

Применение анализа формальных понятий для интеллектуальной поддержки принятия решений

К. И. Пахомова¹, А. В. Коробко¹

¹Сибирский Федеральный Университет, г. Красноярск, Россия

¹Институт вычислительного моделирования СО РАН, г. Красноярск, Россия

Аннотация. Статья посвящена изучению и систематизации современных направлений интеллектуальной поддержки принятия решений, использующих метод анализа формальных понятий (АФП). Рассмотрен опыт современных исследователей применения АФП для классификации, формирования рекомендаций, построения онтологий и информационного поиска. Описаны основные категории решаемых задач, используемые свойства алгебраической решетки и возможности метода АФП. Определены основные преимущества применения метода для развития авторской теории интеллектуальной поддержки оперативной аналитической обработки гетерогенных данных на основе интегральной OLAP-модели.

Ключевые слова: искусственный интеллект, анализ формальных понятий, OLAP, поддержка принятия решений, аналитическая модель.

DOI 10.14357/20718594190405

Введение

Современные тренды развития информационной инфраструктуры непрерывно расширяют цифровое пространство, как на территории нашей страны, так и за рубежом. Вычислительные и емкостные мощности современных устройств позволяют накапливать большие объемы разнородной информации, полученной в ходе производственной или повседневной человеческой деятельности. Необходимость оперативного доступа к желаемой информации увеличивает нагрузку на вычислительные машины и увеличивает время ожидания обработки пользовательских запросов. По этой причине ученые разрабатывают новые подходы для структурирования и анализа данных больших объемов и разных форматов. Повышение эффективности поддержки принятия решений

путем оперативной аналитической обработки и предоставления релевантной информации из больших массивов разнородных данных на текущее время является весьма актуальной задачей развития информационных технологий.

Методы искусственного интеллекта хорошо зарекомендовали себя во многих областях деятельности человека, требующих интеллектуальной поддержки принятия решений: маркетинг, медицинская диагностика, распознавание образов и др. Интеллектуальные системы становятся умными ассистентами, сопровождающими и направляющими пользователя. Среди множества математических подходов области искусственного интеллекта активно набирает популярность метод анализа формальных понятий (АФП) [3], который основывается на классической теории алгебраических решеток. Анализ формальных понятий совмещает в себе мощный математический аппарат и философию

✉ Пахомова Кристина Игоревна. E-mail: krahamova@yandex.ru

фию, которые нашли широкое применение в решении прикладных задач. Стремительное развитие подхода обусловлено изящностью теории и простотой реализации алгоритмов. Данный метод получил широкое применение в задачах классификации, информационного поиска, построения онтологий, формирования рекомендаций.

Одним из новых направлений применения АФП стала теория интегрального OLAP-моделирования гетерогенных данных. Авторами статьи [1] предложен подход к построению интегральной OLAP-модели, объединяющей данные из разнородных источников и представляющей их в виде единой решетки кубов-понятий. Кубы-концепты решетки содержат группы показателей и измерений из разных источников, которые гарантированно могут быть совместно проанализированы. Решетка обеспечивает компактное представление больших данных, в части разнообразия структур и форматов, и создает основу для выполнения произвольных аналитических запросов к интегральной OLAP-модели. В настоящее время актуальной является задача разработки методов и алгоритмов интеллектуальной поддержки оперативной аналитической обработки разнородных данных на основе интегральной OLAP-модели, представленной в виде решетки.

Целью настоящей статьи является определение направлений для развития авторского метода интегрального аналитического моделирования. Для достижения поставленной цели был выполнен содержательный анализ и систематизация существующих задач поддержки принятия решений, использующих АФП, и определены возможности метода и его ограничения для интеллектуальной поддержки оперативной аналитической обработки разнородных данных на основе интегральной OLAP-модели.

Статья состоит из двух частей, где в первой приводится краткое введение в АФП, а именно определяются формальный контекст и формальные понятия. Во второй части описываются основные задачи поддержки принятия решений, в которых АФП нашел свое применение. В каждом примере анализируются основные свойства решетки понятий и оцениваются преимущества применения подхода к задаче интеллектуальной поддержки принятия решений на основе интегральной OLAP-модели, построенной с применением теории АФП. В Заключе-

нии формулируются основные направления развития авторской теории интеллектуальной поддержки анализа больших данных с использованием преимуществ теории АФП.

1. Анализ формальных понятий

Анализ формальных понятий (АФП) [3] — это метод анализа данных, который является прикладной ветвью алгебраической теории решеток. АФП позволяет пользователю анализировать зависимости между подмножествами объектов и признаков. В качестве отношения «объект - признак» могут выступать данные, которые описывают любую область человеческой деятельности. Пары взаимосвязанных множеств объектов и признаков образуют решетку понятий, частично упорядоченную отношением «быть более общим понятием чем».

