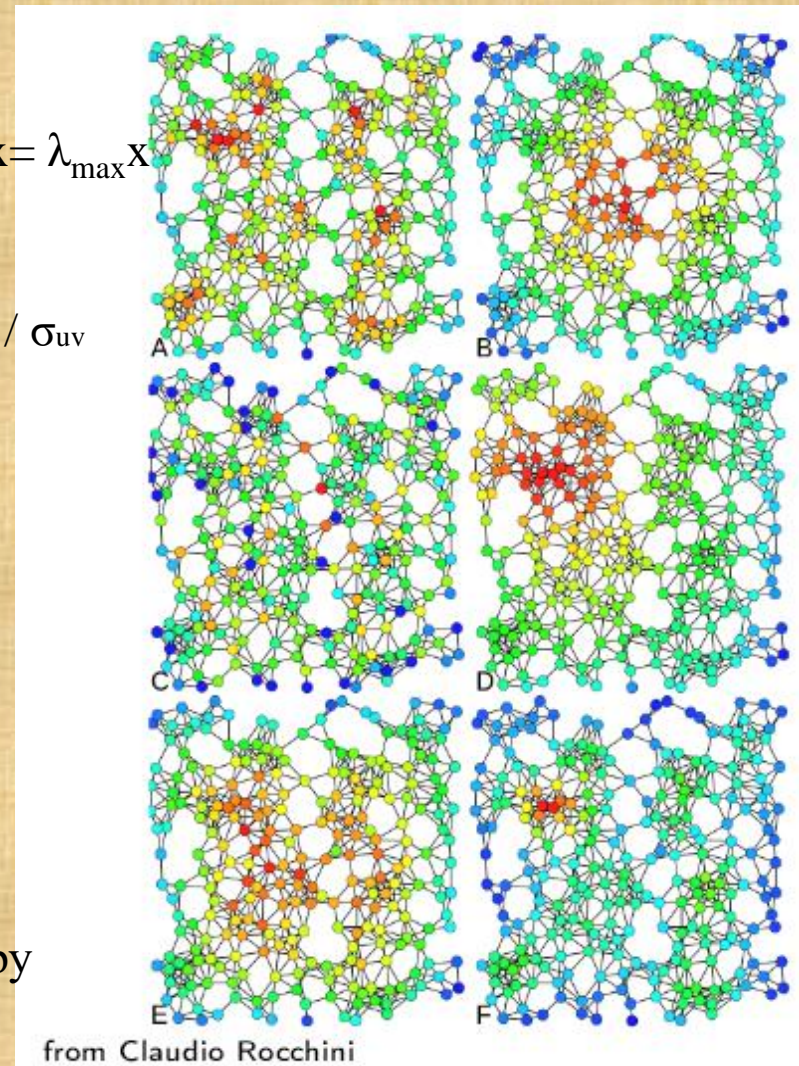
Центральность PageRank: $x_i = \alpha \sum_j A_{ij} x_j / n_j + \beta$

Рассматриваются степени соседей

Центральность по Кацу: $x = \alpha Ax + \beta 1$

Обобщенная центральность по собственному вектору

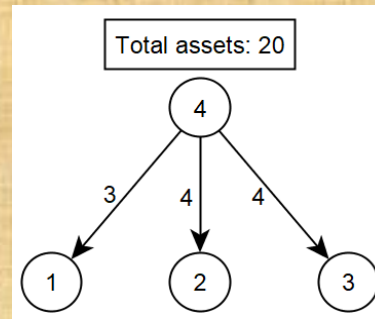
(штрафы, нет ненулевых решений)



