Секция 3

МАТЕМАТИЧЕСКИЕ МЕТОДЫ СТЕГАНОГРАФИИ

УДК 003.26

СТЕГАНОГРАФИЯ И СТЕГОАНАЛИЗ В ВИДЕОФАЙЛАХ

О.В. Моденова

В настоящее время в связи с ростом объемов информации и увеличением пропускной способности каналов связи все большую актуальность приобретает вопрос сокрытия информации в видеопоследовательностях. Передача цифрового видео в последние годы является типичным событием и не вызывает подозрения. Например, сервис YouTube насчитывает сотни миллионов видеофайлов, причем один и тот же видеоматериал встречается в разных форматах и разном качестве. Огромное количество видеофайлов размещается в P2P-сетях.

Были рассмотрены некоторые особенности использования форматов видеофайлов для сокрытия информации. Несмотря на то, что существует большое количество видеоформатов, на практике для сокрытия информации используются форматы MPEG-2 и MPEG-4. На русском языке практически отсутствуют публикации, посвященные встраиванию информации в другие видеоформаты [1–3].

В [2] описаны три способа внедрения информации в файлы формата MPEG-2: встраивание на уровне коэффициентов, на уровне битовой плоскости и за счет энергетической разницы между коэффициентами. Рассмотрим преимущества и недостатки этих способов.

Метод встраивания информации на уровне коэффициентов. Биты скрываемой информации встраиваются в коэффициенты дискретного косинусного преобразования (ДКП). Главной проблемой модификации коэффициентов ДКП в сжатом потоке видео является накопление сдвига или ошибок. Искажения, вызванные изменением коэффициентов ДКП, могут распространяться во временной и в пространственной областях. Поэтому для компенсации искажений добавляют специальный сигнал. В силу ограничения на битовую скорость, при внедрении изменяются только 10–20 % коэффициентов ДКП. При использовании данного метода скрываемая информация сохраняется при фильтровании, зашумлении (аддитивным шумом) и дискретизации.

Метод встраивания информации на уровне битовой плоскости. Этот метод отличается высокой пропускной способностью и небольшой вычислительной сложностью. Но есть и существенный недостаток: информация, встроенная таким образом, может быть легко удалена. При повторном наложении последовательности бит качество видео ухудшится незначительно, а скрываемая информация будет уничтожена.

Метод встраивания информации за счет энергетической разницы между коэффициентами. В основе этого метода лежит дифференциальное встраивание энергии (ДЭВ). Сложность алгоритма ДЭВ незначительно выше сложности метода встраивания на уровне битовой плоскости и значительно ниже сложности метода, основанного на корреляции с компенсацией ошибок предсказания. Метод ДЭВ может быть применен не только к видеоданным MPEG, но и к другим алгоритмам сжатия видео. Информация встраивается путем удаления нескольких коэффициентов ДКП, и

это имеет свои преимущества. Во-первых, в сжатый поток видеоданных не надо ничего добавлять, можно обойтись без повторного сжатия восстановленного потока видео. Во-вторых, удаление высокочастотных коэффициентов будет уменьшать размер стегообраза потока сжатых видеоданных по сравнению с исходным потоком. Алгоритм ДЭВ вносит в видео несколько меньше искажений, чем метод встраивания информации на уровне битовой плоскости. Для удаления скрытой информации требуется проведение более сложных вычислительных операций, чем встраивание новой произвольной битовой последовательности.

Если размер скрываемой информации небольшой по сравнению с объемом контейнера, то можно вносить изменения не в каждый кадр, а с некоторым интервалом. Это уменьшит накопление ошибок и может затруднить обнаружение факта сокрытия информации с помощью методов статистического стегоанализа.

Как правило, для сокрытия информации в видеопоследовательностях используются методы, использующие только видеопоток. Практически нет информации о сокрытии информации в аудиосигнале видеофайлов, хотя это позволило бы скрывать больший объем информации. В видеофайлах может быть упаковано различное количество звуковых дорожек, от 1 до 8 каналов для нескольких языков или нескольких вариантов переводов.

