

Министерство образования и науки Российской Федерации
Федеральное государственное образовательное учреждение
высшего профессионального образования
«Тамбовский государственный технический университет»

На правах рукописи

ЛЫОНГ ХАК ДИНЬ

АНАЛИТИЧЕСКИЕ И ПРОЦЕДУРНЫЕ МОДЕЛИ ДЛЯ
ИНФОРМАЦИОННОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ГРАФИЧЕСКИХ
ОБЪЕКТОВ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Специальность 05.25.05
Информационные системы и процессы

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
В.В. Алексеев

ТАМБОВ-2014

СОДЕРЖАНИЕ

	Стр.
Введение.....	4
1 Современное состояние вопроса распознавания графических объектов в информационных системах.....	8
1.1 Анализ подходов к решению задачи распознавания графических объектов в информационных системах.....	8
1.1.1 Анализ особенностей систем распознавания с «обучением» и без «обучения».....	8
1.1.2 Анализ методов распознавания графических объектов, основанных на кластеризации.....	16
1.1.3 Анализ цифровых способов обработки графических объектов.....	20
1.2 Обзор программных средств распознавания графических объектов.....	23
1.3 Методы оценки информационной потребности пользователя	25
1.3.1 Традиционные методы оценки информационной потребности пользователя.....	25
1.3.2 Методы оценки информационной потребности пользователя на основе пертинентности.....	28
1.4 Выводы по главе 1. Постановка задачи на исследование.....	33
2 Аналитические модели для информационных систем распознавания графических объектов	35
2.1 Модели информационных процессов при распознавании графических объектов.....	35
2.2 Аналитическая модель распознавания графических объектов на основе сплайнов.....	
2.3 Формализация информационной потребности пользователя с помощью коллокаций	39
2.4 Аналитическая модель процесса поиска сведений,	

	результаты которого обладают требуемой пертинентностью.	51
2.5	Выводы по главе 2.....	56
3	Процедурные модели для информационных систем распознавания графических объектов.....	58
3.1	Процедурная модель решения задачи распознавания графических объектов	58
3.2	Процедурная модель распознавания графических объектов на основе сплайнов.....	59
3.3	Моделирование процесса распределения ресурсов в информационных системах	63
3.4	Модель решения задачи построения структур баз данных.....	75
3.5	Выводы по главе 3.....	80
4	Практическое применение разработанных моделей и оценка эффективности их применения.....	81
4.1	Применение моделей в интеллектуальной системе образовательного назначения.....	81
4.2	Информационная система распознавания графических объектов для судебной психиатрии.....	85
4.3	Оценка эффективности применения разработанных моделей	
4.3.1	Оценка живучести информационной системы с использованием системы matlab.....	94
4.3.2	Исследование вычислительных характеристик применения разработанных моделей.....	99
4.4	Выводы по главе 4.....	100
	Заключение.....	102
	Список использованных источников.....	104
	Приложение А. Пример графического объекта.....	119
	Приложение В. Копии актов использования результатов исследования.....	120

ВВЕДЕНИЕ

Актуальность темы исследования.

В последнее время исследователями много внимания уделяется вопросам построения и использования информационных систем (ИС) работающих с графическими объектами. Кроме того, в ряде информационных систем, в частности образовательного, метеорологического или медицинского назначения информация, необходимая специалисту для принятия решения заключена в, иногда еле различимых, изменениях графического объекта, в сравнении с эталонным объектом. Поиск этих отличий сопровождается, как правило, наличием неопределенности, обусловленной недостаточным качеством сканирования графических объектов; шумами и фоном различной природы; недостаточной квалификацией специалистов, а в ряде случаев, целенаправленными действиями по нарушению режимов работы информационных систем. Уменьшение влияния неопределенности возможно различными способами: повышением качества сканирования объектов, помехозащищенности, повышением качества распознавания графических объектов и т.п.

Практика показала, что, зачастую, пользователь не знает, как лучше выявить отличия между исследуемым графическим объектом и, например, эталонным. Этим обусловлено невысокое качество результатов поиска, даже при использовании таких известных и зарекомендовавших себя ИС, как Гугл или Яндекс. Поэтому на смену релевантности приходит понятие пертинентность – соответствие полученной информации информационной потребности. Однако, для поиска на основе пертинентности, необходимо детально формализовать информационную потребность пользователя.

В работе представлены результаты исследования направленные на повышение эффективности функционирования информационных систем на основе созданных моделей, применение которых обеспечивает повышение качества распознавания графических объектов в условиях неопределенности.

Степень разработанности темы исследования. Проблемам повышения эффективности функционирования информационных систем на основе формализации информационной потребности пользователя, повышения качества распознавания графических объектов посвящены работы Попова Э.В., Поспелова Г.С., Поспелова Д.А., Шенка Р., и др.

Исследованию распознавания образов посвящен ряд работ, в частности, эти вопросы нашли отражение в трудах Анисимова Б.В., Айвазяна С.А., Безрука В.М., Бокса Дж., Вапника В.Н., Горелика А.А., Дженкинса Г., Дуда Р., Круглова В.В., Небабина В.Г., Патрика Э., Скрипкина В.А., Фукунаги К., Червоненкиса А.Я., и др. Эти работы внесли существенный вклад в развитие теории и практики исследуемой предметной области. Однако их содержание не в полной мере отражает особенности распознавания графических объектов, главными из которых является малый объем априорных данных относительно эталонных описаний распознаваемых образов, сложность создания оптимальных алгоритмов распознавания сигналов на основе адекватных вероятностных моделей по совокупности показателей качества. Остаются не рассмотренными вопросы улучшения пертинентности поиска сведений за счет качества распознавания графических объектов в условиях неопределенности. В связи с этим в ряде работ предлагается использовать коллокации для поиска и кластеризации информационных объектов.

В процессе исследования сформулированы выводы, свидетельствующие, в частности, о недостаточности аппарата математической статистики для построения эффективных моделей поиска и кластеризации сведений.

Таким образом, практическая задача заключается в развитии информационных систем на основе повышения пертинентности результатов поиска сведений за счёт повышения качества распознавания графических объектов в условиях неопределенности.

Объект исследования – процесс распознавания графических объектов.

Предметом исследования являются модели распознавания графических объектов в условиях неопределенности, обеспечивающие заданное качество их распознавания.

Целью исследования является повышение эффективности информационных систем за счёт применения разработанных моделей, обеспечивающих заданное качество распознавания графических объектов и оценки пертинентности его результатов в условиях неопределенности.

Научная задача исследования заключается в необходимости разработки моделей распознавания графических объектов в условиях неопределенности на основе определения информационной потребности пользователя с применением математического аппарата теории нечетких множеств.

Для достижения цели исследования и решения научной задачи необходимо решить частные задачи:

1. Провести анализ предметной области. Выполнить обоснование и выбор критериев соответствия сведений информационной потребности пользователя.
2. Разработать аналитические модели распознавания графических объектов и оценки пертинентности этого процесса.
3. Разработать процедурную модель решения задачи распознавания графических объектов.
4. Оценить эффективность, полученных в ходе исследования, моделей и живучесть системы, построенной с применением разработанных моделей.

Теоретическая и практическая значимость. Теоретическая значимость результатов исследования заключается в развитии: теории распознавания образов на решение задач распознавания графических объектов в условиях неопределенности путем построения моделей, обеспечивающих повышение качества распознавания графических объектов; теории информационного поиска на основе предложенного критерия и модели оценки пертинентности результатов распознавания графического объекта.

Практическая значимость результатов исследования заключается в программной реализации разработанных моделей и возможности определения

количественных показателей эффективности функционирования ИС, в различных предметных областях, на основе формализации информационной потребности пользователя в виде совокупности некоторых коллокаций.

Методология и методы исследования. Для решения поставленных задач и разработки соответствующих моделей применены методы теорий информационного поиска, нечетких множеств, распознавания образов, сетей Петри.

Результаты, выносимые на защиту, и их научная новизна:

- аналитическая модель распознавания графических объектов, отличающаяся построением системы уравнений для определения коэффициентов огибающих сплайны на выделенных временных интервалах, решение которой формирует набор признаков, характеризующих графический объект;
- аналитическая модель процесса распознавания графического объекта, отличающаяся введением параметров «полнота» и «точность», обобщающей временной F -нормы, включающей функции принадлежности для формирования оценки пертинентности;
- процедурная модель решения задачи распознавания графических объектов, отличающаяся определением динамических характеристик распознаваемого объекта на основе кубических сплайнов;
- результаты оценки эффективности применения разработанных моделей и оценки живучести ИС с использованием системы MATLAB.

Степень достоверности и апробация результатов исследования. *Достоверность научных результатов* обеспечивается полнотой анализа предметной области и подтверждается корректным применением математического аппарата теории информационного поиска, нечетких множеств, распознавания образов, сетей Петри, согласуемостью полученных теоретических результатов с результатами вычислительных экспериментов.

Основные результаты работы докладывались и обсуждались на: Международной научно-практической конференции «Техника и безопасность

объектов уголовно-исполнительной системы-2011» (г. Воронеж); VIII Всероссийской научно-практической конференции «Математические методы и информационно-технические средства», 2012 (Краснодар); Международной научно-практической конференции «Техника и безопасность объектов уголовно-исполнительной системы», 2013 (г. Воронеж); Всероссийской научно-практической конференции «Актуальные проблемы деятельности подразделений уголовно-исполнительной системы», 2013 (г. Воронеж); Международной научно-технической конференции «Современные информационные технологии», 2014 (г. Пенза); XIV Международной конференции «Информатика: проблемы, методология, технологии», 2014 (г. Воронеж).

Результаты исследования использованы в информационной системе, применяемой в учебном процессе для формирования образовательного контента в соответствии с профессионально-ориентированными требованиями направления подготовки и при контроле качества формируемых у студентов профессиональных компетенций в ФГБОУ ВПО «Тамбовский государственный технический университет». Аналитические и процедурные модели распознавания графических объектов использованы при разработке концепции и структуры информационной интеллектуальной системы для автоматизации процессов принятия решений о состоянии пациента в ООО «Медтехника», что подтверждено актами использования.

По результатам выполненных исследований опубликовано 32 печатных работ, в том числе 7 статей в научных изданиях, рекомендуемых ВАК Минобрнауки РФ.

В публикациях, написанных в соавторстве, лично автору принадлежат результаты: анализа предметной области [3, 4, 12, 17, 18, 30]; формулировки постановок задач [9, 10, 14, 19, 20, 22]; построения и анализа моделей, выносимых на защиту [6, 7, 24, 28, 29, 31, 32]; результаты применения [1, 2, 23, 25-27] и оценки эффективности использования моделей [5, 8, 11, 13, 15, 16, 21].

Структура и объем работы. Диссертационная работа состоит из введения,

четырех глав, заключения и списка использованных источников. Общий объем работы составляет 120 страниц основного текста, включая 31 рисунок и 22 таблицы. Список использованных источников включает 127 наименований.

Основные результаты исследования, соответствуют п. 7 паспорта специальности 05.25.05 – Информационные системы и процессы.

Работа выполнена в рамках приоритетных научных направлений стратегического развития Института автоматизации и информационных технологий, Научно-образовательного центра моделирования и управления информационными процессами и системами и информационной безопасности ФГБОУ ВПО «ТГТУ», научных школ ФГБОУ ВПО «ТГТУ» и ФГБУН «Институт радиотехники и электроники им. В.А.Котельникова» РАН, а также при поддержке госконтракта № П292 ФЦП «Научные и научно-педагогические кадры инновационной России на 2009–2013 годы».

В первой главе **«Современное состояние вопроса распознавания графических объектов в информационных системах»** проведен анализ подходов к решению задачи распознавания графических объектов в информационных системах, рассмотрены методы и модели распознавания графических объектов, основанные на различных принципах.

Во второй главе **«Аналитические модели для информационных систем распознавания графических объектов»** представлено описание разработанных аналитических моделей распознавания графических объектов на основе сплайнов, аналитической модели процесса поиска сведений, результаты которого обладают требуемой пертинентностью, построения пертинентного запроса к информационной системе на основе математического аппарата теории нечетких множеств.

В третьей главе **«Процедурные модели для информационных систем распознавания графических объектов»** представлены процедурные модели решения задачи распознавания графических объектов распознавания графических объектов на основе сплайнов, а также результаты моделирования процесса распределения ресурсов в информационных системах для решения

задач распознавания графических объектов и решения задачи определения структуры баз данных.

В четвертой главе **«Практическое применение разработанных моделей и оценка эффективности их применения»** приведены примеры их реализации в ИС различного назначения, представлены результаты оценки эффективности ИС.

ГЛАВА 1. СОВРЕМЕННОЕ СОСТОЯНИЕ ВОПРОСА РАСПОЗНАВАНИЯ ГРАФИЧЕСКИХ ОБЪЕКТОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

1.1 Анализ подходов к решению задачи распознавания графических объектов в информационных системах

В литературе, в том числе и в ГОСТ, представлено достаточно большое количество определений информационной системы [2, 5, 28, 126, 128]. При проведении исследования считается, что информационная система (ИС) это эрготехническая система, предназначенная для отображения и преобразования некоторой части реального мира в интересах удовлетворения информационных потребностей пользователя. Составляющими ИС являются: техническая часть - вычислительные и коммуникационные подсистемы; эргатическая часть – пользователь, а также программное, лингвистическое и другие виды обеспечения, информационные ресурсы.

1.1.1 Анализ особенностей систем распознавания с «обучением» и без «обучения»

Многообразие алгоритмов и способов распознавания, достаточно полно представлено в трудах [7-9, 12, 18, 27, 46, 46, 97, 101, 111, 116, 117, 124]. В них даны основные понятия модели, постановка задач распознавания, изложены базовые подходы к решению задач построения моделей для распознавания различных классов объектов, способов их описания, используемых сигналов при различных подходах и алгоритмах идентификации. В этих источниках содержатся многие аспекты задачи распознавания, которые являются актуальными при изучении проблем построения и всестороннего анализа моделей процессов или систем распознавания графических объектов.

Анализ литературы [2, 3, 5, 6, 12, 19, 20, 36, 37, 41, 47, 50, 51, 91, 94, 115, 120, 122, 125, 127] показал, что центральное место в решении поставленной задачи занимают процедуры подтверждения и выбора моделей, причем особое внимание уделяется методам инструментальных переменных, а задача распознавания, в общем случае, сводится к определению оператора модели, преобразующего входные воздействия объекта в выходные величины. Оператор объекта является его аналитической формализацией, т.е. аналитической моделью объекта, и может быть определен в соответствующих пространствах функций. Операторы могут характеризоваться разными структурой и характеристиками, и соответственно, задача распознавания объекта может иметь различные постановки.

Анализ предметной области показал, что для человека не вызывает затруднений соотнести новый объект или явление к известному ему объекту или явлению. Опытный специалист может без труда определить причину неисправности в технической сфере, геолог распознает местоположение полезных ископаемых, врач – в медицине ставит диагноз, и т.п., , эксперт-криминалист распознает сходные почерки, археолог установит исторический период найденного артефакта и т. д. За период своего существования человечество выработало два основных метода обучения: объяснение и обучение на примерах[5, 11, 42, 103, 104, 106, 133]. С развитием информационных систем различного назначения и электронно-вычислительной техники сами методы не претерпели существенных изменений, но появилась возможность автоматизации процесса распознавания явлений и объектов, что существенно повышает достоверность принимаемых на этой основе решений и существенно уменьшает время принятия решений.

С развитием ИС и средств вычислительной техники появилась возможность автоматизации процессов распознавания различных объектов, проходящих, например, в системах технической или медицинской диагностики. В подобных системах результат диагностики – графический объект, к сожалению «зашумленный» помехами. В таких системах информация о

характерных признаках распознаваемых объектов определяется на основе прямых измерений по данным технических средств. Современные методы распознавания основаны, в основном на том, что существенные признаки, позволяющие отнести графический объект к тому или иному классу (группе) рассчитываются в результате обработки реализаций динамически меняющегося графического объекта. В ряде информационных систем информация о таких признаках получается в результате обработки основе косвенных измерений. Такие системы широко распространены в метрологии, метеорологии, медицине, и других областях. В этом случае значения параметров существенных признаков не измеряются непосредственно, а рассчитываются, по известным соотношениям, из результатов измерения других, функционально связанных с ними параметров. В литературе выделяется три группы таких признаков [8, 86, 101].

К первой группе относят признаки, используемые в системах распознавания нижнего уровня. Они предназначены для определения признаков второго уровня.

Ко второй группе относят признаки, непосредственно используемые в системах распознавания второго уровня. Эти признаки, в свою очередь, предназначены для определения признаков третьего уровня.

Третью группу признаков составляют вторичные признаки, используемые в системах распознавания верхнего уровня.

Основной особенностью систем распознавания без обучения является то, что первоначальной информации достаточно для определения группы классов, составления «словаря признаков» и на основе обработки признаков распознаваемого объекта отнести его к одному из определенных классов.

В отличие от систем без обучения в обучающихся системах первоначальной информации достаточно для определения группы классов и построения словаря признаков, но недостаточно для описания классов на языке признаков.

Наиболее распространенный класс систем распознавания – обучающихся на стадии формирования, т.е. идет работа с «учителем» [19, 20, 41 и др.]. Сущность метода заключается в многократном представлении системе распознавания объектов выделенных классов с указанием принадлежности к классу.

В самообучающихся системах первоначальной информации достаточно для определения словаря признаков [8, 20, 51101 и др.].

Основной особенностью обучающихся и самообучающихся систем распознавания объектов заключается в том, что исходной информации в процессе обучения или самообучения бывает недостаточно. Это усложняет процесс распознавания графических объектов. Учитывая, что целью обучения или самообучения является формирование количества информации, необходимого для обучения системы распознавания. Например, в ряде систем используют понятие «меры близости», получаемые измерением расстояния между характерным участком распознаваемого объекта и эталонами классов для построения алгоритмов распознавания [18, 20, 47 и др.]. Таким образом, как показал проведенный выше анализ, применение методов распознавания, основанных на определении меры близости, предполагает измерение расстояний от «центра» класса до значения параметра, характеризующего существенный признак объекта.

При построении вероятностных систем распознавания применяются вероятностные методы, основанные на статистике [1, 16, 33, 48, 105 и др.]. В общем случае применение вероятностных методов распознавания также основано на определении разности, в вероятностном смысле, между характерными признаками распознаваемых объектов и классами, к которым эти объекты можно отнести.

В последнее время получили развитие системы распознавания графических объектов основанные на логических, в том числе и нечетких, методах распознавания [22, 24, 55, 89, 98, 99 и др.]. В общем случае в тех предметных областях, где существует возможность выявления логических

связей, их применение показало высокую эффективность. Применение таких систем наиболее предпочтительно для распознавания графических объектов в условиях неопределенности.

Характерной особенностью лингвистических систем распознавания является использование специальных грамматик или порождающих языков, на которых описывается объект и группа классов [101, 102, 107-109, 113, 114].

Существуют также комбинированные системы распознавания объектов в которых применяются либо комбинации вышеперечисленных, либо специально разработанные методы вычисления оценок [8, 13, 27, 41 и др].

Проведенный анализ предметной области позволил сделать вывод о том, что основными этапами процесса распознавания графических объектов, реализуемыми практически во всех системах, являются:

- Формулировка требований к результатам измерений, получаемым в процессе сбора экспериментальных данных;
- определение эталонных классов;
- обоснование модели распознавания объектов;
- определение допустимой погрешности и обоснование критериев качества распознавания;
- выбор способа оценки степени соответствия модели экспериментальным данным;
- определение процедуры верификации модели.

Выбор метода распознавания определяет выбор аналитических моделей, при этом ни один из методов распознавания, как показал анализ предметной области, не является универсальным, а используется в отдельных областях применения.

В настоящее время существует большое количество методов и моделей распознавания. На базе общего адаптивного подхода [20, 111], разработаны оптимальные и квазиоптимальные модели распознавания статических объектов, отличающиеся простотой и эффективностью. Исследовались общие теоретические вопросы, связанные с адаптивными алгоритмами

идентификации, их сходимость и скорость сходимости [35, 101, 111], потенциальные возможности таких алгоритмов, влияние корреляции в помехах измерений на структуру оптимальных алгоритмов. Новое направление в этой области связано с созданием стабильных (робастных) способов оценивания [3, 14, 18, 40]. Построенные на основе этих идей алгоритмы распознавания позволяют, в некотором смысле, оптимальным образом использовать имеющуюся априорную информацию о задаче. В этих работах приведены соответствующие алгоритмы для задач распознавания статических объектов, а в работах [26, 41] - для задач фильтрации и распознавания динамических объектов.

Для распознавания многомерных, нелинейных объектов, для которых заранее неизвестна форма зависимости «вход-выход», разработаны методы, основанные на кусочной аппроксимации [8, 19, 26, 43, 53], методы определения структуры [33, 34, 92, 93].

В ряде работ [6-8, 10, 12 и др.] представлены методы решения задач распознавания объектов с использованием специализированных программ. Они основываются на сложных дифференциальных уравнениях, а процесс распознавания заключается в нахождении значений параметров, минимизирующих отклонения измеренных значений от эталонных в какой-либо метрике.

1.1.2 Анализ методов распознавания графических объектов, основанных кластеризации

Методы распознавания графических объектов на основе кластеризации достаточно подробно описаны в [7-9, 12, 20, 26, 27, 34, 41, 97, 101, 102]. Анализ содержания данных источников позволил сделать вывод о том, что при возможности упорядочения или индексации выборки многомерных данных в процессе ее создания поставленная задача распознавания объекта решается

достаточно эффективно, так как образы распознаваемых объектов имеют сопутствующие описания, отражающие свойства и смысловое содержание признаков. Полученное описание обрабатывается известными методами классификации лингвистических данных.

Большое количество проблем требуется разрешать в случае отсутствия какого-либо индекса у графического объекта, позволяющего выполнить процедуру классификации в соответствии с заданными критериями. В настоящее время, как показал анализ литературы, представленной выше, существуют частные методы и модели решения проблем, связанных с отсутствием логических индексов для классификации или упорядочивания по заданным критериям. Область применения этих методов и моделей ограничена работой с каким-либо одним типом данных, например, как в информационной системе CBIR (Content Based Image Retrieval) [89], предназначенная для выборки образов на основе их содержания по соответствующему запросу. Основным достоинством такой системы заключается в относительной простоте реализации и, главное, имеет возможность работы с графической информацией. Недостатком этой системы является то, что по причине достаточно примитивных ассоциативных обобщающих свойств, выбранные образы объекта достаточно близки к эталонному по своей структуре, а по смысловому содержанию находятся на невысоком уровне.

Автоматизированные системы распознавания, основанные на нейронных сетях, как показал анализ литературы [24, 25, 39, 52, 110, 123], показывают достаточно хорошие результаты применения ассоциативных методов распознавания графических объектов. Основной недостаток этих систем заключается в ограниченном «объеме памяти». В связи с этим они находят ограниченное применение для решения задачи построения эффективных информационных систем распознавания графических объектов, основанных на сходстве смыслового содержания образов.

Как показали результаты анализа предметной области, основным функциональным ядром кластерных методов распознавания является модель

кластеризации образов. В такой модели непосредственно решается задача разбиения исходного множества признаков графического объекта на кластеры, содержащие множество его существенных признаков. В таких моделях структура данных существенно зависит от учтенных свойств, выбора меры различия, обоснованности критериев и способа реализации модели. В этих моделях, кластерная структура данных на множестве признаков графического объекта выявляется в процессе применения соответствующего алгоритма в сочетании с выбранной мерой близости. Анализ литературы [27, 34, 97, 101, 116] показал, что существует три основных группы схем кластеризации: последовательные, иерархические и схемы, основанные на оптимизации, например, стоимостной функции.

Основными представителями схем кластеризации последовательной группы являются BSAS (basic sequential algorithmic scheme – основная последовательная алгоритмическая схема) и MBSAS (modified sequential algorithmic scheme – модифицированная основная последовательная алгоритмическая схема) или алгоритм leader [137, 140], являющийся характерным примером, в котором проявляются основные принципы последовательных схем. Факторами, ограничивающими эффективность этих схем, как показал анализ литературы, являются: трудности в определении значения порога (приходится находить методом проб и ошибок) для четкого выявления кластерной структуры данных; зависимость результата от последовательности поступления образов для кластеризации (в процессе кластеризации образов поступающих в начале последовательности не учитывается полная информация о кластерной структуре). Для преодоления указанных недостатков, в частности выбора порога разработана двух-пороговая система с двумя порогами (two-threshold sequential algorithmic scheme – TTAS) [137]. Применение системы обработки с двумя порогами позволяет системе достаточно четко определить границы, так называемой, «серой зоны». Основным недостатком алгоритма, реализующего эту схему, является большое количество вычислительных ресурсов по сравнению с другими. Кроме того, для

улучшения результатов распознавания часто применяются процедуры слияния близких кластеров и переназначения некоторых образов в более подходящие кластеры после процедуры кластеризации [140]. Последовательные алгоритмические схемы кластеризации используются в нейросетевых моделях распознавания образов, например таких как ART2 Гроссберга [137], а также в сети MaxNet, которая является обобщением сети Хэмминга. В силу достаточно простой процедуры масштабируемости эти схемы успешно применяются при распознавании объектов представленных постоянным потоком данных, поступающих в реальном времени.

В иерархических алгоритмических схемах достаточно широко применяются генетические алгоритмы, применяемые в качестве процедуры порождения [101, 116, 125]. Они распространены для проведения исследований и интерпретации результатов, например, в социальной сфере и биологической таксономии.

Развиваются системы распознавания графических объектов, основой которых являются, так называемые, агломеративные алгоритмы [124]. В них исходная кластеризация включает множество кластеров, каждый из которых содержит один образ. В результате выполнения последовательных процедур кластеризации на каждом шаге порождаются новые классы кластеров, образующиеся путем вложения двух кластеров. Этот процесс заканчивается на шаге, в процессе выполнения которого получена группа кластеров, содержащая все множество признаков объекта в виде одного кластера. Проблемой применения этих алгоритмов является то, что при порождении неудачной кластеризации, например в результате неудачного слияния кластеров, на какой-либо итерации, то изменить это на более поздних итерациях невозможно, так как кластеры больше не разделяются.

В некоторых источниках [7, 8, 131] представлены, так называемые, разделительные алгоритмы кластеризации, при применении которых решается обратная задача – получение множества кластеров из одного, содержащего все

множество образов распознаваемого объекта. Такие алгоритмы позволяют получить наилучшую кластеризацию из нескольких подготовленных вариантов.

В иерархических алгоритмах на основе теории матриц в качестве рабочего объекта используется матрица близости [9, 23, 53].

Заканчивая анализ методов и алгоритмов кластерного анализа, применяемых для распознавания объектов нельзя не отметить, что разработаны методы кластеризации с участием человека [124, 125].

Последняя из основных групп алгоритмических схем распознавания объектов основанная на оптимизации некоторой стоимостной функции с использованием различных вычислительных методик [20].

Результаты проведенного анализа данной группы алгоритмических схем позволили сделать вывод о том, что основным недостатком рассмотренных методов и моделей является большие вычислительные затраты, обусловленные необходимостью пересчета иерархии кластеризации при изменении исходной выборки признаков распознаваемого объекта.

1.1.3 Анализ цифровых способов распознавания графических объектов

В моделях, реализующих цифровое распознавание графических объектов, применяется его представление в виде матрицы пикселей $f(m_1, m_2)$, $0 \leq m_1 \leq M_1 - 1$, $0 \leq m_2 \leq M_2 - 1$. Процесс распознавания объекта, в общем случае, заключается в выполнении какого-либо преобразования матрицы пикселей для формирования нового набора его числовых характеристик или получения нового, преобразованного образа объекта $g(n_1, n_2)$, $0 \leq n_1 \leq N_1 - 1$, $0 \leq n_2 \leq N_2 - 1$. Преобразованию, как правило, подвергаются значения параметров, характеризующих существенные признаки объекта или их расположение (индексы) и выполняются как с матрицей в целом или признаком, так и с группой признаков. При поэлементной обработке объектов значение признаков каждого элемента переводится в новое значение в соответствии с заданной

функциональной зависимостью. На практике, для снижения вычислительных затрат, переходят к табличному заданию функции преобразования. При этом существенно упрощается алгоритм преобразования. Достоинствами таких моделей являются высокое быстродействие и гибкость процедуры обработки. Недостатком - является невысокая точность результатов, обусловленная конечными значениями размера таблицы [89].

В наиболее полном виде результаты исследований по переводу графического объекта, представленного непрерывной функцией, в дискретный графический объект приведены в [95, 11, 117]. В этих работах показано, что дискретизация объекта по пространственным переменным переводит его в таблицу выборочных значений.

Другой способ замены непрерывного объекта дискретным заключается в выборе какой-либо системы ортогональных функций, вычислении коэффициентов представления объекта по этой системе (по этому базису) и замены ими объекта. Многообразие базисов обуславливает многообразие дискретных образов объекта, представленного непрерывной функцией. В [111], доказываемая, что наиболее часто применяемым является способ периодической дискретизации, в частности дискретизации с прямоугольным растром. Он является разновидностью способа ортогонального базиса, поскольку в качестве элементов используются сдвинутые дельта-функции [111].

