



OSTIS-2011

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822

О ПРЕДСТАВЛЕНИИ ЗНАНИЙ С ПОМОЩЬЮ СЕМАНТИЧЕСКИХ СЕТЕЙ В ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ ПО ИССЛЕДОВАНИЮ ФОЛЬКЛОРНЫХ ТЕКСТОВ

Н.Д. Москин (*moskin@karelia.ru*)

Петрозаводский государственный университет, г.Петрозаводск, Россия

В работе рассматриваются методы извлечения, хранения и анализа знаний о языковой структуре фольклорных текстов на основе семантических сетей, а также их реализация в интеллектуальной системе «Фольклор».

Ключевые слова: интеллектуальная система, семантические сети, фольклорный текст, XML.

Введение

К настоящему времени специалистами, изучающими народную культуру России и ближнего зарубежья, накоплены большие по объему коллекции фольклорных текстов, которые хранятся как в печатной, так и в электронной форме в виде реляционных баз данных. При анализе подобных коллекций часто возникают достаточно сложные задачи, к которым можно отнести жанровую дифференциацию и атрибуцию текстов, поиск в текстах устойчивых языковых конструкций (мотивов), реконструкцию текстов и т.д. Однако традиционные методики (например, мотивный анализ, сравнительно-типологический метод, контент-анализ) оказываются не всегда достаточными для проведения подобных исследований.

На наш взгляд, адекватной моделью для представления знаний о фольклорном тексте является семантическая сеть. Не случайно в своей работе известный фольклорист А. Т. Хроленко пишет: «Все элементы текста, а именно, основные персонажи, предметы и явления окружающего мира образуют некую систему, представляющую собой фольклорную модель мира, которая стремится к замкнутости, полноте и устойчивости. Можно говорить о том, что она представляет собой своеобразную семантическую «сеть», узлами которой являются опорные однозначные слова со всей их семантической информацией» [Хроленко, 1992].

С 2001 года на кафедре информатики и математического обеспечения Петрозаводского государственного университета ведется работа над исследованием фольклорных коллекций с теоретико-графовой формализацией текстов [Москин, 2006а]. Одним из результатов работы является разработка интеллектуальной системы «Фольклор» в среде визуального программирования Delphi 7.0. Изначально она проектировалась как проблемно-ориентированная система, предназначенная для сравнительного анализа коллекции беседных песен Заонежья конца XIX – начала XX века. Однако впоследствии программа была модифицирована таким образом, что позволила проводить исследование других коллекций, в том числе коллекций исторических текстов, на основе различных теоретико-графовых моделей.

В данной работе описывается, какие виды семантических сетей используются в системе для представления знаний о языковой структуре фольклорных текстов, рассматриваются методы их анализа, вопросы хранения и публикации знаний с помощью технологии XML, структура и основные модули программы «Фольклор».

1. Представление знаний об языковой структуре фольклорных текстов с помощью семантических сетей

Одной из особенностей программы является то, что пользователь может самостоятельно определить вид теоретико-графовой модели, т. е. выбрать систему объектов и отношений между ними в соответствии с целями и задачами исследования. Рассмотрим основные типы моделей, на основе которых проводился анализ коллекций.

В современной теории синтаксиса хорошо известны так называемые деревья зависимостей. Здесь считается заданным анализ предложения, устанавливающий подчинительные связи между словами согласно постулатам традиционной грамматики и грамматики зависимостей. Результаты анализа представляются в виде графов, вершины которых соответствуют словам, а дуги соединяют их в соответствии с синтаксическими связями. При графическом изображении деревьев зависимостей обычно учитывают порядок слов самого предложения. В качестве примера рассмотрим первое предложение былины «Илья Муромец» в записи П. Н. Рыбникова:

Старый казак Илья Муромец
Поехал на добром коне
Мимо Чернигов-град:
Под Черниговым силушки черным-черно,
Черным-черно, как черна ворона.

Изображая графы, будем обозначать: ● – полнозначные слова, ○ – предлоги, ⊕ – сочинительные союзы. Сплошной линией обозначим подчинение слов, пунктиром – однородность слов, сплошной и пунктиром – сочинительную связь (рис. 1). Подобные графы обычно применяются в описаниях языков со свободным порядком слов (например, русского).

