
СЕТЕВОЙ АНАЛИЗ

М.А. Сафонова, Б.Е. Винер
(Санкт-Петербург)

СЕТЕВОЙ АНАЛИЗ СОЦИТИРОВАНИЙ ЭТНОЛОГИЧЕСКИХ ПУБЛИКАЦИЙ В РОССИЙСКИХ ПЕРИОДИЧЕСКИХ ИЗДАНИЯХ: ПРЕДВАРИТЕЛЬНЫЕ РЕЗУЛЬТАТЫ ¹

Авторы статьи заинтересованы в выявлении когнитивной структуры российской этнологии, т.е. исследовательских областей внутри этой дисциплины, с помощью сетевого анализа журнальных ссылок из российских научных журналов за 2005–2009 гг. В данном исследовании выяснялось насколько успешны используемые авторами алгоритмы анализа для решения поставленной ими задачи. Выявлено, что для поставленной задачи алгоритм ван-Донгена оказался более эффективным, чем алгоритм Ньюмена.

Ключевые слова: когнитивная структура науки, сетевой анализ, российская этнология.

Мария Андреевна Сафонова – кандидат социологических наук, доцент кафедры социологии НИУ ВШЭ (Санкт-Петербургский филиал). E-mail: safonovam@yandex.ru.

Борис Ефимович Винер – кандидат социологических наук, старший научный сотрудник Социологического института РАН. E-mail: wienerras@yandex.ru.

¹ Исследование выполнено при финансовой поддержке РГНФ в рамках проекта проведения научных исследований («Когнитивная структура российской этнологии (по данным журнальных цитирований за 2005–2009 гг.)»), проект № 12-03-00030. Данные по ссылкам к 25 наиболее престижным российским социологическим и этнологическим журналам за 2005–2009 гг. любезно предоставлены Научной электронной библиотекой *eLIBRARY.RU*.

Введение

Британский социолог Р. Уитли ввел понятие когнитивной структуры науки, но не дал достаточно четкого определения этого феномена. К когнитивным структурам Уитли относит «парадигмы» Т. Куна, «теории» К. Поппера, «исследовательские программы» И. Лакатоса. Институционализированную когнитивную структуру он связывает прежде всего с научной специальностью¹ [1]. Но когнитивная структура науки может включать как крупные, так и мелкие исследовательские области. Науковед И.В. Маршакова-Шайкевич² в качестве иллюстрации когнитивной структуры науки приводит выпускаемые Институтом научной информации (Филадельфия) карты науки пяти уровней, включающие на нижнем уровне «кластеры “ядерных” публикаций, образующих исследовательские фронты», а на самом высоком уровне – глобальную карту областей исследования в естественных и социальных науках, охватывающие всю мировую совокупность научных ссылок за определенный период [2, с. 210–219]. В основу создания подобных карт положен подсчет количества социцированных (т.е. совместных цитирований) двух конкретных документов цитирующими статьями³.

Уитли в первую очередь интересовала ситуация на нижних уровнях, поскольку его «личный опыт изучения ученых, специализирующихся в области физики твердого тела, показал малую пригодность используемых философами и социологами категорий для выявления того, чем действительно заняты ученые в данной области и как развивается их деятельность» [Ibid.]. Именно опи-

¹ Под научной специальностью в американской социологии принято понимать исследовательскую область, оформленную как секция Американской социологической ассоциации. Данное понятие можно применять также к исследовательским областям внутри других научных дисциплин, но способы их определения вне социологии американские исследователи не приводят.

² Свои ранние работы она публиковала под фамилией Маршакова.

³ Маршакова-Шайкевич предлагает называть цитируемые документы базовыми, а цитирующие – проспективными [3, с. 6].

сание когнитивной структуры научной дисциплины на нижних уровнях позволяет получить детальную картину исследуемых ее представителями объектов и интересующих их проблем.

К настоящему времени в изучении когнитивной структуры науки оформились два подхода. Макроподход использует данные из баз *Web of Science*, которые обрабатываются с помощью сетевого анализа. При этом выборка строится на основе использования полной базы журнальных ссылок за определенный год или квартал. Таким образом, в выборку попадают данные по всем существующим на сегодняшний день научным дисциплинам. Далее возможны два варианта действий. В первом случае исследователь включает в базу данных ссылки журнальных публикаций на другие журналы без различения статей (например, [3]). Во втором случае в сетевом анализе используются ссылки авторов статей уже не на журналы, а на конкретные публикации в других изданиях.

Представители микроподхода используют данные из одного или нескольких журналов по одной дисциплине или субдисциплине за ряд лет. Так, специалист по управлению персоналом М. Гмюр изучал социцитирования в области *organization studies* на основе ссылок к статьям в пяти журналах по менеджменту за 15 лет с помощью различных статистических процедур, большинство из которых представляют собой вариации кластерного анализа [5].

В данной работе авторы для поиска исследовательских областей внутри российской этнологии используют элементы обоих подходов, применяя два алгоритма, разработанных в рамках сетевого анализа, к данным выборки российских этнологических и междисциплинарных журналов и сборников за пять лет. Главная цель авторов – выяснить, насколько успешно используемые алгоритмы справляются с задачей выделения групп, соответствующих исследовательским областям, в совокупности имеющихся у нас ссылок¹.

¹ На данном этапе работы мы ограничили себя получением предварительного перечня исследовательских областей из которых состоит современная российская

Прежде чем начать описывать проделанную нами работу считаем важным определить, что мы здесь понимаем под исследовательской областью. Уитли полагает, что исследовательская область представляет собой совокупность нескольких проблемных ситуаций, а научная специальность объединяет несколько исследовательских областей [1, р. 229–230]. Проблемная ситуация может иметь в качестве основы один из трех принципов: сходство исследуемых феноменов; исследуемое вещество или «система»; использование общих методик. В качестве примеров проблемных ситуаций Уитли приводит соответственно сверхпроводимость, изучение аморфных материалов, применение жидкого гелия в физике низких температур (в радиобиологии он говорит о радиационном методе) [1, р. 230–231]. Исследовательская область характеризуется тем, что вырабатывает общее понимание относительно того, что описывают проблемные ситуации, входящие в данную область [1, р. 234]. Мы полагаем, что для данного проекта будет верным считать исследовательскими областями совокупности ссылок, которые нам удастся идентифицировать, во-первых, в качестве объектов этнологического изучения (этнические, антропологические, лингвистические, религиозные и другие общности разного таксономического уровня; термины, характеризующие особенности культуры этих общностей), во-вторых, как объекты теоретической рефлексии этнологов и специалистов из смежных дисциплин, или, в третьих, как проблемы, изучаемые сходными методами, которые применяются в этнологии и смежных дисциплинах.

этнологии. Ознакомившись с нашими рабочими материалами, петербургский этнолог Т.Б. Щепанская высказала мысль, что описанию каждой из этих областей, или по крайней мере нескольких областей, включающих близкую тематику, можно посвятить самостоятельную статью. Однако к этой задаче мы планируем перейти лишь после окончательной отработки используемых методов.

