

На правах рукописи

ТОЛКАЧЕВА Елена Викторовна

**СИСТЕМА АВТОМАТИЗАЦИИ ФОРМИРОВАНИЯ
ТЕХНОЛОГИЧЕСКОЙ ДОКУМЕНТАЦИИ НА ОСНОВЕ
МОДИФИЦИРОВАННОГО МЕТОДА ПОИСКА
АССОЦИАТИВНЫХ ПРАВИЛ**

**Специальность: 05.13.12 – Системы автоматизации
проектирования (промышленность)**

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук**

Омск 2012

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Сибирская государственная автомобильно-дорожная академия (СибАДИ)»

Научный руководитель: кандидат технических наук, доцент
Семенова Ирина Ивановна

Официальные оппоненты: доктор технических наук, профессор,
Браилов Иван Григорьевич, профессор
кафедры «Прикладная механика» ФГБОУ
ВПО «Сибирская государственная
автомобильно-дорожная академия (СибАДИ)»

кандидат технических наук,
Малахов Иван Игоревич, заведующий
кафедрой «Специальных технических
дисциплин» Омского института водного
транспорта (филиал) ФБОУ ВПО
«Новосибирская государственная академия
водного транспорта»

Ведущая организация: Федеральное государственное бюджетное
учреждение высшего профессионального
образования «Омский государственный
университет путей сообщения»

Защита диссертации состоится 04 мая 2012 г. в 14⁰⁰ ч. на заседании
объединенного диссертационного совета ДМ 212.250.03 при ФГБОУ ВПО
«Сибирская государственная автомобильно-дорожная академия
(СибАДИ)» по адресу: 644080, г. Омск, пр. Мира, 5, зал заседаний.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО
«Сибирская государственная автомобильно-дорожная академия
(СибАДИ)» по адресу: 644080, г. Омск, пр. Мира, 5, зал заседаний.

Отзывы на автореферат направлять по адресу: 644080, г. Омск,
пр. Мира 5, тел., факс: (3812) 65-03-23, e-mail: Arkhipenko_m@sibadi.org

Автореферат разослан 03 апреля 2012 г.

Ученый секретарь объединенного
диссертационного совета ДМ 212.250.03,
кандидат технических наук



М.Ю. Архипенко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Анализ затрат времени предприятий, реализующих этапы жизненного цикла изделия промышленного производства как сложного технического объекта, в том числе научно-исследовательские и опытно-конструкторские работы (НИОКР), конструкторскую и технологическую подготовку производства (ТПП), показал, что более 70 % занимает производственная деятельность, из них только 5-10 % (Зыков О.) – непосредственный выпуск изделий, а остальное время и затраты связывают с подготовительной частью, в том числе, технологической. В условиях многократного использования освоенных технологических решений значительный ресурс тратится на поиск таких решений в накопленных архивах и перенос в документацию на новое изделие. Анализ процессов НИОКР и ТПП омских и др. предприятий, в частности, ОАО «Механический завод «Калачинский», выявил следующие проблемы: низкая эффективность использования накопленных слабоструктурированных архивов конструкторской и технологической документации, созданной вне приобретенных позже инструментальных средств, для разработки комплектов на новое изделие; отсутствие потокового ввода и разбора данных документов на элементарные смысловые связанные единицы из архива конструкторской и технологической документации; необходимость подбора и переноса элементов из конструкторской документации в поисковые подсистемы используемых САПР для нахождения технологических решений; отсутствие ранжирования найденных технологических данных и необходимость «вручную» выбирать и переносить результаты поиска в формы технологической документации; низкий уровень переносимости технологической документации, созданной в одной системе в другую. Эти проблемы приводят к снижению оперативности подготовки комплекта технологической документации на изготовление сложного изделия промышленного производства.

Развитие современных систем автоматизации формирования технологической документации идет по пути интеллектуализации (Бурдо Г.Б., Венцов Н.Н., Головицина М.В., Евгеньев Г.Б., Кондаков А.И., Курейчик В.М., Лебедев Б.К., Палюх Б.В., Тарасов В.Б. и др.). Анализ работ показал, что вопросам интеллектуального анализа данных с целью извлечения знаний из накопленных баз данных разработанной конструкторской и технологической документации в совокупности уделяется мало внимания. Одним из зарекомендовавших себя методов является поиск ассоциативных правил, нашедший свое применение в других областях (Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И., Чубукова И.А. и др.). Накопленный положительный опыт позволяет говорить об актуальности применения этого метода к задаче автоматизации создания технологической документации и оказывает влияние на определение цели и задач, требующих решения.

Целью диссертационной работы является повышение эффективности процесса формирования технологической документации на изготовление сложных изделий промышленного производства на основе методов интеллектуального анализа данных.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Выбрать и обосновать критерий эффективности системы автоматизации формирования технологической документации на изготовление сложных изделий промышленного производства.

2. Разработать методику автоматизации формирования базы данных электронного архива технической документации (БД ЭАТД) на базе архива разработанных проектов.

3. Разработать методику извлечения инженерных знаний из базы данных электронного архива технической документации.

4. Разработать методику автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства.

5. Разработать структуру, алгоритмы и программное обеспечение системы автоматизации формирования технологической документации, а также выполнить сравнительную оценку эффективности предложенных решений.

Объектом исследования является процесс формирования технологической документации на изготовление сложных изделий промышленного производства.

Предметом исследования являются закономерности процесса формирования технологической документации на изготовление сложных изделий промышленного производства.

Методы исследований. В диссертации приведены результаты исследований, полученные с использованием современных достижений в области систем автоматизации проектирования, интеллектуального анализа данных, проектирования технологических процессов, проектирования баз данных и теории алгоритмов.

Научная новизна диссертационной работы:

- на основе методов проектирования баз данных предложена модель отображения совокупности данных конструкторской и технологической документации в БД ЭАТД, отличающаяся возможностью потокового ввода и разбора данных, содержащихся в документах, из накопленных архивов технической документации;

- предложен модифицированный и адаптированный метод поиска ассоциативных правил, направленный на получение совокупности устойчивых закономерностей между конструктивными элементами и технологическими решениями, отражающихся в базе инженерных знаний;

- на основе методов интеллектуального анализа данных предложена модель отображения совокупности конструктивных элементов для нового изделия в совокупность технологических решений, отличающаяся применением базы инженерных знаний;

- предложена структура системы автоматизации формирования технологической документации, отличающаяся организацией функций поиска технологических решений для анализируемой конструкторской документации в базе инженерных знаний, а также включением модуля актуализации базы инженерных знаний на основе данных новых комплектов конструкторской и технологической документации.

Практическая ценность и внедрение результатов

Практическая ценность заключается в разработанном программном обеспечении (ПО) системы автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства, позволяющем сократить сроки разработки комплекта документации. Предложенное ПО может быть внедрено в технологических отделах предприятий машиностроительной отрасли либо использовано для интеграции с применяемой САПР в качестве подсистемы, расширяющей функционал, а также в учебном процессе при изучении дисциплин «Основы автоматизации проектирования изделий в машиностроении», «Основы технологии машиностроения, производства и ремонта подъемно-транспортных, строительных и дорожных машин» и др. Совокупность предложенных алгоритмов может быть использована при модернизации действующих САПР с целью повышения конкурентных преимуществ.