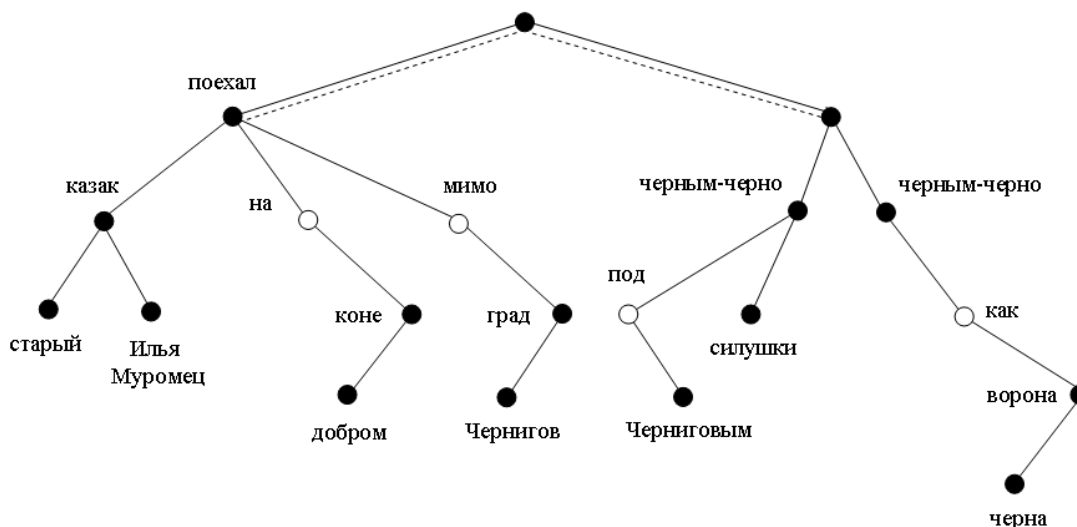


Рисунок 1 - Дерево зависимостей первого предложения былины «Илья Муромец».

Деревья зависимостей отражают стилистические характеристики текстов, что в дальнейшем может использоваться при решении проблем жанровой дифференциации и атрибуции фольклорных произведений.

В основе другой модели лежит понятие мотива, который, по выражению Б. Н. Путилова, является «узловой категорией художественной организации произведения фольклора». Содержательную основу мотива можно представить в виде помеченного мультиграфа, в узлах которого находятся основные персонажи песни, животные, явления природы, предметы обихода и т. д. [Москин, 2006b] Между объектами устанавливаются связи двух видов: локальные и глобальные, соответствующие синтагматическим и парадигматическим отношениям в тексте (пример обобщенного графа мотива представлен на рис. 5). Задача построения графа мотива является достаточно сложной, требующей привлечения знаний эксперта-фольклориста. В

некоторых случаях у разных экспертов могут возникнуть противоположные мнения относительно того, существует ли определенная связь в графе, к какой группе принадлежит объект, каким образом определить границы фольклорного мотива. Эту важную информацию можно отразить в графе, используя нечеткие объекты и отношения, определив значения функции принадлежности в зависимости от квалификации экспертов. Если связать графы мотивов, объединив одинаковые персонажи в одну вершину, то подобную структуру можно изобразить в виде единого графа сюжета песни.

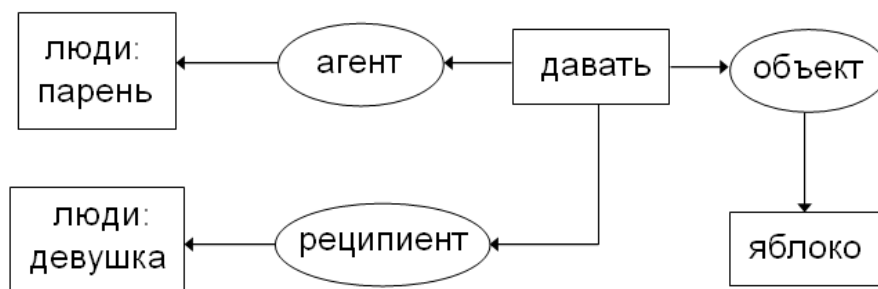


Рисунок 2 – Концептуальный граф предложения

Также для моделирования семантики фольклорного произведения можно использовать концептуальные графы [Люгер, 2005]. Это конечные, связанные, двудольные графы, узлы которых представляют понятия (концепты) или концептуальные отношения (на рис. 2 они изображены прямоугольниками и эллипсами соответственно). Каждый концептуальный граф определяет одно высказывание. Существенным преимуществом подобного представления является простота определения отношений любой арности.

