НОВІ ЗАСОБИ КІБЕРНЕТИКИ, ІНФОРМАТИКИ, ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ ТА СИСТЕМНОГО АНАЛІЗУ

А.К. ЮДИН, Ю.К. ЗИАТДИНОВ, А.Н. ВОРОНИН, А.В. ИЛЬЕНКО

УДК 658.52

МЕТОД ОПРЕДЕЛЕНИЯ ИНФОРМАТИВНЫХ СОСТАВЛЯЮЩИХ НА ОСНОВАНИИ ПОСТРОЕНИЯ ПОСЛЕДОВАТЕЛЬНОГО ПРАВИЛА ПРИНЯТИЯ РЕШЕНИЯ

Аннотация. Исследован новый метод определения информативных составляющих сигнала при введении заданной функции неопределенности из условий построения последовательного правила принятия решения. На основе проведенных исследований определена возможность остановки процесса декодирования в случае соответствия полученных результатов установленным порогам принятия решения.

Ключевые слова: статистические правила принятия решений, минимально-достаточное количество информации, информативные параметры сигналов, идентификация, функция неопределенности.

ВВЕДЕНИЕ

В информационно-коммуникационных системах и сетях важное место занимают методы достоверного возобновления кодовых конструкций информационных потоков. При решении таких задач должны выполняться количественные ограничения на время обработки и передачу данных, на обеспечение достоверности информации и т.д. Современные методы направлены на использование статистических решающих правил при декодировании данных на основании накопления минимально достаточного количества информации. Процедуры принятия решения должны обеспечивать заданную вероятность идентификации полной кодовой конструкции на основе увеличения достоверности, а также достоверности переданных данных. Однако увеличение вероятности возобновления полной кодовой конструкции (при сокращении времени на обработку информационного потока и его декодирование) возможно при организации процедуры определения и дальнейшем использовании наиболее информативных параметров переданного сигнала.

Данный подход дает возможность не только возобновить полную кодовую последовательность при отмеченных ограничениях, но и сформировать последовательные правила для принятия окончательного решения.

Цель настоящей статьи — разработка нового метода определения информативных составляющих сигнала при введении заданной функции неопределенности на основании построения последовательного правила принятия решения. На основе проведенных исследований необходимо определить возможность остановки процесса декодирования в случае соответствия полученных результатов установленным порогам принятия решения исходя из накопления наиболее информативных составляющих.

На базе математического моделирования разработанного статистического метода в процессе работы сформируем графические зависимости апостериорной вероятности правильной идентификации полной кодовой конструкции от увеличивающейся меры количества информации при решении многоальтернативных задач. Покажем, какую меру количества информации вносит каждая информативная составляющая спектра. Для оценки качества построенной процедуры принятия решения на базе использования информативных частот сформируем зависимость правильной идентификации кодовой конструкции от соотношения сигнал/шум.

РАСЧЕТ МЕРЫ НЕОПРЕДЕЛЕННОСТИ ДЛЯ ИДЕНТИФИКАЦИИ КОДОВОЙ КОНСТРУКЦИИ

В теории идентификации при построении статистических правил на базе достаточной меры количества информации широко используются методы, которые вводят в процедуру формирование правила — функцию неопределенности $f(x)$. Исходя из условий отсутствия априорной информации о входном сигнале функция неопределенности должна обобщить процедуру и непосредственно правило при уменьшении или полном устранении информационной неопределенности. Следует отметить, что в теории идентификации существуют разные типы функций $f(x)$. В качестве основных типов функций неопределенности в современной теории принятия решений выступают квадратичная $f(x)=ax^2+bx+c$ (a, b, c — постоянные коэффициенты, $x\in[0,1]$) и экспоненциальная $f(x)=xe^{\alpha x+\beta}$ $(\alpha\le 0)$ функции, которые при соответствующих условиях принимают участие при построении мер неопределенности $\Psi(x)=\varphi(f(x)/x)$ и обобщенных мер количества информации $I_k(x)=1-\Psi(x)$ на уровне решений Котельникова, Шеннона, Кульбака, Байеса, Фишера. Расчет обобщенной меры количества информации сводится к трансформации априорного распределения $p_k(x)$ в апостериорное $q_k(x)$. Процесс получения информации $I_k(x)$ должен привести к уменьшению или полному снятию неопределенности $\Psi(x)$. Количество информации можно представить как неопределенность, которую возможно снять при трансформации априорного распределения. Тогда если обобщенную меру неопределенности по априорным вероятностям ввести таким образом, что $\Psi(p_k)=\varphi(f(p_k)/p_k)$, а по апостериорным вероятностям — согласно выражению $\Psi(q_k)=\varphi(f(q_k)/q_k)$, то мерой количества информации можно считать выражение $I_k(x)=1-\Psi(q_k)$ [1, 4–6].