Часть работы выполнялась в рамках аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы» по теме «Развитие теории проектирования систем управления базами моделей в системах поддержки принятия решений» (этап «Разработка методов управления базами моделей и технологии автоматизации проектирования баз моделей для различных классов систем поддержки принятия решений», 2009), № гос. регистрации 01200803500. Результаты работы апробированы и приняты к использованию в ОАО «Механический завод «Калачинский», а также внедрены в учебном процессе ФГБОУ ВПО «СибАДИ».

На защиту выносятся:

1. Методика автоматизации формирования базы данных электронного архива технической документации на базе архива разработанных проектов.
2. Методика извлечения инженерных знаний из базы данных электронного архива технической документации.
3. Методика автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства.
4. Структура системы автоматизации формирования технологической документации на изготовление сложных изделий промышленного производства.

Апробация работы. Основные научные результаты и выводы, полученные в диссертационной работе, докладывались и обсуждались на всероссийской 65-й научно-технической конференций ФГБОУ ВПО «СибАДИ» (с международным участием) (Омск, 2011); на XVII международной научно-технической конференции «Информационные системы и технологии ИСТ-2011» (Н.Новгород, 2011); на всероссийской научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР-2011» (Томск, 2011); на научно-исследовательском семинаре «Искусственный интеллект и его приложения» (Магнитогорск, 2011); на международном информационном конгрессе «МИК-2010» (Омск, 2010); на межвузовской научно-практической конференции «Информационные технологии и автоматизация управления» (Омск, 2009); на 63-й научно-технической конференции ГОУ «СибАДИ» (Омск, 2009).

Предложенные алгоритмы зарегистрированы в ОФЭРНиО.

Публикации. По теме диссертации опубликовано 12 работ, в том числе 3 статьи в журналах, рекомендованных ВАК.

Структура работы. Диссертационная работа изложена на 160 страницах машинописного текста и включает в себя введение, четыре главы основного материала, заключение, библиографический список из 116 наименований, приложения на 28 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность проводимых исследований, сформулированы цели и задачи работы, указаны методы исследования, изложены научная новизна и практическая ценность результатов.

В первой главе проведен анализ текущего состояния в области автоматизации формирования технической документации, рассмотрен состав и структура комплекта конструкторской и технологической документации на изготовление сложного изделия промышленного производства в электронном файловом архиве технической документации. С использованием подходов, изложенных в международных стандартах по программной инженерии, выявлены потребности инженеров-технологов при создании документации на изготовление сложных изделий промышленного производства, определены функции, подлежащие автоматизации, проведен анализ существующих на рынке систем автоматизации для создания технологической документации и определены проблемные зоны в уровне автоматизации. Сделан выбор и обоснование критериев эффективности, на основании которых будет проводиться оценка предложенных решений. Выполнена постановка задач исследования.

Результаты анализа позволили сделать вывод о том, что комплексное решение выявленных проблем видится в разработке системы автоматизации, в основу работы которой лягут методы интеллектуального анализа данных, которые за счет выявления и хранения закономерностей в виде правил позволяют на несколько порядков сократить время на поиск и выбор технологических решений.

Для повышения эффективности процесса создания технологической документации по уже сформированному комплекту конструкторской документации, учитывая специфику и потребности методов интеллектуального анализа данных необходимо проработать и предложить комплекс решений по этапам, представленным на рисунке 1. В системе необходимо создать базу инженерных знаний, которая бы в себе хранила накопленные устойчивые закономерности (правила), связывающие элементы конструкторской и технологической документации, а также учитывающие последовательности операций в них. Для создания такой базы знаний требуется база данных, содержащая структурированный набор взаимосвязанных данных конструкторской и технологической документации. А для создания и наполнения такой базы данных необходим инструментарий по обработке накопленных на предприятиях комплектов технической документации в виде файлового архива, которыми пользуются технологи при создании новых комплектов документации, при этом должен учитываться факт разработки документации в разных системах. Анализ данных предприятий подтвердил факт наличия таких архивов, что дало основание для реализации данной работы.

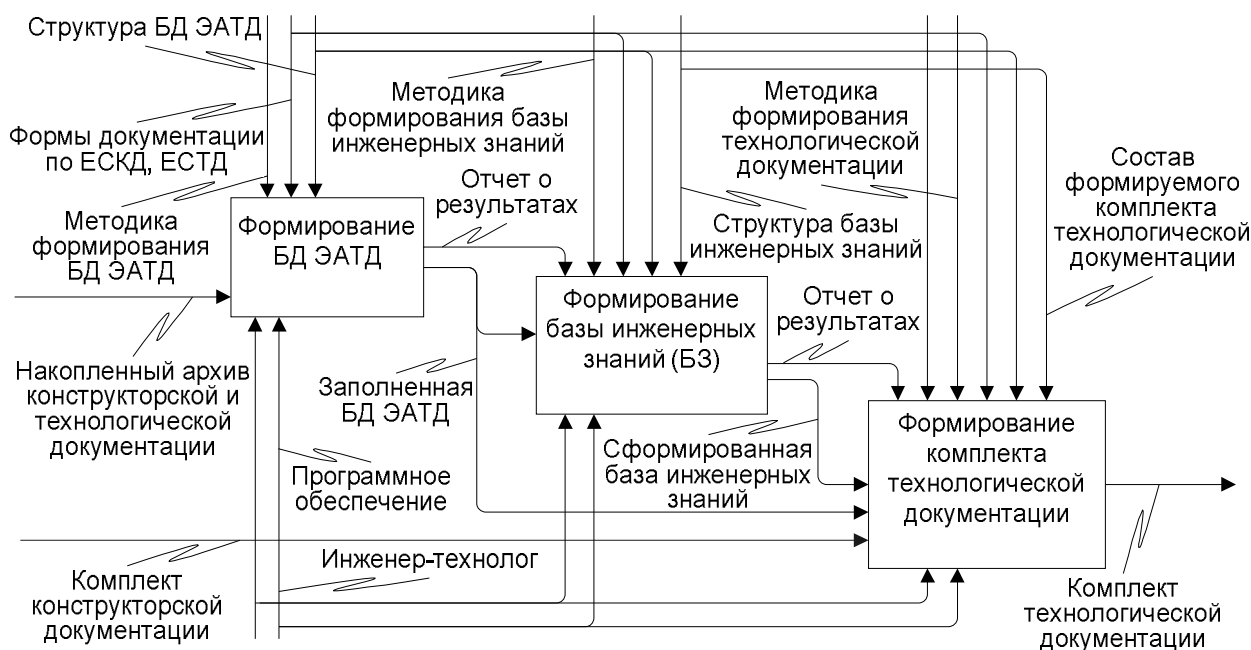


Рисунок 1 – Схема процесса автоматизации формирования технологической документации

Для оценки эффективности решений, предлагаемых в работе, был проведен выбор и обоснование критериев согласно работам Липаева В.В., Андона Ф.И., Коваля Г.И., Коротуна Т.М., Лаврищевой Е.М., Сулова В.Ю., Тейера Т., Липова Р., Нельсона Э., Соммервила И. в пользу следующего набора: временная эффективность, ресурсоемкость, гибкость, модифицируемость и полнота реализации. Так как не все метрики данного набора критериев имеют метрические шкалы, то в комплексе целесообразно сравнительную оценку эффективности предлагаемого и существующих решений выполнять по методу Саати как:

$$\max_i Q_i, Q_i = \sum_{j=1}^n \alpha_j \cdot q_{ij} \quad (1)$$

где Q_i – значимость каждой альтернативы, i – номер альтернативы, α_j – нормализованные величины/приоритеты метрик, j – номер метрики, n – количество метрик, по которым ведется сравнение, q_{ij} – нормализованные величины/приоритеты значений j -ой метрики i -ой альтернативы.