При переносе принципов сокрытия информации в неподвижных изображениях и в аудиофайлах нужно учитывать особенности, связанные со способами кодирования цифрового видео. Во время скрытия данных в видеопоследовательностях возникают трудности, так как одной из составляющих алгоритмов компрессии видеоинформации (в дополнение к компрессии неподвижного кадра) является кодирование векторов компенсации движения.

Для исправления возникающих ошибок при восстановлении информации из сжатого видео можно использовать помехоустойчивое кодирование. Например, использование сверточного кода с декодером Витерби обеспечивает достаточно высокую вероятность восстановления.

Применение вейвлет-преобразований и преобразований ДКП лучше подходит в случае необходимости защиты информации от активного злоумышленника, так как эти алгоритмы хорошо отделяют существенные детали от второстепенных.

Основной задачей стегоанализа является определение факта наличия скрытого сообщения в предполагаемом контейнере. Решается эта задача путем изучения статистических свойств сигнала. Например, распределение младших бит сигналов имеет, как правило, шумовой характер. Стегоаналитик проверяет соответствие реально наблюдаемой статистики ожидаемой. Обычно для этих целей используется критерий хи-квадрат.

Далее контейнер подвергается атакам, которые могут быть направлены на удаление или подмену скрываемой информации. Атаки применяются и в частотной, и в пространственной областях видеопоследовательностей. Основные типы атак на видеоконтейнер можно разделить на:

1) перекодирование видео с использованием алгоритмов сжатия с потерями;

2) изменение порядка кадров исходной видеопоследовательности (частный случай — удаление одного или нескольких кадров);

3) геометрические преобразования (всевозможные аффинные преобразования).

В целях анализа современного развития данного направления стеганографии был проведен поиск и исследование работы программ, реализующих сокрытие информации. В широком доступе была обнаружена лишь MSU StegoVideo, которая позволяет

встраивать в видеопоследовательность произвольный файл. При создании программы были проанализированы популярные кодеки и подобрано преобразование кадра, обеспечивающее наименьшие искажения и потери данных при сжатии видеофайла. Для исправления возникающих ошибок используется помехоустойчивое кодирование (сверточный код с декодером Витерби).

В работе рассматриваются основные методы встраивания информации в видеофайлы формата MPEG-2, проводится анализ, сравнение и обобщение этих методов для других видеоформатов. Рассматриваются особенности хранения аудиосигнала в видеофайлах и приводятся методы, использующие этот сигнал для сокрытия информации. Исследуются возможности компрометации реализованных алгоритмов с помощью методов статистического анализа.

ЛИТЕРАТУРА

1. *Аграновский А. В., Девянин П. Н., Хади Р. А., Черемушкин А. В.* Основы компьютерной стеганографии. М.: Радио и связь, 2003.
2. *Грибунин В. Г., Оков И. Н., Туринцев И. В.* Цифровая стеганография. М.: СОЛОН-Пресс, 2002.
3. *Зырянов А. В.* Методы защиты авторских прав с использованием цифровых водяных знаков в видеоконтейнерах формата MPEG // Вестник Томского госуниверситета. Приложение. 2007. № 23. С. 142–156.

УДК 621.391.037.372

О ПРАВИЛЕ ВЫБОРА ЭЛЕМЕНТОВ СТЕГАНОГРАФИЧЕСКОГО КОНТЕЙНЕРА В СКРЫВАЮЩЕМ ПРЕОБРАЗОВАНИИ

Е. В. Разинков, Р. Х. Латыпов

Стеганография — это наука о скрытой передаче информации, достигаемой встраиванием секретного сообщения в цифровой объект, называемый стеганографическим контейнером [1]. В качестве контейнеров обычно используются цифровые изображения, аудио- и видеофайлы. Результат встраивания — стего — передается по каналу связи, контролируемому нарушителем. Основная задача нарушителя состоит в определении наличия встроенной в перехваченный цифровой объект информации [2, 3].