Из изложенного следует, что задача определения меры количества информации сводится к нахождению меры неопределенности, которая, в свою очередь, находится выбором функции неопределенности. Эта функция используется при построении математической модели процедуры принятия решения. Вопрос сводится к определению условий, которые накладываются на нее: функция должна монотонно увеличиваться по мере увеличения количества задействованных в процедуре возможных состояний, т.е. должна увеличиваться апостериорная вероятность правильной идентификации на основании последовательного накопления итераций обработки информативных параметров сигнальной функции.

При разработке статистической процедуры принятия решения в задачах возобновления кодовой конструкции (кодовое слово в виде сформированной пачки последовательности радиоимпульсов) необходимо ввести функцию неопределенности $f(x)$, которая удовлетворяет условиям обеспечения накопления достаточной меры количества информации $I_k(x)=1-\Psi(q_k)$ по достижении заданной вероятности правильной идентификации. Данная задача есть многоальтернативной и приобретает наибольший интерес при появлении в пространстве слаборазличимых кодовых слов. Слаборазличимыми альтернативными являются гипотезы, которые имеют слаборазличимые параметры кодовых конструкций, т.е. отличаются на нуль, одну, две или три единицы Хеммингова расстояния [2]. Под кодовыми конструкциями будем понимать информационные сигналы, представленные кодовым словом, которое сформировано на базе методов помехоустойчивого канального кодирования в виде последовательности радио- или видеоимпульсов. Наиболее часто цифровые сигналы представляются, как 32-–64-битовая архитектура, а количество гипотез при этом увеличивается к $N \to 2048$. Под информативными параметрами принятого сигнала будем понимать спектральное представление последовательности кодовых слов. Данный вид представления параметров сигнала является наиболее информативным для формирования порогов принятия решения на базе минимально-достаточного количества информации. Для спектрального представления сигналов используем прямое преобразование Фурье, т.е. данный сигнал будет представляться в виде суммы гармонических колебаний с разными частотами.

Далее покажем зависимость меры количества информации от выбранного типа функции неопределенности. На основе проведенных исследований докажем следующее: для построения последовательной процедуры принятия решения на основе минимально достаточного количества информации и расчета апостериорной вероятности правильной идентификации довольно использовать информативные частоты спектра.

Предложенная задача состоит в том, чтобы найти функцию неопределенности и рассчитать на ее основе обобщенную меру количества информации для решения задач идентификации слаборазличимой битовой архитектуры. Введем как функцию неопределенности $f(x)$ для решения задачи идентификации слаборазличимых гипотез соответствующее спектральное представление принятого информационного сигнала $S_i(\omega_j)$. Проанализируем и покажем, что использование спектрального представления кодового слова как функции неопределенности $f(x) \Leftrightarrow S_i(\omega_j)$ будет наиболее приемлемым для построения математической процедуры принятия решения на основании увеличения обобщенной меры количества информации. Введенная функция неопределенности предназначена для уменьшения информационной неопределенности и обеспечения достаточной меры количества информации. Данная функция в дальнейшем будет использоваться при построении математической процедуры принятия решения, т.е. расчета меры неопределенности, меры количества информации и апостериорной вероятности правильной идентификации кодовой конструкции.