Выборка

Наша генеральная совокупность включает все опубликованные в России на русском языке монографии и статьи в сборниках и журналах по этнологии за 2005–2009 гг. К сожалению, получить ссылки ко всему названному массиву литературы станет возможно только в перспективе, если все бумажные носители будут оцифрованы и к ним будет предоставлен свободный доступ. Наша выборка включает 71 журнал и продолжающийся сборник, издававшиеся в 2005–2009 гг., причем ряд из них имеют лакуны¹. Пристатейные библиографические списки мы получили с сайта Научной электронной библиотеки *elibrary.ru*, с доступных нам сайтов журналов, не представленных на названном ресурсе, а также с бумажных носителей из Библиотеки Академии наук и ее филиалов. Для составления списков интересующих нас журналов мы использовали издаваемый ИНИОН библиографический указатель «Новая литература по социальным и гуманитарным наукам. История. Археология. Этнология», издаваемую книжной палатой «Летопись периодических и продолжающихся изданий» за соответствующие годы, а также такие представленные на сайте РИНЦ инструменты поиска журналов, как дисциплинарная принадлежность, ключевые слова в названии журнала и соответствующая рубрике ГРНТИ тематическая рубрика 03.61.00 «Этнография и историческая антропология». Кроме того, мы обращали внимание на наличие слов «этнология», «этнография», «антропология» и производных от них в названии издания. Число собственно этнологических журналов и сборников составило 15². Еще шесть жур-

¹ Данная выборка является подвыборкой из нашей большой выборки, включающей 156 журналов и 9 продолжающихся сборников по социологии и этнологии за тот же период.

² Журналы: «Антропологический форум», «Антропологический форум online», «Манифестация», «Расы и народы», «Этнографическое обозрение», «Этнографическое обозрение online», «Этножурнал», «Живая старина» и «Традиционная

налов являются междисциплинарными¹. В названиях большинства остальных журналов используются слова «вестник», «известия», «проблемы». При пороге в четыре цитирования выборка включает 1 073 названия базовых публикаций, процитированных в общей сложности 6 871 раз в 1 641 проспективных статьях.

База данных для компьютерных вычислений представляет собой прямоугольную матрицу², отображающую отношения между двумя видами узлов: цитирующими статьями и цитируемыми документами³. Далее эта матрица путем умножения (прямоугольная матрица умножается на ее транспоненту) преобразована в квадратную матрицу, где отображены только цитируемые документы. Таким образом, финальная матрица, которая обрабатывалась в ходе статистического анализа – это квадратная матрица, в которой в качестве узлов представлены цитируемые документы. Количество социцированных отображено в ячейке матрицы, находящейся на пересечении соответствующих столбца и строки. В матрице представлено 1 073 документа (узла) и 37 798 связей между ними (ребер). Для того чтобы обнаружить представленные в сети подгруппы, мы использовали алгоритмы поиска подгрупп, разработанные в рамках сетевого анализа. К их рассмотрению мы и переходим.

культура» также включены в этнологическую подвыборку, поскольку РИНЦ включает их в тематическую рубрику «03.61.00 Этнография и историческая антропология». Сборники: «Алгебра родства», «Studia Ethnologica», «Материалы полевых исследований МАЭ РАН», «Полевые исследования Института этнологии и антропологии РАН», «Сборник МАЭ», «Этнологические исследования по шаманству и иным традиционным верованиям и практикам».

¹ «Ab Imperio», «Археология, этнография и антропология Евразии», «Вестник археологии, антропологии и этнографии», «Диаспоры», «Журнал социологии и социальной антропологии», «Мир России: социология, этнология».

² Отредактированную базу данных в формате Excel можно получить от авторов по личному запросу.

³ В сетевом анализе такой вид матриц принято называть бимодальными.

Модель коммунитарной структуры сети

В ходе работы мы пытались выявить основную топологическую (т.е. пространственную) модель в общей структуре цитирований в нашей матрице. Одним из топологических свойств сети является наличие либо отсутствие коммунитарной структуры (*community structure*) [6], т.е. присутствие/отсутствие в сети структурных блоков¹, агенты которых имеют больше связей с агентами внутри своего блока, чем с агентами других подобных образований². Статьи и учебники по сетевому анализу традиционно представляют как графическую, так и матричную форму коммунитарной структуры. В графе на *рис. 1* представлены связи между несколькими узлами³. Интуитивно понятно, что здесь наблюдатель может выделить две группы – А и В. Между узлами, принадлежащими к одной группе, связей больше, чем между узлами, принадлежащими к разным группам. Такая модель взаимодействия отображается в графической и матричной форме, традиционно воспроизводимой в статьях и учебниках по сетевому анализу. На *рис. 2* тот же граф представлен в виде матрицы.

На представляющем блок-матрицу⁴ *рис. 3*, пунктир показывает границы ячеек⁵. Внутри одного блока, т.е. ячейки, находятся узлы,

¹ Такие блоки разные авторы называют группами, сообществами или кластерами, используя эти термины как синонимы. Причем данные термины могут использоваться для описания групп агентов в любых сетях, независимо от того, состоят они из людей [6] или, например, из функционально связанных белков [7].

² Такая идеально-типическая модель взаимодействия лежит в основе многих алгоритмов поиска групп в сети, например, алгоритма поиска фракций [8], метода Гирвана-Ньюмена [6], алгоритма Ньюмена [9], алгоритма ван-Донгена [10], лувеновского метода [11]. Подробнее об алгоритмах поиска групп в сети см. [12; 13; 14].

³ Граф – это «фигура, состоящая из точек – вершин, соединенных прямолинейными отрезками-ребрами» [15, с. 13].

⁴ Блок-матрицы и блок-модели используются, чтобы показать взаимодействие между выделенными в ходе исследования группами агентов (узлов).

⁵ В данном случае это матрица размером 2x2.

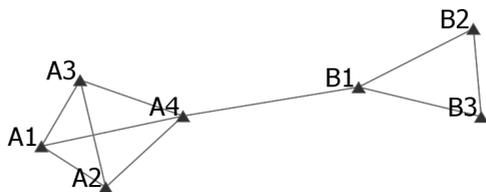


Рис. 1. Модель коммунитарной структуры в графической форме

	A1	A2	A3	A4	B1	B2	B3
A1	1	1	1	1	0	0	0
A2	1	1	1	1	0	0	0
A3	1	1	1	1	0	0	0
A4	1	1	1	1	1	0	0
B1	0	0	0	1	1	1	1
B2	0	0	0	0	1	1	1
B3	0	0	0	0	1	1	1

Рис. 2. Модель коммунитарной структуры в матричной форме

	A	A	A	A	B	B	B
A1	1	1	1	1			
A2	1	1	1	1			
A3	1	1	1	1			
A4	1	1	1	1	1		
B1				1	1	1	1
B2					1	1	1
B3					1	1	1

Рис. 3. Блок-модель с предполагаемыми группами

выделенные нами в одну группу. Блок-матрица позволяет наблюдать интенсивность взаимодействия такими между группами. Мы можем предположить, что наблюдаем коммутитарную структуру в случаях, когда значения, соответствующие присутствию отношений (в данном случае «1»), концентрируются преимущественно на главной диагонали блок-матрицы, т.е. внутри блока, а значения, соответствующие отсутствию отношений, – за пределами главной диагонали (т.е. за пределами блока)¹. Алгоритмы поиска групп в сетевом анализе ищут такой вариант разделения сети на группы, который в максимальной степени соответствует описанной выше модели коммутитарной структуры.

