А.М.Кричевский

# Нейронные сети в задачах компьютерной стеганографии

*Some variants of neural network implementation of steganographical protection are considered.*

При передаче информации одной из важнейших является задача обеспечения конфиденциальности информации. Для ее решения можно использовать различные методы криптографической защиты. Однако криптографическая защита не является достаточной для обеспечения секретности информации, поскольку зашифрованное сообщение будет легко обнаружено. Необходимо не только зашифровать, но и скрыть зашифрованную информацию, для чего можно использовать методы стеганографии. Под стеганографией подразумевается техника сокрытия некоторой секретной информации в больших информационных массивах таким образом, чтобы непосвященный наблюдатель не мог заметить существования этой информации.

Методология компьютерной стеганографии основана на замене несущественных или неиспользуемых массивов данных компьютерных файлов необходимой конфиденциальной информацией, называемой цифровым водяным знаком или цифровой меткой. В результате обработки файла-оригинала методами стеганографии получают файл, сохраняющий свое функциональное назначение, практически неотличимый человеком от оригинала, но содержащий секретную информацию, что позволяет идентифицировать принадлежность этого файла или передать секретную информацию.

Процесс стеганографии в общем случае состоит из следующих этапов [1,2].

1. **Подготовительные операции:**
   - создание цифровой метки, т. е. графического или текстового файла, который необходимо скрыть.
   - выбор файла-контейнера, в который необходимо внедрить цифровую метку. Этот файл, как правило, должен быть во много раз больше, чем метка.
2. **Селекция носителя** - выбор списка элементов оригинала, хранящегося в файле-контейнере, используемых для внедрения цифровой метки, представленной в виде шумоподобного сигнала. Программа или устройство, обеспечивающие селекцию носителя, называется селектором носителя.
3. **Генерация** шумового сигнала псевдослучайной последовательностью, независимой от носителя и определяемой секретными ключами (преобразование метки в шум).
4. **Добавление** сгенерированного шума к выбранному носителю.

Процесс извлечения метки, в свою очередь, состоит из следующих этапов.

5. Селекция носителя (см. выше).
6. Фильтрация (отбор, извлечение) шумового сигнала.
7. Преобразование шума в метку.

В качестве примера рассмотрим способ внедрения цифровой метки в черно-белое изображение, основанный на манипуляциях шумом, вносимым в небольшие фрагменты изображения-контейнера. После подготовительных операций на этапе селекции носителя исходное изображение-контейнер разбивается на небольшие фрагменты размером $pxq = N$ точек, где $p$-высота, а $q$-ширина фрагмента.

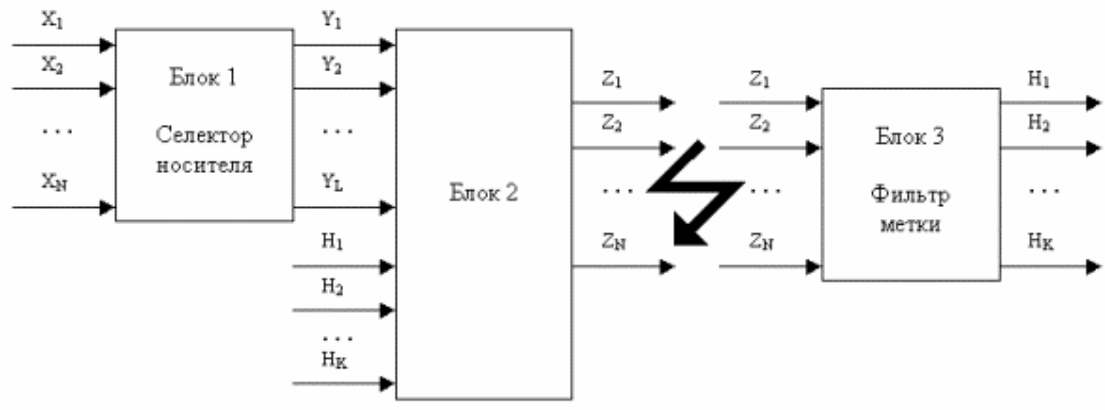
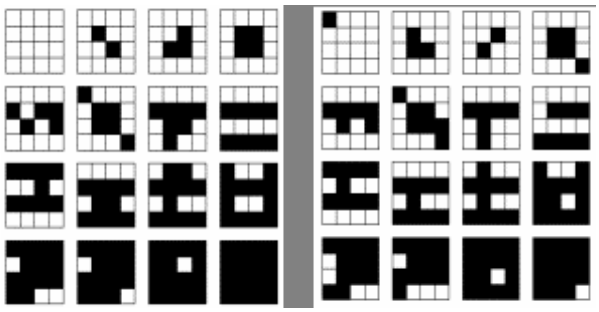



Рис.1. Схема процесса стеганографической защиты черно-белых изображений

Блок 1 (селектор-носителя) получает на входе бинарные значения кодов цвета точек $X_1,X_2,...,X_N$ выделенного фрагмента исходного изображения (рис.1). Соответствие точки фрагмента и переменной $X_i$ задается произвольным способом по желанию разработчика и может служить дополнительной защитой метки.

