АЛГОРИТМ ФОРМИРОВАНИЯ ИНФОРМАЦИОННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ-МЕТОК ДЛЯ ИСПОЛЬЗОВАНИЯ В АЛГОРИТМАХ ВСТРАИВАНИЯ ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ

Митекин В.А.

Институт систем обработки изображений РАН

Аннотация

В статье предложен новый алгоритм генерации последовательностей-меток, используемых при встраивании стойких цифровых водяных знаков (ЦВЗ), показано соответствие разработанного алгоритма основным требованиям к последовательностям данного типа, в том числе требованиям вычислительной сложности и устойчивости к коллизиям. Также в ходе вычислительного эксперимента установлено, что разработанный алгоритм генерации последовательностей-меток позволяет достигать значительно большей, чем у существующих алгоритмов, устойчивости к атаке методом прямого перебора.

Ключевые слова: цифровой водяной знак, ЦВЗ, оптический ортогональный код, коллизия, метод прямого перебора, код без запятых.

Введение

Задача генерации информационных последовательностей специального вида (последовательностей-меток) на основе задаваемого пользователем ключа (так называемого стеганографического ключа) стала актуальной в связи с широким распространением систем встраивания цифровых водяных знаков (ЦВЗ), применяемых для защиты авторских прав на цифровые мультимедийные данные. Генерируемые на основе стеганографического ключа последовательности-метки используются, в частности, в качестве одномерного или двумерного аддитивного ЦВЗ в схеме с «информированным» корреляционным детектором [1], а также в качестве устойчивых к искажениям «меток синхронизации» для первичного обнаружения ЦВЗ [2]. Кроме того, генерируемые на основе ключа последовательности используются для построения набора ортогональных базисных функций в схеме встраивания стойкого ЦВЗ [3] и для формирования подобластей суммирования в алгоритмах семейства «Patchwork» [4]. От выбора подобных информационных последовательностей-меток в значительной степени зависит как устойчивость встраиваемого ЦВЗ к искажениям, так и его способность противостоять преднамеренным атакам с целью удаления, замены или подделки ЦВЗ. Подобная область применения накладывает ряд специфичных требований к генерируемым последовательностям-меткам и к алгоритму их формирования на основе пользовательского ключа. К таким требованиям можно отнести (согласно [5, 6, 7, 8]):

1. Возможность установления однозначного соответствия (биекции) между любым допустимым значением ключа пользователя P и соответствующей этому ключу последовательностью-меткой C. Далее будем предполагать, что все возможные последовательности-метки, сгенерированные на основе допустимых ключей, образуют множество $\hat{\mathbf{C}}$.

2. Возможность генерации последовательности-метки на основе произвольно выбранного ключа P за фиксированное время, не зависящее от значения ключа, без необходимости генерировать или хранить все множество $\hat{\mathbf{C}}$ или его часть.

3. Возможность генерировать последовательность-метку C большой длины (до нескольких тысяч бит), что необходимо для обеспечения устойчивости ЦВЗ к аддитивному шуму при использовании корреляционного детектора ЦВЗ, а также для устойчивости ЦВЗ к ряду специфичных атак, направленных на приближенное вычисление последовательности C прямым перебором [9].

4. Фиксированное значение «минимального циклического расстояния» λ (минимальное расстояние Хэмминга с учётом циклического сдвига, [10, 11, 12]) между любыми двумя последовательностями-метками из множества $\hat{\mathbf{C}}$. Данное требование позволяет избежать так называемых коллизий при определении авторства ЦВЗ (коллизией в данном случае называется ситуация, когда детектор ЦВЗ вместо последовательности C', ранее встроенной в качестве ЦВЗ, обнаруживает последовательность C'', отличную от C'). Кроме того, данное требование позволяет обеспечивать устойчивость ЦВЗ к широкому классу атак, направленных на приближенное вычисление ЦВЗ [9]. Использование минимального циклического расстояния λ вместо расстояния Хэмминга в данном случае обусловлено требованием устойчивости ЦВЗ к кадрированию и повороту (при встраивании ЦВЗ в области преобразования Фурье-Меллина [2, 13, 14, 15, 16]).

5. Мощность пространства $|\hat{\mathbf{C}}|$, достаточная для защиты от вычисления последовательности-метки методом прямого перебора (так называемая brute-force attack). В данной статье будем предполагать, что мощность $|\hat{\mathbf{C}}|$, требуемая для защиты от атак методом прямого перебора, должна составлять не менее 2^{40} ключей.

6. Возможность генерировать последовательности с фиксированным весом (соотношением числа единиц и нолей в бинарной последовательности). Данное требование обусловлено необходимостью обеспечивать фиксированное соотношение сигнал/шум при встраивании ЦВЗ в спектральной области [1] независимо от выбранного пользователем ключа.