Самое главное при выборе способа замены объекта, представленного непрерывной функцией дискретным образом заключается в определении условий, при которых такая замена не сопровождается потерей информации, содержащейся в объекте, представленном непрерывной функцией [33].

В большинстве случаев вид графического объекта искажен помехами, обусловленными как влиянием некоторых участков самого объекта, например, при анализе объектов на сложном фоне (рисунок 1.1, приложение А), так и преднамеренными помехами. Это затрудняет как распознавание объекта человеком-оператором, так и его автоматическое распознавание. Поэтому при

цифровой обработке графических объектов необходимо устранять влияние тех участков графического объекта, которые «зашумляют» его основную часть.

Как правило ослабление действия помех достигается применением процедуры фильтрации в пространственной и частотной областях [124]. При пространственной фильтрации графического объекта преобразование выполняется непосредственно над значениями его отсчетов. Результатом фильтрации является оценка полезного сигнала объекта. Анализ литературы [95, 111] показал, что функция пространственных координат, представляющая графический объект, изменяется медленнее, чем функция, описывающая помеху. Поэтому при оценке полезного сигнала в каждой точке графического объекта процедуру фильтрации целесообразно применять и к окрестности этой точки. В результате формируется множество отсчетов, описывающее конфигурацию окрестности точки отсчета графического объекта.

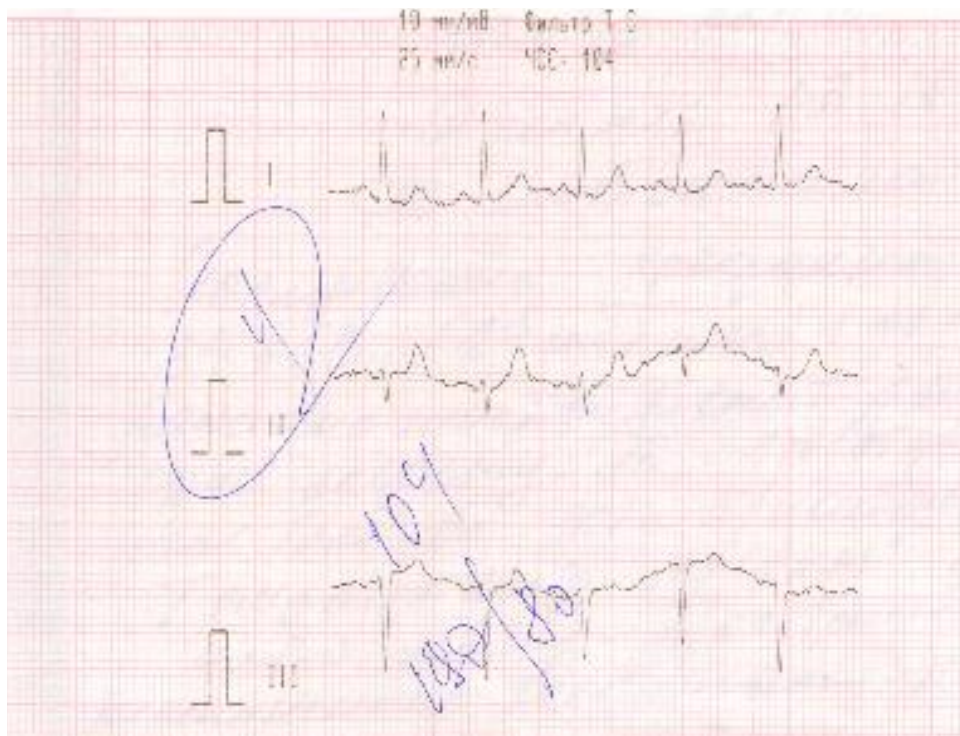


Рисунок 1.1 – Зашумленный графический объект (фрагмент электроэнцефалограммы)

Таким образом, основная идея процедуры фильтрации основывается на использовании, как данных текущей точки, так и ее окрестности. В этом проявляется существенное отличие фильтрации от рассмотренных выше поэлементных процедур. Процедура фильтрации не является поэлементной процедурой обработки графического объекта. Тем не менее, сложность нахождения точных решений приводит к необходимости создания различных приближенных методов и процедур обработки графических объектов в интересах их распознавания.

В практике цифровой обработки объектов широко используется масочная фильтрация [111] и пороговая обработка по яркости [26].

Анализ предметной области показал, что методы и модели распознавания базируются на строго формализованном описании каждого класса объектов или явлений на языке признаков. Различные типы моделей рассматриваемых объектов, систем или процессов используются на стадии создания систем управления этими объектами и на стадии их эксплуатации. Это обуславливает актуальность решения задачи разработки моделей для информационных систем распознавания графических объектов в условиях неопределенности.

1.2 Обзор программных средств распознавания графических объектов

В настоящее время создано множество программных средств, применяемых при решении задачи распознавания графических объектов.

Одним из таких программных средств является комплекс «ГРАД» [129]. Он обеспечивает выполнение большинства расчетов, связанных с проточной частью газотурбинных и комбинированных двигателей и установок, работающих на любом топливе и окислителе. С его помощью обеспечивается возможность идентификации математической модели двигателя по результатам испытаний на стенде или в полете.

Программный комплекс «ДИЗЕЛЬ-РК» [130] отличается возможностью

идентификации основных характеристик системы путем пошагового решения системы разностных уравнений сохранения энергии, массы, а также уравнения состояния, записанных для открытых термодинамических систем.

В биологии и медицине очень сильно выражена тенденция приложения точных математических методов и ресурсов компьютерного оборудования для исследования процессов, встречающихся в организме человека и, в частности, в его главной физиологической системе - в системе циркуляции крови. Программный комплекс моделирования кровообращения [132]. Возможности этого комплекса позволяют оценивать параметры математической модели реального организма на основе их идентификации. Это позволяет прогнозировать поведение исследуемого организма в различных режимах (физические нагрузки, сердечная недостаточность, атеросклероз, гипоксия и т. д.). Результаты, полученные при использовании программной системы применяются специалистами в различных предметных областях, включая образовательный процесс, медицину, физиологические исследования, например, нагрузок пилотов, космонавтов, спортсменов, альпинистов и т.д.

В Тамбовском государственном техническом университете «Экспертная система энергосберегающего управления» [32] позволяет решать широкий круг исследовательских, производственных и учебных задач анализа и синтеза оптимальных управляющих воздействий на множестве состояний функционирования исследуемой системы. Основное назначение системы – идентификация математической модели объекта управления, проверка выполнения условий существования решения задачи, определение видов функций оптимального управления, нахождение параметров управляющих воздействий, проверка выполнения интегральных ограничений на запас энергоресурсов, формирования алгоритмов и данных для записи в микропроцессорное управляющее устройство.

Пакет расширения «System Identification Toolbox» программного пакета «MatLab» [56] включает средства для создания математических моделей различных, в том числе и линейных динамических объектов, на основе

наблюдаемых входных и выходных данных. Удобный графический интерфейс упрощает как предварительную обработку данных, так и диалоговый процесс идентификации модели.

Все проанализированные программные комплексы представляют собой различные в функциональном плане пакеты программ, позволяющие создавать математические модели реальных объектов и проводить их распознавание. Область применения данных программных комплексов ограничена использованием моделей объектов в конкретной предметной области, что не всегда позволяет их применять для решения задач распознавания графических объектов в других отраслях. В случае существования программного средства для решения абсолютно любых задач распознавания, такой пакет будет представлять экспертную систему с интеллектуальным интерфейсом и распределенными базами знаний.

1.3 Методы оценки информационной потребности пользователя

1.3.1 Традиционные методы оценки информационной потребности пользователя

Задача оценки информационной потребности пользователя возникает в связи с необходимостью оценки качества информационного поиска, в том числе и распознавания графических объектов.

Известно, что основными целями информационного поиска [139] являются:

1. Поиск необходимых сведений об объекте и установление его наличия в системе источников.
2. Поиск информационных источников, в которых есть или может содержаться нужная информация.
3. Поиск фактических сведений, содержащихся в информационных источниках.

Эти цели определяют три основных вида информационного поиска: библиографический, документальный и фактографический.

При оценке степени достижения целей и эффективности видов информационного поиска необходимо принимать во внимание влияние многих факторов и критериев, образующих соответствующую группу или класс в силу их большого разнообразия. Совокупность этих классов или групп факторов и критериев образуют систему, на которой основано построение модели распознавания графических объектов для конкретной задачи. В этом смысле система факторов и критериев является отправной точкой для определения необходимых характеристик и особенностей данного информационного поиска.

При организации вариантов ответов на информационный запрос в существующих системах реализованы модели максимально направленные на удовлетворение потребности пользователя. Информация о требуемой потребности извлекается, как правило, из содержания самого запроса. При этом в таких системах варианты ответов на запрос удовлетворяют, принятым смысловым и формальным релевантным критериям [140]. Тем не менее, в вариантах ответов на запрос, формируемых по этим критериям зачастую содержится одна и та же информация или система направляет пользователя в различных вариантах ответа к одному и тому же источнику. Таким образом, основные требования к функционированию информационной системы, а именно: единственности, полноты и непротиворечивости, не выполняются. Причина этого видимо кроется в недостаточной объективности релевантных критериев.

В последнее время, в различного рода информационных системах, в качестве критериев, которым должен соответствовать результат формирования ответа на запрос пользователя, наряду с релевантными критериями, стали применяться критерии, основанные определениях полноты и точности [59, 73, 128 и др.].

Точность варианта ответа на запрос пользователя к информационной системе определяется как отношение количества релевантных вариантов ответа, сформированных системой, к общему количеству вариантов ответов:

$$P_{recision} = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|} ,$$

где D_{rel} – множество релевантных вариантов ответов на запрос пользователя;

D_{retr} – полное множество вариантов ответов на запрос пользователя, сформированных информационной системой.

Полнота варианта ответа на запрос пользователя к информационной системе определяется как отношение количества релевантных вариантов ответов на запрос пользователя к информационной системе к общему количеству релевантных вариантов ответов, имеющихся в базе данных информационной системы:

$$P_{recal} = \frac{|D_{rel} \cap D_{retr}|}{|D_{rel}|} .$$

Эффективной мерой для совместной оценки точности и полноты является F -норма, которая определяется как взвешенное гармоническое среднее точности (P) и полноты (R):

$$F = \frac{1}{\alpha \frac{1}{P} + (1-\alpha) \frac{1}{R}}, \alpha \in [0,1].$$

Обычно F -норму записывают в виде

$$F = \frac{(\beta^2+1)PR}{\beta^2P+R}, \beta^2 = \frac{1-\alpha}{\alpha}, \beta^2 \in [0, \infty].$$

Использование сбалансированной F -нормы не является обязательным: при $0 < \beta < 1$ предпочтение отдаётся точности, а при $\beta > 1$ больший вес приобретает полнота [89].

Для оценки качества удовлетворения информационной потребности пользователя применяется также такой показатель как коэффициент потерь определенный как количественная характеристика результатов информационного поиска, рассчитываемая как отношение числа не представленных пользователю релевантных вариантов ответов, в результате действий в системе по их формированию, к общему числу всех возможных релевантных вариантов ответов [28, 141].

Немаловажным, с точки зрения быстроты принятия решения, является такой показатель как оперативность поиска вариантов ответов, определенный как время, затрачиваемое на выполнение поиска, с момента формирования запроса и принятия решения на формирование вариантов ответов до момента получения результатов по запросу.

Таким образом, результаты анализа методов, критериев и показателей оценки эффективности информационного поиска показали, что, понятие степени соответствия запроса является субъективным, а степень соответствия зависит от конкретного человека, оценивающего результаты выполнения запроса.

1.3.2 Методы оценки информационной потребности пользователя на основе пертинентности

В настоящее время основными показателями пертинентности результатов работы информационных систем на запросы пользователя, являются [28, 134, 141]:

пертинентностная точность выдачи ($\mathcal{Q}_V$);

подлинная пертинентность документа (Ψ);

максимальная пертинентность документов выдачи (Ψ_{max});

кумулятивная пертинентность выдачи Ψ_{KUM} ;

векторная пертинентность выдачи ($\vec{\eta} = (\Psi, N_{View})$);

средние значения указанных показателей, вычисляемые по 32-м точечным значениям: $\bar{\vartheta}_V^{[32]}, \bar{\psi}^{[32]}, \bar{\psi}_{\max}^{[32]}, \bar{\Psi}^{[32]}, \bar{\eta}^{[32]} = (\bar{\Psi}^{[32]}, \bar{N}_{View}^{[32]})$.

Кроме того, существует методика оценки пертинентности на основе экспертных суждений и таких показателей степени контентной эквивалентности как [135, 136, 141]:

Степень емкостной контентной эквивалентности

$$\rho_V(A_1, A_2) = \begin{cases} \frac{V_{\min}}{V_{\max}}, & \text{если } V_{\max} > 0; \\ 1, & \text{если } V_{\min} = V_{\max} = 0. \end{cases}$$

Степень относительной контентной эквивалентности

$$\rho_C(A_1, A_2) = \begin{cases} \frac{V_{\text{Com}}}{V_{\min}}, & \text{если } V_{\min} > 0; \\ 0, & \text{если } V_{\min} = 0. \end{cases}$$

Степень абсолютной контентной эквивалентности

$$\rho_A(A_1, A_2) = \begin{cases} \frac{V_{\text{Com}}}{V_{\max}}, & \text{если } V_{\max} > 0; \\ 0, & \text{если } V_{\max} = 0. \end{cases}$$

Степень эквивалентности по обобщенной ранговой корреляции

$$\rho_B(A_1, A_2) = \begin{cases} 1 - \frac{\sum_{k=1}^{\eta} (r_k^{[1]} - r_k^{[2]})^2}{S_{\max}}, & \text{если } \eta > 0; \\ 0, & \text{если } \eta = 0. \end{cases}$$

Степень эквивалентности по условной ранговой корреляции

$$\rho_{SB}(A_1, A_2) = \begin{cases} 1 - \frac{3 \sum_{k=1}^{\eta} (p_k^{[1]} - p_k^{[2]})^2}{\eta(\eta^2 - 1)}, & \text{если } \eta > 1; \\ 0, & \text{если } \eta \leq 1. \end{cases}$$

Степень эквивалентности по плотности общих элементов

$$\rho_{cd}(A_1, A_2) = \begin{cases} \frac{\min(k_{Last} - k_{First}, l_{Last} - l_{First})}{\max(k_{Last} - k_{First}, l_{Last} - l_{First})}, & \text{если } \max(k_{Last} - k_{First}, l_{Last} - l_{First}) > 0; \\ 0, & \text{если } \max(k_{Last} - k_{First}, l_{Last} - l_{First}) = 0. \end{cases}$$

Анализ литературы [129-131] показал, что для повышения релевантности формируемого множества вариантов ответа на запрос пользователя сами потребности различают по видам. Так, под текущими информационными потребностями пользователя понимается совокупность таких показателей, как: иметь представление о происходящих, в данный момент, событиях; своевременное получение информации о новых достижениях в интересующей предметной области; потребность оперативного получения информации, необходимой для принятия решения. Под ретроспективными информационными потребностями понимается информация, обеспечивающая, например, начало выполнения новой темы, т.е. это такая информация, которая наиболее полно обеспечивает пользователя о предыдущих результатах исследований по данной теме [28, 130].

Решение задачи построения моделей распознавания графических объектов, невозможно без рассмотрения свойств информации [126, 129, 136]: объективности, достоверности, полноты, актуальности, адекватности.

Поскольку задачу получения информации человечество решает многие столетия, то увеличение объема информационных ресурсов, потенциально доступных исследователям представляется разработать аналитические и процедурные модели для информационных систем распознавания графических объектов, обеспечивающие требуемую пертинентность.

Однако в литературе, например в [36, 52, 76, 95, 134], отмечается невысокое качество результатов поиска, даже при использовании таких известных и зарекомендовавших себя информационно-поисковых систем (ИПС), как Гугл или Яндекс. Одной из причин является неправильная или неполная формулировка запросов к ИПС. Пользователь обычно формулирует запрос на естественном языке и зачастую не знает или не понимает, какие слова лучше использовать в запросе, а также особенности конкретной поисковой

машины и уж гораздо реже использует специализированный язык запросов ИПС. Поэтому на смену релевантности – соответствию результатов поиска запросу, приходит понятие пертинентность – соответствие полученной информации информационной потребности [36, 52, 95, 117]. Однако, для поиска на основе пертинентности необходимо формализовать информационную потребность пользователя. Поэтому, предложен подход к формализации информационной потребности на основе теории нечётких множеств.

В работах [21, 75, 86, 98, 102] обосновывается понятие коллокации. Интерес к коллокациям повысился, когда стало ясно, что методы поиска, базирующиеся на появлении слов в тексте или признаков объекта подошли к пределу своих возможностей, а эффективный семантический анализ трудоёмок и недостаточно исследован [52].

Ряд исследователей предложил использовать коллокации для поиска и кластеризации текстовых корпусов и экспериментально показал эффективность метода [7, 9, 98]. В этих работах показано применение аппарата математической статистики для формализации и обработки коллокаций, что, по нашему мнению, привело к упрощению первоначальной модели до простейшей, двухтермовой (состоящей из двух слов) коллокации. Как показанов п 1.1 и 1.2 математический аппарат статистической теории целесообразно применять в условиях, когда можно сделать большое число проб (замеров), часть из которых завершиться положительным результатом. Однако что считать положительным результатом, когда коллокация это некоторый закон взаимного расположения термов? На каком расстоянии друг от друга должны располагаться термы, чтобы считаться коллокацией, а на каком их уже нельзя считать таковой (под расстоянием между словами мы понимаем количество слов между ними)? И являются ли два экземпляра коллокации, отличающиеся наличием некоторого слова в промежутке между словами коллокации, экземплярами одной и той же коллокации. Эти и другие вопросы говорят о недостаточности функционала аппарата математической статистики для

построения эффективной теории поиска и кластеризации текстовых сведений на основе коллокаций. Поэтому, в ряде работ [68, 73, 75, 102] предложено, для создания «коллокационной» модели применить теорию нечётких множеств. Это позволяет задать коллокацию и нечётким образом определить расстояние между термами в объекте. Кроме того математический аппарат теории нечётких множеств позволяет не просто выявлять принадлежность некой группы слов множеству коллокаций, но и ставить в соответствие таким объектам значение функции принадлежности множеству коллокаций. То есть, предложено считать произвольную часть графического объекта принадлежащей нечёткому подмножеству соответствующей коллокации, а степень принадлежности будет тем больше, чем ближе сравниваемые части объекта к друг другу. В результате формализуется информационная потребность пользователя в виде совокупности значений соответствующего кортежа элемента коллокации, значения лингвистической переменной и лингвистической переменной. В результате вычисляется функция принадлежности, формализующая пертинентность результата распознавания по признаку, соответствующему коллокации.

Обосновано, что подход к построению аналитических и процедурных моделей распознавания графических объектов, представлению и оценке информационной потребности пользователя, при их использовании, необходимо осуществлять с позиций теории нечётких множеств.

Обосновано, что показателем соответствия сведений информационной потребности пользователя при организации их эффективного поиска является пертинентность, оценка пертинентности результатов поиска – важная составляющая задачи повышения эффективности функционирования ИС.

1.4 Выводы по главе 1. Постановка задачи на исследование

На основании результатов анализа принципов построения и функционирования средств распознавания графических объектов в

информационных системах, научно-методического аппарата подходов к решению задачи распознавания графических объектов в условиях неопределенности можно сделать следующие основные выводы:

1. В целом из анализа предметной области можно сделать вывод об отсутствии соответствующих моделей для распознавания графических объектов и оценки качества этого процесса. Определено место разрабатываемых моделей на множестве существующих моделей и методов распознавания, представленное на рисунке 1.5.



Рисунок 1.5 – Место разработанных моделей на множестве существующих методов

В настоящее время решение задачи распознавания графических объектов производится методами, не полностью учитывающими особенности распознавания графических объектов в условиях неопределенности, главными из которых является малый объем априорных данных относительно эталонных описаний распознаваемых образов, сложность создания оптимальных

алгоритмов распознавания сигналов на основе адекватных вероятностных моделей по совокупности показателей качества.

2. Подход к построению аналитических и процедурных моделей распознавания графических объектов, представлению и оценке информационной потребности пользователя, при их использовании, необходимо осуществлять с позиций теории нечётких множеств.

3. Современные исследования в области информационного поиска требуют перехода от оценки поиска по релевантности – соответствия результатов поиска информационному запросу, к оценке по пертинентности – соответствию результатов распознавания информационной потребности пользователя. Показателем качества распознавания графического объекта является пертинентность, оценка пертинентности результатов – важная составляющая задачи повышения эффективности функционирования ИС.

4. Для достижения цели исследования и решения поставленной научной задачи необходимо, с применением методологии теорий информационного поиска, нечетких множеств, распознавания образов решить частные задачи:

- разработать аналитические модели распознавания графических объектов и процесса поиска сведений, результаты которого обладают максимальной пертинентностью.

- разработать процедурную модель решения задачи распознавания графических объектов.

- оценить эффективность, полученных в ходе исследования, моделей и живучесть системы, построенной с применением разработанных моделей.

ГЛАВА 2. АНАЛИТИЧЕСКИЕ МОДЕЛИ ДЛЯ ИНФОРМАЦИОННЫХ СИСТЕМ РАСПОЗНАВАНИЯ ГРАФИЧЕСКИХ ОБЪЕКТОВ

2.1 Модели информационных процессов при распознавании графических объектов

Как было показано в первой главе, наиболее перспективным подходом для построения математических моделей информационных процессов при распознавании графических объектов в настоящее время является теория нечетких множеств, тем не менее, построенные логико-лингвистические модели приводят к необходимости введения лингвистических переменных и использования большого числа термов, что влечет за собой большие вычислительные затраты. Для преодоления этого недостатка перспективным и очень интересным является метод, описанный в [44, 120], который основан на том, что результатом работы правил логико-лингвистической модели является не терм, а алгебраическая зависимость. Данный подход использован для построения комбинированной и обобщенной моделей информационных процессов распознавания графических объектов.

Комбинированная аналитическая модель имеет вид [58, 62]:

Правило i : Если $z_1(t)$ есть F_{i1} и ... и $z_g(t)$ есть F_{ig} , то

$$\begin{aligned} \dot{x}(t) &= A_i x(t) + B_i u(t) + w(t) \\ y(t) &= C_i x(t) + v(t), \text{ для } i = 1, 2, \dots, L \end{aligned} \quad (2.1)$$

где $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \in R^{n \times 1}$ - вектор состояний, $u(t) = [u_1(t), u_2(t), \dots, u_m(t)]^T \in R^{m \times 1}$ - вектор управления, $w(t) = [w_1(t), w_2(t), \dots, w_n(t)]^T \in R^{n \times 1}$ - влияние внешней среды, $y(t)$ - выходы системы, $v(t)$ - погрешность измерения, F_{ij} - лингвистические значения исходных параметров $z_1(t), z_2(t), \dots, z_g(t)$, $A_i \in R^{n \times m}$, $B_i \in R^{n \times m}$; L - число правил в рассматриваемой логико-лингвистической модели.

Использование качественной информации, представленной в логико-лингвистической модели с соответствующими термами, для формализации

которых применяются функции принадлежности, дает возможность перейти от модели (2.1) к обобщенной модели:

$$\dot{x}(t) = \frac{\sum_{i=1}^L \mu_i(z(t)) [A_i x(t) + B_i u(t)]}{\sum_{i=1}^L \mu_i(z(t))} + w(t) = \sum_{i=1}^L h_i(z(t)) [A_i x(t) + B_i u(t)] + w(t), \quad (2.2)$$

$$y(t) = \frac{\sum_{i=1}^L \mu_i(z(t)) [C_i x(t)]}{\sum_{i=1}^L \mu_i(z(t))} + v(t) = \sum_{i=1}^L h_i(z(t)) [C_i x(t)] + v(t), \quad (2.3)$$

где

$$\begin{aligned} \mu_i(z(t)) &= \prod_{j=1}^g F_{ij}(z_j(t)), \\ h_i(z(t)) &= \frac{\mu_i(z(t))}{\sum_{i=1}^L \mu_i(z(t))}, \\ z(t) &= [z_1(t), z_2(t), \dots, z_g(t)] \end{aligned} \quad (2.4)$$

и $F_{ij}(z_j(t))$ - функция принадлежности $z_j(t)$ в F_{ij} .

Приняв для всех t

$$\mu_i(z(t)) \geq 0 \text{ и } \sum_{i=1}^L \mu_i(z(t)) > 0, \text{ при } i = 1, 2, \dots, L,$$

$$\text{получим } h_i(z(t)) \geq 0, \text{ при } i = 1, 2, \dots, L \quad (2.5)$$

$$\text{и } \sum_{i=1}^L h_i(z(t)) = 1. \quad (2.6)$$

Считается, что известна эталонная модель процесса, имеющая вид:

$$\dot{x}_3(t) = A_3 x_3(t) + r(t), \quad (2.7)$$

где $x_3(t)$ - исходное состояние; A_3 - асимптотически устойчивая матрица; $r(t)$ - ограниченные контрольные входные данные.

Предполагается, что $x_3(t)$ представляет желательную траекторию, протекания информационного процесса для $x(t)$ при $t \geq 0$. В этом случае процесс распознавания рассматривается как процесс слежения за рассогласованием $x(t) - x_3(t)$, который реализуется следующим образом:

$$\frac{\int_0^{t_f} \{ [x(t) - x_{\mathcal{O}}(t)]^T Q [x(t) - x_{\mathcal{O}}(t)] \} dt}{\int_0^{t_f} \tilde{w}(t)^T \tilde{w}(t) dt} \leq \rho^2 \quad (2.8)$$

или

$$\int_0^{t_f} \{ [\mathbf{x}(t) - \mathbf{x}_r(t)]^T Q [\mathbf{x}(t) - \mathbf{x}_r(t)] \} dt \leq \rho^2 \int_0^{t_f} \tilde{w}(t)^T \tilde{w}(t) dt, \quad (2.9)$$

где $\tilde{w}(t) = [v(t), w(t), r(t)]^T$, $w(t)$ - влияние условий неопределенности, $v(t)$ - ошибки измерения; t_f - время существования реализации; Q - положительно определенная матрица весов; ρ - предписанный уровень ослабления.

Нечеткая модель слежения представляет собой совокупность следующих правил:

Правило слежения i : Если $z_1(t)$ есть F_{i1} и ... и $z_g(t)$ есть F_{ig} , то

$$\dot{\hat{x}}(t) = A_i \hat{x}(t) + B_i u(t) + L_i (y(t) - \hat{y}(t)), \quad (2.10)$$

где L_i - выигрыш слежения для i -го правила и $\hat{y}(t) = \sum_{i=1}^L h_i(z(t)) C_i \hat{x}(t)$.

Нечеткая система слежения представлена следующим образом:

$$\dot{\hat{x}}(t) = \sum_{i=1}^L h_i(z(t)) [A_i \hat{x}(t) + B_i u(t) + L_i (y(t) - \hat{y}(t))]. \quad (2.11)$$

Обозначим

$$e(t) = x(t) - \hat{x}(t). \quad (2.12)$$

Дифференцируя (2.12), получим

$$\begin{aligned} \dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) = \\ &= \sum_{i=1}^L \sum_{j=1}^L h_i(z(t)) h_j(z(t)) [A_i x(t) + B_i u(t) + w(t)] - [A_i \hat{x}(t) + B_i u(t) + L_i C_j (x(t) - \hat{x}(t)) + L_i v(t)] = \\ &= \sum_{i=1}^L \sum_{j=1}^L h_i(z(t)) h_j(z(t)) [(A_i - L_i C_j) e(t) - L_i v(t)] + w(t). \end{aligned} \quad (2.13)$$

Нечеткая модель слежения представлена совокупностью правил:

Правило управления j : Если $z_1(t)$ есть F_{j1} и ... и $z_g(t)$ есть F_{jg} , то

$$u(t) = K_j [\hat{x}(t) - x_r(t)], \text{ при } i = 1, 2, \dots, L. \quad (2.14)$$

При этом нечеткий закон управления имеет вид:

$$u(t) = \frac{\sum_{j=1}^L \mu_j(z(t)) [K_j(\hat{x}(t) - x_r(t))]}{\sum_{j=1}^L \mu_j(z(t))} = \sum_{j=1}^L h_j(z(t)) [K_j(\hat{x}(t) - x_r(t))]. \quad (2.15)$$

Для определения нечеткой системы слежения и закона управления, необходимо чтобы исходные параметры $z(t)$ были измеримыми. Однако, если эти параметры неопределенны, а известны лишь их оценки $\hat{z}(t)$, то ситуация становится более сложной.