Запишем функцию неопределенности, как функцию соответствующего амплитудно-частотного спектра пачки видеоимпульсов:

$$f(x) \Leftrightarrow S_i(\omega_j) = A_m \tau \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot l \cdot \frac{T}{2}\right]}{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{T}{2}\right]} \right| \cdot \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]}{\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]} \right| =$$

$$= S(l, \omega_j) \cdot B(l, \omega_j), \qquad (1)$$

где T — период, τ — длительность импульсов, ω_j — основная частота дискретного спектра пачки ($j = 1, \ldots, N$ — текущий номер спектральных составляющих), $S(l, \omega_j)$ — функция амплитудно-частотного спектра одиночного импульса в пачке, l — количество импульсов в пачке, $B(l, \omega_j)$ — функция частоты, которая не зависит от формы импульсов и определяется лишь их числом и периодом следования [3].

РАЗРАБОТКА ПРОЦЕДУРЫ ПРИНЯТИЯ РЕШЕНИЯ НА БАЗЕ ИНФОРМАТИВНЫХ ЧАСТОТ

После проведенного анализа определим методы нахождения информативных частот (составляющих) спектра. Для построения последовательной процедуры принятия решения введем понятие ширины «частотных окон» Δf_i в задачах идентификации слаборазличимых кодовых слов. Под «частотным окном» понимаем эффективную ширину спектра сигнала из условий энергетического или

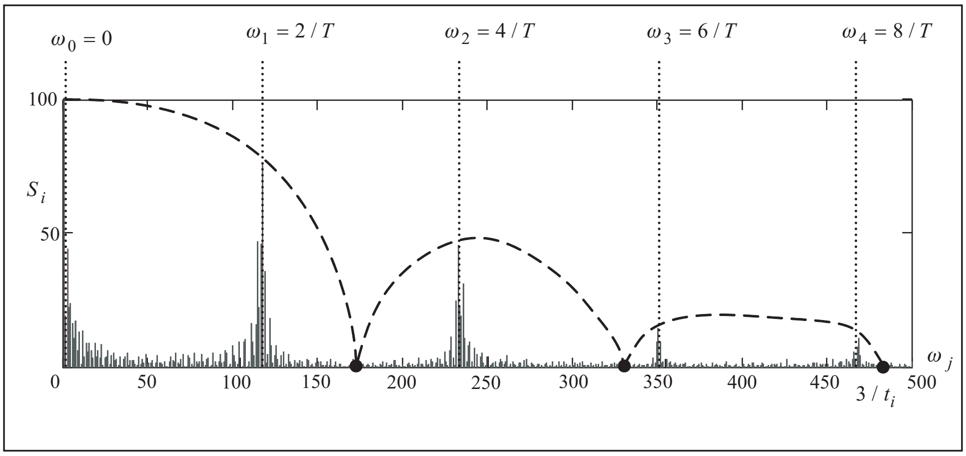

Рис. 1. График амплитудно-частотного спектра 32-битной кодовой конструкции

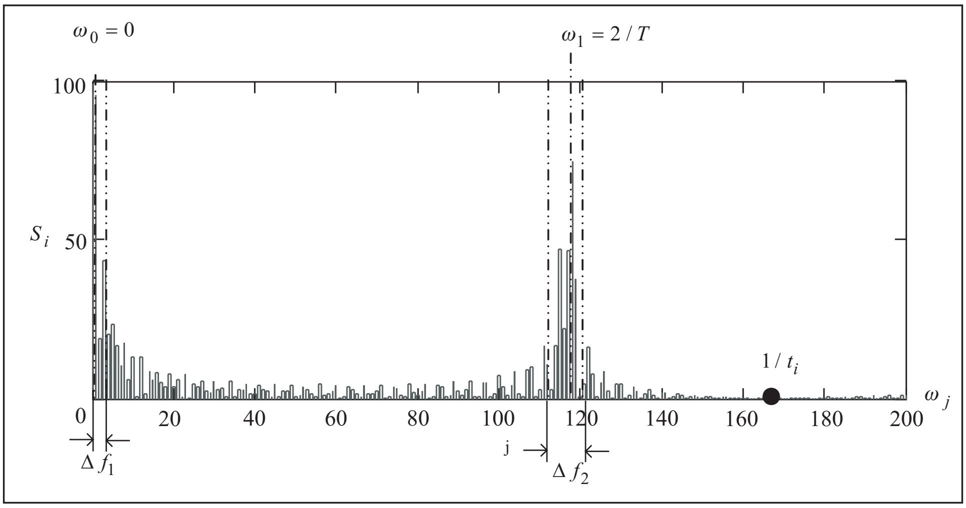

Рис. 2. Графическое отображение эффективной ширины спектра (частотного окна) с центрами в $\Delta\omega_j$