В ходе анализа данных мы пользовались двумя алгоритмами, которые, во-первых, эффективны при работе с графами размером более 1 тысячи узлов, во-вторых, не требуют от пользователя предварительных гипотез о количестве и размере групп, в-третьих, способны разделить сеть на непересекающиеся группы. Такими качествами обладают алгоритмы Ньюмена и ван-Донгена.

Алгоритм группировки Ньюмена

В сетевом анализе поиск подгрупп в сети опирается на гипотезу о том, что в анализируемой структуре существует естественное, обусловленное характеристиками развития этой структуры, разделение на сообщества. В действительности, в конкретном сообществе такого разделения может и не быть. Поэтому одна из задач методологов – разработка алгоритмов, которые, во-первых, могут продемонстрировать невозможность разделения сети на группы, а во-вторых, позволяют рассчитать показатели, демонстрирующие степень соответствия полученной в ходе сетевого анализа группировки идеально-типической коммутитарной структуре, описанной ранее.

¹ Ноли на репрезентации матрицы обычно не указываются.

В 2006 г. физик М. Ньюмен предложил алгоритм, включающий в определение коммутитарной структуры еще одно важное наблюдение. Ньюмен говорит, что недостаточно просто посчитать количество связей внутри блоков и за их пределами. Необходимо показать, что количество связей, полученное в результате подсчета, не случайный результат. С точки зрения Ньюмена, хорошее разделение сети на блоки – это такое разделение, где число связей между блоками меньше, чем число, ожидаемое в ситуации случайного распределения связей по блокам и графу, или число связей внутри блоков больше, чем число связей, ожидаемое в ситуации случайного распределения связей по графу. Эту идею Ньюмен формализует с помощью показателя модулярности (*modularity*).

Алгоритм работает следующим образом. Пусть наша исходная квадратная матрица обозначена как матрица A . Исследователь конструирует матрицу модулярности, обозначаемую как матрица B . Ее размер равен размеру матрицы A (т.е. то же количество строк и столбцов). Ячейки этой матрицы (B_{ij}) заполняются следующим образом: B_{ij} – это значение, которое стояло в корреспондирующей ячейке исходной матрицы (A_{ij}) минус число ожидаемых ребер между вершинами i и j в случае, если бы ребра были разбросаны между вершинами случайно. Значение числа ожидаемых ребер между вершинами i и j в этом случае рассчитывается по формуле:

$$\frac{k_i k_j}{2m},$$

где k_i и k_j – валентность¹ вершин i и j , а m – общее число ребер в графе (т.е. в матрице A). Таким образом, формула приобретает вид:

$$B_{ij} = \frac{A_{ij} - k_i k_j}{2m}.$$

¹ Валентность (*англ.* degree) узла в сетевом анализе – это число связей, которое имеет узел. Термины «валентность» и «степень» в русскоязычной терминологии для сетевого анализа являются синонимами.

Затем для матрицы B ведется поиск главного (самого большого положительного) собственного значения (*eigenvalue*) и собственного вектора матрицы, соответствующего этому значению. Членство в группе приписывается узлу сети по знаку, соответствующему узлу ячейки найденного собственного вектора. Благодаря этой процедуре сеть разделяется на две группы: для членов одной из них знаки в ячейках собственного вектора будут положительными, для членов другой – отрицательными. Если у матрицы нет положительных собственных значений, ее главный собственный вектор будет выглядеть как $(1, 1, 1 \dots)$. В этом случае невозможно разделить узлы на две группы. С точки зрения Ньюмена, такой результат вполне удовлетворителен, поскольку означает, что в матрице не наблюдается коммунитарной структуры. Если же матрица B имеет положительное главное собственное значение, для нее рассчитывается значение показателя модулярности, т. е. число связей, попадающих внутрь наблюдаемых в сети групп, минус число связей, которое попало бы внутрь групп в искусственной сети с идентичными параметрами, но сгенерированной так, что эти связи распределены по сети случайным образом [9, р. 2]. Вычисление общей модулярности матрицы с помощью поиска главного собственного значения и соответствующего этому значению собственного вектора получило название метода оптимальной модулярности [Ibid.]. Значение показателя модулярности для матрицы меняется от -1 до 1 . Положительные значения свидетельствуют о наличии в сети коммунитарной структуры.

В случае, когда у матрицы существует положительное собственное значение, и матрица разделима, алгоритм запускается вновь для каждой найденной подгруппы и, соответственно, субматрицы. Если субматрица разделяема, то из нее снова выделяют две подгруппы и рассчитывают модулярность. Если модулярность субматрицы увеличивает общую модулярность первичной матрицы, то выделенные подгруппы принимают; если модулярность субматрицы уменьшает общую модулярность (т.е. у матрицы нет

положительных собственных значений), то от такого разделения отказываются. Затем эту процедуру повторяют вновь для каждой из полученных на предыдущем шаге групп. У такого алгоритма существует естественная конечная точка: в тот момент, когда разделение очередной субматрицы невозможно, т.е. не приносит увеличения общей модулярности, алгоритм завершает свою работу. Таким образом, алгоритм предлагает новое определение того, что есть «сообщество» в сети: это неделимый субграф [9, р. 3–4].

Следующая операция – поиск среди всех узлов такого узла, который при перемещении в другую группу дает увеличение модулярности, и перемещение данного узла в найденную группу. Затем таким же образом последовательно перебираются все остальные узлы в сети. Основное условие при перемещении: каждый узел перемещается только один раз. Когда все узлы перемещены, находится такое промежуточное состояние системы, когда достигается максимальное значение модулярности [9, р. 5]. Алгоритм встроен в несколько пакетов для сетевого анализа¹ и показал свои преимущества в сравнении с другими методами при работе с сетями, содержащими более тысячи узлов².

Алгоритм ван-Донгена

Алгоритм кластеризации ван-Донгена основан на свойствах случайных графов Маркова (и поэтому имеет название *Markov clustering*) и был предложен специалистом в области биоинформатики С. ван-Донгеном в 2000 г. [10], а затем интегрирован в ряд пакетов для анализа сетевых данных, в том числе и используемый нами *UCINET*³.

¹ В нашем случае использовался пакет ORA (<http://www.casos.cs.cmu.edu/projects/ora/>).

² Ньюмен показал, что он работает эффективнее, чем, например, метод Гирвана-Ньюмена.

³ <http://www.analytictech.com/ucinet/>

Основная модель, на которой построен этот алгоритм, предполагает в сети наличие структуры, похожей на ту, что отображена на *рис. 1* и *2*, т.е. такой структуры, где много связей внутри кластера и мало связей между кластерами. Для наглядности мы предлагаем воспользоваться следующим примером. Представим, что мы находимся в городе, где площади – это узлы, а улицы, виадуки и тоннели – ребра, или связи между ними. Пусть в этом городе есть два соседства¹ (*A* и *B* на *рис. 1*), которые развивались независимо друг от друга, и только после их формирования возникла потребность в перемещении между ними, но она не слишком насущная – соседства соединила только одна транспортная артерия. Поэтому, если турист начинает прогулку на случайной площади этого города, и перемещается от одной площади (узла) к другой по существующей транспортной артерии (связи), то у него больше шансов остаться внутри одного соседства (кластера), чем выйти за его пределы. Кроме того, он может сделать круг по площади, и остаться на месте (петля), все в том же соседстве.

