

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 29.07.2022 18:19:57

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

Методические указания к практическим занятиям

Практическая работа № 1

Задание выполняется индивидуально.

Отчет оформляется на русском языке в формате Word. К отчету прилагается файл выгрузки в формате RIS.

Программное обеспечение:

SCIMAT – программное обеспечение для наукометрического анализа на основе совместной встречаемости слов

<https://sci2s.ugr.es/scimat/software/v1.1.04/SciMAT-v1.1.04.zip>

Java: <http://java-runtime.ru/download> (при необходимости)

Требования к выгрузке:

Перед выгрузкой убедитесь, что публикации отсортированы по числу цитирований (от большего к меньшему).

Формат RIS, 2000 публикаций.

Приветствуется анализ публикаций российских ученых или ученых ЮФУ, но при условии, что можно выгрузить не менее 400 публикаций.

Число временных периодов (годы, пятилетки, десятилетки): не менее двух, оптимально 3-7

Содержание отчета:

Ф.И.О., группа

Направление исследований

Чем обусловлен выбор направления исследований (научные интересы, программа магистратуры, хайп и т.д.)

Как направление соотносится со "Стратегией развития информационного общества в РФ на 2017-2030 гг." и "Прогнозом научно-технологического развития России. Информационно-коммуникационные технологии"?

Индекс(ы) цитирования, в которых велся поиск

Ключевые слова. Структура поискового запроса (И, ИЛИ).

Краткая характеристика результатов поиска: число публикаций и цитирований, распределение по годам (привести диаграмму)

Самые цитируемые (ядерные) публикации, не менее 5 (привести список в виде библиографических ссылок)

Ведущие организации, не менее 5

Ведущие журналы, не менее 5

Анализ исследовательских фронтов (кластеров) с помощью стратегической диаграммы (привести настройки SciMAT, скриншот Overlapping map с окна Longitudinal view, скриншот окна Period view)

Характеристика структуры кластеров ключевых слов (скриншот окна Cluster's network), 3-5 кластеров

Анализ жизненного цикла исследовательских фронтов по квадрантам в разрезе периодов, 3-5 фронтов

Практическая работа № 2

Задание выполняется индивидуально либо совместно двумя студентами.

Подготовить презентацию (~10 слайдов) технологии из гартнеровского цикла за 2021 г. либо 2020 г.

Требования к презентации:

1. Титульный лист / Cover Slide
2. Оглавление / Contents
3. Глоссарий ключевых терминов / Glossary
4. Краткое изложение технологии (1-2 страницы) / Technology in Brief
5. Фундаментальная теория, математические методы, законы (2-3 страницы) / Methodology
6. Текущее состояние исследований, нерешенные проблемы / State of the Art and Open Issues
7. Лидеры индустрии, стартапы / Industry Leaders, Startups
8. Список источников (не менее одной книги или учебника, не менее 2-3 статей) / Bibliography

Дополнительный теоретический материал к занятиям

Кластеризация фронтов исследований

Во многих ситуациях удобнее работать не с отдельными исследовательскими фронтами, а с кластерами фронтов, т.е. метакластерами (кластерами кластеров), иерархическая структура которых представляет собой модель науки в целом – *карту науки*. Исследовательские фронты агрегируются в научные специальности; те, в свою очередь – в более крупные исследовательские направления и области знания, а все в совокупности формируют иерархическую когнитивную структуру современной науки.

Пошаговый многоуровневый подход к построению иерархической структуры кластеров работает по принципу «выход предыдущего шага поступает на вход следующего шага» [Бредихин, 2013]:

На первом уровне: на входе – документы, на выходе – кластеры первого уровня C_1 .

На втором уровне: на входе – кластеры первого уровня, на выходе – кластеры кластеров C_2 и т.д.

Как правило, самый высокий уровень кластеризации в картах науки – пятый (C_5), называемый макроуровнем. Он представляет собой глобальную карту областей исследований.

На первом уровне анализируются социтирования двух документов, на втором и последующих уровнях – социтирования документов, включенных в кластеры. При этом используется индекс Солтона – косино-нормированная частота социтирования:

$$S_s(i, j) = \frac{\text{coc}(i, j)}{\sqrt{\text{cit}(i) \times \text{cit}(j)}}$$

Коэффициенты кластеризации следует варьировать между уровнями, предупреждая формирование макрокластеров с аморфной структурой [Бредихин, 2013].

При организации работы с результирующей картой науки целесообразно организовать иерархическое развертывание (drill down).

В контексте нашего исследования, когда выгрузка фронтов осуществляется в рамках одной предметной области и после предварительной фильтрации множество фронтов представляется обозримым для манипуляции, терминальным можно считать уровень C_2 , максимум C_3 . Фактически, имеет

место концентрации графа $C(X)$, в котором каждому кластеру фронтов сопоставлена своя вершина (рис. 16).

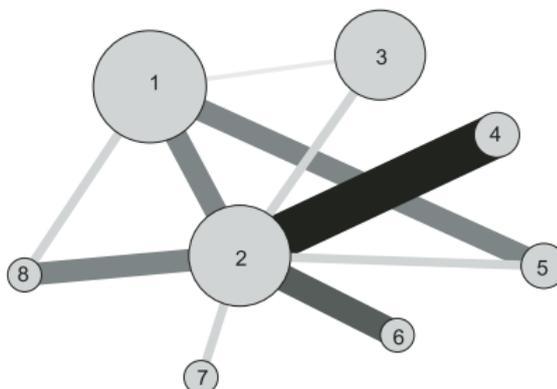


Рис. 16. Концентрация графа – сеть кластеров [Собо, 2012]

В ходе эксперимента экспертам было предложено объединить исследовательские фронты в группы (кластеры) на основе субъективного критерия сходства. Критерий сходства и названия групп заранее не оговаривались.

Как результат, в 12 кластеров был определен 81 фронт (из 100), причем 7 фронтов оказались отнесены одновременно к двум, а 1 фронт (R004 ТЕОРИЯ И ПРИЛОЖЕНИЯ ИМПУЛЬСНЫХ И ВЕРОЯТНОСТНЫХ НЕЙРОННЫХ СЕТЕЙ) – к трем кластерам (рис. 17).



Рис. 17. Распределение фронтов по кластерам

Даже несмотря на то, что данные о цитировании «ядерных» публикаций выгружались из Scopus, результаты выполненной в Gephi кластеризации фронтов с помощью силового алгоритма Force Atlas 2 в целом согласуются с кластеризацией, выполненной экспертами.

