УДК 621.391

К.С. Васюта, С.А. Щербинин

Харьковский университет Воздушных Сил имени Ивана Кожедуба, Харьков

МЕТОД СКРЫТОЙ ПЕРЕДАЧИ ЦИФРОВОЙ ИНФОРМАЦИИ НА ОСНОВЕ НЕЛИНЕЙНО "АВТОРЕГРЕССИОННОГО УСЛОВНОГО ГЕТЕРОСКЕДАСТИЧЕСКОГО" ПРЕОБРАЗОВАННОГО СТОХАСТИЧЕСКОГО ПРОЦЕССА

В работе проанализировано возможность применения нелинейно преобразованного стохастического "авторегрессионного условного гетероскедастического" процесса для повышения структурной скрытности передаваемой информации. Предложен метод оценки показателей качества модели скрытия информации, а так же приведены зависимости вероятности правильной оценки бита информации от отношения сигнал/шум на входе приемника при разных значениях длины элементов кодируемой последовательности и при разных значениях объема усредняемой информации.

Ключевые слова: *авторегрессионное условное гетероскедастическое, коэффициент автокорреляции, скрытность.*

Введение

Увеличение объема обслуживаемых абонентов, а также необходимости повсеместного обеспечения возрастающих требований к качеству обслуживания QoS (Quality of Service) привело к переходу от статических средств обработки информации (электрических схем с заранее запрограммированным алгоритмом обработки) к динамическим (программные продукты с адаптивным изменением протоколов обработки входной информации). Данный факт обусловил экспоненциальное повышение требований к электромагнитной совместимости средств связи, а так же конфиденциальности передаваемой информации.

Поскольку, наиболее адекватными методами решения проблемы электромагнитной совместимости и увеличения количества обслуживаемых абонентов являются методы кластерного (сотового) разбиения то и обеспечение конфиденциальности передаваемой информации должно базироваться на тех же протокольных решениях.

С целью одновременного решения этих задач было предложено использовать в качестве опорных сигналов линейно преобразованные стохастические последовательности с кратковременной памятью, которые обладают сверхширокополосностью, что позволяет применять их в системах «криптованой» сотовой связи, а широкий "спектр" стохастических аттракторов, позволяет обеспечить структурную скрытность и скрытие факта передачи информации.

В общем случае класс моделей с нелинейными структурами должен быть значительно шире по своему составу, чем класс линейных моделей [1], что дает возможность предположить, что данный вид сигналов позволяет обеспечить расширения классов сложных аттракторов тем самым повысить качество обработки и скрытности передаваемой информации.

Таким образом, целью работы является синтез метода скрытой передачи и обработки цифровой информации на основе нелинейно преобразованных стохастических последовательностей, для повышения структурной скрытности и скрытия факта передачи информации.

Результаты исследований

Идея формирования нелинейных моделей временных рядов заключается в представлении ряда X_t в виде нелинейной функции, аргументами которой являются текущее и предшествующие значения стандартизованной случайной переменной ξ_t, свойства которой соответствуют белому шуму, $\xi_t \sim N(0,\ 1)$, $M[\xi_t,\ \xi_{t-i}] = 0,\ i > 0$:

$$X_t = f(\xi_t,\ \xi_{t-i},\ \xi_{t-2}\ldots). \qquad (1)$$

При этом в выражении (1) используется функция $f(.)$, образованная суммой двух функций, одна из которых выражает условное математическое ожидание переменной X_t на прошлые значения ошибки, а другая – условную дисперсию:

$$X_t = g(\xi_t,\ \xi_{t-i},\ \xi_{t-2}\ldots) + \\ +\xi_t \cdot h(\xi_t,\ \xi_{t-i},\ \xi_{t-2}\ldots), \qquad (2)$$

где функции $g(.)$ и $h(.)$ могут быть как линейными, так и нелинейными.