информационного вклада его гармоничных составляющих при расчете апостериорной вероятности правильной идентификации. Тогда суммарное значение окон, которые используются при идентификации, определяется выражением $\Delta F = \sum_{i=0}^{n} \Delta f_i$. Проведем графический и аналитический анализ процедуры формирования частотного окна, определим информативные составляющие спектра частот для l-битной ($l = 32$) кодовой конструкции со слаборазличимыми параметрами (рис. 1).

Из изложенного следует, что информативными частотами являются спектральные составляющие, которые формируются вблизи основных частот дискретного спектра кодовых конструкций $\Delta\omega_j = j/T$. Данные частоты являются центрами частотных окон $\Delta\omega_j = j/T$, $j = 0, 2, 4, \ldots, n$. Представим эффективную ширину одного частотного окна: $\Delta f_i = \omega_j \pm \frac{2}{l \cdot T} = \frac{j}{T} \pm \frac{2}{l \cdot T} = \frac{j \cdot l \pm 2}{l \cdot T}$.

Далее будем рассматривать только три частотных окна, которые концентрируют основную энергию сигнала. Эффективная ширина одного окна, где находится основная частота дискретизации $\Delta\omega_j$, графически отображена на рис. 2.

Тогда, используя введенную функцию неопределенности (1), построим обобщенную меру неопределенности, на ее основе рассчитаем минимально достаточное количество информации $I_k(x)$, которое используется при формировании статистического правила принятия решения и рассчитывается для каждой гипотезы $\{H_k\}$:

$$I_k[P(H_k / S_i(\omega_j))] = 1 - \Psi[P(H_k / S_i(\omega_j))] =$$

$$= 1 - \varphi\left(\frac{f[P(H_k / S(l, \omega_j) \cdot B(l, \omega_j))]}{P(H_k / S(l, \omega_j) \cdot B(l, \omega_j))}\right). \quad (2)$$

Нахождение условной плотности распределения определится формулой

$$W[S_i(\omega_j) / H_k] = \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_k}} \exp\left\{-\frac{(H_k / S(l, \omega_j) \cdot B(l, \omega_j) - m_{ij})^2}{2\sigma_k^2}\right\}. \quad (3)$$

Тогда общее байесовское выражение для нахождения апостериорной вероятности правильной идентификации информационного сигнала при условии введения новой функции неопределенности примет вид

$$P[H_k / S_i(\omega_j)] = \frac{\prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_k}} \exp\left\{-\frac{(S(l, \omega_j) \cdot B(l, \omega_j) - m_{ij})^2}{2\sigma_k^2}\right\}}{\sum_{i=1}^{N} \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left\{-\frac{(S(l, \omega_j) \cdot B(l, \omega_j) - m_{ij})^2}{2\sigma_i^2}\right\}}. \quad (4)$$

Авторами предложен метод определения информативных частот спектра на базе введения рассмотренной функции неопределенности в расчете меры количества информации с учетом присутствия в кодовых последовательностях N слаборазличимых кодовых слов.

