**Лунев Кирилл Владимирович Теоретико-графовые алгоритмы выявления семантической близости между понятиями на основе анализа наборов ключевых слов взаимосвязанных объектов**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Лунев Кирилл Владимирович

Введение

Глава 1. Методы и средства анализа информации с использованием

ключевых слов

1.1 Библиографический обзор

1.1.1 Методы определения близости между парой слов естественного языка

1.1.2 Методы определения близости между объектами в графах знаний

1.1.3 Графовые методы кластеризации слов естественного языка

1.1.4 Выводы из библиографического обзора

1.2 Методология

1.3 Экспертное оценивание качества результатов программных реализаций

Глава 2. Определение смысловой близости пары ключевых слов

2.1 Модель семантической близости WordContSim

2.1.1 Построение графа ключевых слов

2.1.2 Контекстная модель определения семантической близости

для пары ключевых слов

2.1.3 Алгоритм вычисления значения контекстной близости по коллекции ключевых слов

2.1.4 Тестовые испытания

2.1.5 Выводы

2.2 Использование методов машинного обучения для улучшения

модели близости слов. Модель WordMLSim

2.2.1 Методы формирования обучающей выборки

2.2.2 Признаковое описание модели машинного обучения

2.2.3 Тестовые испытания

2.3 Выводы

Стр.

Глава 3. Определение смысловой близости пары наборов ключевых

слов

3.1 Модель определения смысловой близости наборов ключевых слов

3.2 Алгоритм определения уровня близости пары наборов, основанный на переборе всех пар ключевых слов

3.3 Оптимизированный алгоритм определения близости пары наборов

3.4 Тестовые испытания

3.5 Выводы

Глава 4. Приложения моделей близости ключевых слов

4.1 Модель семантической кластеризации ключевых слов

4.1.1 Модель полного контекстного графа ключевых слов

4.1.2 Модель и алгоритм построения усеченного контекстного графа ключевых слов

4.1.3 Модель кластеризации усеченного контекстного графа

4.1.4 Алгоритм кластеризации усеченного контекстного графа

4.1.5 Тестовые испытания

4.2 Определение тематической направленности объекта информационной системы по набору ключевых слов

4.2.1 Определение степени абстрактности слова

4.2.2 Алгоритм определения тематических ключевых слов

4.2.3 Алгоритм выбора тематики объекта

4.2.4 Тестовые испытания

4.2.5 Выводы

4.3 Решение задачи поиска экспертов

4.3.1 Постановка задачи

4.3.2 Процедура поиска экспертов

4.4 Построение тезауруса ключевых слов по коллекции наборов

4.5 Реализация поиска по ключевым словам на базе собранного тезауруса синонимов

4.6 Решение задачи поиска экспертов для графов знаний

4.6.1 Выборка данных

4.6.2 Тестовые испытания

4.6.3 Выводы

Стр.

4.7 Соответствие программного модуля интеллектуального анализа

на основе ключевых слов предъявляемым требованиям

4.8 Выводы

Заключение

Список литературы

Приложение А. Требования к качеству программной системы

анализа ключевых слов

А.1 Функциональные требования

А.2 Надежность

А.3 Практичность

А.4 Эффективность

А.5 Сопровождаемость

А.6 Мобильность

Приложение Б. Самые абстрактные по смыслу слова для каждой

меры центральности

Приложение В. Найденные в коллекции документов тематические теги