Основные понятия

- Каждая вершина определяется вектором параметров, скажем, $\bar{\alpha}^i = (\alpha_1^i, \dots, \alpha_s^i)$ для каждой вершины $i \in V$. Эта возможность была упомянута Ньюманом [Newman 2003]: “И вершины или дуги могут иметь множество свойств, числовых или иных, связанных с ними”.
- Для социальных сетей, где дуги представляют людей, эти параметры могут быть возрастом респондента, ее годовым доходом, уровнем образования, населением поселения, где она живет, политическими взглядами (скажем, на левой-правой шкале), etc.
- В банковской сети, где вершины являются банками, этими параметрами могут быть общие активы, общая сумма кредитов, рискованные кредиты, просроченные кредиты, etc.
- Каждая вершина ij также описывается набором параметров $\bar{w}_{ij} = (w_{ij}^1, \dots, w_{ij}^r)$, которые могут представлять различные детали соединения.

Пример



Малая банковская система

$q_4 = 0.5 * \text{совокупные}$
активы = 10

Пусть совокупные активы банка 4 (то есть кредитор в сети) равны 20. Допустим, кредитор предоставляет кредит трем другим банкам (банки 1, 2 и 3) и страдает от убытка, если заемщики не будут выплачивать 50% своих общих активов (отсюда, $q_4 = 20 \cdot 50\% = 10$). В этом случае каждый отдельный заемщик 1, 2, 3 с кредитами $w_{41} = 3$, $w_{42} = 4$, $w_{43} = 4$ не создают проблему для банка 4, т.к. $w_{4k} < q_4$ для $k = 1, 2, 3$. Тем не менее, $w_{41} + w_{42} + w_{43} > q_4$.

Критическая группа. Ключевые вершины.

Простой путь.

Определение. Критическая группа $\Omega(j)$ для узла j является подмножеством вершин, чье групповое воздействие превышает некоторый порог q_j . Более формально $\Omega(j) \subseteq V$ является критической группой узла j , если

$$\sum_{i \in \Omega(j)} w_{ij} \geq q_j$$

Определение. Вершина является ключевой для некоторой группы, если ее исключение из этой группы делает группу некритической. В одной группе может быть несколько ключевых вершин. Формально $\Omega^p(j) \subseteq \Omega(j)$ является подмножеством ключевых вершин группы $\Omega(j)$, если

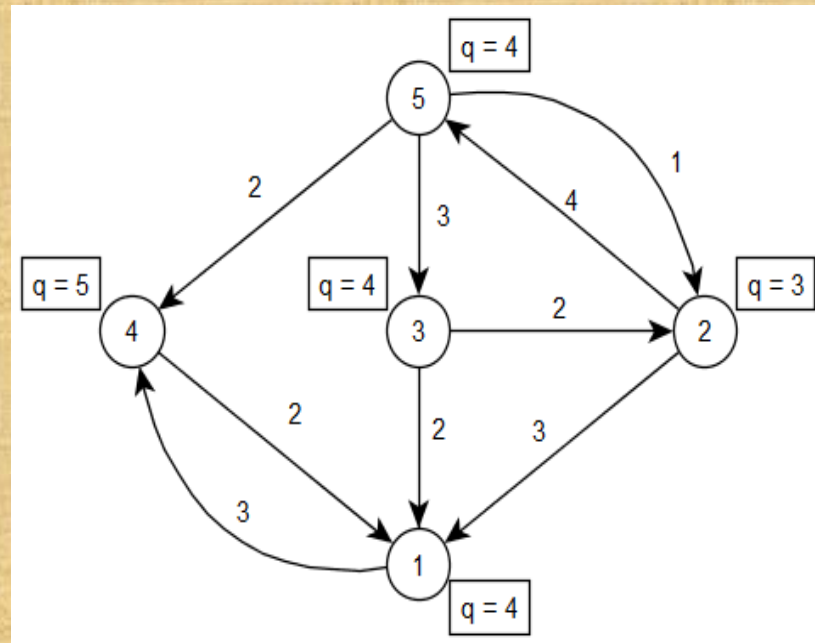
$$\forall k \in \Omega^p(j) \quad \sum_{i \in \Omega(j) \setminus \{k\}} w_{ij} < q_j$$

Определение. Простой путь между вершинами i и j в графе G представляет собой последовательность ребер, которая соединяет их и содержит различные вершины. Точнее,

$$(i, k_1), \quad (k_1, k_2), \quad (k_2, k_3), \quad \dots, \quad (k_{s-1}, j), \\ i \neq k_1 \neq \dots \neq k_{s-1} \neq j$$

- простой путь длины s , обозначаемый как $P_{i-j}^t(s)$, где t - номер пути.

