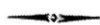


MIESIĘCZNIK
POŁOSKI.

Т о м I.
Р о к 1818.

*«Вестник Полоцкого государственного университета»
продолжает традиции первого в Беларуси литературно-
научного журнала «Месячник Полоцкий».*



ВЕСНІК ПОЛАЦКАГА ДЗЯРЖАУНАГА УНІВЕРСІТЭТА
Серыя С. Фундаментальныя навукі

У серыі С навукова-тэарэтычнага часопіса друкуюцца артыкулы, якія прайшлі рэцэнзаваўне, змяшчаюць новыя навуковыя вынікі ў галіне інфарматыкі і матэматыкі, фізікі.

ВЕСТНИК ПОЛОЦКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
Серия С. Фундаментальные науки

В серии С научно-теоретического журнала публикуются статьи, прошедшие рецензирование, содержащие новые научные результаты в области информатики и математики, физики.

HERALD OF POLOTSK STATE UNIVERSITY
Series C. Fundamental sciences

Series C includes reviewed articles which contain novelty in research and its results in computer science and mathematics, physics.

Адрес редакции:

Полоцкий государственный университет, ул. Блохина, 29, г. Новополоцк, 211440, Беларусь
тел. + 375 (214) 53 34 58, e-mail: vestnik@psu.by

Отв. за выпуск В.А. Груздев, С.Г. Ехилевский, Э.М. Пальчик.

Редакторы Р.Н. Авласенок, О.В. Мальцева. Техн. редактор Р.Н. Авласенок.

Подписано к печати 27.09.2013. Бумага офсетная 70 г/м². Формат 60×84¹/₈. Ризография.

Усл. печ. л. 13,02. Уч.-изд. л. 15,69. Тираж 100 экз. Заказ 1612.

ИНФОРМАТИКА

УДК 621.391.8

МАСКИРОВАНИЕ RGB-ВИДЕОКАДРОВ СИНХРОННЫМ И АДАПТИВНЫМ ШУМОВЫМ RGB-ВИДЕОКАДРОМ

д-р техн. наук, проф. В.К. ЖЕЛЕЗНЯК, А.В. БАРКОВ
(Полоцкий государственный университет)

Предложен метод формирования маскирующей помехи для защиты видеосигнала от утечки по техническим каналам. Видеосигнал обладает рядом характерных особенностей, которые надо учитывать при решении задачи защиты от утечки по техническим каналам. Исследования показали возможность синхронного накопления зашумленного статического видеокadra, которое значительно улучшает отношение сигнал/шум. В этой связи обоснована необходимость создания синхронной адаптивной помехи для маскирования статических видеокadров. Предложен способ формирования маскирующих видеокadров с учетом того, что видеокadры на экране могут быть статическими (неподвижными) и динамическими (подвижными), содержат крупноплановые и мелкодетальные элементы.

Введение. Работа средств вычислительной техники сопровождается излучениями, которые являются источниками информационного сигнала и способствуют образованию каналов утечки информации. Источниками образования каналов утечки могут быть дисплеи, накопители, принтеры, плоттеры, каналообразующая аппаратура и др. В составе персонального компьютера имеется специальная видеосистема, предназначенная для формирования изображений, наблюдаемых на экране монитора. Ее основу составляют специализированные схемы для генерирования электрических сигналов, управляющих монитором, – видеоадаптер, который формирует сигналы, управляющие той информацией, которая выводится на экран монитора, видеоадаптер также может служить источником утечки видеосигнала. Однако основным источником высокочастотного электромагнитного излучения является дисплей, в особенности созданный на базе электронно-лучевой трубки.

Видеосигнал при передаче изображений неподвижных объектов обладает характерным свойством – он периодичен. При достаточно длительной передаче неподвижного изображения видеосигнал может считаться периодическим с частотой видеокadров $f_{\text{кадр}}$. Периодичность видеосигнала позволяет проводить его синхронное накопление с целью улучшения отношения сигнал/шум (ОСШ). Время накопления зашумленного видеокadra устанавливается выигрышем по отношению количества накопленных видеокadров к накопленным шумовым кадрам пропорционально \sqrt{n} , где n – количество накоплений, что соответствует данным исследования авторов и исследованиям в University of Cambridge Computer Laboratory автором Dr Markus Kuhn «Security Limits for Compromising Emanations» [1].

Возможность восстановления информации видеодисплейного модуля [2, с. 121; 3 – 5] за счет сигнала в канале утечки обуславливает необходимость развития средств защиты объектов информатизации.

Маскирование видеосигнала требует учета специфики его структуры, ему присущи следующие свойства: видеосигнал синхронный, обусловлен наличием синхроимпульсов; широкая полоса частот; зависимость сложности обнаружения и выделения от свойств самого видеоизображения. Видеоизображение на экране может быть статическим (неподвижным) и динамическим (подвижным), содержать крупноплановые и мелкодетальные элементы. Известные маскирующие генераторы в основной массе представляют генераторы динамического белого шума и не учитывают синхронность данных, не дают возможность формирования синхронных с маскируемым сигналом статических маскирующих помех для достижения положительного технического эффекта.

В связи с возможностью восстановления видеосигнала в канале утечки и специфики его структуры целью данной работы является формирование маскирующей помехи видеосигнала; задачей – разработать и предложить метод формирования маскирующей помехи с учетом особенностей видеосигнала.

Поставленная задача решается формированием активной маскирующей помехи. Активными помехами называются радиосигналы, создаваемые специальными радиопередатчиками и предназначенные для ухудшения или исключения нормальной работы радиоэлектронных средств. Активные маскирующие помехи создают на входе приемника подавляемого радиоэлектронного средства фон, который затрудняет обнаружение полезных сигналов, их распознавание и определение параметров. Как правило, маскирующие помехи линейно суммируются с сигналом на входе приемника и поэтому называются аддитивными. В задаче защиты видеосигнала помехи подаются в канал утечки и суммируются с видеосигналом.

Предложен способ формирования маскирующих видеокадров с учетом того, что видеокадры на экране могут быть статическими (неподвижными) и динамическими (подвижными), содержать крупноплановые и мелкодетальные элементы.

Восстановление видеосигнала синхронным накоплением. Ряд сигналов (например, видео, цифровой) являются синхронными. При наличии информации о периоде сигнала есть возможность его синхронного накопления (периодического усреднения) и улучшения отношения сигнал/шум (ОСШ) [6, с. 70 – 78]. Накопление заключается в суммировании выборочных значений (импульсов) наблюдаемого колебания, которое в общем случае представляет смесь сигнала и шума [7, с. 157].

При обработке (синхронном накоплении) статических видеокадров их амплитуды складываются синхронно по линейному закону, т.е. при накоплении N кадров амплитуда видеосигнала увеличивается в n раз. При несинхронном накоплении шумов, не сформированных в шумовой видеокадр, их амплитуда складывается по среднеквадратичному закону, т.е. при аналогичном времени, соответствующем накоплению n видеокадров, амплитуда шумового маскирующего сигнала увеличится в \sqrt{n} раз [1].

Количество накопленных видеокадров зависит от ряда факторов [8, с. 96]: когда содержимое экрана остается стабильным в течение периода времени T , то может быть получено до fvT кадров (fv – частота кадров, или вертикальное отклонение частоты экрана); периодическое усреднение видеосигнала может быть успешным, только если частота fv определена с высокой точностью. Накопление видеосигнала возможно при наличии данных синхронизации.

Синхронизацию восстанавливают из принятого сигнала. Современное развитие видеосистем отображения информации, в отличие от ранее разработанных, в которых частота синхроимпульсов привязывалась к частоте сети, основано на том, что частота синхроимпульсов задается кварцевым генератором.

В современных ПЭВМ развертка монитора содержит несколько частот, например, некоторые режимы работы стандарта VGA: 640×480 60 Гц; 640×480 72 Гц; 640×480 75 Гц; 800×600 60 Гц; 800×600 75 Гц.

Для восстановления синхронизации видеосигнала авторами предложен метод восстановления синхроимпульсов с априорно неизвестными параметрами в каналах утечки информации и оценки неизвестных их параметров [9 – 10]. Восстановление синхроимпульсов позволяет проводить синхронное накопление видеосигнала.

Для неподвижного (статического) изображения на экране время накопления видеокадра T (с), находящегося в статике, равно количеству накопленных кадров, то есть

$$T = nt_{\text{кадр}}, \quad (1)$$

где n – количество накопленных кадров; $t_{\text{кадр}}$ – длительность одного кадра видеосигнала (с).

Выигрыш определяют отношением сигнала накопленных синхронных статических кадров к накопленным несинхронным случайным шумам в кадрах $n/\sqrt{n} = \sqrt{n}$.

Покажем синхронное накопление на примере зашумленного видеосигнала. Видеосигнал тестового черно-белого изображения (рис. 1, а) длительностью 30 с динамически зашумлен белым гауссовым шумом. Производим накопление зашумленных видеокадров (рис. 1, б) и нормируем итоговый видеокадр (рис. 1, в).

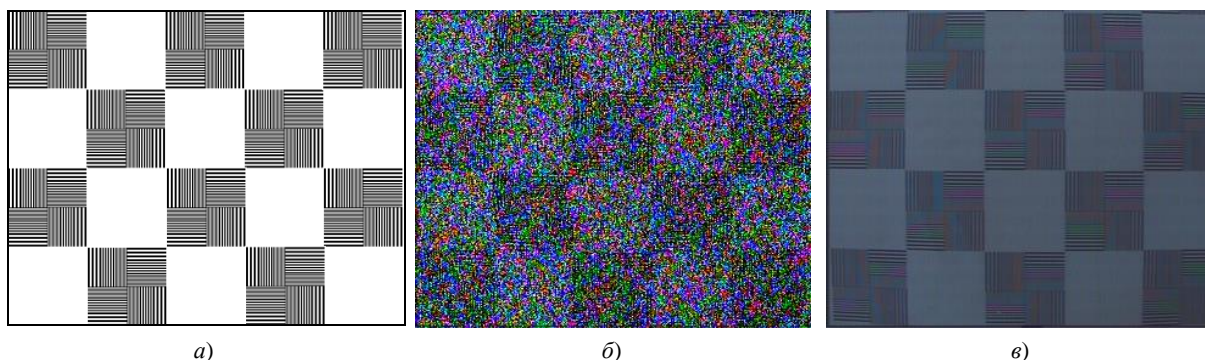


Рис. 1. Кадр тестового черно-белого изображения:

- а – исходный тестовый кадр (шахматное поле с горизонтальным и вертикальным заполнением линиями);
- б – зашумленный динамическим шумом кадр тестового (а) изображения (длительность видеосигнала 30 с);
- в – накопленный видеокадр (б) видеосигнала

В данном примере сигнал помехи представляет собой динамически изменяющийся белый шум, который различен для каждого видеокадра, изображение на экране монитора находилось в статике, т.е. было неподвижным.

