Уменьшение ошибок извлечения встроенной информации в стеганографической системе защиты информации со слепым декодером с минимальным изменением пикселей изображения и его максимальном заполнении

**Человек:** Предметом исследования является стеганографическая система защиты информации (ССЗИ) на основе метода прямого расширения спектра с минимально возможным изменением пикселей изображения и его максимальном заполнении. Особое внимание в статье уделяется проблеме создания ССЗИ со слепым декодером, поскольку оригинальное изображение сильно влияет на возможность правильного извлечения встроенной информации. Авторами подробно рассматриваются такие аспекты темы как применение процесса авторегрессии и корректирующих кодов для уменьшения количества ошибок извлечения встроенной информации. Методологической основой исследования являются методы теории помехоустойчивого кодирования, методы обработки цифровых изображений, методы математической статистики. Новизна исследования заключается в том, что определен оптимальный размер блока (32х32 пикселя), при котором двумерный процесс авторегрессии (АР) позволяет получить максимальный прирост в количестве правильно извлеченных встроенных бит. Кроме того исследована эффективность применения БЧХ кода, кода Рида-Соломона и LDPC кода для восстановления встроенных данных. Показано, что LDPC код с алгоритмом "мягкого" принятия решения Belief propagation является наиболее эффективным для уменьшения количества ошибок извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.

**Key words:** защита информации, стеганография, прямое расширение спектра, цифровые изображения, скрытая передача информации, помехоустойчивые коды, процесс авторегресии, БЧХ код, код Рида-Соломона, LDPC код

=================================

**FastText\_KMeans\_Clean:** В модели со слепым декодером (компонент ССЗИ, который извлекает встроенную информацию) на приемной стороне оригинальное изображение неизвестно, т.е. декодер слепой, поэтому считается, что полученное сообщение зашумлено двумя видами помех – оригинальное изображение и преднамеренная или непреднамеренная помеха, возникшая в результате передачи стегоконтейнера. Для получения графика эффективности применения двумерного АР процесса в зависимости от размера блока изображения найдем отношение разницы количества бит извлеченных правильно с применением и без применения двумерного АР процесса (область 1 на рис. 1) к количеству бит извлеченных правильно без применения двумерного АР процесса (область 2 на рис. 1). При α=0.05 вероятность стирания p =0.33, поэтому пропускная способность данного двоичного стирающего канала C =0.67. Если же на рис. 3 рассматривать вероятность ложного обнаружения α, при котором достигается минимум коэффициента битовых ошибок, т.е. при α=0.5, то при данной вероятности ложного обнаружения α ошибка извлечения обусловлена только ошибкой неправильного извлечения бит. Определена величина скорости передачи кода R =0.5 с помощью теоремы Шеннона и графика коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса.

**Key words part:** 0.6

=================================

**FastText\_KMeans\_Raw/:** Единственное различие между отправленным и полученным сообщениями обусловлено преднамеренной или непреднамеренной помехой, возникшей в результате передачи стегоконтейнера. В модели со слепым декодером оригинальное изображение сильно влияет на возможность правильного извлечения встроенной информации, поэтому при построении ССЗИ с моделью со слепым декодером необходимо применять средства, решающие данную проблему. Для получения графика эффективности применения двумерного АР процесса в зависимости от размера блока изображения найдем отношение разницы количества бит извлеченных правильно с применением и без применения двумерного АР процесса (область 1 на рис. 1) к количеству бит извлеченных правильно без применения двумерного АР процесса (область 2 на рис. 1). При маленьком размере блока длина ПСП также маленькая и его авто– и взаимокорреляционные характеристики не похожи на аналогичные характеристики белого шума, поэтому ПСП не попадает в остаток АР процесса. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса. Для LDPC кода возьмем следующие параметры: длина кодового слова n =255, количество информационных символов k =127. Если сравнить результаты, полученные с помощью LDPC кода с алгоритмом Belief propagation, с результатами, полученными с помощью БЧХ кода и с помощью кода Рида-Соломона, то видно, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.