В работе предлагается общий метод повышения стойкости и пропускной способности стеганографических систем.

На стойкость стеганографической системы критическое влияние оказывает правило выбора элементов стеганографического контейнера, модифицируемых в процессе встраивания информации. Под элементом контейнера будем понимать атомарную часть цифрового объекта, модифицируемую в процессе встраивания информации (яркости цветовых компонент пикселов, коэффициенты JPEG-преобразования, коэффициенты вейвлет-преобразования и т. д.).

Задача состоит в построении метода оптимального выбора элементов контейнера для встраивания информации — метода, позволяющего максимизировать либо стойкость стеганографической системы при заданном размере скрываемого сообщения, либо пропускную способность стегосистемы при заданной стойкости.

Различные элементы контейнера могут быть объединены в непересекающиеся группы таким образом, что элементы одной группы будут иметь схожие свойства и одинаковое распределение.

Рассматриваем контейнер как набор из m групп элементов. Каждая группа характеризуется количеством k_i содержащихся в ней элементов и их распределением. Обозначим через C_i область допустимых значений элементов контейнера, входящих в i-ю группу. Предполагается, что модификация одного элемента i-й группы позволяет встроить q_i бит, $q_i = \lfloor \log_2 |C_i| \rfloor$. Таким образом, рассматриваем цифровой объект (контейнер, стего) в виде набора векторов элементов контейнера $c_1^i c_2^i \dots c_{k_i}^i$, $c_j^i \in C_i$, $i = \overline{1,m}$.

Обозначим через x_i количество модифицируемых элементов i-й группы, $0 \leqslant x_i \leqslant k_i$, $\sum x_i q_i = n$.

Пусть $f_i(c)$ — функция плотности распределения элементов i-й группы неизмененного стеганографического контейнера. Скрываемая информация имеет высокую энтропию, так как часто бывает зашифрованной и/или сжатой. Это свойство скрытого сообщения позволяет найти функцию плотности распределения элементов i-й группы контейнера со встроенной информацией — $\bar{f}_i(c, x_i)$, где x_i — количество измененных элементов:

$$\bar{f}_i(c, x_i) = \frac{k_i - x_i}{k_i} f_i(c) + \frac{x_i}{k_i} \cdot \frac{1}{|C_i|}.$$

Обозначим через $P(S)$ вероятность того, что в качестве контейнера будет выбран цифровой объект S:

$$P(S) = \prod_{i=1}^{m} \prod_{j=1}^{k_i} f_i\left(c_j^i\right).$$

Вероятность $\bar{P}(S)$ того, что в результате встраивания информации будет получено стего S, вычисляется аналогично:

$$\bar{P}(S) = \prod_{i=1}^{m} \prod_{j=1}^{k_i} \bar{f}_i\left(c_j^i, x_i\right).$$

Изложенное выше позволяет оценить стойкость стеганографической системы с помощью информационно-теоретического подхода и относительной энтропии (расстояния Кулльбака — Ляйблера) [4]:

$$D(P||\bar{P}) = \sum_S P(S) \log_2 \frac{P(S)}{\bar{P}(S)}.$$

Чем меньше величина $D\left(P||\bar{P}\right)$, тем выше стойкость стегосистемы. Задача оптимального распределения скрываемого сообщения в стеганографическом контейнере сводится к нахождению такого вектора $\{x_i\}_i$, $0 \leqslant x_i \leqslant k_i$, $\sum x_i q_i = n$, при котором величина $D\left(P||\bar{P}\right)$ была бы минимальной. Эта задача может быть решена, если функции $f_i(c)$ известны.

Полученные в работе результаты позволяют значительно повысить пропускную способность стеганографической системы при фиксированной стойкости или повысить стойкость стегосистемы при заданной пропускной способности. Цель последующих исследований состоит в адаптации предложенной модели к распространенным форматам изображений, аудио- и видеофайлов, что позволит создавать более совершенные стеганографические системы.

