

В. М. Ефременкова, Н. Н. Миловидов,
В. А. Цветкова

ОСОБЕННОСТИ ФОРМИРОВАНИЯ МАССИВА ПАТЕНТНОЙ ИНФОРМАЦИИ В БД ВИНТИ ПО ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКЕ

Проанализированы статистические характеристики потока патентных документов, отражаемых в БД ВИНТИ по вычислительной технике. Приведена динамика распределения патентных документов по странам-создателям и языкам их опубликования в массивах 1981—1992 гг., что может быть полезно при оценке тенденций развития направлений разработок средств вычислительной техники. Дано сравнение потоков патентных документов, отражаемых в БД ВИНТИ и данных службы INPADOC по отдельным направлениям изобретательской активности в этой области.

Описания изобретений к патентам, заявкам и авторским свидетельствам являются ценным источником научно-технической информации, позволяющим проводить анализ состояния и перспектив развития техники и технологии в определенной области знания. Патентная информация составляет около 25—30% общего объема научно-технической информации и увеличивается ежегодно примерно на миллион документов [1]. В России информационной базой, в которой наиболее полно отражен поток патентных документов, являются БД на магнитных лентах НПО «ПОИСК» или данные службы INPADOC. На основе массива службы INPADOC выпускались аналитико-статистические обзоры тематических направлений изобретательской активности Великобритании, СССР, США, Франции, ФРГ, Швейцарии и Японии по 21 подразделу Международной классификации изобретений (МКИ) в 1981—1989 гг. [2].

ВИНИТИ производит отбор патентной информации, поступающей в фонд Всероссийской патентной технической библиотеки (ВПТБ), в соответствии с рубрикаторм ВИНТИ [3]. В РЖ ВИНТИ реферировются полные тексты описаний изобретений из 22 стран: патенты Австрии, Австралии, Венгрии, Ирландии, Канады, Нидерландов, Норвегии, Польши, Румынии, Финляндии, Югославии; заявки Великобритании, Дании, Франции, ФРГ, Швеции, Швейцарии, Японии; авторские свидетельства СССР (России), Болгарии, Чехии и Словакии (ранее Чехословакии); с 1985 г. добавились заявки ЕПВ (Европейского патентного ведомства) и РСТ (Международного договора о патентной информации) на 17 языках: английском, болгарском, венгерском, датском, итальянском, немецком, норвежском, нидерландском, польском, румынском, русском, сербо-хорватском, финском, французском, чешском, шведском, японском. При формировании массива патентных документов следует учитывать, что в результате того, что заявки не проходят экспертизы и, следовательно, содержат опережающую информацию по сравнению с патентами, их поток в последние годы резко возрос.

Важным свойством документальной информационного потока (ДИП) является отражение характерных особенностей предметной области, выделяющих ее из ряда других областей. Предметная область может быть формально описана с помощью различных классификационных систем: рубрикаторов отрасли, индексов МКИ, НКИ и др. [3—5]. Объем ДИП, отражаемый во вторичных изданиях на бумажных или магнитных носите-

лях, формируется в соответствии с правилами отбора первичных документов. В ВИНТИ принят принцип первоначального отбора патентных документов по кодам отраслей Государственного рубрикатора НТИ [4]; после чего осуществляется экспертная оценка документа ведущими специалистами отрасли для определения целесообразности отражения его в проблемно-ориентированных БД (РЖ) [6].

Машиночитаемая БД «Вычислительная техника», генерируемая ВИНТИ с 1976 г., предназначена для абонентов магнитоленточной службы и пользователей автоматизированной информационно-поисковой системы ВИНТИ и может быть предоставлена для информационного поиска на МЛ и в режиме теледоступа. Ретрофонд БД «Вычислительная техника» типа БКР (библиография+ключевые слова+реферат) с 1976 по 1992 гг. насчитывает 195,6 тыс. документов, при этом ретрофонд патентной информации составляет 47,5 тыс. документов. В режиме теледоступа в Централизованной системе баз данных (ЦСБД НТИ) БД «Вычислительная техника» предоставлялась потребителям с 1981 по 1992 гг. с ретрофондом 170,4 тыс. документов.