Еще один вид графов – текстовые сети – описывается в работе [Скороходько, 1983] (пример подобной модели приводится в четвертом параграфе статьи). Для построения этой и других теоретико-графовых моделей текстов в системе «Фольклор» используется следующая пошаговая автоматизированная процедура:

- Шаг 1: выбор параметров построения графа. Пользователю предлагается выбрать текст из коллекции и установить группы объектов и связей.
- Шаг 2: определение объектов в тексте. На этом этапе в тексте последовательно выделяются слова, которые пользователь может добавить в список объектов (рис. 3). Слова отбираются по заранее заданному критерию: например, по принадлежности к некоторой части речи. Если объект состоит из нескольких слов (или словосочетания), то их можно объединить в один объект.
- Шаг 3: разбиение объектов на группы. На этом этапе определяются названия объектов и их принадлежность к той или иной группе, а также задаются «фиктивные» объекты (т. е. объекты, не связанные со словами текста).
- Шаг 4: определение связей в тексте. На этом шаге устанавливаются связи между объектами и названия связей (рис. 4). При этом объекты, между которыми устанавливаются отношения, можно выбирать либо во всем тексте, либо в некотором фрагменте. При определении связи пользователь может обратиться за помощью к экспертной системе, которая с помощью уточняющих вопросов оценивает вероятность существования связи между объектами.
- Шаг 5: разбиение связей на группы. На этом шаге устанавливается принадлежность связей к той или иной группе. При этом структура связей в тексте может задаваться в виде матрицы смежности или матрицы инцидентности.

В процессе работы пользователь может скорректировать полученный граф (удалить или добавить связи и объекты, изменить их тип, порядок, внутреннюю иерархию и т. д.).

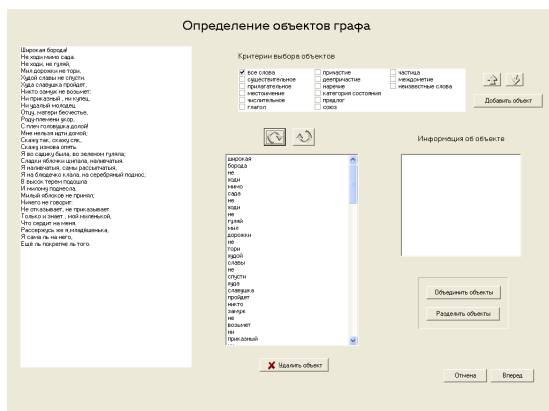


Рисунок 3 – Определение объектов графа

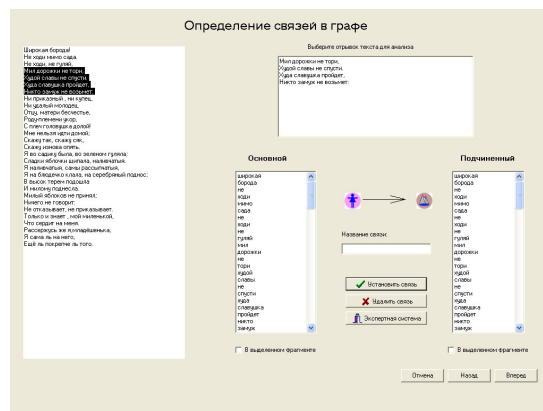


Рисунок 4 – Определение связей графа

2. Методы структурирования и анализа полученных знаний

2.1. Обобщенные нечеткие графы мотивов и задача поиска

Одной из важных задач, возникающих при работе с фольклорной коллекцией, является проблема обнаружения в текстах схожих мотивов и их классификация. Как показано в [Москин, 2010b], задачу обнаружения мотивов можно свести к задаче поиска схожего по структуре подграфа, предварительно определив обобщенный граф мотива с нечеткими объектами и связями на основе нескольких подобных мотивов, записанных разными собирателями или в разных местах (рис. 5).

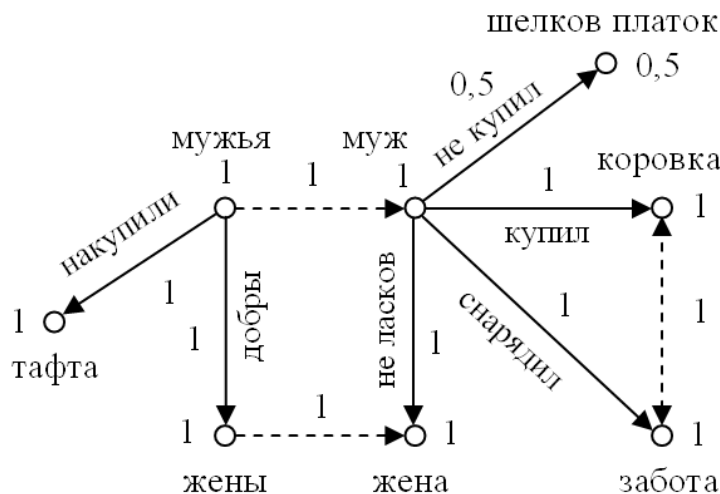


Рисунок 5 – Обобщенный граф мотива с нечеткими объектами и связями

В программе искомый мотив можно задать двумя способами: либо пользователь самостоятельно определяет объекты и связи, либо выделяет в фольклорном тексте границы мотива и программа автоматически строит граф. Алгоритм поиска описывается в работе [Москин, 2010b]. Данный метод можно усовершенствовать, дополнив его поиском по ключевым словам. Также можно ввести ограничения на принадлежность объектов к определенной группе, на тип и порядок появления связей в тексте.