Ближнее взаимодействие. Пример.



Список критических групп и ключевых вершин в этих группах для вершины 1

Критические группы, $\Omega_k(1)$ где k – номер группы	Ключевые вершины, $\Omega_k^p(1)$ где k – номер группы
$\Omega_1(1) = \{2, 3\}$	$\Omega_1^p(1) = \{2, 3\}$
$\Omega_2(1) = \{2, 4\}$	$\Omega_2^p(1) = \{2, 4\}$
$\Omega_3(1) = \{3, 4\}$	$\Omega_3^p(1) = \{3, 4\}$
$\Omega_4(1) = \{2, 3, 4\}$	$\Omega_4^p(1) = \{\emptyset\}$
$\Omega_5(1) = \{2, 3, 5\}$	$\Omega_5^p(1) = \{2, 3\}$
$\Omega_6(1) = \{2, 4, 5\}$	$\Omega_6^p(1) = \{2, 4\}$
$\Omega_7(1) = \{3, 4, 5\}$	$\Omega_7^p(1) = \{3, 4\}$
$\Omega_8(1) = \{2, 3, 4, 5\}$	$\Omega_8^p(1) = \{\emptyset\}$

Прямые и косвенные взаимодействия

Нормализованные интенсивности прямых воздействий и косвенных воздействий через одну промежуточную вершину.

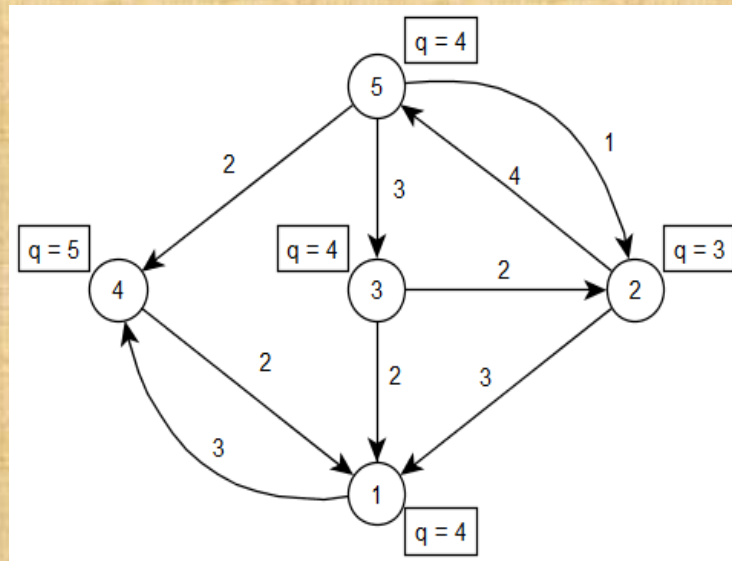
Обратите внимание, что косвенное воздействие рассчитывается только в том случае, если существует прямая связь с вершиной 1.

$p_{ij} = \frac{w_{ij}}{\sum_k w_{kj}}$ – интенсивность прямого воздействия вершины i на вершину j ,

$$p_{ihj} = \begin{cases} \frac{w_{ih}}{\sum_k w_{kj}}, & \text{if } w_{ij} > 0, \quad w_{hj} > 0, \quad w_{hj} \geq w_{ih}, \quad i \neq h \neq j, \\ \frac{w_{hj}}{\sum_k w_{kj}}, & \text{if } w_{ij} > 0, \quad w_{hj} > 0, \quad w_{hj} < w_{ih}, \quad i \neq h \neq j, \\ 0, & \text{иначе.} \end{cases} \quad (4)$$

p_{ihj} – интенсивность косвенного воздействия вершины i на вершину j через вершину h .