В связи с бурным развитием вычислительной техники (число публикаций ежегодно возрастает на 6—8%) происходит выделение программистских и аппаратурно-технических аспектов освещения исследований и разработок, в 1987 г. было предпринято разделение БД «Вычислительная техника» на две БД: «Программное обеспечение» объемом около 8 тыс. документов, включающую материалы по теоретическим основам программирования, вопросам системного программного обеспечения, и БД «Вычислительные машины и системы» объемом около 9 тыс. документов, содержащую публикации по теоретическим основам вычислительной техники (ВТ), технологии и оборудованию для производства средств ВТ, описанию узлов и устройств ВТ. Распределение ДИП по видам первоисточников объединенной БД по вычислительной технике представлено в табл. 1. В БД (РЖ) «Программное обеспечение» в основном отражаются статьи из журналов, сборников, трудов конференций, конгрессов и т. д. и книги — около 97%, доля патентных документов составляет около 1% (табл. 2). В БД (РЖ) «Вычислительные машины и системы» — около 40—55% статей из журналов, сборников, книг и около 55—40% патентных документов (табл. 3).

Таблица 1

Количественное распределение ДИП по видам первоисточников в БД по вычислительной технике в массивах 1981—1992 гг.

| Вид первоисточника | Количество документов в массивах 1981—1992 гг. | |
|-------------------------------|--|------|
| | абс. ц. | % |
| Статьи из журналов | 86273 | 50,6 |
| Статьи из сборников | 28272 | 16,6 |
| Книги | 3802 | 2,2 |
| Патентные документы | 47494 | 27,9 |
| Депонированные рукописи | 3972 | 2,3 |
| Стандарты | 161 | 0,1 |
| Диссертации | 115 | 0,07 |
| Журнал в целом и новый журнал | 381 | 0,23 |

Суммарный массив — 170 470 документов

Таблица 2

Количественное распределение ДИП по видам первоисточников в БД «Программное обеспечение» в массивах 1987—1992 гг.

| Вид первоисточника | Количество документов в массивах 1987—1992 гг. | |
|-------------------------------|--|------|
| | абс. ц. | % |
| Статьи из журналов | 37498 | 68,4 |
| Статьи из сборников | 13436 | 24,5 |
| Книги | 1659 | 3,0 |
| Патентные документы | 420 | 0,8 |
| Депонированные рукописи | 1524 | 2,8 |
| Стандарты | 31 | 0,06 |
| Диссертации | 52 | 0,1 |
| Журнал в целом и новый журнал | 182 | 0,34 |

Суммарный массив — 54802 документа

Таблица 3

Количественное распределение ДИП по видам первоисточников в БД «Вычислительные машины и системы» в массивах 1987—1992 гг.

| Вид первоисточника | Количество документов в массивах 1987—1992 гг. | |
|-------------------------------|--|------|
| | абс. ц. | % |
| Статьи из журналов | 26001 | 45,1 |
| Статьи из сборников | 6179 | 11,1 |
| Книги | 802 | 1,4 |
| Патентные документы | 22369 | 40,3 |
| Депонированные рукописи | 895 | 1,6 |
| Стандарты | 74 | 0,12 |
| Диссертации | 21 | 0,03 |
| Журнал в целом и новый журнал | 140 | 0,25 |

Суммарный массив — 55 481 документ

Отбор патентных документов для БД (РЖ) по вычислительной технике, как уже отмечалось, осуществляется экспертами в соответствии с рубриками БД «Программное обеспечение» и «Вычислительные машины и системы» по следующим ядерным классам МКИ: 1. G 02F 3/00 — оптические логические элементы, G 02F 7/00 — оптические аналого-дискретные преобразователи; G 06D — гидравлические и пневматические цифровые вычислительные машины; G 06F — цифровые вычислительные машины, устройства для обработки цифровых данных; G 06G — аналоговые вычислительные машины; G 06I — комбинированные вычислительные машины; G 06K — распознавание, представление и воспроизведение данных, манипулирование носителями информации, носители информации; 2. G 11B — накопление информации, основанное на относительном перемещении носителя записи и преобразователя; G 11C — запоминающие устройства статического типа; 3. H 03K — импульсная техника.