Из (2) следует, что условное математическое ожидание X_t определяется как $M[X_t/\xi_{t-1}, \xi_{t-2},\ldots] = g(\xi_{t-1}, \xi_{t-2},\ldots)$ в силу свойств переменной ξ_t $(M[\xi_t] = 0,\ M[\xi_t, \xi_{t-i}] = 0,\ i = 1,2,\ldots)$, а условная дисперсия

$$D\left(X_t/\xi_{t-1},\xi_{t-2},\ldots\right) = M\left[X_t - g\left(\xi_{t-1},\xi_{t-2},\ldots\right)\right]^2 =$$
$$= M\left[\xi_t^2\right]\cdot M\left[h^2\left(\xi_{t-1},\xi_{t-2},\ldots\right)\right] = h^2\left(\xi_{t-1},\xi_{t-2},\ldots\right).$$

Все условные центральные моменты более высоких порядков переменной X_t определяются как

$$M\left[X_t - g(.)\right]^n = M\left[\xi_t^n\right]\cdot h^n(.).$$

Выражение (2) можно рассматривать и как результат разложения функции (1) в ряд Тейлора в точке $\xi_t = 0$ при заданных значениях $\xi_{t-1}, \xi_{t-2},\ldots$

$$X_t = f\left(0,\xi_t,\xi_{t-1},\xi_{t-2}\ldots\right) + f_1\left(0,\xi_t,\xi_{t-1},\xi_{t-2}\ldots\right) + \frac{1}{2}\xi_t^2 f_2\left(0,\xi_t,\xi_{t-1},\xi_{t-2}\ldots\right). \quad (3)$$

Пренебрегая в выражении (3) слагаемыми, содержащими ξ_t^n, $k > 1$, получим:

$$g\left(\xi_{t-1},\xi_{t-2}\ldots\right) = f\left(0,\xi_{t-1},\xi_{t-2}\ldots\right), \quad (4)$$

$$h\left(\xi_{t-1},\xi_{t-2}\ldots\right) = f_1\left(0,\xi_{t-1},\xi_{t-2}\ldots\right). \quad (5)$$

Из (3) следует, что все множество нелинейных моделей может быть разделено на две группы: модели с нелинейным условным математическим ожиданием и модели с нелинейной условной дисперсией.

Поскольку в работе [2] уже было показано, что модели с линейной структурой условного математического ожидания могут применяться для повышения структурной скрытности передаваемой информации, то очевидно ожидать, что и модели с нелинейным условным математическим ожиданием так же могут быть применимы для скрытия факта передачи информации.

Для исследования данного факта и упрощения математических выкладок ниже будет рассматриваться только модель типа ARCH (авторегрессионное условное гетероскедастического) первого порядка, которая характеризуется нелинейной условной дисперсией. Описать ее можно следующей математической моделью:

$$X_t = \xi_t\sqrt{\alpha_0 + \alpha_1\xi_{t-1}^2}\ . \quad (6)$$

где ξ_t – белый Гауссовский шум с нулевым математическим ожиданием и единичной дисперсией, α_k – параметры модели.

В данном случае условное, как и безусловное математическое ожидание, равно нулю, $g(.) = 0$, и функция $h(.) = 0$ определяется выражением $\sqrt{\alpha_0 + \alpha_1\xi_{t-1}^2}$, так что условная дисперсия $D\left(X_t/\xi_{t-1},\xi_{t-2},\ldots\right) = \alpha_0 + \alpha_1\xi_{t-1}^2$ также зависит от квадрата предыдущего значения ошибки.

Для анализа свойств модели ARCH было проведено моделирование с разными значениями параметра моделей $\alpha_0,\alpha_1 = \text{var}$ при условии что $\alpha_0 + \alpha_1 = \text{const}$, данное условие обеспечивает неразличимость сигналов в рамках спектрального и корреляционного анализа рис. 1. Из рисунка видно, что нелинейное разрушение случайного порождающего процесса модели ARCH практически не влияет на структуру спектра порождающего процесса, но приводит к формированию в фазовой плоскости двух секторов вдоль оси ординат, которыми ограничивается область возможных значений поведения модели. Так же наблюдается зависимость образования сектора ограниченности от соотношения параметров α_0,α_1. При возрастании отношения α_1 / α_0 глубина сектора ограниченности увеличивается, а угол сектора уменьшается.