Требования 1 и 4 особенно критичны в случаях, когда ЦВЗ применяется для установления несанкционированного распространения мультимедийных данных и для подтверждения авторства. В данных случаях предполагается, что каждому автору (правообладателю) поставлен в соответствие уникальный ключ встраивания. На этапе встраивания ЦВЗ сформированная на основе ключа последовательность C используется для маркировки всех мультимедийных данных данного автора. Коллизия, т.е. фактически, неоднозначность в определении детектором последовательности C, а следовательно, и неоднозначность в определении авторства, является в данном случае недопустимой.

1. Обзор существующих алгоритмов генерации информационных последовательностей-меток

Как показано в [6, 8, 9], подавляющее большинство существующих алгоритмов встраивания ЦВЗ, предназначенных для подтверждения авторства и установления фактов несанкционированного распространения данных, при формировании последовательностей-меток не предусматривают защиту от коллизий и атак при формировании. Наиболее распространённым подходом в данном случае является использование в качестве метки произвольной бинарной последовательности заданной длины.

Некоторые из существующих алгоритмов (например, алгоритм [6]) предполагают использование M – последовательностей полного периода в качестве меток, что позволяет избежать коллизий при формировании последовательностей, но не обеспечивает достаточной мощности $\left|\hat{\mathbf{C}}\right|$ для защиты от атак, направленных на вычисление последовательности-метки.

Кроме того, в работе [3] предложен алгоритм ортогонализации набора псевдослучайных двумерных изображений-шаблонов, используемых в качестве меток при встраивании ЦВЗ. Данный метод также позволяет предотвратить коллизии при извлечении ЦВЗ, но в то же время требует генерации и хранения всего множества $\hat{\mathbf{C}}$, что накладывает значительные ограничения на мощность $\left|\hat{\mathbf{C}}\right|$ множества возможных последовательностей-меток.

Следует отметить, что рассмотренные выше требования 1 и 4 обуславливают задачу формирования последовательностей, схожую с задачей формирования так называемого «оптического ортогонального кода», используемого в системах связи с кодовым разделением доступа (CDMA) [17, 18]. Схожая задача также решается при формировании кодовых слов, различимых с учётом циклического сдвига ([10, 11, 12]). Данные классы последовательностей удовлетворяют требованиям 1 и 4, но могут не соответствовать остальным представленным требованиям.

Рассмотрим существующие алгоритмы формирования «оптических ортогональных кодов» и кодовых слов, различимых с учётом циклического сдвига, с точки зрения их соответствия приведённым выше требованиям. Существующие алгоритмы могут быть разбиты по принципу формирования последовательностей на следующие классы.

1. Комбинаторные методы генерации [17, 18] обладают высокой вычислительной эффективностью, но в то же время обеспечивают относительно малую мощность $\left|\hat{\mathbf{C}}\right|$ (см. табл. 1) и позволяют установить однозначное соответствие между стеганографическим ключом и формируемой последовательностью - меткой только путём формирования полного множества $\hat{\mathbf{C}}$, что противоречит требованиям 1 и 2.

2. Алгебраические методы, основанные на вычислениях в полях Галуа [17, 19, 20, 21, 22], позволяют эффективно генерировать кодовые последовательности большой длины и не требуют формирования и хранения всего множества $\hat{\mathbf{C}}$, но обеспечивают при этом мощность $\left|\hat{\mathbf{C}}\right|$, недостаточную для защиты от атак методом прямого перебора (см. табл. 1). Следовательно, данный класс алгоритмов не удовлетворяет требованиям 2 и 5.

3. Методы, основанные на «прореживании» MDS кодов (в большинстве случаев кодов Рида-Соломона или БЧХ-кодов, [23, 24]) обеспечивают достаточную мощность $\left|\hat{\mathbf{C}}\right|$, но позволяют установить однозначное соответствие между стеганографическим ключом и формируемой последовательностью - меткой только путём формирования полного множества $\hat{\mathbf{C}}$, что противоречит требованиям 1 и 2. Кроме того, что является особенно важным при генерации последовательностей большой длины, методы данного класса позволяют генерировать последовательности только длиной $2^m - 1$, где m – целое.

Далее в статье предложен новый метод формирования последовательностей-меток, удовлетворяющий приведённым выше требованиям 1 - 6 и обеспечивающий большую, чем известные методы, мощность $\left|\hat{\mathbf{C}}\right|$ множества возможных последовательностей-меток.

2. Алгоритм формирования последовательностей-меток на основе БЧХ-кодов

Предлагаемый алгоритм основан на совместном применении бинарных избыточных кодов с известным минимальным расстоянием Хэмминга между кодовыми словами (например, БЧХ-кодов) и так называемых кодов без запятых [25, 26]. Предлагаемый алгоритм формирования последовательности-метки на основе пользовательского ключа может быть условно разделён на 2 этапа.