Прямое и косвенное взаимодействия



Прямые интенсивности воздействий с вершиной 1 равны $p_{21} = 3/7$; $p_{31} = 2/7$; $p_{41} = 2/7$ и косвенные интенсивности влияний с вершиной 1 равны $p_{321} = 2/7$.

Интенсивности влияний ключевых вершин

Зная критические группы, ключевые вершины и нормализованные интенсивности соединений, можно вычислить интенсивности в пределах каждой критической группы для каждой ключевой вершины узла.

Мы также обращаем внимание на размер группы (чем больше группа, тем меньше вклад, который имеет каждый ключевая вершина).

Следовательно, введем интенсивность влияния ключевой вершины i на j внутри группы $\Omega_k(j)$, обозначаемой $f(i, \Omega_k(j))$

$$f(i, \Omega_k(j)) = \frac{p_{ij} + \sum_{h \in \Omega_k(j)} p_{ihj}}{|\Omega_k(j)|}, i \in \Omega_k^p(j).$$

Интенсивность воздействия на вершину 1 внутри группы $\Omega_k(j)$

Критические группы, $\Omega_k(1)$ где k - число групп	Ключевые вершины, $\Omega_k^p(1)$ где k - число групп	Интенсивность влияния ключевого узла i на узел j в критической группе $\Omega_k(j) \rightarrow f(i, \Omega_k(j))$
$\Omega_1(1) = \{2, 3\}$	$\Omega_1^p(1) = \{2, 3\}$	<ul style="list-style-type: none"> • $f(2, \Omega_1(1)) = \frac{3/7}{2} = 3/14$ • $f(3, \Omega_1(1)) = \frac{2/7+2/7}{2} = 4/14$
$\Omega_2(1) = \{2, 4\}$	$\Omega_2^p(1) = \{2, 4\}$	<ul style="list-style-type: none"> • $f(2, \Omega_2(1)) = \frac{3/7}{2} = 3/14$ • $f(4, \Omega_2(1)) = \frac{2/7}{2} = 2/14$
$\Omega_3(1) = \{3, 4\}$	$\Omega_3^p(1) = \{3, 4\}$	<ul style="list-style-type: none"> • $f(3, \Omega_3(1)) = \frac{2/7}{2} = 2/14$ • $f(4, \Omega_3(1)) = \frac{2/7}{2} = 2/14$
$\Omega_4(1) = \{2, 3, 4\}$	$\Omega_4^p(1) = \{\emptyset\}$	-
$\Omega_5(1) = \{2, 3, 5\}$	$\Omega_5^p(1) = \{2, 3\}$	<ul style="list-style-type: none"> • $f(2, \Omega_5(1)) = \frac{3/7}{3} = 3/21$ • $f(3, \Omega_5(1)) = \frac{2/7+2/7}{3} = 4/21$
$\Omega_6(1) = \{2, 4, 5\}$	$\Omega_6^p(1) = \{2, 4\}$	<ul style="list-style-type: none"> • $f(2, \Omega_6(1)) = \frac{3/7}{3} = 3/21$ • $f(4, \Omega_6(1)) = \frac{2/7}{3} = 2/21$
$\Omega_7(1) = \{3, 4, 5\}$	$\Omega_7^p(1) = \{3, 4\}$	<ul style="list-style-type: none"> • $f(3, \Omega_7(1)) = \frac{2/7}{3} = 2/21$ • $f(4, \Omega_7(1)) = \frac{2/7}{3} = 2/21$
$\Omega_8(1) = \{2, 3, 4, 5\}$	$\Omega_8^p(1) = \{\emptyset\}$	-