ЛИТЕРАТУРА

1. *Simmons G. J.* The Prisoners' Problem and the Subliminal Channel // CRYPTO83 — Advances in Cryptology, August 22–24, 1984. P. 51–67.

2. *Wayner P.* Disappearing Cryptography, Second Edition — Information Hiding: Steganography and Watermarking. Elsevier, 2002. 413 p.
3. *Cox I., Miller M., Bloom J., et al.* Digital Watermarking and Steganography. Elsevier, 2008. 593 p.
4. *Cachin C.* An Information-Theoretic Model for Steganography // LNCS. 1998. V. 1525. P. 306–318.

УДК 681.511:3

СТЕГОСИСТЕМЫ ИДЕНТИФИКАЦИОННЫХ НОМЕРОВ, УСТОЙЧИВЫЕ К АТАКЕ СГОВОРОМ[1]

Т. М. Соловьёв, Р. И. Черняк

В связи с бурным развитием медиаиндустрии в настоящее время все более актуальной становится задача защиты интеллектуальной собственности от противоправных действий. Ежегодно медиапиратство наносит колоссальные убытки видео- и аудиоиндустриям. Основной статьей дохода кинокомпаний по-прежнему является прокат фильмов в кинотеатрах, в то время как современные сервисы IPTV, Internet TV и другие остаются в стороне. Такая ситуация во многом обуславливается высокими рисками утечки премьерного фильма и, как следствие, снижения интереса к нему у пользователей.

В настоящее время для защиты от копирования и несанкционированного использования медиаконтента широко применяется такой класс цифровых водяных знаков (ЦВЗ), как идентификационные номера (ИН).

ЦВЗ могут содержать некоторую информацию о собственнике материала или о месте и времени его производства.

В случае применение ИН в контейнер, предназначеный каждому пользователю, внедряется персональный номер, позволяющий контролировать дальнейший путь этого контейнера. Если пользователь окажется медиапиратом и начнет распространение своей копии, то идентификационный номер позволит быстро определить его.

Согласно терминологии, используемой в работе [1], множества ИН называются *стегосистемами идентификационных номеров*. При этом, помимо типичных атак для ЦВЗ, таких, как перекодирование, аффинные и другие преобразования, для стегосистем ИН существует очень опасная *атака сговором*.

Под атакой сговором понимается следующее. Злоумышленник побитно сравнивает имеющиеся у него копии некоторого медиаданного, содержащие различные ИН, и заключает, что биты, в которых сравниваемые данные различаются, суть биты ИН. Затем он устанавливает эти биты в некоторые значения так, чтобы полученный ИН, называемый *ложным*, не совпадал ни с одним из использованных при сравнении. При этом злоумышленник преследует одну из следующих целей: уничтожить ИН либо изменить его таким образом, чтобы он идентифицировал кого-то другого.

Данная работа является продолжением работы [2]. Предлагается решение для противостояния атаке сговором. Продолжается исследование структуры стегосистем идентификационных номеров, устойчивых к данной атаке. Определяется наиболее опасный случай атаки сговором — *мажорирующая атака*. Обсуждается проблема идентификации группы пиратов с помощью полученного ими ложного ИН.

[1]Работа выполнена в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 гг. (гос. контракт № П1010).

В работе [2] было показано, что для успешного противостояния атаке сговором необходимо использовать *допустимые* множества.

Определение 1. Множество $W_n \subset \{0,1\}^n$ (или W, если длина не существенна) называется допустимым, если каждому подмножеству $P \subseteq W_n$ взаимно-однозначно сопоставляется наименьший интервал $I(P)$, его покрывающий.

Далее интервал $I(P)$ будем называть *соответствующим* множеству P.

Замечание 1. В рамках решаемой задачи, достаточно ограничиться рассмотрением таких W_n, для которых $I(W_n) = \underbrace{- - \ldots -}_{n}$.

Некоторые свойства допустимых множеств

Свойство 1. Если W_n допустимое, то $|W_n| \leqslant n$.