**Key words part:** 0.8

=================================

**FastText\_PageRank\_Clean/:** Существуют три основные модели ССЗИ: со слепым декодером; с информированным декодером, с дополнительной информацией. Единственной помехой остаются преднамеренные или непреднамеренные шумы, возникающие в результате передачи стегоконтейнера. Коэффициент битовых ошибок можно считать вероятностью возникновения ошибки. Пропускная способность C двоичного симметричного канала определяется по следующей формуле [5]: C = 1+ p log2 p + (1- p )log2 (1- p ), где p – вероятность ошибки. Коды Рида–Соломона являются недвоичными БЧХ кодами. Это связано с тем, что код Рида-Соломона может исправлять локализованные пакеты ошибок. Здесь ошибки распределены по всей последовательности извлеченных бит. Декодирование LDPC кода в бинарном стирающем канале является простым.

**Key words part:** 0.6333333333333333

=================================

**FastText\_PageRank\_Raw/:** Существуют три основные модели ССЗИ: со слепым декодером; с информированным декодером, с дополнительной информацией. Единственной помехой остаются преднамеренные или непреднамеренные шумы, возникающие в результате передачи стегоконтейнера. Коэффициент битовых ошибок можно считать вероятностью возникновения ошибки. Пропускная способность C двоичного симметричного канала определяется по следующей формуле [5]: C = 1+ p log2 p + (1- p )log2 (1- p ), где p – вероятность ошибки. Коды Рида–Соломона являются недвоичными БЧХ кодами. Это связано с тем, что код Рида-Соломона может исправлять локализованные пакеты ошибок. Здесь ошибки распределены по всей последовательности извлеченных бит. Декодирование LDPC кода в бинарном стирающем канале является простым.

**Key words part:** 0.6333333333333333

=================================

**Mixed\_ML\_TR/:** В данной работе исследуется эффективность применения двумерного процесса авторегрессии (АР) в зависимости от размера блока изображения и корректирующих кодов для правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. На рис. 1 представлен график зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения с применением и без применении двумерного АР процесса. Для получения графика эффективности применения двумерного АР процесса в зависимости от размера блока изображения найдем отношение разницы количества бит извлеченных правильно с применением и без применения двумерного АР процесса (область 1 на рис. 1) к количеству бит извлеченных правильно без применения двумерного АР процесса (область 2 на рис. 1). Из рис. 2 видно, что применение двумерного АР процесса наиболее эффективно при размере блока 32х32 пикселя. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. Предложено представлять изображение последовательно в виде двоичного стирающего канала передачи информации и двоичного симметричного канала передачи информации, а стертые и неправильно принятые биты данных восстанавливать с помощью корректирующих кодов.

**Key words part:** 0.7

=================================

**MultiLingual\_KMeans/:** В данной работе исследуется эффективность применения двумерного процесса авторегрессии (АР) в зависимости от размера блока изображения и корректирующих кодов для правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. Для получения графика эффективности применения двумерного АР процесса в зависимости от размера блока изображения найдем отношение разницы количества бит извлеченных правильно с применением и без применения двумерного АР процесса (область 1 на рис. 1) к количеству бит извлеченных правильно без применения двумерного АР процесса (область 2 на рис. 1). Из рис. 2 видно, что применение двумерного АР процесса наиболее эффективно при размере блока 32х32 пикселя. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. Предложено представлять изображение последовательно в виде двоичного стирающего канала передачи информации и двоичного симметричного канала передачи информации, а стертые и неправильно принятые биты данных восстанавливать с помощью корректирующих кодов.

