

ГОСУДАРСТВЕННЫЙ КОМИТЕТ
СОВЕТА МИНИСТРОВ СССР
ПО НАУКЕ И ТЕХНИКЕ

АКАДЕМИЯ НАУК
СОЮЗА СОВЕТСКИХ
СОЦИАЛИСТИЧЕСКИХ РЕСПУБЛИК

ВСЕСОЮЗНЫЙ ИНСТИТУТ НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

НАУЧНО-ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

СЕРИЯ 2.
ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ
И СИСТЕМЫ

6

МОСКВА - 1973

СИСТЕМА СВЯЗЕЙ МЕЖДУ ДОКУМЕНТАМИ, ПОСТРОЕННАЯ НА ОСНОВЕ ССЫЛОК (по указателю «Science Citation Index»).

И. В. Маршакова

1. В практике информационной работы возникают некоторые общие задачи построения новых систем классификаций или внесения частных изменений в уже существующие. Д. Бернал в книге «Наука в истории общества» писал, что классификация сама по себе становится первым шагом на пути к пониманию новых групп явлений. Их надо упорядочить, прежде чем можно будет с ними что-то сделать. По мнению специалистов в области информатики [1—4], суть всей проблемы распределения научной информации составляют три компонента: классификация, конденсация и оценка данных.

Одним из путей изучения информационных процессов, и в частности классификации документов, может быть анализ библиографических ссылок. Наблюдение за цитируемостью научных публикаций позволяет проследить за развитием той или иной области науки во времени, за проникновением ее в смежные области. Д. Прайс отмечал, что цитирование работ образует сеть, связывающую все работы в единый комплекс. Каждая статья возникает на фундаменте других статей и сама в свою очередь становится одним из отправных моментов для следующей.

По мнению Д. Прайса [5], система научных ссылок позволяет вскрыть природу научных исследований в целом. Эту идею Д. Прайс иллюстрирует соотношением между работами и их цитированием для «замкнутой» области, состоящей из 100 работ, число которых экспоненциально растет с нормальной скоростью (7% в год).

Д. Прайс отмечает, что в любой классификации, кающихся исследовательского фронта, останется немалая часть литературы, которая не подходит полностью ни к одной из рубрик классификации, и что большая часть работ достаточно крепко увязана друг с другом через ссылки.

С точки зрения задачи классификации определенной области знания важно изучить цитируемость всех публикаций, связанных с данным направлением науки. Полученные результаты количественной оценки цитируемости должны в данном случае подвергаться дополнительному содержательному анализу.

С помощью указателя библиографических ссылок «Science Citation Index» (SCI) можно провести формальный анализ развития некоторой области науки, позволяющий следить за развитием отдельного научного направления, его классификацией и реклассификацией. Метод построения классификации с помощью указателя ссылок (SCI) базируется на том, что указывая в библиографии опубликованную ранее работу, специалист определяет концептуальную, т. е. предметную, связь между двумя статьями. Иначе говоря, он

классифицирует свою статью с помощью библиографических ссылок.

2. Указатель «Science Citation Index» Института научной информации (Филадельфия, шт. Пенсильвания, США) создан в 1963 г. Основоположником указателя считается Ю. Гарфильд [4,6].

SCI представляет собой перечень источников, содержащий библиографические ссылки, а также перечень самих библиографических ссылок. По сути SCI — это сборник указателей: Source Index — указатель источников ссылок; Citation Index — указатель ссылок; Patent Citation Index — указатель ссылок патентов; Permutation Subject Index — пермутационный предметный указатель.

Выпуск «Citation Index» (который использовался в данной работе) построен на ссылках и не имеет хронологических ограничений. Таким образом, опубликованные ранее работы регистрируются в свете современного их использования.

Указатель источников «Source Index» по существу является указателем авторов ко всем статьям, опубликованным в журналах, охватываемых SCI за определенный год, но сам указатель ссылок «Citation Index» не является указателем авторов и не должен применяться как таковой. Это указатель конкретных документов, цитируемых другими документами, образующими указатель источников, он организован по фамилиям первых авторов (для удобства потребителей.).

