

АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ СТЕГАНОГРАФИЧЕСКОГО МЕТОДА ДЛЯ ФАЙЛОВ-КОНТЕЙНЕРОВ SVG ФОРМАТА

Защита цифровых электронных документов является критически важной частью, необходимой для их безопасного хранения и передачи. Одним из способов обеспечения безопасности информации является использование технологии цифрового водяного знака (ЦВЗ), которая может использовать стеганографические методы. Эти методы предполагают скрытое внедрение сообщения в документ [1].

Для различных форматов цифровых документов требуется разработка специфических методов защиты из-за уникальных характеристик каждого формата. Формат масштабируемой векторной графики (SVG) позволяет легко манипулировать объектами благодаря своей структуре, основанной на языке разметки XML. Этот язык описывает графические объекты с помощью координат ключевых точек в значениях атрибутов и предоставляет широкий выбор элементов для внедрения информации. Для оценки эффективности метода целесообразно провести анализ производительности операции встраивания сообщения в контейнер.

Рассмотрен метод внедрения информации в практически любую фигуру файла формата SVG с использованием программного обеспечения. Метод, описанный в работе [2], позволяет внедрять данные в пустое пространство изображения, используя уникальные свойства описания фигур векторной графики. Это возможно благодаря тому, что объекты, созданные вне рабочей области, где находится геометрический объект, не видны пользователю, и это свойство используется для внедрения скрытого сообщения. В данном методе также используется дополнительное вычисление для встраивания информации вместо прямого встраивания закодированных и возможно зашифрованных данных.

Для исследования, исходя из представленного описания, целесообразно сосредоточиться на изучении и подборе потенциально оптимальных позиций, в которых будет содержаться информация, а также проверке влияния удаленности центральной точки от которой производится обход точки от ключевой точки. На первом этапе производилась проверка скорости работы алгоритма в зависимости от выбранных позиций.

Для первого исследования были установлены следующие параметры. Позиция центральной точки (30000000, 30000000), значения ограничения для видимой области составляли (0, 0, 21000000, 29700000) а значение ключевой точки было (29999127, 30000205) и внедряемой сообщение «89 conference». Для каждой комбинации значащих позиций угла проводилось 100 измерений. В данном исследовании использовались позиции с 3 по 9. Результаты замеров представлены в таблице 1.

Таблица 1 – Зависимость времени внедрения информации от числа значимых позиций для ключевой точки (29999127, 30000205)

Число значимых позиций	Время
1	17720
2	13762
3	51973
4	411850
5	4147540
6	471790447

Исходя из полученных данных можно заметить, что оптимальным числом позиций для встраивания является 2 исходя их временных затрат. На рис. 1 представлено распределение времени для двух позиций.

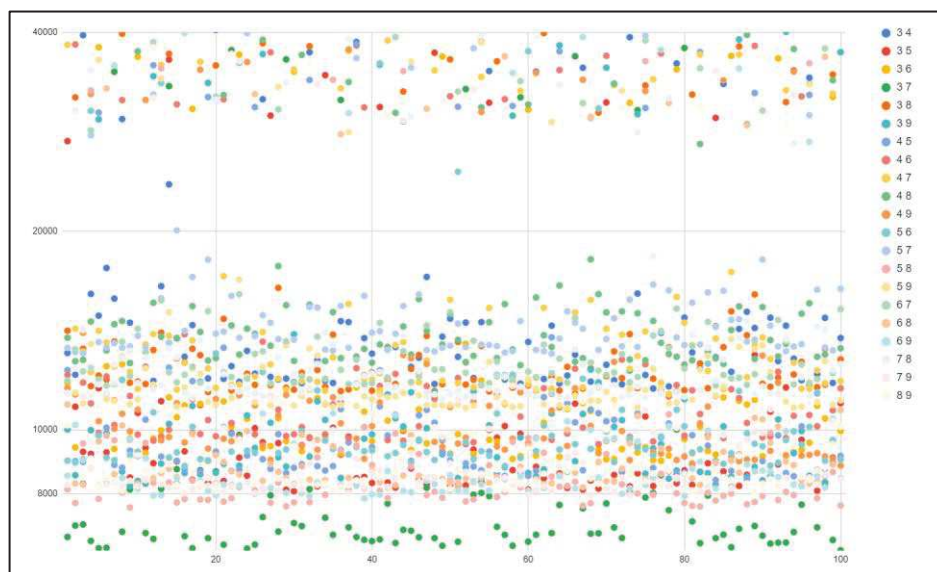


Рисунок 1 – Распределение времени для двух значащих позиций

Исходя из полученных результатов на рисунке можно заметить, что часть комбинаций значащих позиций выполняется заметно быстрее среднего значения. Данными комбинациями в данном случае являются (3,7), (5,8), (7,9) и (6,8) при значении которых время выполнения меньше 8500 тиков, что приблизительно 40 процентов быстрее.