На первом шаге генерации пользовательский ключ, представленный в виде бинарной последовательности фиксированной длины $P = p_1 p_2 p_3 \ldots p_n$, где p_i i-й бит последовательности, преобразуется

путём избыточного кодирования во временную бинарную последовательность $T_{pass} = t_1 t_2 t_3 ... t_{n+k}$, где t_i – i-й бит последовательности T_{pass}, а $t_{n+1} ... t_{n+k}$ – избыточные биты (всего k избыточных бит). Для избыточного кодирования должен быть использован алгоритм кодирования с известным минимальным расстоянием Хэмминга между кодовыми словами (БЧХ-код). Таким образом, всегда будут существовать 2^n различных пользовательских ключей P длины n и соответствующие им 2^n временных последовательностей T_{pass} длины $(n+k)$.

На втором шаге генерации последовательность-метка C длиной $(n+k)\cdot s$ бит формируется следующим образом.

Пусть $\hat{\mathbf{S}} = \left\{ S^{00}, S^{01}, S^{10}, S^{11} \right\}$ – код без запятых ([25, 26]), состоящий из четырёх кодовых слов с длиной кодового слова s и весом (числом единиц) w_S, а $M = m_1 m_2 m_3 ... m_{n+k}$ – M-последовательность полного периода или ее фрагмент, где m_i – i-й бит последовательности M, $1 \le i \le n+k$. Тогда последовательность-метка $C = c_1 c_2 c_3 ... c_{(n+k)\cdot s}$ длиной $(n+k)\cdot s$ бит (c_j - j-й бит последовательности C) формируется в соответствии с соотношением

$$\left\{ c_j c_{j+1} ... c_{j+s-1} \right\} = \begin{cases} S^{00} если\ t_i = 0\ и\ m_i = 0 \\ S^{01}, если\ t_i = 1\ и\ m_i = 0 \\ S^{10}, если\ t_i = 0\ и\ m_i = 1 \\ S^{11}, если\ t_i = 1\ и\ m_i = 1 \end{cases}, \quad (1)$$

$где\ j = i \cdot s, 1 \le i \le n+k.$

Таким образом, последовательность-метка C формируется поблочно, при этом в качестве очередного блока последовательности используется одно из четырёх кодовых слов кода $\hat{\mathbf{S}}$. При этом мощность множества $\hat{\mathbf{C}}$ всех последовательностей-меток длины $(n+k)\cdot s$ бит будет равна числу различных пользовательских ключей длины n, то есть $\left|\hat{\mathbf{C}}\right| = 2^n$. Следует учесть, что код $\hat{\mathbf{S}} = \left\{ S^{00}, S^{01}, S^{10}, S^{11} \right\}$ и последовательность $M = m_1 m_2 m_3 ... m_{n+k}$ одинаковы для всех возможных пользовательских ключей и, таким образом, могут быть вычислены заранее для снижения вычислительной сложности алгоритма формирования последовательности-метки.

Далее будет показано, что любые две последовательности-метки из 2^n возможных являются различимыми с учётом их возможного циклического сдвига и образуют «оптический ортогональный код».

Определим величину **минимального циклического расстояния** между двумя последовательностями C' и C'' как минимум расстояния Хэмминга между ними, т.е. $\lambda = \min_{\Delta h} \sum_{h=1}^{L} c'_h \oplus c''_{(h+\Delta h) \bmod L}$, где λ – минимальное циклическое расстояние, $\oplus$ – побитовая операция «исключающее ИЛИ», L – длина последовательностей C' и C'', а величина циклического сдвига Δh удовлетворяет условию $0 \le \Delta h < L$. Величина λ равна нулю только в том случае, если C' и C'' побитово совпадают при некотором значении Δh циклического сдвига C' относительно C''. Далее будет показано, что $\lambda > 0$ для любой пары последовательностей-меток из $\hat{\mathbf{C}}$, и, следовательно, множество $\hat{\mathbf{C}}$ является «оптическим ортогональным кодом».

Рассмотрим две произвольных последовательности-метки C' и C'' длиной $(n+k)\cdot s$, величина циклического сдвига Δh между которыми удовлетворяет условию $0 \le \Delta h < (n+k)\cdot s$. В этом случае все допустимые значения Δh могут быть разделены на три группы.

1. $\Delta h = 0$. Пусть в этом случае $T'_{pass} = t'_1 t'_2 t'_3 ... t'_{n+k}$ и $T''_{pass} = t''_1 t''_2 t''_3 ... t''_{n+k}$ – временные последовательности, использованные для формирования последовательностей-меток C' и C'' соответственно. Тогда для любых C' и C'' минимальное циклическое расстояние λ_0 между ними равно расстоянию Хэмминга l между T'_{pass} и T''_{pass}, умноженному на минимальное расстояние Хэмминга h_S между словами кода $\hat{\mathbf{S}}$, т.е. $\lambda_0 = l \cdot h_S$. Так как T'_{pass} и T''_{pass} были получены путём избыточного кодирования, то минимальное значение величины l для любых T'_{pass} и T''_{pass} известно, при этом $l \ge 1$. Аналогично, исходя из свойств кода без запятых, $h_S \ge 1$.