После преобразования система принимает следующий вид:

$$\dot{\tilde{x}}(t) = \sum_{i=1}^L h_i(z(t)) \sum_{j=1}^L h_j(z(t)) [\tilde{A}_{ij} \tilde{x}(t) + \tilde{E}_i \tilde{w}(t)]. \quad (2.16)$$

Обозначим

$$\begin{aligned} \tilde{A}_{ij} &= \begin{bmatrix} A_i - L_i C_j & 0 & 0 \\ -B_i K_j & A_i - B_i K_j & -B_i K_j \\ 0 & 0 & A_r \end{bmatrix}, \\ \tilde{x}(t) &= \begin{bmatrix} e(t) \\ x(t) \\ x_r(t) \end{bmatrix}, \quad \tilde{w}(t) = \begin{bmatrix} v(t) \\ w(t) \\ r(t) \end{bmatrix}, \\ \tilde{E}_i &= \begin{bmatrix} -L_i & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}. \end{aligned} \quad (2.17)$$

С учетом (2.17) система из (2.16) примет вид

$$\dot{\tilde{x}}(t) = \sum_{i=1}^L h_i(z(t)) \sum_{j=1}^L h_j(z(t)) [\tilde{A}_{ij} \tilde{x}(t) + \tilde{E}_i \tilde{w}(t)]. \quad (2.18)$$

С учетом начального состояния система слежения за параметрами в (9) может быть модифицирована как

$$\begin{aligned} \int_0^{t_f} \{(\tilde{x}(t) - \tilde{x}_r(t))^T Q (\tilde{x}(t) - \tilde{x}_r(t))\} dt = \\ = \int_0^{t_f} \tilde{x}^T(t) \tilde{Q} \tilde{x}(t) dt \leq \tilde{x}^T(0) \tilde{P} \tilde{x}(0) + \rho^2 \int_0^{t_f} \tilde{w}^T(t) \tilde{w}(t) dt, \end{aligned} \quad (2.19)$$

где $\tilde{P}$ - симметричная положительно определенная матрица весов и

$$\tilde{Q} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & Q & -Q \\ 0 & -Q & Q \end{bmatrix}.$$

Далее необходимо определить нечеткий закон управления (2.15) с гарантируемым слеживанием за параметрами в (2.1) для всех $w(t)$. После того, необходимо минимизировать ρ^2 так, чтобы величина ошибки слежения за параметрами, определенная из (2.19), была минимальна.

Отметим, что система, имеющая вид:

$$\dot{\tilde{x}}(t) = \sum_{i=1}^L h_i(z(t)) \sum_{j=1}^L h_j(z(t)) \tilde{A}_{ij} \tilde{x}(t) \quad (2.20)$$

квадратично устойчива [62, 66, 87]. Следовательно, величина ошибки слежения минимальна для заданного ρ^2 . Это обстоятельство является основанием для распознавания графического объекта.

2.2 Аналитическая модель распознавания графических объектов на основе сплайнов

В общем случае построение модели для конкретного объекта из класса линейных моделей требует по результатам измерений входного $x(t)$ и выходного $y(t)$ сигналов отнесения данного объекта к определенному классу объектов. Если динамические характеристики объекта описываются оператором A , то при наличии результатов измерений входной и выходной случайных функций (переменных) задача идентификации сводится к определению некой оценки $\tilde{A}$ оператора A .

Оптимальный оператор объекта в классе всех возможных операторов по критерию минимума среднего квадрата ошибки, называемый также регрессией выходной реализации $y(t)$ относительно входной $x(t)$, определяется следующим образом:

$$\tilde{y}(t) = \tilde{A}x(s) = M\{y(t) | x_s t, s \in T\}. \quad (2.21)$$

Для модели линейного объекта оптимальная оценка импульсной характеристики $g(t,s)$ по критерию минимума среднего квадрата ошибки определяется из уравнения

$$r_{yx}(t, \sigma) = \int_{t-T}^t g(t, s) r_{xx}(t, \sigma) ds, \quad (2.22)$$

где T - интервал времени наблюдения, $r_{xx}(t, \sigma)$ - корреляционная функция входной реализации сигнала $x(t)$, а $r_{yx}(t, \sigma)$ - взаимно-корреляционная функция реализаций сигналов $y(t)$ и $x(t)$ [62].

Когда случайные функции $x(t)$ и $y(t)$ являются стационарными и стационарно связанными, оптимальная оценка оператора определяется из уравнения

$$r_{yx}(t) = \tilde{A} r_{xx}(t - \tau), \quad (2.23)$$

а импульсная характеристика стационарной линейной системы из интегрального уравнения Фредгольма первого рода

$$r_{yx}(t, \sigma) = \int_0^{\infty} g(\tau) r_{xx}(t - \tau) d\tau, \quad -\infty < t < \infty. \quad (2.24)$$

Для решения задачи распознавания при использовании (2.24) производится классификация априорной информации [62, 65, 66, 84]. Классы определяются качественным поведением $r_{xx}(t)$ и $r_{yx}(t)$.

В ИС корреляционные функции $r_{xx}(t)$ и $r_{yx}(t)$ поступают в виде цифрового графического образа объекта или в виде набора значений $\{t_i, f(t_i)\}$ [87, 88].

Представление и хранение функций в ИС реализовано с помощью кубических сплайнов. Это обеспечивает преобразование набора $\{t_i, f(t_i)\}$ в множество соответствующих кубических сплайнов $\{S_i(t)\}$. Для этого задается сетка $a = t_0 < t_1 < t_2 < \dots < t_n = b$ и соответствующие значения $f(t_i)$. Расстояние между смежными узлами: $h_i = x_i - x_{i-1}$, $i = \overline{1, n}$. В соответствии с [62, 63, 87, 88] функция $S(t)$ удовлетворяет следующим условиям:

1. На каждом отрезке $[t_{i-1}, t_i]$ функция $S(t)$ является полиномом третьей степени.

2. Функция $S(t)$, ее первая $S'(t)$ и вторая $S''(t)$ производные непрерывны на сегменте $[a, b]$.

$$3. \quad S(t_i) = f(t_i) = f_i, \quad i = \overline{1, n}.$$

4. На концах сегмента $[a, b]$ функция $S''(t)$ удовлетворяет условиям $S''(a) = S''(b) = 0$. Однако следует отметить, что на концах сегмента $[a, b]$ могут быть заданы в принципе и другие условия, например: $S''(a) = A$, $S''(b) = B$.

Таким условиям удовлетворяет только один сплайн [62, 87].

Для построения сплайна необходимо определить коэффициенты полиномов третьей степени на каждом из отрезков $[t_{i-1}, t_i]$. Для этого сопоставим отрезку $[t_{i-1}, t_i]$ полином $S_i(t)$, для удобства записанный в виде

$$S_i(t) = a_i + b_i(t - t_i) + \frac{c_i}{2}(t - t_i)^2 + \frac{d_i}{6}(t - t_i)^3,$$

$$t \in [t_{i-1}, t_i], \quad i = \overline{1, n}.$$

Непрерывность сплайна в узлах t_i , $i = \overline{1, n-1}$ и выполнения условия (2.23) при $i = 0$ дает

$$b_i h_i - \frac{c_i}{2} h_i^2 + \frac{d_i}{6} h_i^3 = f_i - f_{i-1}, \quad i = \overline{1, n}. \quad (2.25)$$

Непрерывность первой и второй производной сплайна означает, что

$$c_i h_i - \frac{d_i}{2} h_i^2 = b_i - b_{i-1}, \quad i = \overline{2, n}, \quad (2.26)$$

$$d_i h_i = c_i - c_{i-1}, \quad i = \overline{2, n}. \quad (2.27)$$

Граничные условия (2.24) дают еще два уравнения:

$$\begin{cases} S_1''(t_0) = S_1''(a) = c_1 - d_1 h_1 = 0, \\ S_n''(t_n) = S_n''(b) = c_n = 0. \end{cases} \quad (2.28)$$

Решение системы линейных уравнений (2.25)-(2.28) позволяет определить значения b_i , c_i , d_i , $i = \overline{2, n}$.

В [120] представлены таблицы откликов линейных систем на возмущения видов $r_{xx}^{T,1}(t) = Ae^{-\alpha|t|}$, $r_{xx}^{T,2}(t) = Ae^{-\alpha|t|} \cos \omega t$, $r_{xx}^{T,3}(t) = (A + A_1 |t|)e^{-\alpha|t|}$, а соответствующие

им корреляционные функции входа и выхода $r_{yx}^T(t)$ представлены в ИС в виде

$$S_{r_{yx}}^{T,j,l}(t) = \begin{cases} S_{r_{yx}}^{T,j,l}(t) \text{ при } t \geq 0, \\ S_{r_{yx}}^{T,j,l}(t) \text{ при } t < 0; \end{cases} \text{ где } j = \overline{1,3}.$$

В общем случае $S(t) = \{S_i(t)\}$, $i = \overline{1,k}$, где k – количество классов распознавания, к одному из которых относится распознаваемый объект.

Таким образом, соотнесение графического объекта $x_i(t)$ к определенному классу K_i , $i = \overline{1,n}$ осуществляется набором правил, фрагмент которых имеет вид:

$$\begin{aligned} & \text{IF } r_{yx}^1(t) \text{ is } M_1^1 \text{ and } r_{yx}^1(t) \text{ is } M_1^1 \dots r_{xx}^1(t) \text{ is } M_n^1 \\ & \text{THEN } S^1(t) = a_i^1 + b_i^1(t - t_i) + \frac{c_i^1}{2}(t - t_i)^2 + \frac{d_i^1}{6}(t - t_i)^3 \text{ AND } K_i = s_{i_1}^{p_{i_1}}, t \in [t_{i-1}, t_i], i = \overline{1,n}; \end{aligned}$$

•
•
•

$$\begin{aligned} & \text{IF } r_{yx}^k(t) \text{ is } M_1^k \text{ and } r_{yx}^k(t) \text{ is } M_1^k \dots r_{xx}^k(t) \text{ is } M_n^k \\ & \text{THEN } S^k(t) = a_i^k + b_i^k(t - t_i) + \frac{c_i^k}{2}(t - t_i)^2 + \frac{d_i^k}{6}(t - t_i)^3 \text{ AND } K_i = s_{i_1}^{p_{i_1}}, t \in [t_{i-1}, t_i], i = \overline{1,n}. \end{aligned} \quad (2.29)$$

В выражении (2.29): M_i^k – значение соответствующего элемента термножества.

Таким образом, значения корреляционных функций, полученных экспериментально (верхний индекс «э» $r_{xx}^{\text{э}}(t)$ и $r_{yx}^{\text{э}}(t)$ преобразуются в соответствующий набор признаков с применением процедуры построения сплайнов, а соотнесение графического объекта $x_i(t)$ к определенному классу K_i , $i = \overline{1,n}$ осуществляется набором правил (2.29).

Построенная модель может быть применена для формализации информационно-управляющих воздействий, т.е. для принятия решения по результатам распознавания графического объекта. Для построения формализации используется p нечетких правил, i -ое правило имеет следующий вид:

$$\begin{aligned} & \text{Правило } i : \text{IF } f_1(x(t)) \text{ is } M_1^i \text{ and } \dots \text{ and } f_p(x(t)) \\ & \text{is } M_p^i \text{ THEN } \dot{x}(t) = A_1 x(t) + B_1 u(t), \end{aligned} \quad (2.30)$$

где M_{α}^i - нечеткий элемент правила i , соответствующий функции $f_{\alpha}(x(t))$, которая является функцией неопределенного параметра системы, $\alpha = 1, 2, \dots, \Psi$, $i = 1, 2, \dots, p$, Ψ - целое положительное; $A_i \in R^{n \times n}$ и $B_i \in R^{n \times m}$ - константы и входные матрицы, соответственно; $x(t) \in R^{n \times 1}$ - вектор состояний системы и $u(t) \in R^{m \times 1}$ - вектор управлений. Тогда система уравнений (2.1) будет иметь вид:

$$\dot{x}(t) = \sum_{i=1}^p w_i(x(t))(A_i x(t) + B_i u(t)), \quad (2.31)$$

$$\sum_{i=1}^p w_i(x(t)) = 1, w_i(x(t)) \in [0, 1], \text{ для всех } i \quad (2.32)$$

$$w_i(x(t)) = \frac{\mu_{M_1^i}(f_1(x(t))) \cdot \mu_{M_2^i}(f_2(x(t))) \cdot \dots \cdot \mu_{M_{\Psi}^i}(f_{\Psi}(x(t)))}{\sum_{k=1}^p (\mu_{M_1^k}(f_1(x(t))) \cdot \mu_{M_2^k}(f_2(x(t))) \cdot \dots \cdot \mu_{M_{\Psi}^k}(f_{\Psi}(x(t))))}, \quad (2.33)$$

и является нелинейной функцией от $x(t)$ и $\mu_{M_{\alpha}^i}(f_{\alpha}(x(t)))$, которое, в свою очередь, неизвестно и зависит от неопределенности параметра в системе. Нечеткая система управления реализуется набором правил:

$$\begin{aligned} \text{Правило } j : IF \ g_1(x(t)) \text{ is } N_1^j \text{ and } \dots \text{ and } g_{\Omega}(x(t)) \\ \text{is } N_{\Omega}^j \text{ THEN } u(t) = G_j x(t) + r, \end{aligned} \quad (2.34)$$

где N_{β}^j - нечеткий элемент правила j , соответствующий функции $g_{\beta}(x(t))$, содержащей неопределенный параметр нелинейной системы, $\beta = 1, 2, \dots, \Omega$, $j = 1, 2, \dots, c$, Ω - целое положительное; $G_j \in R^{m \times n}$ - выгода обратной связи правила j ; $r \in R^{n \times 1}$ - контрольный входной вектор.

Таким образом, аналитическая модель распознавания графического объекта, определяемая соотношениями (2.22) – (2.29), позволяет успешно формировать решения и информационно-управляющие воздействия, в системах, основанных на знаниях, на функционирование которых оказывают влияние факторы неопределенности.

Для того, чтобы определить эффективность применения разработанной модели необходимо формализовать информационную потребность пользователя.

2.3 Формализация информационной потребности пользователя с помощью коллокаций

Как было показано в главе 1, сегодня интернет содержит в себе огромные объёмы сведений, которые могут помочь как функционированию объектов экономики, так и деятельности отдельно взятого человека. Однако, невысокое качество результатов распознавания графических объектов в условиях неопределенности, не позволяет гарантировать достоверное принятие решения. Одной из причин является неправильная или неполная формулировка запросов к ИПС. Пользователь обычно формулирует запрос на естественном языке и зачастую не знает, какие слова лучше использовать в запросе, а также особенности конкретной поисковой системы и гораздо реже использует специализированный язык запросов ИПС. Поэтому на смену релевантности – соответствию результатов поиска запросу, как было показано в главе 1 приходит понятие пертинентность – соответствие полученной информации информационной потребности [28, 59, 75]. Однако, для поиска на основе пертинентности необходимо формализовать информационную потребность пользователя. Поэтому предложен подход к формализации информационной потребности на основе теории нечётких множеств. Основным понятием при этом подходе является, как показано в п. 1.3 понятие коллокации. Теория нечётких множеств позволила задать коллокацию и нечётким образом определить расстояние между термами в объекте. Кроме того теория нечётких множеств позволяет не просто выявлять принадлежность некой группы признаков объекта множеству коллокаций, но и ставить в соответствие таким объектам значение функции принадлежности множеству коллокаций [59, 68, 76]. То есть, считается, что произвольная группа признаков объекта принадлежит нечёткому подмножеству соответствующей коллокации, а степень принадлежности будет тем больше, чем ближе эти признаки.

Понятие коллокации определено как взаимное расположение признаков объекта на определённых расстояниях друг от друга (задаваемых

промежутками), которое присуще соответствующему классу объектов или их характерных участков.

Для построения модели введены следующие понятия и множества:

Пусть B – бинарное множество. $B = \{0,1\}$;

S – множество всех термов, в исследуемом представлении графического объекта: $S = \{s_1, s_2, \dots, s_n\}$, $|S| = n$.

Под термами понимаются значимые части только «важных» признаков объекта, которые служат для классификации [59]. В этом случае под элементом

$$s_{i_1}^{p_{i_1}} s_{i_2}^{p_{i_2}} s_{i_3}^{p_{i_3}} \dots s_{i_{l-1}}^{p_{i_{l-1}}} s_{i_l}^{p_{i_l}}, \quad (2.35)$$

где $s_{i_j} \in S$, $p_{i_j} \in B$, $1 \leq i_j \leq n$, $j = \overline{1, l}$, $l \in N$, будем называть коллокацией.

Замечание. Элемент вида (2.35) задаёт порядок термов в коллокации, а p определяет присутствие или отсутствие некоторого признака.

Рассмотрим в качестве примера коллокацию, получаемую при применении аналитической модели распознавания графического объекта (выражение 2.29), $s_{i_1}^{p_{i_1}} s_{i_2}^{p_{i_2}}$ где s_{i_1} = «превышение», $p_{i_1} = 1$, а s_{i_2} и p_{i_2} – «норма» и «0» соответственно. Этой коллокацией задаются участки графического объекта, в которых терм «норма» не встречается после терма «превышение», то есть, по сути, исключаем устойчивое сочетание признаков «превышение нормы». Если же $p_{i_1} = 0$, то мы наоборот задаём рассматриваемое словосочетание.

Однако такое представление коллокаций не даёт нам возможности судить о расстоянии между термами в ней, другими словами, мы не можем ответить на вопрос насколько близко терм «превышение» должен быть расположен к терму «норма». Для решения этого вопроса введена нечёткость в виде лингвистических переменных R_H и A_R .

Лингвистическая переменная R_H представлена кортежем:

$$R < r, Tr, N, Gr, Mr >, \quad (2.36)$$

где r = «расстояние между термами» - имя лингвистической переменной R_H ;

$Tr = \{\text{«большое»}, \text{«среднее»}, \text{«малое»}\}$ – терм-множество значений лингвистической переменной R_H ;

Gr – синтаксическое правило, порождающее названия переменной R представляет собой процедуру образования новых термов с помощью связок $Svr\{\text{«и»}, \text{«или»}, \text{«скорее чем»}\}$ и модификаторов $Mor\{\text{«не»}, \text{«очень»}, \text{«крайне»}, \text{«где-то»}, \text{«наверное»}\}$.

Пусть $svr \in Sv\{ \text{«скорее чем»} \}$, а t_1 и $t_2 \in Tr$, тогда синтаксическое правило построения имеет вид: $t_1 svr t_2$.

Например, пусть $t_1 = \text{«большое»}$, $t_2 = \text{«среднее»}$, а $svr = \text{«или»}$, тогда $t_1 svr t_2 = \text{«расстояние между термами большое или среднее»}$. Если же $svr = \text{«скорее чем»}$, то синтаксическое правило имеет вид $\text{«скорее» } t_1 \text{ «чем» } t_2$, что при тех же значениях t_1 и t_2 равно $\text{«расстояние между термами скорее большое, чем среднее»}$.

Рассмотрим теперь элемент произвольный элемент $mor \in Mor$. Семантическое правило, для произвольного терма $t \in Tr$ имеет вид: $mort$. Например, при $t = \text{«большое»}$, а $mor = \text{«не»}$, $mort = \text{«небольшое»}$.

Суперпозиция связок и модификаторов представляет собой последовательное их применение друг к другу. Например, к выражению, полученному в предыдущем примере, мы можем применить связку «скорее чем» по отношению к t_2 , получив при этом нечёткое высказывание $\text{«расстояние между термами скорее небольшое, чем среднее»}$.

Mr – семантическое правило, которое ставит в соответствие каждой нечёткой переменной её смысл, то есть для каждого нечёткого множества $Y \in Gr(Tr)$ определяет $\mu_Y : N \rightarrow [0; 1]$.

Для простоты обработки на ЭВМ будем использовать непрерывную аппроксимацию функции для задания μ_Y , вида $\mu_Y^A : Z \rightarrow [0; 1]$ такую, что $\forall i \in N$ $\mu_Y^A(i) = \mu_Y(i)$. Из-за того, что все преобразования Mr над непрерывными функциями эквивалентны преобразованиям над значениями этих функций в каждой точки их области определения, такая аппроксимация никак не отразится

на результате, вычислений, связанных с этими переменными. По сути, просто для того, чтобы не хранить таблицу значений функции μ_Y , мы храним μ_Y^A и с её помощью вычисляем значения μ_Y в целых точках области определения μ_Y^A .

Сформулируем семантическое, правило Mr , определив необходимые функции и преобразования, позволяющие построить μ_Y^A для $\forall Y \in Gr(Tr)$. Пусть μ_{Tr1}^A , μ_{Tr2}^A , μ_{Tr3}^A - функции принадлежности, формализующие термы «большое», «среднее», «маленькое» соответственно. Прежде чем задавать их, введём семантические правила, для связок и модификаторов. Связки «и», «или», а также модификатор «не», зададим архимедовыми нормой $T(., .)$, конормой $S(., .)$ и отрицанием $n(.)$. В качестве отрицания $n(.)$ возьмём классическое выражение:

$$n(\mu) = 1 - \mu. \quad (2.37)$$

Модификаторы «крайне», «очень», «где-то» и «наверное» зададим классически [12] с помощью возведения функции принадлежности, к которой они применены, в степени 4; 2; 0,5; и 0,25 соответственно. Будем считать связку «скорее t_1 , чем t_2 », где t_1 и $t_2 \in Tr$, эквивалентной выражению « t_1 или, наверное, t_2 », тогда $\mu_{\{\text{«скорее } t_1, \text{ чем } t_2\}} = S(\mu_{t_1}, \text{наверное}(\mu_{t_2}))$.

Для определения функций μ_{Tr1}^A , μ_{Tr2}^A , μ_{Tr3}^A пользователю предлагается сформулировать с помощью терм-множеств Tr и Gr относительно нескольких контрольных примеров. В результате этой операции, и благодаря тому, что все модификаторы аналитически обратимы, мы получим значения рассматриваемой тройки функций принадлежности в нескольких точках.

Осуществим линейную интерполяцию полученных точек для каждой из функций μ_{Tr1}^A , μ_{Tr2}^A , μ_{Tr3}^A . Будем искать сами функции, как полиномиальную аппроксимацию соответствующей линейной интерполяции, являющейся кусочно-гладкой и, взяв в качестве меры близости метрику $C^1_{[0; 1]}$ [59, 73, 75].

Для того чтобы определить подходящую степень для многочлена μ_{Tri}^A , $i = \overline{1,3}$, будем последовательно аппроксимировать наши точки многочленами всё большей и большей степени, до тех пор, пока ошибка аппроксимации не станет

больше для многочлена с большей степенью, выбор μ_{Tri}^A , алгоритм аппроксимации остановит на предыдущем многочлене.

Однако полученные полиномиальные аппроксимационные кривые ($Appr_{Tri}^A, i = \overline{1,3}$) не совпадают по области определения с μ_{Tri}^A . Определим μ_{Tri}^A следующим образом:

$$\mu_{Tri}^A = \begin{cases} 0, & Appr_{Tri}^A(x) \leq 0, \\ Appr_{Tri}^A(x), & 0 < Appr_{Tri}^A(x) < 1, \\ 1, & Appr_{Tri}^A(x) \geq 1; \end{cases} \quad (2.38)$$

Лингвистическая переменная A_R :

$$A_R = \langle a, Ta, H, Ga, Ma \rangle, \quad (2.39)$$

где a – «ожидаемая степень пертинентности коллокации»;

$Ta = \{\text{«положительное»}, \text{«нейтральное»}, \text{«отрицательное»}\}$;

Ga – синтаксическое правило, порождающее названия переменной a (Ga аналогично Gr , за тем лишь исключением, что терм «нейтральное» может быть использован только в связках и не сочетается с модификаторами);

Ma – семантическое правило, которое ставит в соответствие каждой нечёткой переменной её смысл, то есть для каждого нечёткого множества $E \in Ga(Ta)$ определяет число $\mu_E : H \rightarrow [0; 1]$.

При определении семантического правила для Ma будем руководствоваться классическим пониманием термина «нейтрально», которое предполагает, что $\mu_{\{\text{«нейтральное»}\}} = 0,5$. Будем считать, что $\mu_{\{\text{«положительное»}\}}$, равно как и $\mu_{\{\text{«отрицательное»}\}}$ равноудалены от границ области допустимых значений функции 1 и 0 соответственно и от $\mu_{\{\text{«нейтральное»}\}}$. То есть $\mu_{\{\text{«положительное»}\}} = 0,75$, а $\mu_{\{\text{«отрицательное»}\}} = 0,25$.

Синтаксические правила Ga аналогичны синтаксическим правилам Gr с учётом того, что не все синтаксические конструкции допустимые в Gr разрешены в Ga .

Пусть $h_l \in H_l$ – произвольный элемент вида (2.35) количество термов, в котором равно l . Поставим h_l в соответствие $l-1$ значений лингвистической переменной R , обозначим его вектором $R_l(R_1, R_2, \dots, R_{l-1})$, в том смысле, что μ_{R_i} задаёт расстояние между i -тым и $i+1$ -ым элементом коллокации h_l , $i = \overline{1, l-1}$.

С синтаксической точки зрения $\langle h_l, R_l \rangle$ представимо в виде высказывания: «признак (терм) $s_{i_2}^{p_{i_2}}$ следует за признаком (термом) $s_{i_1}^{p_{i_1}}$ на расстоянии μ_{R_1} и признак (терм) $s_{i_3}^{p_{i_3}}$ следует за ними на расстоянии μ_{R_2} и признак (терм) $s_{i_4}^{p_{i_4}}$ следует за ними на расстоянии μ_{R_3} и ... и признак (терм) $s_{i_l}^{p_{i_l}}$ следует за ними на расстоянии $\mu_{R_{l-1}}$ ». Тогда с точки зрения семантики функция принадлежности μ_R вычисляется как:

$$\mu_{h_l}(X) = \bigwedge_{i=\overline{1, l-1}} (\mu_{R_i}(x_i)), \quad (2.40)$$

где $x_i \in N$, $i = \overline{1, l-1}$, $X = \{x_1, x_2, \dots, x_{l-1}\}$.

Кортеж значений $q = \langle h_l, R_l, A_R \rangle$ определен как информационная потребность пользователя. Каждая такая тройка характеризует мнение пользователя о некоторой коллокации. Пользователь формулирует свою информационную потребность в виде нечётких высказываний «Если терм $s_{i_2}^{p_{i_2}}$ $r_{yx}^1 = R_1 s_{i_1}^{p_{i_1}}$ и терм $s_{i_3}^{p_{i_3}}$ $r_{yx}^{21} = R_2$ и терм $s_{i_4}^{p_{i_4}}$ $r_{yx}^3 = R_3$ и ... и терм $s_{i_l}^{p_{i_l}}$ $r_{yx}^{l-1} = R_{l-1}$, то $A_R = \langle K_i \rangle$ ».

На основе элемента $q = \langle h_l, R_l, A_R \rangle$ можем построить некоторую функцию принадлежности (μ_W) характеризующую позитивное или негативное влияние появления некоторой коллокации на ожидаемую пользователем пертинентность. Будем искать её как функционал:

$$W: (\mu_{h_l}(\cdot), \mu_{A_R}) \rightarrow [0; 1]. \quad (2.41)$$

Семантические требования к функционалу (2.41) удобно выразить в форме тождеств:

1. $W(\mu_{h_i}(X), 0) \equiv n(\mu_{h_i}(X))$, $\forall X \in N^{l-1}$ – так как близость μ_{A^*} к 0 означает крайне отрицательное влияние коллокации на ожидание пертинентности пользователя, что в свою очередь указывает на крайне положительное влияние отрицания функции принадлежности на пертинентность результатов запроса;

2. $W(\mu_{h_i}(X), 0,5) \equiv 0,5$, $\forall X \in N^{l-1}$ – так как близость μ_{A^*} к 0,5 означает нейтральное влияние коллокации на ожидание пертинентности пользователя;

3. $W(\mu_{h_i}(X), 1) = \mu_{h_i}(X)$, $\forall X \in N^{l-1}$ – так как близость μ_{A^*} к 1 означает крайне положительное влияние коллокации на ожидание пертинентности пользователя.

Зафиксируем $X \in N^{l-1}$. Тогда, если μ_{A^*} пробегает отрезок $[0, 1]$, $W(\mu_{h_i}(X), \mu_{A^*})$ пробегает отрезок $[n(\mu_{h_i}(X)), \mu_{h_i}(X)]$. Поскольку нет объективных причин утверждать, что скорость $W(\mu_{h_i}(X), \mu_{A^*})$ на различных отрезках $[0, 1]$ различна, то считается, что при фиксированном X , $W(\mu_{h_i}(X), \mu_{A^*})$ задаётся прямой на $\mu_{A^*} \in [0; 1]$. Тогда согласно уравнению прямой, проходящей через 2 точки [23], получим:

$$W(\mu_{h_i}(X), \mu_{A^*}) = \mu_{A^*} \cdot \mu_{h_i}(X) + (1 - \mu_{A^*}) \cdot n(\mu_{h_i}(X))$$

или, в силу (2.37)

$$W(\mu_{h_i}(X), \mu_{A^*}) = \mu_{A^*} \cdot \mu_{h_i}(X) + (1 - \mu_{A^*}) \cdot (1 - \mu_{h_i}(X)) \quad (2.42)$$

Правильность (2.42) подтверждает и тот факт, что данная прямая удовлетворяет второму требованию к функционалу (2.41).