**Key words part:** 0.7

=================================

**Multilingual\_PageRank/:** В модели с дополнительной информацией перед встраиванием данных используется информация о самом контейнере таким образом, что он больше не является помехой для передачи встроенных данных. Результаты получены при встраивании информации во все непересекающиеся блоки разного размера при вероятности ложной тревоги (ложного обнаружения) α=0.05. Выделенные точки представляют усредненные значения, полученные из стандартных тестовых изображений Airplane, Lena, Peppers и Sailboat on lake. При маленьком размере блока длина ПСП также маленькая и его авто– и взаимокорреляционные характеристики не похожи на аналогичные характеристики белого шума, поэтому ПСП не попадает в остаток АР процесса. Поэтому выбираем наименее возможную длину кода n =7, для этой длины кода единственная стандартная длина информационного слова k =4. Коды Рида–Соломона являются недвоичными БЧХ кодами. Из рис. 5 видно, что результаты, полученные с применением кода Рида-Соломона аналогичны результатам, полученным с применением БЧХ кода. Для LDPC кода возьмем следующие параметры: длина кодового слова n =255, количество информационных символов k =127. Из рис. 6 и 7 видно, что применение алгоритма Belief propagation является более эффективным, что обусловлено "мягким" принятием решения.

**Key words part:** 0.7

=================================

**RuBERT\_KMeans\_Without\_ST/:** В данном методе встраивание информации происходит с минимально возможными параметрами в результате аддитивного сложения непересекающихся блоков изображения определенного размера с псевдослучайной последовательностью (ПСП), которая связана со встраиваемыми битами. Единственной помехой остаются преднамеренные или непреднамеренные шумы, возникающие в результате передачи стегоконтейнера. В данной работе исследуется эффективность применения двумерного процесса авторегрессии (АР) в зависимости от размера блока изображения и корректирующих кодов для правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. На рис. 1 представлен график зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения с применением и без применении двумерного АР процесса. При увеличении размера блока изображения длина ПСП также увеличивается и его авто– и взаимокорреляционные характеристики все больше начинают быть похожими на аналогичные характеристики белого шума и поэтому все больше ПСП попадает в остаток АР процесса. Это связано с тем, что ошибки неравномерно распределены по всей последовательности извлеченных бит, так как в одном блоке кодового слова длиной n =7 бит может быть исправлено не более одного неправильно излеченного бита или не более двух стираний. Коды Рида–Соломона являются недвоичными БЧХ кодами. Определена величина скорости передачи кода R =0.5 с помощью теоремы Шеннона и графика коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса.

**Key words part:** 0.7333333333333333

=================================

**RuBERT\_KMeans\_With\_ST/:** Из рис. 1 видно, что падение эффективности применения АР процесса при большом размере блока связано с ограничением сверху области 1 (прирост правильно извлеченных бит) общим количеством встроенных бит (извлекается правильно вся встроенная информация), а снизу – возрастающим количеством правильно извлеченных бит полученных без применения АР процесса. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α для исследования применения БЧХ кода. Восстанавливаем все стертые элементы кодового слова, так чтобы это удовлетворяло проверочному уравнению [8]: sHT =0 (mod 2), где H – матрица проверки на четность LDPC кода, s – кодовое слово.

**Key words part:** 0.6

=================================

**RUBERT\_page\_rank\_Without\_ST/:** В модели со слепым декодером оригинальное изображение сильно влияет на возможность правильного извлечения встроенной информации, поэтому при построении ССЗИ с моделью со слепым декодером необходимо применять средства, решающие данную проблему. Чтобы попасть в остаток АР процесса ПСП должны иметь авто- и взаимокорреляционные характеристики похожие на аналогичные характеристики белого шума. Ошибка извлечения встроенной информации обусловлена не обнаружением встроенной ПСП (соответствующие биты считаются стертыми) и (или) неправильным обнаружением встроенной ПСП (соответствующие биты считаются ложно обнаруженными). Коэффициент битовых ошибок можно считать вероятностью возникновения ошибки. Коды Рида–Соломона являются недвоичными БЧХ кодами.

**Key words part:** 0.6

=================================

**RUBERT\_page\_rank\_With\_ST/:** Существуют три основные модели ССЗИ: со слепым декодером; с информированным декодером, с дополнительной информацией. Коэффициент битовых ошибок можно считать вероятностью возникновения ошибки. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса. Коды Рида–Соломона являются недвоичными БЧХ кодами. Исследована эффективность применения для восстановления встроенных данных с помощью БЧХ кода, кода Рида-Соломона и LDPC кода.