Случайные «прогулки» по такому городу, или графу, показывают, что «поток» бесцельно перемещающихся (праздношатающихся) людей будет иметь тенденцию циркулировать внутри соседств [10, р. 6]. Если нам нужно обнаружить эти соседства, мы будем искать места, где потоки бесцельно прогуливающихся туристов имеют тенденцию задерживаться. Для того чтобы имитировать такую случайную прогулку, ван-Донген предлагает использовать марковский процесс с дискретным временем [Ibid.]².

Как объяснить фланерам, оказавшимся в таком городе, что они должны прогуливаться так, будто следуют топологическому образцу цепи Маркова? «Случайное событие» в данном случае – это возможность оказаться на одной из площадей нашего города. Теперь представим, что с каждой из площадей начинают свою

¹ В английском мы употребили бы слово *neighborhood*.

² О марковском процессе см., например: [16, с. 386–440].

прогулку по одному туристу. Фланер с площади $A1$ (рис. 1), должен выбирать улицу, по которой он продолжит путь случайно, будто у него нет эстетических или каких-либо других предпочтений. Вероятность того, что он, пройдя одну улицу (одно ребро графа), останется на площади $A1^1$ или попадет на площади $A2$, $A3$ и $A4$ равна 0,25. Вероятность того, что фланер попадет на улицу $B1$, $B2$ или $B3$ равна 0. Если другой фланер начал свой путь на площади $A4$, то вероятность, что он попадет на площади $A1$, $A3$, $A4$ и $B1$ равна 0,25, а на площади $B2$ и $B3$ равна 0. Для иллюстрации данного случая можно создать матрицу (назовем ее $T0$) переходных вероятностей², где будет зафиксирована каждая из начальных площадей в момент времени $t0$, а также вероятности переходов фланеров на другие площади. Это и есть фиксированное настоящее (нулевое испытание) фланеров. Подобную матрицу можно создать и для момента времени $t1$ (возведя матрицу переходных вероятностей $T0$ в квадрат, т.е. умножив на саму себя [10, р. 8], и получив матрицу $T1$). Последняя операция предоставляет каждому из фланеров больше свободы, и каждый из них может прогуливаться, пользуясь не только улицами, но и дворами, что позволяет соединять ранее несвязанные между собой области графа. Матрицы переходных вероятностей для первых четырех моментов показаны на рис. 4.

Можно заметить, что в матрицах «ранних переходов» ($T0$, $T1$) групповая структура хорошо видна: в матричной сетке начального распределения вероятностей ($T0$) есть разреженная зона, заполненная нулями (выделено полужирным). Групповая структура видна

¹ Добавление петли в алгоритме позволяет решить проблемы, которые возникают из-за природы алгоритма: поскольку мы искусственно симулируем поток, то 0 в ячейке на диагонали может создавать зависимости в матрице переходных вероятностей (единицы потока будут воспринимать 0 как разрыв в структуре). Добавление 1 на главную диагональ – нейтральное решение.

² Вес всех исходящих связей при суммировании дает 1 (стохастическая матрица по столбцам).

T0		T1													
		A1	A2	A3	A4	B1	B2	B3	A1	A2	A3	A4	B1	B2	B3
A1	0,25	0,25	0,25	0,25	0,20	0,00	0,00	0,00	A1	0,24	0,24	0,24	0,19	0,05	0,00
A2	0,25	0,25	0,25	0,20	0,00	0,00	0,00	A2	0,24	0,24	0,24	0,19	0,05	0,00	0,00
A3	0,25	0,25	0,25	0,20	0,00	0,00	0,00	A3	0,24	0,24	0,24	0,19	0,05	0,00	0,00
A4	0,25	0,25	0,25	0,20	0,25	0,00	0,00	A4	0,24	0,24	0,24	0,24	0,11	0,08	0,08
B1	0,00	0,00	0,00	0,20	0,25	0,33	0,33	B1	0,05	0,05	0,05	0,09	0,28	0,31	0,31
B2	0,00	0,00	0,00	0,00	0,25	0,33	0,33	B2	0,00	0,00	0,00	0,05	0,23	0,31	0,31
B3	0,00	0,00	0,00	0,00	0,00	0,25	0,33	B3	0,00	0,00	0,00	0,05	0,23	0,31	0,31

T2		T3													
		A1	A2	A3	A4	B1	B2	B3	A1	A2	A3	A4	B1	B2	B3
A1	0,22	0,22	0,22	0,19	0,07	0,03	0,03	0,03	A1	0,19	0,19	0,19	0,17	0,10	0,08
A2	0,22	0,22	0,22	0,19	0,07	0,03	0,03	0,03	A2	0,19	0,19	0,19	0,17	0,10	0,08
A3	0,22	0,22	0,22	0,19	0,07	0,03	0,03	0,03	A3	0,19	0,19	0,19	0,17	0,10	0,08
A4	0,23	0,23	0,23	0,21	0,13	0,11	0,11	0,11	A4	0,21	0,21	0,21	0,20	0,15	0,14
B1	0,07	0,07	0,07	0,11	0,24	0,28	0,28	B1	0,10	0,10	0,10	0,12	0,20	0,23	0,23
B2	0,02	0,02	0,02	0,06	0,21	0,26	0,26	B2	0,06	0,06	0,06	0,08	0,17	0,20	0,20
B3	0,02	0,02	0,02	0,06	0,21	0,26	0,26	B3	0,06	0,06	0,06	0,08	0,17	0,20	0,20

Рис. 4. Матрицы переходных вероятностей для первых четырех переходов (T0, T1, T2 и T3)

и после первого шага ($T1$): ячейки, которые отображают вероятность перехода внутри кластера, заполнены числами с большими значениями, а ячейки, которые отображают вероятность перехода между кластерами, заполнены числами с меньшими значениями. На ранних шагах поток имеет тенденцию «накапливаться» внутри кластеров. Однако с течением времени (по мере того как мы возводим каждую новую матрицу переходных вероятностей в степень) этот эффект исчезает, и «поток» равномерно распределяется по матрице (значения в ячейках $T2$ и $T3$ выравниваются).

В своем алгоритме ван-Донген предлагает после первого шага, т.е. после возведения матрицы переходных вероятностей $T0$ в квадрат, усилить вероятность переходов внутри кластеров и уменьшить вероятность переходов между ними за счет возведения в положительную степень каждого столбца матрицы переходной вероятности (рис. 5)¹. Затем каждая из ячеек нормализуется относительно суммы по столбцу [10, p. 8].

Возведение матрицы переходных вероятностей в степень для отображения вероятностей перехода в следующей временной точке ван-Донген называет операцией «экспансия». А возведение в положительную степень каждой ячейки матрицы вероятностей и нормализацию столбца – процедурой «инфляция» [Ibid.].

Операции экспансии и инфляции поочередно повторяются до тех пор, пока не будет достигнута конвергенция, т.е. переход матрицы в такое стабильное состояние, когда при умножении матрицы на саму себя она перестает изменяться. Все ненулевые значения в строках такой матрицы, получившей название идемпотентной, одинаковы, а сама матрица представляет граф с несколькими компонентами, или подсетями, которые не связаны друг с другом. Они и есть искомые кластеры. Экспериментальным путем ван-Донген показал, что конвергенция чаще всего достигается; впрочем ему встречались и

¹ Обычно в степень 2. Чем больше степень, тем выше «гранулярность» матрицы, т.е. количество итоговых кластеров, которые будут получены.