Заменим в (2.42) $\mu_{h_i}(X)$ согласно (2.40), а $W(\mu_{h_i}(X), \mu_{A^*})$ на μ_W , которая определяется с помощью функционала W . Тогда (2.42) примет вид:

$$\mu_W = \mu_{A^*} \cdot \frac{T(\mu_{R_i}(x_i))}{i=1, l-1} + (1 - \mu_{A^*}) \cdot (1 - \frac{T(\mu_{R_i}(x_i))}{i=1, l-1}) \quad (2.43)$$

Введённые лингвистические переменные позволяют формализовать информационную потребность пользователя в виде совокупности значений $q = \langle h_l, R_l, A_R \rangle$. Функция принадлежности, вычисленная на основе (2.43),

формализует пертинентность распознавания по признаку соответствующему коллокации h_l .

Рассмотренный подход к аналитическому представлению информационной потребности пользователя с помощью теории нечётких множеств позволяет построить модель процесса распознавания графического объекта с оценкой пертинентности результата распознавания. Информационная потребность пользователя формализуется значениями лингвистических переменных, определяющих взаимное расположение, интересующих пользователя признаков графического объекта на основе нечёткого показателя соответствия признаков объекта одному из заданных классов, что соответствует удовлетворению требуемой информационной потребности пользователя.

2.4 Аналитическая модель процесса поиска сведений, результаты которого обладают требуемой пертинентностью

Введённые, в п. 2.3 лингвистические переменные, позволяют формализовать информационную потребность пользователя в виде совокупности значений $\langle h_l, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{l-1}, A_R \rangle$, где h_l – некоторая коллокация. Для определения пертинентности выше (2.40) получено выражение, которое в развернутой форме имеет вид:

$$\mu_{h_l}(x_1, x_2, \dots, x_{l-1}) = \mu_{A^*} \cdot \prod_{i=1, l-1} T(\mu_{R_i}(x_i)) + (1 - \mu_{A^*}) \cdot (1 - \prod_{i=1, l-1} T(\mu_{R_i}(x_i))), \quad (2.44)$$

где μ_{A^*} - степень принадлежности, формализующее значение лингвистической переменной A_R , а μ_{R_i} - функция принадлежности, определяющая семантику лингвистической переменной $\mathbf{R}_i$, которая в свою очередь формализует расстояние между признаками объекта, а соответственно, между i -тым и $(i+1)$ -ым термами коллокации h_l . Функция $T(.)$ – T -норма.

Функция принадлежности, вычисленная на основе (2.44), формализует пертинентность объекта по признаку, соответствующему коллокации h_l . Необходимо разработать правило нечёткого вывода, которое на основе этих

функций, фактически оценивающих пертинентность по конкретному признаку, позволит определить уровень пертинентности всего объекта. Использование степени пертинентности объекта в качестве его ранга в процессе формирования выдачи приведёт к существенному повышению качества принятия решения на основе распознавания графического объекта.

Поскольку решением задачи распознавания является некоторое множество $D^* \in U_D$, для которого существует q^* , такой, что $D^* = f(q^*)$, то для решения задачи необходимо ввести критерий эффективности ее решения. Как показано в главе 1 основой такого критерия являются значения полноты и точности.

С помощью полученных, при применении аналитической модели распознавания графических объектов признаков и коэффициентов, представленной в п. 2.2, рассчитываются следующие характеристики информационного поиска:

$$r = \frac{a}{a+c} \text{ - полнота поиска;} \quad (2.45)$$

$$p = \frac{a}{a+b} \text{ - точность поиска;} \quad (2.46)$$

$$acc = \frac{a+d}{a+b+c+d} \text{ - коэффициент аккуратности;}$$

$$err = \frac{b+c}{a+b+c+d} \text{ - коэффициент ошибки.}$$

Наиболее значимыми характеристиками являются полнота и точность.

Тогда, в качестве критерия эффективности информационного поиска следует принять двухпараметрическую свёртку полноты и точности, названную называется F -нормой [73, 75, 85]:

$$F = \frac{1}{\alpha \cdot \frac{1}{p} + (1-\alpha) \cdot \frac{1}{r}}, \quad (2.47)$$

где α – коэффициент, определяющий предпочтения тому или иному параметру $\alpha \in [0; 1]$, или

$$F = \frac{(1 + \beta^2)p \cdot r}{\beta^2 \cdot p + r}, \beta^2 = \frac{1 - \alpha}{\alpha} \quad (2.48)$$

В общем случае α определяется выражением:

$$\alpha(x) = \begin{cases} 0, x < a, \\ \left(\frac{x-a}{b-a}\right)^{\frac{\gamma}{1-\gamma}}, a \leq x \leq b, \\ 1, x > b; \end{cases} \quad (2.49)$$

где γ – заданная пользователем точность от 0 до 1 (полнота $1 - \gamma$); a – задаёт минимальное число результатов, при которых точность ещё важна, а b – максимальное количество результатов поиска при которых имеет значение полнота. В частном случае, когда можно принять $\alpha = 0.5$ или при $\beta = 1$, F -норма придаст одинаковый вес точности и полноте. Т.е. наблюдается явление «баланса», а F -норму назовем сбалансированной и обозначим как F_1 -норма (нижний индекс – величина β). Выражение для сбалансированной F -нормы имеет вид:

$$F_1 = \frac{2}{\frac{1}{p} + \frac{1}{r}} \quad (2.50)$$

В дальнейшем рассматривается сбалансированная F -норма.

Введём на множестве D – результатов распознавания нечёткое подмножество A – множество пертинентных признаков. Множество A задано отображением $\mu_A : D \rightarrow [0;1]$.

Введём понятие степень пертинентности. Это ожидаемая пользователем степень пертинентности распознавания объекта, удовлетворяющая запросу q . Будем обозначать её $\mu_A(q)$.

В процессе исследования считалось, что существует некоторый произвольный запрос $q' \in Q$ – множество всех запросов. Знак «`» означает удовлетворение условия релевантности. Результатами запроса q' будут только результаты, удовлетворяющие условиям, заданным в запросе, поэтому можно считать результаты данного запроса релевантными логическому высказыванию,

которое он собой представляет. Пусть $D' = f(q')$. Тогда получаем, значения, используемого в качестве мощности в теории нечётких множеств, кардинального числа [15, 17, 44]:

$$|D' \cap A| = \sum_{d \in D'} \mu_A(d) \quad (2.51)$$

$$|D' \cap A| = \sum_{d \in D \setminus D'} \mu_A(d) \quad (2.52)$$

На основе проведённых рассуждений, по аналогии с (2.45) и (2.46) представим полноту и точность поиска на основе пертинентности следующими формулами:

$$r = \frac{\sum_{d \in D'} \mu_A(d)}{\sum_{d \in D'} \mu_A(d) + \sum_{d \in D \setminus D'} \mu_A(d)}, \quad (2.53)$$

$$p = \frac{\sum_{d \in D'} \mu_A(d)}{|D'|} \quad (2.54)$$

Тогда в силу (2.50) F -норма примет вид:

$$F = \frac{2}{\frac{|D'|}{\sum_{d \in D'} \mu_A(d)} + \frac{\sum_{d \in D'} \mu_A(d) + \sum_{d \in D \setminus D'} \mu_A(d)}{\sum_{d \in D'} \mu_A(d)}}$$

или

$$F = \frac{2 \cdot \sum_{d \in D'} \mu_A(d)}{|D'| + \sum_{d \in D'} \mu_A(d) + \sum_{d \in D \setminus D'} \mu_A(d)} \quad (2.55)$$

По определению $\mu_A(q)$ есть ожидаемое пользователем значение степени принадлежности результата распознавания, удовлетворяющего q , к множеству пертинентных результатов. Если пользователь прав, то $\mu_A(q)$ есть среднее значение $\mu_A(d), d \in D'$. Тогда

$$\mu_A(q) = \frac{1}{|D'|} \sum_{d \in D'} \mu_A(d)$$

или

$$\sum_{d \in D'} \mu_A(d) = |D'| \cdot \mu_A(q). \quad (2.56)$$

Аналогично

$$\sum_{d \in D \setminus D'} \mu_A(d) = (N - M) \cdot \mu_A(\bar{q}). \quad (2.57)$$

Тогда (2.55) примет вид:

$$F = \frac{2 \cdot |D'| \cdot \mu_A(q)}{|D'| + |D'| \cdot \mu_A(q) + (N - |D'|) \cdot \mu_A(\bar{q})}$$

или, разделив числитель и знаменатель на $|D'|$, при условии, что он не равен нулю, получим:

$$F = \begin{cases} 0, |D'| = 0, \\ \frac{2 \cdot \mu_A(q)}{1 + \mu_A(q) + \left(\frac{N}{|D'|} - 1\right) \cdot \mu_A(\bar{q})}, |D'| > 0 \end{cases}$$

или

$$F = \begin{cases} 0, |D'| = 0, \\ \frac{2 \cdot \mu_A(q)}{1 + \mu_A(q) - \mu_A(\bar{q}) + \frac{N}{|D'|} \cdot \mu_A(\bar{q})}, |D'| > 0 \end{cases}. \quad (2.58)$$

Согласно (2.58) получили некоторую функцию F , зависящую от q , для которой верно: чем больше значение функции F , тем более предпочтительно использовать данный запрос, для удовлетворения информационной потребности пользователя. Тогда решение поисковой задачи можно выразить формулой:

$$D^* = f(\arg \max_{q \in Q} (F(q))) \quad (2.59)$$

Выражения (2.35)-(2.59) определяют аналитическую модель процесса распознавания графических объектов. Применение данной модели позволяет

определить требуемое значение пертинентности, найденного с помощью модели, описанной в п.2.2, результата распознавания графического объекта. Получаемые значения функции F дают возможность оценивать эффективность результата распознавания по данному запросу, т.е. определять требуемую пертинентность.

Для реализации разработанных аналитических моделей необходимо построить соответствующие процедурные модели и провести исследования по определению требований к необходимым информационным ресурсам и особенностям применяемых баз данных.

2.5 Выводы по главе 2

В результате исследования информационных процессов, происходящих в ИС при распознавании графических объектов, с применением построенных комбинированной и обобщенной моделей процессов подтверждено, что результатом работы правил логико-лингвистической модели является не терм, а алгебраическая зависимость.

Доказана необходимость введения лингвистических переменных для построения соответствующих моделей распознавания графических объектов в условиях неопределенности, обеспечивающих заданное качество их распознавания. Создана методологическая основа для построения аналитических моделей распознавания графических объектов и процесса поиска сведений, результаты которого обладают требуемой пертинентностью.

Построена аналитическая модель распознавания графических объектов на основе сплайнов основным отличием которой, от существующих, заключается в том, что для определения значений коэффициентов огибающих сплайнов, на выделенных временных интервалах, осуществляется построение системы уравнений решение которой формирует набор признаков, характеризующих графический объект, а отнесение признаков к одному из заданных классов осуществляется посредством выполнения определенного набора правил.

Построенная модель обеспечивает повышение качества распознавания графических объектов и ее целесообразно применять для формализации информационно-управляющих воздействий, т.е. для принятия решения по результатам распознавания графического объекта.

Показана возможность применения значений признаков, полученных при применении аналитической модели распознавания графических объектов в условиях неопределенности, для вычисления пертинентности. Обосновано введение лингвистических переменных для формализации информационной потребности пользователя в виде кортежа, элементы которого характеризуют мнение пользователя о некоторой коллокации.

На этой основе построена аналитическая модель процесса поиска сведений, результаты которого обладают требуемой пертинентностью, отличительной особенностью которой является введение параметров «полнота», «точность», обобщающей временной F -нормы, рассчитываемых на основе понятия «коллокация».

Исследование разработанных моделей показало, что для обеспечения их реализации в ИС распознавания графических объектов необходимо построить процедурную модель распознавания графических объектов, решить задачи распределения информационных ресурсов и построения структуры базы данных ИС в процессе распознавания графических объектов.

ГЛАВА 3. ПРОЦЕДУРНЫЕ МОДЕЛИ ДЛЯ ИНФОРМАЦИОННЫХ СИСТЕМ РАСПОЗНАВАНИЯ ГРАФИЧЕСКИХ ОБЪЕКТОВ

3.1 Процедурная модель решения задачи распознавания графических объектов

Эффективное решение задачи распознавания графических объектов, т.е. построение соответствующих моделей для применения в информационных системах необходимо, в конечном счете, для принятия решений на управление состоянием системы. Как показано в п. 2.1 процесс распознавания объекта рассматривается как процесс слежения за рассогласованием $x(t) - x_0(t)$ (выражение 2.8), поэтому для получения лучших результатов распознавания необходимо решать задачу оптимизации информационных процессов в системе относительно предписанного уровня ослабления (ρ), введенного в п. 2.1:

$$\min_P \rho^2, \quad (2.60)$$

с учетом ограничений: $\tilde{P} > 0$ и $\tilde{A}_{ij}^T \tilde{P} + \tilde{P} \tilde{A}_{ij} + \frac{1}{\rho^2} \tilde{P} \tilde{E}_i \tilde{E}_i^T \tilde{P} + \tilde{Q} < 0$,

где $\tilde{P}$ - симметричная положительно определенная матрица весов;

A_{ij}, E_i – определены выражением (2.17).

Как доказано в [65] система, описываемая уравнением (2.20) устойчива, если

$$V(t) = \tilde{x}^T(t) \tilde{P} \tilde{x}(t). \quad (2.61)$$

В результате дифференцирования (2.61) и с учетом ограничений, наложенных в (2.60), получим [65]

$$\dot{V}(t) = \sum_{i=1}^L \sum_{j=1}^L h_i(z(t)) h_j(z(t)) \tilde{x}^T(t) [\tilde{A}_{ij}^T \tilde{P} + \tilde{P} \tilde{A}_{ij}] \tilde{x}(t) < 0. \quad (2.62)$$

Получение общего решения $\tilde{P} = \tilde{P}^T > 0$ оптимизационной задачи (2.60) аналитическим путем является сложной проблемой. Однако, при представлении (2.60) как задачу минимизации системы линейных матричных неравенств (СЛМН), вариант решения которой представлен в [31, 65, 66], нетрудно

построить процедурную модель вычисления значений параметров, удовлетворяющих неравенству (2.62), подробно описана в [65,66].

Основное содержание процедурной модели решения задачи распознавания графических объектов, в интересах принятия решений на управление состоянием системы представлено этапами:

1. Выбор функций принадлежности и формирование нечетких правил в (2.1);
2. Задание начального значения ρ^2 .
3. Решение СЛМН для расчета значений параметров, удовлетворяющих неравенству (2.62).
4. Вычисление параметров для решения (2.9);
5. Уменьшение ρ^2 и повторение шагов 3-5 до тех пор, пока не будут найдены значения параметров, удовлетворяющие выполнению условия (2.60).
6. Построение нечеткой системы слежения (2.11).
7. Построение нечеткой системы управления (2.15).

Таким образом, подход, позволяющий осуществить переход от комбинированных математических моделей информационных процессов к обобщенным моделям, ставить и решать задачу управления и осуществлять оценку устойчивости, реализован.

3.2 Процедурная модель распознавания графических объектов на основе сплайнов

В п. 2.2 представлена аналитическая модель распознавания графических объектов на основе сплайнов. Для облегчения процесса ее реализации построена процедурная модель распознавания графических объектов на основе сплайнов (рисунок 3.1), которая реализуется в процессе выполнения следующих действий [54, 73, 85]:

1. Получение экспериментальных кривых $r_{xx}^{\mathfrak{z}}(t)$, $r_{yx}^{\mathfrak{z}}(t)$, и определение кубических сплайнов $S_{r_{yx}}^{\mathfrak{z}}(t)$, интерполирующих функцию $r_{yx}^{\mathfrak{z}}(t)$.
2. Определение параметров функций $r_{xx}^{\mathfrak{z}\Pi,1}(t) = Ae^{-\alpha|t|}$, $r_{xx}^{\mathfrak{z}\Pi,2}(t) = Ae^{-\alpha|t|} \cos \omega t$, $r_{xx}^{\mathfrak{z}\Pi,3}(t) = (A + A_1 |t|)e^{-\alpha|t|}$, при которых значение функции погрешности
$$e(t) = \sum_{t=t_0}^T (r_{xx}^{\mathfrak{z}}(t) - r_{xx}^{\mathfrak{z}\Pi,j}(t))^2$$
 минимально.
3. Соотнесение вида $r_{xx}^{\mathfrak{z}}(t)$ с одним из $r_{xx}^{\mathfrak{z}\Pi}(t)$. Выбирается такой j -ый вид функции $r_{xx}^{\mathfrak{z}\Pi,j}(t)$, при котором достигается минимум функции погрешности $r_{xx}^{\mathfrak{z}}(t)$ от $r_{xx}^{\mathfrak{z}\Pi,j}(t)$.

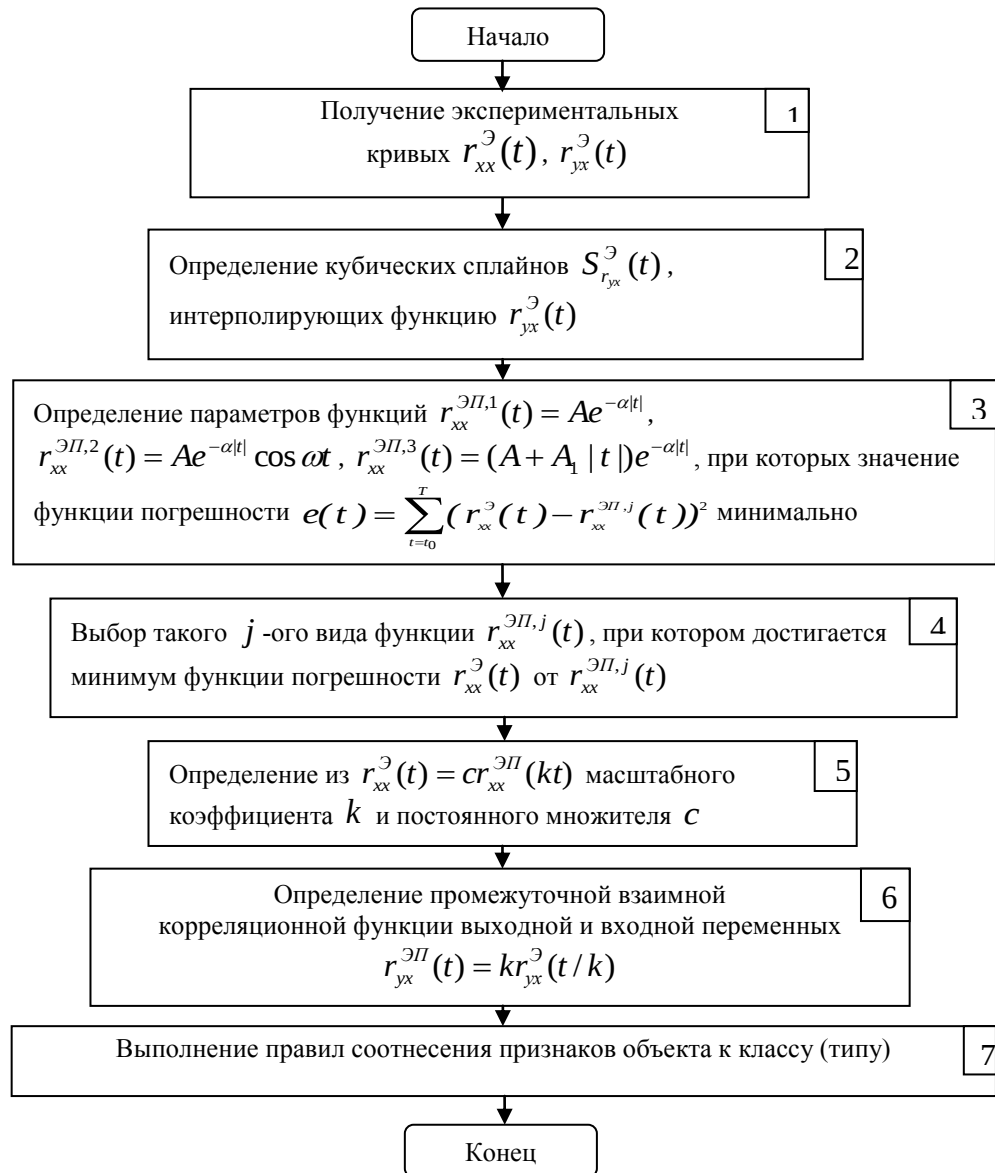


Рисунок 3.1 – Блок-схема алгоритма процедурной модели

4. Определение из $r_{xx}^{\mathfrak{O}}(t) = cr_{xx}^{\mathfrak{OII}}(kt)$ масштабного коэффициента k и постоянного множителя c .

5. Определение промежуточной взаимной корреляционной функции выходной и входной переменных $r_{yx}^{\mathfrak{OII}}(t) = kr_{yx}^{\mathfrak{O}}(t/k)$.

6. Определение вида $r_{yx}^T(t)$, близкой по форме к $r_{yx}^{\mathfrak{OII}}(t)$ в аналогично п.3, и соответствующих коэффициентов дифференциального уравнения объекта

$$r_{yx}'''(t) + a_1 r_{yx}''(t) + a_2 r_{yx}'(t) + a_3 r_{yx}(t) = b_0 r_{xx}''(t) + b_1 r_{xx}'(t) + b_2 r_{xx}(t).$$

7. Выполнение правил соотнесения признаков объекта к классу (типу) обеспечивающих требуемую пертинентность результатов распознавания.

Применение процедурной модели позволяет провести распознавание объекта, т.е. определить его динамические характеристики объекта и построить его математическую модель.

Для получения данных, для реализации первого этапа процедурной модели распознавания графических объектов целесообразно выполнение процедуры, описанной в [43, 85, 120].

Суть ее заключается в следующем:

Для решения задачи распознавания графиков функций в информационной системе разработана специальная процедура. Рассмотрим ее этапы.

Этап 1. Дискретизация непрерывного представления графического объекта.

Дискретизация непрерывного представления графического объекта является обязательной операцией перед их обработкой. В ходе нее обычно используется представление графического объекта в памяти в виде матрицы пикселей.

Замену непрерывного представления графического объекта дискретным можно выполнить различными способами. Однако наиболее употребительным является периодическая дискретизация, в частности, дискретизация с прямоугольным растром. Если $x_H(t_1, t_2)$ - непрерывное представление

графического объекта, а $x(i_1, i_2)$ - соответствующее ему дискретное, полученное из непрерывного путем прямоугольной дискретизации, то .

$$x(t_1, t_2) = \sum_{i_1} \sum_{i_2} x(i_1, i_2) \frac{\sin \left[\frac{\pi(t_1 - i_1 \Delta t_1)}{\Delta t_1} \right]}{\frac{\pi(t_1 - i_1 \Delta t_1)}{\Delta t_1}} \frac{\sin \left[\frac{\pi(t_2 - i_2 \Delta t_2)}{\Delta t_2} \right]}{\frac{\pi(t_2 - i_2 \Delta t_2)}{\Delta t_2}}.$$

Этап 2. Снижение шума.

В практике цифровой обработки объектов распознавания широко используется масочная фильтрация, основанная на применении множества весовых коэффициентов, заданных во всех точках окрестности текущей точки объекта. Это обеспечивает выполнение условия сохранения средней яркости точек объекта, а, следовательно, выходной сигнал будет соответствовать диапазону входного сигнала.

Этап 3. Бинаризация представления объекта распознавания.

Операция порогового разделения (бинаризация) заключается в сопоставлении значения яркости каждого пикселя графического объекта с заданным значением порога. Выбор соответствующего значения пороговой величины дает возможность выделения на объекте областей определенного вида. Основная цель состоит в выборе такого порога, который минимизирует среднюю ошибку от принятия решения о принадлежности данного пикселя объекту или фону.

Альтернативный вариант уменьшения постороннего шума при бинаризации дает метод Отсу [117, 120].

Шаг 1. Автоматическое определение порога бинаризации по нормированной гистограмме яркостей объекта по формуле:

$$p_i = n_i / N,$$

где N — общее число пикселей на объекте; n_i — число пикселей с уровнем яркости i .

Шаг 2. Процедура простого порогового разбиения представления объекта.

Диапазон яркостей делится на два класса с помощью порогового значения уровня яркости.

Этап 4. Бинарная морфология.

Ошибки классификации «чёрных» и «белых» пикселей устраняются последовательным применением четырёх операций бинарной морфологии: расширение, эрозия, открытие и замыкание.

Этап 5. Распознавание сложных геометрических объектов.

Алгебраические кривые хорошо зарекомендовали себя в качестве инструмента представления сложных объектов. Отсутствие требований к порядку точек, описывающих границы объекта, а также возможность получения геометрических инвариантов позволяют использовать алгебраические кривые в задачах представления и распознавания образов.

3.3 Моделирование процесса распределения ресурсов в информационных системах распознавания графических объектов

Качество разработки ИС напрямую влияет на эффективность и безопасность ее функционирования. В связи с этим вопросы моделирования распределения ресурсов в ИС являются актуальной задачей при разработке моделей распознавания графических объектов для нее.

Анализ научных публикаций и результатов исследований, посвященных проблематике решения задач выбора и распределения ресурсов (ЗВРР) в сложных системах [5, 32, 57, 60, 70], позволяет говорить о двух уровнях решения данной задачи. Это является следствием того, что в процессе ВРР в ИС параметрического синтеза системы ВРР, т.е. решения задачи определения ресурсов ИС на ее каждый элемент, необходимо проводить анализ функционирования данной ИС, что связано с осуществлением контроля для заданного процесса, реализуемого всегда во времени. Анализ функционирования системы и ее синтез представляют собой различные процессы, хотя и взаимосвязанные между собой. Синтез СВРР основан на параметрах, описывающих технические или технологические характеристики системы, а следовательно частные задачи, которые соответствуют этих процессам, разделяются на статические и динамические [5]. Применительно к

ИС означает это означает, что ЗВР технических ресурсов является статической задачей, а информационных – динамической.

На рисунке 3.2 представлена специфика и методы ВРР, которые обеспечивают функционирование ИС и отражены классы моделей обобщенной ЗВРР [5].

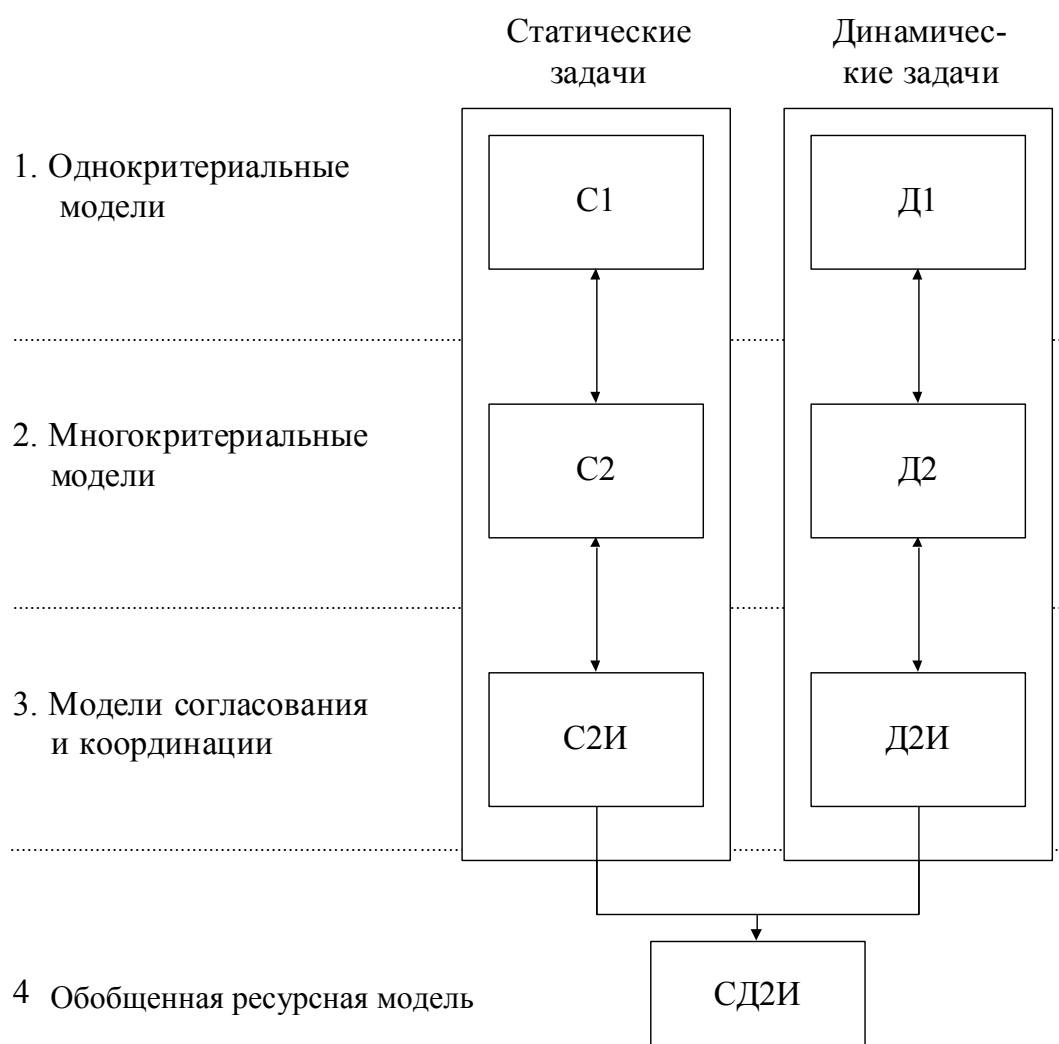


Рисунок 3.2 - Классификация моделей решения ЗВРР

На рисунке 3.2 модели решения статических и динамических задач с одним критерием представлены блоками «С1» и «Д1»; модели с многими критериями – блоками «С2» и «Д2»; эти же модели, при наличии условия применения процедуры согласования и координации интересов различных уровней, реализуемых, когда невозможно сформировать области допустимых решений задачи, обозначены блоками «С2И» и «Д2И»; обобщенная ресурсная

модель (ОРМ), которая учитывает особенности решения всех выше приведенных частных моделей, обозначена на рисунке как блок «СД2И».