С точки зрения философии науки, формальное понятие является естественным способом описания группы объектов через их характерные свойства, которое определяется объемом — величиной всех объектов, принадлежащих понятию и содержанием — совокупностью всех признаков, общих для объектов [2-4]. Первая задача АФП заключается в формировании формального контекста в виде тройки из множеств объектов, признаков и отношений их инцидентности. Обычно формальный контекст представляют в виде объектно-признаковой таблицы, в которой столбцам соответствуют признаки, а строками — объекты. Наличие крестика в таблице на пересечении столбца и строки означает вхождение соответствующей пары «объект, признак» в отношение инцидентности. Воздействуя на контекст оператором Галуа, представленным парой отображений объектов на признаки, и признаков на объекты, получаем (формальные) понятия. Для упорядочивания понятий строится решетка понятий. Вершинами решетки соответствуют (формальные) понятия, с соответствующими им объемами и содержаниями, а ребрами — отношения соседства между самими понятиями.

Адаптивность и актуальность представленного Б. Гантером и Р. Вилле [3] математического аппарата отражена в работах современных ученых, занимающихся развитием искусственного интеллекта, математики, логики и т.д. Помимо устоявшегося теоретического аппарата данный подход нашел свое применение в практических задачах.

На сегодняшний день данное направление активно развивается учеными разных стран. В России научная школа АФП сложилась под руководством профессора С.О. Кузнецова [5-8, 12, 14] на базе Национального исследовательского университета «Высшая школа экономики». В [5, 6, 12, 14] приводится подробное описание построения решеток формальных понятий и их использование на прикладном уровне. Авторы [9] агрегировали труды ученых в совершенно разных направлениях исследований и показали многообразие решаемых задач с использованием АФП путем построения решетки понятий, в которой в качестве объектов выступают научные работы, а признаками являются современные подходы искусственного интеллекта. Теория АФП успешно применяется для классификации, информационного поиска, построения онтологий, формирования рекомендаций и прочие [5-7].

2. Разнообразие задач поддержки принятия решений

2.1. Интегральное аналитическое моделирование

Оперативная аналитическая обработка данных основывается в первую очередь на представлении пользователю информации в виде, который стимулирует естественные когнитивные способности человека, отображая результаты выполнения оперативных аналитических запросов к накопленным данным. Многомерное представление данных обеспечивает искомую прозрачность за счет представления интересующего фрагмента информации в виде OLAP-куба. С ростом объемов накапливаемых данных, числа доступных источников информации и разнообразия используемых форматов ранее предложенный способ организации данных перестает отвечать современным требованиям аналитики.

В ответ на все возрастающий интерес к анализу разнородных и больших данных авторами [1] предложен подход к построению интегральной OLAP-модели, позволяющей объединить данные, относящиеся к различным областям деятельности человека, независимо от исходного формата хранения данных и представлять их в виде единой (общей) многомерной модели. Анализ функциональных зависимостей в исходных схемах хранения данных позволяет ав-

томатически выявлять возможные аналитические связи между показателями и измерениями, не нарушающие непротиворечивости данных. С целью построения интегральной OLAP-модели множества показателей и измерений может быть интерпретированы как формальный контекст с учетом замены бинарного отношения «объект-признак», традиционного в рамках теории анализа формальных понятий, на отношение аналитической сопоставимости между измерениями и показателями. Переопределение классических операторов Галуа в терминах интегральной OLAP-модели позволяет построить решетку, вершины которой являются OLAP-кубами (кубами-концептами), элементы которых гарантированно могут участвовать в одном аналитическом запросе без потери достоверности информации.

По сути, полученная решетка кубов-понятий представляет интегральную OLAP-модель, объединяющую гетерогенные источники данных на логическом уровне, без физического перемещения и трансформации. Предложенный подход является теоретическим фундаментом для выявления структурных особенностей объединяемых источников, построения и объединения их многомерных аналитических моделей и формирования единой модели, обеспечивающей представление всех возможных аналитических сочетаний. Однако, несмотря на математическую строгость и элегантность решетки кубов-понятий, нерешенными остаются вопросы удобства оперирования (навигации, управления) столь обширной моделью, построения и выполнения аналитических запросов к ней. Одним из ключевых направлений развития теории интегрального аналитического моделирования на сегодняшний день является поиск решений, обеспечивающих интуитивное манипулирование большими моделями с целью оперативного анализа данных. Актуальным остается ряд задач, связанных с поддержкой исследования и анализа данных, представленных большой моделью:

- поддержка формирования целевых аналитических запросов к обширной (большой) модели – интеллектуальная поддержка формирования точных запросов в условиях большого числа возможных вариантов;
- поддержка исследования аналитической модели – проведение аналитических экспериментов, обнаружение скрытых связей и новых аналитических сочетаний в гетерогенном информационном пространстве;

– оперативное выполнение распределенных пользовательских запросов – получение и слияние данных из разрозненных и разнородных источников, учитывая их большой объем.

Дальнейшее развитие авторской технологии интегрального аналитического моделирования сопряжено с изучением примеров успешного применения метода АФП для интеллектуального анализа данных через призму поставленных задач. Эффективность анализа гетерогенных данных на основе интегральной OLAP-модели определяется возможностями решетки кубов-понятий в части интеллектуальной поддержки принятия решений.