Во второй главе проведен анализ структуры и форматов электронных форм документов согласно ЕСКД и ЕСТД, определен круг подзадач, которые необходимо решить для осуществления разбора документов, извлечения и сохранения данных из документов в БД ЭАТД. Для обоснования выбора способа разбора строк выполнена оценка вычислительной сложности процедуры разбора строк в документах. Предложена методика автоматизации формирования БД ЭАТД на базе архива разработанной технической документации на изготовление сложных изделий промышленного производства, описанная в теоретико-множественном представлении. Также в главе описан предложенный алгоритм и подпрограмма заполнения БД ЭАТД на базе архива разработанных проектов с оценкой оперативности работы.

В целом задача формирования БД ЭАТД решается построением отображения $f: R_1 \rightarrow R_2$, где R_1 – неструктурированное множество данных, содержащихся в

конструкторской и технологической документации; R_2 – структурированное множество атрибутов таблиц БД ЭАТД. В существующих системах автоматизации формирования технологической документации отображение в БД ЭАТД данных из документов, созданных вне системы или в других системах, для которых не разработан инструментарий обмена данными, осуществляется путем ручного заполнения карточек документов. Для повышения степени автоматизации предложена методика автоматизации формирования БД ЭАТД на базе архива разработанных проектов, в основу которой положена спроектированная структура БД ЭАТД и совокупность правил отражения данных из документов в атрибуты таблиц БД ЭАТД.

Анализ сборочных чертежей и чертежей деталей в составе комплектов конструкторской документации показал, что чертежи содержат текстовую и графическую часть. Текстовая часть отражает технические требования, которые включают сведения об используемых материалах и др. Графическая часть отражает конструкцию детали/сборочной единицы. На выбор способа представления чертежей в БД ЭАТД повлияли следующие ограничения: каждая САД-система имеет свою библиотеку разработанных примитивов и/или элементов чертежа и компонентов приложений; при большом количестве на рынке САД-систем и их версий существует недостаток программ-анализаторов чертежей. Необходимость разработки переносимой системы привело к отказу от выделения из чертежей разных форматов элементов чертежа и компонентов приложений, в результате анализируются совокупности примитивов, извлекаемые из чертежей, представленных в формате обмена данными. Таким образом, структура признакового описания чертежа детали C_1 имеет вид:

$$C_1 = C_M \cup C_G \cup C_L,$$

где $C_M = \{m \mid m - \text{наименование материала}\}$; $C_G = \{(g,c) \mid g - \text{тип геометрического примитива, } c - \text{количество на чертеж}\}$; $C_L = \{(l,v) \mid l - \text{тип размера, } v - \text{величина}\}$.

Поскольку сборочные единицы представляют собой совокупность деталей, при анализе сборочных чертежей не имеет смысла выделять геометрические примитивы, поэтому структура признакового описания сборочного чертежа C_2 имеет вид $C_2 = C_M \cup C_L$.

Анализ текстовых документов в составе комплектов конструкторской (спецификации, ведомости покупных изделий) и технологической (ведомости оснастки, ведомости оборудования, ведомости материалов, ведомости специфицированных норм расхода материалов, карты технологических процессов, маршрутные карты, операционные карты, технико-нормировочные карты) документации показал, что содержимое каждого документа представляет собой последовательность взаимосвязанных записей. Каждая запись является набором значений атрибутов. Совокупность атрибутов записи задается в соответствии с формой документа по ЕСКД/ЕСТД, множество значений каждого атрибута определяется его типом. Анализ форм и правил заполнения документации показал, что связи между группами записей в документах имеют иерархическую древовидную структуру. Корневая вершина дерева соответствует данным о детали/сборочной единице, вершины k -го уровня соответствуют записям k -ой группы данных. Таким образом, структура признакового описания D_i содержания перечисленных типов документов имеет следующий вид:

Таким образом, модель отображения текстовой конструкторской и технологической документации в БД ЭАТД может быть представлена в виде суперпозиции отображений:

$$f_2 \circ f_1 : R_1 \rightarrow R_2, f_1 : \{(i_n, d, r, c)\} \rightarrow \{(k, j)\}, f_2 : \{(k, j)\} \rightarrow \{\tau, \alpha\}, \quad (2)$$

где n – порядковый номер документа в архиве, i_n – тип n -го документа, d – данные в r -й строке, c -м столбце таблицы n -го документа; $n = \overline{1, n_{R_1}}$, n_{R_1} – количество документов в архиве; $1 \leq i_n \leq 10$; $r = \overline{1, r_n}$, $c = \overline{1, c_n}$, r_n и c_n – количество строк и столбцов в таблице n -го документа; τ – порядковый номер таблицы БД ЭАТД, α – порядковый номер атрибута τ -й таблицы; $\tau = \overline{1, \tau_{R_2}}$, τ_{R_2} – количество таблиц БД ЭАТД; $\alpha = \overline{1, a_\tau}$, a_τ – количество атрибутов в τ -й таблице.

Некоторые атрибуты в документах неатомарные, поэтому отражение их значений в БД ЭАТД требует разбора строк для выделения значений атомарных атрибутов. К неатомарным относятся атрибуты следующих типов: наименование стандартного изделия, наименование материала, наименование оборудования, наименование технологического оснащения. Значения атрибутов первых трех типов могут содержать ссылки на нормативные документы. Значения атрибутов четвертого типа могут содержать маркировку. Структура значений атрибутов A_l перечисленных типов имеет вид:

$$A_l = \{(n_l, t_l, v_l) | n_l \in N_l, t_l \in T_l, v_l \in V_l\}, k = \overline{1, 4},$$

где n_l – наименование, N_l – справочник наименований для атрибута l -го типа; t_1, t_2, t_3 – тип нормативного документа, t_4 – тип маркировки оборудования, $T_1 = T_2 = T_3 = \{\text{ГОСТ Р, ОСТ, ТУ, ...}\}$, $T_4 = \{\text{тип, марка, модель, ...}\}$; v_1, v_2, v_3 – обозначение нормативного документа, v_4 – обозначение маркировки оборудования, V_1, V_2, V_3 – справочник нормативных документов, V_4 – справочник маркировок оборудования.

В соответствии с описанной структурой признакового описания конструкторской и технологической документации для восстановления, и отражения выявленных последовательностей взаимосвязанных данных из архива технической документации спроектирована структура БД ЭАТД.

Целью является формирование для каждого документа архива признакового описания содержимого документа и отражение его в БД ЭАТД с сохранением связей между элементами признакового описания документов.

Чертежи представлены в формате обмена данными. Текстовая конструкторская и технологическая документация представлена в текстовых форматах (doc, xls, pdf и др.). Документация прошла все виды контроля.

Ниже приведено пошаговое описание предложенной методики формирования БД ЭАТД на базе архива разработанных проектов.

Шаг 1. Для всех $n = \overline{1, n_{R_1}}$ документов F_n архива выполнить шаг 2.

