

## ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертацию Гомзина Андрея Геннадьевича

«Методы и программные средства определения значений стационарных демографических атрибутов пользователей социальных сетей», представленную на соискание учёной степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Определение значений демографических атрибутов пользователей является одной из задач анализа социальных сетей. Значения пола, возраста, рода деятельности, семейного положения и других атрибутов пользователей имеют важное значение в различных прикладных задачах, начиная от рекламы в коммерческих компаниях, заканчивая анализом демографических тенденций общества на государственном уровне. Социальные сети предоставляют пользователям возможность вести персональные страницы, которые могут содержать информацию о значениях демографических атрибутов пользователей. На практике лишь небольшая доля значений указывается в профилях, кроме того, не всем указанным значениям можно доверять. Отсюда возникает задача предсказания значений демографических атрибутов пользователей социальных медиа по доступным данным и необходимость специальных программных средств для решения этой задачи.

Актуальность работы обусловлена выявленными недостатками существующих методов предсказания значений демографических атрибутов. Отмечаются проблемы таргетированного сбора текстов сообщений для заданных пользователей. Кроме того, в работе описываются случаи, когда существующие методы неверно предсказывают значения атрибутов при использовании социального графа. Исследование направлено на устранение выявленного недостатка методов предсказания значений атрибутов с использованием социального графа.

В процессе научной работы Гомзин А.Г. провёл тщательный обзор предметной области, начиная с ранних работ, посвящённых разработке инструментов для автоматического анализа текстов личных дневников и электронных писем, и заканчивая инструментами анализа текстов и социального графа в современных социальных сетях. В тексте работы явно обозначены недостатки существующих методов. Был предложен и обоснован подход для предсказания значений атрибутов по социальному графу, заключающийся в повышении значимости связи с вершиной, у которой большинство соседей имеют схожее значение демографического атрибута. Такой подход ранее не использовался в научных работах. В рамках подхода было разработано 5 методов предсказания значений демографических атрибутов пользователей социальных сетей по социальному графу. Было проведено экспериментальное сравнение разработанных методов с существующими аналогами и показано превосходство

разработанных методов над аналогами по качеству предсказания значений атрибутов. Была проведена теоретическая оценка вычислительной сложности методов и представлено фактическое время работы реализованных в рамках программной системы методов. На основе исследования вычислительной сложности, времени и качества работы методов были даны рекомендации к их применению. В работе приведено описание программной системы, которая позволяет как проводить экспериментальные исследования методов определения значений демографических атрибутов, так и решать прикладные задачи в составе программных комплексов анализа социальных сетей.

Результаты работы опубликованы в 5 печатных изданиях, получено 5 свидетельств о регистрации программ для ЭВМ. Гомзиным А.Г. были успешно сделаны доклады на нескольких конференциях, в том числе в 2021 году на международной конференции «55th Annual Conference on Information Sciences and Systems (CISS)».

В процессе работы Гомзиным А.Г. было продемонстрировано умение работать с большим объемом литературы, навыки ведения научного исследования, компетенции в сфере программирования. Результаты работы применяются в научно-исследовательских проектах отдела «Информационных систем» ИСП РАН. Кроме того, Гомзин А.Г., будучи аспирантом ВМК МГУ, активно работал со студентами, руководил подготовкой их курсовых и дипломных работ.

В целом научная работа Гомзина А.Г. является законченным научным исследованием, имеющим теоретическое и практическое значение. Диссертация соответствует всем требованиям, предъявляемым ВАК к диссертационным работам на соискание учёной степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, а её автор, Гомзин Андрей Геннадьевич, заслуживает присуждения ему этой степени.

Научный руководитель:  
зав. отделом ИСП РАН, к.ф.-м.н.  
31.03.2021

Турдаков Денис Юрьевич