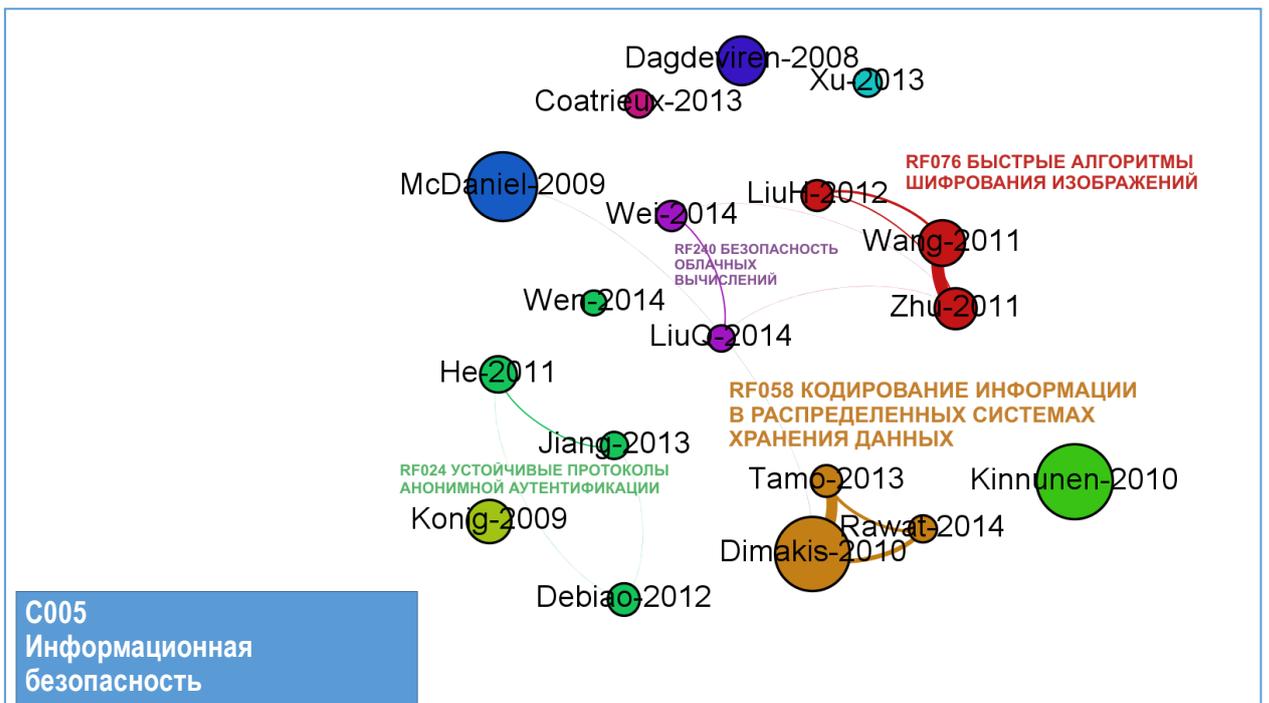


Рис. 19. Кластер фронтов «Информационная безопасность»

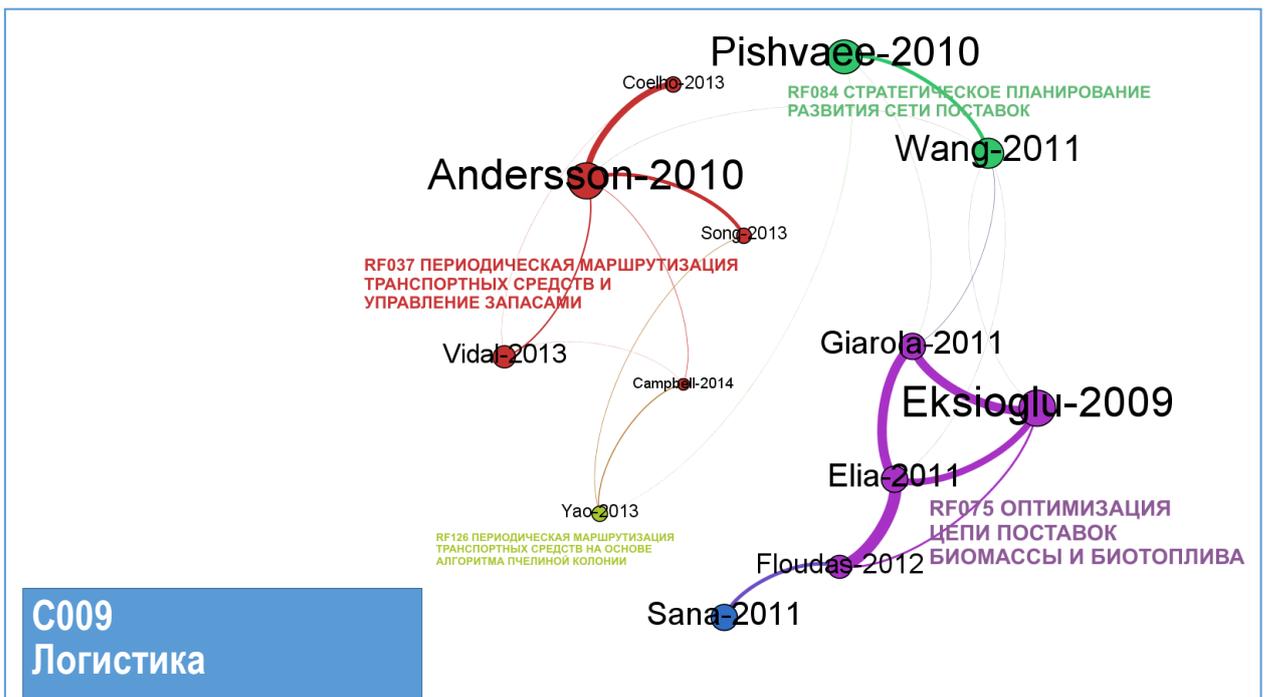


Рис. 20. Кластер фронтов «Логистика»

Малую плотность графа и наличие нескольких компонент связности (изолированных вершин) на рис. 20 можно объяснить тем, что в число «ядерных» публикаций фронта (в представленной для анализа выборке) вошло мало статей (иногда одна).

Другим перспективным подходом к кластеризации фронтов является сочетание методов на основе цитирования с методами на основе совместной встречаемости слов.

[Cobo et al., 2011] предложили категоризировать кластеры в виде стратегической диаграммы (strategic diagram) на основе метрик центральности (Callon's centrality) и плотности (Callon's density) по Каллону [Callon, 1991].

Центральность по Каллону (или внешняя связность) является мерой взаимодействия вершин одного кластера с вершинами в других кластерах. Рассчитывается по формуле:

$$c = 10 * \sum e_{uv} ,$$

где u – элементы, принадлежащие кластеру, v – элементы в других кластерах.

Плотность по Каллону (или внутренняя связность) является мерой внутренней связности вершин кластера. Рассчитывается по формуле:

$$d = 100 * \frac{\sum e_{ij}}{n},$$

где i и j – элементы, принадлежащие кластеру, n – общее число элементов.

Для построения стратегической диаграммы значение центральности откладывают по оси абсцисс, а значение плотности – по оси ординат (рис. 21).

Квадрант I («Мейнстрим») в правом верхнем углу содержит движущие, актуальные темы (кластеры), которые одновременно обладают внутренней когерентностью и являются центральными в сети.

Квадрант II («Башня из слоновой кости») в левом верхнем углу содержит развитые, но изолированные, самостоятельные темы (кластеры).