Представление в таком виде ЗВРР позволяет перейти к разработке ОРМ, которая включает процедуры решения ЗВРР, применяемые на рассмотренных уровнях обеспечения функционирования ИС, которые учитывают использование нескольких критериев в соответствующих частях модели и позволяют реализацию процедур согласования и координации различных уровней, которые участвуют в ресурсном процессе.

Процесс реализации ЗВРР во время синтеза ИС заключен в том, что каждому элементу синтезируемой ИС назначаются ресурсы, определенных видов и объемов.

Рассмотрим только те методы математического моделирования, в которых используются знания о целях системы.

Решение задач анализа функционирования системы, как правило, затруднено при реализации моделей оптимизации, поскольку применяемые в ней критерии и ограничения чаще всего заданы не аналитические, т.е. вербально, моделирующими алгоритмами, графиками и т.д. Для решения таких задач используется метод имитационного моделирования.

Описанный вид задач является самостоятельным этапом решения ЗВРР, включающий в себя не только процедуры имитационного моделирования, но и иные процедуры поиска оптимальных динамических параметров ИС.

Последовательность событий, происходящих в ИС, при моделировании ИС представляет динамику моделей этих ИС. Следует отметить о различии типов моделей, которые описывают динамику функционирования объекта исследования [2, 5, 57, 60]. К первому типу следует отнести модели, которые носят событийный характер. При этом, схема моделирования, соответствующая данному типу, отмечает моменты начала и окончания событий, моделируемых на интересующем интервале времени. Второй тип моделей, соответствует таким моделям, при которых рассматриваемый интервал времени делится на заданные интервалы, каждый из которых рассматривается без учета

обязательности наличия изучаемых событий. Результатом проводимого анализа является выделение модели систем, которые, используя пошаговый или событийный способ, позволяют провести имитацию функционирования объектов, что может быть представлено с использованием агрегативных систем, автоматных моделей, сети Петри или модели системной динамики, модели теории расписаний, а для стохастических систем - модели массового обслуживания. Все представленные модели позволяют применять различные процедуры, которые позволяют решать задачи имитационного моделирования.

Отметим, что чаще всего процесс оптимизации характеристик и параметров ИС заключен в поиске оптимального расписания функционирования синтезируемой ИС. При этом расписание является совокупностью моментов времени начала выполнения операций ИС и их окончания. Последовательность всех операций, выполняемых в ИС, и время их выполнения являются исходными данными для процесса поиска оптимального расписания. При решении такой задачи осуществляется поиск параметров, которые определяют показатель эффективности функционирования синтезируемой ИС. В случае использования в данной задаче большого количества варьируемых переменных невозможно применить точные алгоритмы решения, поскольку данная задача относится к NP – полным задачам.

Основываясь на вышесказанном, построим обобщенную модель системы выбора и распределения ресурсов. Считаем, что ИС есть множество $Q = \{Q_1, \dots, Q_m\}$. Под целью W будем понимать множество $W(X)$, $W = \{W(X) : W(X) = (W_1(X), \dots, W_s(X)); W(X) \subset W\}$ состояний $X(t) = \{X_1(t), \dots, X_m(t)\}$ системы Q , $(X(t) \in W(X))$, где t - время. Цель исследования задает необходимость создания и развития системы. Для описания множества $W(X)$ необходимо сформировать ограничения, которые в данных условиях являются предпочтительными для ЛПР, на вектор $X(t)$ и вектор оценки свойств цели $E(W)$. Вектор $E(W)$ оценивается как с использованием свойств множеств, так характеристик цели. Считаем, что цель $W(X)$ является

количественно измеримой на множестве состояний $Q(X)$, при условии задания на $Q(X)$ вещественной функции полезности $\Gamma_X(E(X)) - \Gamma_X(E)$ такая, что если $X_i(t) \stackrel{W(X)}{\succ} X_j(t)$, то $\Gamma_X(E(X_i(t))) < \Gamma_X(E(X_j(t)))$, символ $\stackrel{W(X)}{\succ}$ означает «лучше в смысле $W(X)$ ». Вектор $X(t)$ оценивается эффективностью $E(W)$. Так как для достижения целевых состояний $X(t) \in W(X)$ можно использовать различные стратегии, то существует, в общем случае, такое множество стратегий $\Omega(Y)$, где $Y(t) = \{Y_1(t), \dots, Y_k(t)\}$, $Y(t) \in \Omega(Y)$, которое реализует цель $W(X)$ [5].

Граф $G_{TC} = (E, Z)$, являющийся связанным графом, в котором выполняется условие $E = \bigvee_{e_j \in E} Z(e_j)$, $E = \{e_j, j \in J\}$, $J = \{1, \dots, M\}$, $M = |E|$ – множество вершин графа G_{TC} ; Z – определяет соответствие между парами вершин G_{TC} , $\bar{Z}(e_j) = \{e_j\} \cup Z(e_j) \cup Z[Z(e_j)] \cup \dots$ – транзитивное замыкание.

Отметим принципиальный характер наличия такого свойства графа G_{TC} , как связанность. Оно характеризует структуру ИС в не зависимости от ее целей и назначения, поскольку связанный граф всегда обладает путем из вершины $e_i \in E$ в вершину $e_j \in E$.

Если $Q_j, Q_i \in Q, i, j = \overline{1, M}, i \neq j$, связаны между собой благодаря использованию общего ресурса $d \in D$, D – множество ресурсов, которые поступают на входы элементов ИС, то $Q_j = H^{ij} Q_i$, H^{ij} – нуль-единичная матрица связи Q_j с Q_i через использование общего ресурса. Процедурная ресурсная модель функционирования ИС представлена на рисунке 3.3.

В схеме $P_Q = \{p_m\}$ – множество результатов работы ИС.

ЗВРР ИС (задача Π) состоит из двух задачах: B – задача выбора из множества ресурсов, которые поступают на входы элементов ИСКВ, вектора такого ресурса, который может быть использован, и P – задача распределения вектора ресурсов выбранного среди элементов ИС, который обеспечивает требуемое функционирование.

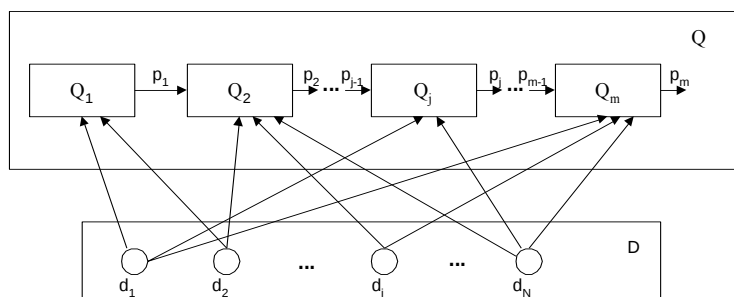


Рисунок 3.3 – Блок-схема функционирования процедурной ресурсной модели ИС

Представленная выше постановка задачи выбора и распределения ресурсов в информационной системе распознавания графических объектов соответствует, основным принципам системного подхода: единства функционально-целевых и причинно-следственных отношений на всех этапах процесса распознавания графического объекта и управления им.

В силу, как показано в гл. 1, наличия различных систем распознавания, в различных предметных областях определяемых спецификой применения ИС, целесообразно исследовать различные модели выбора и распределения ее ресурсов, содержание механизмов ресурсного взаимодействия ее элементов. Для этого, в рамках общей постановки задачи выбора и распределения ресурсов выделено множество самостоятельных ресурсных действий, определяемых условиями ресурсного взаимодействия между элементами системы (рисунок 3.4).

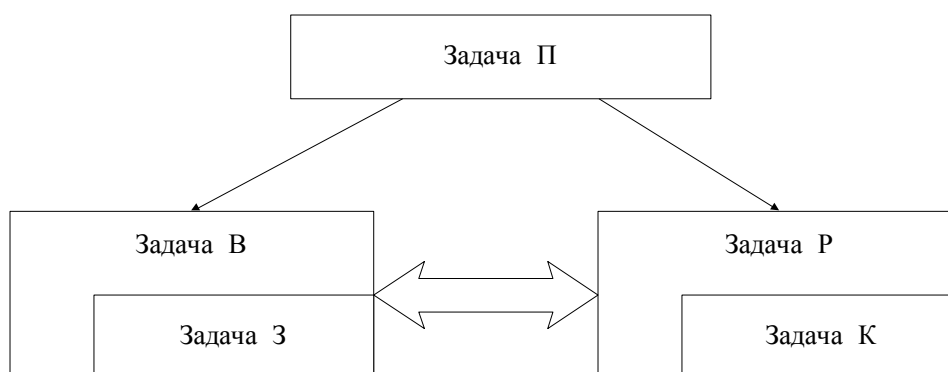


Рисунок 3.4 – Структура представления ЗВВР

При решении задачи выбора ресурсов могут возникнуть актуальные вопросы, связанные с проблемой замещения ресурсов (задача *З*). При решении задачи распределения ресурсов может возникнуть ситуация, когда некоторые элементы ИСКВ в процессе достижения своих локальных целей, взаимодействуя друг с другом и с внешней средой, вступают между собой в конфликт из-за общего ресурса (задача *К*).

Важной задачей при создании любой сложной системы, в том числе и системы выбора и распределения ресурсов (СВРР), является этап формирования ее структуры. Кроме того, проблема выбора и распределения ресурсов требует анализа состояния ИС, в которое она пришла при использовании поступающего на входы ее элементов ресурса. На сегодняшний день не учитывается, что при решении задачи *П*, в дальнейшем, в структуру СВРР возможно придется вносить некоторые, заранее не определенные изменения, связанные с решением дополнительных задач, например, с возможным замещением одного ресурса другим, возникновением ресурсного конфликта и др.

Подобные проблемы заставляют решать задачу синтеза структуры СВРР в условиях неопределенности. Возникшее противоречие между необходимостью учета условий неопределенности и невозможностью их задания в явном виде в определенной мере можно устранить, если воспользоваться подходом, описанным в [5] в котором разработчики сложной системы ориентируются не на синтез готовой структуры, а на создание некоторой обобщенной модели будущей структуры. Такая модель должна включать описание инвариантных характеристик ИС, а также механизмы генерации конкретной структуры (например, выбор и распределение ресурсов ИС в условиях замещения, ресурсного конфликта и т.п.), использование которых позволит настроить модель распределения ресурсов ИС распознавания графических объектов.

Пусть система выбора и распределения ресурсов ИС предназначена для решения ресурсной задачи *П*, которую будем именовать основной. Помимо

этого для системы с целью создания совокупности механизмов MX , обеспечивающих функционирование ИС, может быть сформулирована и дополнительная задача $П_d$ (например, выбор и распределение ресурсов в условиях возможного замещения одного вида ресурса другим или в условиях возможного ресурсного конфликта и др.). Тогда систему представим в следующем виде [5]:

$$S = \langle П, П_d, ИСКВ, MX \rangle. \quad (3.1)$$

Указанные механизмы MX могут включать процедуры порождения MX_n , развития MX_p и изменения MX_u как системы в целом, так и отдельных ее элементов, т.е.

$$MX = \langle MX_n, MX_p, MX_u \rangle \quad (3.2)$$

Главная цель создания таких механизмов состоит в том, чтобы обеспечить выполнение разнообразных переходов от одних качественных состояний ИС к другим. Все это позволяет заложить в систему определенные свойства гибкости и адаптируемости. Для описания подобных переходов рассмотрим некоторые возможные состояния ИС: $ИС^\wedge$ - до - ИС, т.е. ИС, из которой порождается ИС; $ИС^\vee$ - после - ИС, т.е. ИС, образующаяся после изменения элемента ИС; $ИС_{t1}$ - ИС, реализующая задачу $П$ на время t_1 ; $ИС_{t2}$ - ИС, реализующая задачу $П$ на время t_2 , тогда процедуры MX_n , MX_p , MX_u будут соответствовать следующим отображениям:

$$MX_n : ИС^\wedge \rightarrow ИС ; MX_p : ИС_{t1} \rightarrow ИС_{t2} ; MX_u : ИС \rightarrow ИС^\vee \quad (3.3)$$

Выделим также случай, когда MX оказывают воздействие не на всю ИС, а на некоторую его часть. Обозначим: MX' - совокупность механизмов из MX , воздействующих на часть ИС, например, механизмы замещения одного вида ресурса другим; $ИС^{+M}$, $ИС^{-M}$ - наборы элементов системы, на которые соответственно распространяется и не распространяется действие механизмов MX' , при этом $ИС = ИС^{+M} \cup ИС^{-M}$ и $ИС = ИС^{+M} \cap ИС^{-M} = \emptyset$.

Используя введенные выше описания для возможных состояний ИС, а также механизмов MX , можно определенным образом классифицировать варианты системы [5]:

$$S_1 = \langle \Pi, ИСКВ, MX \rangle, \quad (3.4)$$

$$S_2 = \langle \Pi, \Pi_\partial, ИС, MX \rangle, \quad (3.5)$$

$$S_3 = \langle \Pi, \Pi_\partial, ИС, MX' \rangle. \quad (3.6)$$

Если механизмы MX' и MX при этом включают процедуры порождения, то для систем классов (3.5), (3.6) выделяются исходные (или порождающие) системы:

$$S_2 = \langle \Pi, \Pi_\partial, ИС^\wedge, MX \rangle, \quad (3.7)$$

$$S_3 = \langle \Pi, \Pi_\partial, ИС^\wedge, MX' \rangle. \quad (3.8)$$

Системы классов (3.7), (3.8) могут рассматриваться как обобщенные модели функционирующих систем, т.е. систем классов (3.5), (3.6). В этом случае при их разработке необходимо создание таких механизмов MX' и MX , которые обеспечат последовательный переход от $ИС^\wedge$ к $ИС$. При этом для реализации механизмов типа MX_n необходимо установить четкое соответствие между описаниями $ИС^\wedge$ и средствами представления ИС.

Основная ресурсная задача Π применительно к настоящему исследованию включает задачи выбора и распределения, $\Pi = \{B, P\}$, а дополнительная задача Π_D может состоять из задач выбора и распределения ресурсов ИС в условиях замещения и ресурсного конфликта, т. е. $\Pi_D = \{Z, K\}$. Для реализации задачи Π в целом необходимо, чтобы система выбора и распределения ресурсов была 2-х уровневая, т. е. $S = \{S_\partial, S_p\}$, где S_∂, S_p – системы соответственно выбора и распределения ресурсов: $S_\partial = \{S_{\partial m}\}$, $S_p = \{S_{pk}\}$. Тогда, если принять, что все три типа систем (S, S_∂, S_p) принадлежат к рассмотренным классам (3.7), (3.8), то в качестве обобщенной модели синтезированной системы выбора и распределения ресурсов может выступать кортеж:

$$S = \langle \{B, P\}, \{Z, K\}, \{S_\partial, S_p\}, ИС, MX \rangle. \quad (3.9)$$

Исключительное многообразие практических ситуаций, определяемых спецификой ИС (задача II), требует рассмотрения различных моделей выбора и распределения ресурсов в условиях возможного замещения и ресурсного конфликта, которые должны строиться только на основе учета содержательных механизмов ресурсного взаимодействия элементов ИС.

Рассматривая эти модели с системных позиций можно сделать вывод о том, что, роль решения задачи выбора и распределения ресурсов ИС будет определяться необходимостью выделения для ИС некоторых видов и объемов ресурсов, которые потребуются ей для достижения поставленных перед ней целей. Функции же этих задач будут связаны с таким формированием матрицы ресурсов d , чтобы при этом выполнялись все ограничения рассматриваемой задачи $f_k(d)$, $k = \overline{1, K}$, а также, чтобы полученное решение ресурсной задачи было оптимальным.

Так как рассматриваются многокритериальные задачи выбора и распределения ресурсов ИС, в качестве исходной будем использовать модель [5]:

$$\begin{aligned} Q(d) = (q_1(x), \dots, q_s(x)) &\xrightarrow{d \in \Phi} Opt, \\ \Phi: f_k(d) &\leq 0, k = \overline{1, K}, \\ d = \|d_{ij}\| &\left(i = \overline{1, n}; j = \overline{1, m} \right). \end{aligned} \quad (3.10)$$

где Opt - оператор векторной оптимизации;

$Q(d)$ - вектор частных критериев качества q_μ ($\mu = \overline{1, s}$) ресурсного распределения;

Φ - область допустимых решений ресурсной задачи. При этом $d' \in \Phi$ - условие парето-оптимального решения.

Совокупность всех возможных решений ресурсной задачи d' образует множество Парето, которое и является формальным решением рассматриваемой задачи. Для выбора же наиболее предпочтительного решения d'' необходимо получение дополнительной информации.

При рассмотрении модели функционирования ИС (рисунок 3.3) функции полезности Q_j и Q_i имеет вид [5]:

$$\Gamma_j(Q_j) = \Gamma_j[d_j, R_j(d_j, c_j)], \quad (3.12)$$

$$\Gamma_i(Q_i) = \Gamma_i[d_i, R_i(d_i, c_i)], \quad (3.13)$$

где $d_j = (d_{j1}, d_{j2}, \dots, d_{jN})$, $d_i = (d_{i1}, d_{i2}, \dots, d_{iN})$ – векторы ресурсов из D , поступающих на вход Q_j , Q_i соответственно; $|D| = N$ – мощность множества D ; R_j, R_i – глобальная реакция Q_j , Q_i ; $C_j = \{c_j\}$, $C_i = \{c_i\}$ – множество глобальных состояний Q_j , Q_i соответственно.

Тогда связь между d_j и d_i подсистем Q_j , $Q_i \in Q$ также можно записать через матрицу использования ресурсов:

$$d_j(Q_j) = H^{ij} d_i(Q_i). \quad (3.14)$$

Обозначим через $T_{ij} = H^{ij} d_i(Q_i)$. Тогда (3.14) преобразуется к виду

$$d_j(Q_j) = T_{ij} [d_i(Q_i)]. \quad (3.15)$$

Соотношение (3.15) позволяет отслеживать изменения ресурсов на входе Q_i для всех $d_i \in D$, $i = \overline{1, N}$. В общем случае формирования новых реализаций ресурсов на входе $\forall Q_j \in Q$ можно представить в виде $d = T(d)$, где $T = [T_{ij}]$ – матрица размером $N \times N$ и $T = H^{ij} d_i$, $i, j = \overline{1, N}$.

Изучая теперь каждую из подсистем Q_j в морфологическом плане (используя принцип декомпозиции) с той или иной позиции и применяя к ней изложенную выше структурную формализацию, можно построить механизмы распределения ресурсов $\Gamma(d)$ как между элементами Q , так и с внешней средой.

Используя предложенный выше подход к формулированию модели выбора и распределению ресурсов ИС, можно в общем виде определить последовательность ее реализации. Она включает следующие три основных уровня: решение ресурсной задачи на этапе синтеза ИС, решение ресурсной задачи на этапе функционирования ИС, а также получение окончательно согласованного решения общей ресурсной задачи Π в условиях замещения, векторной оценки, критериев, стратегий и ресурсов. Третий этап здесь

требуется для того, чтобы эксперт или программа принятия решения (ППР) могли накапливать все полученные ранее результаты решения частных задач и тем самым формировать общее множество решений $\{d, z\}$.

В силу того, что в общей модели должна быть предусмотрена возможность возврата на более ранний этап общего процесса решения ресурсной задачи, процедурную модель решения ресурсной задачи для ИС отразить в виде схемы (рисунок 3.5).

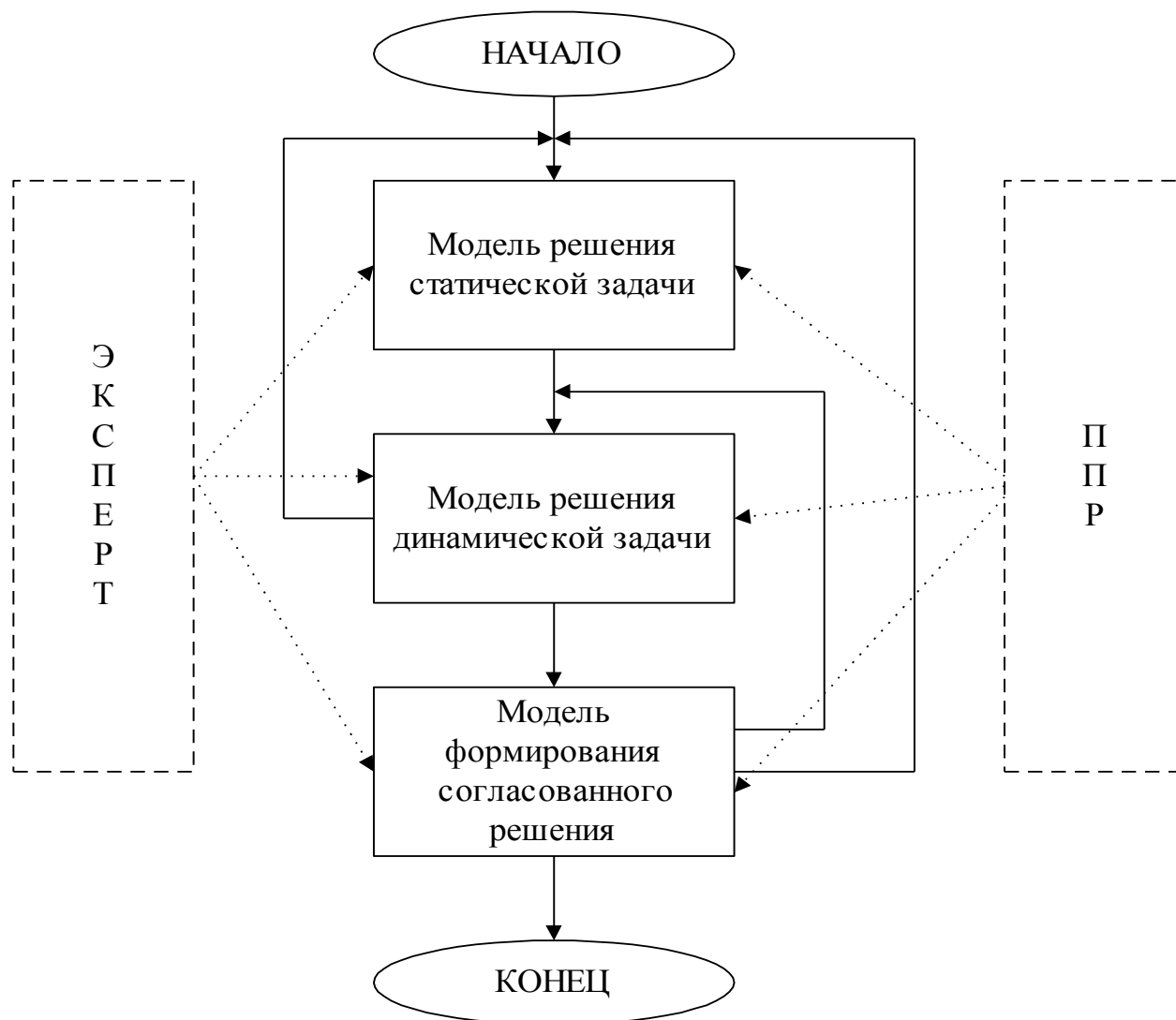


Рисунок 3.5 - Схема решения общей ресурсной задачи

Взаимодействие статической и динамической моделей в рамках общего процесса может осуществляться различными способами, что во многом определяется спецификой рассматриваемых задач, степенью их формализованности, размерностью и другими факторами.

3.4 Решение задачи построения структур баз данных

Поскольку реализация разработанных моделей распознавания графических объектов требует, наряду с другими условиями, выполнения условия эффективного обмена данными между собой, проведено исследование возможных вариантов решения задачи построения структур баз данных для обеспечения эффективности функционирования ИС и связанных с ней приложений. Основой для решения этой задачи, явились методы и способы построения логических структур баз данных, согласно которым этот процесс рассматривается как процесс поиска оптимальных вариантов отображения канонических структур в такие логические структуры, которые обеспечивают оптимальное значение заданного критерия эффективности их разработки или использования и удовлетворяют основным требованиям и ограничениям, накладываемым на логическую структуру [78, 108, 119, 138].

Исходным, для решения задачи построения структуры базы данных (БД), является выражение:

$$G(C,U) \xrightarrow{\theta^*} G(N,L) \quad (3.16)$$

где $G(C,U)$ – граф канонической структуры БД ($C = \{c_i / i = \overline{1, I}\}$ – множество информационных элементов ИС; $U = \{(c_i, c_{i'}) / i, i' = \overline{1, I}\}$ – множество взаимосвязей между ее элементами);

$G(N,L)$ – граф логической структуры БД ($N = \{n_j / j = \overline{1, J}\}$ – множество логических записей, $L = \{(n_i, n_{j'}) / j, j' = \overline{1, J}\}$ – множество взаимосвязей между ними).

В [119] исходные данные задачи синтеза оптимальных логических структур БД представлены в детерминированном виде. Здесь, в соответствии с результатами исследования, представленными выше, часть исходных данных представлена в нечетком виде. Исходными данными для решения задачи построения БД являются параметры канонических структур.

Для решения поставленной задачи рассмотрены методы расчета значений суммарного числа просматриваемых указателей связи в БД,

сравниваемых экземпляров групп $\beta_{ii'}^k$ и средних значений суммарного числа выбираемых, при поиске экземпляров групп $p_{ii'}^k$. Известно, что значения этих параметров зависят от структуры запроса [108].

Выражения для расчета данных параметров должны включать возможные варианты поиска экземпляров i' -й группы, лежащей на (ii') -м пути доступа соответствующего (k -го) запроса. Считается, что соответствующий путь доступа к дереву поиска от k -го запроса направлен от главной группы C_i к подчиненной группе $C_{i'}$ ($f_{ii'}^k = 1$). При предположении, что на предыдущем шаге поиска из C_i было отобрано некоторое подмножество ее экземпляров r_i^* , $r_i^* = \overline{1, r_i}$ и по условию k -го запроса из $C_{i'}$ требуется выбрать запрашиваемое подмножество $r_{i'}^*$ ее экземпляров ($r_{i'}^* = \overline{1, r_{i'}}$), возможны следующие варианты поиска:

1. $r_i^* = r_{i'}^* = 1$ - нужно по одному экземпляру из C_i поиска экземпляра из $C_{i'}$.
2. $r_i^* = 1, r_{i'}^* = \delta_{ii'}$ - нужно по одному экземпляру главной группы найти все подчиненные ему экземпляры группы.
3. $r_i^* = 1, 2 \leq r_{i'}^* \leq \delta_{ii'}$ - нужно по экземпляру из C_i группы найти некоторое подмножество подчиненных ему экземпляров из $C_{i'}$.
4. $2 \leq r_i^* \leq r_i, r_{i'}^* = r_i^*$ - нужно по подмножеству экземпляров из C_i найти все подчиненные ему экземпляры из $C_{i'}$.
5. $2 \leq r_i^* \leq r_i, r_{i'}^* = r_i^* \delta_{ii'}$ - нужно по подмножеству экземпляров группы C_i найти все подчиненные им экземпляры группы $C_{i'}$.
6. $2 \leq r_i^* \leq r_i, 2 \leq r_{i'}^* \leq \delta_{ii'}$ - нужно по подмножеству экземпляров из C_i найти подмножество подчиненных им экземпляров из $C_{i'}$.
7. $r_i^* = r_i, r_{i'}^* = r_i \delta_{ii'} = r_i$ - нужно по всем экземплярам из C_i найти все экземпляры из $C_{i'}$.

Для каждого варианта поиска определяется среднее суммарное число просматриваемых указателей связи и сравниваемых экземпляров групп $\beta_{ii'}^k$.

Очевидно, что $\beta_{ii'}^k = \beta_{ii'}^{ky} + \beta_{ii'}^{kcp}$, где $\beta_{ii'}^{ky}$ – среднее число просматриваемых указателей связи; $\beta_{ii'}^{kcp}$ – среднее число сравниваемых экземпляров групп.

Для всех вариантов поиска справедливы соотношения: $\beta_{ii'}^{ky} = \beta_{ii'}^{kcp}$ и $\beta_{ii'}^k = 2\beta_{ii'}^{kcp}$, поэтому для расчета $\beta_{ii'}^{kcp}$ целесообразно использовать процедуры определения среднего числа сравниваемых элементов списка.

Для первого варианта поиска имеем: $\beta_{ii'}^{kcp} = \sum_{n=1}^{\delta_{ii'}} n p_{ii'}$, где $p_{ii'}$ – вероятность обнаружения n -го подчиненного экземпляра из c_i . При выполнении условия равновероятности результатов из c_i , то $p_{ii'} = 1/\delta_{ii'}$ и тогда

$$\beta_{ii'}^{kcp} = \sum_{n=1}^{\delta_{ii'}} n \frac{1}{\delta_{ii'}} = \left[\frac{\delta_{ii'}+1}{2} \right].$$

Отсюда $\beta_{ii'}^k = \left[\frac{\delta_{ii'}+1}{2} \right]$.