Шаг 2. Если F_n в формате САД, то перейти к шагу 3, иначе к шагу 4.

Шаг 3. Если в F_n чертеж детали, то $P'(F_n)$, иначе если в F_n сборочный чертеж, то $P''(F_n)$, иначе перейти к следующему n .

Шаг 4. Если F_n в текстовом формате, то для всех $i = \overline{1, 10}$: если в F_n –

документ i -го типа, то $i_n = i$ и выполнить шаги 5-8 и перейти к следующему n .

Шаг 5. Для всех позиций $r = \overline{1, r_n}$, $c = \overline{1, c_n}$ на форме документа F_n выполнить шаги 6-8.

Шаг 6. Считать блок данных d из r -й строки, c -м столбца таблицы документа F_n .

Шаг 7. Определить номер k группы записей и номер j атрибута данных d :
 $(k, j) = f_1((i_n, d, r, c))$

Шаг 8. Если j -й атрибут k группы записей атомарный, то определить $(\tau, \alpha) = f_2((k, j))$, записать d в α -й атрибут τ -й таблицы и перейти на шаг 1, иначе выполнить шаги 9-15.

Шаг 9. Определить тип l атрибута данных d .

Шаг 10. Если в d найден $t \in T_l$, то перейти к шагу 11, иначе к шагу 15.

Шаг 11. Если для t множество $V \neq \emptyset$ ($V \subseteq V_l$), то шагу 12, иначе к шагу 15.

Шаг 12. Если в d найден $v \in V$, то шаг 13, иначе к шагу 15.

Шаг 13. Если для v множество $N \neq \emptyset$ ($N \subseteq N_l$), то шагу 14, иначе к шагу 15.

Шаг 14. Если в d найден $n \in N$, то $d_{ij}^k = (n, t, v)$ записать в БД ЭАТД и перейти на шаг 6, иначе перейти к шагу 15.

Шаг 15. Если в d найден $n \in N_l$, то $d_{ij}^k = (n, ,)$ записать в БД ЭАТД, иначе передать d для анализа инженеру-технологу; перейти на шаг 6.

Здесь P' , P'' – функции разбора графической и текстовой части чертежей деталей и сборочных единиц и записи результатов разбора в БД ЭАТД.

Для практической реализации процесса заполнения БД ЭАТД был предложен алгоритм и создана подпрограмма, которая на входе принимает путь к каталогу, с которого начинается поиск файлов технической документации, результаты работы сохраняются в спроектированной БД ЭАТД, содержащей 9 схем и 59 таблиц. Проведен вычислительный эксперимент с объемом файлового архива 8 Мб и количеством документов в архиве 200, спецификации и технологическая документация содержит в среднем 20 записей на документ, частота процессора 2,4 ГГц, оперативная память 2 Гб. Время работы в среднем на анализ и разбор одного документа составило 35 с.

Ключевые отличия от существующих систем состоят в добавлении функций потокового автоматизированного анализа и разбора текстовой части конструкторских и технологических документов со структурированием взаимосвязей между полученными данными без привлечения инженера-технолога для поддокументного ввода данных из накопленного архива в БД ЭАТД.

В третьей главе проведен анализ современных алгоритмов поиска ассоциативных правил, сформулирована постановка задачи поиска ассоциативных правил применительно к проблеме формирования базы инженерных знаний, обоснована необходимость формирования секвенций, учитывающих порядок выполнения технологических операций. Выбран базовый метод для адаптации и модификации под существующую задачу. Разработана методика формирования базы инженерных знаний на основе БД ЭАТД, описанная в теоретико-множественном представлении. Также в главе описан алгоритм и подпрограмма формирования базы инженерных знаний.

Обобщенная задача поиска ассоциативных правил состоит в нахождении групп объектов из некоторого множества $I = \{i_1, i_2, \dots, i_n\}$, часто встречающихся в большом количестве наборов объектов $D = \{T_1, T_2, \dots, T_m\}$, наборы $T_j \subseteq I$ называются транзакциями. Пусть $F \subseteq I$ – произвольный набор объектов множества I , $D_F = \{T_j \in D \mid F \subseteq T_j\} \subseteq D$ – множество транзакций, в которые входит набор F . Отношение $\text{supp}(F) = |D_F|/|D|$ называется поддержкой набора F . Набор $F = X \cup Y$ задает правило $X \rightarrow Y$ («если X , то Y »), при этом X называется консеквентом, а Y – антецедентом. Достоверностью правила называется отношение $\text{conf}(X \rightarrow Y) = \text{supp}(X \cup Y) / \text{supp}(X)$. Правило $X \rightarrow Y$ называется достоверным, если $\text{conf}(X \rightarrow Y)$ больше некоторого заданного значения conf_{\min} . Задача поиска ассоциативных правил сводится к отысканию всех наборов объектов $F = X \cup Y$, задающих достоверные правила.

В случае, когда принципиальное значение имеет порядок следования анализируемых объектов в транзакциях, задачу поиска ассоциативных правил обобщают путем введения понятия последовательности. Последовательностью называется упорядоченное множество объектов $S \subseteq I: S = \{i_1, \dots, i_p, \dots, i_q, \dots\}$, где $p < q$. Говорят, что последовательность S входит в транзакцию T , если $S \subseteq T$ и объекты, входящие в S , входят в T с сохранением порядка. Задача поиска ассоциативных правил в упорядоченном множестве объектов (когда наборы X , Y содержат последовательности) называется задачей секвенциального анализа (Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.).

БД ЭАТД представляет собой совокупность описаний сборочных единиц в составе сложных изделий промышленного производства. Каждое описание содержит сведения из конструкторской и технологической документации на изделие, т.е. является подмножеством множества объектов:

$$I = X_1 \cup X_2 \cup X_3 \cup Y_1 \cup Y_2 \cup Y_3,$$

где X_1 – сборочные единицы, X_2 – детали, X_3 – стандартные изделия; Y_1 – материалы, Y_2 – технологические операции, Y_3 – оснастка и оборудование.

Целью является поиск закономерностей в составе БД ЭАТД и структурирование их в составе базы инженерных знаний.

На выбор способа решения повлияли следующие ограничения. Поскольку составление технологической документации осуществляется на базе разработанной конструкторской документации, то искомые закономерности можно рассматривать как ассоциативные правила $X \rightarrow Y$, которые имеют следующую структуру: для деталей $X \subseteq X_3$, для сборочных единиц, не имеющих в составе других сборочных единиц $X \subseteq X_2 \cup X_3$, для остальных сборочных единиц $X \subseteq X_1 \cup X_2 \cup X_3$; $Y \subseteq Y_1 \cup Y_2 \cup Y_3$.

Порядок выполнения технологических операций (переходов) накладывает ограничения на метод поиска ассоциативных правил, а именно, требует модификации процедуры вычисления поддержки и достоверности.

Кроме того, поскольку информация в БД ЭАТД частично структурирована (технологические операции связаны с деталями, оборудование и материалы связаны с технологическими операциями и др.), то при генерации кандидатов в правила необходимо учитывать взаимосвязи, сохраненные в БД ЭАТД. С

учетом перехода на новые технологии производства технологические операции, выполненные для изготовления деталей в более позднее время, должны иметь приоритет по сравнению с технологическими операциями, выполненными для изготовления аналогичных деталей ранее.