Данное свойство налагает ограничение на количество пользователей системы. К примеру, для того чтобы обеспечить идентификационными номерами сеть из восьми миллионов пользователей, длина каждого из этих номеров должна быть не менее мегабайта. Такой объем дополнительного материала является существенным, и его дальнейшее увеличение может создавать проблемы на этапе внедрения.

Свойство 2. Все допустимые множества максимальной мощности ($|W_n| = n$) имеют вид $O_1(a)$, где $O_1(a) = \{x \in \{0,1\}^n : d_H(a,x) = 1\}$, $d_H(a,x)$ — расстояние по Хэммингу между векторами a и x.

Иными словами, любое допустимое множество максимальной мощности — это сфера радиуса один с некоторым вектором a в качестве центра.

Матричное представление стегосистем ИН

Рассмотрим допустимое множество W_n мощности k. Представим его в виде булевой матрицы $\|W\|_{k\times n}$, строками которой будут векторы из множества W_n. Матрицы, соответствующие допустимым множествам, будем называть допустимыми:

$$W_n = \begin{pmatrix} w_1[1] & w_1[2] & \ldots & w_1[n] \\ w_2[1] & w_2[2] & \ldots & w_2[n] \\ \vdots & \vdots & \ddots & \vdots \\ w_k[1] & w_k[2] & \ldots & w_k[n] \end{pmatrix}_{k\times n}.$$

Исходя из свойств 1 и 2, целесообразно рассматривать матрицы, для которых $k < n$.

Определим следующие операции над допустимыми матрицами:

1) перестановка столбцов и строк;
2) инверсия столбца;
3) удаление повторяющихся столбцов.

Утверждение 1. Применение операций 1–3 никак не влияет на свойство допустимости.

Определение 2. Матрицы A и A' назовем эквивалентными, если они могут быть получены друг из друга путем применения определенных выше операций.

Определение 3. Допустимые множества W и W' назовем эквивалентными, если соответствующие им допустимые матрицы эквивалентны.

Утверждение 2. В каждом классе эквивалентности существует матрица $W_n = (E_k A_{n-k})$, где

$$E_k = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}_{k \times k}, A = \|A\|_{k \times (n-k)}.$$

Определение 4. Множество W называется *сильно допустимым*, если удаление любого столбца в соответствующей ему матрице влечет потерю свойства допустимости.

Утверждение 3. Матрица, соответствующая любому сильно допустимому множеству, эквивалентна E_k.

Утверждение 4. Пусть W_n — допустимое множество. Тогда существует матрица $W'_n = (E_k A)$, эквивалентная матрице W_n, в котрой подматрица A не является допустимой.

Замечание 2. Из утверждений 2 – 4 следует, что свойство допустимости основывается на наличии подматрицы, эквивалентной единичной. Оставшаяся часть не влияет на допустимость и может быть выбрана произвольно.

Идентификация злоумышленников по ложному ИН

Рассмотрим сильно допустимое множество, заданное матрицей E_k. В каждом столбце этой матрицы все элементы, за исключением одного, равны нулю. Значит, единица в какой-либо компоненте ложного ИН может появиться в том и только в том случае, если соответствующий пользователь принимал участие в атаке сговором. В соответствии с этим, идентифицировать участников сговора по построенному ими ложному ИН возможно во всех случаях, кроме одного — когда ИН состоит из всех нулей. Наблюдая единицу в i-й компоненте ложного ИН, заключаем, что i-й пользователь участвовал в сговоре. При инвертировании i-го столбца в матрице E_k на злоумышленника укажет единственный ноль в i-м столбце. Покомпонентно просматривая ложный ИН, можно сделать вывод о степени вины каждого участника. Злоумышленниками окажутся пользователи с номерами i, такими, что $w'[i] = \overline{f}_{\text{maj}}(w_1[i], w_2[i], \dots, w_k[i])$, где w'— ложный ИН, а f_{maj} — мажоритарная функция.

Согласно утверждению 2, описанный способ идентификации злоумышленников может быть использован в любом допустимом множестве.