Распределение потока патентных документов по странам-создателям и языкам их опубликования в массивах 1981—1992 гг. представлено в табл. 4, 5 и на рисунке. В БД (РЖ) отражаются патентные документы 23 стран и с 1987 г. патенты Европейского и Международного патентных ведомств на 16 языках. Патенты составляют 28% (13,5 тыс.) от общего массива патентных документов, заявки — 22% (10,4 тыс.) и авторские свидетельства — 50% (23,5 тыс.).

Изучение динамики распределения количества патентных документов в массивах 1981—1992 гг. выявило неравномерность их распределения по годам: при этом максимумы и минимумы для различных стран приходятся на разные годы. Максимумы для массива документов России (СССР) наблюдались в 1981, 1983 и 1989 гг.;

Таблица 4

Распределение потока патентных документов, отраженных в БД ВИНТИ по вычислительной технике, по странам-создателям

| № п/п | Страна-создатель | Количество документов в массивах 1981—1992 гг. | |
|-------|------------------|--|------|
| | | абс. ц. | % |
| 1 | СССР | 22732 | 47,9 |
| 2 | США | 11853 | 25,0 |
| 3 | Япония | 6845 | 14,4 |
| 4 | Германия | 2863 | 6,0 |
| 5 | Франция | 876 | 1,8 |
| 6 | Великобритания | 711 | 1,5 |
| 7 | Чехия и Словакия | 585 | 1,2 |
| 8 | Польша | 285 | 0,6 |
| 9 | Болгария | 236 | 0,5 |
| 10 | ЕПВ | 127 | 0,3 |
| 11 | Румыния | 77 | 0,2 |
| 12 | Венгрия | 66 | 0,14 |
| 13 | Швейцария | 52 | 0,1 |
| 14 | Швеция | 33 | 0,07 |
| 15 | РСТ | 29 | 0,06 |
| 16 | Нидерланды | 28 | 0,06 |
| 17 | Австрия | 27 | 0,06 |
| 18 | Канада | 23 | 0,04 |
| 19 | Австралия | 22 | 0,04 |
| 20 | Финляндия | 5 | 0,01 |
| 21 | Югославия | 4 | 0,01 |
| 22 | Дания | 3 | 0,01 |
| 23 | Норвегия | 2 | |

Суммарный массив — 47 484 патентных документа

Таблица 6
Распределение потока патентных документов, отраженных в БД ВИНТИ по вычислительной технике, по языкам опубликования

| № п/п | Язык | Количество документов в массивах 1981—1992 гг. | |
|-------|------------------|--|------|
| | | абс. ц. | % |
| 1 | Русский | 22732 | 47,9 |
| 2 | Английский | 12765 | 26,9 |
| 3 | Японский | 6845 | 14,4 |
| 4 | Немецкий | 2942 | 6,2 |
| 5 | Французский | 876 | 1,8 |
| 6 | Чешский | 585 | 1,2 |
| 7 | Польский | 285 | 0,6 |
| 8 | Болгарский | 236 | 0,5 |
| 9 | Румынский | 77 | 0,2 |
| 10 | Венгерский | 66 | 0,14 |
| 11 | Шведский | 33 | 0,07 |
| 12 | Нидерландский | 28 | 0,06 |
| 13 | Финский | 5 | 0,01 |
| 14 | Сербо-хорватский | 4 | 0,01 |
| 15 | Датский | 3 | 0,01 |
| 16 | Норвежский | 2 | |

США — в 1983, 1987, 1989 и 1991 гг.; Японии — 1984 и 1987 гг.; Германии — 1983, 1987 и 1991 гг.; для остальных стран на кривых распределения имели место небольшие осцилляции массивов, характерные для распределений документов Великобритании и Франции. Неравномерность отражения патентных документов по годам их отражения в БД может быть связана как со спецификой их оформления — экспертизой, так и с вопросами политики и экономики различных стран (степень секретности данных), конъюнктурой рынка и прочее.