Квадрант III («Хаос») в левом нижнем содержит либо новые, развивающиеся темы либо затухающие, которые одновременно не являются сильно связанными и центральными в сети.

Наконец, квадрант IV («Фургон») в правом нижнем углу содержит темы с низкой внутренней когерентностью, но сильными связями с другими темами (кластерами). Это общие и надпредметные темы.



Рис. 21. Стратегическая диаграмма кластеров

Стратегическая диаграмма позволяет анализировать жизненный цикл кластеров [Callon, 1993; Liu, 2014; Kostakos, 2015]. Темы зарождаются (Квadrant III), затем стабилизируются (Квadrant IV), становятся мейнстримом (Квadrant I) и затухают (снова Квadrant III). Темы могут никогда не достичь Квadrанта I, но могут отправиться напрямую в Квadrant II, они могут надолго задержаться в Квadrante III или Квadrante IV. Но основная идея жизненного цикла – «здоровое» движение тем по квадрантам графика.

Предметом дискуссий последнее время является то, насколько критично для области знаний отсутствие движущих, мейнстримовых тем (кластеров) [Liu, 2014; Kostakos, 2015].

Мы построили стратегическую диаграмму для фронта «Беспроводные нательные сенсорные сети» (Wireless Body Area Network, WBAN).

Можно видеть, что в 2012-2013 годах: начал расти кластера «Ad Hoc сети»; в число изолированных кластеров вошли исследования на тему «Приемопередатчик»; в число общих, надпредметных тем – «Безопасность»; в числе движущих, мейнстримовых тем были «Удаленный мониторинг пациентов» и вопросы применения стандарта «IEEE 802.15.4», который определил физический слой и управление доступом к среде (Media Access Control, MAC) для беспроводных персональных сетей с низким уровнем скорости.

В 2014-2015 годах: тема «Удаленный мониторинг пациентов» стабилизировалась в Квадранте IV; в число развитых, изолированных кластеров вошли «Биосенсоры»; в мейнстриме остается кластер «IEEE 802.15.4», а именно вопросы энергоэффективности, моделирования каналов и протокола ZigBee (рис. 24); в движущие кластеры вырвалась тема «Активного старения» (Ambient Assisted Living), тесно связанная с вопросами здравоохранения, телемедицины и распознавания активности (рис. 25).

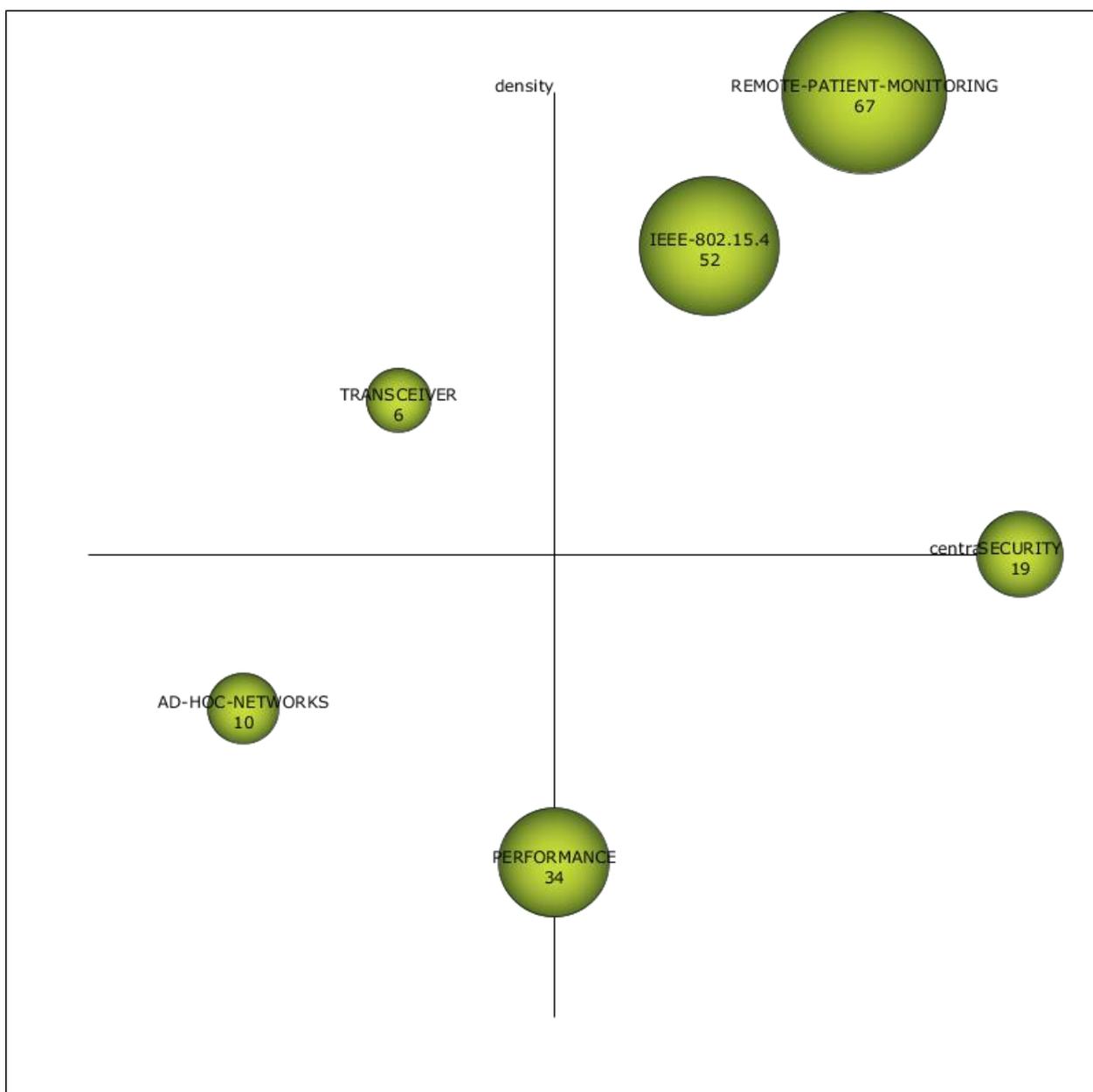


Рис. 22. Стратегическая диаграмма для фронта «Беспроводные нательные сенсорные сети» в 2012-2013 гг.