**Key words part:** 0.6333333333333333

=================================

**RUSBERT\_KMeans\_Without\_ST/:** На рис. 2 представлен график эффективности применения двумерного АР процесса в зависимости от размера блока изображения. Поэтому в [3] предложено использовать пересечение блоков для увеличения объема передаваемой информации кроме того там же определено, что пересечение блоков не влияет на количество правильно извлеченных бит. При α=0.05 вероятность стирания p =0.33, поэтому пропускная способность данного двоичного стирающего канала C =0.67. Если же на рис. 3 рассматривать вероятность ложного обнаружения α, при котором достигается минимум коэффициента битовых ошибок, т.е. при α=0.5, то при данной вероятности ложного обнаружения α ошибка извлечения обусловлена только ошибкой неправильного извлечения бит. При α=0.5 вероятность ошибки p =0.03, поэтому пропускная способность данного двоичного стирающего канала C =0.8. Извлечение встроенной информации осуществляется при маленькой вероятности ложного обнаружения (α=0.05), поскольку это позволяет определить наличие или отсутствие встраивания, поэтому скорость передачи кода должна удовлетворять условию: R <0.67. Из данного условия выбираем R =0.5. Особенностью БЧХ кода является снижение корректирующей способности при увеличении длины кодового слова [6].

**Key words part:** 0.6333333333333333

=================================

**RUSBERT\_KMeans\_With\_ST/:** В модели со слепым декодером (компонент ССЗИ, который извлекает встроенную информацию) на приемной стороне оригинальное изображение неизвестно, т.е. декодер слепой, поэтому считается, что полученное сообщение зашумлено двумя видами помех – оригинальное изображение и преднамеренная или непреднамеренная помеха, возникшая в результате передачи стегоконтейнера. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. При этом в кодовом слове количество возможных исправлений неправильно излеченных бит t =1 или количество возможных исправлений стираний l =2. Если сравнить результаты, полученные с помощью LDPC кода с алгоритмом Belief propagation, с результатами, полученными с помощью БЧХ кода и с помощью кода Рида-Соломона, то видно, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.

**Key words part:** 0.8

=================================

**RUSBERT\_page\_rank\_Without\_ST/:** График эффективности применения двумерного АР процесса в зависимости от размера блока изображения. При маленьком размере блока длина ПСП также маленькая и его авто– и взаимокорреляционные характеристики не похожи на аналогичные характеристики белого шума, поэтому ПСП не попадает в остаток АР процесса. Это объясняет невысокую эффективность применения АР процесса при маленьком размере блока изображения. Ошибка извлечения встроенной информации обусловлена не обнаружением встроенной ПСП (соответствующие биты считаются стертыми) и (или) неправильным обнаружением встроенной ПСП (соответствующие биты считаются ложно обнаруженными). Коды Рида–Соломона являются недвоичными БЧХ кодами.

**Key words part:** 0.6

=================================

**RUSBERT\_page\_rank\_With\_ST/:** График эффективности применения двумерного АР процесса в зависимости от размера блока изображения. Размер блока, при котором применение двумерного АР процесса наиболее эффективно составляет 32х32 пикселя. Коды Рида–Соломона являются недвоичными БЧХ кодами. Из рис. 6 и 7 видно, что применение алгоритма Belief propagation является более эффективным, что обусловлено "мягким" принятием решения. Объяснена небольшая эффективность применения двумерного АР процесса при маленьком и большом размере блока изображения.

**Key words part:** 0.5333333333333333

=================================

**Simple\_PageRank/:** В [1,2] определено, что лучше использовать двумерный, а не одномерный АР процесс для извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. При α=0.05 вероятность стирания p =0.33, поэтому пропускная способность данного двоичного стирающего канала C =0.67. Если же на рис. 3 рассматривать вероятность ложного обнаружения α, при котором достигается минимум коэффициента битовых ошибок, т.е. при α=0.5, то при данной вероятности ложного обнаружения α ошибка извлечения обусловлена только ошибкой неправильного извлечения бит. На рис. 4 представлено три графика: первый график (обозначен цифрой 1) представляет собой зависимость коэффициента битовых ошибок BER от вероятности ложного обнаружения α при отсутствии использования двумерного АР процесса и корректирующих кодов для извлечения встроенных бит; второй график (обозначен цифрой 2) с использованием только двумерного процесса авторегрессии для извлечения встроенных бит; третий график (обозначен цифрой 3) с использованием двумерного АР процесса и БЧХ кода для извлечения встроенных бит. Если сравнить результаты, полученные с помощью LDPC кода с алгоритмом Belief propagation, с результатами, полученными с помощью БЧХ кода и с помощью кода Рида-Соломона, то видно, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. Определена величина скорости передачи кода R =0.5 с помощью теоремы Шеннона и графика коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса.

