

Афонин С.А., к.ф.-м.н.

Краткое описание возможных направлений курсовых работ

кафедра вычислительной математики, доцент
НИИ механики МГУ, ведущий научный сотрудник

20 апреля 2020 г.

Область моих интересов можно разделить на две группы: задачи, которые имеют прикладной характер, и задачи, которые его лишены.

Направления первой группы связаны с системой учета и анализа научных публикаций ИСТИНА (150 тысяч пользователей):

- обработка текстов на естественном языке;
- анализ социальных сетей;
- информационная безопасность.

Вторая группа включает:

- теорию формальных языков и
- моделирование когнитивных процессов.

Некоторые задачи обработки и интеллектуального анализа текстов:

- выполнение поисковых запросов «сложнее, чем у Гугла»;
- описание предметной области (ключевые слова, примеры документов, формальные онтологические модели);
- тематический анализ: принадлежность документа теме, выявление более «мелких» тем;
- выявление трендов: «В этом месяце заговорили о возможности применения графена в программировании!»;
- оценка тональности.

Методы решения: комбинация методов *вычислительной семантики* (нейронные сети) и методов *основанных на знаниях* (онтологии).

Анализ социальных сетей

Научные публикации как соцсеть: посты (статьи), авторы, лайки (цитирование), сообщества (соавторы, коллеги), ресурсы (журналы).

Цель: разработать методы поиска «хороших» документов, авторов, коллективов, конференций. Критерии оценки могут включать:

- наличие у автора «известных» публикации;
- принадлежность автор «признанному» сообществу;
- наличие ссылок на ключевые работы в области;
- усредненную оценку других документов в этом источнике;
- актуальность и новизну публикации (да, это оценивают).

Методы: вычисление различных характеристик на графах очень большого размера. Объем данных: десятки тысяч документов в сутки.

Логическое разграничение доступа: проверить, что запрашиваемая пользователем операция с объектом соответствует политике безопасности (заданной в виде текста).

Задачи:

- разработать формальную модель и язык описания политики безопасности;
- разработать систему, которая проверяет, что пользователь имеет право выполнять операцию;
- разработать методы доказательства корректности политики безопасности.

Условия доступа должны допускать использование свойства объекта и субъекта доступа: «Нельзя удалять чужие файлы больше 10 мегабайт.»

- Теория формальных языков:
 - методы описания множеств, элементами которых являются формальные языки;
 - алгоритмы факторизации регулярных языков;
 - автоматные методы анализа политик информационной безопасности.
- Моделирование когнитивных процессов:
 - разработка алгоритмов решения комбинаторных задач, соответствующих результатам психологических исследований их решения человеком.
- Самообучающиеся системы.

Требования к студентам

Требования носят субъективный характер.

- Заинтересованность.
- Готовность тратить на курсовую работу больше времени, чем это обычно требуется.
- Готовность регулярно встречаться.
- Склонность к программированию, знание C/C++, готовность изучать новые языки.
- Отсутствие задолженности по курсу «Работа на ЭВМ и программирование» (желательно).

Афонин Сергей Александрович

к.ф.-м.н., доцент кафедры вычислительной математики, в.н.с. НИИ механики МГУ

serg@msu.ru

serg.tk/contacts/

Научные интересы: обработка текстов на естественном языке, информационный поиск, теория формальных языков, моделирование когнитивных процессов, анализ социальных сетей, информационная безопасность.

Публикации <https://istina.msu.ru/profile/safonin/>