Зашумленный видеокادر (см. рис. 1, б) показывает, что крупноплановые элементы изображения – клетки шахматного поля – могут быть различимы на отдельном кадре; мелкодетальные элементы – горизонтальные и вертикальные линии различной толщины внутри клеток – не различимы. После накопления видеосигнала длительностью 30 с статического изображения с динамическим шумом получаем четкое изображение (см. рис 1, в), в котором восстановлены мелкодетальные элементы, заполнение клеток четко просматривается.

Статическое изображение, наблюдаемое на экранах видеодисплейных модулей, исключает выделение сигнала видеокадра из шумов только при придании статических свойств видеосигнала маскирующей помехе. Синхронное накопление не позволит получить требуемое улучшение ОСШ, если помеха будет накапливаться по тому же закону, что и видеосигнал, что требует решения задачи создания синхронной адаптивной помехи в виде шумового RGB-видеокадра.

Формирование тестового видеокадра. Для определения эффективности маскирования видеосигнала необходимо оценить изображение, полученное при обработке и восстановлении видеосигнала в канале утечки [11] с использованием тестового сигнала, представляющего последовательность черных и белых горизонтальных полос на экране видеодисплейного модуля (гармонический сигнал), что не позволяет определить тонкую структуру видеокадра, определяемую крупноплановыми и мелкодетальными элементами. Это обосновывает необходимость формирования тестового видеокадра.

Из анализа видеосигналов [11] для контроля каналов утечки предложен тестовый видеокادر черно-белого (рис. 1, а) и цветного (рис. 2, а) изображения в виде шахматного поля, шахматные клетки (на основе элемента изображения программы «Nokia Monitor Test») включают горизонтальные и вертикальные линии различной толщины. В отличие от известных [11], изображение в виде шахматного поля позволяет определить тонкую структуру видеокадра, определяемую крупноплановыми и мелкодетальными элементами, а также учитывает цветовые компоненты RGB-видеосигнала.

Маскирование RGB-видеокадров синхронным и адаптивным шумовым RGB-видеокадром. Качественное маскирование видеосигналов активными помехами требует создания широкополосных шумовых сигналов, используемых, например, в качестве генераторов маскирующих помех, адаптируемых к свойствам маскируемых сигналов. Маскирующие активные помехи должны с максимальной эффективностью [12, с. 14 – 18] затруднять обнаружение сигналов при минимальных энергетических ресурсах, чтобы не нарушать требования по электромагнитной совместимости. Видеосигналы являются синхронными. Это позволяет их накапливать. Маскирующие их шумы таким свойством не обладают [13, с. 55 – 58].

Известен ряд генераторов шума, из которых, например, «Гном-3» – генератор шума, стационарный для защиты помещений и объектов электронно-вычислительной техники от утечки конфиденциальной информации за счет побочных электромагнитных излучений компьютеров и другой оргтехники. Генератор «Гном-3» и аналогичные ему не учитывают синхронность данных и не дают возможность формирования синхронных с маскируемым сигналом статических маскирующих помех для достижения положительного технического эффекта.

Ограничения существующих способов обусловили задачу: сформировать требования к маскирующим шумам, которые, так же как и видеосигнал, должны обладать свойством синхронности.

Предложен метод маскирования статических видеосигналов синхронным и адаптивным шумовым RGB-видеокадром, заключающийся в формировании синхронных с видеосигналами статических шумовых видеокадров, учитывающими особенности синхронности статического видеокадра. Время накопления видеокадра, находящегося в статике, равно количеству накопленных кадров (1). Выигрыш, определяемый отношением сигнала накопленных синхронных статических кадров к сигналу накопленных несинхронных шумов, равен $n/\sqrt{n} = \sqrt{n}$. Накопление видеосигнала, зашумленного динамическим шумом, улучшает ОСШ по сравнению с зашумленным синхронными и адаптивными шумовыми RGB-видеокадрами пропорционально \sqrt{n}/\sqrt{k} (n – число видеокадров, k – количество смен шумового видеокадра).

Проведены экспериментальные исследования синхронного накопления зашумленного динамическим шумом и статическим шумовым RGB-кадром видеосигнала.

Исследовано маскирование цветного и черно-белого тестового изображения.

Результаты эксперимента по синхронному накоплению видеосигнала длительностью 30 с динамически зашумленного тестового цветного изображения (рис. 2, а) показывают восстановление RGB-видеокадра, где различимы крупноплановые элементы в виде клеток и мелкодетальные линии внутри клеток поля (рис. 2, б), а также цветовые компоненты видеокадра.

Результатом синхронного накопления статических видеокадров, синхронно зашумленных адаптивным шумовым RGB-видеокадром (рис. 2, в), является ухудшение восстановления изображения статических видеокадров.

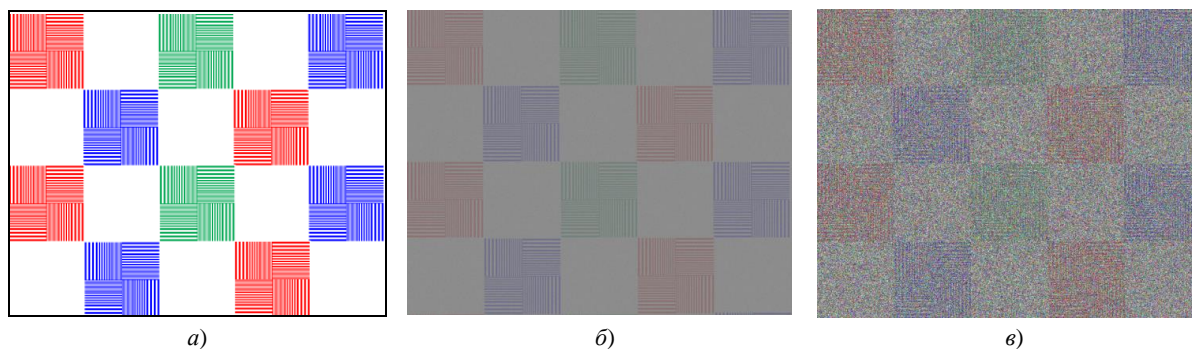


Рис. 2. Кадр тестового цветного изображения:

- a* – исходный тестовый кадр (шахматное поле с горизонтальным и вертикальным заполнением линиями);
b – накопленный видеокادر видеосигнала (*a*), зашумленного динамическим шумом;
в – накопленный кадр цветного тестового изображения, зашумленного статическим адаптивным шумовым RGB-видеокадром

Предложена схема реализации устройства формирования маскирующей помехи в виде шумового RGB-видеокадра для защиты видеосигнала от утечки по техническим каналам [14], которая включает:

- формирователь маскирующей помехи (содержит два идентичных формирователя накопленного шумового видеокadra);
- синхронный переключатель;
- формирователь хаотической импульсной последовательности;
- формирователь шумового RGB-видеокадра;
- генератор синхросмеси (может работать в автономном режиме либо в режиме внешнего запуска от внешнего синхронизатора);
- генератор синхронизации;
- внешний синхрогенератор.

Особенность маскирующих помех заключается в том, что формируется пиксельный (точечный) шумовой RGB-видеокадр, подавляющий мелкодетальные элементы изображения. Контурные крупноплановых изображений разрушаются низкочастотными составляющими хаотических импульсных последовательностей, являющихся элементами шумовых RGB-видеокадров.

Вывод

Предложен и обоснован метод маскирования статических и динамических RGB-видеокадров, который учитывает особенности видеосигнала:

- 1) шумовой RGB-видеокадр является синхронным и адаптивным с видеосигналом, что обеспечивает лучшее качество маскирования и исключает ухудшение ОСШ по сравнению с динамическим пропорционально \sqrt{n}/\sqrt{k} ;
- 2) предложено тестовое изображение, которое учитывает тонкую структуру видеокadra, включающую крупноплановые и мелкодетальные элементы;
- 3) экспериментальные исследования подтверждены разработанным алгоритмом синхронного накопления видеокadров, демонстрируют преимущества маскирования статического RGB-видеокадра синхронным шумовым RGB-кадром, который по сравнению с динамическим шумом затрудняет выделение видеокadra, уменьшает его разборчивость и повышает защищенность видеосигнала пропорционально \sqrt{n}/\sqrt{k} .

ЛИТЕРАТУРА

1. Markus, G. Kuhn Security Limits for Compromising Emanations / G. Markus [Electronic resource] Computer Laboratory, University of Cambridge, 15 JJ Thomson Avenue, Cambridge CB3 0FD. – Mode of access: <http://www.cl.cam.ac.uk/~mgk25/ches2005-limits.pdf>. – Date of access: 02.09.2013.
2. Петраков, А.В. Основы практической защиты информации / А.В. Петраков. – М.: Радио и связь, 1999. – 368 с.
3. Wim van Eck: Electromagnetic Radiation from Video Display Units: An Eaves-dropping Risk? Computers & Security. – 1985. – Vol. 4. – P. 269 – 286.
4. Tanaka, H. A trial of the interception of display image using emanation of electromagnetic wave / H. Tanaka, O. Takizawa, A. Yamamura // Journal of the National Institute of Information and Communications Technology. – 2005. – № 52(1/2). – P. 213 – 223.