2.2. Классификация и кластеризация семантических сетей

Другой интересной задачей является сравнительный анализ семантических сетей, построенных на основе текстов. В теории графов существует много определений типов графов, которые отражают их особенности: планарность, связность, ацикличность и т. д. Подобная классификация удобна тогда, когда графы имеют случайную структуру и не «загружены» дополнительной семантикой. Поэтому часто этих понятий оказывается недостаточно для проведения глубокого исследования и получения в результате новых знаний об объектах-оригиналах.

Один из способов определения сходства графов связан с использованием различных числовых инвариантов (например, параметра семантической связности текста, цепочечного коэффициента [Скороходько, 1983] и т. д.). Другой подход основан на использовании подграфовой метрики. За рубежом данное направление получило название «graph matching». Эти методы нашли свое применение в обработке изображений, химии, молекулярной биологии, дактилоскопии и т. д. На множестве графов задается расстояние, которое позволяет оценить насколько те или иные структуры «похожи» друг на друга. Среди них можно выделить следующие метрики [Москин, 2009a]:

- Мера на основе максимального общего подграфа;
- Мера на основе минимального общего надграфа;
- Мера на основе операций редактирования.

На основе этих методов в системе «Фольклор» была реализована процедура автоматического разбиения графов на группы (кластеры). При этом пользователь может выбрать тип расстояния между объектами и между кластерами. Результаты анализа выводятся в левом окне в виде списка групп текстов (рис. 6). Также предусмотрена усложненная процедура, включающая детальное отображение всех промежуточных числовых данных и графическое изображение кластеров.

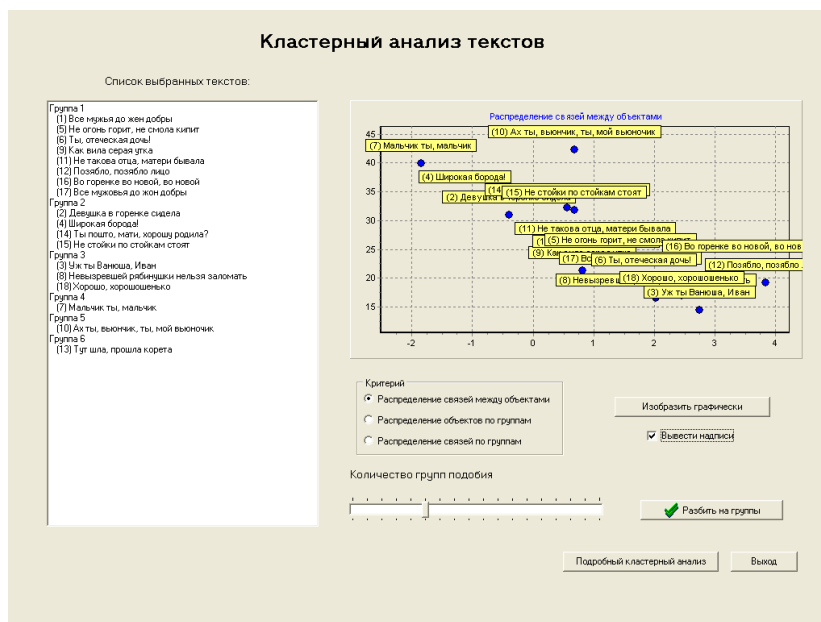


Рисунок 6 – Процедура кластерного анализа

Для представления совокупности близких графов одним, который содержал бы основную информацию обо всех структурах, рассмотрим понятие среднего графа («median graph») [Москин, 2009a]. Пусть на множестве $S = \{G_1, G_2, \dots, G_m\}$ задано расстояние $d(G_i, G_j)$. Тогда назовем средним графом на множестве S :

$$\hat{G} = \arg \min_{G \in S} \sum_{i=1}^m d(G, G_i). \quad (1)$$

Если граф рассматривается на всем множестве допустимых графов $S \subset U$, то назовем такой граф обобщенным средним графом:

$$\bar{G} = \arg \min_{G \in U} \sum_{i=1}^m d(G, G_i). \quad (2)$$

Средние графы можно использовать для представления текстов определенного вида или жанра, а также при определении инварианта для нескольких фольклорных текстов и анализа отклонений.