Надо сказать, что SCI обходит лингвистический барьер (в рамках латинского алфавита во всяком случае), поскольку он организован не по словам, а по фамилиям авторов. При введении в систему указателя документов на других языках, например на русском, японском и пр., возникает проблема транслитерации, которая представляет меньше затруднений, чем перевод. SCI указывает связи любой данной статьи со всеми публикуемыми статьями, причем не обязательно в той же области. Отсутствие размежевания между отдельными дисциплинами является по мнению Дж. Бернала, одним из главных достоинств указателя ссылок Ю. Гарфилда.

3. В практике научометрии и классификации документов известное распространение получил метод, разработанный М. Кесслером в Массачусетском технологическом институте и названный методом библиографической связи. В основе этого метода лежит принцип выделения взаимосвязи между двумя документами по числу общих цитируемых документов. Технический аппарат метода прост: 1) отдельная библиографическая ссылка, используемая в двух статьях называется одной единицей связи между ними; 2) несколько статей образуют связанную группу G_A , если каждый член группы

имеет по меньшей мере одну связь с данной исследуемой статьей P_0 ; 3) сила связи между P_0 и любым членом G_A измеряется числом единиц связи (n) между ними.

М. Кесслер, предлагая метод «библиографической связи» [7], исходил из гипотезы, что библиография технических статей дает автору возможность показать интеллектуальную среду, в которой он работает; и если две статьи содержат сходную библиографию, между ними имеется скрытое родство.

Как уже говорилось, в основе метода, разработанного М. Кесслером, лежит связь документов по числу общих цитируемых документов в библиографиях работ. По этому методу две работы прочно связаны и эта связь не меняется от новых поступлений, т. е. не зависит от реального расширения информационного массива во времени. Исходя из этого, библиографическую связь по Кесслеру можно назвать ретроспективной. Связь любой новой работы определяется также раз и навсегда со всеми предыдущими. Именно поэтому методика М. Кесслера удобна для работы с оперативными поступлениями.

Общая проблема исследования изменений классификации позволяет поставить несколько иную задачу: изучить связи документов по общим цитирующим работам. Два документа a и b будут считаться связанными, если существует достаточное количество работ, ссылающихся одновременно на документ a и документ b . Ясно, что в момент появления работ a и b ничего нельзя сказать, связаны они друг с другом или нет; картина проясняется по мере появления ссылок на эти работы. С появлением новых работ в какой-то области науки связь ретроспективных работ этой области по предлагаемому методу меняется; такая связь документов может быть названа проспективной (в отличие от библиографической связи по Кесслеру — ретроспективной). Проспективная связь документов максимально зависит от развития науки и больше интересна с точки зрения научометрии и анализа эволюции классификации.

Технический аппарат предлагаемого метода проспективной связи документов по сравнению с ретроспективной связью более сложен. Построим математическую модель классификации информационного массива, а затем полученные результаты интерпретируем на процедуру выделения групп документов этим методом.

4. Рассмотрим множество документов, представляющих информационный массив

$$\mathcal{M} \{n_1, n_2, \dots, n_m\}.$$

Выделим на этом множестве два подмножества

$$B \subset \mathcal{M} \text{ (1) и } P \subset \mathcal{M} \text{ (2)}$$

Зададим отношение $n_i f n_j$, определенное на парах $\langle n_i, n_j \rangle$, которое означает, что документ n_j цитирует документ n_i так, что

$$(Vn_i)(n_i \in B) \rightarrow (\exists n_j) n_i f n_j$$

и

$$(Vn_j)(n_j \in P) \rightarrow (\exists n_i) n_i f n_j$$

Отношение f нерефлексивно, антисимметрично и не-транзитивно.

Другими словами, подмножество $B \subset \mathcal{M}$ состоит из цитируемых документов, а подмножество $P \subset \mathcal{M}$ состоит из цитирующих документов. Очевидно, что один и тот же документ $n_i \in \mathcal{M}$ может входить в подмножества (1) и (2).

Будем называть цитирующие документы перспективными, а цитируемые документы, которые будут положены в основу классификации — базовыми документами.

Естественно, что один и тот же документ n_i множества M может выступать в роли проспективного и базового элемента подмножества $B \subset \mathcal{M}$ и $P \subset \mathcal{M}$.