2.2. Информационный поиск

Задача информационного поиска описывает процесс выделения из неструктурированной текстовой информации ту, которая соответствовала бы конкретному критерию (теме) или заранее утвержденному поисковому запросу. В настоящее время одним из самых популярных методов решения проблемы поиска «интересной» информации является латентно-семантический анализ (Latent semantic analysis, LSA). Также распространение получила порождающая модель латентного размещения Дирихле (Latent Dirichlet allocation, LDA), основанная на классификации множества текстов по различным тематикам.

Еще в 80-е годы были выдвинуты основные подходы, решающие проблему информационного поиска с помощью АФП. В своей обзорной статье [10] авторы излагают в хронологическом порядке развитие теории АФП в задачах информационного поиска. Помимо теоретической составляющей АФП, где авторы подчеркивают уникальность подхода, также детально описывают рекомендации для эффективного проектирования приложений: детальный анализ задачи информационного поиска, интеграция с передовыми методами информационного поиска, использование оценочных метрик, тестирование исходных данных [11].

Таким образом, задача информационного поиска в случае АФП решается следующим образом: формируется формальный контекст, в котором множество объектов – это разнообразие текстов «мешок слов» и признаков, это набор ключевых слов. Тогда решетка будет состоять из упорядоченных концептов, включающих некоторый набор текстов, которым свой-

ственен один набор ключевых слов. Соответственно, анализируя решетку сверху вниз, а именно переходя от вершины к вершине, мы можем определить источники информации, удовлетворяющие поисковому запросу. С помощью АФП решаются задачи анализа сообщений микроблогов или запросов в системах интернет-поиска. Так, в статье [12] описана задача информационного поиска с акцентом на аспекты визуализации, подходы машинного обучения, интеллектуальный анализ данных и текстов.

Преимущества метода АФП в задачах информационного поиска заключаются в следующем: преодолении ситуаций с отсутствием результатов поиска и избытком информации; возможности совместить режим выполнения запроса и навигации; возможности учесть иерархические зависимости и близость по значению между ключевыми словами (терминами); гибком поиске за счет определения приоритета признаков и ограничения пространства поиска.

2.3. Классификация

Классификация является одним из подходов в теории машинного обучения и распознавания образов. Суть задачи заключается в разделении объектов по схожим признакам на заранее известное количество классов. На данный момент предложен ряд алгоритмов классификации, основанный на теории АФП. ДСМ-метод генерации гипотез, представляющий модель автоматического обучения по положительным и отрицательным примерам, изначально был сформулирован В.К. Финном в [13], как теория правдоподобного вывода на основе бесконечнозначной логической теории первого порядка с кванторами по коротким переменным длины. Установление соответствия между ДСМ-гипотезами и формальными понятиями привело к взаимному влиянию теории ДСМ-метода и АФП [6, 7, 14]. Метод порождения ДСМ-гипотез описывается в работах [13, 14], где авторы определяют основное понятие гипотезы как некоторый набор признаков, который описывает объекты конкретного класса. При постановке задачи устанавливаются «положительные», «отрицательные» и «недоопределенные» формальные контексты, которые определяют отношения объектов к определенному классу: положительному или отрицательному, по отношению к целевому признаку. В свою очередь, нейтральный контекст характеризует недоопределенный объект,

то есть объект, который нужно определить к положительному или отрицательному классу. Далее рассчитываются формальные понятия для положительных, отрицательных и нейтральных контекстов, где для первого определяются положительные формальные объем и содержание, для второго – отрицательные объем и содержание, для третьего – соответственно, нейтральные. Таким образом, при классификации необходимо определить вхождение содержания нейтрального контекста в положительное или отрицательное содержание целиком и полностью. При условии, если в нейтральном содержании находятся как положительные, так и отрицательные признаки, устанавливаем объект как «отказ от классификации». В работах [15, 17] авторы применяют подход, основанный на реальных данных, как пример «заболевание печени», где выборка содержит объекты – пациенты и признаки – результаты анализов из 6 показателей (средний объем эритроцитов, щелочная фотосинтеза, аламиновая трансаминаза, средний объем употребляемого алкоголя в крови и т.д.). Авторы подчеркивают, что данный метод не уступает другим методам классификации, но существуют явные проблемы со скоростью работы алгоритма. В [18] представлена новая таксономия методов классификации, основанных на АФП. Для их сравнения приводятся две сводные таблицы – исчерпывающие методы классификации и комбинаторные, в которых алгоритмы классификации сравниваются по следующим параметрам: концептуальная структура, тип данных, концепция, вычислительная сложность.

В свою очередь, одним из наиболее распространенных алгоритмов классификации, использующих решетку понятий, является Ruleminer. Он предназначен для поиска логических закономерностей в данных. С этой целью используется вся решетка понятий, как пространство поиска для вывода логических правил. Таким образом, на вход алгоритма подается решетка и обучающая выборка, состоящая из множества объектов и признаков, причем каждый объект обладает определенным количеством признаков. Каждый объект в соответствии с его множеством признаков помечается меткой класса, которая указывает на принадлежность конкретного объекта к определенному классу. Тогда при определении объекта к определенному классу генерируется правило, которое проверяется на соответствие признаков.

Данный алгоритм был разработан в 1995 году. Автор [19] приводит псевдокод подхода, а также сравнивает точность с другими алгоритмами классификации на разных наборах данных.