2. Δh делится на s без остатка. В данном случае минимальное циклическое расстояние между словами C' и C'' определяется свойствами используемой последовательности M и величиной h_S. Пусть $\Delta h = \Delta i \cdot s$ – величина циклического сдвига C' относительно C'', где Δi – положительное целое число, $1 < \Delta i \le n+k$. Тогда блоку $c'_j c'_{j+1} ... c'_{j+s-1}$ последовательности C' при вычислении минимального циклического расстояния будет соответствовать блок $c''_{j+\Delta h} c''_{j+\Delta h+1} ... c''_{j+\Delta h+s-1}$ последовательности C''. В соответствии с соотношением (1) блок $c''_{j+\Delta h} c''_{j+\Delta h+1} ... c''_{j+\Delta h+s-1}$, сформированный на основе значений m_i и t'_i, и блок $c'_j c'_{j+1} ... c'_{j+s-1}$, сформированный на основе значений $m_{i+\Delta i}$ и $t''_{i+\Delta i}$, побитово совпадут только в том случае, если $m_i = m_{i+\Delta i}$. Ис-

ходя из свойства циклического сдвига M-последовательности[1], условие $m_i = m_{i+\Delta i}$ будет выполнено для $\frac{(n+k)}{2}$ блоков вне зависимости от выбранных C' и C''. Таким образом, минимальное циклическое расстояние $\lambda_{кратн}$ в данном случае будет определяться соотношением $\lambda_{кратн} = h_S \frac{(n+k)}{2}$.

3. Δh не делится на s без остатка [26, 27]. В данном случае минимальное циклическое расстояние $\lambda_{нк}$ между C' и C'' определяется базовой характеристикой различимости h_c кода без запятых $\hat{\mathbf{S}}$. Как следует из определения кода без запятых, каждый блок $c'_j c'_{j+1} ... c'_{j+s-1}$ слова C', являющийся согласно (1) словом из $\hat{\mathbf{S}}$, будет отличаться от блока $c''_{j+\Delta h} c''_{j+\Delta h+1} ... c''_{j+\Delta h+s-1}$ как минимум в h_c битах. Следовательно, минимальное циклическое расстояние между C' и C'' в данном случае определяется как $\lambda_{нк} = h_c(n+k)$, причём $h_c \geq 1$.

Исходя из рассмотренных случаев, можно утверждать, что для произвольных C' и C'' и для любого целого Δh, удовлетворяющего неравенству $0 \leq \Delta h < s(n+k)$, минимальное циклическое расстояние C' и C'' определяется соотношением

$$\lambda = \min(\lambda_0, \lambda_{кратн}, \lambda_{нк}) =$$
$$= \min(l \cdot h_s, h_s \frac{(n+k)}{2}, h_c(n+k)).$$

Таким образом, соответствие предложенного алгоритма требованию фиксированного значения «минимального циклического расстояния» λ между любыми двумя последовательностями-метками.

Предлагаемый алгоритм генерации последовательностей-меток также обладает следующими преимуществами по сравнению с существующими алгоритмами.

Во первых, мощность пространства допустимых ключей ($\left|\hat{\mathbf{C}}\right| = 2^n$ при длине кодового слова $s \cdot (n+k)$ бит) значительно превышает мощность, обеспечиваемую существующими алгоритмами, что делает предлагаемые последовательности-метки более устойчивыми к атакам методом прямого перебора. Сравнение с некоторыми существующими алгоритмами по числу допустимых ключей при заданной длине последовательности-метки приводится в табл. 1.

Во-вторых, предложенный алгоритм позволяет формировать последовательность-метку на основе произвольного ключа без необходимости генерировать всё множество $\hat{\mathbf{C}}$ или его часть. При этом существует однозначное соответствие (биекция) между любым допустимым значением ключа пользователя P и соответствующим этому ключу кодовым словом из $\hat{\mathbf{C}}$.

В-третьих, предлагаемый алгоритм позволяет эффективно генерировать последовательности большой длины. Наиболее вычислительно сложная операция – БЧХ-кодирование – требует $O(n \cdot k)$ операций, при этом код $\hat{\mathbf{S}}$ и последовательность M одинаковы для всех кодовых слов и могут быть вычислены заранее.

В-четвёртых, алгоритм позволяет формировать последовательности-метки с фиксированным весом при том условии, что все слова кода $\hat{\mathbf{S}}$ также имеют фиксированный вес.

Таким образом, предложенный алгоритм удовлетворяет требованиям 1 - 6 и, следовательно, обладает по сравнению с существующими алгоритмами рядом преимуществ, важных с практической точки зрения при реализации систем встраивания стойких ЦВЗ.

В следующем разделе приведены результаты сравнительного исследования стойкости различных типов последовательностей-меток к атакам методом прямого перебора.