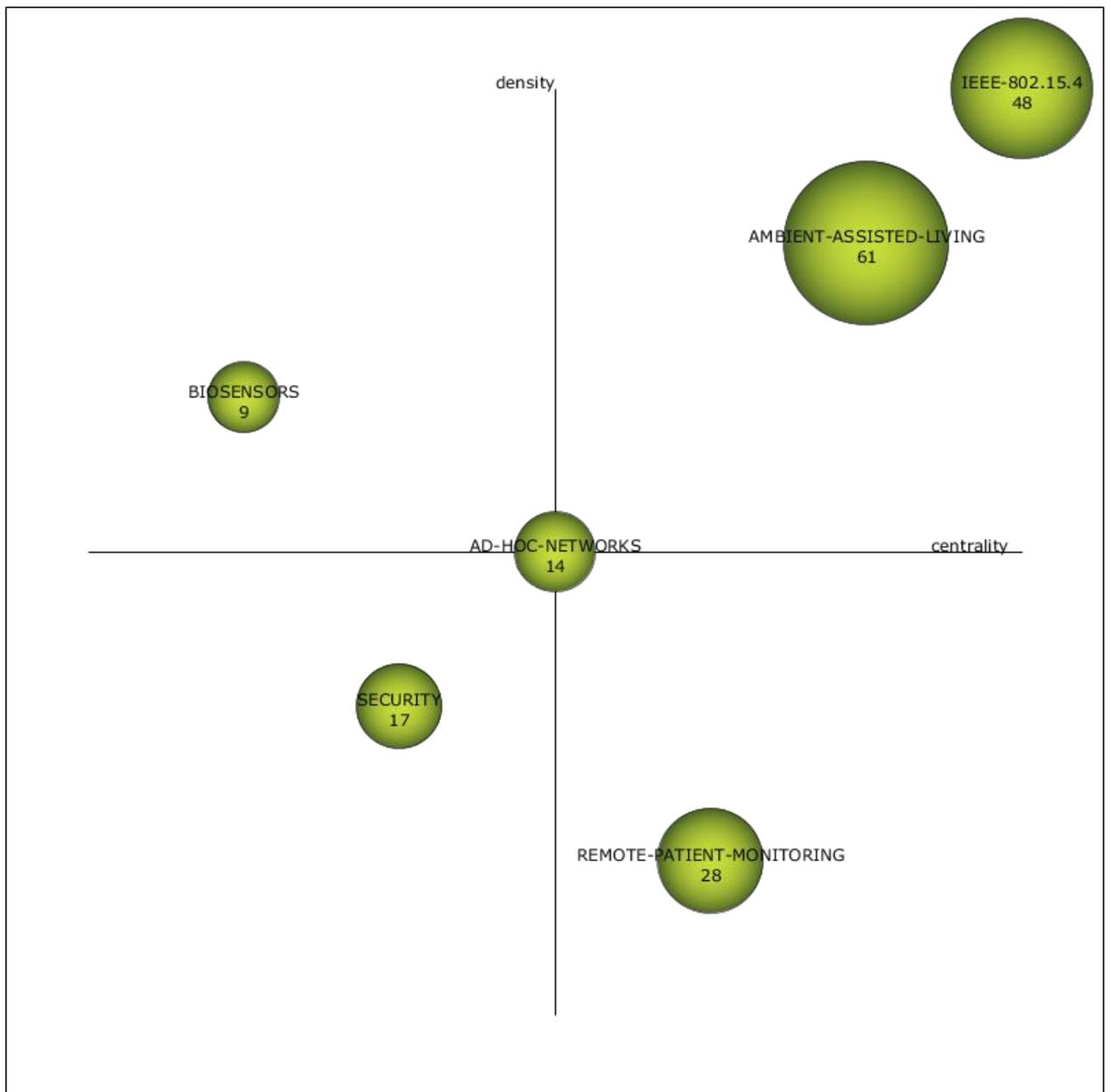


Рис. 23. Стратегическая диаграмма для фронта «Беспроводные нательные сенсорные сети» в 2014-2015 гг.

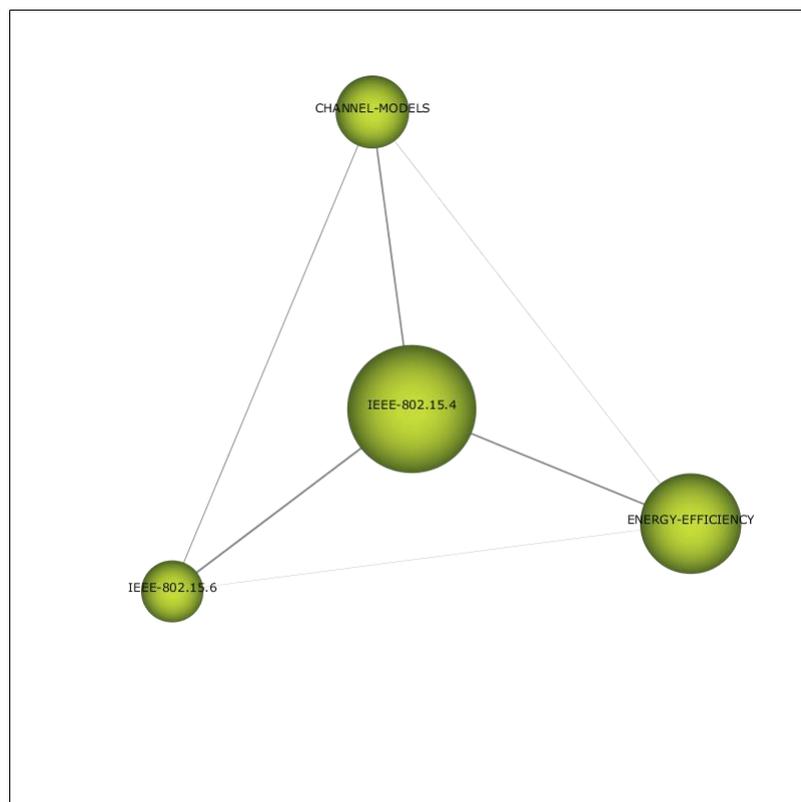


Рис. 24. Связи кластера «IEEE 802.15.4»

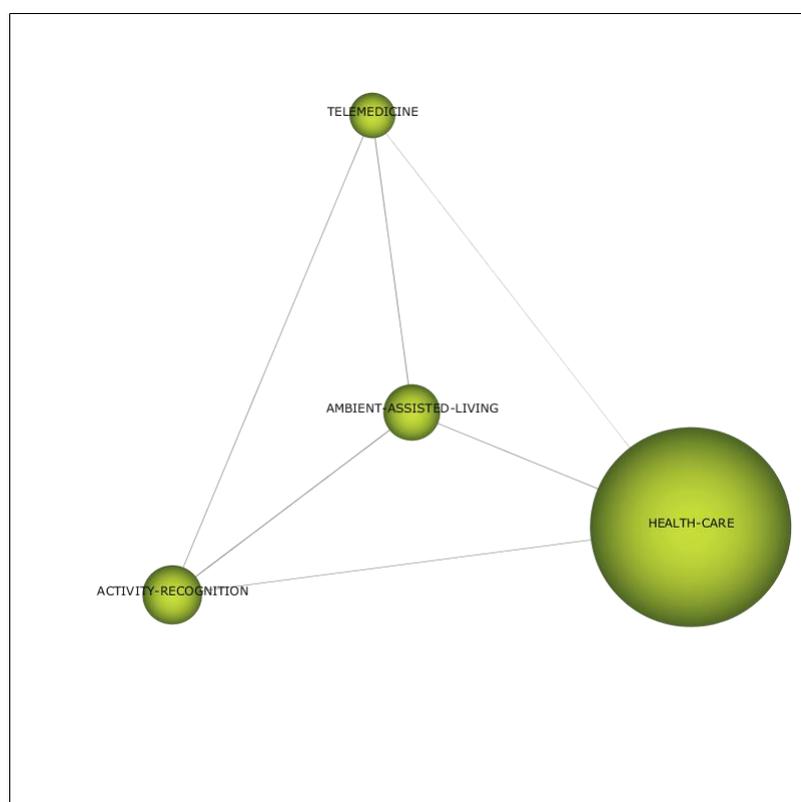


Рис. 25. Связи кластера «Ambient Assisted Living»