На основе проведенных исследований построим график роста вероятности правильной идентификации в зависимости от количества использованных информативных спектральных составляющих при многоальтернативных задачах [2] (рис. 3). Очевидно, что каждая спектральная составляющая вносит определенную меру количества информации $I_k[P(H_k / S_i(\omega_j))]$, таким образом увеличивая апостериорную вероятность правильной идентификации $P(H_k / S_i(\omega_j)$ кодовой конструкции:

$$I_k\left\{\frac{\prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_k}} e^d}{\sum_{i=1}^{N} \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i}} \cdot e^f}\right\} > V_{\min}, \quad (5)$$

где

$$d = -\frac{\left(A_m \tau \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot l \cdot \frac{T}{2}\right]}{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{T}{2}\right]} \right| \cdot \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]}{\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]} \right| - m_{ij}\right)^2}{2\sigma_k^2},$$

$$f = -\frac{\left(A_m \tau \left| \frac{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot l \cdot \frac{T}{2}\right]}{\sin\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{T}{2}\right]} \right| \cdot \left| \frac{\sin\left[\left((\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]}{\left[\left(\omega_j \pm \frac{2}{l \cdot T}\right) \cdot \frac{\tau}{2}\right]} \right| - m_{ij}\right)^2}{2\sigma_i^2},$$

$V_{\min}$ — порог принятия решения.

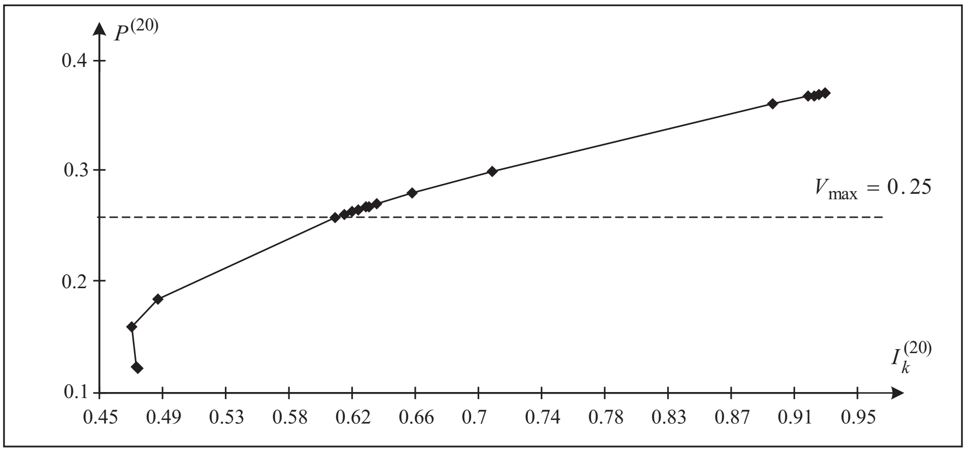

Рис. 3. Зависимость вероятности правильной идентификации от обобщенной меры количества информации при использовании 20 гармонических составляющих

На основании полученных результатов можно констатировать, что 20 гармонических составляющих вносят такую же меру количества информации, как и вся выборка, состоящая из 2049 гармонических составляющих. Апостериорная вероятность правильной идентификации при использовании всей выборки принимает значение $P^{(2049)}(H_k / S_i(\omega_j)) = 0.39$, а при использовании информативных составляющих трех спектральных окон — значение $P^{(20)}(H_k / S_i(\omega_j)) = 0.37$.

Не большая, но достаточная вероятность порогов правильной идентификации соответствует установленным условиям возобновления слаборазличимых кодовых слов. При решении стандартной задачи декодирования информационного потока данных пороги будут принимать стандартные значения на уровне решений Котельникова, Шеннона, Кульбака, Байеса и Фишера [2]. На основе полученных выражений (3)–(5) построим графическую зависимость вероятности правильной идентификации от растущей меры количества информации при использовании 20 гармонических составляющих (см. рис. 3).

Для оценки качества нововведенного метода построим график зависимости вероятности правильной идентификации от соотношения сигнал/шум: $b = \sigma_s / \sigma_n$. Сравним вероятности правильной идентификации при использовании в процедуре всех составляющих спектра и трех частотных окон с 20 информативными частотами. Графики построены с учетом условий идентификации 30 гипотез (кодовых слов) при использовании 20 наиболее информативных составляющих частотного спектра (рис. 4).