2.3. Применение агрегированных графов для обобщения информации

Часто экспериментальные графы имеют достаточно сложную структуру связей, поэтому возникает задача, насколько важны те или иные элементы и можно ли их отбросить при дальнейшем анализе. Для решения этой задачи можно использовать алгоритмы агрегации графов. Данным направлением математики занимались в 80-е годы такие ученые, как Л. И. Бородин, Э. М. Браверман, В. Л. Куперштох, Б. Г. Миркин, И. Б. Мучник, Г. Ш. Фридман и др. [Москин, 2006b]. Основу методов составляет задача нахождения такого «простого» графа, вершины которого соответствуют подмножествам вершин исходного графа, а ребра соответствуют «основным потокам связей» на исходном графе. Эти методы применяются в социологии (при изучении неформальных структур в человеческих коллективах), в экономике, биологии, в изучении транспортных и коммуникационных систем и т. д.

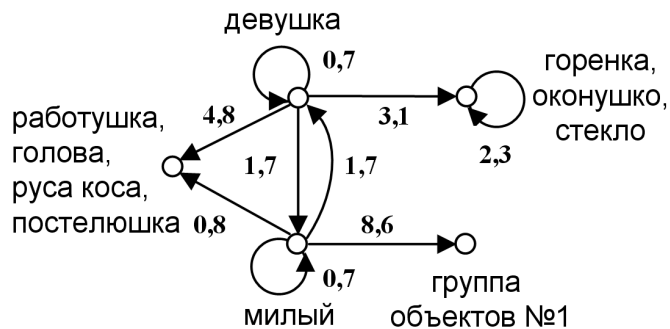


Рисунок 7 - Пример агрегированного графа для песни «Девушка в горенке сидела»

В системе «Фольклор» реализованы два подобных метода [Москин, 2010а]. В первом методе (авторы И. Б. Мучник и Э. М. Браверманн) структура основных связей задается заранее в виде особого графа. В другом методе она ищется в процессе построения агрегированной модели. На рис. 7 представлен пример агрегированного графа для песни «Девушка в горенке сидела» (здесь группа объектов №1 – это «поле», «ясен сокол», «белые руки», «желтые кудри», «резвые ноги» и «высокий терем», т. е. все объекты, связанные с «милым дружкой»).

3. Основные компоненты интеллектуальной системы «Фольклор»

В настоящее время система содержит 562 текста из четырех фольклорных коллекций, знания о которых представлены в виде различных семантических сетей:

- бесёдные песни Заонежья XIX – начала XX века (118 текстов);
- Лужские песни из сборника «Песни городского хора» (32 текста);
- духовные стихи о Голубиной книге из сборника Кириши Данилова и «Собрания народных песен П. В. Киреевского» (88 текстов);
- былины из сборника П. Н. Рыбникова (44 текста);
- описания «народных святых» Нижегородского края (280 текстов).

Общая структура системы изображена на рис. 8. Программа состоит из семи функциональных модулей:

- Модуль ввода и редактирования фольклорных текстов;
- Модуль анализа текстов. Включает процедуры графематического, морфологического и контент-анализа текстов, встроенный морфологический словарь;
- Модуль автоматизированного построения теоретико-графовых моделей;
- Модуль визуализации графов. Включает методы двухмерной и трехмерной визуализации теоретико-графовых моделей, основанные на физических аналогиях, а также процедуру поуровневого изображения деревьев;
- Модуль агрегации графов;
- Модуль поиска графов;
- Модуль классификации и кластеризации теоретико-графовых моделей. Содержит модуль определения параметров графов (коэффициент связности, распределение объектов и связей на группы, коэффициенты регрессии, полученные при аппроксимации рангового распределения объектов по числу связей гиперболической кривой, функциональные веса вершин и др.), модуль вычисления метрик на множестве деревьев, модуль вычисления метрик на основе общих подграфов и операций редактирования для произвольных графов, модуль кластерного анализа.

Более подробное описание инструментов анализа графов можно найти в [Москин, 2010а].

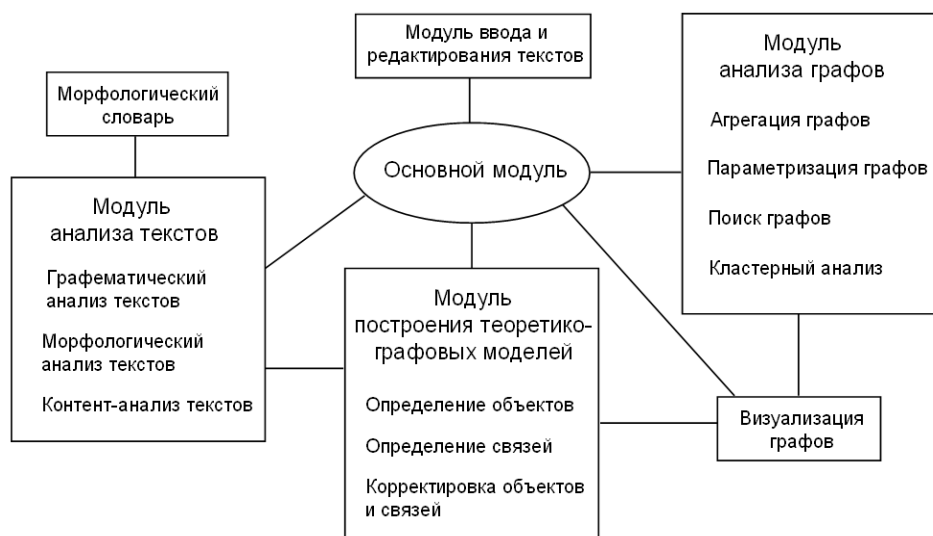


Рисунок 8 – Общая структура системы «Фольклор»