Установление связей между фронтами исследований, объединение, разбиение и отображение фронтов

Исследовательский фронт исчерпывающе определяется множеством его «ядерных» статей. Исследовательский фронт является минимальной структурной единицей тематического строения науки. С позиции теории систем и в контексте настоящего исследования исследовательский фронт в момент его идентификации является для нас входной единицей анализа, условно нечленимым элементарным (атомарным) носителем содержательных свойств системы, при расчленении которого будут потеряны качества элемента. Дополнительное «рафинирование» фронтов путем удаления отдельных «ядерных» публикаций или добавления новых (т.е. декомпозиция и агрегирование фронтов) будет означать введение новых конструктов и по сути подмену объекта анализа. Такое усложнение, на наш взгляд, не является оправданным и необходимым.

Если исследования в рамках фронта являются междисциплинарными и/или «ядерные» публикации охватывают несколько вопросов, целесообразно провести кластеризацию тем исследований (research themes) внутри фронта. Такая техника применялась на ScienceWatch в 2009-2010 годах (рис. 26) и актуальна для фронтов с большим числом «ядерных» статей, однако на нашей выборке их число не превышало 9.

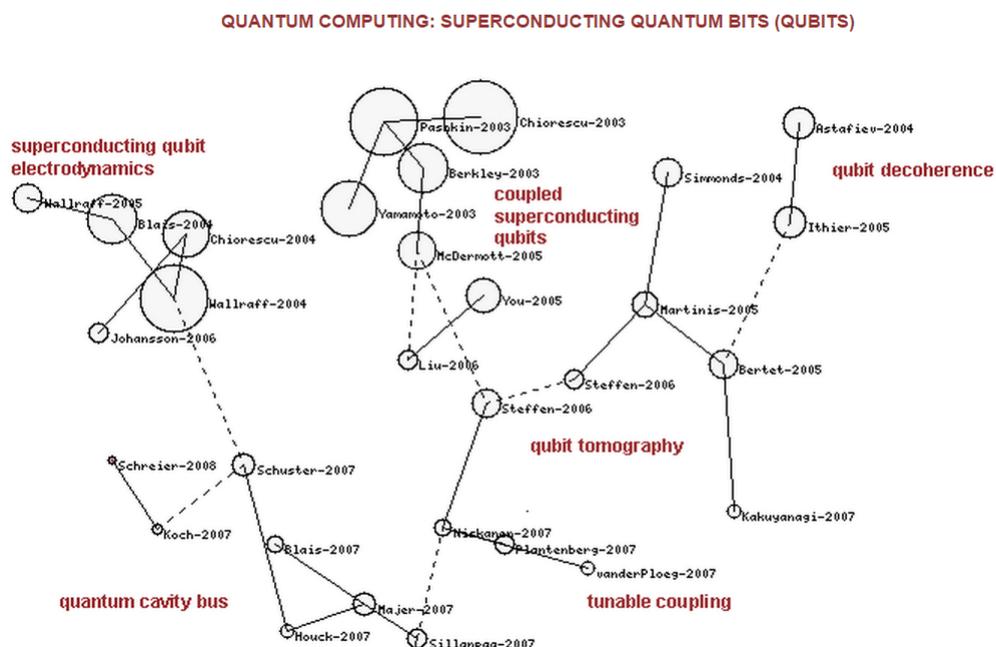


Рис. 26. Аннотации тем исследований внутри одного фронта

[<http://archive.sciencewatch.com/ana/st/hts/rfmap4/>]]

В то же время большой интерес представляет анализ динамики фронтов (их эволюции во времени), как-то: внутренний рост, слияние с другими фронтами, распад на несколько фронтов.

Поэтому в рамках этого параграфа мы обсуждаем не инструменты слияния и разбиения фронтов, а подходы к анализу их естественной (наблюдаемой) динамики, выражаемой в виде роста/сужения, слияния, разбиения.

Для работы с историей фронтов важны нижеследующие пояснения и определения.

Высокоцитируемыми признаются статьи из 1% самых цитируемых работ в каждой из 22 предметных областей ISI в годовой когорте. Ранее мы отмечали, что пик цитирования «ядерных» статей приходится на короткий период времени, а средняя продолжительность времени, в течение которого их продолжают цитировать, существенно меньше, чем у классических статей и тесно связана со скоростью развития соответствующей области исследований. Согласно отчетам Thomson Reuters, общее число фронтов выросло с 8000 в 2013 году до 9700 в 2014 году, из них в области Computer Science – более 400.

База данных Essential Science Indicators обновляется каждые два месяца (6 раз в год). Так, первый выпуск в 2015 году (обозначается 2005-2015_1) включает в себя публикации, вышедшие за предыдущие 10 лет и 2 месяца 2015 года, второй – за 10 лет и 4 месяца 2015 года и т.д. При этом для каждой предметной области пересчитываются пороговые значения цитирования, при которых статья в признается высокоцитируемой (рис. 27). В предметной области «Информатика» (Computer Science) пороговые значения – одни из самых низких: 5 цитирований для статей, опубликованных в 2014 году, 79 цитирований для статей, опубликованных в 2005 году. Более низкие пороги установлены только для предметной области «Математика».

Также устанавливаются пороговые значения (рис. 28) для признания высокоцитируемыми отдельно взятых авторов (1% самых цитируемых авторов), стран (50% самых цитируемых стран), учреждений (1% самых цитируемых учреждений) и журналов (50% самых цитируемых) – так называемых «восходящих звезд» (rising stars). Можно видеть, что в предметной области Computer Science на начало 2015 года автор считается высокоцитируемым, если его работы цитируются не менее 242 раз.

Новыми, зарождающимися фронтами (emerging fronts)¹ являются те фронты, «ядерные» публикации которых не входили ни в один фронт в предыдущих выпусках.