Для различия в одном множестве документов базовых и проспективных обозначим $\{n_b\}$ элементы подмножества $B \subset M$, а $\{n_p\}$ элементы подмножества $P \subset M$. Сейчас можно сказать, что подмножество $B \{n_{b_1}, n_{b_2}, \dots, n_{b_l}, \dots, n_{b_b}\} \subset M$ состоит из базовых элементов, а подмножество $P \{n_{p_1}, n_{p_2}, \dots, n_{p_1}, \dots, n_{p_p}\} \subset M$ — из проспективных элементов. Отношение f запишем как $n_{b_i} f n_{p_j}$. Число элементов подмножества $B \{n_b\} \subset M$ обозначим

$$b(n) = |B \{n_b\}|, \quad (3)$$

а число элементов подмножества $P \{n_p\} \subset M$

$$p(n) = |P \{n_p\}| \quad (4)$$

Ясно, что $b(n)$ — это число базовых работ в рассматриваемом множестве документов M ; а $p(n)$ — число проспективных работ в том же множестве документов M .

Тогда $\{p(n_{b_1}), p(n_{b_2}), \dots, p(n_{b_l}), \dots, p(n_{b_b})\}$, (5)

где $p(n_{b_i})$ — число проспективных работ базового документа $n_{b_i} \in B$ можно рассматривать как признаки базовых документов, отбираемых для классификации.

Сформируем матрицу $M_1 = n_b \times n_p$

$$M_1 = \begin{bmatrix} 1 & 1 & 2 & \dots & j & \dots & p \\ 2 & m_{11} & m_{21} & m_{1j} & \dots & m_{1p} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ i & m_{i1} & m_{i2} & m_{ij} & \dots & m_{ip} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b & m_{b1} & m_{b2} & m_{bj} & \dots & m_{bp} \end{bmatrix}$$

Элемент матрицы m_{ij} может принимать одно из двух значений:

$$m_{ij} = \begin{cases} 0, & \text{означает, что отношение } f \text{ не выполняется} \\ 1, & \text{означает, что } n_{b_i} f n_{p_j} \end{cases} \quad (6)$$

Просуммировав элементы матрицы M_1 в строке $\sum_i m_{ij}$, получим соответственно для каждого базового документа $n_{b_i} \in B$ число проспективных работ $p(n_{b_i}) = |\{n_p\}_i|$.

Обозначим сумму элементов в строке матрицы

$$O(b_i) = \sum_l m_{il} = p(n_{b_i}). \quad (7)$$

Значение $O(b_i)$ будем называть объемом базового документа $n_{b_i} \in B$.

Просуммировав элементы матрицы M_1 в столбце $\sum_j m_{ij}$, получим соответственно для каждого проспективного документа $n_{p_j} \in P$ число базовых работ $b(n_{p_j}) = |\{n_b\}_j|$. Обозначим сумму элементов матрицы в столбце

$$L(p_j) = \sum_i m_{ij} = b(n_{p_j}) \quad (8)$$

$L(p_j)$ назовем весом проспективного документа $n_{p_j} \in P$.

Каждому базовому документу $n_{b_i} \in B$ можно сопоставить число проспективных работ или его объем $n_{b_i} \rightarrow p(n_{b_i}) = O(b_i)$, каждому проспективному документу

n_{p_j} — число базовых работ или его вес $n_{p_j} \rightarrow b(n_{p_j}) = L(p_j)$. Введем следующие пороговые значения для объема базового документа и веса проспективного. Пусть

$$O(b_i) = p(n_{b_i}) \geq q, \quad (9)$$

$$L(p_j) = b(n_{p_j}) \geq p. \quad (10)$$

Условие (9) означает, что для дальнейшего исследования отбираются только те базовые документы, объем которых не меньше (равен или больше) q ; а так как объем базового документа это не что иное как число проспективных работ, то условие (10) означает, что каждая проспективная работа, входящая в (9), должна иметь не меньше p базовых документов.