Методы классификации на основе АФП опираются на базовые свойства решетки формальных понятий, а именно – выявление объектов с одинаковыми свойствами. Разметка объектов относительно принадлежности к определенному классу позволяет определить свойства, характерные (определяющие) для класса, и тем самым обнаружить новые знания. К недостаткам подхода можно отнести сложность интерпретации смешанных понятий, которые содержат объекты, отнесенные к разным классам в обучающей выборке.

2.4. Построение онтологических моделей

Формализация некоторой области знаний в виде концептуальных схем, состоящих из структурированных данных, осуществляется с помощью онтологий. Онтология включает в себя множество понятий – сущностей и отношений между ними. В свою очередь решетка понятий, обладающая таксономическими свойствами, пересекается с онтологической концепцией. В настоящее время популярными задачами, которые может решать анализ формальных понятий являются: извлечение онтологических понятий и выявление иерархий или сетей, построение и уточнение онтологий, исследование атрибутов в целях получения экспертных знаний [20], а также слияние онтологий. В статье [21] представлена унифицированная платформа (фреймворк) для работы с онтологиями среды описания ресурса Resource Description Framework (RDF). RDF является частью концепции семантической паутины и представляет собой модель для представления данных, в особенности – метаданных, и обладает средствами для построения информационных моделей. Данная платформа позволяет эффективно извлекать знания из семантических веб-онтологий и баз данных, контролируя их лаконичность в соответствии с правилами. В работах [22, 23] описывается процесс слияния онтологий с помощью метода снизу-вверх. Каждую из онтологий можно представить в виде формального контекста со множеством объектов и признаков. Суть подхода слияния онтологий заключается в выявлении похожих объектов со схожими признаками что позволяет объединить

онтологии одной предметной области. Далее строится решетка понятий, связывающая концепты исходных онтологий. Данный алгоритм получил название FCA-Merge, который предполагает отношения эквивалентности и подкласс-суперкласс. Полученный результат – решетка понятий анализируется экспертом и используется как руководство для создания объединенной онтологии.

В задаче построения онтологических моделей концепт решетки интерпретируется в философском смысле, объединяя объекты со схожими свойствами. Такой подход позволяет исследовать существующие и создавать новые онтологии, извлекая новые знания о моделируемой предметной области. Решение задачи слияния онтологических моделей с помощью АФП опирается на классические свойства решетки за счет выявления в разных моделях похожих объектов, обладающих схожими признаками.

2.5. Отбор значимых признаков

Классическим примером применения АФП для определения значимых признаков является Теория неточных множеств, предложенная польским ученым З. Павляком в начале 1990-х годов. Суть данной теории заключается в представлении множества объектов в виде множества подмножеств, определенных ключевыми свойствами объектов, а не конкретными объектами. Ученый дает понятие неточного множества как пары двух множеств – нижнего и верхнего приближений, построенных из элементарных множеств объектов. Эта идея является ключевой для решения многих других задач: классификации, оценки зависимостей между признаками и классификацией объектов, определения степени такой зависимости, вычисления важности признаков, сокращения количества признаков и порождения решающих правил по исходным данным.

В силу того, что реальные базы данных обычно содержат избыточную информацию, перед аналитиками стоит задача оценки важности атрибутов. Сокращение малоинформативных признаков позволяет увеличить скорость выполнения пользовательского запроса и предоставить только релевантную информацию. В работах [24–26], благодаря тесной связи между теорией неточных множеств (Rough Set, RST) и АФП, решается задача приближения неточных множеств. Операцию аппроксимации

неточных множеств можно реализовать с помощью АФП, выбирая в решетки формальные понятия, удовлетворяющие определенным условиям. Реализация операции аппроксимации неточных множеств на решетке понятий может быть разделена на 2 шага. На первом необходимо преобразовать многозначную информационную систему с несколькими ключевыми свойствами объектов в формальный контекст с одним ключевым свойством объектов. На втором шаге выполняется операция аппроксимации неточного множества с помощью решетки понятий [27] путем вычисления наименьшей верхней и наибольшей нижней граней.

Рассматривая объем формальных понятий как показатель значимости вершин диаграммы решетки, с помощью АФП можно определить ключевые признаки, характеризующие рассматриваемое множество объектов. Такая интерпретация решетки хорошо сочетается с теорией неточных множеств, позволяя повысить эффективность выполнения запросов к большим данным.

2.6. Формирование рекомендаций

Множество задач интеллектуальной поддержки принятия решений опирается на рекомендательный подход. Действительно, реализация умных ассистентов помогает лицу принимающего решение определиться с выбором фильма, товара, услуги или мероприятия в соответствии с его интересами.

В современных работах ученых наблюдается тенденция к анализу веб-контента, интересов пользователей, взаимосвязей между пользователями. В статье [28] авторы выявляют зависимости в наборе данных из социальных сетей. В предложенном авторами подходе АФП является инструментом формализации данных. В частности, ученые используют логические методы для определения деятельности пользователя в социальной сети. Кроме того, представленная в работе платформа позволяет выявить группы влияния и лидеров в этих группах. В работах [29,30] авторы описывают решение прикладной задачи выявления рекламных ключевых слов, которые могут быть интересны рекламодателю в области электронной коммерции – контекстная Интернет-реклама. В статье [31] описывается задача рекомендации, в которой анализируется суточная активность пользователей и в соответствии с выбранным про-

межутком времени предлагается определенная деятельность, исходя из его личных интересов в конкретный промежуток времени. При решении данной задачи авторы использовали сочетание АФП и теории нечетких множеств.