После того, как мы получим интенсивности воздействий вершины i на вершину j во всех критических группах вершины j , где i является ключевой, нам необходимо агрегировать их в одно значение или, другими словами, в воздействие между вершинами. Модель агрегирования, представленная в [Aleskerov et al., 2014], выглядит следующим образом

$$\chi_i(j) = \sum_k f(i, \Omega_k(j)).$$

Наконец, эти значения нормируются к их сумме

$$\hat{\chi}_i(j) = \frac{\chi_i(j)}{\sum_k \chi_k(j)}.$$

Для нашего примера

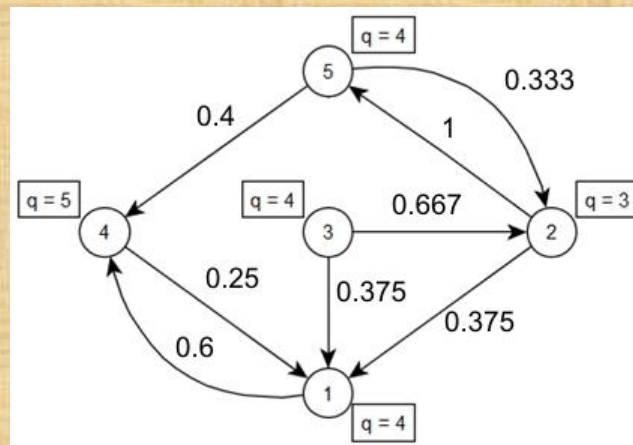
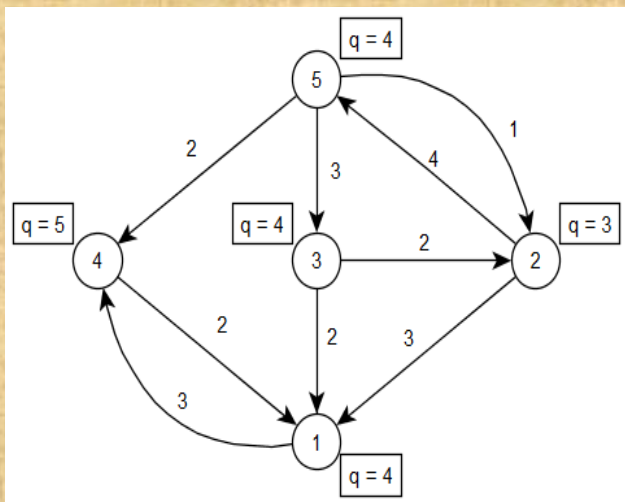
- $\chi_2(1) = \frac{3}{14} + \frac{3}{14} + \frac{3}{21} + \frac{3}{21} = \frac{15}{21}$; $\chi_3(1) = \frac{4}{14} + \frac{2}{14} + \frac{4}{21} + \frac{2}{21} = \frac{15}{21}$;
- $\chi_4(1) = \frac{2}{14} + \frac{2}{14} + \frac{2}{21} + \frac{2}{21} = \frac{10}{21}$; $\chi_5(1) = 0$.

Нормированные значения равны

- $\hat{\chi}_2(1) = \frac{15/21}{15/21+15/21+10/21+0} = \frac{15}{40}$; $\hat{\chi}_3(1) = \frac{15/21}{15/21+15/21+10/21+0} = \frac{15}{40}$;
- $\hat{\chi}_4(1) = \frac{10/21}{15/21+15/21+10/21+0} = \frac{10}{40}$; $\hat{\chi}_5(1) = 0$.

Полученные значения могут интерпретироваться как новые веса на дугах между i -ой вершиной и j -ой вершиной: это окончательное модифицированное воздействие вершины i на вершину j , которое включает как прямое, так и косвенное воздействие индивидуально или через группы с другими вершинами с учетом индивидуальных параметров вершин.