4. О хранении и публикации знаний с помощью технологии XML

Для хранения фольклорных текстов и их теоретико-графовых моделей, а также для дальнейшей публикации методики и результатов исследования в сети Интернет, необходимо разработать специальный формат данных. На наш взгляд, наиболее удобным решением является использование расширяемого языка разметки XML. Данные в XML хранятся в упорядоченном виде, тексты отличаются простым синтаксисом, обеспечены бесплатными и общедоступными сопутствующими технологиями. Кроме того, именно XML является универсальным форматом для хранения и передачи информации в сетевых приложениях. В настоящее время эта технология все чаще находит свое применение в гуманитарных исследованиях: в качестве примеров можно привести стандарт для описания исторических событий HEML (<http://www.heml.org/>), проект МЕР для исследования документов по истории США (<http://mep.cla.sc.edu/>), описание сказочных сюжетов с помощью языка FolkML, методику анализа средневековых документов комплекса «Moscovitica-Ruthenica» и др. Недавно в Петрозаводском государственном университете начал осуществляться проект специализированной виртуальной среды для научных исследований,

ориентированной на текстологический анализ списков литературных памятников вопросно-ответной формы (апокрифа «Беседа трёх святителей»), где разметка текстов осуществляется с помощью формата TEI, также основанного на технологии XML [Варфоломеев и др., 2010].

Для описания графов и теоретико-графовых моделей в настоящее время существует несколько общепризнанных стандартов, основанных на технологии XML. Одним из предшественников таких форматов является язык GML (Graph Modelling Language), который появился в результате работы, начатой на конференции «Graph Drawing-1995» в Пассау и завершённой на «Graph Drawing-1996» в Беркли. GML до сих пор поддерживается многими прикладными программами и библиотеками для работы с графами. Позднее появились стандарты GraphXML, GraphML, XGMML (eXtensible Graph Markup and Modeling Language), GXL (Graph eXchange Language) и др. Однако все перечисленные форматы предназначены для описания произвольных графов и графовых моделей, не привязанных к тексту. Поэтому для формального описания, хранения и изучения теоретико-графовых моделей текстов нами был разработан язык разметки TextGML (Textual Graph Modelling Language) на основе XML. Формат TextGML позволяет хранить сам текст, его характеристики, объекты и связи теоретико-графовой модели, их свойства: тип, упорядоченность, внутреннюю иерархию и др. [Москин, 2009b].

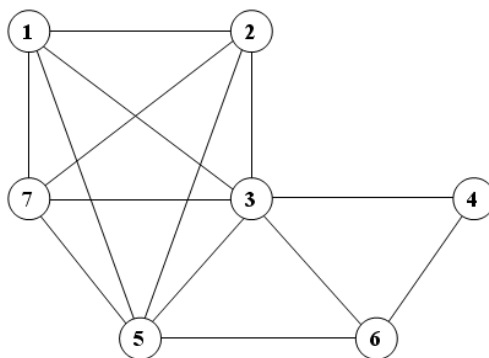


Рисунок 9 – Текстовая сеть рассказа «Уличная торговля»

Заметим, что в качестве объектов графа могут выступать как слова, так и словосочетания, предложения, абзацы и более крупные элементы текста. Применим данную технологию для описания текстовой сети на примере рассказа «Уличная торговля» [Скороходько, 1983]. Вершинами данного графа (рис. 9) являются предложения текста, а в качестве критерия семантической связи между предложениями выступает повторение одних и тех же существительных (таких как «Мугдхабуддхи», «сандал», «уголь») при условии замены личных местоимений их антецедентами. На языке TextGML описание графа выглядит следующим образом:

```

<text name="Уличная торговля" type="рассказ">
  <text_parameter>
    <parameter id="p1" name="source">Повести, сказки, притчи древней Индии</parameter>
    <parameter id="p2" name="publish_year">1964 год</parameter>
  </text_parameter>
  <graph id="g1" name="текст" directed="false">
    <node id="n1" name="sentence" fuzzy="1">Был когда-то у одного богатого купца сын и звали его Мугдхабуддхи.</node>
    <node id="n2" name="sentence" fuzzy="1">Поехал однажды Мугдхабуддхи торговать на остров Катаха и нагрузил корабль, помимо всего прочего, изрядным грузом сандалового дерева.</node>
    <node id="n3" name="sentence" fuzzy="1">Вот все товары он уже продал, а на сандал нет покупателя и нет!</node>
    <node id="n4" name="sentence" fuzzy="1">Никто в той стране знать не знал, слухом не слыхивал, для чего сандал употребляют.</node>
    <node id="n5" name="sentence" fuzzy="1">Проведая он, что тамошние люди у дровосеков уголь покупают.</node>
    <node id="n6" name="sentence" fuzzy="1">Тогда обращает глупец весь сандал в уголь, сбывает его по цене угля и торжествующий возвращается домой.
  