3. Экспериментальное исследование стойкости последовательностей-меток к атакам методом прямого перебора

Для сравнения разработанного алгоритма с существующими по числу последовательностей $\left|\hat{\mathbf{C}}\right|$ рассмотрим случай генерации последовательности-метки большой длины (например, $L \approx 2^{14}$ бит, что приблизительно соответствует встраиванию последовательности-метки в спектр изображения размером 4096×4096 пикселей). Далее для ряда существующих алгоритмов оценим число возможных последовательностей-меток в наборе $\hat{\mathbf{C}}$, а также величину минимального циклического расстояния между последовательностями каждого набора.

Для формирования последовательностей-меток разработанным алгоритмом был использован следующий код без запятых

$$S^{00} = \{000101000\}, S^{01} = \{001000100\},$$
$$S^{10} = \{010000010\}, S^{11} = \{100000001\},$$
где $h_S = 2, h_c = 1$.

Результаты сравнения различных алгоритмов формирования последовательностей-меток представлены в табл. 1.

Далее была экспериментально исследована стойкость последовательностей-меток к коллизиям и атакам методом прямого перебора на примере алгоритма встраивания стойких ЦВЗ в области ДПФ

[1] Свойство циклического сдвига *M*-последовательности предполагает, что результатом побитовой операции «исключающее ИЛИ» между любой *M*-последовательностью и её циклически смещённой копией также является *M*-последовательностью.

([1]). В данном случае бинарная последовательность-метка аддитивно встраивается в спектр изображения-контейнера. Далее детектор ЦВЗ вычисляет коэффициент взаимной корреляции между известной ему последовательностью-меткой и каждой строкой спектра изображения-контейнера. Решение о наличие ЦВЗ принимается путём сравнения коэффициента корреляции с фиксированным пороговым значением. Подобная схема обнаружения ЦВЗ широко известна и называется схемой с информированным детектором. Следует учитывать, что в подобной схеме пороговые значения детектора выбираются таким образом, чтобы обеспечить некоторую устойчивость детектора к малым искажениям ЦВЗ (возникающим, например, при сжатии с потерями изображения-контейнера) и тем самым снизить вероятность ошибки 1-го рода (пропуск изображения с ЦВЗ). В то же время такой способ повышения устойчивости детектора к малым искажениям неизбежно приведёт к появлению коллизий, т.е. ситуаций, когда детектор обнаружит наличие ЦВЗ не только при использовании исходной последовательности-метки, но и при использовании схожих с ней (по критерию минимального циклического расстояния) последовательностей. Как уже было сказано ранее, при использовании ЦВЗ для установления авторства подобная ситуация является недопустимой.

С другой стороны, стеганографическая стойкость систем встраивания ЦВЗ определяется устойчивостью к атакам методом прямого перебора. Атака прямым перебором для схемы информированного детектора заключается в том, что нарушитель, перебирая возможные последовательности-метки, должен найти хотя бы одну, при использовании которой детектор даст утвердительный ответ о наличии ЦВЗ. Предполагая, что атакующий перебирает все возможные ключи в случайном порядке, и учитывая возможность коллизий, сложность атаки перебором будет определяться как соотношение $N_{br} \approx \dfrac{\left|\hat{\mathbf{C}}\right|}{2 \cdot \bar{N}_{collision}}$, где $\left|\hat{\mathbf{C}}\right|$ – число возможных последовательностей-меток, $\bar{N}_{collision}$ – экспериментально вычисленное среднее число коллизий, т.е. среднее количество последовательностей-меток из $\hat{\mathbf{C}}$, при использовании которых вместо исходной последовательности C' детектор ЦВЗ даёт утвердительный ответ.

Таким образом, с точки зрения стойкости к атакам прямым перебором алгоритм генерации последовательностей меток должен обеспечивать **максимизацию величины $|\hat{\mathbf{C}}|$, гарантируя в то же время отсутствие коллизий при заданном пороге обнаружения ЦВЗ**.

На основе данного критерия было проведено экспериментальное сравнение различных алгоритмов формирования последовательностей-меток, в том числе алгоритма, предложенного в данной работе. Для проведения эксперимента был использован существующий корреляционный детектор ЦВЗ [1] с длиной последовательности-метки $L \approx 500$ бит. Для сравнения были выбраны алгоритмы, обеспечивающие наибольшее число возможных последовательностей при заданной длине (согласно табл. 1). В ходе экспериментального исследования последовательности-метки синтезировались выбранным алгоритмом на основе случайного пароля и встраивались в изображение-контейнер, используя алгоритм [1], после чего изображение-контейнер подвергалось сжатию с потерями (JPEG с показателем качества $Q = 80$). Для полученного сжатого изображения экспериментально оценивалось число коллизий $N_{collision}$ путём перебора до 10000 «ложных» последовательностей-меток из $\hat{\mathbf{C}}$, имеющих малое значение минимального циклического расстояния λ относительно уже встроенной в данное изображение последовательности. Для вычисления $\bar{N}_{collision}$ полученное число коллизий усреднялось по всем тестовым изображениям со встроенной последовательностью-меткой (200 цветных изображений размером 4096×4096 пикселей, перебор 10000 «ложных» последовательностей-меток производился независимо для каждого изображения).

