

УДК 004.056

**РАЗРАБОТКА СИСТЕМЫ ОБНАРУЖЕНИЯ ФИШИНГОВЫХ ВЕБ-ПРИЛОЖЕНИЙ  
НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ  
ФАКТОРИЗАЦИИ МАТРИЦ**

**Челноков Н.П.** (Санкт-Петербург, Университет ИТМО)

**Научный руководитель – к.т.н., доцент Менщиков А.А.**  
(Санкт-Петербург, Университет ИТМО)

В настоящее время развитие веб-приложений и веб-технологий, используемых в разработке, происходит с большой скоростью. Вместе с технологиями разработки происходит и совершенствование техник и методов атак, позволяющих потенциальным злоумышленникам получить доступ к критичным данным пользователей сети Интернет (например, данные платежных карт). В данной работе представлена система обнаружения фишинговых веб-приложений, в основе которой лежит классификация методами машинного обучения, выполненная с использованием комплексного набора метрик.

**Введение.** В существующих работах представлены различные способы обнаружения фишинговых веб-приложений с использованием машинного обучения. Данные технологии в большинстве основываются на анализе следующих характеристик веб-приложений: URL-адреса, домены, сертификаты и хэш-суммы. Данные метрики не являются лишними или устаревшими, однако возможности, доступные злоумышленникам, расширяются вследствие разработки новых веб-технологий и обнаружения новых уязвимостей в старых. Таким образом, имеющиеся подходы не обеспечивают достаточный уровень безопасности при анализе веб-приложений на наличие фишинг-контента.

**Основная часть.** Атаки и техники, которые используются в фишинговых веб-приложениях, являются вариативными и обладают большим количеством мутаций. Вследствие данного обстоятельства, существующие системы имеют низкую точность в обнаружении данных веб-приложений. В данной работе проведен анализ современных техник, которые злоумышленники используют для создания и внедрения фишинг-контента в веб-приложения. Помимо основных метрик веб-приложений (например, URL-адрес и домен) были рассмотрены комплексные и динамические, такие как подгружаемый код, связанные с приложением страницы, исходный код, динамический контент и другие. На основе комплекса критериев был выполнен сравнительный анализ классификаторов для решения задачи обнаружения фишинговых веб-приложений, а также методов машинного обучения, основанных на использовании факторизации матриц.

**Выводы.** Разработанная система позволяет выполнять более точный и эффективный анализ веб-приложений на наличие фишинг-контента, а также является гибкой и расширяемой вследствие использования методов машинного обучения и уникальных метрик. Дальнейшее расширение системы не представляет сложности и позволяет адаптироваться к изменяющимся внешним условиям. Данная система является эффективным и перспективным решением, которое может использоваться в любых современных системах обнаружения фишинговых веб-приложений.

Челноков Н.П. (автор)

Подпись

Менщиков А.А. (научный руководитель)

Подпись