¹ <http://archive.sciencewatch.com/dr/erf/>

Field	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Agricultural Sciences	114	100	89	76	59	52	36	25	14	6	2
Biology & Biochemistry	229	205	185	164	137	107	76	51	28	9	2
Chemistry	182	164	149	139	121	108	86	60	34	11	2
Clinical Medicine	218	189	166	143	121	96	71	48	26	9	2
Computer Science	79	73	81	74	60	51	37	24	13	5	2
Economics & Business	161	130	115	86	67	54	36	22	11	4	2
Engineering	88	84	80	67	62	51	39	26	15	6	2
Environment/Ecology	193	173	154	136	112	89	74	48	24	8	3
Geosciences	156	148	120	110	96	72	58	38	20	7	3
Immunology	306	261	258	207	183	137	100	66	37	12	2
Materials Science	157	141	138	120	111	100	78	55	30	10	2
Mathematics	66	58	53	48	39	33	23	16	9	4	2
Microbiology	210	192	155	150	134	104	68	48	26	8	2
Molecular Biology & Genetics	435	373	375	308	262	199	148	94	50	14	3
Multidisciplinary	987	847	651	573	544	498	307	207	103	25	2
Neuroscience & Behavior	253	233	204	180	147	119	84	56	29	10	2
Pharmacology & Toxicology	170	164	141	128	104	80	59	40	22	8	2
Physics	156	138	130	125	105	91	69	50	28	10	2
Plant & Animal Science	131	117	103	87	76	61	44	30	16	6	2
Psychiatry/Psychology	205	183	156	126	102	83	56	35	19	7	3
Social Sciences, general	119	103	89	70	57	46	32	22	12	5	3
Space Science	228	209	188	157	146	119	95	66	40	15	2

Рис. 27. Пороговые значения для статьи

Быстро растущими (fast moving)² являются фронты с максимальным в процентном отношении (для их предметной области) приростом числа «ядерных» публикаций и их цитирований. Фронт может увеличиться в размере либо в результате внутреннего роста, либо в результате слияния с одним или несколькими другими фронтами.

² <http://archive.sciencewatch.com/dr/fmf/>

Field	Author	Country	Institution	Journal
Agricultural Sciences	327	710	1481	1216
Biology & Biochemistry	775	479	5155	5351
Chemistry	1267	786	5477	4585
Clinical Medicine	1582	2639	1704	3255
Computer Science	242	165	2267	881
Economics & Business	295	155	3223	849
Engineering	360	506	1549	1319
Environment/Ecology	556	965	3069	1998
Geosciences	869	847	4396	1727
Immunology	833	1163	3548	5937
Materials Science	820	482	3182	1421
Mathematics	257	264	3155	1071
Microbiology	594	673	4317	3441
Molecular Biology & Genetics	1709	822	9434	5706
Multidisciplinary	1402	401	7921	161
Neuroscience & Behavior	1081	312	4475	5365
Pharmacology & Toxicology	454	580	2670	3826
Physics	6480	1191	10828	3547
Plant & Animal Science	487	950	1924	1770
Psychiatry/Psychology	631	228	3388	1846
Social Sciences, general	302	816	1048	565
Space Science	4107	363	24140	1243

Рис. 28. Пороговые значения для авторов, стран, учреждений и журналов

Быстро развивающимися статьями (fast breaking papers)³ называются такие высокоцитируемые статьи, число цитирований резко (больше, чем для остальных статей в их предметной области) выросло по сравнению с предыдущим выпуском, выходящим раз в два месяца.

Актуальными, «горячими» статьями (hot papers)⁴ называются публикации, которые вышли за последние два года и чаще других (99,9-й перцентиль) цитировались за последние два месяца. Другими словами, это публикации, которые активно, чаще остальных (для их предметной области) цитировались сразу после выхода.

Пороговые значения для признания статьи «горячей» приводятся на рис. 29. Для предметной области «Информатика» это более 4 цитирований в январе-феврале 2015 года.

С 2008 по 2012 год Sciencewatch.com⁵ регулярно брал интервью у авторов высокоцитируемых статей, которые вошли в зарождающийся либо быстро растущий фронт.

³ <http://archive.sciencewatch.com/dr/fbp/>

⁴ <http://archive.sciencewatch.com/about/met/core-hp/>

⁵ <http://www.sciencewatch.com/>

Field	2014-1	2014-2	2014-3	2014-4	2014-5	2014-6	2015-1
Agricultural Sciences	4	5	4	4	3	3	4
Biology & Biochemistry	9	8	8	7	5	4	4
Chemistry	12	13	12	10	7	4	3
Clinical Medicine	9	9	8	7	5	4	4
Computer Science	5	5	7	4	3	3	4
Economics & Business	4	4	3	3	3	2	4
Engineering	6	6	5	5	4	4	3
Environment/Ecology	11	13	7	12	7	4	7
Geosciences	6	7	4	4	3	3	5
Immunology	10	9	10	8	5	4	3
Materials Science	11	11	14	9	7	5	3
Mathematics	3	3	3	3	3	3	2
Microbiology	8	8	5	5	4	7	3
Molecular Biology & Genetics	12	14	10	8	7	5	4
Multidisciplinary	24	14	13	23	17	4	2
Neuroscience & Behavior	9	8	7	6	5	4	3
Pharmacology & Toxicology	6	6	5	5	4	4	3
Physics	10	10	8	7	5	4	3
Plant & Animal Science	4	5	5	4	3	3	5
Psychiatry/Psychology	6	6	5	4	3	3	6
Social Sciences, general	4	4	4	4	3	3	6
Space Science	10	11	7	10	8	16	4

Рис. 29. Пороговые значения для «горячих» статей

В методологии, опубликованной⁶ Е. Гарфилдом в 1994 году, каждому фронту присваивался уникальный идентификатор (серийный номер) вида 92-3056 (UPTAKE OF SURFACTANT PROTEIN-B; CASEIN KINASE-II; CATALYTIC SUBUNITS)], где 92 – год (1992), данные за который использовались при идентификации фронта, 3056 – случайное четырехзначное число, а в круглых скобках – список ключевых слов. Кроме того, рассчитывался вес публикации – число цитируемых ею «ядерных» статей фронта. На рис. 30 общее число цитирований «ядерных» статей фронта RF=92-4812 равно 8144, одновременно 5 «ядерных» статей фронта были процитированы в 4 публикациях, 4 «ядерные» статьи – в 55 публикациях и т.д.

E3	8144	RF=92-4812	(PUTATIVE ANAEROBICOPROPORPHYR...
E4	6200	RF=92-4812 001	(PUTATIVE ANAEROBIC COPROPORPHYR...
E5	1519	RF=92-4812 002	(PUTATIVE ANAEROBIC COPROPORPHYR...
E6	366	RF=92-4812 003	(PUTATIVE ANAEROBIC COPROPORPHYR...
E7	55	RF=92-4812 004	(PUTATIVE ANAEROBIC COPROPORPHYR...
E8	4	RF=92-4812 005	(PUTATIVE ANAEROBIC COPROPORPHYR...

Рис. 30. Пример хранения статистик для фронтов исследований

⁶ <http://wokinfo.com/essays/research-fronts/>

Сегодня фронтам в каждом выпуске базы данных Essential Science Indicators также присваивается четырехзначный номер.

Таким образом, выгрузки следует производить с периодичностью 1 раз в два месяца по мере выхода новых выпусков базы данных Essential Science Indicators, присваивая фронтам идентификаторы вида 2005-2010_1-XXXX, где XXXX – номер фронта в выпуске 2005-2010_1.

Для анализа эволюции фронтов во времени воспользуемся методами темпорального (temporal) и продольного (longitudinal) анализа [Cobo, 2012], а затем представим результаты с помощью карты эволюции (evolution map) [Cobo et al., 2011a] и графа перекрытия (overlapping items graph) [Price & Gurse, 1975; Small, 1977]. Заметим, что такой подход применим как к отдельным фронтам, так и к кластерам фронтов.