Пороговые значения детектора в ходе экспериментального исследования были фиксированы таким образом, чтобы обеспечить заданную устойчивость детектора к сжатию изображения с потерями (вероятность ошибки первого рода $P_{10} < 0{,}001$ при обнаружении ЦВЗ в сжатом изображении).

Результаты приведённого экспериментального сравнения собраны в табл. 2. Из таблицы следует, что наибольшую сложность для перебора и одновременно отсутствие коллизий при указанных выше параметрах детектора обеспечивают последовательности-метки на основе кода БЧХ $(56, 29, 11)$. Существующие последовательности ([21], [22]), хотя и гарантируют отсутствие коллизий, но обеспечивают гораздо меньшую сложность атаки прямым перебором.

Исходя из полученных экспериментальных результатов, можно сделать следующий вывод. Использование известных алгоритмов формирования последовательностей-меток ([21], [22]) позволяет избежать коллизий за счёт максимизации величины λ, но не является предпочтительным с точки зрения представленного выше критерия стойкости. В то же время предложенный алгоритм генерации последовательностей-меток позволяет в широких пределах изменять число возможных последовательностей заданной длины и тем самым достигать высокой сложности к атаке перебором, одновременно сохраняя устойчивость к коллизиям при выбранных параметрах детектора (пороговых значениях обнаружения ЦВЗ). Таким образом, можно утверждать, что разработанный алгоритм по показателю устойчивости к атакам прямого перебора превосходит существующие алгоритмы генерации последовательностей-меток.

Заключение

В статье проанализированы существующие методы встраивания стойких ЦВЗ для установления авторства, сформулированы требования к последовательностям-меткам, используемым при встраивании стойких ЦВЗ. Показано, что существующие алгоритмы формирования таких последовательностей не удовлетворяют в полной мере приведённым требованиям. Предложен новый алгоритм генерации последовательностей, удовлетворяющий всем приведённым требованиям, в том числе требованиям вычислительной сложности и устойчивости к коллизиям. В результате проведённых экспериментальных исследований установлено, что разработанный алгоритм генерации последовательностей-меток также позволяет достигать значительно большей по сравнению с существующими алгоритмами устойчивости к атаке методом прямого перебора.

Благодарности

Работа выполнена при поддержке РФФИ (гранты 09-01-00511-а, 11-07-12059-офи-м-2011) и при частичной поддержке Программы фундаментальных исследований Президиума РАН «Фундаментальные проблемы информатики и информационных технологий», проект 2.12.

Таблица 1. Оценка числа последовательностей-меток заданной длины

Тип последовательности	**Допустимая длина последовательности-метки *L*, бит**	**Теоретическая оценка числа возможных последовательностей $\left\|\hat{\mathbf{C}}\right\|$ в зависимости от *L***	**Допустимая длина последовательности, ближайшая к 2^{14} бит**	**Число возможных последовательностей-меток $\left\|\hat{\mathbf{C}}\right\|$**	**Минимальное циклическое расстояние λ**
Bent[27] (согласно [21])	$L=2^n-1$	$2^{\log_2(L+1)}+1$	*16383*	2^7	*8128*
Udaya[28] (согласно [21])	$L=2^n-1$	$2^{\log_2(L+1)}+1$	*16383*	$2^{14}+1$	*8063*
Chang[22] (согласно [21])	$L=2^n-1$ (n-нечетное)	$2^{2\cdot\log_2(L+1)}$	*8191*	2^{26}	*3967*
Gong[20, 21]	$L=2^n-1$	$2^{\log_2(L+1)\cdot(\log_2(L+1)/2-1)}$	*16383*	2^{84}	*4096*
Разработанный алгоритм на основе кода БЧХ(1820, 555, 301)	$L=s(n+k)$ $s,n,k-$целые	$2^{(L/s-k)}$	*16380*	2^{555}	*602*
Разработанный алгоритм на основе кода БЧХ (1820, 258, 401)	$L=s(n+k)$ $s,n,k-$целые	$2^{(L/s-k)}$	*16380*	2^{258}	*802*
Разработанный алгоритм на основе кода БЧХ (1820, 27, 501)	$L=s(n+k)$ $s,n,k-$целые	$2^{(L/s-k)}$	*16380*	2^{27}	*902*

Таблица 2. Результаты экспериментального исследования устойчивости последовательностей-меток к атакам прямого перебора