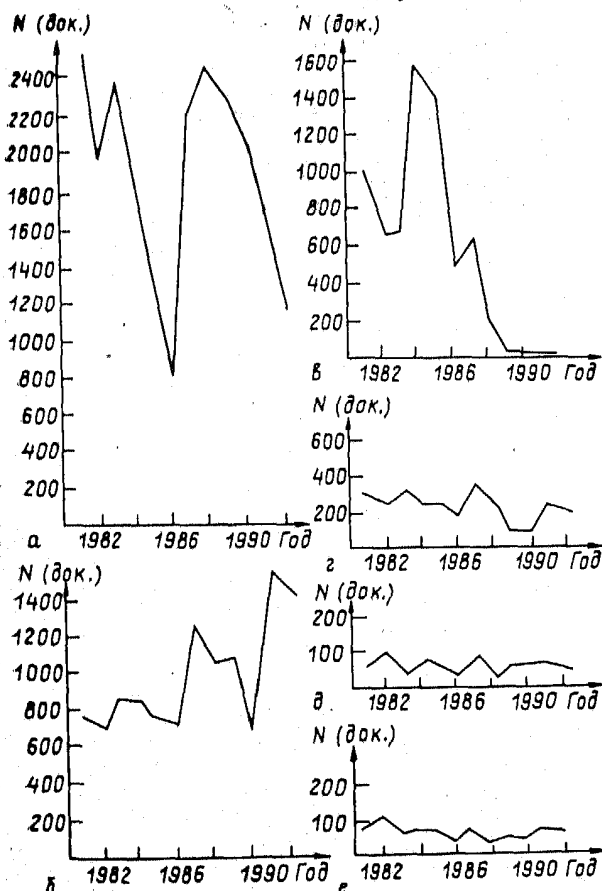
Распределение потока патентных документов по кодам МКК было получено на массивах 1990 и 1991 гг. Как видно из табл. 6, наиболее интенсивно ведутся

Таблица 6

Количественное распределение патентных документов, отраженных в БД INPADOC по вычислительной технике, по странам-создателям и кодам МКК в массивах 1987—1989 гг.

| Код МКК | СССР | США | Великобритания | Франция | Германия | Япония |
|---------|------|------|----------------|---------|----------|--------|
| G 02F | 50 | 954 | 139 | 90 | 105 | 7 650 |
| G 06F | 4872 | 5312 | 572 | 626 | 978 | 49 427 |
| G 06D | 16 | 1 | | | | |
| G 06G | 827 | 174 | 41 | 22 | 22 | 126 |
| G 06J | 49 | 19 | 19 | | 1 | 10 |
| G 06K | 374 | 1111 | 131 | 258 | 245 | 4 914 |
| G 11B | 612 | 4003 | 375 | 274 | 993 | 41 259 |
| G 11C | 917 | 1356 | 70 | 145 | 179 | 5 622 |
| H 03K | 2486 | 1851 | 187 | 172 | 426 | 6 602 |

Суммарный массив — 47 484 патентных документа



Динамика распределения потока патентных документов по годам их отражения в БД ВИНТИ по вычислительной технике: а — авторские свидетельства России (бывш. СССР); б — патенты США; в — заявки Японии; г — патентные документы Германии (ранее заявки ГДР и патенты ФРГ); д — заявки Великобритании; е — заявки Франции

разработки и внедрение в области создания устройств и отдельных узлов вычислительных машин нового поколения (G 06), накопления информации (G 11). Согласно аналитико-статистическим материалам, полученным на массивах 1981—1989 гг. [2], максимальная изобретательская активность в области вычислительной техники наблюдалась в Японии — около 70% общего потока патентной информации; на долю США приходится примерно 12%, СССР — около 7%. При этом основным направлением является G 11B (в 1989 г. вклад СССР был около 2%, США — около 8%, Японии — около 14%, остальных стран — около 12%). Анализ потоков патентных документов, содержащихся в БД ВИНТИ по вычислительной технике, показал, что в этой области ВИНТИ отражает около 15% объема БД INPADOC, при этом для класса H 03K это соотношение несколько лучше — около 23%, для класса G 11C — около 17%. Худшее соотношение — для классов G 02F, это связано с тем, что этот класс имеет более широкую тематическую направленность, чем тематика, описываемых БД.