Для второго варианта поиска $\beta_{ii'}^{kcp} = \delta_{ii'}$, следовательно, $\beta_{ii'}^k = 2\delta_{ii'}$.

Для третьего варианта поиска есть два варианта реализации:

1. Множество признаков, задаваемых в условиях запроса, и множество подчиненных экземпляров из c_i одинаковым образом упорядочены.

2. Данные множества неупорядочены. Тогда либо $\beta_{ii'}^{kcp} = \delta_{ii'}$ и $\beta_{ii'}^k = 2\delta_{ii'}$, т.е. требуется просмотр всего списка подчиненных экземпляров, либо для поиска r_i^* экземпляров i' -й группы, требуется r_i^* обращений к связному списку. При каждом очередном обращении к списку находится один из экземпляров множества r_i^* , после чего осуществляется возврат к началу списка, т. е. к экземпляру главной группы, что обеспечивается использованием комбинированной системы указателей.

Среднее число сравниваемых экземпляров из c_i при прямом просмотре списка равно $\beta_{ii'(np)}^{kcp} = r_i^* \left[\frac{\delta_{ii'}+1}{2} \right]$ и, следовательно, $\beta_{ii'(np)}^k = 2r_i^* \left[\frac{\delta_{ii'}+1}{2} \right]$.

При обратном просмотре осуществляется возврат в начало списка. Количество просматриваемых при этом указателей $\beta_{ii'(об)}^{ky}$ равно $\beta_{ii'(об)}^{ky} = r_i^* - 1$. Поэтому для рассматриваемого случая имеем

$$\beta_{ii'}^k = \beta_{ii'(np)}^k + \beta_{ii'(об)}^k = 2r_i^* \left[\frac{\delta_{ii'}+1}{2} \right] + r_i^* - 1 = r_i^* \left(1 + 2 \left[\frac{\delta_{ii'}+1}{2} \right] \right) - 1. \quad (3.17)$$

Варианты 4, 5, 6 и 7 характеризуются тем, что поиск осуществляется в r_i^* связанных списках, $r_i^* \geq 2$, поэтому для их реализации целесообразно применить операцию возврата в начало анализируемого списка. В этом случае для этих вариантов поиска получим:

$$\beta_{ii'}^k = 2\beta_{ii'}^{kcp} + \beta_{ii'(об)}^{ky},$$

где $\beta_{ii'(об)}^{ky}$ среднее суммарное число указателей связи, необходимых для обратных просмотров связей с целью выбора следующего списка множества r_i^* . Значения $\beta_{ii'}^{kcp}$ для этих вариантов поиска также могут быть получены путем умножения на величину r_i^* значений $\beta_{ii'}^{kcp}$ для 1, 2 и 3-го вариантов, соответственно. Значения $\beta_{ii'}^{ky}$ для этих вариантов поиска будут равны $r_i^* r_i^* - 1$ в силу отсутствия необходимости обратного просмотра связей из искомого экземпляра подчиненной группы, расположенного в последнем по порядку списке. Таким образом, для четвертого варианта поиска имеем

$$\beta_{ii'}^{kcp} = r_i^* \left[\frac{\delta_{ii'+1}}{2} \right], \quad (3.18)$$

$$\beta_{ii'}^k = 2r_i^* \left[\frac{\delta_{ii'+1}}{2} \right] + r_i^* - 1 = r_i^* \left(1 + 2 \left[\frac{\delta_{ii'+1}}{2} \right] \right) - 1 \quad (3.19)$$

Аналитические выражения для остальных вариантов можно получить аналогичным образом.

Значения параметра $p_{ii'}^k$ также зависят от варианта поиска, задаваемого условиями запроса к главной группе. Для расчета параметра $p_{ii'}^k$ целесообразно использовать выражения, полученные для $\beta_{ii'}^{kcp}$, так как при однократном просмотре списка экземпляра связи $p_{ii'}^k = \beta_{ii'}^{kcp}$, а при многократном – значение параметра $p_{ii'}^k$ не больше значения $\beta_{ii'}^{kcp}$. Таким образом, для первого варианта поиска имеем

$$p_{ii'}^k = \left[\frac{\delta_{ii'+1}}{2} \right], \quad (3.20)$$

а для второго варианта - $p_{ii'}^k = \delta_{ii'}$.

Для обоих случаев третьего варианта поиска $p_{ii'}^k = \delta_{ii'}$. Справедливость данного равенства для случая упорядоченных множеств достаточно очевидна, так как просмотр списка осуществляется однократно и, следовательно, $p_{ii'}^k = \beta_{ii'}^{kcp}$. Для случая неупорядоченных множеств $\beta_{ii'}^{kcp} = r_i^* \left[\frac{\delta_{ii'+1}}{2} \right] u r_i^* \geq 2$. Поэтому $\beta_{ii'}^{kcp} > \delta_{ii'}$ и, следовательно, для поиска r_i^* экземпляров группы c_i требуется неоднократный просмотр всего списка, т. е. $p_{ii'}^k = \delta_{ii'}$. Для четвертого варианта имеем: $p_{ii'}^k = r_i^* \left[\frac{\delta_{ii'+1}}{2} \right]$; для пятого варианта: $p_{ii'}^k = r_i^* \delta_{ii'}$; для шестого варианта: $p_{ii'}^k = r_i^* \delta_{ii'}$; для седьмого варианта поиска: $p_{ii'}^k = r_i$.

В моделях и постановках задач построения логических структур БД в качестве исходных данных, помимо количественных характеристик структуры БД и запросов к ней, используются и временные параметры. В связи с тем, что информационные системы, как правило, функционируют в неоднородных локальных вычислительных сетях (ЛВС) на различных по мощности электронных вычислительных машинах (ЭВМ), на которых установлены разные системы управления базами данных (СУБД), то, в таких условиях затруднительно точно определить временные параметры БД, которые зависят от характеристик ЛВС, ЭВМ и СУБД. Поэтому, в качестве нечетких исходных данных задачи построения структуры БД будем применять такие временные параметры, как: среднее время поиска и выборки экземпляра группы $\tilde{t}_t$; среднее время слияния (композиции) групп в логическую запись $\tilde{t}_{cl}$; среднее время загрузки записи в СБД $\tilde{t}_z$; время формирования указателя (адреса) связи $\tilde{t}_y$; среднее время просмотра указателя связи и сравнения экземпляра группы $\tilde{t}_{nc}$; среднее время корректировки одного экземпляра группы данных $\tilde{t}_{кор}$.

Эти временные параметры БД представлены в виде нечетких переменных (L-R) - типа [78, 108]:

$$\mu_{A_q}(\tilde{t}_r) = \begin{cases} L_q \left(\frac{t_r^* - t_r}{M_q} \right), & t_r < t_r^*, \\ R_q \left(\frac{t_r - t_r^*}{M_q} \right), & t_r > t_r^*; \end{cases} \quad r \in \{i, cl, z, y, nc, kor\}, q = \overline{1, 6}, i = \overline{1, I}. \quad (3.20)$$

Для различных вариантов решения задачи значения временных параметров будут определяться соотношениями:

1. $t_r < t_r^*$, для L_q^- :

$$t_r = t_r^* - M_q L_q^- (\mu_{A_q}(\widetilde{t_r})), r \in \{i, сл, з, у, nc, кор\}, q = \overline{1,6}, i = \overline{1,I} \quad (3.21)$$

2. $t_r > t_r^*$, для L_q^- :

$$t_r = t_r^* + M_q L_q^- (\mu_{A_q}(\widetilde{t_r})), r \in \{i, сл, з, у, nc, кор\}, q = \overline{1,6}, i = \overline{1,I}. \quad (3.22)$$

Таким образом, рассмотрены исходные данные задачи построения структур БД. Часть из них представлена в виде нечетких переменных. Данный подход позволит объективнее определить исходные данные задачи построения структур БД и проектировать эффективные БД.

3.5 Выводы по главе 3

В результате применения теории информационного поиска в сочетании с предложенным критерием и разработанными моделями оценки пертинентности при решении задач распознавания графических объектов в условиях неопределенности построена процедурная модель решения задачи распознавания графических объектов, отличающаяся возможностью определения динамических характеристик распознаваемого объекта на основе кубических сплайнов.

Определены три основных уровня решения ресурсной задачи и обоснованы процедурные модели ее решения.

Реализация разработанных моделей предъявляет определенные требования к используемым базам данных. Решение задачи построения структур баз данных (БД) позволило определить структуру БД, обеспечивающую наилучшую производительность.

4. ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ МОДЕЛЕЙ И ОЦЕНКА ЭФФЕКТИВНОСТИ ИХ ПРИМЕНЕНИЯ

4.1 Применение моделей в информационной системе образовательного назначения

В настоящее время в комплексе с традиционными учебно-методическими средствами все шире применяют информационные системы образовательного назначения (ИСОН) [11, 58, 67, 85].

Структура курса обучения в таких системах представляет собой набор информационных ресурсов или учебно-тренировочных задач (УТЗ). Поскольку наполнение этих информационных ресурсов осуществляется различными людьми, то, как правило, и форма их представления различна: текстовые страницы, веб-страницы, ссылки на файлы, тестовые задания, вопросы и т.п.

Один из вариантов представления информационных ресурсов на основе сети Петри представлен на рисунке 4.1.

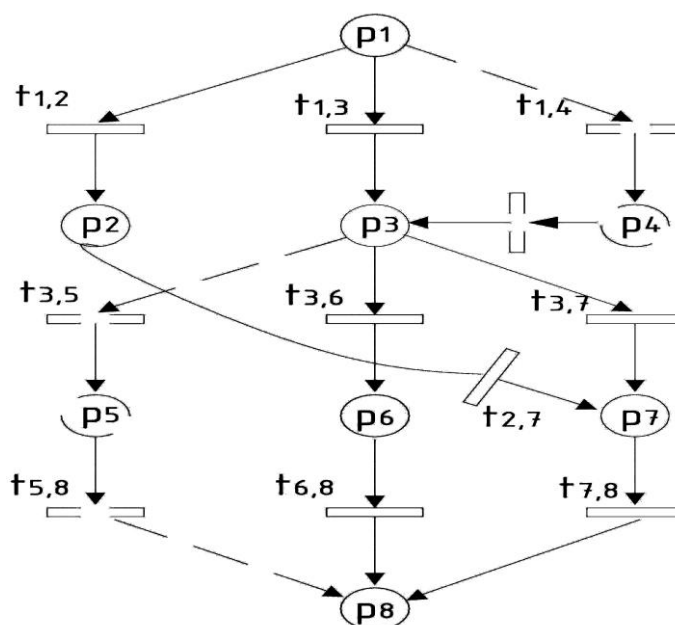


Рисунок 4.1 - Модель структуры курса обучения на основе НВСП.

На рисунке обозначено: $P = \{p_1, p_2, \dots, p_{m(p)}\}$ - множество позиций;
 $T = \{t_1, t_2, \dots, t_{m(p)}\}$ - множество переходов.

Поскольку в последние годы большой популярностью в университетах мира пользуется программное обеспечение MOODLE (Modular Object-Oriented Dynamic Learning Environment), то она выбрана для реализации модели структуры курса обучения на основе сети Петри.

Применение в учебном процессе ИС организации обучения Moodle достаточно подробно рассмотрено, например в [29, 58]. Все подходы, рассмотренные авторами, затрагивают только базовые функции LMS Moodle, но не рассматриваются возможности дополнения системы своими программными модулями. Актуальной проблемой является разработка ИС организации обучения, которая будет обладать функциональными возможностями, позволяющими проведение адаптации структуры курса обучения для каждого пользователя индивидуально

Распространение MOODLE в качестве программного обеспечения с открытым исходным кодом под лицензией GNU GPL позволяет создать на своей основе собственную систему обучения, обладающую всеми необходимыми возможностями функционала системы.

Пример разработки собственной ИС обучения студентов материалам раздела «Теория» описан в [58]. В процессе применения ИС организации обучения студент имеет возможность выбрать курсы обучения, которые ему доступны. Далее он изучает тематический материал, а затем выполняет задания, входящие в рубежный контроль и представляющие собой набор тестовых вопросов, который подготовлены для первоначального прохождения курса (см. рисунок 4.2).

Описанные выше возможности MOODLE являются базовыми. Расширенный функционал программного комплекса мониторинга обучения, разработанный на языке php с использованием системы управления базами данных MySQL, основан на применении сетей Петри и генерации адаптивной структуры курса обучения.

Разработанный программный комплекс мониторинга обучения состоит из компонент, взаимодействуют друг с другом, каждый из которых

представляет неделимый элементарный блок учебного материала. Такой учебный материал представляется в виде текстовой страницы, веб-страницы, пунктом глоссария, ссылки на файл, заданием, вопросом текста и имеет свое собственное состояние. Состояние компоненты можно определить как некую абстракцию соответствующей информации, которая необходима для описания будущих действий. Состояние компоненты находится в зависимости от ее предыстории. Со временем состояние компоненты изменяется. Понятие «состояние» является крайне важным, поскольку оно служит отображением поведения моделируемой системы. Одним из основных свойств действий компоненты системы является совмещенность или параллелизм, т.е. действия компоненты могут производиться одновременно (т.е. параллельно) с действиями других компонент. Например, в системе, представленной на рисунке 4.2, одновременно может происходить процесс авторизации пользователя, процесс чтение другим пользователем веб-страниц или файлов, процесс проведения тестирования для третьего пользователя и т.п.

1	... функция - в программировании особый вид функций, которые объявляются в месте использования и не получают уникального идентификатора для доступа к ним. (Дополните предложение)
Баллов: 1/1	
Ответ:	Анонимная ✓
Отправить	
Верно	
Баллов за ответ: 1/1.	
2	Инструмент сред программирования, компоновщик -
Баллов: 1/1	
Выберите один ответ.	☐ а. создаёт и изменяет исходные символьные файлы, содержащие программу ☒ б. поддерживает совокупность объектных файлов с подпрограммами, типами данных X ☐ в. формирует исполняемый файл ☐ г. переводит программы с языка программирования на язык машинных кодов
Отправить	
Неверно	
Баллов за ответ: 0/1.	
3	Языки программирования по ориентации на сферу применения делятся на (Дополните предложение)
Баллов: 1/1	
Выберите один ответ.	☐ а. процедурные, непроцедурные, функциональные ☒ б. процедурно-ориентированные, проблемно-ориентированные, машинно-ориентированные, объектно-ориентированные ✓ ☐ в. операторные, функциональные, проблемно-ориентированные ☐ г. машинно-ориентированные, объектно-ориентированные, непроцедурные
Отправить	
Верно	
Баллов за ответ: 1/1.	

Рисунок 4.2 – Пример работы подсистемы тестирования

На рисунке 4.3 представлены возможные позиции и переходы в структуре курса обучения пользователя после реконструкции набора учебных элементов на основе нечетких раскрашенных сетей Петри. Набор позиций $p_1, p_2, \dots, p_n$ соответствует компонентам учебного материала. Маркер в позиции p_3 описывает этап прохождения курса обучения одного из пользователей.

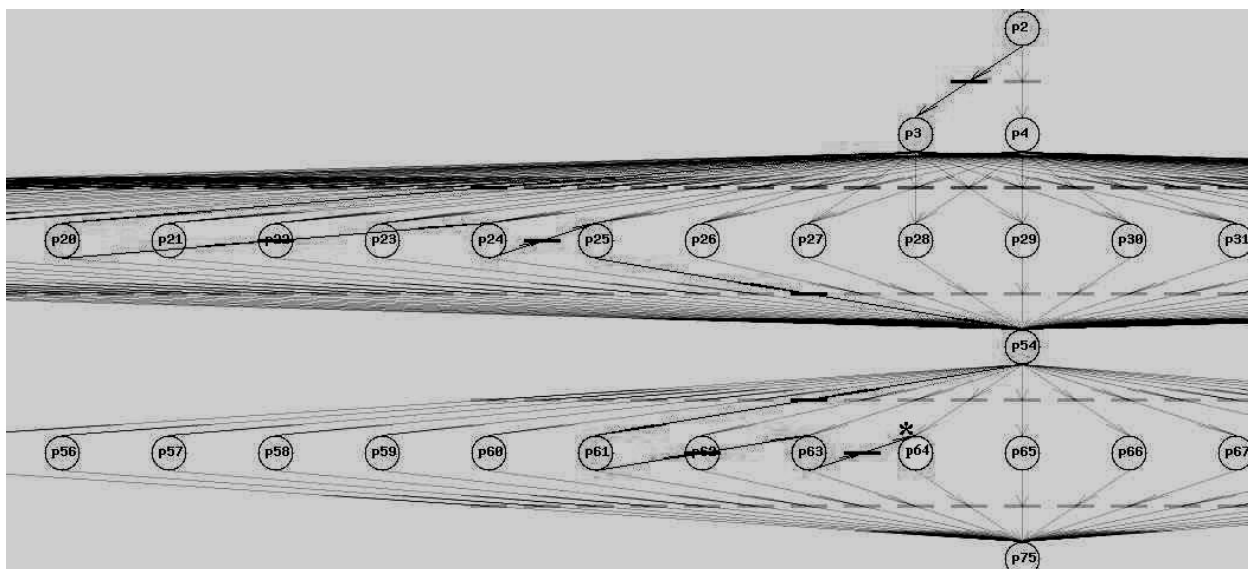


Рисунок 4.3 – Модель прохождения обучения в виде сети Петри

Допустимые в сети цвета маркера определены следующим образом:

$ID = (Id_group, Id_user, Id_req, Full, Number, Id_app)$, где

- Id_group . целое неотрицательное число;
- Id_user . целое неотрицательное число.
- Id_req . целое неотрицательное число.
- $Full$. логического типа с набором значений $\{true, false\}$;
- $Number$. целое неотрицательное число.
- Id_app . целое неотрицательное число.

$Req = \text{list of } (key, value)$, где

- Key . строкового типа.
- $Value$. строкового типа.

$Data = \text{list of } (key, value)$, где также,

- Key . строкового типа.
- $Value$. двоичный набор данных.

Функции цвета, позволяют математически описать параметры маркера, который полностью соответствует данным одного из обучаемых.

Результатом прохождения теста является выдаваемая комплексом мониторинга обучения числовая оценка в количественном и процентом соотношении. Однако оценка не позволяет сделать выводы, которые дадут полную информацию успеваемости студента по предмету. Разработана процедура, позволяющая получить качественную оценку успеваемости студента, и уже на основе этой оценки подготовить адаптированный для данного студента курс обучения [42].

В рамках тематического класса УТЗ соотносятся с проверяемыми положениями учебного материала. Задание позволяет контролировать одно или несколько положений учебного материала. Для повышения качества контроля изучения учебного материала, связанного с распознаванием графических объектов, созданных обучаемым, в эту систему интегрированы программные элементы, построенные с применением разработанных моделей. Результат работы моделей распознавания графических объектов заключается в «окрашивании» набора позиций, соответствующих или не соответствующих успешному прохождению этапа или выполнения задания.

Предложенная модель построения структуры курса обучения отличается от известных тем, что позволяет сформировать, на основе модели уровня подготовки пользователя, оптимальный набор информационных ресурсов. Это дает возможность дальнейшей разработки и моделирования ИСОН. Преимуществом применения сетей Петри при построении модели структуры курса обучения является математически строгое описание модели, что позволяет проводить их анализ с помощью вычислительной техники.

4.2 Экспертные системы в судебно-медицинской экспертизе

Важной задачей судебно-медицинской экспертизы и психиатрии в Российской Федерации является профилактика психических расстройств и их

негативных последствий. Одним из тяжелых социально значимых последствий является совершение общественно опасных деяний и преступлений лицами, страдающими психическими расстройствами.

В связи с этим совершенствование профилактики общественноопасных действий психически больных является одним из перспективных и стратегически важных направлений развития общей и судебной психиатрии.

В современных реалиях работа психиатров, в том числе судебных, связана с анализом огромного количества информации по каждому пациенту. Как следствие, даже опытным психиатрам тяжело учитывать все имеющие вес факторы болезней пациентов, что в значительной степени сказывается на точности и эффективности работы психиатрических больниц. Одним из возможных путей решения данной проблемы является интегрирование инновационных методов обследования в практическую работу специалистов, занятых в области судебной психиатрии, для получения максимального точного решения сложных диагностических вопросов и необходимых выводов в короткие сроки.

В ходе анализа существующих научных трудов о проблеме информационных технологий в психиатрии, было предложено разработать экспертную систему для судебной психиатрии.

Если исходить из теории оценки опасности, вся информация о пациентах делится на три оси «личность – синдром – ситуация», однако в программе вся информация будет представлена в привычном для врача формате истории болезни: социально-демографическая характеристика, анамнез жизни, анамнез болезни, психическое состояние с описанием актуального поведения. Соответственно для обработки в программе и для представления вся личная информация должна быть представлена в двух разных форматах.

В судебной психиатрии выделено 3 наиболее трудоемкие, а, следовательно, нуждающиеся в автоматизации проблемы [81-83]: принятие решения о необходимости постановки пациента на активное диспансерное наблюдение, принятие решения судебно-психиатрической экспертизой о типе принудительного лечения и оценка риска общественной опасности. Решение

каждого из данных вопросов должно быть выделено в отдельный блок программы, каждый из которых будет давать результат, базируемый на собственных метода анализа информации.

Проблема эффективной постановки диагноза выражена в первую очередь огромным количеством синдромов и факторов, представленных, зачастую, графическими объектами (электроэнцефалограммами, реоэнцефалограммами, результатами плетизмографии и фонокадиографии и т.п.). Человеку крайне сложно учесть все эти синдромы и факторы, хотя каждый из этих факторов имеет ту или иную степень влияния на конечный вывод. Следовательно, можно выделить необходимость в разработке новой информационной интеллектуальной системы, использующей классификацию МКБ-10, что позволит применять данную систему в психиатрических больницах. Учитывая эту необходимость, было предложено разработать информационную интеллектуальную систему, рассчитанную на судебную психиатрию. Было выделено три вопроса, требующих автоматизации: принятие решения о необходимости постановки пациента на активное диспансерное наблюдение, принятие решения судебно-психиатрической экспертизой о типе принудительного лечения и оценка риска общественной опасности. Под каждую проблему в системе должен быть выделен отдельный блок. Работа всех блоков объединена одним хранилищем данных, связь с которым должна осуществляться через систему электронного документооборота.

Основываясь на информации о существующих информационных интеллектуальных системах, разработанных в интересах психиатрии, можно сделать вывод о необходимости создания подобной системы для российской судебной психиатрии, основанной на классификации психических расстройств МКБ-10.

Новая экспертная система должна быть представлена тремя блоками обработки личных данных пациента и системой управления пользователями и пациентами. Три базовых блока – это:

1. блок принятия решения о необходимости постановки на активное

диспансерное наблюдение;

2. блок принятия решения о типе принудительного лечения;
3. блок оценки риска общественной опасности.

Все информация о пользователях и пациентах должна храниться в базе данных. Доступ к информации разного уровня должен быть организован в соответствии с политикой избирательного управления доступом RBAC (Role Base Access Control). Иными словами, каждый зарегистрированный пользователь будет отождествлен с одной из трех групп. Пользователи каждой группы должны обладать определенными правами и обязанностями. Так, например, администраторы отвечают за стабильность работы системы, настройку системы и сети, добавление новых пользователей и не касаются медицинских аспектов работы программы.

Диаграмма прецедентов (рисунок 4.4) (use case diagram) отражает отношения между актерами и прецедентами. Тонем выделены элементы диаграммы, в которых использованы разработанные модели. Актерами в данном контексте являются пользователи, относящиеся к одной из следующих групп: администраторы; врачи; ассистенты.



Рисунок 4.4 – Диаграмма прецедентов

Подобное разделение доступа позволит использовать экспертную систему максимально гибко и эффективно.

Основная функция врача при работе с экспертной системой – проведение анализа личных данных пациентов и получение вывода на основе полученных данных. Планируется, что система будет проводить анализ данных пациентов по 3 направлениям, требующим автоматизации. Решение каждого вопроса выделено в отдельный блок программы со своей формой и характерными выводами в виде наглядных диаграмм.

Каждый из блоков использует те или иные данные пациента, относящиеся к одной из 4-х описанных выше категорий (социально-демографическая характеристика, анамнез жизни, анамнез болезни и психическое состояние на момент внесения информации). Блок принятия решения о необходимости постановки на активное диспансерное наблюдение на основе данных пациента в качестве результата работы выдает обоснованное заключение о необходимости включения пациента в группу АДН или вывода из нее. Блок принятия решения о типе принудительного лечения делает заключение о необходимости одного из 4-х типов принудительного лечения:

1. амбулаторное принудительное наблюдение и лечение у психиатра;
2. принудительное лечение в психиатрическом стационаре общего типа;
3. принудительное лечение в психиатрическом стационаре специализированного типа;
4. принудительное лечение в психиатрическом стационаре специализированного типа с интенсивным наблюдением.

Последний блок оценки риска общественной опасности использует данные четырех категорий вкупе с данными о настоящем психическом, психологическом и социальном состоянии пациента делает вывод о возможном выборе более мягкого типа принудительного лечения или полном снятии с него.

Диаграмма развертывания служит для описания конфигурации сети всех обрабатывающих узлов. Конфигурация схемы развертывания системы может отличаться в зависимости от количества персонала и количества вычислительных

узлов (ПК). На рисунке 4.5 представлена возможная схема развертывания экспертной системы. Данная схема подходит для описания конфигурации сети с неограниченным количеством пар «врач-ассистент», одним администратором и принтером при наличии отдельного ПК для каждого пользователя. В то же время любой пользователь может воспользоваться любым свободным ПК и получить доступ к работе со своими пациентами благодаря использованию политики избирательного управления доступом (RBAC).

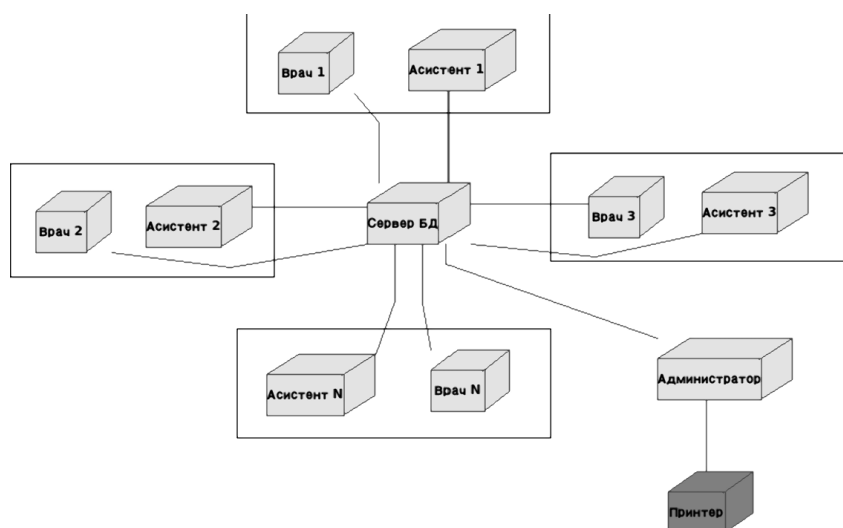


Рисунок 4.5 – Диаграмма развертывания

Диаграмма развертывания иллюстрирует логику разработанной системы, четко расписаны роли различных групп пользователей («Врачи», «Ассистенты», «Администраторы»). Описаны некоторые технические аспекты взаимодействия клиентов

Основываясь на информации о существующих информационных интеллектуальных системах, разработанных в интересах психиатрии, можно сделать вывод о необходимости создания подобной системы для российской судебной психиатрии, основанной на классификации психических расстройств МКБ-10.

В общем виде разрабатываемая система, основанная на клиент-серверной архитектуре должна включать в себя, по меньшей мере, три вычислительных узла (ПК врача, ПК администратора, сервер БД).

Конфигурация схемы развертывания системы может отличаться в зависимости от количества персонала и количества вычислительных узлов.

Основным рабочим элементом экспертной системы является так называемый аналитический блок. Данный блок содержит в себе аналитические и численные модели, составленные на основе существующих методик современной психиатрии. Эти модели используют различные данные, хранящиеся на сервере БД. Структура БД разрабатываемой экспертной системы представлена ниже.

На рисунке 4.6 представлена диаграмма базы данных разрабатываемой системы. Структура базы данных описана в п. 3.4 и была разработана для хранения информации как о пользователях, так и о пациентах со всеми личными данными на одном сервере. База данных состоит из шести таблиц:

1. «users» – содержит информацию о пользователях системы;
2. «permissions» – привязывает пациентов к их лечащим врачам (необходимо для разграничения полномочий RBAC);
3. «persons» – содержит общую информацию о пациентах, включая их паспортные данные и диагноз по МКБ-10;
4. «features» – в этой таблице содержится список характеристик, относящихся к одной из следующих категорий: социально демографические характеристики, анамнез жизни, анамнез болезни, психическое состояние на момент внесения информации;
5. «options» – содержит варианты для характеристик, описанных в таблице «features»;
6. «characteristics» – содержит все характеристики для всех пациентов.