С учетом условий (9) и (10): множество отобранных базовых документов

$$\mathcal{B}'\{n'_{b_1}, n'_{b_2}, \dots, n'_{b_l}, \dots, n'_{b_b}\}, \quad (11)$$

число элементов множества $\mathcal{B}' \subseteq \mathcal{M}$

$$b(n') = |\mathcal{B}'\{n'_b\}|, \quad (12)$$

множество отобранных проспективных документов

$$\mathcal{P}'\{n'_{p_1}, n'_{p_2}, \dots, n'_{p_j}, \dots, n'_{p_p}\}, \quad (13)$$

число элементов множества $\mathcal{P}' \subseteq \mathcal{M}$

$$p(n') = |\mathcal{P}'\{n'_p\}|. \quad (14)$$

Каждому базовому документу $n'_{b_i} \in \mathcal{B}'$ сопоставляется его объем

$$n'_{b_i} \rightarrow O'(b_i) = p(n'_{b_i}), \quad (15)$$

каждому проспективному документу $n'_{p_j} \in \mathcal{P}'$ — вес

$$n'_{p_j} \rightarrow L'(p_j) = b(n'_{p_j}). \quad (16)$$

На множестве $\mathcal{B}' \subseteq \mathcal{M}$ определим отношение толерантности.

Определение.

Отношение толерантности $n'_{b_i} \sim n'_{b_j}$ выполняется тогда и только тогда, когда $\mathcal{P}'(n'_{b_i}) \cap \mathcal{P}'(n'_{b_j}) \neq \emptyset$. Иначе говоря, два базовых документа толерантны, если они имеют общие проспективные работы.

Сопоставим паре $\langle n'_{b_i}, n'_{b_j} \rangle$ силу связи

$$\langle n'_{b_i}, n'_{b_j} \rangle \rightarrow \xi_{ij} = \xi(n'_{b_i}, n'_{b_j}) = |\mathcal{P}'(n'_{b_i}) \cap \mathcal{P}'(n'_{b_j})| \quad (17)$$

Очевидно, что

$$\xi(n'_{b_i}, n'_{b_j}) = \xi(n'_{b_j}, n'_{b_i}); \quad (18)$$

$$\xi_{ii} = \xi(n'_{b_i}, n'_{b_i}) = \xi(n'_{b_i}) = p'(n'_{b_i}) = O'(b_i)$$

Сформируем b -мерную симметричную матрицу сил связи M_2

$$M_2 = \begin{pmatrix} \xi_{11} & \xi_{12} & \dots & \xi_{1l} & \dots & \xi_{1b} \\ \xi_{21} & \xi_{22} & \dots & \xi_{2l} & \dots & \xi_{2b} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \xi_{1l} & \xi_{2l} & \dots & \xi_{ll} & \dots & \xi_{lb} \\ \xi_{2b} & \xi_{lb} & \dots & \xi_{lb} & \dots & \xi_{bb} \end{pmatrix}$$

Сила связи ξ_{ij} корректируется следующим образом

$$\xi'_{ij} = \xi_{ij} - \delta, \quad (19)$$

где δ — порог, зависящий от величины M_a , который находится следующим образом. Для каждой пары

$\langle n'_{b_i}, n'_{b_j} \rangle$ определим величину математического ожидания M_a

$$M_a = \frac{p(n'_{b_i}) \cdot p(n'_{b_j})}{|\mathcal{P}'\{n'_p\}|} = \frac{O'(b_i) \cdot O'(b_j)}{p(n')} \quad (20)$$

Зная M_a и предполагая распределение Пуассона, определяем δ по следующему условию: $P_{>\delta} < 0,05$, где $P_{>\delta}$ — вероятность появления события больше δ раз (табл. 1а):

Таблица 1а

M_a	0,36	0,82	1,44	1,94	2,60	3,28	3,98	4,70	5,42
δ	1	2	3	4	5	6	7	8	9

Введем понятие связанный группы. Пусть $G \subseteq \mathcal{B}'$ — некоторое множество базовых (классифицированных) документов. Тогда

$$\forall n'_{b_i} \in G \sum_j \xi'(n'_{b_i}, n'_{b_j}) > \sum_k \xi'(n'_{b_i}, n'_{b_k}), \quad (21)$$

где $n'_{b_j} \in G$
 $n'_{b_k} \in \mathcal{B}' \setminus G$

Условие (21) можно принять как определение того, что множество G образует связанный группу документов. Выбор групп по условию (21) положен в основу процедуры классификации множества $\mathcal{B}' \subseteq \mathcal{M}$.

