

О Т З Ы В

на автореферат диссертации Исхаковой Анастасии Олеговны «Метод и программное средство определения искусственно созданных текстов», представленной на соискание ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики»

Методы создания искусственных текстов широко применяются для создания информационного контента, позволяя распространять уникальные публикации любого содержания. Проблема анализа и выявления искусственных текстов имеет громадное значение, поскольку на основе использования интернет-технологий эти тексты могут использоваться в различных, в том числе, неправомерных целях. Диссертационная работа А.О. Исхаковой посвящена разработке методов определения искусственных текстов, что позволяет решать задачи автоматизированной фильтрации интернет-контента и идентифицировать вредоносные текстовые сообщения.

Тематика работы представляется актуальной и соответствует специальности 05.13.17 – «Теоретические основы информатики». При решении задач исследования автором используются методы текстовой атрибуции, методы статистического анализа.

Исхаковой А.О. предложен алгоритм формирования инвариантов классов текстов, основанный на использовании качественных и уточняющих их количественных характеристик текста, при этом наборы исследуемых характеристик зависят от конкретной задачи. Инвариант позволяет формализовать отличия текстов одного класса и представляет собой модель данного класса в виде набора значений определенных характеристик текста. Предложенный алгоритм в отличие от существующих снижает риск упущения текстовых характеристик, которые отсутствуют в стандартных наборах, но в конкретном случае могут обладать различительной способностью. С использованием предложенного алгоритма были получены инварианты для определения искусственных текстов, обладающие различительной способностью при определении происхождения текста. Основным результатом работы является метод определения искусственных текстов, отличающийся использованием меры принадлежности исходного текста к известным классам текстов. Кроме того, разработана программа для автоматизации процесса обнаружения и последующей фильтрации искусственных текстов на основе веб-ориентированной технологии. Результаты экспериментов, приведенные в автореферате, показали, что разработанный метод в подавляющем большинстве случаев позволяет определить искусственные тексты.

Компьютерная программа и включенные в нее модули для систем управления контентом внедрены в деятельность ООО «Агентство

медиарешений», ООО «Лингвистические и информационные технологии», а также используются в учебном процессе Томского государственного университета систем управления и радиоэлектроники.

К замечаниям по работе можно отнести следующее: поскольку мера принадлежности входного текста к известным классам определяется на основе функции расстояния, то целесообразно было бы использовать параметрические функции расстояния, что позволило бы за счет настройки параметра на обучающем множестве «усилить» различительные способности метода определения искусственных текстов.

Приведенное замечание не влияет на общее положительное впечатление от диссертационной работы, которая удовлетворяет требованиям ВАК к кандидатским диссертациям. Имеется достаточное количество публикаций, результаты работы внедрены. Считаю, что Исхакова А.О. заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.17 – «Теоретические основы информатики».

Рецензент согласен на включение своих персональных данных в документы, связанные с работой диссертационного совета, и дальнейшую их обработку.

Воронина Ирина Евгеньевна

5 декабря 2016 г

Профессор кафедры программного обеспечения
и администрирования информационных систем,
д.т.н., доцент



И.Е. Воронина

Почтовый адрес: 394018, г. Воронеж, Университетская пл., 1

Телефон: 89036504410

Адрес электронной почты: irina.voronina@gmail.com

федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет» (ФГБОУ ВО «ВГУ»)	
Подпись	<i>Ворониной И.Е.</i>
заверяю	<i>Секретарь Ф.П. ПММ</i>
	Функция
<i>А.С. Смирнов</i>	<i>5.12.16</i>
подпись, расшифровка подписи	