Также особенностью предметной области является необходимость учета тривиальных правил (кандидатов с высокой достоверностью), которые соответствуют типовым конструкторским и технологическим решениям, которые могут быть рекомендованы к применению в новых проектах. Правила с низкой достоверностью соответствуют оригинальным решениям, повторное использование которых в новых проектах может быть нецелесообразным. Таким образом, должно быть установлено минимальное значение достоверности правил conf_{\min} . На практике такое значение определяется методом экспертных оценок. Для сокращения доли участия эксперта предложена методика адаптивного подбора минимального значения достоверности правил, которая позволит в большинстве случаев избежать привлечения экспертов, а в остальных случаях позволит избавить экспертов от необходимости проведения дополнительных вычислительных экспериментов. В основу методики положена проверка статистической гипотезы о нормальном распределении $\text{conf}(X \rightarrow Y)$ и правило трех сигм: если гипотеза о нормальном распределении принимается, conf_{\min} устанавливается согласно правилу трех сигм, иначе требуется привлечение эксперта.

Формирование базы инженерных знаний осуществляется путем применения межмодельного отображения $g: S_1 \rightarrow S_2$ из структуры БД ЭАТД S_1 в структуру базы инженерных знаний S_2 . Ниже приведено пошаговое описание предложенной методики формирования базы инженерных знаний на основе БД ЭАТД. Используемые обозначения: $A = X_1 \cup X_2 \cup X_3 = \{a_1, a_2, a_3, \dots, a_p\}$, $B = Y_1 \cup Y_2 \cup Y_3 = \{b_1, b_2, b_3, \dots, b_q\}$.

Шаг 1. Для всех наборов элементов $X \in \{\{a_1\}, \{a_2\}, \dots, \{a_p\}, \{a_1, a_2\}, \{a_1, a_3\}, \dots, \{a_1, a_2, a_3, \dots, a_p\}\}$ множества A выполнить шаги 2-4.

Шаг 2. Определить множество связанных с элементами набора X элементов $B' = \{b'_1, b'_2, b'_3, \dots, b'_q\} \subseteq B$.

Шаг 3. Для всех наборов элементов $Y \in \{\{b'_1\}, \{b'_2\}, \dots, \{b'_q\}, \{b'_1, b'_2\}, \{b'_2, b'_3\}, \dots, \{b'_1, b'_2, b'_3, \dots, b'_q\}\}$ множества B' выполнить шаги 4-5.

Шаг 4. Определить $\text{conf}(X \rightarrow Y)$ с учетом порядка технологических операций в маршруте и добавить в базу инженерных знаний.

Шаг 5. Если для ранее найденных правил $X' \rightarrow Y'$ выполнены условия $X' \cup Y' \subseteq X \cup Y$, $\text{conf}(X' \rightarrow Y') = \text{conf}(X \rightarrow Y)$, то правила $X' \rightarrow Y'$ удалить из базы инженерных знаний.

Шаг 6. *Актуализация правил.* Для всех правил с идентичными консеквентами $X \rightarrow Y_1, X \rightarrow Y_2, \dots, X \rightarrow Y_n$: $\text{conf}(X \rightarrow Y_i) = \max_{1 \leq j \leq n} \{k_{ij} \cdot \text{conf}(X \rightarrow Y_j)\}$, где

$k_{ij} = 1$, если дата создания i -го документа больше даты создания j -го, иначе $k_{ij} = 0$.

Шаг 7. Для случайной величины $\text{conf}(X \rightarrow Y)$ вычислить выборочное математическое ожидание m_B , выборочное среднеквадратичное отклонение σ_B , центральные эмпирические моменты третьего и четвертого порядка m_3, m_4 , асимметрию $a_s = m_3 / \sigma_B^3$, эксцесс $e_k = m_4 / \sigma_B^4 - 3$, критерий χ^2 Пирсона.

Шаг 8. Если $\chi^2 < \chi_{кр}$, $|a_s| < \delta$ и $|e_k| < \varepsilon$, то $conf_{min} = m_B - 3\sigma_B$, иначе передать эксперту полигон частот встречаемости значений $conf(X \rightarrow Y)$ по интервалам, (используя шкалу стенов «стандартной десятки» Р.Б. Кеттелла: влево и вправо от m_B откладываются интервалы, равные $\sigma_B/2$) и положить $conf_{min}$ равным значению, выбранному экспертом.

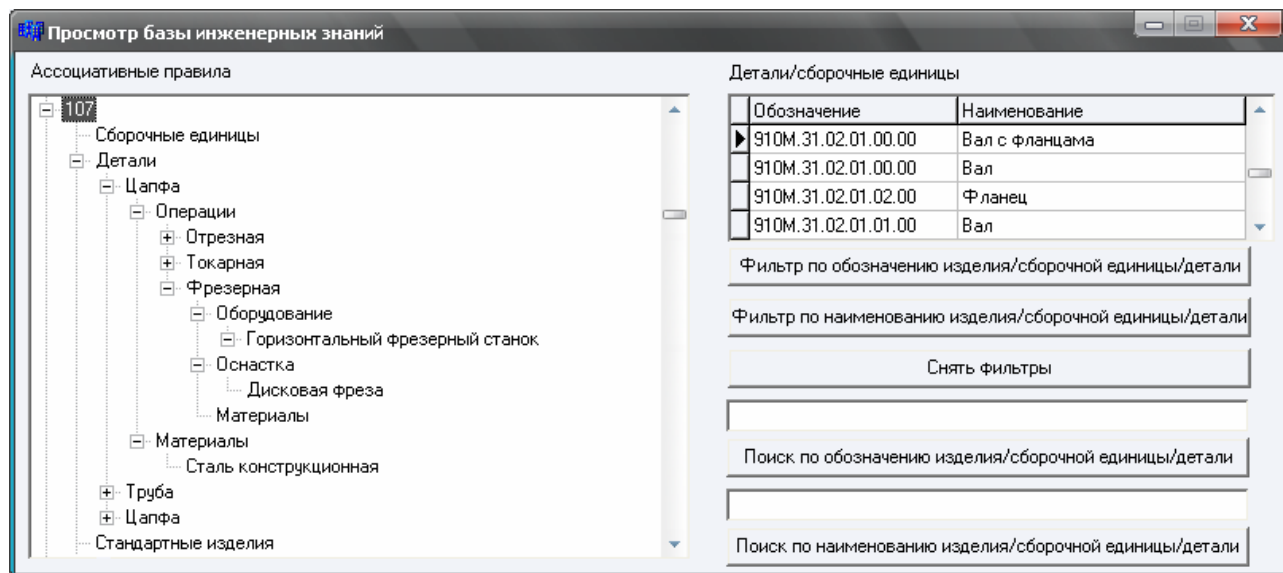


Рисунок 4 – Пример окна просмотра базы инженерных знаний

Были разработаны алгоритм и подпрограмма формирования базы инженерных знаний, для хранения которых создана дополнительно 1 схема и 12 таблиц. Вычислительный эксперимент по формированию базы инженерных знаний проводился по предварительно заполненной БД ЭАТД в результате запуска подпрограммы, описанной во второй главе, время работы которой составило $\approx 1,5$ часа. На рисунке 4 показан пример окна просмотра базы инженерных знаний.