Графическая интерпретация:

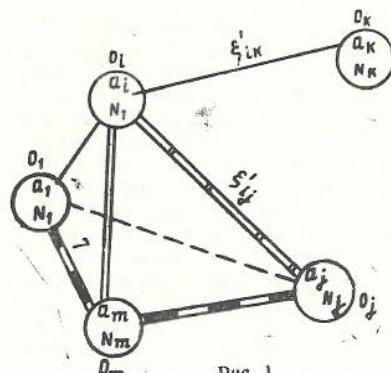
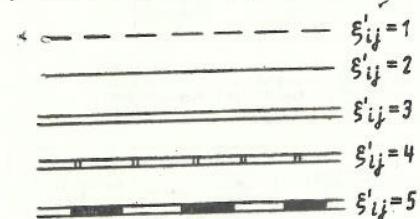


Рис. 1

a_i — последняя цифра года выхода в свет документа;

N_i — порядковый номер документа;
 O_i — объем документа. Радиус вершины графа \sim объему;

ξ'_{ij} — сила связи i и j документов.



Сила связи $\xi'_{ij} > 5$
 указывается цифрой, например,

5. Излагаемая процедура классификации строилась формальным образом, на основании ссылок по указанию «Citation Index», выявленных для информационного массива из РЖ ВИНИТИ «Физика» под обобщенной рубрикой «Генераторы стимулированного излучения» (лазеры) за 1961—1969 гг. Количественное представление массива дано в табл. 1.

Таблица 1

Год	Кол-во документов в РЖ «Физика»	Кол-во документов по годам выхода их в свет (при наличии авторов) по данным РЖ. Реальный массив	Увеличение реального массива, %
1960		7	
1961 № 7—12	19	125	18 раз
1962	188	407	325%
1963	569	549	135%
1964	803	850	155%
1965	1034	1058	124%
1966	1220	1116	105%
1967	1257	1346	119%
1968	1512	1525	113%
1969	1778	897	
1970 № 1—2	257		

Кривая роста реального массива показана на рис. 2.

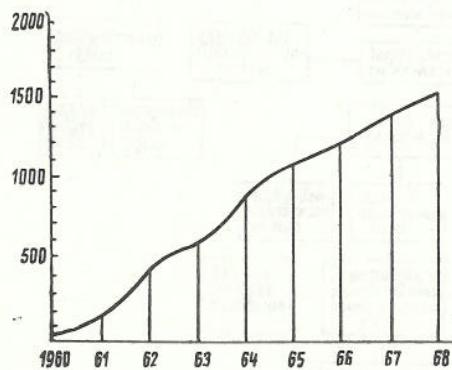


Рис. 2

Общий план процедуры классификации таков: используя условие группировок документов по силе проспективной связи, попытаться разбить массив документов на более или менее автономные группы. Затем расширить массив за счет реального пополнения документами обобщенной рубрики «Генераторы стимулированного излучения» РЖ «Физика» и применить к нему тот же алгоритм выделения автономных групп. После выделения связанных групп, попытаемся содержательно описать эти изменения в классификации, которые происходят при расширении материала данного эксперимента.

Сущность предлагаемого метода классификации сводится к следующему.

1. Два документа n_{b_i} и n_{b_j} считаются связанными, если существует достаточное количество работ, ссылающихся одновременно на документ n_{b_i} и n_{b_j} .
2. Сила проспективной связи между двумя документами n_{b_i} и n_{b_j} измеряется величиной ξ_{ij} (19).

3. Несколько документов образуют связанную группу согласно условию (21).

Материалом данного эксперимента служил информационный массив из 7.881 документа (табл. 2).

Таблица 2

Год	Суммарное количество документов информационного массива на данный год	Количество найденных ссылок по «Citation Index»
1966	4.113	1.200
1967	5.459	1.282
1968	6.984	1.343
1969	7.881	1.332

Эксперимент проводился в два этапа. На первом этапе были найдены ссылки по указателю «Citation Index» за 1966 и 1967 годы; на втором этапе — за 1968 и 1969 годы.

Результаты построенной математической модели классификации информационного массива используем для выделения автономных групп документов. Согласно модели сформируем матрицу M_1 , из которой определим объемы базовых документов и веса проспективных документов. Зададим следующие пороговые значения для объема базового документа $O(b_i)$ и веса проспективного документа $L(p_j)$.

$$O(b_i) \geq 3$$

и

$$40 > L(p_j) \geq 3$$

Очевидно, что проспективный документ, имеющий большой вес, относится к библиографиям и из рассмотрения исключается.