Влияние вершины на вершину для численного примера



Это влияние вершины на вершину может быть преобразовано в одно значение, соответствующее каждой вершине (мера влияния на графике) - индекс SRIC

α_{SRIC} , который введен в [Aleskerov et al., 2014]. Идея заключается в том, что воздействие на конкретную вершину важнее, если эта вершина управляет многими другими вершинами.

α_{SRIC} индекс в соответствии с нормированной сверткой по степеням исхода

$$\alpha_{SRIC}(1) = 0.19; \alpha_{SRIC}(2) = 0.34; \alpha_{SRIC}(3) = 0.22;$$

$$\alpha_{SRIC}(4) = 0.06; \alpha_{SRIC}(5) = 0.19.$$

Другое определение $\alpha_{SRIC}(i)$

Другим вариантом является среднее значение для каждой вершины, т.е.

$$\alpha_{SRIC}(i) = \frac{\sum_{j \in V} \hat{\chi}_i(j)}{|V|}.$$

Для нашего примера

$$\alpha_{SRIC}(1) = 0.12; \alpha_{SRIC}(2) = 0.28; \alpha_{SRIC}(3) = 0.21;$$
$$\alpha_{SRIC}(4) = 0.05; \alpha_{SRIC}(5) = 0.15.$$

Недостатки SRIC

- требует наличия прямой связи для оценки косвенного воздействия, которое обычно приводит к недооценке реального воздействия: в нашем примере вершина 5 может воздействовать на вершину 1 по вершине 4, но SRIC не чувствителен к ней;
- учитывает только соединения по дальним воздействиям и не учитывает воздействие более чем одного промежуточной дуги;
- процедуры формирования групп включают косвенно связанные вершины, и это выглядит нецелесообразным, поскольку эти вершины не будут воздействовать на рассматриваемую.

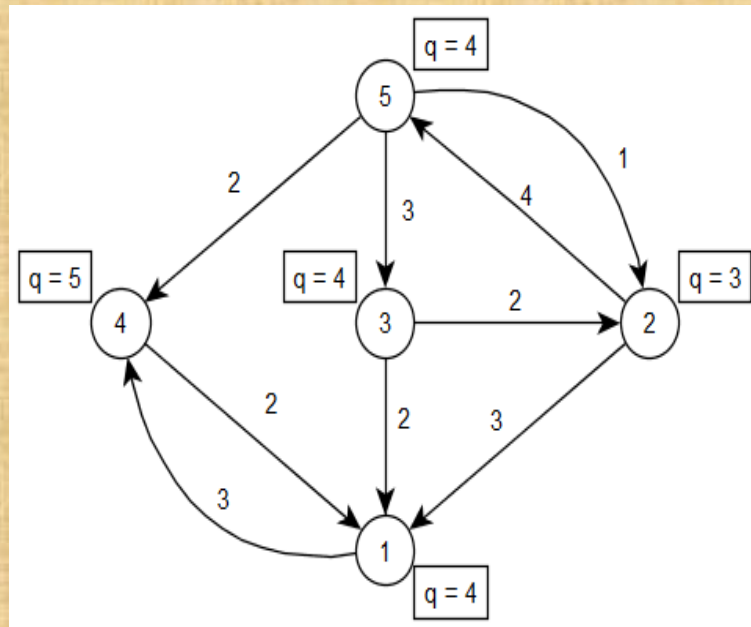
Центральность по дальним взаимодействиям

В нашем подходе мы учитываем только группы прямых соседей. Когда мы вычисляем воздействие вершины i на вершину j , мы формируем все критические группы j , где i является ключевой, и рассмотрим ту, в которой i непосредственно влияет на j наибольшим образом. Формально,

$$c_{ij} = \begin{cases} \max_{\Omega_k(j): i \in \Omega_k^p(j)} \frac{w_{ij}}{\sum_{h \in \Omega_k(j)} w_{hj}}, & \text{if } \exists \Omega_k^p(j): i \in \Omega_k^p(j), \\ 0, & \text{иначе.} \end{cases}$$

Очевидно, $0 \leq c_{ij} \leq 1$. $c_{ij} = 0$ означает, что либо i не имеет прямой связи с j вообще, либо вес на дуге (i, j) слишком мал, а i не является ключевой вершиной в любых критических группах из j . Следовательно, некоторые слабые стрелки могут исчезнуть по сравнению с исходным графом. $c_{ij} = 1$ означает, что i превышает порог j без образования групп.