АФП широко применяется для решения задачи рыночной корзины. В данном подходе в формальном контексте признаками являются товары (предметы), а объектами – чеки (транзакции). По сути, импликация определяет отношения между набором предметов. Пространство поиска решетки позволяет анализировать смежные группы товаров путем просмотра объема и содержания вершин диаграммы решетки – формального понятия.

В рекламной сфере [32] одной из классических задач является подбор релевантных и «горячих» слов, которые могут поднять рейтинг компании в поисковой системе. Допустим, при поиске компанией рекламных терминов, которые описывают определенные товары и услуги, целесообразно найти термины, схожие с тематиками конкурирующих компаний. Таким образом, при анализе данных в роли объектов выступают рекламные компании, а признаками являются рекламные термины (ключевые слова), что по сути является формальным контекстом. Рассмотрение решетки понятий, а именно определенных вершин позволяет сформировать набор правил, по которым будут строиться рекламные компании. Также можно определить самые популярные слова в контексте использования их в рекламных лозунгах. В статье [33] авторы описывают концепцию айсберга, применяемую к решетке формальных понятий, которая представляет часть наборов элементов, располагающихся в верхней части решетки. Рассматриваемая концепция позволяет сформировать ассоциативные правила на основе решетки путем введения ограничения на количество концептов. Этот тип решеток может быть вычислен модифицированной версией алгоритма Next-Closure или алгоритмом Титаника [33, 34].

Задача формирования рекомендаций решается за счет использования отношения частичного порядка на множестве концептов решетки. Импликация между объемами и содержаниями концептов позволяет сформировать пространство поиска и находить объекты со смежными характеристиками. Класс решаемых задач значительно расширяется при совместном исполь-

зовании АФП с теорией нечетких множеств и ассоциативными правилами.

Заключение

Применимость АФП обусловлена его универсальностью. Метод подходит для решения любой задачи, в которой исходные данные могут быть представлены как два разнородных множества, между элементами которых задано отношение, определенное семантикой предметной области.

В данной работе описаны методы интеллектуальной поддержки принятия решений, в которых свойства АФП помогли выявить закономерности между признаками, выполнить их редуцирование, снизить вычислительную нагрузку, осуществить процесс обнаружения новых знаний и визуализировать их. Перечисленные результаты целиком и полностью актуальны и в задаче поддержки оперативного анализа данных интегральной OLAP-модели. Анализ свойств формальной решетки и в целом теории АФП, позволяет, как в рекомендательной задаче предоставить пользователю в соответствии с его запросом возможные интересные для него знания, в задаче неточных множеств облегчить поиск интересной информации с помощью редуцирования незначительных признаков, в задаче классификации выявить группы схожих признаков. Таким образом, изучение теории алгебраических решеток, в частности АФП, позволяет агрегировать основные теоремы и правила, впоследствии применить при решении задачи поддержки оперативного анализа больших данных совокупность знаний, которая используется в задачах интеллектуальной поддержки принятия решений.

В работе был рассмотрен математический метод АФП, основанный на теории алгебраических решеток. На основе приведенных примеров показывается стройность и универсальность подхода при решении задач интеллектуальной поддержки принятия решений, в частности, в каждом конкретном случае авторами подчеркиваются основные свойства решетки понятий. Таким образом, совокупность свойств, примененных при решении описанных примеров, позволяет обнаружить новые знания, снизить вычислительную нагрузку и визуализировать информацию при решении задач интеллектуального анализа больших данных.