Для 1-го этапа число базовых документов (с учетом порога отбора) $b(n')=238$, а число проспективных документов $p(n')=250$.

Для 2-го этапа аналогично $b(n')=442$; $p(n')=600$.

Далее формируется симметричная матрица M_2 , по которой определяется сила связи базовых документов ξ_{ij} (с учетом коррекции (19), (20)).

Зная объемы базовых документов и силу проспективной связи их, строим граф связи документов; вершины графа интерпретируют номера базовых документов, ребра — силу их связи. По условию (21) выделяем автономные группы документов на графе. Для 1-го этапа построен граф 1—13 групп (рис. 3), для 2-го этапа — граф 2—26 групп (рис. 5).

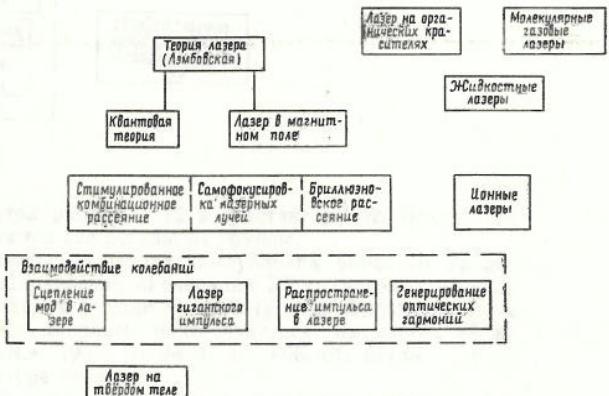


Рис. 3

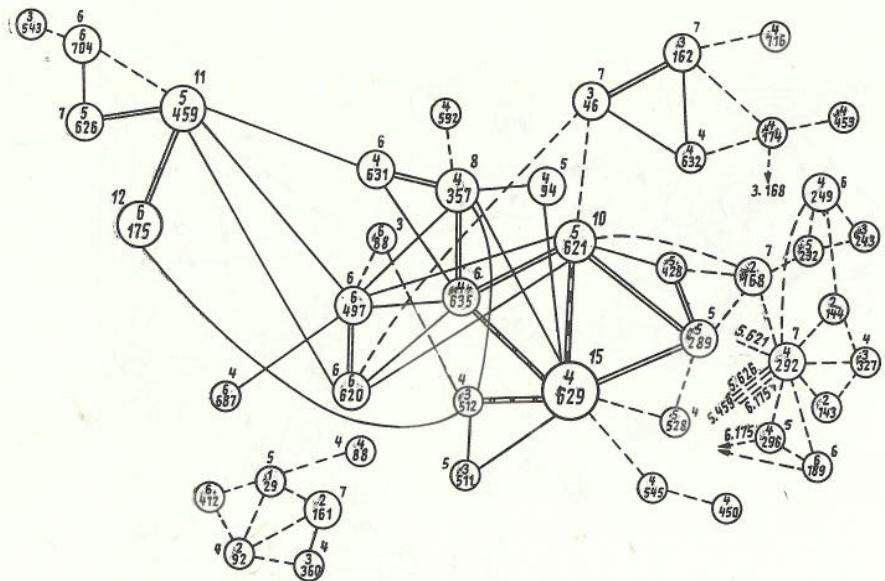


Рис. 4

6. Проследим изменения в классификации, которые происходят на втором этапе, сравнивая схемы (рис. 3 и 5) и фрагменты графа. рис. 4, 6. Рассмотрим автоном-

та 6.175 De Maria A. J., Stetser D. A., Неупан Н. Самоцепление мод в лазерах с насыщающими поглотителями «Appl. Phys. Lett.», 1966, 8, № 7, 174. РЖ 11 Д546 66.