**Key words part:** 0.7666666666666667

=================================

**TextRank/:** В данной работе исследуется эффективность применения двумерного процесса авторегрессии (АР) в зависимости от размера блока изображения и корректирующих кодов для правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. На рис. 1 представлен график зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения с применением и без применении двумерного АР процесса. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. При α=0.5 вероятность ошибки p =0.03, поэтому пропускная способность данного двоичного стирающего канала C =0.8. Извлечение встроенной информации осуществляется при маленькой вероятности ложного обнаружения (α=0.05), поскольку это позволяет определить наличие или отсутствие встраивания, поэтому скорость передачи кода должна удовлетворять условию: R <0.67. Из данного условия выбираем R =0.5. Особенностью БЧХ кода является снижение корректирующей способности при увеличении длины кодового слова [6]. На рис. 4 представлено три графика: первый график (обозначен цифрой 1) представляет собой зависимость коэффициента битовых ошибок BER от вероятности ложного обнаружения α при отсутствии использования двумерного АР процесса и корректирующих кодов для извлечения встроенных бит; второй график (обозначен цифрой 2) с использованием только двумерного процесса авторегрессии для извлечения встроенных бит; третий график (обозначен цифрой 3) с использованием двумерного АР процесса и БЧХ кода для извлечения встроенных бит. В работе определен размер блока изображения, при котором наиболее эффективно применять двумерный АР процесс для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра, это блок размером 32х32 пикселя.

**Key words part:** 0.7333333333333333

=================================

**TF-IDF\_KMeans/:** В данном методе встраивание информации происходит с минимально возможными параметрами в результате аддитивного сложения непересекающихся блоков изображения определенного размера с псевдослучайной последовательностью (ПСП), которая связана со встраиваемыми битами. Недостатком модели с информированным декодером является необходимость передачи оригинального изображения приемной стороне, что усложняет процесс организации скрытой передачи информации, поэтому модель со слепым декодером выгоднее модели с информированным декодером. На рис. 1 представлен график зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения с применением и без применении двумерного АР процесса. При маленьком размере блока длина ПСП также маленькая и его авто– и взаимокорреляционные характеристики не похожи на аналогичные характеристики белого шума, поэтому ПСП не попадает в остаток АР процесса. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса. Для LDPC кода возьмем следующие параметры: длина кодового слова n =255, количество информационных символов k =127. Если сравнить результаты, полученные с помощью LDPC кода с алгоритмом Belief propagation, с результатами, полученными с помощью БЧХ кода и с помощью кода Рида-Соломона, то видно, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. Предложено представлять изображение последовательно в виде двоичного стирающего канала передачи информации и двоичного симметричного канала передачи информации, а стертые и неправильно принятые биты данных восстанавливать с помощью корректирующих кодов.