Основываясь на проведенных исследованиях, можно сделать вывод: трех частотных окон достаточно для построения эффективной процедуры принятия решения на базе накопления достаточного количества информации. Исходя из этого определим возможность остановки последовательной процедуры декодирования на основе обеспечения достаточного количества информации. Таким образом, происходит увеличение скорости идентификации полной кодовой конструкции при заданной достоверности. В результате использования изложенного метода уменьшается время, затраченное на анализ и обработку всего спектра, повышается эффективность и результативность разработанной последовательной процедуры принятия решения за счет использования только информативных составляющих частотного спектра.

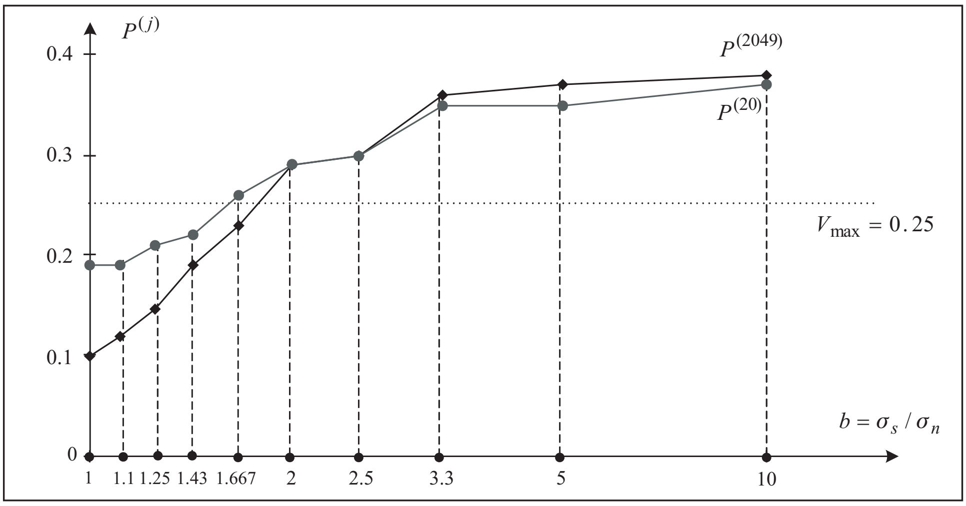

Рис. 4. Графики зависимости вероятности правильной идентификации от соотношения сигнал/шум (30 гипотез)

ЗАКЛЮЧЕНИЕ

Разработан новый метод определения информативных составляющих сигнала при введении заданной функции неопределенности исходя из условий построения последовательного правила принятия решения.

В результате проведенных исследований определена возможность остановки процесса декодирования в случае соответствия полученных результатов установленным порогам принятия решения на основе использования наиболее информативных составляющих спектра кодовой конструкции.

На базе математического моделирования разработанного статистического метода построены графические зависимости апостериорной вероятности правильной идентификации полной кодовой конструкции от увеличивающейся меры количества информации. Определена оценка эффективности изложенной процедуры принятия решения на основе использования информативных частот в зависимости от соотношения сигнал/шум.

СПИСОК ЛИТЕРАТУРЫ

1. Ferguson T.S. Mathematical statistics: A decision theoretic approach. — New York: Academic Press, 1967. — 396 p.
2. Grendár M. Jr., Grendár M. Maximum probability and maximum entropy methods: Bayesian interpretation // Bayesian inference and maximum entropy methods in science and engineering / G. Erickson and Y. Zhai (Eds.). — Melville, 2004. — P. 490–494.
3. Gwo-Hshiung Tzeng, Jih-Jeng Huang. Multiple attribute decision making: methods and applications . — Boca Raton; London; New York: CRC Press, Taylor & Francis Group, 2011. — 340 p.
4. Rice J.A. Mathematical statistics and data analysis. — Belmont: Thomson Books, 2010. — 550 p.
5. Воронин А.Н. Метод многокритериальной оценки и оптимизации иерархических систем // Кибернетика и системный анализ. — 2007. — № 3. — С. 84–92.
6. Юдін О.К., Чунарьова А.В., Луцький М.Г. Підвищення достовірності відновлення інформаційних потоків // Наукоємні технології. — 2010. — Вип. 4. — С. 84–88.

Поступила 11.06.2015