5. Markus, G. Kuhn Eavesdropping attacks on computer displays / G. Markus [Electronic resource] // Computer Laboratory, University of Cambridge, 15 JJ Thomson Avenue, Cambridge CB3 0FD. – Mode of access: <http://www.cl.cam.ac.uk/~mgk25/iss2006-tempest.pdf>. – Date of access: 02.09.2013.
6. Харкевич, А.А. Борьба с помехами / А.А. Харкевич. – изд. 2-е, исправл. – М.: Наука. Гл. ред. физ-мат. лит., 1965. – 275 с.
7. Денисенко, А.Н. Статистическая теория радиотехнических систем / А.Н. Денисенко. – М.: АРИ, 2007. – 200 с.
8. Markus, G. Kuhn: Compromising emanations: eavesdropping risks of computer displays / G. Markus. Chapter 5: Emission limits. Technical Report UCAM-CL-TR-577, University of Cambridge, Computer Laboratory, December 2003.
9. Способ обнаружения периодической импульсной последовательности и оценки ее периода: пат. 17138 Респ. Беларусь, МПК G 01R 23/02 / В.К. Железняк, А.В. Барков; заявитель Полоц. гос. ун-т. – № а20110815; заявл. 10.06.2011.
10. Барков, А.В. Методы разрушения и восстановления импульсов синхронизации в каналах утечки информации / А.В. Барков, В.К. Железняк // Технические средства защиты информации: тез. докл. XI белорус.-рос. науч.-техн. конф., Минск, 5 – 6 июня 2013 г.; редкол.: Л.М. Лыньков (отв. ред.) [и др.]. – Минск: БГУИР, 2013. – С. 29 – 30.
11. Исследование алгоритма обнаружения побочного электромагнитного излучения компьютеров: материал 4-й междунар. конф. и выставки «Цифровая обработка сигналов и ее применения», Москва, февр. 2002 г. – М., 2002. – С. 326 – 329.
12. Помехоустойчивость и эффективность систем передачи информации / А.Г. Зюко [и др.]; под ред. А.Г. Зюко. – М.: Радио и связь, 1985. – 272 с.
13. Железняк, В.К. Защита информации от утечки по техническим каналам: учеб. пособие / В.К. Железняк. – СПб.: ГУАП, 2006. – 155 с.
14. Барков, А.В. Формирование маскирующей помехи для защиты видеосигнала от утечки по техническим каналам периода / А.В. Барков, В.К. Железняк // Вестн. Полоц. гос. ун-та. Серия С. Фундаментальные науки. – 2012. – № 12. – С. 40 – 45.

Поступила 04.09.2013

MASKING OF RGB-VIDEO FRAMES BY SYNCHRONOUS AND ADAPTIVE NOISE RGB-VIDEO FRAME

V. ZHELEZNYAK, A. BARKOV

A method of forming a masking noise to protect the video signal from the drain through technical channels is proposed. The video signal has a number of characteristics that must be taken into account when solving the problem of leakage protection through technical channels. The aim is to form a masking noise signal. The objective is to develop and propose a method of forming a masking noise with the features of the video signal. Studies have shown the possibility of simultaneous accumulation of noisy static video frame, which significantly improves the signal / noise ratio (SNR). In this context, the necessity of creating a synchronous adaptive noise to mask the static video frames is grounded. A method of forming a mask taking into account that the motion pictures on the screen may be static (fixed) or dynamic (moving) comprise major and small detailed elements.

УДК 621.37.037

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ СВЕРХРАЗРЕШЕНИЯ ИСТОЧНИКОВ ИЗЛУЧЕНИЯ НА ОСНОВЕ ПРОЕКЦИОННОГО МЕТОДА

А.О. БЕЗДЕЛЬ, канд. техн. наук, доц. В.Ф. ЯНУШКЕВИЧ, В.М. ЧЕРТКОВ
(Полоцкий государственный университет)

Рассматриваются методы цифровой обработки сигналов, позволяющие получить лучшие характеристики и обеспечивающие высокое качество обработки радиолокационных сигналов. Представлены результаты теоретического анализа сверхразрешения источников излучения с использованием метода «теплового шума». Демонстрируются исследования для одного и двух источников излучения. Проанализированы функции углового разрешения объектов. Даны рекомендации по уменьшению вычислительной сложности алгоритмов обработки.

В задачах радиопеленгации в ряде случаев необходимо путём обработки исследуемого сигнала на конечном интервале наблюдения определить количество источников излучения (или переизлучения), образующих принятый и исследуемый сигнал, и оценить угловые координаты источников.

Актуальность исследования обусловлена тем, что к современному оборудованию предъявляют жесткие ограничения по массе, габаритным размерам и энергопотреблению. Налагаемые ограничения, а также исключительная сложность электромагнитной обстановки приводят к новой, требующей детального исследования задаче синтеза на базе «малозлементной» антенны алгоритмов цифровой обработки, обеспечивающих эффективные характеристики подавления различного рода помех и высокое качество обработки радиолокационных сигналов.

Теоретический анализ. С ростом степени n разрешающая способность адаптивной антенной решетки с выходной характеристикой $\eta_n(\varphi) = [S^H(\varphi)M^{-n}S(\varphi)]^{-1}$ монотонно возрастает, а сама разрешающая функция $\eta_n(\varphi)$ при неограниченном увеличении параметра n стремится к величине $[S^H(\varphi)DS(\varphi)]^{-1}$, где $D = \left(E - \sum_{k=1}^J U_k U_k^H \right)$ – матрица проектирования на шумовое подпространство.

Такой метод можно назвать проекционным и обозначить разрешающую функцию как $N - J$.

Чтобы воспользоваться этим методом, надо уметь оценивать размерность шумового подпространства и строить проекционную матрицу, используя входные сигналы. Оценивание размерности шумового подпространства равносильно оцениванию числа J внешних сигналов, принимаемых антенной решеткой, поскольку размерность шумового подпространства равно $N - J$. Нетрудно понять, что это обстоятельство принципиально отличает проекционный метод от методов Кейпона и «теплового шума», где размерность шумового подпространства не оценивается.

Пусть N -элементная антенная решетка произвольной геометрии принимает сигналы от J источников с угловыми координатами $\varphi_1, \varphi_2, \dots, \varphi_j$ и мощностями $\nu_1, \nu_2, \dots, \nu_j$, а также аддитивные собственные шумы приемных каналов. Проекционный метод можно использовать только в случае, когда число источников меньше числа приемных каналов антенной решетки ($J < N$). Будем также считать, что сигналы и аддитивные шумы представляют собой случайные, нормально распределенные комплексные величины с нулевым средним значением.

Рассмотрим вначале **идеальный случай**, когда число выборок бесконечно ($L \rightarrow \infty$), т.е. известна точная корреляционная матрица M входного процесса. Векторное N -мерное пространство разбивается на два подпространства: сигнальное и шумовое. Первое из них имеет размерность J и соответствует сигнальным собственным числам данной матрицы ($\mu_1 \geq \mu_2 \geq \dots \geq \mu_J \geq 1$). Второе подпространство является ортогональным первому, имеет размерность $N - J$ и соответствует шумовым собственным числам корреляционной матрицы ($\mu_{j+1} = \mu_{j+2} = \dots = \mu_N = 1$).

Корреляционную матрицу можно представить в виде разложения по собственным векторам U_k ($k = 1, 2, \dots, N$). Так как эта матрица имеет кратные собственные числа, то её удобнее представить в другом эквивалентном виде:

$$M = \sum_{k=1}^{N'} \mu_k P_k, \tag{1}$$

где P_k – матрица-проектор на собственное подпространство, соответствующее собственному числу μ_k ; N' – число неравных между собой собственных чисел ($N' \leq J + 1$).

Число неравных между собой собственных чисел N' равно степени минимального многочлена корреляционной матрицы M . Минимальным многочленом матрицы называют многочлен минимальной степени с коэффициентом при старшем члене, равном единице, который является делителем характеристического многочлена [1]. Это значит, что собственные числа матрицы являются простыми корнями минимального многочлена. Таким образом, минимальный многочлен можно записать в виде

$$\psi(\mu) = (\mu - \mu_1)(\mu - \mu_2) \dots (\mu - \mu_{N'}). \quad (2)$$

Минимальный многочлен матрицы обладает следующим интересным свойством. Он аннулирует собственную матрицу, т.е. матрица $\psi(M)$ является нулевой матрицей. Чтобы доказать этот факт, подставим в (2) в качестве аргумента матрицу M , и тогда мы получаем следующее выражение:

$$\psi(M) = (M - \mu_1 E)(M - \mu_2 E) \dots (M - \mu_{N'} E). \quad (3)$$

Первый сомножитель $(M - \mu_1 E)$ способен аннулировать (занулить) любые векторы первого собственного подпространства. Это значит, что любой вектор из этого собственного подпространства, будучи умноженным на матрицу $(M - \mu_1 E)$, дает нулевой вектор. Поскольку в (2) присутствуют все N' сомножителей, то матрица $\psi(M)$ аннулирует любой вектор N -мерного пространства. Это возможно только тогда, когда $\psi(M) = [0]$.

Если в (3) уменьшить число сомножителей до $N'' (N'' < N')$, получим ненулевую матрицу: $\psi_0(M) \neq [0]$. Эта матрица будет аннулировать любой вектор Φ_0 , который принадлежит совокупности выбранных собственных подпространств, т.е. $\psi_0(M)\Phi_0 = [0]$. В частности, если исключить из (2) только один i -й сомножитель, получим матрицу $\psi_{N'}^{(i)}(M)$, которая будет проектировать любой вектор на i -е собственное подпространство. Таким образом, можно построить входящие в (1) матрицы-проекторы $\Pi_k (k = 1, 2, \dots, N')$ в виде матричных многочленов:

$$\Pi_k = \left[\prod_{p=1, p \neq k}^{N'} (M - \mu_p E) \right] \left[\prod_{p=1, p \neq k}^{N'} (\mu_k - \mu_p) \right]^{-1}. \quad (4)$$

Отметим, что соответствующее k -му подпространству собственное число матрицы Π_k равно единице, а ее остальные $(N' - 1)$ собственных чисел равны нулю.