**Key words part:** 0.8333333333333334

=================================

**Текст:** В качестве базового метода для построения ССЗИ в работе используется метод на основе прямого расширения спектра с минимально возможным изменением пикселей изображения при его максимальном заполнении информацией. В данном методе встраивание информации происходит с минимально возможными параметрами в результате аддитивного сложения непересекающихся блоков изображения определенного размера с псевдослучайной последовательностью (ПСП), которая связана со встраиваемыми битами.. Существуют три основные модели ССЗИ: со слепым декодером; с информированным декодером, с дополнительной информацией.. В модели со слепым декодером (компонент ССЗИ, который извлекает встроенную информацию) на приемной стороне оригинальное изображение неизвестно, т.е. декодер слепой, поэтому считается, что полученное сообщение зашумлено двумя видами помех – оригинальное изображение и преднамеренная или непреднамеренная помеха, возникшая в результате передачи стегоконтейнера.. В модели с информированным декодером оригинальное изображение известно на приемной стороне и вычитается из стегоконтейнера. Единственное различие между отправленным и полученным сообщениями обусловлено преднамеренной или непреднамеренной помехой, возникшей в результате передачи стегоконтейнера.. В модели с дополнительной информацией перед встраиванием данных используется информация о самом контейнере таким образом, что он больше не является помехой для передачи встроенных данных. Единственной помехой остаются преднамеренные или непреднамеренные шумы, возникающие в результате передачи стегоконтейнера.. Поскольку в качестве базового используется метод на основе прямого расширения спектра, то возможно построение ССЗИ только с моделью со слепым или с информированным декодером. Недостатком модели с информированным декодером является необходимость передачи оригинального изображения приемной стороне, что усложняет процесс организации скрытой передачи информации, поэтому модель со слепым декодером выгоднее модели с информированным декодером. В модели со слепым декодером оригинальное изображение сильно влияет на возможность правильного извлечения встроенной информации, поэтому при построении ССЗИ с моделью со слепым декодером необходимо применять средства, решающие данную проблему.. В данной работе исследуется эффективность применения двумерного процесса авторегрессии (АР) в зависимости от размера блока изображения и корректирующих кодов для правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.. В [1,2] определено, что лучше использовать двумерный, а не одномерный АР процесс для извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра. Определим размер блока изображения, при котором применение двумерного АР процесса наиболее эффективно, т.е. дает максимальный прирост в количестве правильно извлеченной встроенной информации.. На рис. 1 представлен график зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения с применением и без применении двумерного АР процесса. Результаты получены при встраивании информации во все непересекающиеся блоки разного размера при вероятности ложной тревоги (ложного обнаружения) α=0.05. Выделенные точки представляют усредненные значения, полученные из стандартных тестовых изображений Airplane, Lena, Peppers и Sailboat on lake.. Рис. 1. График зависимости правильно извлеченных бит в процентах от общего количества встроенной информации от размера блока изображения. Для получения графика эффективности применения двумерного АР процесса в зависимости от размера блока изображения найдем отношение разницы количества бит извлеченных правильно с применением и без применения двумерного АР процесса (область 1 на рис. 1) к количеству бит извлеченных правильно без применения двумерного АР процесса (область 2 на рис. 1).. На рис. 2 представлен график эффективности применения двумерного АР процесса в зависимости от размера блока изображения.. . Рис. 2. График эффективности применения двумерного АР процесса в зависимости от размера блока изображения. График на рис. 2 показывает на сколько процентов применение двумерного АР процесса позволяет увеличить количество правильно извлеченных бит от количества правильно извлеченных бит без применения двумерного АР процесса.. Из рис. 2 видно, что применение двумерного АР процесса наиболее эффективно при размере блока 32х32 пикселя.. В [1] предложено обнаруживать встроенную ПСП (извлекать встроенные биты) в остатке АР процесса с помощью правила принятия решений инвариантного к аддитивному нормальному шуму неизвестной мощности, при этом ПСП считается детерминированным сигналом (сигналом известной формы). Чтобы попасть в остаток АР процесса ПСП должны иметь авто- и взаимокорреляционные характеристики похожие на аналогичные