Тип последовательности	**Число возможных последовательностей-меток $\left\|\hat{\mathbf{C}}\right\|$**	**Среднее число коллизий на одну последовательность, $\bar{N}_{collision}$**	**Сложность атаки методом прямого перебора, $N_{br}\approx\frac{\left\|\hat{C}\right\|}{\bar{N}_{collision}}$**
Разработанный алгоритм на основе кода БЧХ (56, 44, 5), $L=504$	2^{44}	$\approx 2^{14}$	$\approx 2^{30}$
Разработанный алгоритм на основе кода БЧХ (56, 38, 7), $L=504$	2^{38}	$\approx 2^{9}$	$\approx 2^{29}$
Разработанный алгоритм на основе кода БЧХ (56, 29, 11), $L=504$	2^{29}	Коллизий не обнаружено	2^{29}
Gong [20, 21] $L=511$	2^{12}	Коллизий не обнаружено	2^{12}
Chang [22] $L=511$	2^{18}	Коллизий не обнаружено	2^{18}

Литература (References)

1. **Glumov, N.I.** The New Blockwise Algorithm for Large-scale Images Robust Watermarking / N.I. Glumov, V.A. Mitekin // Proceedings of 20th International Conference on Pattern Recognition. – Istanbul 2010, August 23-26. – P. 1453-1456.
2. **Pereira, S.** Template based recovery of Fourier-based watermarks using log-polar and log-log map / S. Pereira, J.K. O'Ruanaidh, F. Deguillaume, G. Csurka, T. Pun // Proceedings of ICMCS. – 1999. – P. 870-874.
3. **Fridrich, J.** Robust Digital Watermarking Based on Key-Dependent Basis Functions // J. Fridrich, A. Baldoza, R.J. Simard // Information Hiding, Second International Workshop. – 1998. –Vol. 1525. – P. 143-157.
4. **Bender, W.** Techniques for Data Hiding / W. Bender, D. Gruhl, N. Morimoto, A. Lu // IBM Systems Journal. – 1996. – Vol. 35. – P. 313-336.
5. **Seo, J.S.** On the design of template in the autocorrelation domain / J.S. Seo , C.D. Yoo // Proc. SPIE Electronic imaging. – 2002. – P. 305-312.
6. **Das, T.S.** Spread Spectrum based M-ary Modulated Robust Image Watermarking / T.S. Das, V.H. Mankar, S.K. Sarkar // IJCSNS International Journal of Computer Science and Network Security. – 2007. – Vol. 7, N 10. – P. 154-160.
7. **Moreno, O.** New families of arrays in two dimensions for watermarking applications / O. Moreno, A.Z. Tirkel, U. Parampalli, R.G. van Schyndel // IEEE Electronics Letters. – 2010. – Vol. 46. – P. 1500-1502.
8. **Tirkel, A.Z.** A two-dimensional digital watermark / A.Z. Tirkel, R.G. van Schyndel, C.F. Osborne // Asian Conference on Computer Vision. – 1995. – P. 378-383.
9. **Doerr, G.J.** Danger of Low-Dimensional Watermarking Subspaces / G.J. Doerr, J.-L. Dugelay// ICASSP 2004, 29th IEEE International Conference on Acoustics, Speech, and Signal Processing. – 2004. – Vol. 3. – P. 93-96.
10. **Nguyen, Q.A.** Constructions of binary constant-weight cyclic codes and cyclically permutable codes / Q.A. Nguyen, L. Györfi, J.L. Massey // IEEE Transactions Information Theory. – 1992. – Vol. 38. – P. 940-949.
11. **Kumar, P.** New constructions of optimal cyclically permutable constant weight codes / O. Moreno, Z. Zhang, P. Kumar, V. Zinoviev // IEEE Transactions on Information Theory. – 1995. – Vol. 41, N 2. – P. 448-455.
12. **Bitan, S.** Constructions for optimal constant-weight cyclically permutable codes and difference families / S. Bitan, T. Etzion // IEEE Transactions on Information Theory. – 1995. – Vol. 41, N 10. – P. 77-87.
13. **Lin, C.Y.** Rotation, scale, and translation resilient watermarking for images / C.Y. Lin, M. Wu, J.A. Bloom, I.J. Cox, M.L. Miller, Y.M. Lui // IEEE Trans on Image Processing. – 2001. – Vol. 10(5). – P. 767-782.
14. **Kim, B.S.** Robust digital image watermarking method against geometrical attacks / B.S. Kim, J.G. Choi, C.H. Park, J.U. Won, D.M. Kwak, S.K. Oh, C.R. Koh, K.H. Park // Real-Time Imaging. – 2003. – Vol. 9. – P. 139-149.
15. **O'Ruanaidh, J.K.** Rotation, scale, and translation invariant digital image watermarking / J.K. O'Ruanaidh, T. Pun // Proceedings of ICIP. –1997. – Vol. 1. – P. 536-539.
16. **O'Ruanaidh, J.K.** Rotation, scale, and translation invariant spread spectrum digital image watermarking / J.K. O'Ruanaidh, T. Pun // Signal Processing. – 1998. – Vol. 66. – P. 303-317.
17. **Richardson, D.J.** Optical Code Division Multiple Access Communication Networks: Theory and Applications / H. Yin, D.J. Richardson. – Springer, 2009. – 382 p.
18. **Chung, W.** Optical orthogonal codes: Design, analysis, and applications / W. Chung, J. Salehi, V. Wei // IEEE Transactions on Information Theory. – 1989. – Vol. 35, N 3. – P. 595-604.
19. **Moschim, E.** Some optical orthogonal codes for asynchronous CDMA systems / A.D. Neto, E. Moschim // IEEE Global Telecommunications Conference. – 2002. – Vol. 3. – P. 2065-2068.
20. **Gong, G.** New designs for signal sets with low cross correlation, balance property, and large linear span: GF(p) case / G. Gong // IEEE Transactions Information Theory. – 2002. – Vol. 48, N 11. – P. 2847-2867.
21. **Gong, G.** A new binary sequence family with low correlation and large size / N.Y. Yu, G. Gong //IEEE Transactions Information Theory. – 2006. – Vol. 52. – P. 1624-1636.
22. **Chang, A.** On a conjectured ideal autocorrelation sequence and a related triple-error correcting cyclic code / A. Chang, P. Gaal, S.W. Golomb, G. Gong, T. Helleseth, P.V. Kumar // IEEE Trans. Inform. Theory. – 2000. – Vol. 46, N 2. – P. 680-687.
23. **Kuribayashi, M.** How to Generate Cyclically Permutable Codes From Cyclic Codes / M. Kuribayashi, H. Tanaka // IEEE Transactions on Information Theory. – 2006. – Vol. 51, N 10. – P. 4660-4663.
24. **Chatterjee, P.K.** Optical Orthogonal Codes Using Error Correcting Codes / M. Choudhary, P.K. Chatterjee, J. John. – 21 November 2007.
25. **Golomb, S.W.** Comma-free codes / S.W. Golomb, B. Gordon, L.R. Welch // Canadian J. Math. – 1958. – Vol. 10, N 2. – P. 202-209.
26. **Eastman, W.L.** On the construction of comma-free codes / W.L. Eastman // IEEE Transactions on Information Theory. – 1965. – Vol. 11, N 2. – P. 263-267.
27. **Olsen, J.D.** Bent-function sequences / J.D. Olsen, R.A. Scholtz and L.R. Welch // IEEE Transactions Information Theory. – 1982. – Vol. 28, N 6. – P. 858-864.
28. **Udaya, P.** Polyphase and frequency hopping sequences obtained from finite rings / P. Udaya // Kanpur: Dept. Elec. Eng., Indian Inst. Technol. – 1992. – Ph.D dissertation.