Т1	Т1 после инфляции													
	A1	A2	A3	A4	B1	B2	B3	A1	A2	A3	A4	B1	B2	B3
A1	0,24	0,24	0,24	0,19	0,05	0,00	0,00	A1	0,25	0,25	0,25	0,20	0,01	0,00
A2	0,24	0,24	0,24	0,19	0,05	0,00	0,00	A2	0,25	0,25	0,25	0,20	0,01	0,00
A3	0,24	0,24	0,24	0,19	0,05	0,00	0,00	A3	0,25	0,25	0,25	0,20	0,01	0,00
A4	0,24	0,24	0,24	0,24	0,11	0,08	0,08	A4	0,25	0,25	0,25	0,32	0,06	0,02
B1	0,05	0,05	0,05	0,09	0,28	0,31	0,31	B1	0,01	0,01	0,01	0,05	0,38	0,33
B2	0,00	0,00	0,00	0,05	0,23	0,31	0,31	B2	0,00	0,00	0,00	0,01	0,26	0,33
B3	0,00	0,00	0,00	0,05	0,23	0,31	0,31	B3	0,00	0,00	0,00	0,01	0,26	0,33

Рис. 5. Матрица переходных вероятностей до и после процедуры инфляции

противоположные случаи [10, p. 101–122]. Невозможность достичь конвергенции наблюдается при отсутствии групповой структуры в изучаемой сети. Таким образом, отсутствие конвергенции свидетельствует о наличии однородной сети, неделимой на отдельные группы.

Для операции «инфляция» можно использовать возведение в разные степени – 2, 3, 4 и т.д.)¹. Бóльшая степень дает бóльшую гранулярность получаемой индемпотентной матрицы (т.е. мы получим больше кластеров небольшого размера).

Алгоритм работает с неориентированными графами (симметричными матрицами), как взвешенными, так и невзвешенными. Для упрощения объяснения в нашем примере мы использовали бинарную матрицу; однако в случае со взвешенной матрицей «поток» будет иметь тенденцию задерживаться на более широких транспортных артериях (т.е. связях с большим весом). Алгоритм устойчив к шумам, например к случайным пропускам в базе данных. Использование алгоритма обусловлено как типом данных, которые имелись у нас (симметричная и взвешенная матрица), так и общей логикой поиска и упорядочения информации в практике исследовательской работы.

Мы можем применять данный алгоритм при анализе социцированных исходя из того, что при работе с литературой ученый «застревает» в границах определенных теоретических или предметных областей: рано или поздно ссылки в работах, которые он читает, начнут повторяться, и он будет проходить по этим ссылкам заново, как проходит «поток» по матрице, где существуют плотные соседства и разрывы между ними. Таким образом, основное важное для наших интерпретаций различие между алгоритмами заключается в том, что алгоритм Ньюмена иерархичен, а алгоритм

¹ Сам показатель степени, в которую возводят значения матричных ячеек, в сетевом анализе называется фактором инфляции (*Gamma inflation factor*). Создатели UCINET рекомендуют использовать фактор инфляции в диапазоне от 1,5 до 3. При этом по умолчанию при запуске алгоритма марковской кластеризации этом пакете используется 2. Мы решили не менять эту установку.

кластеризации Маркова – нет. Группы, которые выделяет алгоритм Ньюмена, могут быть представлены в виде иерархической структуры, где каждая группа на предыдущем шаге была частью более крупной (за исключением самого первого шага). Группы, которые выделяет алгоритм ван-Донгена, являются набором неупорядоченных по отношению друг к другу формирований.

За использованием этих двух алгоритмов стоят разные гипотезы об особенностях образцов социитирования. Модель, заложенная в алгоритм Ньюмена предполагает, что может иметь место деление сети на две крупные обособленные группы, в которых мы в дальнейшем можем найти иерархию исследовательских областей или традиций¹. «Модель», получаемая благодаря алгоритму ван-Донгена, не предполагает, что есть какой-то упорядоченный образец взаимодействия между группами.

Результаты

Мы начали обработку данных с использования алгоритмов Ньюмена и ван-Донгена применительно к нашей исходной квадратной матрице социитирований (порог социитирований 1, при котором сохранены все узлы и связи). Затем оба алгоритма были повторены для случаев, когда из матрицы сначала были удалены связи с весом 1 и изолированные узлы (порог социитирований 2), а после этого связи с весом два и вновь появившиеся изолированные узлы (порог социитирований 3) (табл. 1).

¹ Скажем, можно предположить, что сеть социитирований из социологических публикаций конца XIX – начала XX вв. распадется на две группы работы, практически не ссылающихся друга на друга, а именно: статьи представителей французской и немецкой национальных социологических школ. Каждая из них могла в свою очередь делиться на подгруппы, занимающиеся изучением сходных явлений. Например, и немецкие, и французские социологи изучали такое явление, как религия, но при этом немецкие и французские социологии религии могли быть не связаны друг с другом через социитирования.

Таблица 1
 ГРУППЫ СОЦИТИРОВАННИЙ, ВЫЯВЛЯЕМЫЕ С ПОМОЩЬЮ АЛГОРИТМОВ ВАН-ДОНГЕНА И
 НЬЮМЕНА ПРИ РАЗНЫХ ПОРОГАХ СОЦИТИРОВАННИЙ

Порог социци- рований	Число		Алгоритм					
	узлов	связей	Ньюмена		ван-Донгена			
			Общее количество групп, найденных алгоритмом	Идентифицированные группы	Неидентифицированные группы и группы не отвечающие введенным нами критериям	Общее количество групп, найденных алгоритмом	Идентифицированные группы	Неидентифицированные группы и группы не отвечающие введенным нами критериям
1	1073	37798	7	6	1 (1)	59	24	35 (21)
2	968	8300	27	11	6 (3)	107	34	73 (59)
3	600	2272	59	18	41 (36)	90	33	57 (50)

Примечание. В скобках указано количество групп, которые состояли из двух-трех узлов (документов).

Для каждого из порогов в *табл. 1* указаны характеристики, соответствующей ему сети (число узлов и связей), а также общее количество групп, найденных алгоритмами, число групп, поддающихся идентификации с учетом введенных нами критериев¹, и число групп, не поддающихся идентификации или несоответствующих этим критериям. При этом подавляющее большинство последних включает группы из двух-трех публикаций. Пример идентификации группы социтирований (в данном случае это «Этнология самодийцев») представлен в *табл. 2* (см. Приложение).

Табл. 3 (см. Приложение) содержит описание выделенных групп ссылок, представляющих идентифицированные нами по этим ссылкам исследовательские области. Причем области эти довольно устойчивы, поскольку обнаруживаются либо обоими алгоритмами, либо одним алгоритмом, но на всех трех порогах.

Дроби в ячейках описывают отношение количества публикаций, соответствующих названию группы, к общему количеству публикаций в группе. Отсутствие дроби в ячейке обозначает, что на данном пороге группа с тематическим содержанием, обозначенным в заглавии, не была обнаружена. Далее мы проверяли, сличая число одинаковых публикаций, как группы, включавшие 50 публикаций и менее, соотносятся с группами, полученными тем же методом при более высоком и более низком пороге, и как они соотносятся с группами, выявленными с помощью другого метода на том же пороге, что и проверяемая группа. Например, в совокупность исследовательских областей «Этнология неславянских народов Европейской России» входит область «Этнология удмуртов». При использовании алгоритма ван-Донгена на пороге 2

¹ В предыдущих публикациях мы полагали, что можно надежно определить характер группы, если она включала на менее 5 публикаций и не менее 50 ссылок на эти публикации. Подробнее см.: [17, с. 154]. В данной статье, учитывая предварительный характер полученных результатов, с целью экономии времени мы подсчитывали лишь количество публикаций в группе.