С помощью Блока 1 для каждого фрагмента исходного изображения-контейнера подбирается наиболее близкий к нему фрагмент из набора А (рис.2а). Код этого фрагмента в виде двоичных сигналов $Y_1,Y_2,...,Y_L$ с выходов Блока 1, а также цифровая метка в виде двоичных сигналов $H_1, H_2,...,H_K$ подаются на соответствующие входы Блока 2, который в соответствии с некоторыми правилами на основе кодов метки и фрагмента, выдает на выходе фрагмент из набора А или В (рис. 2б), соответствующий поданному на его вход коду фрагмента. Набор В выбирается таким, чтобы его элементы были максимально похожи на соответствующие элементы набора А, но незначительно отличались от них. Кроме того, не допускается наличие в наборах двух одинаковых элементов.



а) б)

Рис.2. Наборы фрагментов изображений

Таким образом, Блок 2 осуществляет генерацию и добавление шумового сигнала к выбранному носителю. Из фрагментов, полученных на выходе Блока 2, точки которых кодируются цветами, соответствующими уровням сигналов $Z_1,Z_2,...,Z_N$, составляется изображение, подобное оригиналу, но содержащее цифровую метку.

Для идентификации принадлежности файла, содержащего цифровую метку, необходимо выполнить извлечение метки, которое производится путем разбиения изображения на фрагменты и извлечения из полученных фрагментов частей метки Блоком 3.

Процесс стеганографической защиты, представленный на рис.1, может быть реализован различными способами. Однако наиболее перспективной является реализация

на основе нейронных сетей, так как они обладают большой способностью к ассоциациям и могут обучаться эффективно классифицировать изображения.

Стеганографическая защита может быть реализована посредством следующих методов.

Вариант 1. В качестве Блока 1 используется дискретный многослойный персептрон (МСП), который обучается с помощью алгоритма обратного распространения ошибки (ОРО) на основе набора А, а в качестве Блоков 2 и 3 – ассоциативная память, основанная на нейронной сети Хопфилда, обученная на основе набора В и значений метки.

Вариант 2. В качестве объединения Блоков 1 и 2 используется дискретный многослойный персептрон, который обучается на основе наборов А (подается на вход) и В (подается на выход), а также значений метки (подается на вход). В качестве Блока 3 используется многослойный дискретный персептрон или нейросетевая память Хопфилда, которые обучаются на основе набора В и значений метки. Этот способ является наиболее простым, но в тоже время менее надежным и более длительным в обучении.

Рассмотрим принцип действия и обучения МСП и сети Хопфилда.

Как известно [3], в нейронных сетях процесс программирования заменяется процессом обучения. МСП, относящийся к классу прямонаправленных сетей, обычно состоит из трех слоев нейронов (вычислительных элементов, которые подобны их биологическому прототипу - реальному нейрону): входного, скрытого и выходного. Сущность обучения такой сети сводится к минимизации ошибки между реальным значением выходного сигнала $y$ и требуемым (целевым) сигналом $t$. Минимизация ошибки достигается в результате адаптации весов $w$ между слоями сети посредством метода ОРО. Алгоритм предложенного метода показан на рис.3.