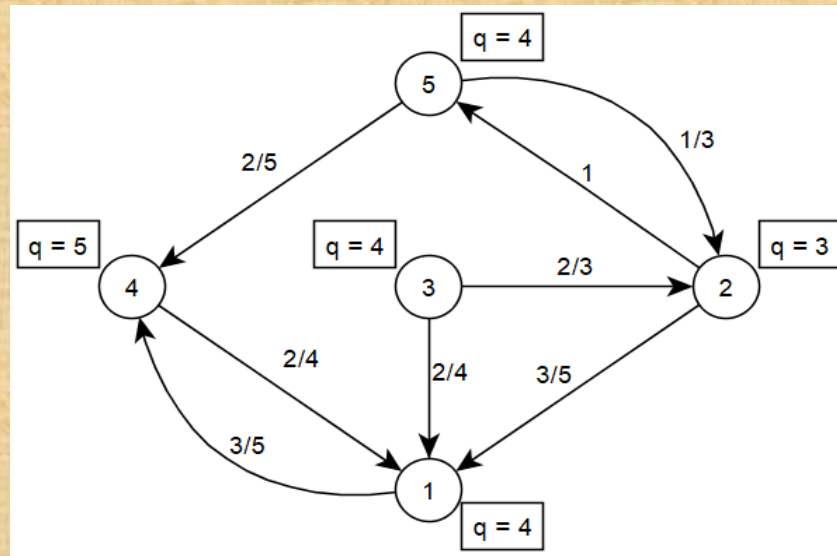
Наш пример



Критические группы и ключевые вершины для каждой вершины

i (число вершин)	Критические группы, $\Omega_k(i)$ где k – число групп	$\sum_{h \in \Omega_k(j)} w_{hj}$	Ключевые вершины, $\Omega_k^p(i)$ где k – число групп
1	$\Omega_1(1) = \{2, 3\}$	5	$\Omega_1^p(1) = \{2, 3\}$
	$\Omega_2(1) = \{2, 4\}$	5	$\Omega_2^p(1) = \{2, 4\}$
	$\Omega_3(1) = \{3, 4\}$	4	$\Omega_3^p(1) = \{3, 4\}$
	$\Omega_4(1) = \{2, 3, 4\}$	7	$\Omega_3^p(1) = \{\emptyset\}$
2	$\Omega_1(2) = \{3, 5\}$	3	$\Omega_1^p(2) = \{3, 5\}$
3	–	–	–
4	$\Omega_1(4) = \{1, 5\}$	5	$\Omega_1^p(4) = \{1, 5\}$
5	$\Omega_1(5) = \{2\}$	4	$\Omega_1^p(5) = \{2\}$

Прямые взаимодействия для численного примера



Косвенное взаимодействие

Одна вершина также может косвенно воздействовать на другую вершину. Таким образом, чтобы построить матрицу полного воздействия, мы рассмотрим график прямых воздействий и проанализируем все простые пути между дугами (Определение 3). В качестве примера мы оцениваем воздействие вершины 3 на вершины 1.

Существует 3 простых пути от вершины 3 до вершины 1, перечисляемого в таблице.