THE ALGORITHM FOR LARGE-SCALE IMAGES ROBUST WATERMARKING USING BLOCKWISE PROCESSING

V. A. Mitekin

Image Processing Systems Institute of the RAS

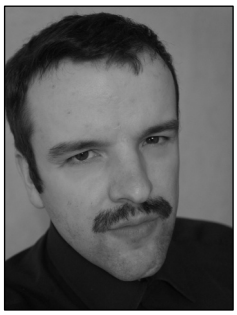
Abstract

A new algorithm for watermarking sequence generation is proposed, allowing collision-free robust watermarking using correlation-based watermark detector. The proposed algorithm provides considerably large number of codewords (password-based watermarking sequences) which makes it more robust to a brute-force watermark extraction attacks compared to known algorithms for watermarking sequence generation. Also, proposed algorithm fulfills the set of application-

specific requirements motivated by needs of watermarking scheme computational effectiveness, robustness etc.

Key words: digital watermarking, optical orthogonal codes, cyclically permutable codes, watermark collision.

Сведения об авторе



Митекин Виталий Анатольевич родился в 1983 году. В 2006 году окончил Самарский государственный аэрокосмический университет (СГАУ) по специальности «Прикладная математика и информатика». В 2009 году защитил диссертацию на соискание степени кандидата технических наук. В настоящее время работает научным сотрудником в Институте систем обработки изображений РАН. Круг научных интересов включает обработку изображений и распознавание образов, стеганографию и стегоанализ, криптографию. Имеет 21 публикацию, в том числе 7 статей.

E-mail: *mitekin@smr.ru* .

Vitaly Anatolyevich Mitekin (b. 1983) graduated from the S.P. Korolyov Samara State Aerospace University (SSAU), majoring in Applied Mathematics and Informatics in 2006. He received his Candidate in Technical Sciences degree from Samara State Aerospace University in 2009. Currently he works as the researcher at the Image Processing Systems Institute of the Russian Academy of Sciences. He has 21 publications, including 7 articles . His scientific interests include image processing and recognition, steganography and steganalysis, cryptography.

Поступила в редакцию 17 сентября 2011 г.

Дизайн: Я.Е. Тахтаров. Оформление и верстка: М.А. Вахе, С.В. Смагин и Я.Е. Тахтаров.

Подписано в печать 15.12.2011 г. Усл. печ. л. 12,47.
Отпечатано в типографии ООО «Предприятие «Новая техника».
Заказ № 11/4. Тираж 320 экз. Печать офсетная. Формат 62x84 1/8.