На рис. 31 представлены все возможные варианты эволюции фронтов:

1. Фронт A^1 эволюционировал во фронт A^2 в результате внутреннего роста – среди таких фронтов можно выделить быстро растущие (fast moving).
2. Фронт B^1 затухающим и деградировал во фронт B^2 .
3. Фронт C^1 в результате эволюции распался на фронты C^2 и D^2 .
4. Фронты D^1 и E^1 объединились во фронт E^2 .
5. Фронт F^1 прекратил существование.
6. Фронт F^2 – новый, зарождающийся фронт (emerging front).

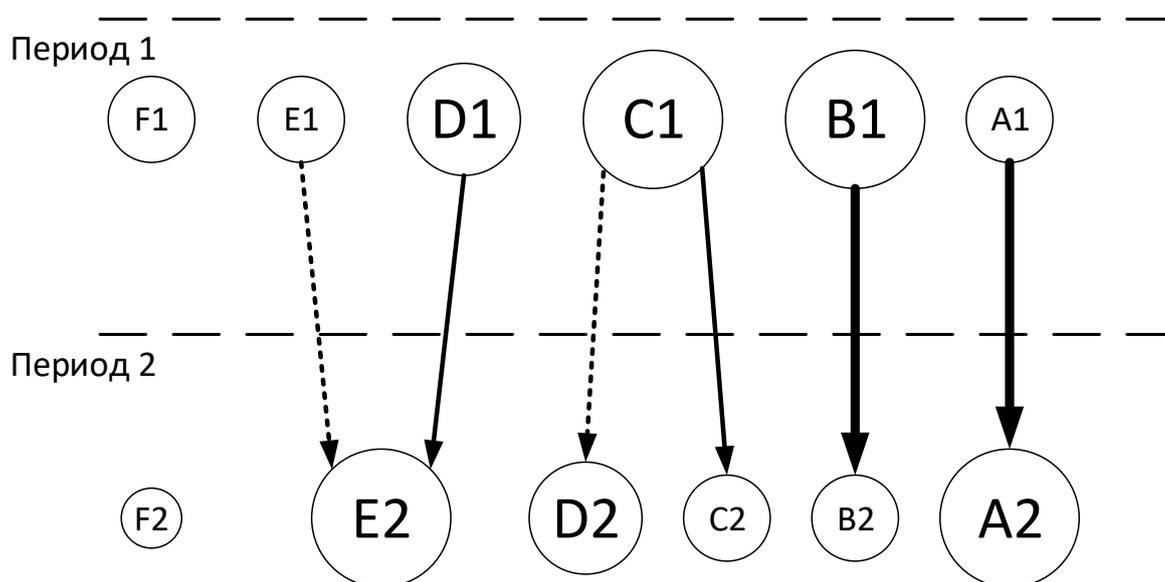


Рис. 31. Варианты эволюции исследовательских фронтов

Сплошная линия показывает, что связанные фронты имеют в основе существенную общую часть. Пунктирная линия означает, что общие элементы несущественны. Толщина линий пропорциональна индексу включения

(Inclusion Index) [Peters & van Raan, 1993]. Размер вершины пропорционален числу документов, ассоциированных с фронтами (это могут быть только «ядерные» публикации либо «ядерные» публикации и цитирующие их работы).

Граф перекрытия на рис. 32 отражает динамику между двумя соседними периодами. Вершины соответствуют периодам и помечены числом документов, ассоциированных с фронтом. Горизонтальная линия между вершинами помечена числом документов, которые являются общими для двух периодов. В скобках – индекс стабильности (Stability Index). Исходящая стрелка вверху помечена числом документов фронта, вошедшими в первый период, но не вошедшими во второй. Входящая стрелка помечена числом документов, которые впервые появились во втором периоде.

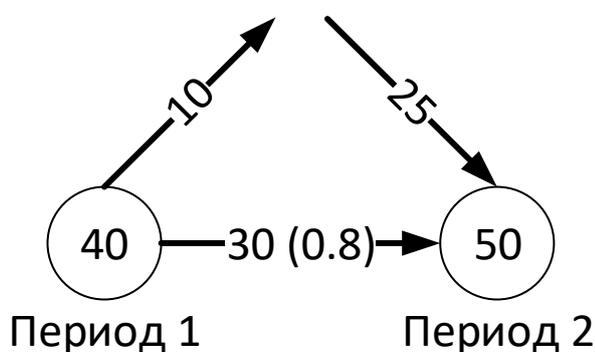


Рис. 32. Граф перекрытия

В работе [Palla, 2007] рассматривается схожая задача эволюции социальных групп на примере сетей соавторства и телефонных звонков.

На рис. 33 показаны варианты эволюции социальных групп.

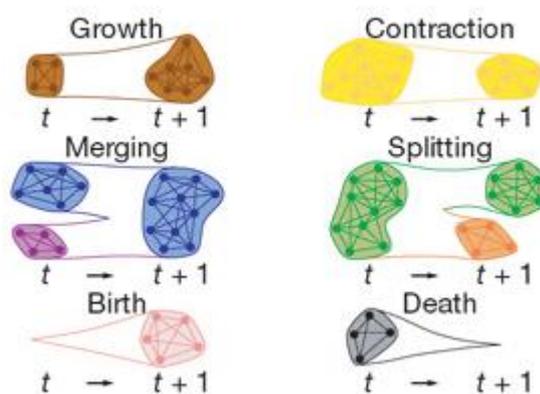


Рис. 33. Варианты эволюции социальных групп [Palla, 2007]

Авторы предлагают следующий алгоритм. Для каждой пары следующих друг за другом периодов t и $t+1$ построить объединение графов как объединение связей из двух сетей (рис. 34), и извлечь структуру сообществ методом взаимопроникновения клик (clique percolation method, CPM) [Palla, 2007]. Обнаруживаемые этим методом сообщества характеризуются тем, что они могут частично совпадать (overlap), а члены сообществ связаны между собой через сильно связанные подмножества вершин.

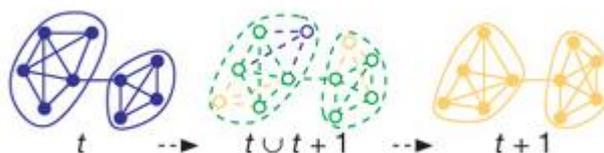


Рис. 34. Объединение графов для двух периодов [Palla, 2007]

Каждое сообщество из t или из $t+1$ будет содержаться точно в одном сообществе в новом графе, поскольку в результате добавления связей в сеть сообщества могут только расти, сливаться или оставаться неизменными. Если сообщество в новом графе содержит одно сообщество из t и одно сообщество из $t+1$, между ними существует соответствие. Если сообщество в новом графе содержит более одного сообщества из любых периодов, между сообществами устанавливается соответствие в порядке убывания степени перекрытия их вершин, вычисляемой по формуле (для сообществ A и B):

$$C(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Заметим, что данных выгрузки оказалось недостаточно для апробации рассмотренных подходов и собрать необходимые данные вручную не представляется возможным.