Построение матриц-проекторов $\Pi_k (k = 1, 2, \dots, N')$ в принципе решает поставленные выше задачи обработки сигналов. При этом важную роль играет матрица-проектор Π_{noise} на шумовое подпространство, которую можно получить из (4), полагая $k = N'$ и $\mu_{N'} = 1$.

В результате будем иметь

$$\Pi_{noise} = \left[\prod_{p=1}^{N'-1} (M - \mu_p E) \right] \left[\prod_{p=1, p \neq k}^{N'-1} (1 - \mu_p) \right]^{-1}. \quad (5)$$

В этом выражении матрица проектирования на шумовое подпространство выражена через корреляционную матрицу и ее собственные числа. Таким образом, мы избежали сложной с вычислительной точки зрения процедуры вычисления собственных векторов корреляционной матрицы.

Теперь, зная матрицу проектирования на шумовое подпространство, можно определить число источников по формуле

$$J = N - Sp(\Pi_{noise}), \quad (6)$$

где $Sp(\cdot)$ – след матрицы. Значение $Sp(\Pi_{noise})$ совпадает с размерностью шумового подпространства.

Направления на источники сигналов отождествляются согласно проекционному методу с максимумами функции:

$$\eta_z(\varphi) = [S^H(\varphi)\Pi_{noise}S(\varphi)]^{-1}. \quad (7)$$

Выражение в квадратных скобках (7) имеет физический смысл части выходной мощности антенной решетки, соответствующей подпространству собственного шума. Если текущий угол φ у вектора $S(\varphi)$ совпадает с направлением на любой источник сигнала $(\varphi_1, \varphi_2, \dots, \varphi_j)$, то этот вектор будет принадлежать сигнальному подпространству, а его проекция на шумовое подпространство будет равна нулю. Поэтому в этой точке функция $\eta_\infty(\varphi)$ имеет особенность типа 0^{-1} . По пикам функции $\eta_\infty(\varphi)$ определяются угловые положения на источники сигналов.

В случае одного источника сигналов ($J=1$) корреляционная матрица M имеет два собственных числа $\mu_1 > 1$ и $\mu_2 = 1$. Проектор Π_{noise} на шумовое подпространство и функция $\eta_\infty(\varphi)$ будут равны:

$$\Pi_{noise} = (M - \mu_1 E)(1 - \mu_1), \quad (8)$$

$$\eta_\infty(\varphi) = (1 - \mu_1) [S^H(\varphi)MS(\varphi) - \mu_1]^{-1}. \quad (9)$$

Здесь проекционный метод полностью эквивалентен согласованной обработке, которая является оптимальной и сводится к измерению выходной мощности $S^H(\varphi)MS(\varphi)$ при сканировании лучом антенны. При наличии двух и более источников сигналов ($J > 2$) метод согласованной обработки перестает быть оптимальным и отличается от проекционного метода, который дает точные значения параметров сигналов.

Рассмотрим теперь случай, когда число выборок входного процесса конечно и равно L . Тогда вместо точной корреляционной матрицы мы должны использовать ее максимально правдоподобную оценку $\hat{M} = \frac{1}{L} \sum_{j=1}^L Z_j Z_j^H$. Элементы оценочной корреляционной матрицы являются случайными величинами и имеют совместное распределение плотности вероятности Уишарта [2; 3].

В силу случайного характера вероятность появления кратных собственных чисел ничтожно мала. Поэтому при $L \geq N$ матрица \hat{M} имеет N случайных положительных собственных чисел $\mu_1 > \mu_2 > \dots > \mu_N > 0$. Шумовое собственное число, равное единице и имеющее кратность $N - J$ для точной корреляционной матрицы, расщепляется на $N - J$ простых собственных чисел оценочной корреляционной матрицы \hat{M} . Разброс шумовых собственных чисел возрастает с уменьшением числа выборок L . Они могут принимать значения, существенно меньшие единицы. При $L < N$ выборочная матрица \hat{M} является вырожденной и имеет L положительных собственных чисел, а количество $N - L$ ее собственных чисел равно нулю. Подпространство, соответствующее нулевым собственным числам, ортогонально подпространству выборочных векторов $Z(l)$ и называется ядром матрицы \hat{M} .

Теперь многочлен (2) состоит из N сомножителей и имеет вид

$$\psi(\mu) = (\mu - \mu_1)(\mu - \mu_2) \dots (\mu - \mu_N), \quad (10)$$

то есть является характеристическим многочленом корреляционной матрицы \hat{M} степени N . Отсюда следует, что $\psi(M) = [0]$, и линейно независимыми будут N степенных матриц $E, \hat{M}, \hat{M}^2, \dots, \hat{M}^{N-1}$.

Таким образом, при переходе от точной корреляционной матрицы M к ее оценке \hat{M} число линейно независимых матриц увеличивается от N' до N . Это обусловлено влиянием собственного шума и появлением множества шумовых собственных чисел вместо одного. Поэтому для оценки числа и параметров источников сигналов нельзя использовать выражения (5) – (7), основанные на проекционных матрицах (4), а необходимо сначала оценить степень минимального многочлена корреляционной матрицы \hat{M} .

Для оценки степени минимального многочлена в [4] предложено использовать статистику вида

$$I^{(m)} = \min \left\| I^{(m)}(\hat{M}) \right\|^2; \quad I^{(m)}(\hat{M}) = E - \sum_{n=1}^m b_n \hat{M}^n. \quad (11)$$

Степень минимального многочлена находится из решения пороговой задачи с использованием априорной информации о статистических свойствах собственного шума. Если собственный шум имеет единичную корреляционную матрицу, то априори можно утверждать, что при отсутствии внешних сигналов степень минимального многочлена $N' = 1$. Полагая в (11) $m = 1$ и считая, что выборки берутся только из собственного шума, находится среднее $\langle I^{(1)} \rangle$ и дисперсия σ_1 для статистики $I^{(1)}$.

Определяется порог p_{target} , который можно принять равным, например, $p_{target} = \langle I^{(1)} \rangle$ или $p_{target} = \langle I^{(1)} \rangle + \sigma_1$. Величина порога зависит от числа выборок L и числа N элементов антенной решетки. При $L \rightarrow \infty$ значение $p_{target} \rightarrow \infty$.

При наличии не только собственного шума, но и сигналов статистика $I^{(1)}$ будет превышать значение порога. Так как с ростом m от 1 до N функционал (11) убывает до нуля, то всегда существует такое значение $N = \hat{N}'$, при котором $I^{(1)} < p_{target}$. Оценка \hat{N} принимается за оценку степени минимального многочлена, так как дальнейшее уменьшение $I^{(m)}$ будет обусловлено влиянием шумовых собственных чисел. Получаемая при этом матрица $I^{(m)}(\hat{M})$ будет наиболее близкой к нулевой матрице.

Теперь, используя оценку степени минимального многочлена, можно построить алгоритм для оценки числа источников сигналов и их параметров. Матрица $I^{(m)}(\hat{M})$ представляет собой матричный многочлен степени m . Соответствующий ему алгебраический многочлен имеет вид

$$\psi^{(m)}(\mu) = 1 - \sum_{n=1}^m b_n \mu_n = \prod_{n=1}^m (1 - \zeta_n \mu), \quad (12)$$

где ζ_n – величина, обратная корню многочлена $\psi^{(m)}(\mu)$.

Многочлен такого типа получается из минимального многочлена (2) путем его деления на не равное нулю число $(-1)^n \mu_1 \mu_2 \dots \mu_m$. В соответствии с (12) статистику (11) запишем в другом виде:

$$I^{(m)} = \min_{\zeta_n} \left\| I^{(m)}(\hat{M}) \right\|^2, \quad I^{(m)}(\hat{M}) = \prod_{n=1}^m (E - \zeta_n \hat{M}), \quad (13)$$

где минимизация $I^{(m)}$ достигается за счет выбора чисел ζ_n .

Аналогично (3) $I^{(m)}(\hat{M})$ имеет ряд сомножителей и будет аннулировать любой вектор при $m = N$, так как $I^{(m)}(\hat{M}) = [0]$. При $m < N$ векторы будут аннулироваться в некоторую сферу минимального размера. Минимальными будут и значения собственных чисел $I^{(m)}(M)$, так как критерий (13) можно записать в виде

$$I^{(m)} = \min_{\zeta_n} \sum_{i=1}^N \prod_{n=1}^m (1 - \zeta_n \mu_i)^2. \quad (14)$$

Условия минимума этого функционала можно записать как

$$dI^{(m)} / d\zeta_n = 0, \quad n = 1, 2, \dots, m. \quad (15)$$

Отсюда получим следующую систему нелинейных уравнений для коэффициентов ζ_n $n = 1, 2, \dots, m$ [4]:

$$\zeta_n = Sp \left[\hat{M} \prod_{i=1, i \neq n}^m (E - \zeta_i \hat{M})^2 \right] \left\{ Sp \left[\hat{M}^2 \prod_{i=1, i \neq n}^m (E - \zeta_i \hat{M})^2 \right] \right\}^{-1}. \quad (16)$$

Эта система решается итерационным путем для каждого m . Причем m значений ζ_n , вычисленных для функционала $I^{(m)}$, являются начальными приближениями для $I^{(m)}$ вычисления $(m+1)$ чисел ζ_n для функционала $I^{(m+1)}$. При $m=1$ итераций не требуется, и из (16) сразу имеем, что $\zeta_1 = Sp(\hat{M}) / Sp(\hat{M}^2)$.

Полученный таким образом набор чисел ζ_n дает оценки величин, обратных собственным числам корреляционной матрицы M . Поэтому оценками собственных чисел будут величины ζ_n^{-1} . Наименьшее из них дает оценку шумового собственного числа, а другие дают оценки сигнальных собственных чисел.

Теперь аналогично (4) построим оценки матриц-проекторов в виде

$$P_k = \left[\prod_{p=1, p \neq k}^N (E - \zeta_k \hat{M}) \right] \left[\prod_{p=1, p \neq k}^N \left(1 - \frac{\zeta_p}{\zeta_k} \right) \right]^{-1}. \quad (17)$$

Выражение (17) перейдет в (5) при $L \rightarrow \infty$, если учесть, что $\hat{N}' \rightarrow N'$, $\hat{M} \rightarrow M$, $\zeta_p \rightarrow \mu_p^{-1}$, $\zeta_k \rightarrow \mu_k^{-1}$, и сделать простые алгебраические преобразования.