Простые пути между вершинами 3 и 1

t , номер пути	s , длина пути	$P_{3-1}^t(s)$
1	1	$3 \xrightarrow{2/4} 1$
2	2	$3 \xrightarrow{2/3} 2 \xrightarrow{3/5} 1$
3	4	$3 \xrightarrow{2/3} 2 \xrightarrow{1} 5 \xrightarrow{2/5} 4 \xrightarrow{2/4} 1$

Агрегирование по путям

Чтобы оценить воздействие через разные пути, мы вычисляем силу каждого пути и объединяем его в одно значение. Существует несколько методов этой процедуры. Например, оценку прочности пути можно выполнить, умножив эти оценки или взяв минимальную оценку

$$f_{mult}(P_{i-j}^t(s)) = c_{ik_1(t)} \times c_{k_1(t)k_2(t)} \times \dots \times c_{k_{s-1}(t)j},$$

или

$$f_{min}(P_{i-j}^t(s)) = \min\{c_{ik_1(t)}, c_{k_1(t)k_2(t)}, \dots, c_{k_{s-1}(t)j}\}$$

Первый вариант рассматривается как объединенная вероятность, где первая вершина в этой цепочке непосредственно воздействует на вторую с «вероятностью» $c_{ik_1(t)}$, вторая вершина воздействует на третью с «вероятностью», $c_{k_1(t)k_2(t)}$, и т.д., и все эти события независимы.

Второй вариант рассматривается в виде пропускной способности, когда вершина i не может воздействовать на вершину j через путь P_{i-j}^t больше чем на минимальное воздействие этого пути.

Агрегирование по путям

Поскольку между вершинами существует много путей, нам необходимо агрегировать попарное взаимодействие в одно значение. Выбор правил зависит от проблемы, поскольку они учитывают различные функции. Например, общее взаимодействие между двумя вершинами может быть рассчитано как максимальное значение по путям или суммарное значение по путям.

Косвенное воздействие узла 3 на узел 1 на разные пути

t , номер пути	P_{3-1}^t	$f_{mult}(P_{3-1}^t)$	$f_{min}(P_{3-1}^t)$
1	$3 \xrightarrow{2/4} 1$	$2/4$	$2/4$
2	$3 \xrightarrow{2/3} 2 \xrightarrow{3/5} 1$	$2/3 \times 3/5 = 2/5$	$3/5$
3	$3 \xrightarrow{2/3} 2 \xrightarrow{1} 5 \xrightarrow{2/5} 4 \xrightarrow{2/4} 1$	$2/3 \times 1 \times 2/5 \times 2/4 = 2/15$	$2/5$
Итоговое воздействие		$1/2$	$2/3$

Для большинства реальных сетей очень дальние взаимодействия не имеют никакого значения. Принимая во внимание этот факт, мы можем ограничить длину рассмотренных путей некоторым параметром s и не изучать пути длиной больше, чем s . Практически этот параметр обычно не превышает 4 (большее число не приведет к резкому изменению результатов).

Воздействие вершины на вершину для численного примера

Вершины	1	2	3	4	5
1	0	0	0	0.6	0
2	0.6	0	0	0.6	1
3	0.5	0.667	0	0.6	0.667
4	0.5	0	0	0	0
5	0.4	0.333	0	0.4	0

Мы получили одну версию матрицы дальнего взаимодействия, которая содержит информацию о попарном влиянии. Как и в подходе с ближними взаимодействиями, мы можем применять различные правила агрегирования для получения общего влияния вершин. Например, в соответствии с нормированной сверткой по степеням исхода индекс α_{LRIC} равен:

$$\alpha_{LRIC}(1) = 0.10; \alpha_{LRIC}(2) = 0.32; \alpha_{LRIC}(3) = 0.35; \\ \alpha_{LRIC}(4) = 0.07; \alpha_{LRIC}(5) = 0.16.$$

Наиболее важными элементами в этом примере являются вершины 2 и 3.

Индекс влияния при учете дальних взаимодействий

Общее косвенное влияние: комбинации методов

		Агрегирование путей	
		Сумма влияния путей	Влияние пути для наихудшего случая
Влияние пути	Умножение прямого влияния	Сумма путей	Maximal Path Influence
	Минимальное прямое влияние	–	MaxMin