

«Система защиты авторства материалов образовательных курсов на онлайн-платформах»

Ванин Дмитрий Михайлович

Введение

- После покупки злоумышленник может совершить незаконное копирование
- В блоге онлайн школы «Dr. Study» продемонстрирован явный запрос на решение проблемы кражи материалов курсов
- С 2020 года растёт популярность запроса «слив курса» в поисковой системе Google

Целевая аудитория

Составители курсов, размещающие материалы на платформах в сети Интернет

Проблема

Проблема идентификации субъекта, совершившего незаконное копирование материалов курса: текста и изображений.

Аналоги

- Система watermark платформы GetCourse
- Steghide

Целевая аудитория

Авторы курсов, распространяющие свои материалы в сети Интернет

Цель

Создание системы защиты авторства материалов,
в виде расширения, предназначенного для
использования с LMS

Задачи

- Обеспечение обнаружения копирования текстовой информации
- Обеспечение обнаружения копирования растровых изображений

Используемые материалы

- Язык программирования python
 - Чёткий синтаксис
 - Популярный
- Библиотека pillow для обработки изображений
 - Достаточный функционал
 - Активная поддержка

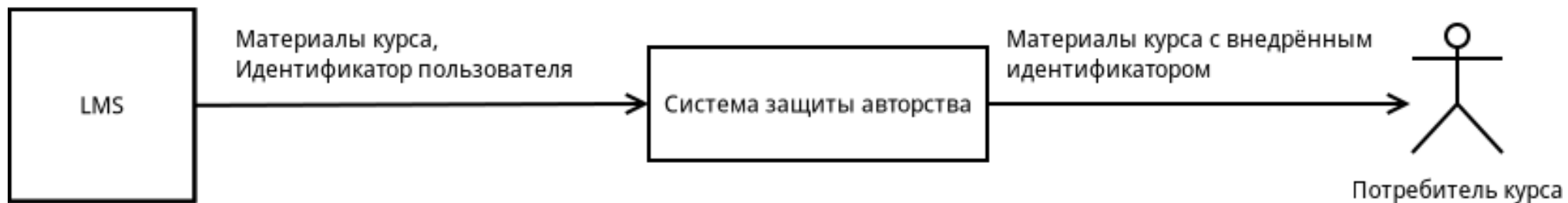
Этапы работы над проектом

1. Разработка алгоритмов, исполняющих задачи
2. Реализация алгоритмов
3. Создание расширения для LMS Moodle
4. Написание пользовательской документации
5. Проведение функционального тестирования
6. Проведение пользовательского тестирования

Тестирование

- Функциональное
 - Запуск реализации алгоритма на определённых данных и сверка результата с ожидаемым
- Пользовательское
 - Группа, сформированная из целевой аудитории, опробует проект, после чего даст обратную связь

Схема работы ПО



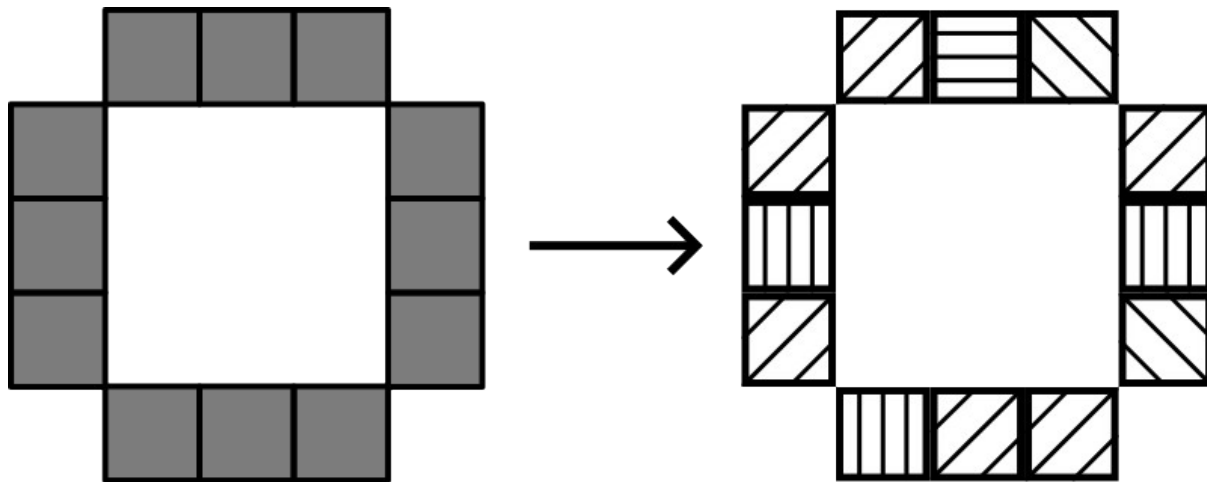
Алгоритм для текста

1. В текст внедряются символы, идентичные по форме с оригинальными, но при этом имеющие другое значение в таблице символов Unicode
2. Для каждого текста составляется словарь синонимов для союзов, существительных и прилагательных
3. С помощью n -ичного кода (от 2 до 16) происходит кодирование подписанного идентификатора в тексте, с алфавитом на основе символов из первого пункта
4. С помощью n -ичного кода (от 2 до 16) происходит кодирование подписанного идентификатора в тексте, с алфавитом на основе синонимов из второго пункта

Алгоритм для растровых изображений

1. Подписанный идентификатор пользователя записывается в наименее значащие биты цветов картинки
2. Изображение разбивается на области 4x4, 8x8 16x16, области, в которых сплошное заполнение не менее 70%, заменяются штриховкой.
 1. Новая заливка представляет собой разную штриховку, с помощью которой n-ичным кодом кодируется подписанный идентификатор пользователям

Демонстрация алгоритма



$$7946629_{10} = 132110012011_4$$



Вывод

Проект решит проблему идентификации пользователей, совершающих незаконное копирование, не требуя при этом значительных ресурсов и проведения дополнительных мероприятий.

Так как происходит подпись идентификатора ЭЦП по ГОСТ 34.10-2018, по статье 6 ФЗ «Об электронной подписи» от 06.04.2011 N 63-ФЗ идентификатор пользователя может быть признан сведением о правонарушителе, после чего злоумышленник может быть наказан по статье 1301 ГК РФ.

Использованные источники

Dr Study ищет решение для обеспечения конфиденциальности своих курсов URL // Dr Study: <http://dr-study.online/blog.php> (Дата обращения: 07.02.2024)

Статистика по запросу «слив курса» // Google Trends: <https://trends.google.com/trends/explore?date=all&geo=RU&q=%D1%81%D0%BB%D0%B8%D0%B2%20%D0%BA%D1%83%D1%80%D1%81%D0%B0&hl=en> (Дата обращения: 07.02.2024)

Спасибо за внимание!
Готов ответить на ваши вопросы