характеристики белого шума. При маленьком размере блока длина ПСП также маленькая и его авто– и взаимокорреляционные характеристики не похожи на аналогичные характеристики белого шума, поэтому ПСП не попадает в остаток АР процесса. Это объясняет невысокую эффективность применения АР процесса при маленьком размере блока изображения. При увеличении размера блока изображения длина ПСП также увеличивается и его авто– и взаимокорреляционные характеристики все больше начинают быть похожими на аналогичные характеристики белого шума и поэтому все больше ПСП попадает в остаток АР процесса.. Из рис. 1 видно, что падение эффективности применения АР процесса при большом размере блока связано с ограничением сверху области 1 (прирост правильно извлеченных бит) общим количеством встроенных бит (извлекается правильно вся встроенная информация), а снизу – возрастающим количеством правильно извлеченных бит полученных без применения АР процесса. Рост количества правильно извлеченных бит полученных без применения АР процесса связан с увеличением отношения сигнал/шум, которое для широкополосных систем прямо пропорционально длине ПСП, где сигнал – встроенная ПСП, шум – пиксели изображения.. Размер блока, при котором применение двумерного АР процесса наиболее эффективно составляет 32х32 пикселя. Чем больше размер блока, тем меньше бит информации можно встроить при фиксированном размере изображения. Поэтому в [3] предложено использовать пересечение блоков для увеличения объема передаваемой информации кроме того там же определено, что пересечение блоков не влияет на количество правильно извлеченных бит.. Из рис. 1 видно, что применение двумерного АР процесса при размере блока 32х32 пикселя не позволяет полностью избавиться от ошибок извлечения встроенной информации. Ошибка извлечения встроенной информации обусловлена не обнаружением встроенной ПСП (соответствующие биты считаются стертыми) и (или) неправильным обнаружением встроенной ПСП (соответствующие биты считаются ложно обнаруженными). Стирание и ложное обнаружение соответствует представлению изображения в виде двоичного стирающего и двоичного симметричного канала передачи информации.. В двоичном стирающем канале переданный бит 0 или 1 с вероятностью P принимает некоторое третье значение E. При этом если принят символ 0 или 1, то вероятность ошибочного приема равна нулю.. В двоичном симметричном канале на вход поступают двоичные биты, для каждого из них имеется вероятность P того, что бит получен неправильно (ложно) и вероятность 1-P того, что бит получен правильно.. Восстановление стертых и неправильно излеченных бит исследовалось с помощью корректирующих кодов: Боуза–Чоудхури–Хоквингема (БЧХ коды), Рида–Соломона (РС коды) и с низкой плотность проверок на четность (LDPC коды).. Для применения корректирующих кодов необходимо определить параметры: n – длина кода; k – количество информационных бит; R = k /n – скорость передачи кода.. Для определения скорости передачи кода R воспользуемся теоремой Шеннона для канала с шумами, которая гласит что, если скорость передачи сообщений меньше пропускной способности канала связи (R <C ), то существуют коды и методы декодирования, что вероятность ошибки стремиться к нулю при длине блока стремящегося к бесконечности [4].. На рис. 3 представлен график зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса для извлечения встроенных бит при размере блока 32х32 пикселя и пересечением 28х28 пикселей. Коэффициент битовых ошибок можно считать вероятностью возникновения ошибки.. Рис. 3. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса. На рис. 3 при вероятности ложного обнаружения α=0.05 ошибка извлечения в основном обусловлена стираниями битов сообщения. Пропускная способность C двоичного стирающего канала определяется по следующей формуле [5]: C =1-p , где p – вероятность стирания. При α=0.05 вероятность стирания p =0.33, поэтому пропускная способность данного двоичного стирающего канала C =0.67.. Если же на рис. 3 рассматривать вероятность ложного обнаружения α, при котором достигается минимум коэффициента битовых ошибок, т.е. при α=0.5, то при данной вероятности ложного обнаружения α ошибка извлечения обусловлена только ошибкой неправильного извлечения бит. Пропускная способность C двоичного симметричного канала определяется по следующей формуле [5]: C = 1+ p log2 p + (1- p )log2 (1- p ), где p – вероятность ошибки. При α=0.5 вероятность ошибки p =0.03, поэтому пропускная способность данного двоичного стирающего канала C =0.8. Извлечение встроенной информации осуществляется при маленькой вероятности ложного обнаружения (α=0.05), поскольку это позволяет определить наличие или отсутствие встраивания, поэтому скорость передачи кода должна удовлетворять условию: R <0.67. Из данного условия выбираем R =0.5.. Особенностью БЧХ кода является