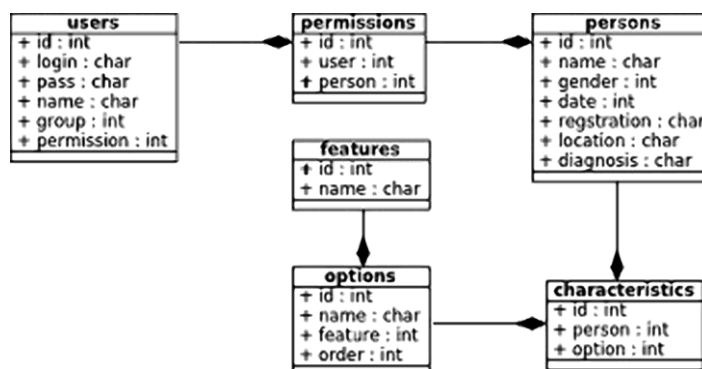


Рисунок 4.6 – Диаграмма БД

Все характеристики условно поделены на три группы: социально-демографические характеристики, анамнез жизни, анамнез болезни. В соответствии с личным опытом врачей и существующими методиками решения следующих медицинских задач: принятия решения о необходимости постановки на активное диспансерное наблюдение, принятия решения о типе принудительного лечения, оценка риска общественной опасности было предложено хранить данные пациентов в представленной ниже форме. При этом сами характеристики хранятся в таблице «features», а варианты характеристик в таблице «options».

Диаграмма классов является одной из диаграмм статической структуры. Она описывает структуру системы, показывая ее классы.

Разрабатываемая экспертная система состоит из следующих классов:

- AuthForm – класс формы авторизации;
- SettingsForm – класс формы настроек и редактирования пользователей;
- MainForm – класс главного окна;
- MathModel – класс аналитических блоков;
- InfoForm – класс формы диалога заполнения характеристик пациента;
- PatientForm – класс формы редактирования общих данных пациентов;
- ResultForm – класс формы вывода результатов работы аналитических блоков;
- User – класс пользователя;
- Patient – класс пациента;
- DB – класс работы с БД.

В данной работе не описаны классы, работающие с аналитическими блоками (MathModel и ResultForm). Во всех классах форм (кроме MainForm) функция Show() создает новую форму, а функция Hide() ее скрывает. В классе MainForm функция Show() единожды создает главную форму, а функция Close() закрывает ее.

AuthForm. Функция `Autorization()` производит сверку введенных в форму данных с данными пользователя в БД посредством класса DB.

SettingsForm. Функция `AddUser()` вносит данные нового пользователя в БД, функция `DelUser()` удаляет данные существующего пользователя из БД. Функция `SaveSettings()` сохраняет данные настроек сети и БД в файл.

MainForm. Функция `UpdList()` главной формы обновляет отображение списков пользователей. Функции `AddPatient()` главной формы создает окно добавления нового пациента. Функция `EditPatient()` создает окно редактирования выбранного пациента. Функция `Estimation()` выбирает один из трех аналитических блоков и запускает процесс вычислений. Функция `Settings()` создает окно настроек.

InfoForm. Функция `Dialog()` создает диалог заполнения характеристик пациента.

PatientForm. Функция `Add()` добавляет нового пациента, функция `Edit` обновляет общие данные пациента, функция `Info()` создает форму InfoForm и запускает функцию `Dialog()`, функция `Del()` удаляет выбранного пользователя.

Класс User загружает функцией `Read()` данные пользователя и содержит их в своих переменных.

Класс Patient загружает данные выбранного пациента функцией `Read()` и хранит их в своих переменных. Функцией `Info()` загружается подробная характеристика пациента, а функция `Close()` очищает все переменные класса Patient.

Класс DB используется остальными классами для связи с БД. В нем содержатся все запросы.

Данная работа содержит две важные диаграммы, описывающие статическую структуру экспертной системы. Приведенные диаграммы подробно описывают все классы этой системы и базовые функции, структуру базы данных и информацию, необходимую для заполнения характеристик пациентов. Эта работа максимально подробно описывает статическую структуру разрабатываемой

Анализ результатов исследования, проведенных в данном подразделе,

показал, что типовая начальная конфигурация экспертной системы состоит из 7 элементов (узлов). Поэтому представляет интерес оценить живучесть такой системы.

4.3 Оценка эффективности применения разработанных моделей

4.3.1 Оценка живучести информационной системы для распознавания графических объектов с использованием системы matlab

В инфраструктуре современного информационно-индустриального общества сетевые структуры занимают одно из ключевых мест. Типичными задачами, возникающими перед аналитиками и инженерами-проектировщиками, являются расчет распределения потоков внутри сети при воздействии на сетевую структуру неблагоприятных внешних факторов, выявление «узких мест» и т. д. Такой структуре присуще свойство уязвимости [1 из статьи 6_лыонг].

Применительно к сетевым структурам живучесть – это способность системы выполнять основные функции во время атак, повреждений и аварийных ситуаций.

Использование искусственных нейронных сетей (ИНС) как инструмента для исследования живучести является оправданным решением, так как нейронная сеть способна решать задачи, в которых неизвестны закономерности развития ситуации и зависимости между входными и выходными данными. Традиционные математические методы и экспертные системы в таких случаях недостаточно эффективны.

Для оценки живучести исследуемой структуры использован подход к оценке живучести локальных вычислительных сетей с использованием распределенных информационных систем, описанный в [38, 61, 64, 67, 69, 71, 72].

Суть процедуры, реализующей данный подход, заключается в следующем.

Для значительного уменьшения времени расчета живучести ИС применена технология распределенных систем, которая позволяет создать географически распределенные вычислительные инфраструктуры, которые объединяют разнородные ресурсы и реализуют возможность коллективного доступа к этим ресурсам. Принципиальной новизной этих технологий является объединение ресурсов путем создания компьютерной инфраструктуры нового типа, обеспечивающей глобальную интеграцию информационных и вычислительных ресурсов на основе сетевых технологий и специального программного обеспечения промежуточного уровня (middleware), а также набора стандартизованных сервисов (служб) для обеспечения надежного совместного доступа к географически распределенным информационным и вычислительным ресурсам: отдельным компьютерам, кластерам, хранилищам информации и сетям. Однако необходимо было разработать математический аппарат распределения вычислений в такой системе для наиболее максимального ускорения проведения вычислений, связанных с оценкой живучести ЛВС.

В настоящее время существуют готовые пакеты прикладных программ (ППП) для работы с ИНС. Примером может служить Neural Network Toolbox (NNT) [2], функционирующий под управлением ядра системы MATLAB [3, 4], что делает его таким популярным. К недостаткам данного пакета можно отнести следующее:

- некорректное обучение сети в GUI режиме для некоторых алгоритмов обучения.
- достаточно сложный консольный вариант работы с ППП NNT для простых пользователей.
- цена. MATLAB стоит около \$3000, NNT – около \$2000.

Создана и обучена ИНС с помощью системы MATLAB R2010a с ППП NNT версии 6.0.4.

Сгенерировано и экспортировано 500 обучающих топологий ИС в формат данных MATLAB. Примеры сгенерированных тестовых топологий сетевых структур с порядковыми номерами 1, 30 и 60 представлены на рисунке 4.7.

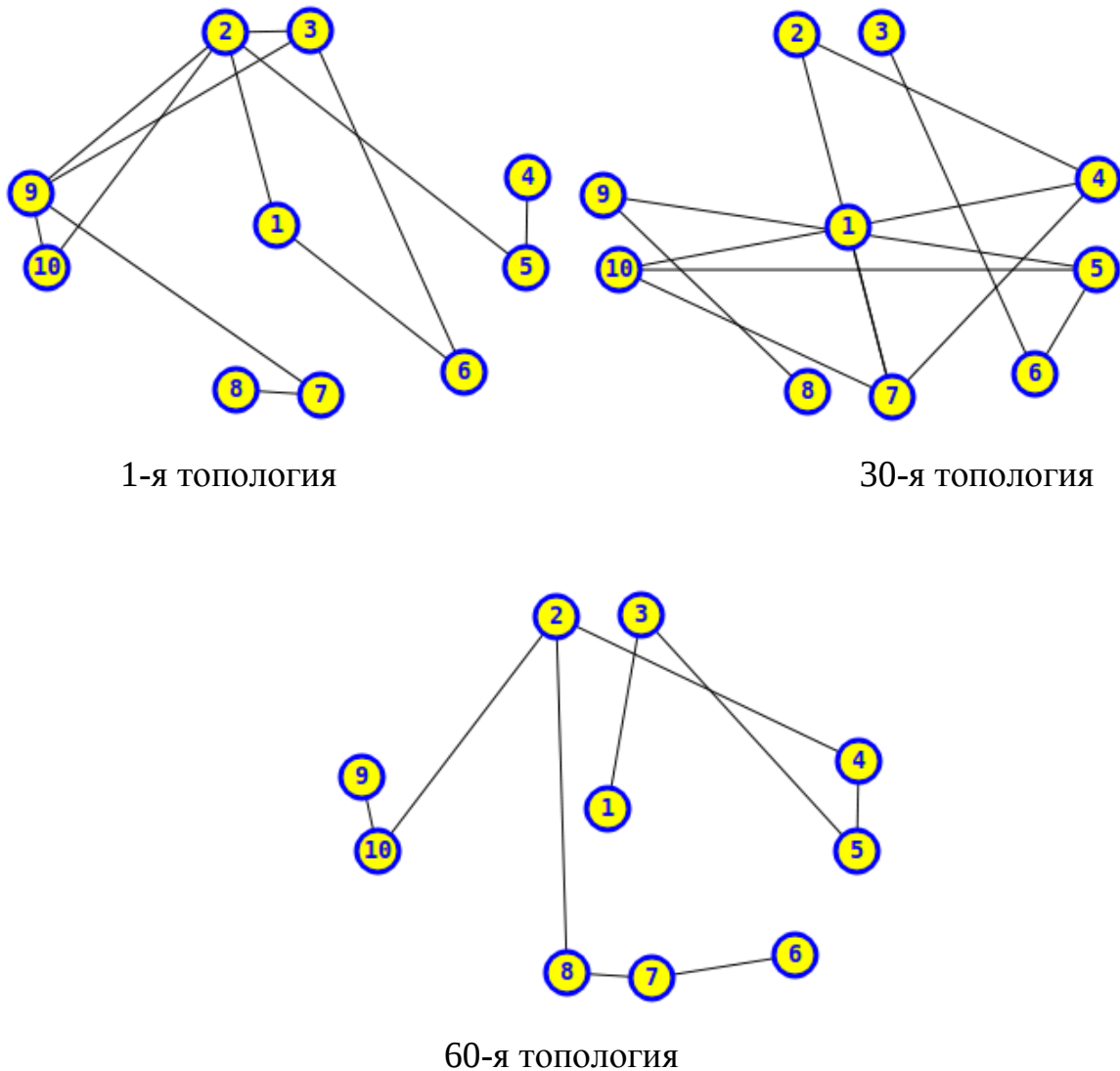


Рисунок 4.7 – Примеры тестовых топологий

Введены обучающие данные в среду MATLAB:

```
>> INP = [0 0 1 1 ... ];
```

Создана ИНС (сеть с прямой передачей сигнала, количество нейронов в скрытом слое – 21, выход – 1, сигмоидальная функция активации нейронов в скрытом и выходном слое, алгоритм обучения – пороговый обратного распространения ошибки):

```
>> net = newff(minmax(INP),[21,1],
{'logsig', 'logsig'},'trainrp');
```

Задано условие останова – достижение 5000 эпох:

```
>> net.trainParam.epochs = 5000; Обучаем сеть:
```

```
>> net=train(net,INP,OUTP).
```

Интерфейс модуля NNT в режиме обучения представлен на рисунке 4.8.

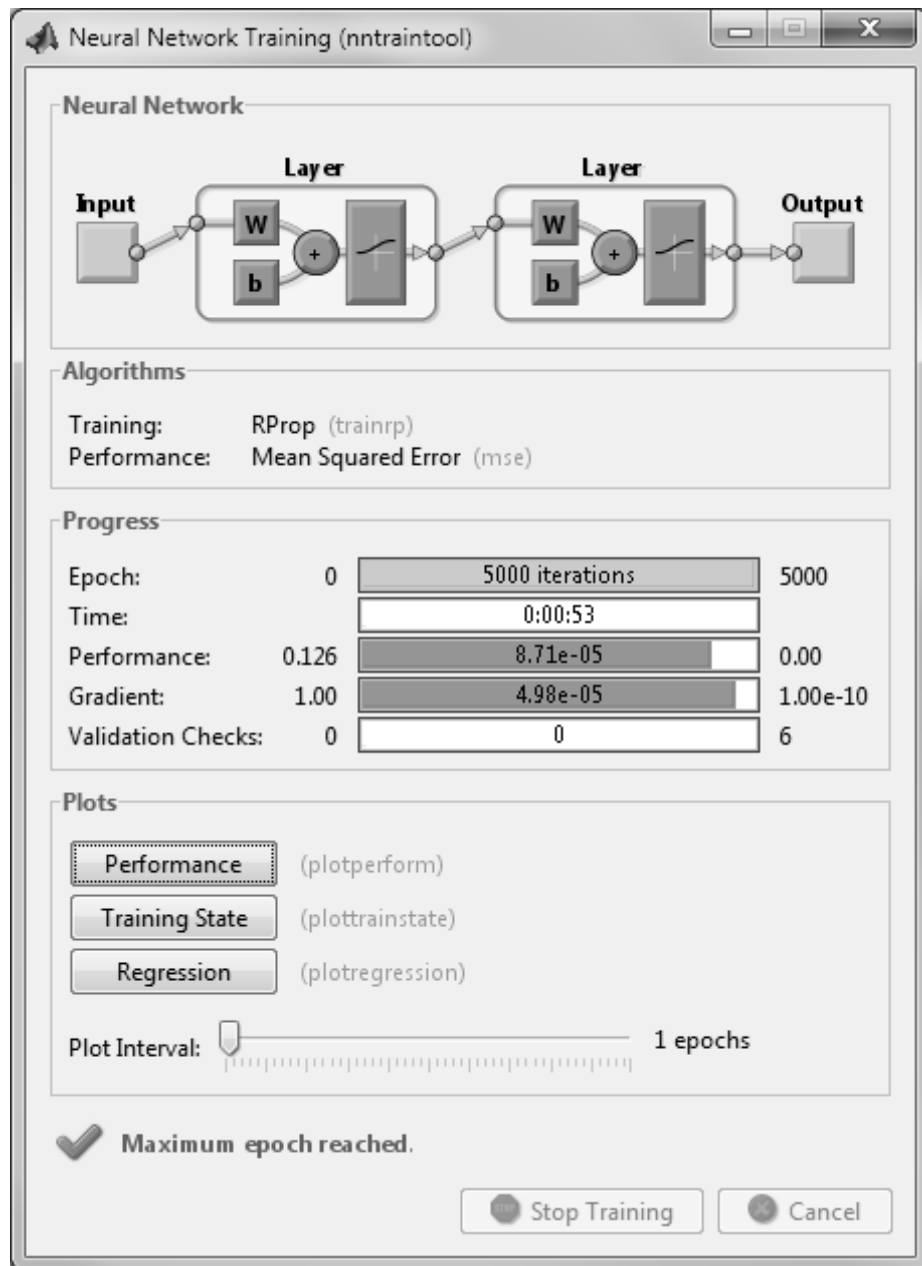


Рис. 4.8 – Интерфейс модуля NNT в режиме обучения

Процесс обучения занял 53 с, среднеквадратическая ошибка (MSE) снизилась примерно до 10^{-4} .

Введем сгенерированный тестовый набор:

```
>> INP_TEST = [1 1 1 0 ... ];
```

Оценим живучесть тестового набора:

```
>> sim(net,INP_TEST);
```

В переменную `ans` система запишет выход оценки ИНС.

На рисунке 4.9 представлен сравнительный анализ результатов оценки живучести для сгенерированных топологий сетевых структур, полученных с помощью ИНС, созданной и обученной на базе ППП NNT, и с помощью точного полиномиального расчета.

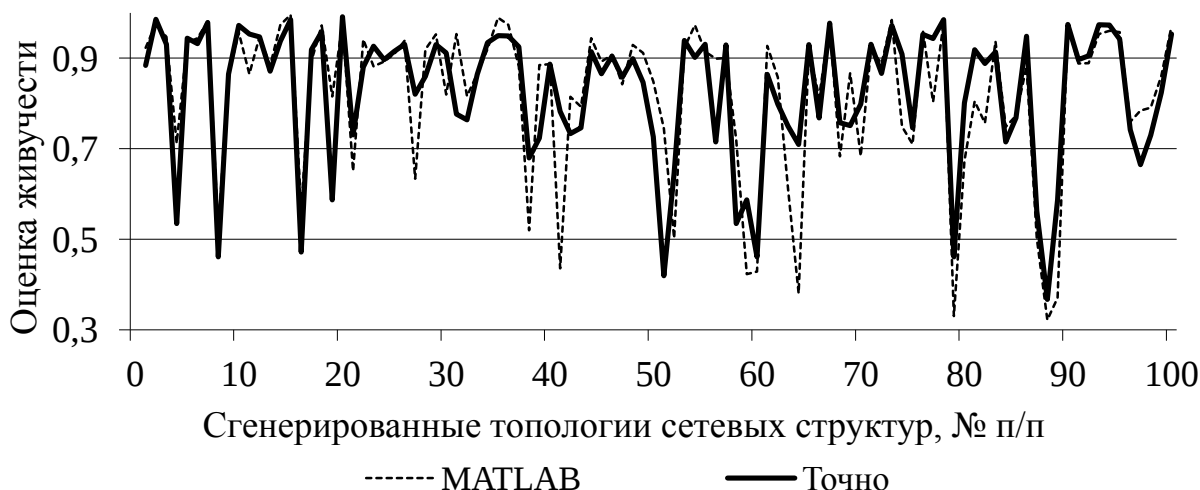


Рисунок 4.9 – Сравнительный анализ результатов оценки живучести

Несмотря на то, что MSE в ППП NNT достаточно мала, как показывает график (рисунок 4.9), обученная ИНС недостаточно адекватна для данного тестового набора.

Для обучения ИНС была выбрана сетевая структура с 7-ю узлами. Для реальных же сетей, где количество узлов измеряется десятками тысяч, система MATLAB является малоэффективной, т.к. имеет ограниченные размеры памяти для работы с матрицами, векторами и другими структурами.

Результаты моделирования показали, что внедрение разработанных моделей в ИС не ухудшает ее живучесть.

Сравнительный анализ результатов ППП NNT и библиотеки FANN показал, что в отличие от бесплатной библиотеки FANN которая при прочих равных условиях вполне адекватно оценивает живучесть, здесь требуется как увеличение обучающих топологий сетевых структур, так и количество эпох обучения, следовательно и дополнительные затраты времени обучения. Но, несмотря на это, система MATLAB является одной из лидирующих систем в

области решения нейросетевых задач ввиду своей популярности, простоты работы с матрицами и векторами, наглядного интерфейса ППП NNT для быстрого создания и обучения нейронных сетей без каких-либо знаний языков программирования.

4.3.2 Исследование вычислительных характеристик применения разработанных моделей

Вычислительные характеристики это, прежде всего, количество машинного времени и памяти которое требуется алгоритму для расчета результата. Для оценки данных характеристик использована выборка объемом 5000 образов, преобразованных в векторы по методике изложенной в []. В качестве сравнения использована одна из традиционных и относительно новых систем распознавания образов, основанная на применении механизма кластеризации [107], работа которой также смоделирована в пакете Matlab [96]. Образы, на основе которых построены обрабатываемые векторы, набраны таким образом, чтобы включать векторы большого количества параметров объекта.

Графики длительности процессов в зависимости от размера выборки (M), по которой ведется поиск, представлены на рисунке 4.10.

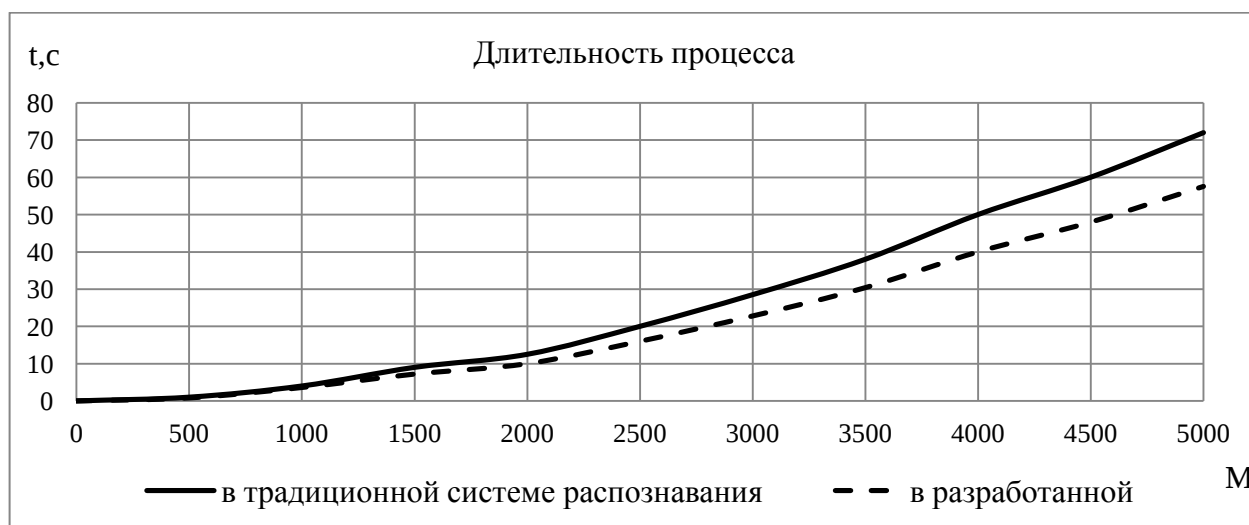


Рисунок 4.10 – Зависимость длительности процесса распознавания (t) от размера выборки (M)

Очевидно, что процесс распознавания значительно быстрее в системе с применением разработанных моделей. При размере поисковой выборки $M=5000$ он достигает 72 секунды для кластерной системы распознавания, и 57 секунд для системы на основе разработанных моделей. При этом требуемый объем памяти для хранения данных меньше примерно в 11 раз. Это обусловлено тем, что для расчета результата распознавания в системах, основанных на динамических моделях наподобие моделей Хопфилда требуется расчет матрицы корреляций векторов прототипов модели. Данный процесс требует наибольшее количество памяти при распознавании. В данном случае, при размере выборки $M=5000$, требуется расчет матрицы размером 5000×5000 , то есть использованная в эксперименте модель распознавания требует при таком размере выборки памяти для хранения 25 000 000 значений с плавающей запятой.

В свою очередь в разработанных моделях требуется произвести 1500 расчетов мер различия, в каждом из которых также требуется считать матрицу корреляций. При оценке сверху требуется хранение $1502 \times 4 \times 4 = 24032$ значений с плавающей запятой плюс $1502 \times 1502 = 2256004$ значений для итогового конкурса распознавания, итого примерно в 11 раз меньший объем памяти для хранения промежуточных данных. Это, естественно, грубая оценка минимума отношений объемов необходимой памяти.

В результате, на данном практическом примере, показано, что процесс распознавания на основе разработанных моделей протекает значительно быстрее и требует меньшего объема памяти для расчетов по сравнению с традиционными динамическими системами распознавания.

4.4 Выводы по главе 4

В результате исследования эффективности применения разработанных моделей доказано, что внедрение их в ИСОН позволяет на 18 % повышает эффективность ее функционирования, а более качественное, по сравнению с

существующими системами, распознавание таких графических объектов, как электроэнцефалограммы, фонокардиограммы, реоэнцефалограммы позволило повысить качество принимаемых заключений судебно-психиатрической экспертизы на 15 %.

В процессе проведения исследования была проведена оценка влияния введенных, в структуру информационной системы, моделей на ее живучесть. Для оценки живучести исследуемой структуры применен подход к оценке живучести внутренней структуры ИС с использованием распределенных информационных систем. Процесс расчета оценки живучести ИС с применением разработанных моделей проведен с использованием системы MATLAB R2010a с ППП NNT версии 6.0.4. Результаты моделирования показали, что внедрение разработанных моделей в ИС не ухудшает ее живучесть.

Сравнительная оценка длительности процесса распознавания графических объектов, при размере выборки $M=5000$ показала, что его длительность достигает 72 секунды для традиционной системы распознавания, и 57 секунд для системы с применением разработанных моделей, а требуемый объем памяти для хранения данных меньше примерно в 11 раз.

ЗАКЛЮЧЕНИЕ

1. Проведенный анализ предметной области на основе системного описания задач распознавания графических объектов в информационных системах, позволил выделить системные компоненты и определить общесистемные характеристики процесса распознавания графических объектов, благодаря чему обеспечивается построение эффективных информационных систем, а также осуществление обоснования и выбор критериев соответствия сведений информационной потребности пользователя.

2. Разработана аналитическая модель распознавания графических объектов, отличающаяся использованием соответствующего множества кубических сплайнов для перевода графического объекта в соответствующий набор признаков и позволяющая, в автоматизированном режиме, определять характеристики, обеспечивающие качество распознавания объектов.

3. Разработана аналитическая модель процесса распознавания графического объекта, отличающаяся учетом понятий полнота и точность поиска, а также взвешенной F -нормы и определяющая требуемую пертинентность результатов процесса.

4. Разработана процедурная модель решения задачи распознавания графических объектов, отличающаяся определением динамических характеристик распознаваемого объекта на основе кубических сплайнов.

5. Оценка эффективности применения, полученных в ходе исследования моделей, показала, что их применение позволяет повысить эффективность ИС образовательного назначения на 18 %, а ИС, применяемых в психиатрии – на 15 %. Оценка живучести ИС с учетом применения разработанных моделей показала, что их внедрение в систему не ухудшает ее живучесть. Сравнительная оценка длительности процесса распознавания графических объектов, при размере выборки $M=5000$ показала, что его длительность достигает 72 секунды для традиционной системы распознавания, и 57 секунд

для системы с применением разработанных моделей, а требуемый объем памяти для хранения данных меньше примерно в 11 раз.

Рекомендации и перспективы дальнейшей разработки темы. Разработанные модели целесообразно применять в информационных системах распознавания графических объектов образовательного и специального назначения для принятия решений, обоснованность которых зависит от качества распознавания графических объектов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей. - М.: Финансы и статистика, 1985. - 487 с.
2. Алгазинов Э.К. Анализ и компьютерное моделирование информационных процессов и систем / Э.К. Алгазинов, А.А. Сирота. М.: Диалог-МИФИ, 2009. - 416 с.
3. Александров А.Г. Оптимальные и адаптивные системы. - М.: Высшая школа. 1986. - 262 с.
4. Аль-Тамими Н. Ю. Оценка важности ресурсов компьютер ной сети / Н. Ю. Аль-Тамими // Вестник Воронежского института высоких технологий. – 2010. – № 6. – С. 36–39.
5. Алексеев В.В., Корыстин С.И., Малышев В.А., Сысоев В.В. Моделирование информационного воздействия на эргатический элемент в эрготехнических системах. М.: Стенсвил, 2003. – 163 с.
6. Amar K. Das. Computers in psychiatry: a review of past programs and an analysis of historical trends. Psychiatric Quarterly, Vol.3, No.4, 2002.
7. Андреев. А.М., Метод кластеризации документов текстовых коллекций и синтеза аннотаций кластеров / А.М. Андреев, Д.В. Берёзкин, В.В. Морозов, К.В. Симаков // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды десятой всероссийской научной конференции– Переславль-Залесский, 2008. – С. 25-31.
8. Анисимов, Б.В. Распознавание и цифровая обработка изображений / Б.В. Анисимов, В.Д. Курганов, В.К. Злобин // М.: Высшая школа, 1983. - 295 с.
9. Афонин А.А., Кластеризация текстовых коллекций: помощь при содержательном поиске и аналитический инструмент /А.А. Афонин, М.Г. Крейнс // Интернет-порталы: содержание и технологии. 2006. – С. 510-537.
10. Батыршин, И.З. Основные операции нечёткой логики и их обобщения. – Казанский ГТУ, 2001 — 102 с.

11. Башмаков А.И., Башмаков И.А. Разработка компьютерных учебников и обучающих систем. – М: Информационно-издательский дом «Филинь», 2003. – 616 с.
12. Безрук, В.М. Теоретические основы проектирования систем распознавания сигналов для автоматизированного радиоконтроля / В.М. Безрук, Г.В. Певцов // Монография. – Харьков: Коллегиум, 2006. – 430с.
13. Бендат Дж., Пирсол А. Применения корреляционного и спектрального анализа. - М.: Мир, 1983. - 312 с.
14. Бесекерский В. А. Цифровые автоматические системы. - М.: Наука, 1976. - 576 с.
15. Блюмин С. Л. Модели и методы принятия решений в условиях неопределенности / С. Л. Блюмини, И. А. Шуйкова. – Липецк: ЛЭГИ, 2001. – 138 с.
16. Бокс, Дж. Анализ временных рядов / Дж. Бокс, Г. Дженкинс // М.: Мир, 1974, вып. 1, 2. - 406 с.
17. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. – М.: Горячая Линия – Телеком, 2007.- 284с.
18. Быстрые алгоритмы в цифровой обработке изображений./ Т. С. Хуанг, Дж-О. Эклунд, Г. Дж. Нуссбаумер и др.; Под ред. Т. С. Хуанга: - М.: Радио и связь, 1984. - 224 с.
19. Вапник, В.Н. Алгоритмы и программы восстановления зависимостей / В.Н. Вапник, Т.Г. Глазкова, В.А. Кошечев // М.: Наука, 1984. – 816 с.
20. Вапник, В.Н. Теория распознавания образов / В.Н. Вапник, А.Я. Червоненкис / М.: Наука, 1974. - 416 с.
21. Влавацкая М.В. Понятие коллокации и коллигации в диахроническом рассмотрении: [Электронный источник] – <http://ffl.nspu.ru/files/konf/konf-2011-15.pdf>.
22. Вольтерра В. Теория функционалов, интегральных и интегро-дифференциальных уравнений. - М.: Наука, 1982. - 304 с.

