УДК 004.896

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СЕРВИСА ВЫЯВЛЕНИЯ УТЕЧЕК ПЕРСОНАЛЬНЫХ ДАННЫХ НА ОСНОВЕ НЕЙРОСЕТЕВОЙ МОДЕЛИ

Бондаренко А.Л., Трофимов М.И., Ткаченко С.Ф.
Научный руководитель – кандидат технических наук Менисов А.Б.
(Военно-космическая академия имени А.Ф.Можайского)

Аннотация. В современных условиях стремительно растет и развивается киберпреступность. Злоумышленники используют персональные данные сотрудников организаций на подготовительных и начальных стадиях компьютерных атак. В докладе представлена схема разработанного программного комплекса для выявления сущностей персональных данных, основанный на объединении нескольких алгоритмов машинного обучения: нейросетевой рекуррентной архитектуры двунаправленной долгой краткосрочной памяти, кодировщика персональных данных и метода ближайших соседей. Результаты оценивания показателей эффективности в сравнении с современным средством обработки текстов естественного языка (Spacy) показали перспективы практического применения программного комплекса.

Введение. Современный этап развития общества характеризуется возрастающей ролью информационной сферы, представляющей собой совокупность данных, информационной инфраструктуры, субъектов, осуществляющих сбор, формирование, распространение и использование данных, а также системы регулирования возникающих при этом общественных отношений. Информационная сфера, являясь системообразующим фактором жизни общества, активно влияет на состояние политической, экономической, оборонной и других составляющих национальной безопасности. В свою очередь, развитие информационных технологий привело к преобразованию понимания личного пространства и частной жизни. Процессы, ранее происходившие в физическом (реальном) мире, перетекли в онлайн-среду: электронная торговля, поисковые сервисы, социальные сети, распространение планшетов и смартфонов, дающих людям возможность постоянно находиться онлайн. Вследствие этого объемы персональных сведений, которые человек раскрывает и выкладывает в глобальную сеть, как и персональных данных граждан, собираемых и систематизируемых разными учреждениями и ведомствами, увеличились многократно. Таким образом, одной из значимых информационных угроз является возможность использования персональных данных на подготовительных и (или) начальных этапах атак на информационную. Вместе с тем, модели, методы, технологии, технические и программные средства выявления и удаления персональных данных из открытых источников информации являются недостаточно совершенными из-за низкой результативности использования методов поиска такой информации данных, а также отсутствием глобального охвата открытого сегмента информации. В рамках настоящего исследования проблемная ситуация сформулирована как необходимость обеспечения эффективного выявления угроз утечек персональных данных с html-страниц на основе разработки программного комплекса выявления и удаления персональных данных из открытых источников информации.

Основная часть. Инновационность программного комплекса, заключается в модуле выявления, включающий новый подход применение трех алгоритмов машинного и глубинного обучения:

1) нейросетевой рекуррентной архитектуры двунаправленной долгой краткосрочной памяти (BLSTM), которая показала улучшение качества выявления сущностей для решения других задач;

2) кодировщика персональных данных для создания образов данных;

3) метода ближайших соседей для классификации образов на основе сигнатурных;

Несмотря на высокую точность выявления персональных данных, трудно судить о ее релевантности. С этой целью было выделено 2 класса данных: относящиеся к объектам контроля и являющиеся утечкой, нейтральные данные, не несущие ущерба. Впредь под объектом контроля будет пониматься физическое лицо, проходящее проверку на утечку данных. Для решения задачи на выходе нейросетевой модели был размещен классификатор поточной информации, включающий в себя:

1) кодировщик персональных данных;

2) метод ближайших соседей;

Объединение кодировщика собранных данных и метода ближайших соседей, включает в себя следующие этапы:

1. создание выборки контроля и такое же количество образов нейтральной данных на основе модели векторизации считающего вектора;
2. кодирование собранной персональных данных на основе обученной модели считающего вектора;
3. обучение модели классификатора ближайших соседей на сигнатурных образах;
4. классификации собранных персональных данных.

Во время обучения модель запоминает сигнатурные образы объектов контроля, формируя в пространстве локализованные подмножества.

Для классификации каждого из полученных векторов данных необходимо последовательно выполнить следующие операции:

1. вычислить расстояние до каждого из известных образов
2. отобрать K образов, расстояние до которых минимально

Выводы. Огромный объем неструктурированных данных сети Интернет, распространяемый ежедневно, вызывает потребность в разработке эффективных методов поиска и извлечения данных. Парирования утечек персональных данных – сложная задача классификации для текстов естественного языка, которая еще более усложняется при применении к html-страницам из-за особых свойств и сложной структуры. В докладе представлен новый подход глубокого обучения для выявления персональных данных, который доказал свою эффективность по сравнению с другими.

Основная цель разработки программного комплекса – предоставить более детализированные результаты для практического применения в области обработки естественного языка и информационной безопасности. В разработанном подходе использована нейросетевая технология двунаправленной долгой краткосрочной памяти в сочетании с многоязычным универсальным кодировщиком предложений, а также метода ближайших соседей.