Анализируя отчеты Thomson Reuters за 2013 и 2014 год, можно предположить частичную взаимосвязь фронтов «Алгоритмы дифференциальной эволюции и меметические вычисления» (2013) и «Метод роя частиц и другие виды оптимизации» (2014), но отчеты не содержат список «ядерных» публикаций, который необходим для проверки нашей гипотезы.

В этой связи первоочередной задачей является организация периодической выгрузки данных для работы с историей фронтов.

Список литературы

[Baba, 2013] Baba, Y., Kashima, H. Statistical quality estimation for general crowdsourcing tasks (2013) AAAI Workshop - Technical Report, WS-13-18, pp. 8-9.

[Brumfield, 2012] Brumfield, B. Quality Control for Crowdsourced Transcription. (2012) <http://manuscripttranscription.blogspot.ru/2012/03/quality-control-for-crowdsourced.html>

[Callon, 1991] Callon, M. (1991). Techno-economic Networks and Irreversibility. In J. Law (Ed.) A Sociology of Monsters? Essays on Power, Technology and Domination, Sociological Review Monograph. London, Routledge. 38: 132-161.

[Callon, 1993] Callon, M. (1993). Variety and irreversibility in networks of technique conception and adoption. In D. Foray and C. Freeman (Eds.) Technology and the Wealth of Nations: The Dynamics of Constructed Advantage. London, Pinter Publishers: 232-268.

[Cobo, 2012] Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., Herrera, F. SciMAT: A new science mapping analysis software tool (2012) Journal of the American Society for Information Science and Technology, 63 (8), pp. 1609-1630.

[Fichman, 2012] Fichman, P. How Many Answers Are Enough? Optimal Number of Answers for Q&A Sites (2012) Social Informatics, Lecture Notes in Computer Science Volume 7710, pp 260-274.

[Flach, 2010] Peter A. Flach, Sebastian Spiegler, Bruno Golénia, Simon Price, John Guiver, Ralf Herbrich, Thore Graepel, and Mohammed J. Zaki. 2010. Novel tools to streamline the conference review process: experiences from SIGKDD'09. SIGKDD Explor. Newsl. 11, 2 (May 2010), 63-67. DOI=10.1145/1809400.1809413 <http://doi.acm.org/10.1145/1809400.1809413>

[Hung, 2013] Quoc Viet Hung, N., Tam, N.T., Tran, L.N., Aberer, K. An evaluation of aggregation techniques in crowdsourcing (2013) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8181 LNCS (PART 2), pp. 1-15.

[Kostakos, 2015] Vassilis Kostakos. 2015. The big hole in HCI research. interactions 22, 2 (February 2015), 48-51. DOI=10.1145/2729103 <http://doi.acm.org/10.1145/2729103>

[Lange, 2012] Lange, R., Lange, X. Quality control in crowdsourcing: An objective measurement approach to identifying and correcting rater effects in the social evaluation of products and services (2012) AAAI Spring Symposium - Technical Report, SS-12-06, pp. 32-37.

[Lerman, 2015] Lerman, K., Yan, X., and Wu, X.-Z. The Majority Illusion in Social Networks. <http://arxiv.org/abs/1506.03022>

[Liu, 2008] Liu, Y., Li, S., Cao, Y., Lin, C.Y., Han, D., Yu, Y. Understanding and summarizing answers in community-based question answering services (2008) Coling 2008 - 22nd International Conference on Computational Linguistics, Proceedings of the Conference, 1, pp. 497-504

[Liu, 2012] Liu, Q., Peng, J., Ihler, A. Variational inference for crowdsourcing (2012) Advances in Neural Information Processing Systems, 1, pp. 692-700.

[Luon, 2012] Luon, Y., Aperjis, C., Huberman, B.A. Rankr: A mobile system for crowdsourcing opinions (2012) Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 95 LNICST, pp. 20-31.

[Mei, 2007] Mei Q., X. Shen, C. Zhai. 2007. Automatic labeling of multinomial topic models. In Proc. of KDD'07.

[Na, 2002] Na, S. H., Kang, I. S., Lee, S. Y., and Lee, J. H. Question answering approach using a WordNet-based answer type taxonomy (2002) In Proceedings of the TREC 2002 Conference, pp. 512-519.

[Palla, 2007] Palla, G., Barabási, A.-L., Vicsek, T. Quantifying social group evolution (2007) Nature, 446 (7136), pp. 664-667.

[Pasca, 2001] Marius A. Pasca and Sandra M. Harabagiu. 2001. High performance question/answering. In Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '01). ACM, New York, NY, USA, 366-374.

[Singh, 2012] Singh, P., Lasecki, W.S., Barelli, P. and Bigham, J.P. HiveMind: A Framework for Optimizing Open-Ended Responses from the Crowd. UR CSD Technical Report, 2012.

[Snell, 2015] Snell, R. Menage a Quoi? Optimal Number of Peer Reviewers. (2015) PLoS ONE 10(4): e0120838. doi:10.1371/journal.pone.0120838

[Vuurens, 2011] Vuurens, J., et al.: How much spam can you take? an analysis of crowdsourcing results to increase accuracy. In: CIR (2011)

[Zhou, 2012] Zhou, D., Platt, J.C., Basu, S., Mao, Y. Learning from the wisdom of crowds by minimax entropy (2012) Advances in Neural Information Processing Systems, 3, pp. 2195-2203.

[Аналитический доклад, 2011] Аналитический доклад по результатам выполнения первого этапа НИР по теме «Актуализация долгосрочного прогноза важнейших направлений научно-технологического развития на период до 2030 года» // Национальный исследовательский институт «Высшая школа экономики», 2011 год.

[Бредихин, 2013] Бредихин С. В., Кузнецов А. Ю., Щербакова Н. Г. Анализ цитирования в библиометрии. – Новосибирск: ИВМиМГ СО РАН, НЭИКОН, 2013. – 344 с

[Карелин, 2012] Карелин В.П., Протасов В.И. Квалиметрический подход к формированию группы экспертов и определению их компетентностей при организации коллективного интеллекта // Вестник Таганрогского института управления и экономики. – №3. – С. 46-50.

[Косова, 2014] Косова М.В., Кушнерук С.П. Проблемы описания интегральной развивающейся терминотерминологической системы (предметная область «Управление документами») // Вестник Волгоградского государственного университета. Серия 2. Языкознание. – 2014. – №2 (21). – С. 130-135.

[Куракова, 2015] Куракова Н.Г., Григорьев О.Г., Тихомиров И.А., Девяткин Д.А. Оценка соответствия мировому уровню исследований в условиях самоизоляции российской науки: проблемы и возможные решения // Экономика науки. - №1, 2015. – С. 6-14.

[Новиков, 2013] Управление большими системами / Сборник трудов. Специальный выпуск 44 – Наукометрия и экспертиза в управлении наукой / [под ред. Д.А. Новикова, А.И. Орлова, П.Ю. Чеботарева]. – М.: ИПУ РАН, 2013. – 568 с.