Для проверки полученных результатов был произведён замер для

тех же параметров с изменяемой ключевой точкой. Результаты замера представлены в таблице 2.

Таблица 2 – Зависимость времени внедрения информации от числа значимых позиций для четырех ключевых точек

Ключевая точка	1	2	3	4	5
29993145, 29996451	7461	13147	54799	387044	3049986
30066732, 29917452	6699	13083	55770	447703	3351237
29248340, 30838147	19763	34889	79801	432563	3487119
39650041, 27590540	1163072	1928516	3502522	4178132	6925221

Полученные замеры времени для дополнительных значений ключевых точек позволяют сделать вывод о оптимальности в большинстве случаев в использовании двух значащих цифр в значении угла. Кроме того, с ростом удалённости ключевой точки от позиции центральной точки можно наращивать число значащих позиций в значении угла без сильной потери производительности вычислений. Для близко расположенных точек переход с двух значащих символов до трех, а также с трех до четырёх приводит к сильной потере производительности.

Также был проеден тест производительности с увеличенным сообщением, представляющим собой строку «89th Scientific and Technical Conference of the Faculty» для ранее упомянутых пяти ключевых позиций с сохранением параметров центральной точки, и видимой области. Результаты представлены в таблице 3.

Таблица 3 – Зависимость времени внедрения информации от числа значимых позиций для пяти ключевых точек с увеличенным сообщением

Ключевая точка	1	2	3	4	5
29999127, 30000205	53459	61728	227130	1778119	16721223
29993145, 29996451	22062	41393	225673	1543548	13460713
30066732, 29917452	22630	43678	233131	1626582	12317551
29248340, 30838147	50859	69818	257013	1609925	12363793
39650041, 27590540	1395836	2638716	4613277	6967642	26370142

Полученные замеры времени для с увеличенным объёмом внедряемой информации позволяют подтвердить предположение о наилучшем соотношении времени выполнения к числу значащих позиций,

кроме того при увеличении сообщения с 13 символов до 55 можно заметить, что зависимость времени выполнения в ряде случаев нелинейная, что обуславливается взаимным расположением точек.

Исходя из полученных результатов можно сделать вывод что время вычисления необходимых точек для внедрения сообщения зависит от взаимного расположения центральной точки и ключевой точки. С ростом удаленности наблюдается увеличение необходимого времени для вычисления необходимых позиций. Кроме того, при выборе значащих позиций следует выбирать значения больше пяти что ускоряет работу алгоритма.

ЛИТЕРАТУРА

1. Урбанович, П.П. Защита информации методами криптографии, стеганографии и обфускации : учеб.-метод. пособие / П.П. Урбанович. – Минск : БГТУ, 2016. – 220 с.

2. Уласевич, Н.И. Методы встраивания информации с использованием тега ELLIPSE в SVG / Н.И. Уласевич, Н.А. Жилияк // Повышение качества жизни и обеспечение конкурентоспособности экономики на основе инновационных и научно-технических разработок: сборник статей VII Международной научно-технической конференции "Минские научные чтения – 2024", Минск, 3-5 декабря 2024 г.: в 3 томах. Том 1. – Минск : БГТУ, 2024. – С. 323–327.

УДК 004.043

Н.И. Гурин, доц.; Н.В. Ржеутская, ст. преп. (БГТУ, г. Минск)

ОСОБЕННОСТИ ПОДГОТОВКИ И ИСПОЛЬЗОВАНИЯ БАЗЫ ЗНАНИЙ В СИСТЕМЕ ТЕСТИРОВАНИЯ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

База знаний, используемая в системе тестирования на естественном языке [1], требует для своего успешного функционирования определенной настройки и предполагает выполнения следующих основных условий: соответствующей настройки своей структуры, подготовку содержания записей семантических триад «субъект»-«отношение»-«объект» и вспомогательных данных, а также разработку правил использования системы при проведении собственно самого тестирования.

Прежде всего, необходимо отметить, что при проверке правильности полученного от тестируемого ответа на заданный ему вопрос, основной сложностью является необходимость определения минимального количества символов в элементе «объект» семантической триады