Мажорирующая атака

При использовании предложенного метода идентификации всегда существует ложный ИН, который не идентифицирует никого:

$$w[i] = f_{\text{maj}}(w_1[i], w_2[i], \dots, w_k[i]), i = 1, 2, \dots, k.$$

Далее будем обозначать его w_{maj}. Возникает закономерный вопрос: всегда ли злоумышленник может построить w_{maj} и существует ли стратегия, позволяющая строить именно его, а не какой-либо случайный вектор из интервала, соответствующего имеющимся у него ИН?

Утверждение 5. Если мощность множества пиратов P больше или равна 2, то w_{maj} принадлежит $I(P)$.

Утверждение 6. $w_{\mathrm{maj}}[i] = f_{\mathrm{maj}}(w_1[i], w_2[i], w_3[i]), i = 1, 2, \ldots, k$, где w_1, w_2, w_3 — любые попарно различные векторы из W.

Замечание 3. Из утверждений 5 и 6 следует, что если мощность множества пиратов больше или равна 3, то злоумышленник всегда может построить ИН, не идентифицирующий никого.

Описаный способ построения ложного ИН назовём мажорирующей атакой. Эта атака — частный случай атаки сговором, характеризующийся строго определенным выбором вектора из $I(P) \setminus P$. Согласно предложенному методу идентификации, построение вектора w_{maj} является единственным способом для группы злоумышленников избежать ответственности в полном объеме, т. е. ни один их них не будет вычислен. В связи с этим дальнейшая задача состоит в разработке узконаправленного метода идентификации, противостоящего мажорирующей атаке.

ЛИТЕРАТУРА

1. *Грибунин В. Г., Оков И. Н., Туринцев И. В.* Цифровая стеганография. М.: Солон-Пресс, 2002.
2. *Стружков Р. С., Соловьёв Т. М., Черняк Р. И.* Цифровые водяные знаки, устойчивые к атаке сговором // Прикладная дискретная математика. Приложение. 2009. № 1. С. 56–59.

УДК 519.651

О ВЫЯВЛЕНИИ ФАКТА ЗАШУМЛЕНИЯ КОНЕЧНОЙ ЦЕПИ МАРКОВА С НЕИЗВЕСТНОЙ МАТРИЦЕЙ ПЕРЕХОДНЫХ ВЕРОЯТНОСТЕЙ

А. М. Шойтов

Задача выявления факта наличия вкраплений в случайных последовательностях исследована в целом ряде работ (см., например, [1 – 4]. При этом обычно предполагаются известными тип и параметры распределения исходной последовательности. Так, в [4] рассмотрена последовательность, полученная по полиномиальной схеме с известными вероятностями исходов, и установлено, что гарантированно обнаружить факт наличия независимых вкраплений возможно только в случае, когда объем вкраплений растет по порядку быстрее корня от длины исходной последовательности. Аналогичный факт установлен в [3] для последовательности, образующей простую цепь Маркова с известной матрицей переходных вероятностей. В тезисах приводится обобщение результата [3] на случай простой цепи Маркова с неизвестной матрицей переходных вероятностей.

Пусть $X = \{X_1, ..., X_n, ...\}$ — простая конечная неразложимая и ацикличная цепь Маркова с N исходами, которые, не ограничивая общности, будем обозначать числами $1, ..., N$, и фиксированной матрицей переходных вероятностей $\Pi = \|\pi_{a,b}\|_{N\times N}$. Соответственно определены стационарные вероятности цепи X, которые обозначим через $(\pi_1, ..., \pi_N)$.

Будем предполагать, что на множестве $A = \{1, ..., N\}$ задана последовательность независимых случайных преобразований $\Phi = \{\varphi_1, ..., \varphi_n, ...\}$, полученных по схеме серий (n — номер серии) так, что для всех $i \neq j$, $i, j = 1, ..., n, ...$, преобразования φ_i и φ_j независимы и каждое из них определяется одной матрицей переходных вероятностей $P = \|p_{a,b}\|_{N\times N}$ по правилу $\mathbf{P}\{\varphi_i(a) = b\} = p_{a,b}$, $i = 1, ..., n, ...$ Определим случайную

последовательность $Z = \{Z_1, ..., Z_n, ...\}$ следующим образом: $Z_i = \varphi_i(X_i)$, $i = 1, ..., n, ...$, и будем писать $Z = \Phi(X)$.