Литература

1. Коробко А.В., Пенькова Т.Г., Поддержка оперативной аналитической обработки данных на основе базы знаний о кубах-концептах. 2011. С. 136–143.
2. Belohlavek R., Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic*. 2004. Vol. 128, pp. 277–298.
3. Ganter B., Wille R. *Formal Concept Analysis*. Springer-Verlag Berlin Heidelberg. 1999.
4. Davey B.A., Priestley H.A., *Introduction to Lattices and Order*. Cambridge: Cambridge University Press. 2002.
5. Poelmans, J., Kuznetsov, S.O., Ignatov, D.I., & Dedene, G., Formal Concept Analysis in knowledge processing: A survey on models and techniques. *Expert Systems with Applications*. 2013. Vol. 40, pp. 6601–6623.
6. Kuznetsov, S.O., & Poelmans, J., Knowledge representation and processing with formal concept analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2013. Vol. 3, pp. 200–215.
7. Kuznetsov S.O., Galois Connections in Data Analysis: Contributions from the Soviet Era and Modern Russian Research. In: B. Ganter, G. Stumme, R. Wille, Eds., *Formal Concept Analysis: Foundations and Applications*, Lecture Notes in Artificial Intelligence (Springer), State-of-the-Art Ser. 2005. Vol. 3626, pp. 196–225.
8. Obiedkov S., Modeling Ceteris Paribus Preferences in Formal Concept Analysis. In: Cellier P., Distel F., Ganter B. (eds) *Formal Concept Analysis*. ICFCA 2013. Lecture Notes in Computer Science. Springer. Berlin. Heidelberg. 2013. Vol. 7880.
9. Богатырев М. Ю., Нуриахметов В. Р., Вакурин В. С. Методы анализа формальных понятий в информационных системах технической поддержки // *Известия ТулГУ. Технические науки*. 2013.
10. Poelmans, J. et al. Text Mining Scientific Papers: A Survey on FCA-Based Information Retrieval Research. In: Petra Perner, Ed., *Proc. 11th Industrial Conference on Data Mining (ICDM 2012)*, Lecture Notes in Computer Science (Springer). 2012. Vol. 7377, pp. 273–287.
11. Carpineto C., Romano G., Using Concept Lattices for Text Retrieval and Mining. In: Ganter B., Stumme G., Wille R. (eds) *Formal Concept Analysis*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg. 2005. Vol. 3626.
12. Ignatov D.I., Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields. In: Braslavski P., Karpov N., Worrning M., Volkovich Y., Ignatov D. (eds) *Information Retrieval. RuSSIR 2014*. Communications in Computer and Information Science. Springer. Cham. 2015. Vol. 505.
13. Финн В. К. О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф. Бэкона – Д. С. Милля // *Семиотика и информатика*. 1983. №20. С. 35–101.
14. Кузнецов С.О. Автоматическое обучение на основе анализа формальных понятий // *Автоматика и телемеханика*. 2001. №10. С. 3–27.
15. Prokasheva O., Onishchenko A., Gurov S. Classification Methods Based on Formal Concept Analysis. *Formal Concept Analysis Meets Informational Retrieval*. Workshop collocated with 35th European Conference on Informational Retrieval (ECIR 2013). 2013. pp. 1–10.
16. Онищенко А. А., Гуров С., И. Классификация на основе АФП и бикластеризации: возможности подхода // *Прикладная математика и информатика: Труды факультета Вычислительной математики и кибернетики*. 2011. Т. 38. С. 77–87.
17. Prokasheva O., Efficiency improvement of the FCA-based classification algorithm. 2013. Vol. 1, pp. 550–556.
18. Trabelsi, M., Meddouri, N., & Maddouri, M., New Taxonomy of Classification Methods Based on Formal Concepts Analysis. *FCA4AI@ECAI*. 2016. Vol. 1703, pp. 113–120.
19. Sahami M., Learning classification rules using lattices (Extended abstract). In: Lavrac N., Wrobel S. (eds) *Machine Learning: ECML-95*. ECML 1995. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), Springer, Berlin, Heidelberg. 1995. Vol. 912, pp. 343–346.
20. Coulet A., Domenach F., Kaytoute M., Napoli A., Using Pattern Structures for Analyzing Ontology-Based Annotations of Biomedical Data. In: Cellier P., Distel F., Ganter B. (eds) *Formal Concept Analysis*. ICFCA 2013. Lecture Notes in Computer Science. Springer. Berlin. Heidelberg. 2013. Vol. 7880.
21. Bertet K. et al. *Semantic Web: Big Data, Some Knowledge and a Bit of Reasoning*. 2017.
22. Stumme, G., Maedche, A., FCA-MERGE: Bottom-Up Merging of Ontologies. *IJCAI*. 2001, pp. 225–234.
23. Cimiano P., Hotho A., Stumme G., Tane J., Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies. In: Eklund P. (eds) *Concept Lattices*. ICFCA 2004. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 2004. Vol. 2961.
24. Benítez-Caballero M.J., Medina J., Ramírez-Poussa E., Attribute Reduction in Rough Set Theory and Formal Concept Analysis. In: Polkowski L. et al. (eds) *Rough Sets*. IJCRS 2017. Lecture Notes in Computer Science. Springer. Cham. 2017. Vol. 10313, pp. 513–525.
25. Liu, M., Shao, M., Zhang, W., & Wu, C., Reduction method for concept lattices based on rough set theory and its application. *Computers & Mathematics with Applications*. 2007. Vol. 53, pp. 1390–1410.
26. Kent R.E., Rough Concept Analysis. In: Ziarko W.P. (eds) *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Workshops in Computing. Springer. London. 1994.
27. Zhi H., Realization of Rough Set Approximation Topological Operations Based on Formal Concept Analysis. *International Journal of Intelligence Science*. 2014. Vol. 4, pp. 65–69.
28. Cordero P. et al. Knowledge discovery in social networks by using a logic-based treatment of implications. *Knowledge-Based Syst*. Elsevier B.V. 2015. Vol. 87, pp. 16–25.
29. Ignatov, D.I., Kuznetsov, S.O., Concept-based Recommendations for Internet Advertisement. 2008. Vol. 433, pp. 157–166.
30. Игнатов Д.И., Кузнецов С.О. Методы разработки данных (Data Mining) для рекомендательной системы Интернет-рекламы // *Одиннадцатая национальная конференция по искусственному интеллекту с международным участием*. М.: Ленанд. 2008.
31. Medina J., Pakhomova K., Ramirez-Poussa E., Recommendation Solution for a Locate-Based Social Network via Formal Concept Analysis. In: Cornejo M., Kóczy L., Medina J., De Barros Ruano A. (eds) *Trends*