```



```

</node> <node id="n7" name="sentence" fuzzy="1">Всюду растрезвонил он, как умно с делами
управился и доставил тем народу потеху немалую.</node>
<link id="l1" source="n1" target="n2" type="S" fuzzy="1"/>
<link id="l2" source="n1" target="n3" type="S" fuzzy="1"/>
<link id="l3" source="n2" target="n3" type="S" fuzzy="1"/>
<link id="l4" source="n3" target="n4" type="S" fuzzy="1"/>
<link id="l5" source="n1" target="n5" type="S" fuzzy="1"/>
<link id="l6" source="n2" target="n5" type="S" fuzzy="1"/>
<link id="l7" source="n3" target="n5" type="S" fuzzy="1"/>
<link id="l8" source="n3" target="n6" type="S" fuzzy="1"/>
<link id="l9" source="n4" target="n6" type="S" fuzzy="1"/>
<link id="l10" source="n5" target="n6" type="S" fuzzy="1"/>
<link id="l11" source="n1" target="n7" type="S" fuzzy="1"/>
<link id="l12" source="n2" target="n7" type="S" fuzzy="1"/>
<link id="l13" source="n3" target="n7" type="S" fuzzy="1"/>
<link id="l14" source="n5" target="n7" type="S" fuzzy="1"/>
</graph>
</text>

```

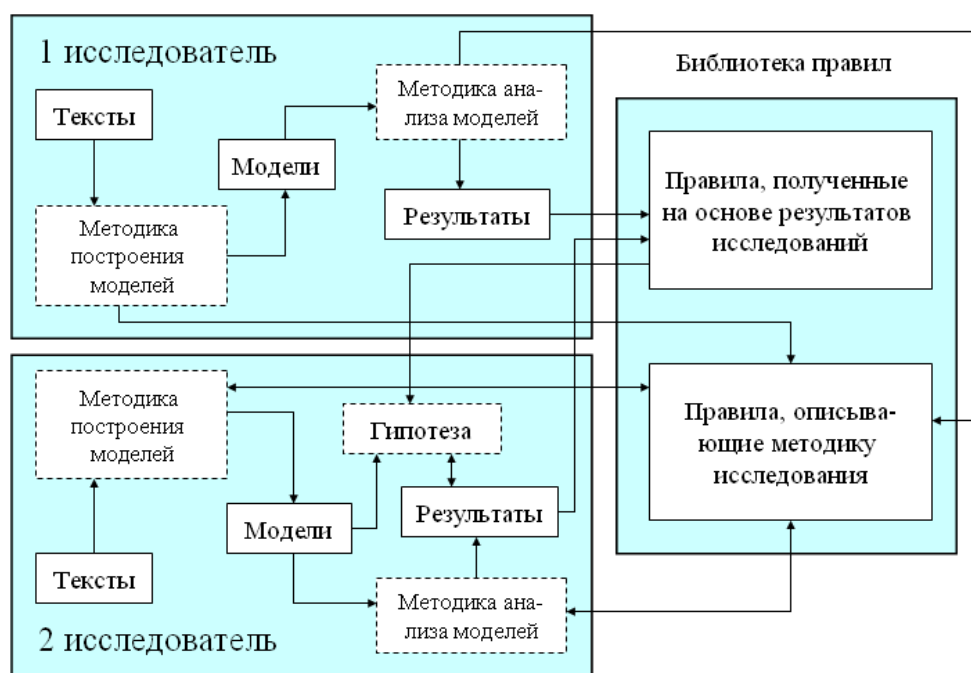


Рисунок 10 – Использование библиотеки правил для хранения результатов и методик исследования