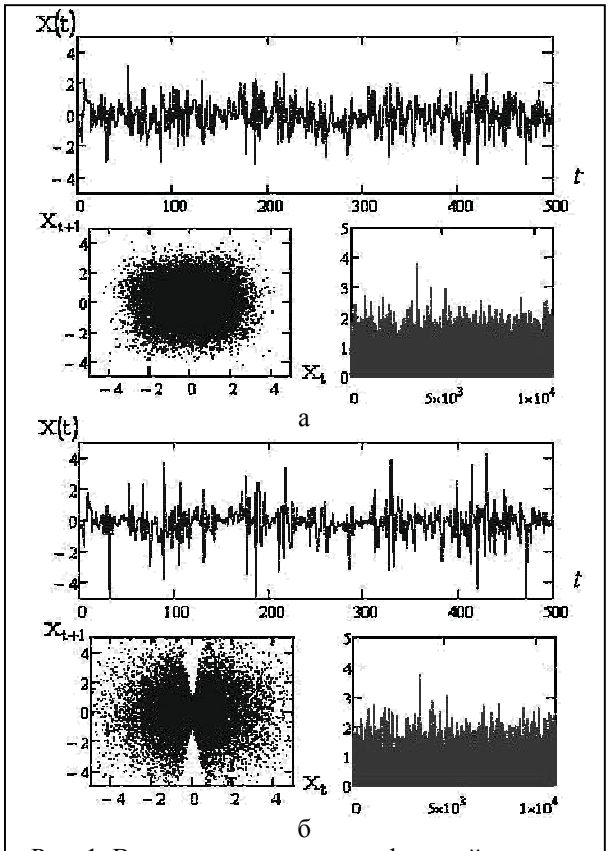

Рис. 1. Временная реализация, фазовый портрет и спектр процесса: а – при α_0=1; α_1=0,1
б – при α_0=0,1; α_1=1

Проведя внесение бинарной последовательности (рис. 2) методом, предложенным в [3]. Получаем временную реализацию и фазовый портрет, представленный на рис. 3.

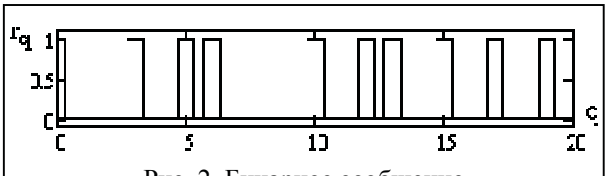

Рис. 2. Бинарное сообщение

Из рисунка видно, что вид передаваемой информации во временной, спектральной и фазовой плоскости не отличим от белого шума, при условии, что количество битов соответствующим "0" и "1" в передаваемой информации равны.

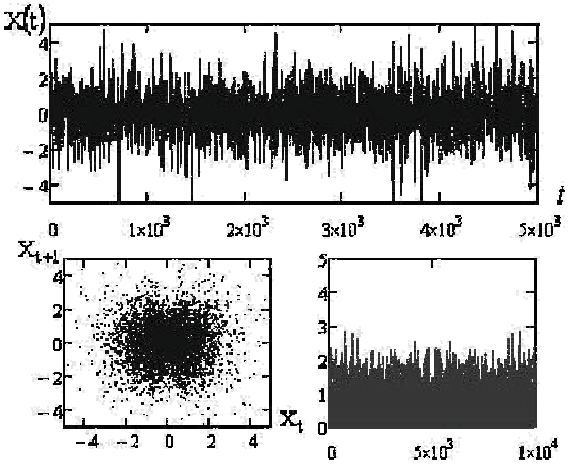

Рис. 3. Временная реализация, фазовый портрет и частотный спектр сигнала на входе приемника

Декодирование принятой информации.

На практике невозможно сформировать стохастические сигналы, идеально удовлетворяющие их математической интерпретации, поскольку всегда существуют неординарные события (перепады напряжения, внутренние шумы передатчика, псевдослучайность формирования порождающего процесса белого шума) вызывающие резкие изменения (скачки) напряжения. Однако подобные события случаются нечасто, и сопровождающие их изменения также достаточно редки, и в самих этих изменениях не прослеживается наличие каких-либо внутренне присущих им закономерностей.

Исходя из этого, значение показателя X_t в момент t может быть определено уравнением

$$X_t = \mu(t) + \nu_t \cdot \xi_t, \qquad (7)$$

где $\xi_t \sim N(0,1)$ – стандартизованная случайная переменная, свойства которой предполагаются соответствующими свойствам "строгого белого шума", ν_t – процесс, образованный положительной случайной переменной, равной условному стандартному отклонению, так что $D(X_t/\nu_t) = \nu_t^2$.