она получила № 6¹. Совокупность публикаций, соответствующая этой ячейке *табл. 3* сопоставлялась с перечнем публикаций, соответствующих порогам 1 (№ 39) и 3 (№ 6) в данном алгоритме, а также порогу 2 (№ 3) в алгоритме Ньюмена. Мы считали, что существует соответствие между ячейками *табл. 3*, если хотя бы 1/3 из публикаций, принадлежащих меньшей сравниваемой ячейки присутствовала в ячейке, с которой она сравнивается.

Из *табл. 3* видно, что алгоритм Ньюмена совершенно неэффективен для наших целей и данных для порога 1 из-за предполагаемого присутствия неустойчивых связей между узлами. На следующих порогах, после удаления этих «шумов» алгоритм позволяет обнаружить большее число групп, которые можно идентифицировать. Причем на пороге 3 таких устойчивых групп больше, чем на пороге 2. Алгоритм Ньюмена на каждом пороге находит меньшее число подгрупп, чем алгоритм ван-Донгена. Мы предполагаем, что подобные различия между результатами обоих методов определяются тем, что подгруппы социцирований не соотносятся друг с другом иерархически, и осуществленное на первых шагах деление расчленяет подгруппы меньшего размера таким образом, что позднее они уже не могут быть найдены.

Алгоритм ван-Донгена более эффективен и для работы с исходной матрицей (порог социцирований 1), и с преобразованными матрицами (пороги социцирований 2 и 3). Поскольку в алгоритм не заложена иерархическая модель деления сети, он позволяет искать более плотные² и «весомые»³ ее участки, не упорядоченные каким-либо образом. Однако даже он на пороге 1 при работе с исходной матрицей социцирований в 1 073 узла среди 59 выделенных групп выявляет огромную группу № 6, в состав которой входят 533 узла, соответ-

¹ Номера приписывались полученным группировкам пакетами *ORA* и *UCINET*.

² Плотность здесь понимается технически – как отношение числа возможных связей к числу существующих.

³ «Весомый» здесь – это производное от веса связи, т.е. количество социцирований.

ствующих документам из разных областей восточнославянской этнологии и фольклора. При первой трансформации матрицы (порог социтирований 2) такого рода гигантские недифференцированные группы исчезают.

Выводы

Наше исследование когнитивной структуры российской этнологии с помощью применения алгоритмов Ньюмена и ван-Донгена к социтированиям, полученным из публикаций в российской научной периодике, демонстрирует возможность использования данных алгоритмов для обозначенных целей. Альтернативой было бы получение группирований, неидентифицируемых с позиций этнологии. При этом алгоритм ван-Донгена позволяет обнаружить большее количество подгрупп, предположительно соответствующих исследовательским областям этой дисциплины. Вместе с тем в ходе исследования мы шли по пути предлагаемому целым рядом исследователей, которые начинали свой анализ данных с некоторого порога цитирования, превышающего 1. Причем данный порог очень высок в макроподходе: например, японский исследовательский коллектив, разработавший карту естественных наук за 2006 г., из 50 тыс. высокоцитируемых работ за 2001–2006 гг. для своего анализа использовал лишь один верхний процент из этой совокупности [18, р. 2]. У представителей микроподхода он заметно ниже: у Гмюра – 42 цитирования [5, р. 34], у Г. Шилдта и Ю. Маттсона – 4 [19, р. 150].

Вместе с тем значительное количество неидентифицированных подгрупп при выполнении алгоритмов Ньюмена и ван-Донгена может вызываться потерей части данных в ходе введения нами порога цитирований 4. Отсечение данных, благодаря введению порога цитирований, представляется в значительной степени волюнтаристским, поскольку неизвестно, каким образом отброшенная совокупность цитирований влияет на результаты различных видов количественного анализа социтирований. Отсюда мы можем сделать предположение

методологического свойства: в ходе анализа социцированных научной периодики необходимо начинать количественный анализ данных с порога цитирований 1, а затем в ходе работы по мере необходимости отсекаать слабые связи за счет повышения порога социцированных. Мы планируем прояснить этот вопрос в будущем.

ЛИТЕРАТУРА

1. Уитли Р. Когнитивная и социальная институционализация научных специальностей и областей исследования // Научная деятельность: структура и институты. М.: Прогресс, 1980. С. 218–256.
2. Маршкова-Шайкевич И.В. Наукометрические методы анализа научной деятельности // Философия науки. Наука как инновационная деятельность. Уфа: Академия ВЭГУ, 2009. С. 149–227.
3. Moody J., Light R. A View from Above: the Evolving Sociological Landscape // American Sociologist. 2006. 37(2). P. 67–86.
4. Маршкова И.В. Система связей между документами, построенная на основе ссылок (по указателю «Science Citation Index») // Научно-техническая информация. Серия 2: Информационные процессы и системы: ежемесячный научно-технический сборник. 1973. № 6. С. 3–8.
5. Gmür M. Co-citation Analysis and the Search for Invisible Colleges: A Methodological Evaluation // Scientometrics. 2003. Vol. 57 (1). P. 27–57.
6. Girvan M., Newman M. E. J. Community Structure in Social and Biological Networks // Proceedings of the National Academy of Science. 2002. Vol. 99 (12). P. 7821–7826.
7. Palla G., Derényi I., Farkas I., Vicsek T. Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society // Nature. 2005. Vol. 435 (7043). P. 814–818.
8. Borgatti S.P., Everett M.G., Freeman L.C. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies, 2002.
9. Newman M. Modularity and Community Structure in Networks // ArXiv preprint physics / 0602124. 2006. P. 1–7. URL: http://140.123.102.14:8080/reportSys/file/paper/joji/joji_7_paper.pdf (дата обращения: 17.12.2012).
10. van Dongen S. Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht. May 2000. URL: <http://igitur-archive.library.uu.nl/dissertations/1895620/inhoud.htm> (дата обращения: 17.12.2012).
11. Blondel V.D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast Unfolding of Community Hierarchies in Large Networks // Journal of Statistical Mechanics: Theory and Experiment. 2008. Vol. 2008 (10). P10008. URL: http://proxy.library.spbu.ru:2277/1742-5468/2008/10/P10008/pdf/1742-5468_2008_10_P10008.pdf (дата обращения: 8.07.2013).

12. *Danon L., Duch J., Diaz-Guilera A., Arenas A.* Comparing Community Structure Identification // Journal of Statistical Mechanics: Theory and Experiment. Vol. 2005 (9). P09008. URL: http://proxy.library.spbu.ru:2277/1742-5468/2005/09/P09008/pdf/1742-5468_2005_09_P09008.pdf (дата обращения 8.07.2013).

13. *Fortunato S.* Community Detection in Graphs // Physics Reports. 2010. Vol. 486 (3-5). P. 75–174.

14. *Porter M. A., Onnela J.-P., Mucha P. J.* Communities in Networks // Notices of the AMS. 2009. Vol. 56 (9). P. 1082–1097.

15. *Оре О.* Графы и их применение. М.: Мир, 1965.