В четвертой главе предложена методика автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства, в основу работы которой положена БД ЭАТД и база инженерных знаний, созданные ранее. Описан алгоритм автоматизации формирования технологической документации на изготовление сложных изделий промышленного производства. Также в главе описана структура и примеры работы предложенной системы автоматизации формирования технологической документации, проведена сравнительная оценка результатов работы системы.

Модель отображения БД ЭАТД и базы инженерных знаний в комплект формируемой технологической документации может быть представлена в виде суперпозиции отображений:

$$h_2 \circ h_1 : C \rightarrow D, \quad h_1 : C \rightarrow S_1 \cup S_2, \quad h_2 : S_1 \cup S_2 \rightarrow D, \quad (3)$$

где C – исходный комплект конструкторской документации, D – формируемый комплект технологической документации, S_1 – структура БД ЭАТД, S_2 – структура базы инженерных знаний.

Ниже приведено пошаговое описание предложенной методики автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства, основанной на использовании

БД ЭАТД и базы инженерных знаний, содержащей ассоциативные правила, отражающие устойчивые закономерности в составе комплектов конструкторской и технологической документации. Входными данными является комплект конструкторской документации на вновь изготавливаемое изделие.

Шаг 1. Выделение данных из конструкторской документации (процедура, ранее описанная в рамках методики автоматизации формирования БД ЭАТД на базе архива разработанных проектов) и формирование на их основе признаков описания X .

Шаг 2. Для всех деталей в составе изделия найти правила $\{X_{u_1} \rightarrow Y_{w_1} \mid u_1=1,2,\dots,U_1, w_1 = 1,2,\dots,W_1\}$, для которых antecedent $X_{u_1} \in X$, где U_1, W_1 – количество antecedentов и консеквентов правил для деталей.

Шаг 3. Для всех сборочных единиц в составе изделия, не содержащих в составе сборочных единиц найти правила $\{X_{u_2} \rightarrow Y_{w_2} \mid u_2=1,2,\dots,U_2, w_2 = 1,2,\dots,W_2\}$, для которых antecedent $X_{u_2} \in X$, где U_2, W_2 – количество antecedentов и консеквентов правил.

Шаг 4. Для остальных сборочных единиц в составе изделия найти правила $\{X_{u_3} \rightarrow Y_{w_3} \mid u_3=1,2,\dots,U_3, w_3 = 1,2,\dots,W_3\}$, для которых antecedent $X_{u_3} \in X$, где U_3, W_3 – количество antecedentов и консеквентов правил.

Шаг 5. Положить консеквент $\bar{Y} = \emptyset$.

Шаг 6. Для каждой детали/сборочной единицы выполнить шаги 7-8.

Шаг 7. Ранжировать правила по убыванию достоверности.

Шаг 8. Для каждого правила: если antecedent правила совпадает по совокупности конструктивных признаков, то добавить консеквент в описание технологии производства \bar{Y} , иначе если следующее правило имеет лучшее совпадение по совокупности признаков, то перейти к следующему правилу, иначе добавить консеквент в описание технологии производства \bar{Y} .

Шаг 9. Для всех $i = \overline{3,10}$ для блока данных $D'_i \subseteq D_i$ для документа i -го типа в составе комплекта технологической документации положить $D'_i = \emptyset$.

Шаг 10. Для всех деталей и сборочных единиц выбрать данные из \bar{Y} в D'_i , $i = \overline{3,10}$.

Шаг 11. Перенести данные из D'_i , $i = \overline{3,10}$ в соответствующие позиции форм технологической документации.

Шаг 12. Если для нормы расхода материала/нормы времени выполнения технологической операции задан идентификатор подпрограммы расчета, то выполнить шаги 13-14, иначе выполнить шаг 15-16.

Шаг 13. Задать значения параметров подпрограммы.

Шаг 14. Выполнить подпрограмму расчета нормы расхода материала/нормы времени выполнения технологической операции.

Шаг 15. Передать расчет инженеру-технологу.

Шаг 16. Записать в БД ЭАТД идентификатор выбранной инженером-технологом подпрограммы расчета норм расхода материала/норм времени выполнения технологической операции.

Шаг 17. Перенести нормы в форму документа.

По завершении инженером-технологом корректировки созданной документации выполняется дообучение базы инженерных знаний на основе данных из нового комплекта технической документации (выполняется пересчет значений поддержки и достоверности правил и корректируются значения минимальной поддержки и достоверности путем применения ранее предложенной методики).

На рисунке 5 приведена структура предложенной системы. Система включает в себя: модуль формирования БД ЭАТД, работающий на основе методики автоматизации формирования электронного архива технической документации на базе архива разработанных проектов (далее методика 1); модуль формирования базы инженерных знаний, работающий на основе одноименной методики (далее методика 2); модуль анализа комплекта конструкторской документации, работающий на основе методики 1 в части анализа и выделения признаков пространства конструкторской документации; модуль поиска ассоциативных правил в базе инженерных знаний, работающий на основе предложенной методики автоматизированного формирования технологической документации на изготовление сложного изделия промышленного производства (далее методика 3); модуль формирования технологической документации, работающий на основе методики 3; модуль пополнения БД ЭАТД, работающий на основе методики 1; модуль дообучения базы инженерных знаний, работающий на основе методик 1 и 2.