Распределение по странам следующее: в БД отражается около 75% авторских заявок России (СССР), около 27% патентов США, 12—13% заявок Франции, Великобритании и патентов ФРГ, около 8—10% заявок Швейцарии и в последние годы около 1—0,5% заявок Японии.

Резкое сокращение числа заявок Японии связано с отказом от стратегии максимальной полноты отражения заявок в РЖ ВИНТИ. Дело в том, что заявки обычно подаются с расчетом на то, чтобы заранее получить полное покрытие всех вариантов будущего изобретения. При этом в общем потоке нередко встречаются пачки заявок (по 30—40 штук) с одинаковым названием (например, «Полупроводниковое ЗУ»), отличающиеся лишь нюансами в технологии, методах реализации и алгоритмах работы.

Кроме того, в связи с ограничением на объемы РЖ и плотность публикаций научные редакции ВИНТИ были вынуждены в 1990—1991 гг. перейти на выборочное освещение наиболее ценных материалов.

В целом содержание подготовленных в ВИНТИ рефератов на патентные документы отличается большой глубиной раскрытия таких аспектов, как актуальность, научная и технологическая значимость, новизна, суть изобретения. Разумеется, при ограничении объема реферата до 1000 знаков, это делается в ущерб патентоведческой и правовой информации о приоритетах, аналогах и прототипах, которая также весьма важна для разработчиков новой техники.

Анализ и сопоставление потоков патентов по отдельным направлениям ВТ позволяет сделать вывод о степени отставания отечественной ВТ в таких областях, как, например, периферийные устройства (накопители на магнитных дисках, принтеры, сканеры и т. п.), технология микроэлектроники. Можно отметить, что в ряде направлений оптоэлектроники изобретательская активность в России не ниже, чем в США и Японии.

ВЫВОДЫ

1. Получены распределения патентных документов в БД ВИНТИ по вычислительной технике по странам—создателям документов, языкам их опубликования и кодам МКИ.

2. Проведенное сравнение потоков патентной информации в БД ВИНТИ по вычислительной технике с БД INPADOS показало, что при экспертном принципе отбора документов в БД ВИНТИ отражается около 75% отечественных и от 27% (США) до 8% (Швейцария) зарубежных патентных документов данных ядерных классов МКИ.

3. В целом массивы патентной информации в БД по вычислительной технике отражают все основные новые технические решения, нововведения и принципиально

новые технологии. По ним можно отследить и спрогнозировать основные тенденции развития средств вычислительной техники. Однако получить исчерпывающий анализ патентной чистоты какой-либо новой заявки практически невозможно, ввиду выборочного представления всего потока информации.

4. Наибольшую ценность БД ВИНТИ по вычислительной технике представляют на этапе поиска новых технических решений и сравнительного анализа альтернатив для таких направлений, в которых отставание отечественной промышленности наиболее сильно. К этим направлениям относятся: периферийные устройства ввода—вывода информации (принтеры, сканеры, дисплеи и т. п.), накопители информации на магнитных и оптических дисках, а также технология полупроводниковых ЗУ.

СПИСОК ЛИТЕРАТУРЫ

1. Encyclopedia of information systems and services. 7th Edition / Ed. Amy Lucas Kathleen, Young Marcacaeio.— Detroit, Michigan: Gale Res. Com.— 1987.
2. Тематические направления изобретательской активности в СССР и промышленно развитых капиталистических странах за 1988 г. Сводный аналитико-статистический обзор.— М.: ВНИИПИ, 1988.
3. РЖ ВИНТИ. Автоматика и вычислительная техника. Сводный том.— М., 1992.— № 1.
4. Государственный рубрикатор научно-технической информации. Изд. 4.— М., 1992
5. Международная классификация изобретений. Четвертая редакция.— М.: ВНИИПИ, 1985.
6. Цветкова В. А., Жданова Г. С., Алексеев В. М., Ланчикова Н. Н. Входящий поток иностранных первоисточников научно-технической литературы ВИНТИ.— М., 1989.— Деп. в ВИНТИ 11.04.89, № 2300—1389.

Материал поступил в редакцию 26.05.94.