16. *Феллер В.* Введение в теорию вероятностей и ее приложения. М.: Мир, 1984. Т. 1. С. 386–440.

17. *Винер Б.Е., Дивисенко К.С.* Когнитивная структура современной российской социологии по данным журнальных ссылок // Журнал социологии и социальной антропологии. 2012. № 4. С. 144–166.

18. *Saka A., Igami M., Kuwahara T.* Science Map 2006: Study on Hot Research Areas (2001–2006) by Bibliometric Method. (Nistep Report No. 110). June, 2008. URL: <http://www.nistep.go.jp/achiev/ftx/eng/rep110e/rep110e.pdf> (дата обращения: 1.07.2009).

19. *Schildt H.A., Mattson J.T.* A Dense Network Sub-grouping Algorithm for Co-citation Analysis and Its Implementation in the Software Tool Sitkis // Scientometrics. 2006. Vol. 67. No 1. P. 143–163.

Приложение

Таблица 2

ЭТНОЛОГИЯ САМОДИЙЦЕВ
ГРУППИРОВАНИЕ №13 (ПРИ ПОРОГЕ СОЦИТИРОВАНИЙ 3)
соотношение источников – 18/30

Васильев В.И. Особенности развития этнических и языковых процессов в этноконтактных зонах Европейского Севера и Северной Сибири (по материалам этнографического обследования северсамодийских народов: ненцев, энцев и иганасан) // Этнокультурные процессы у народов Сибири и Севера. М.: Наука, 1985. С. 65–93
Васильев В.И. Проблемы формирования северсамодийских народностей. М.: Наука, 1979
Вербов Г.Д. Лесные ненцы // СЭ. 1936. № 2. С. 57–70
Вербов Г.Д. Пережитки родового строя у ненцев // Совесткая этнография. 1939. № 2. С. 43–66
Головнев А.В. Говорящие культуры: традиции самодийцев и угров. Екатеринбург: УрО РАН, 1995

<i>Головнев А.В.</i> Историческая типология хозяйства народов Северо-Западной Сибири. Новосибирск: Изд-во Новосиб. гос. ун-та, 1993
Грачева Г.Н. Традиционное мировоззрение охотников Таймыра (на материалах нгансан XIX – начала XX в.). Л., 1983
Долгих Б.О. Очерки по этнической истории ненцев и энцев. М.: Наука, 1970
Евладов В.П. По тундрам Ямала к Белому острову: экспедиция на Крайний Север полуострова Ямал в 1928–1929 гг. Тюмень: ИПОС СО РАН, 1992. «В качестве топографа, этнографа-экономиста, научного сотрудника Всесоюзного Арктического института Евладов участвовал в первых советских экспедициях на полуостров Ямал. Им была создана первая подробная карта Ямала и острова Белого, изучен язык ненцев-оленьеводов, составлен первый русско-ненецкий словарь, выпущен ряд интересных работ» (http://www.nasledie-rus.ru/podshivka/8727.php).
Зуев В.Ф. Описание живущих в Сибирской губернии в Березовском уезде иноверческих остяков и самоедов // Материалы по этнографии Сибири XVIII века (1771–1772) . М.; Л.: Изд-во АН СССР, 1947 (ГИЭ. Нов. сер.; Т. 5)
<i>Иванов С.В.</i> Скульптура народов севера Сибири XIX – первой половины XX вв. Л.: Наука, 1970
Иславин В. Самоеды в домашнем и общественном быту. СПб., 1847
<i>Кастрен М.А.</i> Путешествие Александра Кастрена по Лапландии, Северной России и Сибири, 1838–1844, 1845–1849. Т. 6. М., 1860.
Квашин Ю.Н. Этнические процессы и хозяйство у коренного населения Надымского района Ямало-Ненецкого автономного округа // Вестник археологии, антропологии и этнографии . 2005. № 5. С. 120–130. «Значительную часть коренного населения Надымского района составляют ненцы, предки которых жили на юге Западной Сибири между Обью и Енисеем (http://www.ref.by/refs/17/37764/1.html).
Кушелевский Ю.И. Северный полюс и земля Ямал. СПб.: [Тип. МВД], 1868. «Данное издание, повествующее о жизни и быте самоедов, их обрядах и преданиях, является ценным источником по этнографии. Книга содержит также русско-самоедский словарь, составленный Ю. И. Кушелевским» (http://alfaret.com/item.php?prod=801&subid=105&catid=111).
Лехтисало Т. Мифология юрако-самоедов (ненцев). Томск: Изд-во Том. гос. ун-та, 1998
Народы Сибири / Под ред. М.Г. Левина и Л.П. Потапова. М.; Л., 1956 (Сер. «Народы мира»).
Общественный строй у народов Северной Сибири XVII – начала XX в. М.: Наука, 1970

<i>Патканов С.К.</i> Статистические данные, показывающие племенной состав населения Сибири, язык и род инородцев. СПб., 1911. Т. 2: Тобольская, Томская и Енисейская губернии
Список населенных пунктов Уральской области. Т. 12: Тобольский округ. Свердловск: Изд. орготдела Уралоблисполкома, Уралстатуправления и окружных исполкомов, 1928
Старцев Г. Самоеды (ненца): Историко-этнографическое исследование. Л.: Ин-т народов Севера, 1930
Судьбы народов Обь-Иртышского Севера: (Из истории национально-государственного строительства. 1822–1941 гг.): сб. документов. Тюмень: ИПП «Тюмень», 1994
Харючи Г.П. Традиции и инновации в культуре ненецкого этноса (вторая половина XX века). Томск: Изд-во Томск. ун-та, 2001
Хомич Л.В. Ненцы: Очерки традиционной культуры. СПб.: Русский двор, 1995 [Хомич Л.В. Ненцы: Историко-этнографические очерки. М.; Л.: Наука, 1966]
Хомич Л.В. Проблемы этногенеза и этнической истории ненцев. Л.: Наука, 1976
<i>Хомич Л.В.</i> Современные этнические процессы на Севере европейской части СССР и Западной Сибири // Преобразования в хозяйстве и культуре и этнические процессы у народов Севера. М.: Наука, 1970. С. 28–61
<i>Шренк Л.И.</i> Об инородцах Амурского края. СПб.: Изд. Императорской Академии наук, 1899. Т. 2
<i>Шренк Л.И.</i> Об инородцах Амурского края. СПб.: Изд. Императорской Академии наук, 1903. Т. 3
<i>Шренк А.</i> Путешествие к северо-востоку европейской России через тундры самоедов к Северным Уральским горам. СПб., 1855
<i>Штернберг Л.Я.</i> Гиляки, орочи, гольды, негидальцы, айны: статьи и материалы. Хабаровск: Дальгиз, 1933

Примечание. Жирным выделены документы, тематика которых определяет название таблицы. В ячейки с работами Евладова и Квашнина добавлена информация, помогающая определить содержание этих сочинений, поскольку из их названий это сделать невозможно.

Вероятно, и часть остальных документов посвящена этнологии самодийцев. Но по названиям и информации, найденной в Интернете, установить это не представляется возможным.