Для хранения и публикации полученных в результате исследования закономерностей организации текста, а также правил, по которым строятся различные виды семантических сетей, можно использовать формат RuleML (<http://www.ruleml.org/>), основанный на XML [Варфоломеев и др., 2008]. Рассмотрим пример: пусть в результате исследования оказалось, что в среднем бесёдная песня состоит из 12 объектов и 14 связей. Песни, которые исполнялись в быстром темпе, в основном содержат большое количество объектов и связей (число вершин $m > 14$, а ребер $n > 17$). В эту группу входят семейные и любовные песни. Напротив, медленные песни характеризуются небольшим числом объектов и связей (число вершин $m < 14$, а ребер $n < 17$). В эту группу попали все песни на темы «свадьба» и «игра». Таким образом, можно определить знание следующего вида: «Если в графе песни число вершин $m > 14$ и число ребер $n > 17$, то эта песня с большой вероятностью исполнялась в быстром темпе». На языке RuleML данная закономерность может быть записана следующим образом:

<Implies>	<Rel>больше</Rel>
<head>	<Var>число вершин</Var>
<Atom>	<Var>14</Var>
<Rel>имеет темп</Rel>	</Atom>
<Var>песня</Var>	<Atom>
<Var>быстрый</Var>	<Rel>больше</Rel>
<Var>с большой вероятностью	<Var>число ребер</Var>
</Var>	<Var>17</Var>
</Atom>	</Atom>
<head>	</And>
<body>	<body>
<And>	</Implies>
<Atom>	

Схема взаимодействия нескольких исследователей, использующих общую библиотеку правил для хранения результатов и методик исследования, изображена на рис. 10. Дальнейшее использование данных форматов может быть полезно для организации распределенных научных исследований в сети Интернет и публикации их результатов в рамках деятельности сетевых сообществ исследователей [Варфоломеев и др., 2008].

Библиографический список

[Варфоломеев и др., 2008] Варфоломеев, А. Г. Применение RuleML для представления и вывода знаний о семантической структуре фольклорных текстов, полученных на основе их теоретико-графовых моделей / А.Г. Варфоломеев, Н.В. Каргинова, И.В. Кравцов, Н.Д. Москин // Труды XI национальной конференции по искусственному интеллекту с международным участием КИИ-2008. – Т. 2. – М.: ЛЕНАНД, 2008. – С. 183-191.

[Варфоломеев и др., 2010] Варфоломеев, А. Г. О проекте виртуальной среды для исследования списков «Беседы трёх святителей» / А.Г. Варфоломеев, М.Г. Бабалык, А.В. Пигин // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды XII Всероссийской конференции RCDL'2010. – Казань, 2010. – С. 551-556.

[Люгер, 2005] Люгер, Джордж Ф. Искусственный интеллект: стратегии и методы решения сложных проблем / Джордж Ф. Люгер, 4-е издание. Пер. с англ. – М.: Издательский дом «Вильямс», 2005.

[Москин, 2006а] Москин, Н. Д. Свидетельство об официальной регистрации программы для ЭВМ «Информационная система по исследованию фольклорных коллекций с теоретико-графовой формализацией текстов» № 2006612744. Федеральная служба по интеллектуальной собственности, патентам и товарным знакам (Роспатент), 03.08.2006.

[Москин, 2006б] Москин Н. Д. Теоретико-графовые модели структуры фольклорных текстов, алгоритмы поиска закономерностей и их программная реализация : дис. ... канд. техн. наук : защищена 03.11.2006 / Н. Д. Москин. – Петрозаводск, 2006.

[Москин, 2009а] Москин, Н. Д. Алгоритмы сравнения графов и теоретико-графовых моделей: учебное пособие / Н.Д. Москин. – Петрозаводск: Изд-во ПетрГУ, 2009.

[Москин, 2009б] Москин, Н. Д. Решение задач визуализации и поиска мотивов в электронной библиотеке фольклорных текстов / Н.Д. Москин // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды XI Всероссийской конференции RCDL'2009. – Петрозаводск: КарНЦ РАН, 2009. – С. 465-471.

[Москин, 2010а] Москин, Н. Д. Инструменты исследования текстовых коллекций на основе теоретико-графовых моделей в информационной системе «Фольклор» / Н.Д. Москин // Прикладная информатика, №4(28), 2010. – С. 48-62.

[Москин, 2010б] Москин, Н. Д. Применение нечетких теоретико-графовых моделей в задачах моделирования и поиска песенных мотивов / Н. Д. Москин // Труды Двенадцатой национальной конференции по искусственному интеллекту с международным участием КИИ-2010. Тверь, 20-24 сентября 2010 г. – М.: Физматлит, 2010. Т. 1 – С. 243-251.

[Скороходько, 1983] Скороходько, Э. Ф. Семантические сети и автоматическая обработка текста / Э.Ф. Скороходько. – Киев: Наукова думка, 1983.

[Хроленко, 1992] Хроленко, А. Т. Семантика фольклорного слова / А.Т. Хроленко. – Воронеж: Издательство Воронежского университета, 1992.