Оценка проектора $\hat{\Pi}_{noise}$ на подпространство собственного шума получается при подстановке в (17) в качестве ζ_k наибольшего из полученных чисел ζ_n .

Теперь число действующих источников J можно оценить по формуле

$$J = N - \left[Sp(\hat{\Pi}_{noise}) \right], \tag{18}$$

где $[.]$ обозначает округление до целого числа.

Выражение (18) является аналогом точного выражения (6).

Угловые координаты источников сигналов будут отождествляться с максимумами функции

$$\hat{\eta}_\infty(\varphi) = \left[S^H(\varphi) \hat{\Pi}_{noise} S(\varphi) \right]^{-1}. \tag{19}$$

Выражение (19) является аналогом точной формулы (7). Максимум функции $\hat{\eta}_\infty(\varphi)$ совпадает с минимумом функции $S^H(\varphi) \hat{\Pi}_{noise} S(\varphi)$. В свою очередь, глубина минимума этой функции зависит от того, насколько близки к нулю сигнальные собственные числа матрицы $I^{(m)}(\hat{M})$. Так как в соответствии с критерием (13) эта матрица выбирается наиболее близкой к нулевой матрице, то ее собственные числа будут также наиболее близки к нулю. Поэтому матрица $\hat{\Pi}_{noise}$ наиболее эффективно аннулирует векторы сигнального подпространства.

Пример 1. Допустим, что в пространстве имеется один источник. Этот пример полезно рассмотреть, хотя в случае одного источника вопрос об угловом разрешении не имеет смысла.

На рисунке 1 изображена функция (19). Расчеты выполнялись для линейной эквидистантой антенной решетки с числом элементов $N=16$ и полуволновым расстоянием между элементами, длина выборки входного процесса $L=100$. Предполагалось, что плоская волна от источника приходит по направлению нормали к антенной решетке, т.е. $\varphi_1 = 0$.

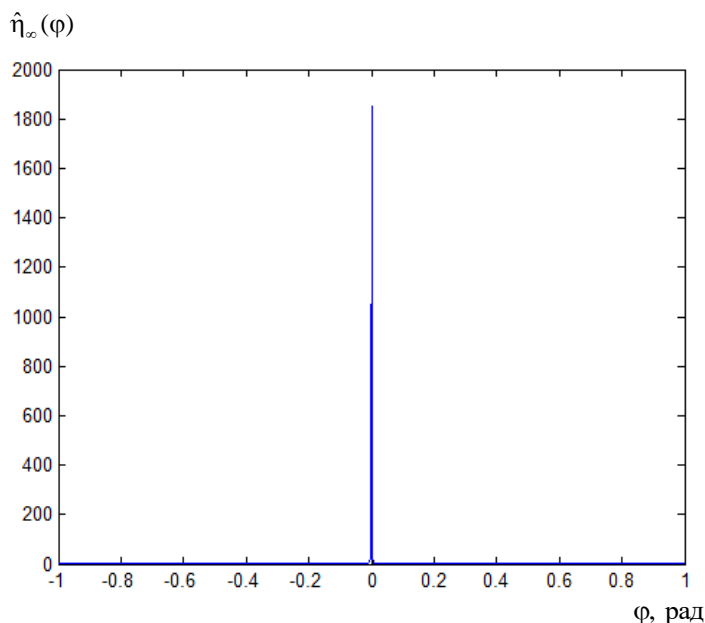


Рис. 1. Зависимость $\hat{\eta}_\infty(\varphi) = \psi(\varphi)$ при наличии одного источника

Пример 2. Допустим теперь, что в пространстве имеются два источника. Мы предполагаем, что оба источника имеют одинаковые средние мощности в каждом элементе антенной решетки. Углы прихода волн полагаем равными: $\varphi_1 = \pi/64$, $\varphi_2 = -\pi/64$.

Как и прежде, предположим, что измерение углов прихода волн выполняется с помощью линейной эквидистантной антенной решетки из 16 элементов ($N = 16$) и полуволновым межэлементным расстоянием ($d/\lambda = 0,5$). Результаты моделирования приведены на рисунке 2.

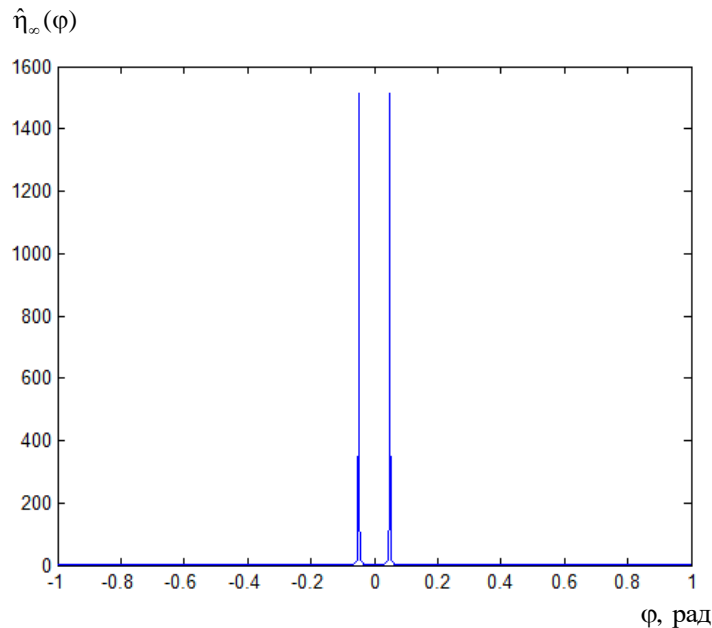


Рис. 2. Зависимость $\hat{\eta}_\infty(\varphi) = \psi(\varphi)$ при наличии двух источников

Проекционный метод, так же как и метод Кейпона, и метод «теплового шума», не требует априорного знания числа источника излучения. Однако проекционный метод является более сложным с вычислительной точки зрения, чем метод Кейпона и метод «теплового шума». Проекционный метод имеет малую чувствительность к точности аппроксимации фоновых шумов. В отличие от вышеприведенных методов, данный метод позволяет определить и количество источников излучения.

Проекционный метод имеет следующие преимущества:

- не требует априорного знания количества источников излучения;
- намного большая разрешающая способность, чем у метода «теплового шума»;
- лучшие точностные показатели, чем у метода «теплового шума»;
- меньшее влияние случайных амплитудно-фазовых ошибок.

Данный метод позволяет оценить количество источников излучения.

Проекционный метод отличается большей вычислительной сложностью алгоритма, чем у метода Кейпона и метода «теплового шума».

Для уменьшения вычислительной сложности данного алгоритма можно воспользоваться симметричностью корреляционной матрицы. Также можно воспользоваться симметричностью матрицы проецирования на шумовое подпространство (рис. 3), что позволит уменьшить количество вычислений вдвое.

1	0.8207	-0.1505	-0.1399	-0.1278	-0.1133	-0.0943	-0.0742	-0.0523	-0.0279	-0.0038	0.0205	0.0444	0.0674	0.0881	0.1083	0.1229
2	-0.1505	0.8575	-0.1409	-0.1315	-0.1197	-0.1031	-0.0855	-0.0657	-0.0427	-0.0200	0.0035	0.0265	0.0493	0.0705	0.0911	0.1073
3	-0.1399	-0.1409	0.8485	-0.1296	-0.1210	-0.1074	-0.0928	-0.0780	-0.0553	-0.0348	-0.0131	0.0081	0.0297	0.0506	0.0707	0.0876
4	-0.1278	-0.1315	-0.1296	0.8663	-0.1211	-0.1108	-0.0994	-0.0857	-0.0676	-0.0496	-0.0299	-0.0108	0.0095	0.0297	0.0493	0.0668
5	-0.1133	-0.1197	-0.1210	-0.1211	0.8948	-0.1123	-0.1044	-0.0942	-0.0791	-0.0640	-0.0467	-0.0299	-0.0114	0.0079	0.0265	0.0443
6	-0.0943	-0.1031	-0.1074	-0.1108	-0.1123	0.8912	-0.1050	-0.0985	-0.0870	-0.0753	-0.0610	-0.0473	-0.0313	-0.0138	0.0030	0.0205
7	-0.0742	-0.0855	-0.0928	-0.0994	-0.1044	-0.1050	0.9087	-0.1021	-0.0942	-0.0862	-0.0752	-0.0646	-0.0513	-0.0357	-0.0208	-0.0039
8	-0.0523	-0.0657	-0.0760	-0.0857	-0.0942	-0.0985	-0.1021	0.9268	-0.0996	-0.0954	-0.0879	-0.0808	-0.0704	-0.0571	-0.0444	-0.0285
9	-0.0279	-0.0427	-0.0553	-0.0676	-0.0791	-0.0870	-0.0942	-0.0996	0.8991	-0.0994	-0.0957	-0.0925	-0.0857	-0.0753	-0.0655	-0.0513
10	-0.0038	-0.0200	-0.0348	-0.0496	-0.0640	-0.0753	-0.0862	-0.0954	-0.0994	0.8975	-0.1032	-0.1038	-0.1004	-0.0929	-0.0861	-0.0736
11	0.0205	0.0035	-0.0131	-0.0299	-0.0467	-0.0610	-0.0752	-0.0879	-0.0957	-0.1032	0.8778	-0.1115	-0.1117	-0.1074	-0.1037	-0.0934
12	0.0444	0.0265	0.0081	-0.0108	-0.0299	-0.0473	-0.0646	-0.0808	-0.0925	-0.1038	0.8968	-0.1132	-0.1232	-0.1221	-0.1213	-0.1132
13	0.0674	0.0493	0.0297	0.0095	-0.0114	-0.0313	-0.0513	-0.0704	-0.0857	-0.1004	-0.1117	-0.1232	0.8937	-0.1325	-0.1349	-0.1294
14	0.0881	0.0705	0.0506	0.0297	0.0079	-0.0138	-0.0357	-0.0571	-0.0753	-0.0929	-0.1074	-0.1221	-0.1325	0.8624	-0.1435	-0.1406
15	0.1083	0.0911	0.0707	0.0493	0.0285	0.0030	-0.0208	-0.0444	-0.0655	-0.0861	-0.1037	-0.1213	-0.1349	-0.1435	0.8806	-0.1519
16	0.1229	0.1073	0.0876	0.0668	0.0443	0.0205	-0.0039	-0.0285	-0.0513	-0.0736	-0.0934	-0.1132	-0.1294	-0.1406	-0.1519	0.8275

Рис. 3. Пример матрицы проецирования на шумовое подпространство

Для увеличения быстродействия можно уменьшить точность определения собственных чисел корреляционной матрицы, однако это приведёт к ухудшению точности оценивания количества источников излучения и точности определения направления на источники излучения.