снижение корректирующей способности при увеличении длины кодового слова [6]. Поэтому выбираем наименее возможную длину кода n =7, для этой длины кода единственная стандартная длина информационного слова k =4. При этом в кодовом слове количество возможных исправлений неправильно излеченных бит t =1 или количество возможных исправлений стираний l =2.. На рис. 4 представлено три графика: первый график (обозначен цифрой 1) представляет собой зависимость коэффициента битовых ошибок BER от вероятности ложного обнаружения α при отсутствии использования двумерного АР процесса и корректирующих кодов для извлечения встроенных бит; второй график (обозначен цифрой 2) с использованием только двумерного процесса авторегрессии для извлечения встроенных бит; третий график (обозначен цифрой 3) с использованием двумерного АР процесса и БЧХ кода для извлечения встроенных бит.. . Рис. 4. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α для исследования применения БЧХ кода. Из рис. 4 видно, что использование кода БЧХ в общем уменьшает ошибку извлечения бит, но не полностью избавляет от ошибок извлечения. Это связано с тем, что ошибки неравномерно распределены по всей последовательности извлеченных бит, так как в одном блоке кодового слова длиной n =7 бит может быть исправлено не более одного неправильно излеченного бита или не более двух стираний.. Коды Рида–Соломона являются недвоичными БЧХ кодами. Корректирующая способность кода Рида-Соломона зависит только от разности длины кодового слова n и длины информационного слова k [7]. В качестве параметров возьмем длину кодового слова n =255, чтобы 8 информационных двоичных бит составляли символ кодового слова, а количество информационных символов k =127.. На рис. 5 представлено три графика: первый и второй график аналогичны графикам из рис. 4; третий график (обозначен цифрой 3) с использованием двумерного АР процесса и кода Рида–Соломона для извлечения встроенных бит.. Рис. 5. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α для исследования применения кода Рида-Соломона. Из рис. 5 видно, что результаты, полученные с применением кода Рида-Соломона аналогичны результатам, полученным с применением БЧХ кода. Это связано с тем, что код Рида-Соломона может исправлять локализованные пакеты ошибок. Здесь ошибки распределены по всей последовательности извлеченных бит.. Для LDPC кода возьмем следующие параметры: длина кодового слова n =255, количество информационных символов k =127.. Декодирование LDPC кода в бинарном стирающем канале является простым. Восстанавливаем все стертые элементы кодового слова, так чтобы это удовлетворяло проверочному уравнению [8]: sHT =0 (mod 2), где H – матрица проверки на четность LDPC кода, s – кодовое слово.. Для декодирования LDPC кода в двоичном симметричном канале передачи информации существует несколько алгоритмов. Так в [9] Галлагер предложил два алгоритма декодирования LDPC кодов: алгоритм с инверсией бита (Bit flip) и алгоритм с распространением доверия (Belief propagation).. График зависимости коэффициента битовых ошибок BER от вероятности ложного обнаружения α с последовательным применением двумерного АР процесса, алгоритма восстановления стертых элементов кодового слова, алгоритма Bit flip представлен на рис. 6, а график с алгоритмом Belief propagation представлен на рис. 7.. Из рис. 6 и 7 видно, что применение алгоритма Belief propagation является более эффективным, что обусловлено «мягким» принятием решения.. Если сравнить результаты, полученные с помощью LDPC кода с алгоритмом Belief propagation, с результатами, полученными с помощью БЧХ кода и с помощью кода Рида-Соломона, то видно, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.. В работе определен размер блока изображения, при котором наиболее эффективно применять двумерный АР процесс для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра, это блок размером 32х32 пикселя. Объяснена небольшая эффективность применения двумерного АР процесса при маленьком и большом размере блока изображения.. Предложено представлять изображение последовательно в виде двоичного стирающего канала передачи информации и двоичного симметричного канала передачи информации, а стертые и неправильно принятые биты данных восстанавливать с помощью корректирующих кодов.. Определена величина скорости передачи кода R =0.5 с помощью теоремы Шеннона и графика коэффициента битовых ошибок BER от вероятности ложного обнаружения α с использованием только двумерного АР процесса.. Исследована эффективность применения для восстановления встроенных данных с помощью БЧХ кода, кода Рида-Соломона и LDPC кода.. Определено, что применение LDPC кода является более эффективным для увеличения количества правильного извлечения встроенной информации стеганографическим методом на основе прямого расширения спектра.