Предполагая, что значения ν_t представляют собой условное стандартное отклонение, являющееся детерминированной функцией от прошлых значений

$$\nu_t = f(X_{t-1}) = \left(\alpha_0 + \alpha_1 (X_{t-1} - \mu)^2\right)^{1/2}, \qquad (8)$$

а также предполагая, что независимость условного стандартного отклонения от уровня значений, но допуская, что переменная ν_t может быть представлена функцией авторегрессии – скользящего среднего

$$\nu_t = \phi(\nu_{t-1}, \nu_{t-2}, \ldots, \eta_t), \qquad (9)$$

ставящей ее уровень в момент t в зависимость от ее значений в прошедшие периоды времени и случайной составляющей η_t. Случайная составляющая η_t отличается от ошибки ε_t, присутствовавшей в моделях значений, хотя их статистические свойства идентичны:

$$\eta_t \sim N(0, \sigma_\eta^2),\ \varepsilon_t = X_t - \mu(t),\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2).$$

Таким образом, общий вид ARCH-модели имеет вид:

$$X_t - \mu = \xi_t \cdot \left(\alpha_0 + \sum_{i=1}^{n} \alpha_i (X_{t-i} - \mu)^2\right)^{1/2}. \qquad (10)$$

С учетом ограничения на порядок модели типа ARCH, проведя замену переменных, $s_t = (X_t - \mu)^2$ получим, что

$$M[s_t / s_{t-1}] = \alpha_0 + \alpha_1 s_{t-1}, \qquad (11)$$

где $M[s_t/s_{t-1}]$ – условное математическое ожидание переменной s_t при известном значении s_{t-1}.

При условии, что дисперсия переменной существует, в соответствии, с выражением, представленном в [1] имеем

$$\alpha_1 = r_1(s_t) \qquad (12)$$

где $r_1(s_t)$ – первый коэффициент автокорреляции ряда s_t, вычисляющийся по следующей формуле

$$r_1 = \frac{\sum_{t=2}^{n} (s_t - \bar{s}_1)(s_{t-1} - \bar{s}_2)}{\sqrt{\sum_{t=2}^{n} (s_t - \bar{s}_1)^2 \sum_{t=2}^{n} (s_{t-1} - \bar{s}_2)^2}}, \qquad (13)$$

где $\bar{s}_1 = \frac{\sum_{t=2}^{n} s_t}{n-1}$, $\bar{s}_2 = \frac{\sum_{t=2}^{n} s_{t-1}}{n-1}$.

Переходя к безусловному математическому ожиданию переменной s_t на основе выражения (11) получим:

$$M[s_t] = \alpha_0 + \alpha_1 M[s_t], \qquad (14)$$

откуда следует, что параметр α_0 находится из соотношения

$$\alpha_0 = M[s_t](1 - \alpha_1), \qquad (15)$$

где $M[s_t] = \frac{1}{T}\sum_t (X_t - \mu)^2$.

Аналогичные рассуждения справедливы и в отношении общей модели (10). Ее коэффициенты приблизительно могут быть оценены на основании системы Юла-Уокера для ряда s_t, а коэффициент α_0 найден из соотношения:

$$\alpha_0 = M[s_t]\left[1 - (\alpha_1 + \alpha_2 + \ldots \alpha_n)\right]. \qquad (16)$$

Найденные оценки далее уточняется на основе использования ММП (метода максимального правдоподобия). В общем случае оценке подлежат следующие параметры: $\mu,\alpha_0,\alpha_1,\ldots,\alpha_n$. Функция правдоподобия в предположении, что ошибка распределена по нормальному закону представляется в следующем виде:

$$L_n\left(\alpha,\hat{\mu}\right) = \prod_{t=n+1}^{T} f\left(X_t/I_{t-1},\alpha,\hat{\mu}\right), \qquad (17)$$

где $\alpha = \left(\alpha_0,\alpha_1,\ldots,\alpha_k\right)$ – вектор оценок параметров;

$\hat{\mu}$ – оценка математического ожидания;

$$f\left(X_t/I_{t-1},\alpha,\hat{\mu}\right) = f\left(X_t/\nu_t\right) = \frac{1}{\nu_t\sqrt{2\pi}} e^{-\frac{1}{2}\left(Y_t-\hat{\mu}\right)^2/\nu_t^2} \qquad (18)$$