Увеличения точности определения количества источников излучения и направления на источники излучения можно достигнуть увеличением точности определения собственных чисел корреляционной матрицы, но это приведет к большему числу итераций нахождения собственных чисел и, как следствие, увеличению вычислительной сложности алгоритма.

Для увеличения быстродействия также можно использовать оценку степени минимального многочлена от прошлого оценивания: если количество источников не изменилось, нам необходимо вычислить только m собственных чисел для функционала I^m .

Для устранения ошибки неправильного оценивания числа источников излучения можно сравнить оценку числа источников излучения со степенью минимального многочлена. Минимальный многочлен имеет j сигнальных собственных чисел и одно шумовое число, равное 1 (при бесконечном числе выборок), следовательно, число источников должно быть на единицу меньше, чем степень минимального многочлена.

Также степень минимального многочлена от прошлого оценивания можно использовать для корректировки его оценивания при резком изменении его значения, что может быть обусловлено появлением, например, импульсных помех большой амплитуды или кратковременного попадания некоторого числа источников излучения за преграду.

Заключение. Проведен анализ проекционного метода для получения сверхразрешения. Разработаны рекомендации, основанные на симметричности матриц, позволяющие уменьшить вычислительную сложность алгоритмов, что дает возможность упростить цифровую реализацию данных алгоритмов. Результаты исследований могут быть использованы в радиопеленгации и радиолокации.

ЛИТЕРАТУРА

1. Воеводин, В.В. Линейная алгебра / В.В. Воеводин. – М.: Наука, 1980. – 400 с.
2. Монзинго, Р.А. Адаптивные антенные решетки: Введение в теорию / Р.А. Монзинго, Т.У. Миллер; пер. с англ. – М.: Радио и связь, 1986. – 448 с.
3. Уидроу, Б. Адаптивная обработка сигналов / Б. Уидроу, С. Стирнз; пер. с англ. – М.: Радио и связь, 1989. – 440 с.
4. Ермолаев, В.Т. Оценивание параметров сигналов, принимаемых антенной решеткой / В.Т. Ермолаев, А.Г. Флакман, А.А. Анурын // Изв. вузов. Радиофизика. – 1996. – Т. 39, № 9. – С. 1144 – 1160.

Поступила 16.04.2013

COMPUTER SIMULATION OF SUPERRESOLUTION OF LIGHT SOURCES ON THE BASIS OF PROJECTION METHOD

A. BEZDEL, V. YANUSHKEVICH, V. CHERTKOV

The results of theoretical analysis of superresolution light sources with the help of “thermal noise” method have been presented. The studies for one or two light sources are given. Recommendations are given to reduce the computational complexity of algorithms. The function of the angular resolution of authorization objects is analyzed. The results of the research can be used in radio position finding and radiolocation.

УДК 621.391

СИНТЕЗ ВЕЙВЛЕТ-МАНИПУЛИРОВАННЫХ СИГНАЛОВ

д-р техн. наук, доц. С.В. ДВОРНИКОВ, В.В. ДЬЯКОНОВ

(Военная академия связи им. Маршала Советского Союза С.М. Буденного, Санкт-Петербург);

С.С. ДВОРНИКОВ

(Санкт-Петербургский государственный политехнический университет);

А.В. ЖЕЛЕЗНЯК

(Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»)

Исследуется проблема синтеза вейвлет-манипулированных сигналов. Полученные результаты свидетельствуют о том, что новые манипуляционные форматы на основе вейвлетов обладают свойствами повышенной помехоустойчивости, обусловленной расширением базы. Учитывая, что спектральные различия двухпозиционных манипуляционных форматов на основе разнотипных вейвлетов связаны с наличием двух ярко выраженных максимумов функции огибающей спектральной плотности, сделан вывод о том, что эти особенности позволят облегчить их демодуляцию.

Введение. В настоящее время вейвлеты широко применяются в задачах, преимущественно связанных с анализом сложных нелинейных сигналов [1]. Это обусловлено возможностью конструирования на их основе иерархических базисов, адаптированных для описания процессов с высокой динамикой изменения параметров. Между тем в [2] предложен синтез фазоманипулированных сигналов, обладающих свойствами повышенной помехоустойчивости по отношению к их классической форме на основе гармонических функций, что открывает новые направления применения вейвлетов в радиотехнике.

Теоретическое обоснование синтеза сигналов на основе вейвлетов. В общем случае, согласно [1], вейвлеты $\psi(t)$ путем целочисленного масштабирования обеспечивают реконструкцию базиса для пространства $L^2(\mathbf{R})$, причем в качестве аналога гармонической частоты используется форма представления вейвлета $\psi(2^k t - \tau)$, в которой τ и k – целые числа ($\tau, k \in \mathbf{I}$). Следовательно, открывается возможность посредством масштабных преобразований ($1/2^k$) и временных сдвигов ($\tau/2^k$) базисного вейвлета $\psi(t)$ описать все комбинации частот и временных интервалов.

Так как норма вейвлета определяется как

$$\|\psi(2^k t - \tau)\|_2 = 2^{-k/2} \|\psi(t)\|_2,$$

то при условии, что исходный вейвлет $\psi(t) \in L^2(\mathbf{R})$ имеет единичную норму, все порождаемые на его основе функции базиса $\{\psi_{\tau k}\}$ вида

$$\psi_{\tau k}(t) = 2^{k/2} \psi(2^k t - \tau), \quad \tau, k \in \mathbf{I} \quad (1)$$

также будут нормированы к единице, т.е. $\|\psi_{\tau k}\|_2 = \|\psi\|_2 = 1$.

Таким образом, вейвлет $\psi \in L^2(\mathbf{R})$ будет ортогональным, если формируемое на его основе семейство $\{\psi_{\tau k}\}$, определенное соотношением (1), представляет ортонормированный базис функционального пространства $L^2(\mathbf{R})$:

$$\langle \psi_{\tau k}, \psi_{l m} \rangle = \delta_{\tau k} \delta_{l m}, \quad (2)$$

причем каждая функция f указанного пространства ($f \in L^2(\mathbf{R})$) может быть представлена в виде ряда

$$f(t) = \sum_{k, \tau=-\infty}^{\infty} c_{k\tau} \psi_{k\tau}(t). \quad (3)$$

В общем случае строгость выполнения условий (2) и (3) предполагает существование обратных преобразований. Однако на практике достаточно выполнения свойств устойчивости и, согласно [3], «приблизительной» ортогональности системы функций разложения. С указанных позиций для базиса (1) в качестве таких условий выступают свойства адекватности расширения и сдвига исходного вейвлета ($1/2^k, \tau/2^k$), $k, \tau \in \mathbf{I}$ и возможность обратной реконструкции с точностью до постоянного множителя,

т.е. вейвлет, используемый для реконструкций, должен удовлетворять условиям базиса Рисса. Кроме того, условиям существования вейвлета соответствуют производные высоких порядков от функции Гаусса [3]:

$$\psi_m(t) = (-1)^m \partial_t^m \left[\exp(-t^2/2) \right]. \tag{4}$$

Здесь $\partial_t^m = \partial^m[\dots]/\partial t^m$, $m \geq 1$.

Таким образом, функции, синтезируемые на основе выражения (4), являются вейвлетами и, следовательно, удовлетворяют условиям формирования на их основе базисов анализа и синтеза сигналов. Действительно, принцип передачи информации базируется на физических различиях радиосигналов, соответствующих различным компонентам информационного алфавита. Так, при амплитудной модуляции указанные различия наблюдаются в изменении амплитуды; при частотной – в номиналах несущих частот; при фазовой – соответственно в позициях фазы.

Очевидно, что в процессе модуляции можно использовать вместо несущих колебаний и вейвлеты. При этом следует учитывать, что, в отличие от гармоник, вейвлеты локализованы не только в частотном, но и во временном пространстве. Следовательно, их применение оправданно в манипуляционных форматах, например, частотных и фазовых, в которых четко локализован единичный элемент сигнала. Тогда при каждой смене позиций можно в качестве несущего колебания использовать новый вид вейвлета. Для двухпозиционных форматов (двойная фазовая манипуляция или двойная частотная манипуляция) обоснованно применение инверсных форм вейвлетов.

Таким образом, вейвлеты, способные формировать функциональные базисы, вполне могут быть использованы для синтеза сигналов частотной и фазовой манипуляции.

Одним из наиболее помехоустойчивых видов манипуляционных форматов является двойная фазовая манипуляция (ФМ-2). Принцип ее реализации состоит в инверсии фазы при смене позиции.

Так, на рисунке 1 показана временная диаграмма тестового сигнала ФМ-2 длительностью в 312 дискретных отчетов. Верхний индекс здесь и далее указывает на базис формирования сигнала; символ Γ – на гармонический базис.