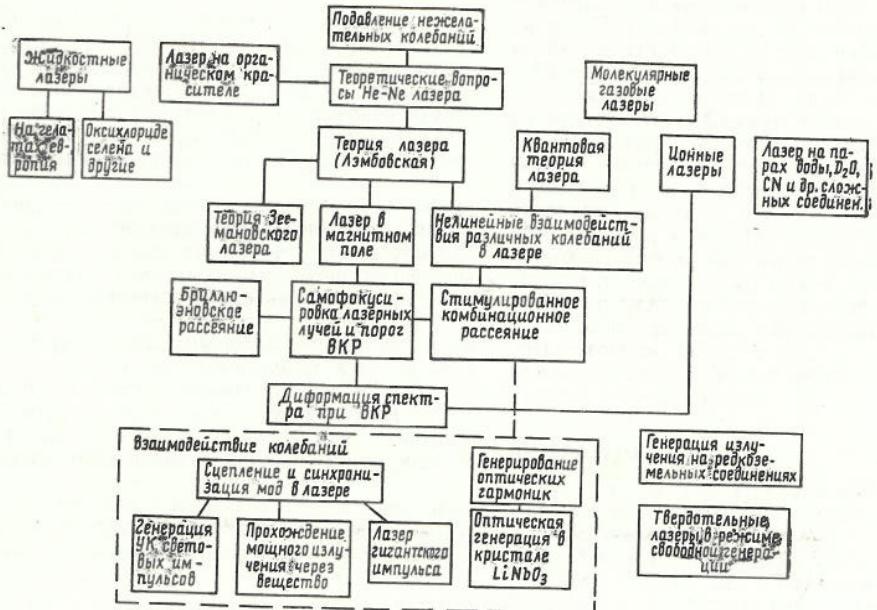


Рис. 5

ный сильносвязанный раздел, относящийся к вопросам взаимодействия колебаний. Сюда входят группы (рис. 3):
— сцепление мод в лазере;
— лазер гигантского импульса;
— распространение импульса в лазере;
— генерирование оптических гармоник.

При реальном расширении массива документов в этом разделе происходят следующие изменения (см. рис. 4 и 6). Раздел увеличивается количественно, одновременно изменяется его структура. Выделяется ядро — рабо-

тва 6.175 De Maria A. J., Stetser D. A., Неупан Н. Самоцепление мод в лазерах с насыщающими поглотителями «Appl. Phys. Lett.», 1966, 8, № 7, 174. РЖ 11 Д546 66.

Объем этой работы с 12 возрастает до 42. Вокруг нее формируются все остальные группы.
Документы последних поступлений образуют группу нового направления «Генерация УК — световых импульсов». Лидер в группе — работа 7.28 Armstrong J. A. Измерение ширины пикосекундных импульсов «Appl. Phys. Lett.», 1967, 10, № 1, 16. РЖ 8Д 877 67. Группа по структуре лучевая.

Группа «лазер гигантского импульса» усиливается в смысле силы связей, к ней добавляется несколько но-

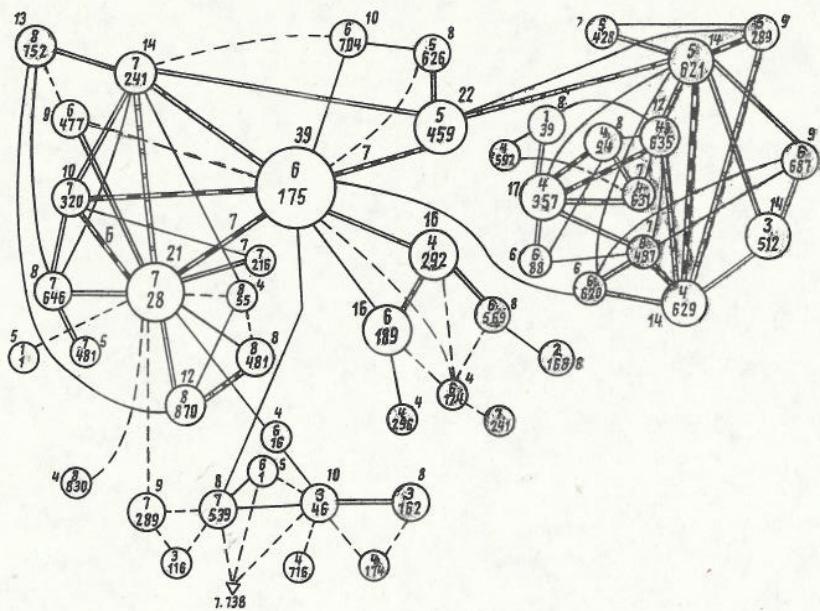


Рис. 6

вых документов той же тематики. Название группы остается.