- in Mathematics and Computational Intelligence. Studies in Computational Intelligence. Springer. Cham. 2019. Vol. 796, pp. 131–138.
32. Lakhall L., Stumme G., Efficient Mining of Association Rules Based on Formal Concept Analysis. In: Ganter B., Stumme G., Wille R. (eds) Formal Concept Analysis. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, 2005. Vol. 3626, pp. 180–195.
 33. Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., & Lakhall, L., Computing iceberg concept lattices with Titan-ic. Data Knowl. Eng. 2002. Vol. 42, pp. 189–222.
 34. Abdullah Z., Saman M.Y.M., Karim B., Herawan T., Deris M.M., Hamdan A.R., FCA-ARMM: A Model for Mining Association Rules from Formal Concept Analysis. In: Herawan T., Ghazali R., Nawi N., Deris M. (eds) Recent Advances on Soft Computing and Data Mining. SCDM 2016. Advances in Intelligent Systems and Computing. Springer. Cham. 2017. Vol. 549, pp. 213–223.

Application of Formal Concept Analysis for Intellectual Decision Support

K. I. Pakhomova¹, A. V. Korobko¹¹

¹ Siberian Federal University, Krasnoyarsk, Russia

¹¹ Institute of Computational Modelling of SB RAS, Krasnoyarsk, Russia

Abstract. The development of smart assistants in the information systems area has led to the increased interest of the scientific to artificial intelligence methods. This article focused on the study of modern directions of intellectual support of decision-making by using the method of algebraic lattice theory that is formal concept analysis (FCA). Most scientists apply this approach for a solution to a wide range of tasks in an artificial intelligence area. In this article, we have learned modern researchers used FCA for classification, recommendations, ontology, information retrieval, and feature selection. For each category of problems were analyzed the advantages of lattice theory in FCA case, also was described the solution of problems the possibilities of the FCA method. On the result, was defined the main advantages of the FCA method which were applied for the development of the author's theory of intellectual support of operational analytical processing of heterogeneous data on the basis of integral OLAP-model.

Keywords: artificial intelligence, formal concept analysis, OLAP, intelligent decision support, analytical model.

DOI 10.14357/20718594190405

References

1. Korobko A.V., Pen'kova T.G. Podderzhka operativnoj analiticheskoy obrabotki dannyh na osnove bazy znaniy o kubah-konceptah, pp. 136–143., 2011.
2. Belohlavek R., Concept lattices and order in fuzzy logic. Annals of Pure and Applied Logic, Vol. 128, pp. 277–298., 2004.
3. Ganter B., Wille R. Formal Concept Analysis. Springer-Verlag Berlin Heidelberg, 1999.
4. Davey B.A., Priestley H.A., Introduction to Lattices and Order. Cambridge: Cambridge University Press, 2002.
5. Poelmans, J., Kuznetsov, S.O., Ignatov, D.I., & Dedene, G., Formal Concept Analysis in knowledge processing: A survey on models and techniques. Expert Systems with Applications, Vol. 40, pp. 6601–6623., 2013.
6. Kuznetsov, S.O., & Poelmans, J., Knowledge representation and processing with formal concept analysis. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 3, pp. 200–215., 2013.
7. Kuznetsov S.O., Galois Connections in Data Analysis: Contributions from the Soviet Era and Modern Russian Research. In: B. Ganter, G. Stumme, R. Wille, Eds., Formal Concept Analysis: Foundations and Applications, Lecture Notes in Artificial Intelligence (Springer), State-of-the Art Ser., Vol. 3626, pp. 196–225., 2005.
8. Obiedkov S., Modeling Ceteris Paribus Preferences in Formal Concept Analysis. In: Cellier P., Distel F., Ganter B. (eds) Formal Concept Analysis. ICFCFA 2013. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Vol 7880, 2013.
9. Bogatyrev M. YU., Nuri Ahmetov V. R., Vakurin V. S. Metody analiza formal'nyh ponyatij v informacionnyh sistemah tekhnicheskoy podderzhki, Izvestiya TulGU, Tekhnicheskie nauki, 2013.
10. Poelmans, J. et al., Text Mining Scientific Papers: A Survey on FCA-Based Information Retrieval Research. In: Petra Perner, Ed., Proc. 11th Industrial Conference on Data Mining (ICDM 2012), Lecture Notes in Computer Science (Springer), Vol. 7377, pp. 273–287., 2012.
11. Carpineto C., Romano G., Using Concept Lattices for Text Retrieval and Mining. In: Ganter B., Stumme G., Wille R. (eds) Formal Concept Analysis. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Vol. 3626, 2005.
12. Ignatov D.I., Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields. In: Braslavski P., Karpov N., Worring M., Volkovich Y., Ignatov D. (eds) Information Retrieval.