Относительно наблюдаемой последовательности Y алфавита A выдвигаются две сложные гипотезы H_0: $Y = X$ и H_1: $Y = \Phi(X)$. Причем при обеих гипотезах, в отличие от [3], будем предполагать, что матрица $\Pi = \|\pi_{a,b}\|_{N\times N}$ неизвестна, а матрица $P = \|p_{a,b}\|_{N\times N}$ также неизвестна, но удовлетворяет ограничению $\lim\limits_{n\to\infty} p_{a,a} = 1$, $a \in A$.

Определим статистику

$$\chi_n^2 = \sum_{a,b,c\in A} \frac{\left(\nu_{abc} - \nu_{ab}\nu_{bc}/\nu_b\right)^2}{\nu_{ab}\nu_{bc}/\nu_b},$$

где $\nu_{abc} = \sum\limits_{i=1}^{n} \mathbf{I}\{Y_i = a, Y_{i+1} = b, Y_{i+2} = c\}$; $\nu_{ab} = \sum\limits_{i=1}^{n} \mathbf{I}\{Y_i = a, Y_{i+1} = b\}$; $\nu_a = \sum\limits_{i=1}^{n} \mathbf{I}\{Y_i = a\}$, $a, b, c \in A$. Статистики типа χ_n^2 применяются для различения гипотез о порядке цепи Маркова (см., например, [5]). Известно, что при гипотезе H_0 при $n \to \infty$ распределение χ_n^2 сходится к распределению хи-квадрат с N степенями свободы.

При сделанных предположениях справедлива следующая теорема.

Теорема 1. Если при $n \to \infty$ матрица переходных вероятностей P преобразований Φ меняется так, что для некоторых $a, b, c \in A$ выполнено условие

$$\sqrt{n} \sum_{\substack{y\in A \\ y\neq b}} p_{y,b}\left(\pi_{b,c} - \pi_{y,c}\right)\left(\pi_y \pi_{a,b} - \pi_b \pi_{a,y}\right) \to \infty,$$

то статистический критерий различения гипотез H_0 и H_1 на основе статистики χ_n^2 является состоятельным.

Следствие 1. Если при $n \to \infty$ для всех $a \neq b$, $a, b \in A$, справедливы оценки $p_{a,b} = f(n)\big(1 + o(1)\big)$ и $\sqrt{n} f(n) \to \infty$, стационарное распределение $(\pi_1, ..., \pi_N)$ цепи X является равномерным и в матрице $\Pi = \|\pi_{a,b}\|_{N\times N}$ есть хотя бы два различных элемента, то критерий различения гипотез H_0 и H_1 на основе статистики χ_n^2 является состоятельным.

Замечание. Статистика χ_n^2 применима для различения гипотез H_0 и H_1 только в случае, когда известно, что последовательность X образует *простую* цепь Маркова.

ЛИТЕРАТУРА

1. *Иванов В. А.* Модели вкраплений в однородные случайные последовательности // Труды по дискретной математике. 2008. Т. 10. С. 18–34.
2. *Пономарев К. И.* Параметрическая модель вкрапления и ее статистический анализ // Дискретная математика. 2009. Т. 21. № 4. С. 148–157.
3. *Filler T., Ker A. D., Fridrich J.* The square root law of steganographic capacity for Markov covers // Proc. SPIE. 2009. V. 7254, 725408. P. 31–47.
4. *Ker A. D.* A capacity result for batch steganography // IEEE Signal Processing Letters. 2007. V. 14(8). P. 525–528.
5. *Ивченко Г. И., Глибоченко А. Ф., Иванов В. А., Медведев Ю. И.* Статистический анализ дискретных случайных последовательностей. М.: МИЭМ, 1984. 92 с.