Таблица 3
УСТОЙЧИВЫЕ ГРУППЫ, ВЫДЕЛЕННЫЕ АЛГОРИТМАМИ НЬЮМЕНА И ВАН-ДОНГЕНА ПРИ ПОРОГЕ ЦИТИРОВАНИЙ В 4 ПУБЛИКАЦИИ

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцированных для алгоритма							
	Ньюмена				ван-Донгена			
	1	2	3		1	2	3	
	Восточнославянская этнология и фольклор							
Восточнославянская этнология и фольклор	№1 206/274	№1 241/274	№1 107/130		№1 159/203	№1 103/122		
Духовная культура восточных славян				№1 35/53				
Русский героический эпос			№19 7/8	№25 3/3		№31 4/7		
Фольклор Русского Севера		№25 4/4			№69 4/5			
Русские предания			№22 4/7			№60 3/3		
Теория фольклора						№10 7/11		
Теоретическая фольклористика				№18 4/4				

Примечание. В таблице не показаны неидентифицированные группы. Полу жирным выделены смежные строки, где имеется перекрестная корреляция между группами.

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социгирований для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования			№50 3/3		№90 3/4	№80 3/3
Фолькорное текстообразование					№18 4/6	
Религиозный компонент в культуре Русского Севера					№64 5/5	
Этнографическая семиотика		№13 4/5				
Этнология Средней Азии и Казахстана и исламоведение						
Исламоведение или Религиозные представления народов Средней Азии			№21 11/22	№26 11/21	№26 7/16	№33 9/18
Религиозные верования народов Средней Азии и Казахстана					№36 9/20	
Этнология узбеков				№35 2/4		

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцированных для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования	Этнология Кавказа					
История адыгов	№5 69/102	№11 50/61	№32 12/13	№55 11/17	№55 11/17	№55 3/6
Этнология адыго-абхазов	№5 69/102	№11 50/61	№32 12/13	№42 18/24	№55 11/17	№56 6/7
Этнология адыгов	№5 69/102	№11 50/61		№28 3/4	№56 6/7	№56 6/7
Этнология Дагестана			№18 6/8	№24 9/17	№31 8/17	№30 6/8
Религия в Дагестане			№23 3/5			
Этнология Северного Кавказа или Этнология Дагестана				№28 6/6	№35 7/12	
Ислам на Кавказе					№61 3/5	
Этнология Сибири	Этнология Сибири					
	№3 134/259		№7 29/29	№8 5/5	№8 8/9	№8 6/6

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социгирований для алгоритма									
	Ньюмена					ван-Донгена				
	1	2	3	1	2	3	1	2	3	
Имя группирования	№4	№7	№10	№13	№13	№15	№17	№13	№15	№26
Этнология народов уральской языковой семьи или Этнология Западной Сибири	32/53	61/96	20/39	4/5	15/19	32/51	11/12	4/5	18/30	11/12
Этнология самодийцев			№34	№58	№59	№58	№58	№58	№58	№58
Религия обских угров			9/13	9/13	11/13	9/13	9/13	9/13	9/13	9/13
Этнология Южной и Восточной Сибири и Дальнего Востока		№6 48/89								
Этнология народов тунгусской группы						№22	№20	№22	№22	№20
Алтайцы и бурханнизм					№37	№51	№64	№51	№51	№64
Этнология алтайцев					8/8	6/6	4/5	6/6	4/5	4/5
Этнология хакасов						№95	№14	№95	№95	№14
						5/5	6/7	5/5	6/7	6/7

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцированных для алгоритма						
	Ньюмена			ван-Донгена			
	1	2	3	1	2	3	
Имя группирования							
Этнология якутов					№21 3/4	№19 3/3	
Этнология народов Северо-Восточной Азии					№24 18/33	№23 9/16	
Религиозные представления и мифология народов Сибири			№9 30/73		№16 27/56	№24 10/22	
Религиозные представления и мифология народов Сибири			№9 30/73		№16 27/56	№45 10/12	
Материальная культура народов Сибири						№43 7/10	
Этнология народов монгольской группы							
История монголов					№72 3/5		
Этнология калмыков			№48 4/4	№57 4/4	№88 4/4	№78 4/4	

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социгирований для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования				№47 3/5	№67 3/4	№66 4/7
Этнология бурят					№34 6/6	№66 4/7
Этнология бурят						
Бурятский шаманизм				№62 4/7		№50 4/5
Этнология неславянских народов Европейской России						
Этнология удмуртов	№3 8/31		№5 7/26	№39 6/7	№6 8/13	№6 5/10
Религиозные верования удмуртов						№54 2/3
Карельская этнография и фольклор	№23 3/3					
Этнология чувашей	№7 2/3			№53 2/5	№77 3/6	№70 3/5
Другие этнологические группирования						
МАЭ (Кунсткамера)	№6 10/19	№5 17/33	№31 7/12	№38 10/14	№32 9/18	№52 7/12

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцирований для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования		№10 49/80				
Шаманство и шаманизм + ислам и доисламские пережитки в Средней Азии, Казахстане и на Кавказе				№23 9/13	№30 9/18	№29 7/12
Шаманство и шаманизм			№17 12/12			
Первобытное общество, системы родства, домонотеистические формы религии и их пережитки???		№4 27/45				
Первобытная религия (НЗ; вДЗ) или Первобытное мышление и первобытная религия (вД2)			№6 5/9		№7 6/11	№7 4/8
Ранние формы религии и их пережитки					№60 4/7	

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социгирований для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования			№26 7/11	№31 8/15	№38 7/14	№42 6/11
Системы родства			№43 2/3	№55 3/5	№81 5/9	№73 2/3
Этническая экология						
Этничность, нации, национализм, этнонациональная политика						
Этничность + социология	№2 150/360	№2 140/192				
Теория этноса				№7 9/11	№5 12/24	№5 9/11
Этнология Л.Н. Гумилева и ее критика				№9 6/11	№9 6/16	№9 6/7
Теории этноса Ю.В. Бромлея и Л.Н. Гумилева			№4 16/25			
Этничность, национализм, этнонациональная политика			№2 25/49	№3 15/29	№3 30/55	№3 14/20

Продолжение табл. 3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцированных для алгоритма					
	Ньюмена			ван-Донгена		
	1	2	3	1	2	3
Имя группирования			№2 25/49	№3 15/29	№3 30/55	№49 1/5
Этничность, национализм, этнонациональная политика						№37 4/5
Историческая память						№2 5/17
Диаспоры				№2 15/35	№2 14/30	
Мигранты		№12 3/3				
Нетнологические группирования						
Социология		№8 33/56				
Цивилизационные подходы			№16 4/5	№22 4/8	№29 3/6	
Характеристики современного общества			№13 3/4	№16 3/5	№19 5/10	№17 3/4
Социальная стратегификация			№40 3/3	№34 6/10	№44 7/8	№69 3/3

Исследовательские области (группы) обнаруженные с помощью алгоритмов Ньюмена и ван-Донгена	Пороги социцированных для алгоритма								
	Ньюмена			ван-Донгена					
	1	2	3	1	2	3	1	2	3
Имя группирования		№16 3/5		№52 3/5	№76 3/5				
Экономическая социология					№37 3/4				
Экономическая социология				№11 4/6	№12 5/8		№11 5/6		
Социология П. Бурдьё				№4 5/14	№4 11/18		№4 4/6		
Социальный конструктивизм			№3 10/12						
Семиотика социальной коммуникации (междисциплинарная)		№9 10/22	№36 3/3	№46 7/9	№65 5/7		№62 3/3		
Молодежные субкультуры		№15 3/3		№51 3/3	№73 3/3				