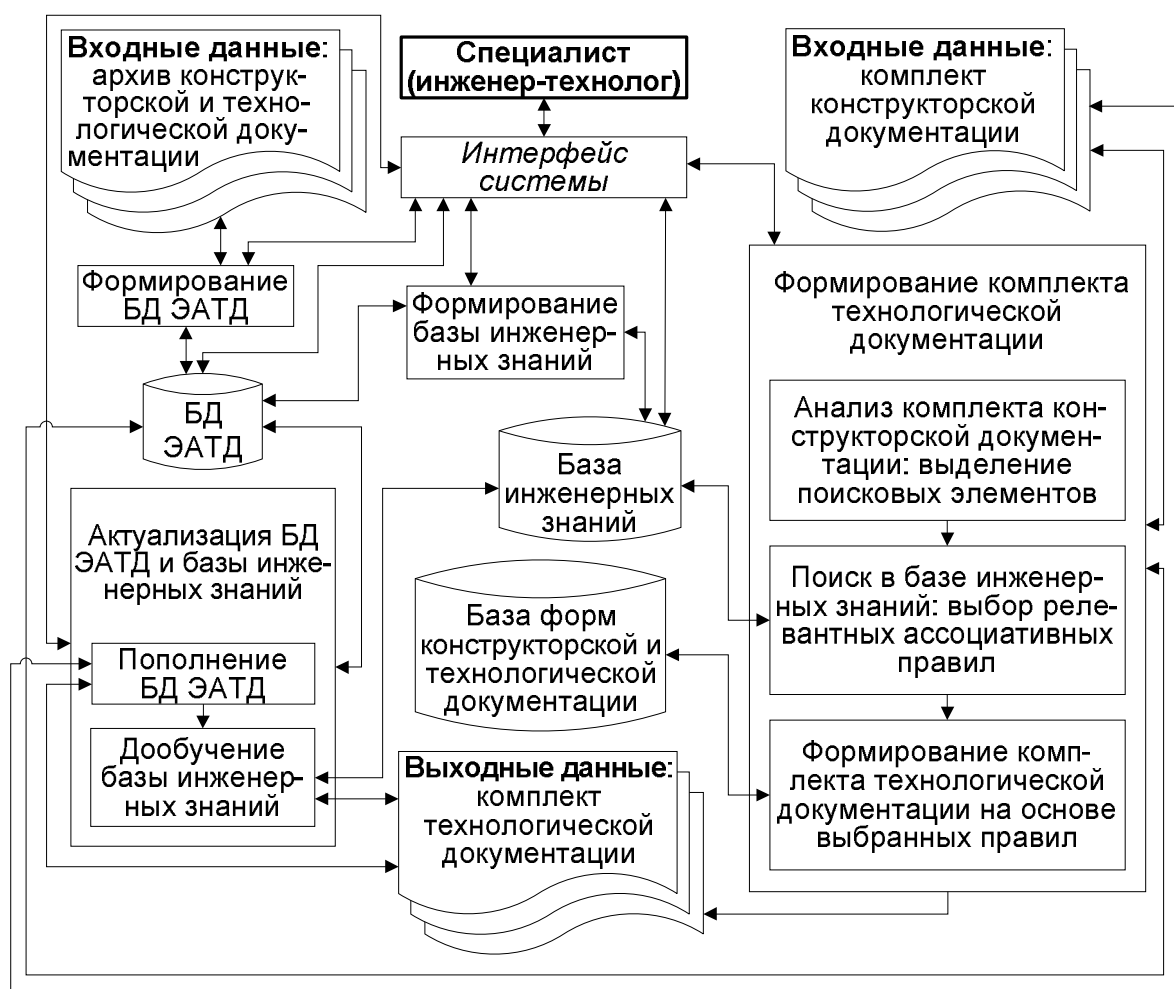


Рисунок 5 – Структурно-функциональная схема системы автоматизации формирования технологической документации

Разработана система автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства с использованием в качестве входных данных предварительно разработанного комплекта конструкторской документации. На рисунке 6 показан пример одной из форм программного обеспечения системы.

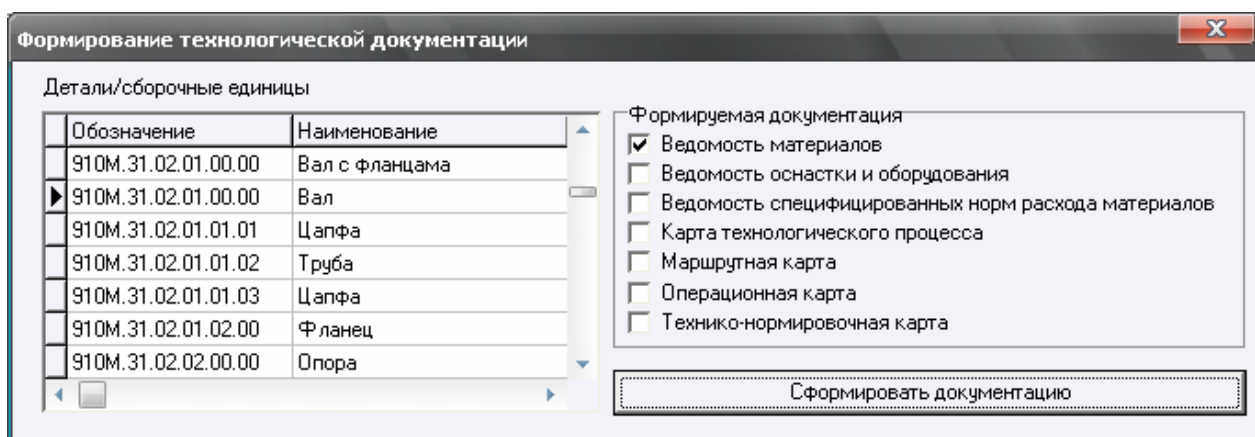


Рисунок 6 – Пример окна выбора состава технологической документации

Для системы спроектированы структуры БД ЭАТД и базы инженерных знаний, содержащие в общей сложности 10 схем и 71 таблицу. В работе приведены примеры и порядок работы с системой.

Был проведен анализ и обоснование выбора метрик для критериев в пользу следующего набора: F_1 – средняя ошибка первого рода для группы связанных атрибутов по записям, %; F_2 – средняя ошибка второго рода для группы связанных атрибутов по записям, %; F_3 – средняя ошибка первого рода для группы вычисляемых атрибутов по записям, %; F_4 – возможность применения системы в различных отраслях промышленности; F_5 – простота внесения изменений в систему; F_6 – среднее время на заполнение одного документа, с; F_7 – время формирования БД ЭАТД и базы инженерных знаний, с на документ; F_8 – среднее время анализа комплекта конструкторской документации на новое изделие, с на документ; F_9 – динамическая ресурсоемкость.

Значения количественных оценок были получены в ходе вычислительного эксперимента (при этом: количество документов в комплекте конструкторской документации 30, объем базы инженерных знаний 2,5 Мб, частота процессора 2,4 ГГц, оперативная память 2 Гб) на примере формирования ведомости материалов: $F_1 \approx 6\%$ (в эксперименте процент неопределенных наименований материалов и связанных с ними показателей); $F_2 \approx 4\%$ (процент неверно определенных наименований материалов и/или связанных с ними показателей); $F_3 \approx 10\%$ (процент неопределенных значений нормы расхода материалов, ЕН, КИМ); F_4 – средняя (так как требуется наполнение другими данными и частично реформирование БД ЭАТД и базы инженерных знаний); F_5 – средняя (в случае изменения стандартов форм документов необходимо корректировать методы их анализа и заполнения); $F_6 \approx 90$ с (в Лоцман:PLM для аналогичных документов порядка 480 с, в Excel для пустой формы ≈ 1800 с); $F_7 \approx 35$ с на документ; $F_8 \approx 30$ с на документ; F_9 – средняя.

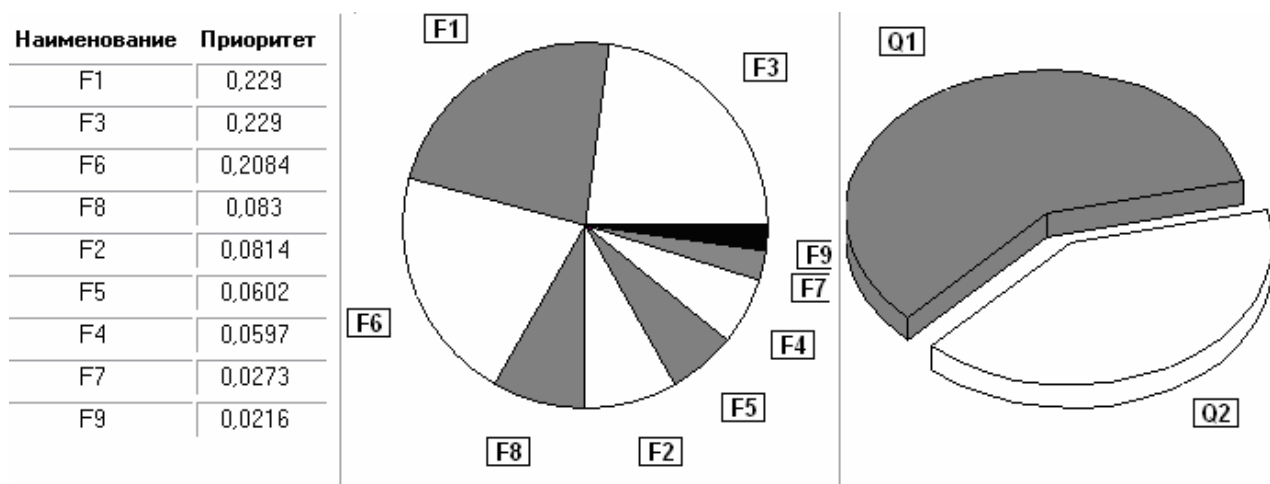


Рисунок 7 – Диаграмма нормализованных приоритетов α_i для выбранных метрик F_i и значимости альтернатив Q_1 и Q_2