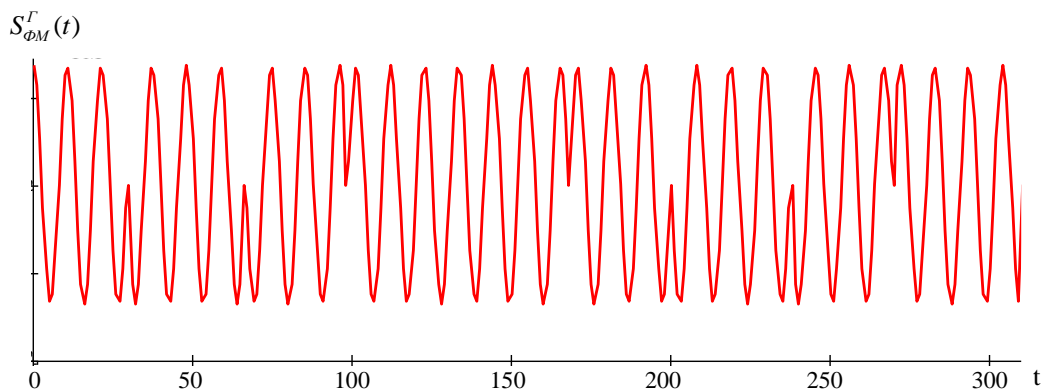


Рис. 1. Временная развертка тестового сигнала ФМ-2 на основе гармоник

Для сигнала ФМ-2 каждый из битовых символов характеризуется сменой фазы несущего колебания. Сигнал ФМ-2 относится к классу узкополосных сигналов, у которых база $B=1$, и согласно [4] определяется как

$$B = \Delta F \tau, \tag{5}$$

где τ – длительность элемента сигнала; ΔF – ширина его спектра.

Параметр B определяет помехоустойчивость. Так, отношение сигнал/шум (ОСШ) на выходе приемника q^2 связано с ОСШ на входе приемника h^2 следующим отношением [4]:

$$q^2 = 2Bh^2, \tag{6}$$

где $q^2 = \widehat{P}_c / \widehat{P}_n$ и $h^2 = P_c / P_n$. Здесь \widehat{P}_c , \widehat{P}_n – соответственно мощность сигнала и мощность помехи на выходе; P_c , P_n – соответственно мощность сигнала и мощность помехи на входе приемника.

Величину h^2 можно рассчитать с позиций энергии сигнала E и спектральной мощности шума N_0 : $h^2 = 2E/N_0$. Соответственно, $E = P_c \tau$, $N_n = P_n / F$.

Учитывая полученные соотношения, в [2] предложено использовать в качестве несущих колебаний при формировании сигнала ФМ-2 ортогональные вейвлеты, представляющие вторую производную от функции Гаусса и рассчитываемые согласно выражению (4) при $m = 2$.

На рисунке 2 показана временная развертка тестового сигнала ФМ-2, сформированная на основе указанного вейвлета. Здесь символ $B-2$ указывает на вейвлет при $m = 2$.

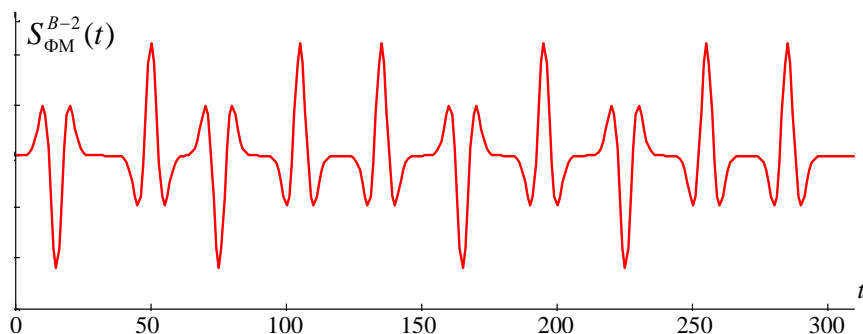


Рис. 2. Временная развертка тестового сигнала ФМ-2 на основе вейвлетов при $m = 2$

Практический интерес такого подхода состоит в том, что переход к вейвлетам ведет к расширению значимой полосы спектра, занимаемой сигналом. Проведенные в [2] исследования указывают на то, что применение в качестве несущих вейвлетов при $m = 2$ ведет к увеличению базы сигналов ФМ-2 в 3,5 раза. Согласно выражению (6) во столько же раз возрастает и помехоустойчивость. Анализ спектров вейвлетов, синтезированных на основе выражения (4), показывает, что увеличение порядка дифференцирования функции, т.е. увеличение значения m , ведет к смещению максимума спектральной плотности на частотной оси при сохранении значимой полосы спектра.

Так, на рисунке 3 показаны функции огибающих спектров вейвлетов, полученных на основе формулы (4) при различных значениях m .

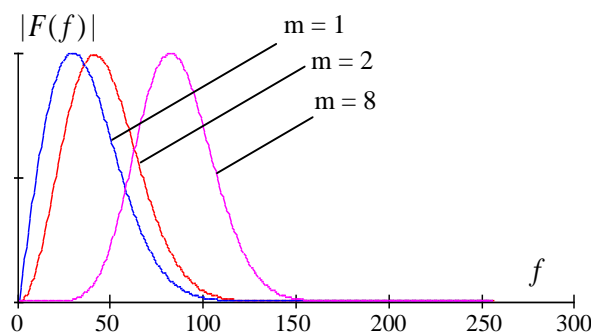


Рис. 3. Функции огибающих спектров вейвлетов при различных значениях m

Полученный результат обусловлен следующим. Сохранение ширины значимой полосы частот объясняется сохранением длительности формируемой функции при различных значениях m . В то же время повышение степени дифференцирования ведет к увеличению числа осцилляций в синтезируемом вейвлете, что равносильно увеличению частоты заполнения одиночного импульса. Следовательно, если при формировании сигнала ФМ-2 использовать в качестве несущих различные вейвлеты, а не инверсные копии одной и той же реализации функции (4), можно достичь еще большего увеличения базы. В частности, предлагается использовать вейвлеты при $m = 1$ и $m = 8$.

Пример подобной реализации для тестового сигнала ФМ-2 показан на рисунке 4, где символ $B-18$ указывает на использование вейвлетов при $m = 1$ и $m = 8$. Такой выбор вейвлет-базиса обусловлен тем, что точка пересечения функций спектральных огибающих для $m = 1$ и $m = 8$ проходит по уровню 0,5 от максимального значения, что обеспечивает рациональное распределение спектральной плотности мощности в действующей полосе частот.

На рисунке 5 показан спектр тестового сигнала ФМ-2 в рекомендуемом вейвлет-базисе с нанесенным уровнем значимой полосы частот.

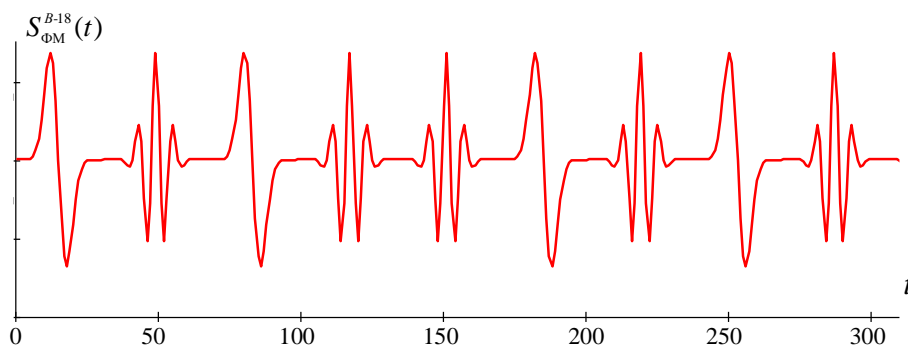


Рис. 4. Временная развертка тестового сигнала ФМ-2 на основе вейвлетов

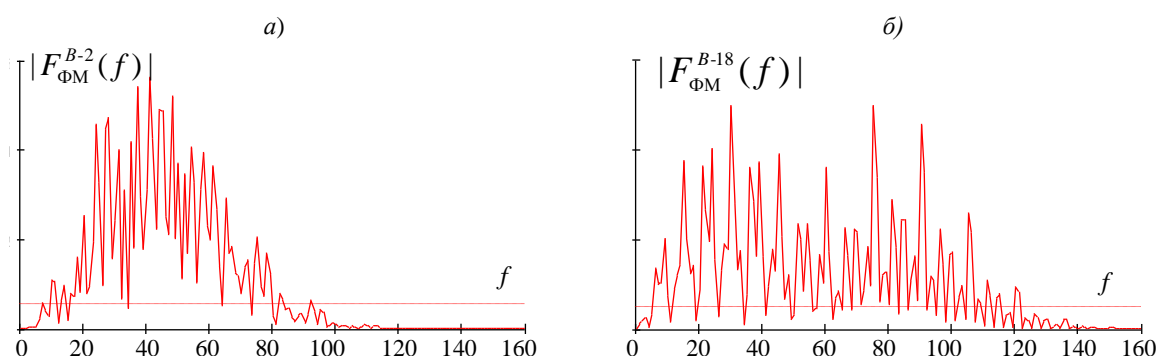


Рис. 5. Спектры сигналов ФМ-2 на основе базисов вейвлетов при $m = 2$ (а); при $m = 1$ и $m = 8$ (б)

В ходе моделирования было установлено, что использование базисов вейвлетов при $m = 1$ и $m = 8$ для синтеза тестового сигнала ФМ-2 ведет к увеличению базы в 4,9 раза по отношению к гармоническому базису и в 1,4 раза по отношению к вейвлет-базису при $m = 2$.

Заключение. Полученные результаты указывают на то, что новые манипуляционные форматы на основе вейвлетов обладают свойствами повышенной помехоустойчивости, обусловленной расширением базы. Учитывая, что спектральные различия двухпозиционных манипуляционных форматов на основе разнотипных вейвлетов связаны с наличием двух ярко выраженных максимумов функции огибающей спектральной плотности, эти особенности позволяют облегчить их демодуляцию. Дальнейшее направление исследования видится в оптимизации выбора пары вейвлетов для синтеза на их основе манипуляционных форматов.

ЛИТЕРАТУРА

1. Wavelets / Eds J.M. Combes, A. Grossmann, P. Tchamitchian. – Berlin: Springer-Verlag, 1989.
2. Способ формирования помехоустойчивых сигналов: пат. Рос. Федерации. – № 2412551 от 20.02.2011.
3. Асафьев, Н.М. Вейвлет-анализ: основы теории и примеры применения / Н.М. Асафьев // Успехи физических наук. – 1996. – Т. 166, № 11. – С. 1145 – 1170.
4. Варакин, Л.Е. Системы связи с шумоподобными сигналами / Л.Е. Варакин. – М.: Радио и связь, 1985. – 384 с.