и

$$I_{t-1} = \left\{Y_1,Y_2,\ldots,Y_{t-1}\right\}.$$

Натуральный логарифм функции (17) равен

$$\ln L_n\left(\alpha,\hat{\mu}\right) = -\frac{1}{2}\left(T-k\right)\ln\left(2\pi\right) - \sum_{t=n+1}^{T}\ln\nu_t - \frac{1}{2}\sum_{t=1}^{T}\left(Y_t-\hat{\mu}\right)^2/\nu_t^2. \qquad (19)$$

Максимизируя выражение (19) по параметрам $\hat{\mu},\alpha_0,\alpha_1,\ldots,\alpha_n$ при условии, что

$$\nu_t^2 = \alpha_0 + \sum_{i=1}^{n}\alpha_i\left(X_{t-i}-\hat{\mu}_{t-i}\right)^2 \qquad (20)$$

получаем систему нелинейных уравнений, выражающих эти параметры через значения процесса X_t. Далее, решая эти уравнения, находятся искомые значения $\hat{\mu},\alpha_0,\alpha_1,\ldots,\alpha_n$.

Опираясь на изложенные выше математические выкладки построены графики зависимости изменения оценки показателя модели от отношения сигнал/шум на входе приемника (рис. 4).

Из рисунка видно, что с возрастанием отношения сигнал/шум на входе приемника все значения оценок параметров модели возрастают по степенному закону, что приводит к неточностям при восстановлении принимаемой информации. Для исключения данной недостатка вычисляется среднее арифметическое полученных значений оценок параметра принимаемой информации (в частном случаи некоторого объема усредняемой информации D), а затем полученное значение вычитается из каждой из полученных оценок.

Принятие решения о наличии бита информации осуществляется на основе критерия отношения правдоподобия, а алгоритм восстановления бинарного сообщения, заданного вектором $\hat{r}$ с числом

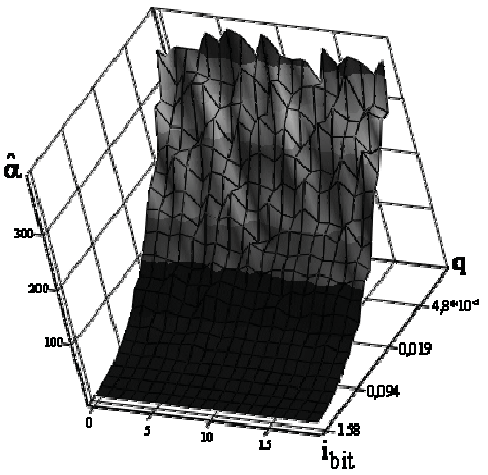

Рис. 4. Зависимости изменения оценки показателя модели от отношения сигнал/шум на входе приемника

элементов L ,можно представить в следующем виде:

$$\hat{\alpha}_i - \frac{1}{T}\sum_{t=i}^{i+T}\partial_t \begin{matrix}\hat{r}=0\\ >\\ <\\ \hat{r}=1\end{matrix} 0, \quad \text{при } T \geq 20, \qquad (22)$$

где T – некоторый объем принимаемой информации.

Оценка элемента (бита) сообщения сводится к проверке гипотезы о наличии или отсутствии в восстановленном сообщении $\hat{r}$ на заданном интервале символа "0" или "1", а качество оценки определяется значением вероятности правильного принятия решения.

На рис. 5 представлены зависимости вероятности правильной оценки $P_{\vec{r}}(L,q) = 1 - P_{err.\vec{r}}(L,q)$ сообщения $\vec{r}$ от значения объема усредняемой информации D бинарного сообщения и отношения сигнал/шум на входе приемника $q = \sigma_S^2 / \sigma_n^2$.

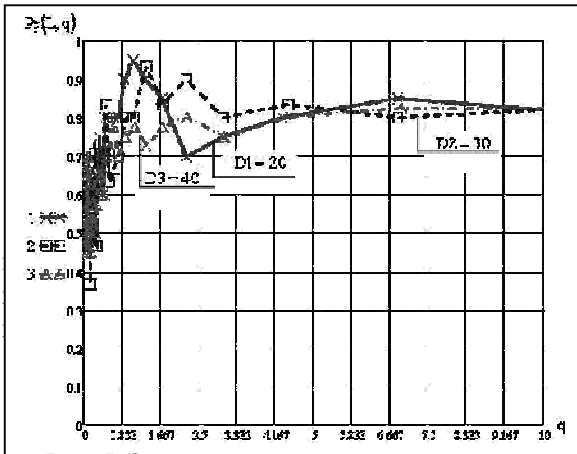

Рис. 5. Зависимости вероятности правильной оценки от отношения сигнал/шум на входе приемника при разных значениях объема усредняемой информации при L=500

Величина $P_{err.\vec{r}} = d(\vec{r}, \hat{\vec{r}}) / Q$ определяет долю ошибок в оценках элементов сообщения и равна отношению расстояния Хемминга $d(\vec{r}, \hat{\vec{r}})$ между передаваемой бинарной последовательностью $\vec{r}$ и её оценкой $\hat{\vec{r}}$ к общему числу L ее элементов.