Группа «Распространение импульса в лазере» переформировывается, приобретает лучевую структуру слабых связей с центром 3.46. Из группы выходят документы 4.632 и 4.453, входят новые: 7.539, 6.1, 6.16, 4.716 и др. Группа связана с работой 6.175 и группой «Генерация УК-световых импульсов». Название группы уточняется «Прохождение мощного когерентного излучения через вещество».

Группа «Генерирование оптических гармоник» заметно разрастается (тематика прежняя «Генерирование оптических гармоник и параметрическое преобразование света»); наблюдается очевидное выделение подгруппы близкой тематики «Оптическая генерация в кристалле LiNbO₃».

Документы этого раздела, имеющие слабые связи на первом графике и неоднозначно относящиеся к какой-либо группе, 4.292, 6.189 и др. концентрируются вокруг работы 6.175. Таким образом, центральная группа с лидером 6.175 растет и определяет следующее тематическое направление «Сцепление и синхронизация мод в лазере».

Рассмотренный пример является типичным для изучаемой реклассификации документов. Изменение классификации, как можно видеть происходит вследствие — укрепления групп, — выделения групп близкой тематики, — образования групп новых направлений.

Устойчивые классификационные группы — группы с сильными проспективными связями; документы, связанные единичными и двоичными связями могут перегруппировываться, выходя за пределы автономного раздела.

Кроме того, реклассификация позволяет проследить исторический аспект развития направлений данной области науки. Группа «Лазер на парах воды, D₂O, CN и других сложных соединениях» появляется только на схеме (рис. 5); научно-исследовательские работы в этом направлении, активно ведущиеся в 1967—1969 гг. являются продолжением работ 1964—1965 гг. Разработки, связанные с рубиновым и He—Ne лазером, начало которых относится к концу 50-х, началу 60-х годов, продолжались и в 1967—1969 гг. Лидером теоретических разработок газовых лазеров, относящихся к 1964 году, является

Лэмб (Lamb W. E. Jr.). Его работа «Теория лазера» — 4.401, объем которой с 46 (граф I рис. 4) увеличивается до 77 (граф 2 рис. 6) может считаться «классической» (по Прайсу) в реестре работ по лазерам.

Рамки журнальной статьи не позволяют проследить все изменения, которые происходят в классификации на втором этапе. Скажем только, что классификация по ссылкам распространяется на работы, имеющие теоретическое и научно-исследовательское направления; популярные обзоры, сообщения рекламного и коммерческого характера, выставки и т. п. выпадают из информационного массива, классифицируемого этим методом. Предложенный способ построения классификации документов (проспективная связь) может оказаться полезным при изучении развития определенной области науки. При такой классификации, как правило, получаются тематические группы. Одновременно в этой классификации отражаются лидеры направлений и выявляются «классические» работы области.

ЛИТЕРАТУРА

1. Bergal J. The Transmission of Scientific Information: a User's Analysis. Intern. Conf. Sci. Inform. Proc. Washington, 1959, 1 p. 77, 1958.
2. Ко блан Г. Проблема научной информации. В сб.: Наука о науке. М., «Прогресс», 1966, с. 94.
3. Прайс Д. Малая наука, большая наука. В сб.: Наука о науке. М., «Прогресс», 1966, с. 86.
4. Garfield E., Sher I. ISI'S Experience With ASCA—A Selective Dissem. System.— «J. Chem. Syst.», 1967, 7, p. 147.
5. Price D. Networks of Scientific papers.— «Science», 1965, 149 (3683), p. 510.
6. Garfield E., Stevens L. Über den Science Citation Index und verwandte Entwicklungen der jüngsten Zeit.— «Nachr. Dokum.», 1965, Bd 16, № 2, S. 130.
7. Kessler M. M. Comparison of the Results of Bibliographic Coupling and Analytic Subject Indexing.— «Amer. Docum.», 1965, 16, No. 3, p. 223.
8. Шрейдер Ю. А. Равенство, сходство, порядок. М., «Наука», 1971.

Статья поступила в редакцию 9 марта 1973 г.