- RuSSIR 2014. Communications in Computer and Information Science, Springer, Cham, Vol. 505, 2015.
13. Finn V. K. O mashinno-orientirovannoj formalizacii pravdopodobnyh rassuzhdenij v stile F. Bekona – D. S. Millya. Semiotika i informatika, № 20, pp. 35–101., 1983.
 14. Kuznecov S.O. Avtomaticheskoe obuchenie na osnove analiza formal'nyh ponyatij. Avtomat. i telemekh, № 10, pp. 3–27., 2001.
 15. Prokasheva O., Onishchenko A., Gurov S. Classification Methods Based on Formal Concept Analysis. Formal Concept Analysis Meets Informational Retrieval. Workshop collocated with 35th European Conference on Informational Retrieval (ECIR 2013), pp. 1–10., 2013.
 16. Onishchenko A.A., Gurov S.I., Klassifikaciya na osnove AFP i biklasterizacii: vozmozhnosti podhoda. Prikladnaya matematika i informatika Trudy fakul'teta Vychislitel'noj matematiki i kibernetiki, Vol. 38, pp. 77–87, 2011.
 17. Prokasheva O., Efficiency improvement of the FCA-based classification algorithm. Vol. 1, pp. 550–556, 2013.
 18. Trabelsi, M., Meddouri, N., & Maddouri, M., New Taxonomy of Classification Methods Based on Formal Concepts Analysis. *FCA4AI@ECAI*, Vol. 1703, pp. 113–120., 2016.
 19. Sahami M., Learning classification rules using lattices (Extended abstract). In: Lavrac N., Wrobel S. (eds) Machine Learning: ECML-95. ECML 1995. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), Springer, Berlin, Heidelberg, Vol. 912, pp. 343–346., 1995.
 20. Coulet A., Domenach F., Kaytoute M., Napoli A., Using Pattern Structures for Analyzing Ontology-Based Annotations of Biomedical Data. In: Cellier P., Distel F., Ganter B. (eds) Formal Concept Analysis. ICFA 2013. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Vol. 7880, 2013.
 21. Bertet K. et al., Semantic Web: Big Data, Some Knowledge and a Bit of Reasoning. 2017.
 22. Stumme, G., Maedche, A., FCA-MERGE: Bottom-Up Merging of Ontologies. *IJCAI*, pp. 225–234., 2001.
 23. Cimiano P., Hotho A., Stumme G., Tane J., Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies. In: Eklund P. (eds) Concept Lattices. ICFA 2004. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Vol. 2961, 2004.
 24. Benítez-Caballero M.J., Medina J., Ramírez-Poussa E., Attribute Reduction in Rough Set Theory and Formal Concept Analysis. In: Polkowski L. et al. (eds) Rough Sets. IJCRS 2017. Lecture Notes in Computer Science. Springer, Cham, Vol. 10313, pp. 513–525., 2017.
 25. Liu, M., Shao, M., Zhang, W., & Wu, C., Reduction method for concept lattices based on rough set theory and its application. *Computers & Mathematics with Applications*, Vol. 53, pp. 1390–1410., 2007.
 26. Kent R.E., Rough Concept Analysis. In: Ziarko W.P. (eds) Rough Sets, Fuzzy Sets and Knowledge Discovery. Workshops in Computing. Springer, London, 1994.
 27. Zhi H., Realization of Rough Set Approximation Topological Operations Based on Formal Concept Analysis. *International Journal of Intelligence Science*, Vol.4, pp. 65–69., 2014.
 28. Cordero P. et al. Knowledge discovery in social networks by using a logic-based treatment of implications, *Knowledge-Based Syst. Elsevier B.V.*, Vol. 87, pp. 16–25., 2015.
 29. Ignatov, D.I., Kuznetsov, S.O., Concept-based Recommendations for Internet Advertisement, Vol. 433, pp. 157–166., 2008.
 30. Ignatov D.I., Kuznecov S.O., Metody razrabotki dannyh (Data Mining) dlya rekomendatel'noj sistemy Internet-reklamy. Odinnadcataya nacional'naya konferenciya po iskusstvennomu intellektu s mezhdunarodnym uchastiem. Moskva, Lenand, 2008.
 31. Medina J., Pakhomova K., Ramírez-Poussa E., Recommendation Solution for a Locate-Based Social Network via Formal Concept Analysis. In: Cornejo M., Kóczy L., Medina J., De Barros Ruano A. (eds) Trends in Mathematics and Computational Intelligence. Studies in Computational Intelligence. Springer, Cham, Vol. 796, pp. 131–138., 2019.
 32. Lakhal L., Stumme G., Efficient Mining of Association Rules Based on Formal Concept Analysis. In: Ganter B., Stumme G., Wille R. (eds) Formal Concept Analysis. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, Vol. 3626, pp. 180–195., 2005.
 33. Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., & Lakhal, L., Computing iceberg concept lattices with Titanic. *Data Knowl. Eng.*, Vol. 42, pp. 189–222., 2002.
 34. Abdullah Z., Saman M.Y.M., Karim B., Herawan T., Deris M.M., Hamdan A.R., FCA-ARMM: A Model for Mining Association Rules from Formal Concept Analysis. In: Herawan T., Ghazali R., Nawi N., Deris M. (eds) Recent Advances on Soft Computing and Data Mining. SCDM 2016. Advances in Intelligent Systems and Computing. Springer, Cham, Vol. 549, pp. 213–223., 2017.