В результате применения метода Саати при сравнении значимостей альтернатив Q_1 (предлагаемая в работе система) и Q_2 (Лоцман:PLM) согласно (1) получили: $Q_1 = 0,592$ при $q_{11} = 0,833$, $q_{12} = 0,833$, $q_{13} = 0,099$, $q_{14} = 0,167$, $q_{15} = 0,500$, $q_{16} = 0,833$, $q_{17} = 0,875$, $q_{18} = 0,750$, $q_{19} = 0,500$; $Q_2 = 0,408$ соответственно (рис. 7).

Значимость Q_1 выше \approx на 19 %, что дает основание считать, что цель работы достигнута, а включение предложенных решений в существующие системы, даст им конкурентные преимущества на рынке систем поддержки жизненного цикла изделий промышленного производства для соответствующей отрасли.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Для сравнительной оценки эффективности существующих систем выбран и обоснован набор метрик критериев, позволяющих дать комплексную оценку при отсутствии информации о внутренней работе алгоритмов систем-аналогов. Данный набор можно использовать при выполнении сравнительной оценки эффективности и других классов систем автоматизации, направленных на создание документации.

2. Разработанная методика автоматизации формирования БД ЭАТД на базе архива разработанных проектов позволяет реализовать потоковый ввод и разбор данных конструкторской и технологической документации, что с одной стороны существенно повышает уровень автоматизации процесса наполнения БД ЭАТД, с другой стороны сокращает до минимума долю участия в процессе инженера-технолога.

3. Разработанная методика извлечения инженерных знаний из БД ЭАТД, реализующая предложенную модификацию и адаптацию метода поиска ассоциативных правил, и полученная база инженерных знаний дают существенный выигрыш по времени при их последующем использовании для создания технологической документации в рамках новых проектов. Для оценки достоверности ассоциативных правил предложенная методика позволила сократить долю участия эксперта в ее оценке, тем самым может быть расширена область применения методов поиска ассоциативных правил.

4. Разработанная методика автоматизации формирования технологической документации на изготовление сложного изделия промышленного производства

позволяет избавить инженера-технолога от выполнения рутинных операций по формированию содержания технологической документации на изготовление сложного изделия промышленного производства при описании полных и неполных аналогов деталей, сборочных единиц.

5. Разработанные структура, комплекс алгоритмов и программное обеспечение системы автоматизации формирования технологической документации, а также проведенные вычислительные эксперименты и сравнительная оценка предложенных решений с существующими методом Саати доказала преимущества использования методов интеллектуального анализа данных при решении задач автоматизации процессов формирования технологической документации. Включение предложенных решений в существующие системы даст им конкурентные преимущества на рынке систем поддержки жизненного цикла изделий промышленного производства для соответствующей отрасли.

ОСНОВНЫЕ ПОЛОЖЕНИЯ ДИССЕРТАЦИИ ОПУБЛИКОВАНЫ В РАБОТАХ

В изданиях, рекомендованных ВАК:

1. Толкачева Е.В., Семенова И.И. Автоматизация синтеза технологических решений и их документирования на основе извлечения инженерных знаний// Вестник Воронежского государственного технического университета. 2011. Т.7. № 4. С. 76-80.

2. Толкачева Е.В., Семенова И.И. Автоматизация формирования электронного архива технической документации на базе файлового архива разработанных проектов// Вестник Воронежского государственного технического университета. 2011. Т.7. № 3. С. 114-116.

3. Толкачева Е.В., Семенова И.И. Подход к автоматизации формирования проектно-сметной документации с использованием методов интеллектуального анализа данных // Системы управления и информационные технологии. 2010. № 1.1 (39). С. 191–194.

В других изданиях:

4. Толкачева Е.В. Разработка структуры системы автоматизации документирования для оперативного создания технологической документации // Материалы Всероссийской научно-технической конференции (с международным участием). Омск: СибАДИ, 2011. Кн. 1. С. 291-295.

5. Толкачева Е.В. Повышение эффективности функционирования автоматизированных систем создания технологической документации на изготовление сложных технических объектов // Информационные системы и технологии ИСТ-2011: Материалы XVII международной научно-технической конференции, [Электронный ресурс]: Н.Новгород, 2011. 1 эл. опт. диск (CD-ROM). С. 244-245.

6. Толкачева Е.В. Поиск ассоциативных правил при автоматизации проектирования и документирования технологических процессов // Научная сессия ТУСУР-2011: Материалы всероссийской научно-технической конференции студентов, аспирантов и молодых ученых, Томск: В-Спектр, 2011. Ч. 4. С. 51-53.

7. Толкачева Е.В. Задача оптимизации выбора производственно-технологического варианта изготовления сложного технического объекта //

Искусственный интеллект и его приложения: сборник материалов научно-исследовательского семинара, Магнитогорск: МаГУ, 2011. С. 39-41.

8. Толкачева Е.В. О подходе к поддержке решений при автоматизированном составлении проектной документации // Материалы международного информационного конгресса «МИК-2010», Правительство Омской области, 2011. С. 591-595.

9. Толкачева Е.В. Возможные пути повышения эффективности информационно-поисковых подсистем САПР // Современные проблемы информатизации в анализе и синтезе технологических и программно-телекоммуникационных систем: Сб. трудов. Вып. 15 Воронеж: "Научная книга", 2010. С. 396-398.

10. Толкачева Е.В. Перспективы использования методов data mining в задаче автоматизированного формирования проектно-сметной документации // Межвузовский сборник трудов молодых ученых, аспирантов и студентов; Вып. 7. Омск: Изд-во СибАДИ, 2010. С. 121-125.

11. Толкачева Е.В. Задача автоматической классификации проектной документации в системах автоматизированного проектирования // Информационные технологии и автоматизация управления: материалы межвуз. науч.-практ. конф. ОмГТУ, Омск: Изд-во ОмГТУ, 2009. С. 124-125.

12. Толкачева Е.В. Методы повышения оперативности разработки проектно-сметной документации // Материалы 63-й научно-технической конференции ГОУ «СибАДИ». Омск: СибАДИ, 2009. Кн. 3. С. 180-183.

Свидетельства о регистрации электронного ресурса

1. Алгоритм автоматизации формирования электронного архива технической документации на базе файлового архива разработанных проектов. Инв. № ОФЭРНиО:16875/ Толкачева Е.В., Семенова И.И.; опубл. 23.03.2011.

2. Алгоритм извлечения инженерных знаний из базы данных электронного архива технической документации. Инв. № ОФЭРНиО:17552/ Толкачева Е.В., Семенова И.И.; опубл. 07.11.2011.

3. Алгоритм автоматизированного формирования технологической документации на изготовление нового сложного технического объекта. Инв. № ОФЭРНиО:17758/ Толкачева Е.В.; опубл. 29.12.2011.