**Менщиков Александр Алексеевич Методы обнаружения и противодействия автоматизированному сбору информации с веб-ресурсов**

ОГЛАВЛЕНИЕ ДИССЕРТАЦИИ

кандидат наук Менщиков Александр Алексеевич

СОДЕРЖАНИЕ

РЕФЕРАТ

SYNOPSIS

ВВЕДЕНИЕ

ГЛАВА 1. СОВРЕМЕННОЕ СОСТОЯНИЕ ПРОБЛЕМЫ АВТОМАТИЗИРОВАННЫХ УГРОЗ БЕЗОПАСНОСТИ ВЕБ-РЕСУРСОВ

1.1. Угрозы и атаки на веб-ресурсы

1.1.1 Атаки на процессы аутентификации и авторизации

1.1.2 Атаки на клиентов

1.1.3 Исполнение кода на веб-сервере

1.1.4 Логические атаки

1.1.5 Атаки на отказ в обслуживании

1.1.6 Спам

1.1.7 Разглашение информации

1.1.8 Автоматизированные атаки

1.2 Автоматизированные угрозы

1.3. Автоматизированный сбор информации с веб-ресурсов

1.4. Классификация веб-роботов

1.4.1 Структура запроса

1.4.2 Назначение веб-роботов

1.4.3 Политика повторного посещения

1.4.4 Мотивация веб-роботов

1.5. Проблемы автоматизированного сбора информации

1.6. Классификация методов обнаружения веб-роботов

1.7. Проблемы обнаружения автоматизированного сбора информации

1.7.1 Идентификация сессий

1.7.2 Изменяющийся User-Agent

1.8. Базовые принципы обнаружения веб-роботов

1.9. Правовое регулирование веб-роботов

1.10. Выводы

ГЛАВА 2. МЕТОДЫ ОБНАРУЖЕНИЯ ВЕБ-РОБОТОВ

2.1. Задача обнаружения веб-роботов

2.2 Анализ поведения пользователей и автоматизированных средств

2.3 Анализ методов обнаружения веб-роботов

2.4 Анализ структуры веб-ресурса

2.4.1 Создание графа связности веб-ресурса

2.5 Анализ структуры веб-роботов

2.5.1 Обобщённая архитектура веб-роботов

2.5.2 Распределённые веб-роботы

2.6. Методы обнаружения веб-роботов на основе структурных характеристик

2.6.1 Использование Марковского алгоритма кластеризации

2.6.2 Исследование шаблонных запросов файлов разных типов

2.6.3 Анализ графа запросов сессии

2.6.4 Байесовский подход к обнаружению

2.6.5 Способ онлайн обнаружения

2.6.6 Предсказание поведения веб-роботов

2.6.7 Использование наивного Байесовского классификатора

2.6.8 Использование деревьев решений

2.6.9 Использование нейронных сетей

2.7. Характеристики, используемые для обнаружения автоматизированных угроз

2.7.1 Метод на основе выделения функциональных признаков страниц

2.7.2 Метод на основе содержимого файлов

2.7.3 Метод на основе типа ресурса

2.8. Задача извлечения информации с веб-ресурса

2.9. Методы обнаружения веб-роботов

2.9.1. Синтаксический анализ логов

2.9.2. Поведенческий анализ

2.9.3. Аналитические методы

2.10. Программные методы обнаружения и противодействия

2.10.1. Блокировка источника запросов

2.10.2. Использование учётных записей

2.10.3. Использование CAPTCHA

2.10.4. Использование сложной JavaScript логики

2.10.5. Динамическое изменение структуры страницы

2.10.6. Ограничение частоты запросов и объема загружаемых данных

2.10.7. Защита отображения данных

2.11 Характеристики обнаружения

2.11.1 Характеристики HTTP запроса

2.11.2 Временные характеристики

2.11.3 Поведенческие характеристики

2.12 Метод обнаружения веб-роботов на основе семантического анализа

посещённых страниц

2.13 Метод обнаружения веб-роботов на основе анализа графа пользовательского поведения

2.13.1 Построение графа веб-ресурса

2.13.2 Характеристики сессий

2.14. Выводы

ГЛАВА 3. РАЗРАБОТКА СИСТЕМЫ ОБНАРУЖЕНИЯ И ПРОТИВОДЕЙСТВИЯ АВТОМАТИЗИРОВАННОМУ СБОРУ ИНФОРМАЦИИ С ВЕБ-РЕСУРСОВ

3.1. Алгоритм обнаружения и противодействия веб-роботам

3.1.1. Этап инициализации

3.1.2. Этап обучения

3.1.3. Этап обнаружения

3.1.4. Этап реагирования

3.2. Методика подготовки данных для обучения системы обнаружения веб-роботов

3.3. Меры по противодействию автоматизированным угрозам

3.4. Вопрос онлайн обнаружения

3.5. Архитектура системы обнаружения и противодействия веб-роботам

3.6. Выводы

ГЛАВА 4. ПРАКТИЧЕСКИЕ АСПЕКТЫ РЕАЛИЗАЦИИ КОМПЛЕКСНОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ И ПРОТИВОДЕЙСТВИЯ

АВТОМАТИЗИРОВАННОМУ СБОРУ ИНФОРМАЦИИ С ВЕБ-РЕСУРСОВ

4.1 Сравнение облачных веб-роботов

4.2 Сравнение промышленных веб-роботов

4.3 Используемые наборы данных

4.4 Архитектура экспериментального стенда

4.5 Архитектура системы реагирования

4.6 Архитектура подсистемы хранения и обработки информации

4.7 Оценка результатов обнаружения веб-роботов

4.8. Выводы

ЗАКЛЮЧЕНИЕ

СПИСОК РИСУНКОВ

СПИСОК ТАБЛИЦ

СПИСОК ЛИТЕРАТУРЫ

ПРИЛОЖЕНИЯ

Приложение А. Акты о внедрении

Приложение В. Результаты интеллектуальной деятельности

Приложение С. Основные публикации по теме диссертации

РЕФЕРАТ