Проведя оценку эффективной длины кодированной последовательности, получим график, представленный на рис. 6.

Рис. 6. Зависимости вероятности правильной оценки от длины вариации длинны кодирующей последовательности при D=40

Из рисунка видно, что наиболее эффективной длины кодированной последовательности является $L = 460 \pm 10\%$.

Выводы

Применение моделей с нелинейными структурами дает возможность расширить состав моделей, которые могут применяться для скрытой передачи информации.

Рассмотренная модель передачи информации на основе модели с нелинейной условной дисперсией имеет ряд преимуществ над моделями с линейным и нелинейным условным математическим ожиданием. Они заключаются в полном сходстве частотного спектра со спектром порождающей случайной последовательности, а так же в малой структурируемости аттракторов при изменении параметров модели тем самым обеспечивается повышенная скрытность в процессе спектрального, корреляционного и нелинейного анализов.

Однако предложенный метод оценки имеет малую стойкость к воздействию помех, что обусловлено применением его в относительно "чистых" каналах передачи информации.

В дальнейшем авторами будет исследована возможность применения более сложных моделей скрытой передачи информации, в состав которых входят, как составляющие условного математического ожидания та и условной дисперсии.

Список литературы

1. Тихомиров Н.П. Эконометрика / Н.П. Тихомиров, Е.Ю. Дорохина // – М.: Изд-во Рос. экон. акад., 2002. 640 с.

2. Васюта К.С. Метод скрытой передачи бинарной информации на основе применения линейнопреобразованного стохастического процесса "скользящего среднего" / К.С. Васюта, С.А. Щербинин / – Х. Системи озброєння і військова техніка. – 2014. – № 1(37). – С. 108-111.

3. PulsON Technology. Time Modulated Ultra-Wideband for Wireless Applications.-Time Domain Corporation, 2000.

Поступила в редколлегию 15.08.2014

Рецензент: д-р техн. наук, проф. П.Ю. Костенко, Харьковский университет Воздушных Сил имени Ивана Кожедуба, Харьков.

МЕТОД ПРИХОВАНОЇ ПЕРЕДАЧІ ЦИФРОВОЇ ІНФОРМАЦІЇ НА ОСНОВІ НЕЛІНІЙНО "АВТОРЕГРЕСІЙНОГО УМОВНОГО ГЕТЕРОСКЕТАСТИЧНОГО" ПЕРЕТВОРЕНОГО СТОХАСТИЧНОГО ПРОЦЕСУ

К.С. Васюта, С.О. Щербінін

В роботі проаналізовано можливість застосування нелінійно перетворено стохастичного "авторегресійного умовного гетероскетастичного" процесу для підвищення структурної прихованості інформації, що передається. Запропоновано метод оцінки показника якості моделі приховання інформації, а також приведені залежності вірогідності правильної оцінки біта інформації від відношення сигнал/завада на вході приймача при різних значеннях довжини елементів послідовності, що кодується, та при різних значеннях об'єму інформації, що усереднюється.

Ключові слова: *авторегресійне умовне гетероскедастичне, коефіцієнт автокореляції, прихованість.*

THE METHOD OF DIGITAL INFORMATION HIDDEN TRANSMISSION BASED ON NONLINEAR "AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTIC" CONVERTED STOCHASTIC PROCESS

K.S. Vasyuta, S.A. Shcherbinin

The paper analyzes the possibility of using nonlinear converted stochastic "autoregressive conditional heteroscedastic" process to improve the structural security of transmitted information. The method for estimating the quality parameters of information hiding model is proposed and also the dependence of the correct evaluation probability of information bits on the signal / noise ratio at the receiver at different length values of the encoded sequence elements and at different values of the averaged information volume.

Keywords: *autoregressive conditional heteroscedastic, autocorrelation coefficient, security.*