23. Выгодский М.Я. Справочник по высшей математике — М: Наука, 2002. – 870 с.
24. Галушкин А.И. Нейроматематика (Проблемы развития). М.: Радиотехника, 2003. 234 с.
25. Галушкин А.И. Решение задач в нейросетевом логическом базисе. // Нейрокомпьютеры. 2006. №2. С .49–70.
26. Гонсалес Р., Вудс Р. Цифровая обработка изображений Москва: Техносфера, 2005. - 1072 с.
27. Горелик А.Л., Скрипкин В.А. Методы распознавания. 2-е издание. – М : Высшая школа, 1984. – 864 с.
28. ГОСТ 7.73-96 Система стандартов по информации, библиотечному и издательскому делу. Поиск и распространение информации. Термины и определения :[Электронный источник] – <http://www.complexdoc.ru/lib/ГОСТ%207.73-96>.
29. Грейди Буч, Джеймс Рамбо, Айвар Джекобсон. Язык UML. Руководство пользователя. М., СПб.: ДМК Пресс, Питер, 2004.
30. Громов Ю.Ю. К проблеме оценки живучести сетевых информационных систем с использованием элементов искусственного интеллекта / Ю. Ю. Громов, О. Г. Иванова, К. А. Набатов, Д. Е. Винокуров // Системы управления и информационные технологии. – 2007. – № 3.2 (29). – С. 232–237.
31. Громов Ю.Ю. Управление информационными процессами в условиях неопределенности / Ю.Ю. Громов, В.Е. Дидрих, О.Г. Иванова, Ю.В. Минин, А.Ю. Громова // Информация и безопасность. – Воронеж: Издательско-полиграфический центр Воронежского государственного университета. – 2011. – № 2. – С. 233–238.
32. Громов Ю.Ю., Набатов К.А., Калинин В.Ф., Сербулов Ю.С., Драчев В.О. Распределение ресурсов сетевых электротехнических систем М.: Машиностроение, 2008.- 214 с.
33. Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И.,

Спектор А.А. Цифровая обработка изображений в информационных системах: Учебное пособие.- Новосибирск: Изд-во НГТУ, 2002. – 352 с.

34. Гонсалес Р., Вудс Р. Цифровая обработка изображений.- Москва: Техносфера, 2005. – 1072 с.

35. Гудвин, Дж. Теория управления: идентификация и оптимальное управление / Дж. Гудвин, К. Спири, Р. Браун // М.: Мир, 1973. – 248 с.

36. Данилкин С.В. Гносеологические подходы к решению задач классификации с позиции теории нечетких множеств // Информационные системы и процессы: сб. научных тр. Тамбов; М.; Баку; Вена: Изд-во «Нобелистика»; 2006, Вып.4. – С.41-49.

37. Дейч А.М. Методы идентификации динамических объектов. – М.: Энергия, 1979. - 240с.

38. Додонов А.Г., Ландэ Д.В. Живучесть информационных систем. – К.: Наук. думка, 2011. – 256 с.

39. Долгов А.А. Создание и обучение искусственной нейронной сети с помощью средства разработки Qt и библиотеки FANN / А.А. Долгов // Вестник воронежского института высоких технологий. Воро- неж, гос. ун-т. – 2010. – Вып. 6. – С. 244–247.

40. Дрейпер Н., Смит Г., Прикладной регрессионный анализ: В 2-х кн. Кн. 1,2./ Пер. с англ. - 2-е изд., перераб. и доп. - М.: Финансы и статистика, 1986. - 366 с.

41. Дуда, Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт / М.: изд-во «Мир», 1978. - 512 с.

42. Живенков А.Н. Разработка модели структуры курса обучения в условиях неопределенности. - Вестник Воронежского института высоких технологий, 2009 №5 – С.123-127

43. Завьялов Ю.С. Методы сплайн-функций / Ю.С. Завьялов, Б.И. Квасов, В.Л. Мирошниченко. – М.: Наука, 1980. – 355с.

44. Заде Л. Понятие лингвистической переменной и её применение к принятию приближённых решений — М: МИР, 1973. – 167 с.

45. Зимарин А.М., Алексеев В.В. Отбор информативных признаков в задачах распознавания образов // Сборник докладов IX Всероссийской научно-технической конференции «Повышение эффективности средств обработки информации на базе математического моделирования», Тамбов: ТВВАИУРЭ. С. 176–181.
46. Идентификация и диагностика систем: учеб. для студ. высш. учеб. заведений/ А.А. Алексеев, Ю.А. Кораблев, М.Ю. Шестопалов. - М.: Издательский центр «Академия», 2009. - 352 с.
47. Кашьяп, Р.Л. Построение динамических стохастических моделей по экспериментальным данным / Р.Л. Кашьяп, А.Р. Рао // М.: Наука, 1983. - 384 с.
48. Кендал М. Временные ряды. - М.: Радио и связь, 1981. - 198 с.
49. Кини Р. Л. Принятие решений при многих критериях: предпочтения и замещения / Р.Л. Кини, Х. Райфа. – М.: Радио и связь, 1981. – 560 с.
50. Клир Дж. Системология. Автоматизация решения системных задач / Пер. с англ. – М.: Радио и связь, 1990. – 544 с.
51. Когаловский М.Р. Перспективные технологии информационных систем. — М.: ДМК Пресс; М: Компания АйТи, 2003. – 288 с.
52. Ландэ Д.В. Интернетика: Навигация в сложных сетях: модели и алгоритмы. – М: Издательство ЛИБРОКОМ, 2009. – 264 с.
53. Ланкастер П. Теория матриц. М.: Наука, 1978. – 280 с.
54. Ларичев О.И. Объективные модели и субъективные решения / О.И. Ларичев. – М.: Наука, 1987. – 141 с.
55. Леденева Т.М. Обработка нечеткой информации: учебное пособие / Т.М. Леденева. – Воронеж: Воронежский государственный университет, 2006. – 233 с.
56. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А.В. Леоненков. – СПб.: БХВ-Петербург, 2005. – 736 с.
57. Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов./ Пер. с англ. - М.: Наука, Гл. ред. физ.-мат. лит., 1986. – 232 с.

58. Лыонг Хак, Д. Оценка надежности средств парирования внешних воздействий / Д. Лыонг Хак, Н.Ю. Аль-Тамими, М.П. Аль-Балуши, М. Ауад, Ю.В. Минин // Техника и безопасность объектов уголовно-исполнительной системы-2011: сборник материалов Международной научно-практической конференции: в 2 т./ФКОУ ВПО Воронежский институт ФСИН России. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2011. – С. 394 – 399.

59. Лыонг Хак, Д. Аналитическая модель распознавания информационного процесса в системе поддержки принятия решений / Д. Лыонг Хак, Ю.В. Минин, В.В. Алексеев, Ю.Ю. Громов.// Вестник Воронежского института МВД России. - 2013. – № 4. - С. 221 – 225.

60. Лыонг Хак, Д. Интеллектуальная информационная система оценки устойчивости функционирования сетевых информационных систем / Д. Лыонг Хак, Н.Ю. Аль-Тамими, М.П. Аль-Балуши, М. Ауад, Ю.В. Минин // Техника и безопасность объектов уголовно-исполнительной системы-2011: сборник материалов Международной научно-практической конференции: в 2 т./ФКОУ ВПО Воронежский институт ФСИН России. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2011. – С. 382 – 388.

61. Лыонг Хак, Д. Интеллектуальная система обучения с адаптивным построением курса обучения на LMS Moodle / Д. Лыонг Хак, Т.О. Авдеева, М. Ауад, М. Аль-Балуши // Приборы и системы. Управление, контроль, диагностика. – 2012. - № 3 – С. 71 – 74.

62. Лыонг Хак, Д. Использование математического аппарата нечеткой логики для определения пертинентности результатов поиска текстовых сведений / Д. Лыонг Хак, Д.В. Поляков, М.П. Аль-Балуши, М. Ауад // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 163.

63. Лыонг Хак, Д. К вопросу о моделировании процесса распределения ресурсов в информационных системах. / Д. Лыонг Хак, К.А. Набатов, Ю.В. Минин, С.Н. Коршиков // Информация и безопасность. - Воронеж:

Издательско-полиграфический центр Воронежского государственного университета. - 2012. – № 4.- С. 461 – 470.

64. Лыонг Хак, Д. К вопросу о показателях живучести сетевых информационных систем / Д. Лыонг Хак, А.И. Елисеев, М.А. Хорохорин, А.А. Долгов // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 74.

65. Лыонг Хак, Д. К вопросу об управлении пространственно распределенными информационными системами / Д. Лыонг Хак, А.А. Долгов, М.А. Хорохорин, Ю.В. Минин, М. Аль-Балуши // Актуальные проблемы деятельности подразделений УИС: сборник материалов Всероссийской научно-практической конференции: в 2 т. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2013. - С. 54 – 58.

66. Лыонг Хак, Д. К вопросу оценки живучести сложных систем с использованием распределенных систем на основе графовой модели / Д. Лыонг Хак, А.И. Елисеев, А.А. Долгов, М.А. Хорохорин // Прикладная математика, управление и информатика: сборник научных работ, 3 – 5 октября 2012 г. – Белгород: БелГУ. – 2012. - Т. 2. – С. 287 – 292.

67. Лыонг Хак, Д. К вопросу построения математической модели кластеризации текстовых сведений / Д. Лыонг Хак, Д.В. Поляков, В.В. Самойлов, М.П. Аль-Балуши // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – С. 164.

68. Лыонг Хак, Д. Некоторые аспекты интеллектуальной информационной системы оценки функционирования сетевых информационных систем / Д. Лыонг Хак, Н.Ю. Аль-Тами, М.П. Аль-Балуши, М. Ауад, Ю.В. Минин // Техника и безопасность объектов уголовно-исполнительной системы-2011: сборник материалов Международной научно-

практической конференции: в 2 т./ФКОУ ВПО Воронежский институт ФСИИ России. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2011. – С. 388 – 394.

69. Лыонг Хак, Д. Обзор современных систем распознавания образов / П.И. Карасев, Д. Лыонг Хак, А.А. Лебедев, В.В. Алексеев // Техника и безопасность объектов уголовно-исполнительной системы: Материалы Международной научно-практической конференции, 3-4 октября 2013 г./ Воронежский институт ФСИИ России – Воронеж: ИПЦ «Научная книга». 2013. –С. 151-153.

70. Лыонг Хак, Д. Определение информационно-управляющих воздействий в системах с использованием качественной информации / Д. Лыонг Хак, Ю. В. Минин, О. Г. Иванова, В. В. Алексеев // Современные информационные технологии. – Пенза: Издательство: Пензенская государственная технологическая академия. – 2013. - № 17 - С. 117 - 121.

71. Лыонг Хак, Д. Определение исходных данных задачи синтеза оптимальных логических структур сетевых баз данных / Д. Лыонг Хак, Е.В. Костерин, А.С. Моисеев // Прикладная математика, управление и информатика: сборник научных работ, 3 – 5 октября 2012 г. – Белгород: БелГУ. – 2012. - Т. 2. – С. 425 - 428.

72. Лыонг Хак, Д. Определение целевой функции поиска текстовых сведений в сети Интернет / Д. Лыонг Хак, Д.В. Поляков, М.П. Аль-Балуши, М. Ауад // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 168.

73. Лыонг Хак, Д. Оценка живучести сетевых структур с использованием системы Matlab / Д. Лыонг Хак, А.А. Долгов, О.Г. Иванова, М. Ауад // Вестник Воронежского института ФСИИ России. - Воронеж: ООО ИПЦ «Научная книга». – 2011. - № 1. – С. 56 - 59.

74. Лыонг Хак, Д. Подход к формированию набора информационных ресурсов образовательного назначения на основе нечетких сетей Петри / Д.

Льонг Хак, О.Г. Иванова, Т.О. Авдеева, М. Ауад, М. Аль-Балуши // Приборы и системы. Управление, контроль, диагностика. – 2012. - № 3 – С. 55 - 58.

75. Льонг Хак, Д. Построение пертинентного запроса к информационно-поисковой машине на основе математического аппарата нечеткой логики / Д. Льонг Хак, Д.В. Поляков, В.В. Самойлов, М. Ауад // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 167.

76. Льонг Хак, Д. Практическое использование экспертных систем в психиатрии / Д. Льонг Хак, В.А. Перфильев, А.В. Клишина, М.Г. Селезнев, А.В. Толмачев // Актуальные проблемы деятельности подразделений УИС: сборник материалов Всероссийской научно-практической конференции: в 2 т. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2013. - С. 136 – 139.

77. Льонг Хак, Д. Применение нечетких сетей Петри для решения задачи адаптивного структурирования информационных ресурсов / Д. Льонг Хак, Т.О. Авдеева, К.М. Копылов, О.Г. Иванова // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 3.

78. Льонг Хак, Д. Применение распределенных информационных систем для оценки живучести нечетких графов / Д. Льонг Хак, М.А. Хорохорин, А.А. Долгов, М. Ауад // Информация и безопасность. - Воронеж: Издательско-полиграфический центр Воронежского государственного университета. - 2012. – № 2. Том 15. - С. 245 – 248.

79. Льонг Хак, Д. Применение цепных дробей для оценки живучести сетевых информационных систем в условиях неопределенности / Д. Льонг Хак, А.И. Елисеев, М.П. Аль-Балуши, М. Ауад // Математические методы и информационно-технические средства: Труды VIII Всероссийской научно-практической конференции, 22-23 июня 2012 г. – Краснодар: Краснодарский университет МВД России, 2012. – 278 с. – С. 71.

80. Лыонг Хак, Д. Процедура идентификации объекта в интеллектуальной информационной системе / Д. Лыонг Хак, В.А. Горин, Ю.В. Минин, В.В. Алексеев // Техника и безопасность объектов уголовно-исполнительной системы: Материалы Международной научно-практической конференции, 3-4 октября 2013 г./ Воронежский институт ФСИН России – Воронеж: ИПЦ «Научная книга». 2013. –С. 182-183.

81. Лыонг Хак, Д. Процедура принятия решений в информационной тренажерной системе при условии ограниченного времени прохождения / Д. Лыонг Хак, А.С. Моисеев, М. Аль-Балуши, Ю.В. Минин // Актуальные проблемы деятельности подразделений УИС: сборник материалов Всероссийской научно-практической конференции: в 2 т. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2013. - С. 106 – 109.

82. Лыонг Хак, Д. Разработка экспертной системы для судебной психиатрии (часть 1) / Д. Лыонг Хак, В.А. Перфильев, А.В. Клишина, М.Г. Селезнев, А.В. Толмачев // Актуальные проблемы деятельности подразделений УИС: сборник материалов Всероссийской научно-практической конференции: в 2 т. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2013. - С. 140 – 143.

83. Лыонг Хак, Д. Разработка экспертной системы для судебной психиатрии (часть 2) / Д. Лыонг Хак, В.А. Перфильев, А.В. Клишина, М.Г. Селезнев, А.В. Толмачев // Актуальные проблемы деятельности подразделений УИС: сборник материалов Всероссийской научно-практической конференции: в 2 т. – Воронеж: ИПЦ «Научная книга». – Т. 1. – 2013. - С. 144- 147.

84. Лыонг Хак, Д. Управление информационными процессами в системах с использованием качественной информации / Д. Лыонг Хак, Ю.В. Минин, А.Ю. Громова, В.В. Родин, М. Ауад // Вестник Воронежского института ФСИН России. - Воронеж: ООО ИПЦ «Научная книга». – 2011. - № 2 – С. 64 – 68.

85. Лыонг Хак, Д. Управление информационными процессами при нечеткой информации / Д. Лыонг Хак, Ю.В. Минин, А.Ю. Громова, В.В. Родин,

М. Ауад // Вестник Воронежского института ФСИН России. - Воронеж: ООО ИПЦ «Научная книга». – 2011. - № 1. – С. 71 – 75.

86. Лыонг Хак, Д. Формализация информационной потребности с помощью коллокаций на основе теории нечетких множеств для пертинентного поиска текстовых сведений. / Д. Лыонг Хак, Ю.Ю. Громов, Д.В. Поляков, П.Б. Шихук // Информация и безопасность. - Воронеж: Издательско-полиграфический центр Воронежского государственного университета. - 2012. – № 2. Том 15. - С. 213 - 218.

87. Люгер Д.Ф. «Искусственный интеллект: стратегии и методы решения сложных проблем», 4-е издание. Пер. с англ. – М.: Издательский дом «Вильямс», 2003.–864 с.

88. Мартынов Н. Н. Matlab 7. Элементарное введение / Н. Н. Мартынов. – М. : Кудиц-Образ, 2005. – 416 с.

89. Manning C., Raghavan P., Schütze H. Introduction to Information Retrieval. — Cambridge University Press, 2008.

90. Медведев В. С. Нейронные сети. MATLAB 6 / В. С. Медведев, В. Г. Потемкин. – 2002. – 496 с.

91. Мелихов А.Н. Ситуационные советующие системы с нечеткой логикой / А.Н. Мелихов, Л.С. Берштейн, С.Я. Коровин. – М.: Наука. Гл. ред. физ.-мат. лит., 1990. – 272 с.

92. Месарович М., Такахара Я. Общая теория систем: математические основы. – М.: Мир, 1978. – 311 с.

93. Методы анализа и синтеза модульных информационно-управляющих систем / Н.А. Кузнецов [и др.]. – М.: ФИЗМАТЛИТ, 2002. – 800 с.

94. Методы классической и современной теории автоматического управления: Учебник в 5-и тт.; 2-е изд., перераб. и доп. Т2: Статистическая динамика и идентификация систем автоматического управления / под ред. К. А. Пупкова и Н.Д. Егупова. - М.: Издательство МГТУ им. Н.Э. Баумана, 2004. – 646 с.

95. Минин Ю.В. Информационная система поддержки принятия решений, оптимизирующих коммерческую деятельность банка: диссертация ... кандидата технических наук: 05.25.05 / Минин Юрий Викторович. Тамбов, ТГТУ, 2008. – 130 с.

96. Надежность информационных систем: учебное пособие. Рекомендовано УМО вузов по университетскому политехническому образованию / Ю.Ю. Громов, О.Г. Иванова, Н.Г. Мосягина, К.А. Набатов – Тамбов: Изд-во ГОУ ВПО ТГТУ, 2010. – 160 с.

97. Небабин, В.Г. Методы и техника радиолокационного распознавания / В.Г. Небабин, В.В. Сергеев // М.: Радио и связь, 1984. – 152 с.

98. Недошивина Е.В., Учет синтаксических связей при поиске коллокаций. // Семинар: Natural Language Processing; 2008 – С. 32-34.

99. Нечеткие множества в системах управления и искусственного интеллекта/ Под ред. Д.А. Поспелова М.: Наука, 1986. – 312 с.

100. Орловский С.А. Проблемы принятия решений при нечеткой исходной информации / С.А. Орловский – М.: Наука, 1981 – 437с.

101. Патрик, Э. Основы теории распознавания образов / Э. Патрик // М.: Сов. радио, 1980 – 408 с.

102. Поляков Д. Математические модели и алгоритмы эффективного поиска текстовой информации на основе кластеризации по нечётким коллокациям, 2013 -

103. Попов, Э. В. Искусственный интеллект. Модели и методы / Э. В. Попов, Д. А. Поспелов, В. Н. Захаров, В. Ф. Хорошевский // М.: Радио и связь, 1990. - 304 с.

104. Попов, Э. В. Искусственный интеллект. Системы общения и экспертные системы. / Э. В. Попова, Д. А. Поспелова, В. Н. Захарова, В.Ф. Хорошевского // М.: Радио и связь, 1990. – 290 с.

105. Поспелов, Г.С. Процедуры и алгоритмы формирования комплексных программ / Г.С. Поспелов, В.А. Ириков // М.: Наука, 1985. – 422 с.

106. Поспелов, Г.С. Искусственный интеллект прикладные системы. /

Г.С. Поспелов, Д.А. Поспелов // М.: Знание, 1985. - № 9. – 48 с.

107. Поспелов, Д.А. Нечёткие множества в моделях управления и искусственного интеллекта. / Д.А. Поспелов // М.: Наука, 1986. – 128 с.

108. Поспелов, Д.А. Логико-лингвистические модели в системах управления. / Д.А. Поспелов // М.: Энергоатомиздат, 1981. – 224 с.

109. Поспелов, Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов / Д.А. Поспелов // М.: Радио и связь, 1989. - 184 с.

110. Питерсон Дж. Теория сетей Петри и моделирование систем. – М.: Мир, 1984 – 368 с.

111. Райбман Н.С. Что такое идентификация? - М.: Наука, 1970. -246 с.

112. Райфа Г. Анализ решений. – М.: Наука, 1977. -402 с.

113. Рассел С., Норвиг П., «Искусственный интеллект: современный подход (AIMA)», 2-е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1424 с.

114. Сейдж Э.П., Мелса Дж.Л. Идентификация систем управления. - М.: Наука, 1974. - 248 с.

115. Семенов А.Д., Артамонов Д.В., Брюхачев А.В. Идентификация объектов управления: Учебн. пособие. - Пенза: Изд-во Пенз. гос. ун-та, 2003. – 211 с.

116. Современные методы идентификации систем/Под ред. П. Эйкхоффа. - М.: Мир, 1983. – 400 с.

117. Сорокин А.И. Разработка алгоритмов распознавания рукописных символов на основе аналитических свойств изображения. Диссерт. на соискание уч. ст. к.ф-м.н., 2010. –153 с.

118. Сысоев В.В., Сербулов Ю.С., Сипко В.В. Теоретико-игровые модели принятия решений многоцелевого управления в задачах выбора и распределения ресурсов/Воронеж. гос. технол. акад. – Воронеж, 2000. – 60 с.

119. Теоретические основы проектирования оптимальных структур распределенных баз данных / В.В. Кульба [и др.]. – М.: Синтег, 1999. – 660 с.

120. Типовые линейные модели объектов управления / С.А. Анисимов,

И.С. Зайцева, Н.С. Райбман, А.А. Яралов; под ред. Н.С. Райбмана. – М.: Энергоатомиздат, 1983. – 264 с.

121. Трухаев Р. И. Модели принятия решений в условиях неопределенности / Р. И. Трухаев. – М. : Наука, 1981. – 258 с.

122. Takagi T. Fuzzy identification of systems and its application to modeling and control / T. Takagi, M. Sugeno // IEEE Trans. Systems, Man and Cybernetics. 1985. - №15(1). - P. 116—132.

123. Филлипс Д. Методы анализа сетей: Пер. с англ. / Д. Филлипс, А. Гарсиа-Диас. – М.: Мир, 1984. – 496 с.

124. Фисенко В.Т., Фисенко Т.Ю. Компьютерная обработка и распознавание изображений/ Учеб. пособие. - СПб: СПбГУ ИТМО, 2008. – 192 с.

125. Фукунага, К. Введение в статистическую теорию распознавания образов / К. Фукунага // М.: Наука, 1979. - 368 с.

126. Шелухин О.И. Моделирование информационных систем / О.И. Шелухин, А.М. Тенякшев, А.В. Осин. - М.: Радиотехника, 2005. - 368 с.

127. Шенк, Р. Обработка концептуальной информации / Р. Шенк // М.: Прогресс, 1980 - 365 с.

128. http://ru.wikipedia.org/wiki/Информационный_поиск#.D0.92.D0.B8.D0.B4.D1.8B_.D0.BF.D0.BE.D0.B8.D1.81.D0.BA.D0.B0 .

129. <http://www.bibliofond.ru/view.aspx?id=135700> .

130. <http://ru.wiktionary.org/wiki/> .

131. <http://www.mirea.ru/files/abstract/yurgaev.doc> .

132. <http://www.bogoslov.ru/text/321597.html> .

133. http://www.benran.ru/Magazin/cgi-bin/Sb_95/pr_95.exe?!33 .

134. <http://poiskbook.kiev.ua/nti05.html> .

135. <http://www.likt590.ru/project/poisk/to.htm> .

136. <http://www.cph.phys.spbu.ru/documents/First/infoPMF/AlgInt1.pdf> .

137. <http://nashol.com/201009214567/>

138. <http://mgup.ru/public/files/153.doc> .

139. http://www.gumfak.ru/bib_html/evristka/evr04.shtml .
140. <http://inftis.narod.ru/is/is-n8.htm> .
141. <http://gpntb.ru/win/book/1/Doc17.HTML> .

Копии актов использования результатов исследования



подтверждаю
 Исследования в директор
 ОАО «Меттехника»
 Проскуряков В.С.
 30 декабря 2013 года

АКТ
 использования результатов диссертационной работы
ЛЫОНГ ХАК ДИНЬ
 на тему: «Аналитические и процедурные модели для информационной системы
 распознавания графических объектов в условиях неопределенности»

Комиссия в составе: председателя – главного инженера Проскурякова Сергея Викторовича; и члена – заместителя главного инженера Ченчик Людмилы Николаевны, составила настоящий Акт о том, что результаты диссертационной работы ЛЫОНГ ХАК ДИНЬ на соискание ученой степени кандидата технических наук, а именно программная реализация построенных

- аналитических моделей распознавания графических объектов и процесса поиска сведений;
- процедурной модели решения задачи распознавания графических объектов

использованы при разработке концепции и структуры информационной интеллектуальной системы, с применением классификаций психических расстройств МКБ-10 для автоматизации наиболее трудоемких задач: принятия решения о необходимости постановки пациента на активное диспансерное наблюдение и типе принудительного лечения; оценки риска общественной опасности пациента по результатам клинических исследований.

Программная реализация разработанных автором моделей позволила автоматизировать процесс распознавания характерных участков таких графических объектов, как электроэнцефалограммы, фонокардиограммы, реоэнцефалограммы, позволила повысить качество принимаемых заключений судебно психиатрической экспертизы на 15 % и существенно сократить время принятия решения экспертом.

Председатель комиссии:

член комиссии:

30 декабря 2013 года



С.В. Проскуряков

Л.Н. Ченчик

Приложение Б (продолжение)

«УТВЕРЖДАЮ»



Министерство образования и науки
Российской Федерации

Федеральное государственное бюджетное
образовательное учреждение
высшего профессионального образования
ФГБОУ ВПО «ТГТУ»

«Тамбовский государственный
технический университет»
392000, Тамбов, Советская, 106
Телеграфный адрес: Тамбов, ТГТУ,
телефон: 631019, факс: (4752) 63-06-43
E-mail: tsu@admin.tstu.ru
ОГРН 1026801156557
ИНН 6831006362, ОКПО 02069289

ИО ректора ФГБОУ ВПО «ТГТУ»,
Заслуженный деятель науки РФ,
д.т.н., профессор



С.И. Дворецкий

« 27 » декабря 2013 г.

№ _____ от _____

На № _____ от _____

АКТ

об использовании научных результатов диссертации Лыонг Хак Динь,
представленную на соискание ученой степени кандидата технических наук по
специальности 05.25.05 – Информационные системы и процессы

Комиссия в составе: председателя – директора Института автоматизации и информационных технологий д.т.н., профессора Громова Ю.Ю. и членов комиссии: заведующей кафедрой Информационных систем и защиты информации к.т.н., доцента Ивановой О.Г.; профессора кафедры Информационных систем и защиты информации д.т.н., профессора Дидриха В.Е. составили настоящий акт об использовании результатов исследования, полученных в диссертационной работе Лыонг Хак Динь.

Основные научные результаты работы внедрены в учебный процесс Тамбовского государственного технического университета по кафедре «Информационные системы и защита информации» при изучении дисциплин:

Модели и методы интеллектуального анализа данных, специальность 230402.68; Компьютерная геометрия и графика 230201.62, а также при выполнении курсовых и выпускных квалификационных работ, разработке и применении автоматизированных программ контроля качества формирования у студентов профессиональных компетенций.

Приложение Б (продолжение)

Применение разработанных в диссертации моделей и их программной реализации способствовало повышению эффективности освоения компетенций студентами, в частности, применение разработанных моделей в информационной системе образовательного назначения, реализованной в программной среде MOODLE, существенно (на 12 %) повысило ее эффективность за счет улучшения распознавания графических объектов при оценке качества формируемых компетенций.

ПРЕДСЕДАТЕЛЬ КОМИССИИ



Ю.Ю. Громов

ЧЛЕНЫ КОМИССИИ:



О.И. Иванова

В.Е. Дидрих