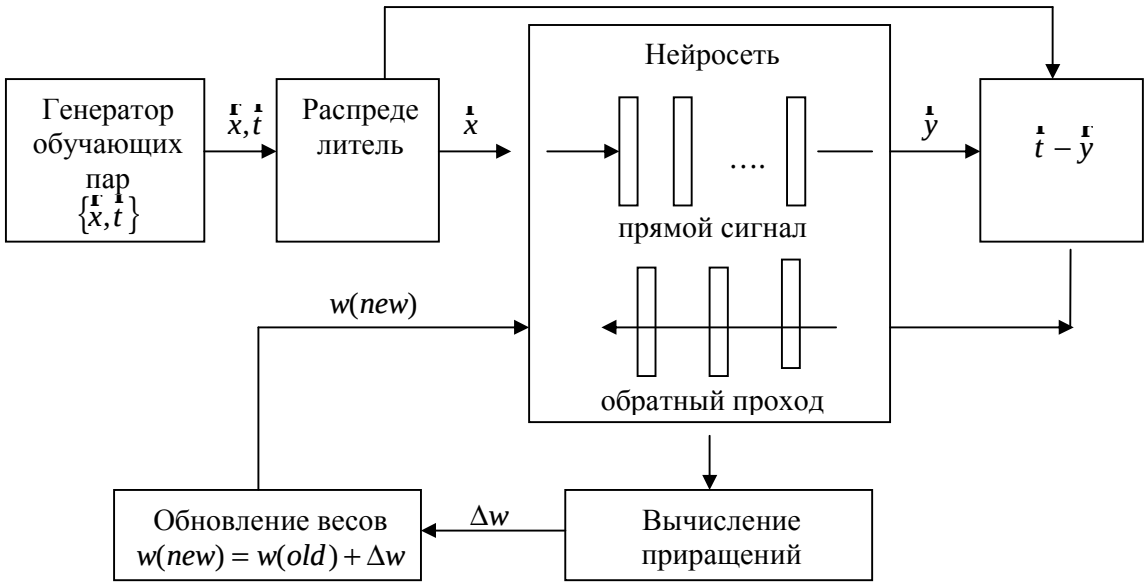



Рис. 3. Вычислительная схема метода обратного распространения ошибки

В зависимости от реализуемого варианта нейросетевой стеганографии входными и целевыми векторами могут быть разные наборы данных.

Сеть Хопфилда, принадлежащая к классу рекуррентных нейронных сетей, играет роль ассоциативной памяти. Главная задача ассоциативной памяти сводится к запоминанию входных (обучающих) выборок таким образом, чтобы при предъявлении новой выборки система могла сгенерировать ответ: какая из запомненных ранее выборок наиболее близка к вновь поступающему образу.

Архитектура сети Хопфилда представляется, как правило, в виде системы с непосредственной обратной связью выхода со входом (рис.4). Характерная особенность сети заключается в том, что выходные сигналы нейронов являются одновременно входными сигналами сети: $x_i(k) = y_i(k-1)$. В классической сети Хопфилда отсутствует связь нейрона с собственным вектором, а матрица весов является симметричной.

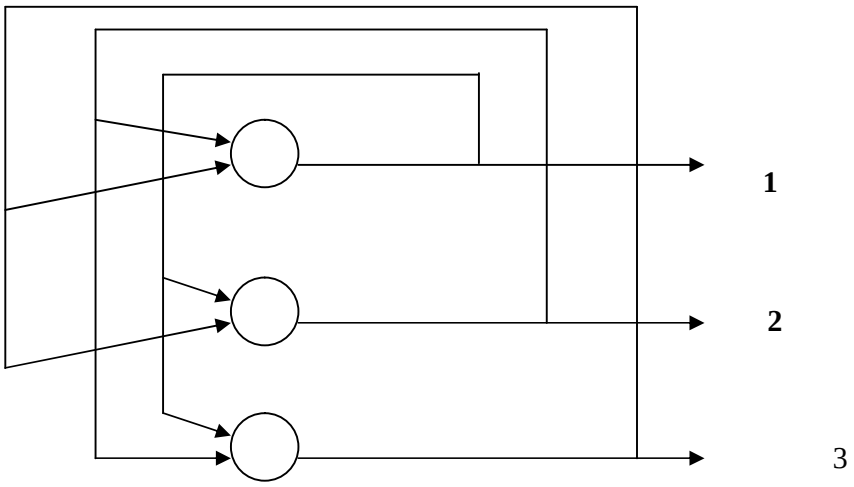



Рис.4. Архитектура сети Хопфилда

В процессе обучения сети формируются зоны притяжения (аттракторы) некоторых точек равновесия, соответствующих обучающим данным. При использовании ассоциативной памяти мы имеем дело с обучающим вектором $\boldsymbol{x}$ или с множеством этих векторов, которые в результате проводимого обучения определяют расположение конкретных аттракторов.

Нейронные сети указанных типов могут справиться с решением задачи стеганографической защиты изображений и позволяют рекомендовать предложенные способы для внедрения и использования на практике. Дальнейшие работы в этом направлении будут включать моделирование предложенных ситуаций.

**Литература**

1. В.И. Дубровин, С.А. Субботин. Защита изображений в распределенных системах передачи информации на основе нейросетевой стеганографии. - http://mimicria.narod.ru/HTML/Stego/Html/a1.htm

2. N.J. Johnson, S.Jajodia. Steganalysis of images created using current stegonography software. (in Lecture Notes in Computer Science, v.1525, Berlin, Springer-Verlag, 1998, pp. 273-289).

3.С. Осовский. Нейронные сети для обработки информации. - М., Финансы и статистика, 2002.