Поступила 14.05.2013

SYNTHESIS OF WAVELET-MANIPULATED SIGNALS

S. DVORNIKOV, V. DYAKONOV, S. DVORNIKOV, A. ZHELEZNYAK

The problem of synthesis of wavelet-manipulated signals is studied. The results obtained state that new manipulation formats on the basis of wavelets have the properties of increased noise immunity caused by expansion of the base. Taking into account that spectral differences of two-position manipulation formats on the basis of polytypic wavelets are connected with the presence of two strongly pronounced peaks of the function of circumflex spectral density, one can make a conclusion that these peculiarities will allow to facilitate their demodulation.

УДК 621.372.037.372

ОБОСНОВАНИЕ ОПТИМАЛЬНОГО СИГНАЛА ДЛЯ ОЦЕНКИ ЗАЩИЩЕННОСТИ ЦИФРОВЫХ КАНАЛОВ УТЕЧКИ ИНФОРМАЦИИ

д-р техн. наук, проф. В.К. ЖЕЛЕЗНЯК, Д.С. РЯБЕНКО
(Полоцкий государственный университет)

Исследуется оптимальный сигнал для оценки защищенности цифровых каналов утечки информации. Технология технической защиты информации формирует обобщенные требования к теории и технике передачи систем сигналов. Рассматривается система сигналов, используемая для передачи информации как совокупность сигналов, объединяемых единым правилом построения. Известно, что помехоустойчивость – одно из основных требований к системе передачи. Предложены оптимальная система сигналов, обеспечивающая максимальную помехоустойчивость при минимальных отношениях энергии бита к спектральной плотности мощности шума в каналах утечки информации, методы оценки защищенности дискретных систем сигналов в каналах утечки информации при воздействии шумов высокого уровня типа белого гауссовского шума, а также выбор и обоснование оптимального сигнала, который позволит оценить защищенность цифровых каналов утечки информации.

Защищенность цифровых сигналов в канале утечки информации основана на формировании исходных требований, учитывающих свойства канала передачи сигнала, характеристики и параметры формируемых сигналов для передачи и их структуру. Сигнал канала передачи оценивают средней мощностью передачи сигнала, его искажением, которое определяется отношением энергии бита к спектральной плотности мощности шума. Обеспечение помехоустойчивости при приеме сигналов является важной задачей. Помехоустойчивость при воздействии шумов обеспечивают когерентным либо некогерентным приемом сигналов, их видами.

Основная часть. Виды сигналов различают по их цифровой модуляции. Нами будут рассмотрены фазоманипулированные сигналы (ФМн), частотно-манипулированные (ЧМн), амплитудно-манипулированные (АМн) и сигналы с квадратурно-амплитудной модуляцией (КАМ).

Системы сигналов обладают пороговым эффектом. В качестве критерия оценки влияния канала передачи на канал утечки информации может служить пропускная способность, которая определяется по формуле К.Е. Шеннона [1]:

$$C = W_c \log_2(1 + S/N). \quad (1)$$

Как следует из формулы (1), пропускная способность гауссовского канала C (бит/с) определяется шириной полосы сигнала W_c (Гц), отношением мощности сигнала S (Вт) к средней мощности шума N (Вт), ограниченного полосой W_c .

Зависимость энергии бита (E_b) к спектральной плотности мощности шума (N_0) от нормированной полосы пропускания канала представлена на рисунке 1 [1, с. 550].

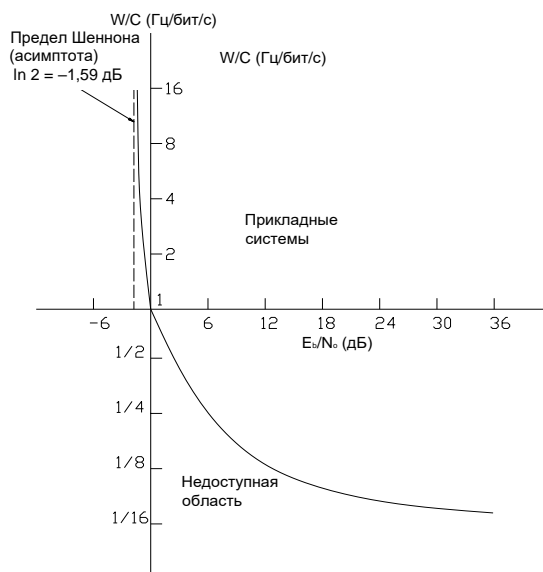


Рис. 1. Зависимость E_b/N_0 от нормированной полосы пропускания канала

При нижнем предельном значении $E_b/N_0 = -1,6$ дБ, определяющем пороговое значение, ни при какой скорости передачи нельзя осуществить безошибочную передачу информации.

Важным критерием являются методы оценки защищенности дискретных систем сигналов в каналах утечки информации при воздействии шумов высокого уровня типа белого гауссовского шума. Помехоустойчивость зависит от расстояния между сигналами $S_i(t)$ и $S_j(t)$, образующими систему [2]:

$$d(S_i, S_j) = \left\{ \int_0^{T_c} [S_i(t) - S_j(t)]^2 dt \right\}^{1/2}, \quad (2)$$

где T_c – длительность сигнала.

Помехоустойчивость системы сигналов увеличивается при увеличении расстояния d . Для достижения большего расстояния коэффициент взаимной корреляции должен быть минимальным:

$$d(S_i, S_j) = [2E(1 - R_{ij})]^{1/2}. \quad (3)$$

Здесь $R_{ij} = \frac{1}{E} \int_0^{T_c} S_i(t) \cdot S_j(t) dt$ – коэффициент взаимной корреляции между сигналами $S_i(t)$ и $S_j(t)$, принимающий значения в пределах от -1 до $+1$; E – энергии сигналов $S_1(t) \cdot S_2(t)$.

Согласно [3], для оптимальной системы сигналов $R = -\frac{1}{m-1}$. Это достигается равенством всех коэффициентов корреляции, т.е. $R_{ij} = R$ для всех i и j .

Под оптимальной следует понимать систему сигналов, обеспечивающую максимальную помехоустойчивость при заданных априорных условиях передачи информации.

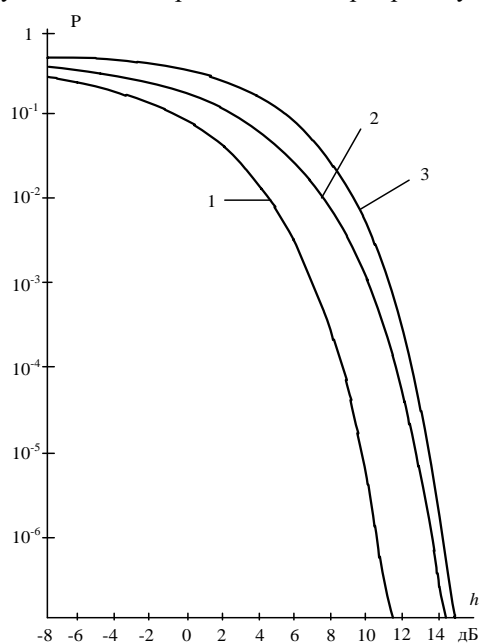


Рис. 2. Зависимость вероятности ошибки P от отношения энергии бита к спектральной плотности мощности шума h

Для детального анализа сигналов необходимо учесть, что отношение сигнал/шум в канале утечки информации гораздо меньше единицы, а их наводки являются несимметричными.

На рисунке 2 представлена зависимость вероятности ошибки P от отношения мощности сигнала к мощности шума h для когерентного 1 приема двоичных ФМ-сигналов, когерентного 2 и некогерентного 3 приема двоичных ЧМ-сигналов [4; 5].

Из рисунка следует, что при малом значении отношения сигнал/шум разброс вероятности ошибок незначителен.

Вероятность ошибки $P_{ош}$ от отношения энергии бита спектральной плотности мощности к мощности шума h для некогерентного приема ЧМ-сигнала представлена зависимостью [1, с. 245]:

$$P_{ош} = \frac{1}{2} \exp\left(-\frac{1}{2} \frac{E_b}{N_0}\right). \quad (4)$$

Зависимость вероятности двоичного ФМ-сигнала от отношения энергии сигнала к спектральной плотности мощности помехи для когерентного приема [3; 6]

$$P_e = 1 - F\left[\sqrt{\frac{2E_b}{N_0}}\right]. \quad (5)$$

Зависимость вероятности ортогональных ЧМ-сигналов на промежутке времени $[0, T]$ при условии равенства энергии сигналов [3]

$$P_e = 1 - F\left[\sqrt{\frac{E_b}{N_0}}\right]. \quad (6)$$

Формирование ортогональных ЧМ- и противоположных ФМ-сигналов основано на определенных закономерностях.

Заданы два гармонических колебания [3]:

$$\left. \begin{aligned} S_k(t) &= A_k \cos(\omega_k t + \varphi_k), \\ S_m(t) &= A_m \cos(\omega_m t + \varphi_m). \end{aligned} \right\} \quad (7)$$

Определяют условия, при которых эти колебания являются ортогональными сигналами в промежутке времени $[0, T]$ [3]:

$$\int_0^T S_k(t) \cdot S_m(t) dt = 0. \quad (8)$$

Для ЧМн-сигналов условие ортогональности соблюдается при равенстве нулю взаимной энергии двух гармонических колебаний $S_k(t)$ и $S_m(t)$ в промежутке времени $[0, T]$ [3]:

$$E(S_k, S_m) = \int_0^T A_k A_m \cos(\omega_k t + \varphi_k) \cos(\omega_m t + \varphi_m) dt = 0. \quad (9)$$

Равенство нулю (9) достигается дополнительными условиями.

Условие 1 – начальные фазы $\varphi_k = \varphi_m = 0$ на промежутке времени $[0, T]$, ортогональность сигналов достигается подбором частот [3]:

$$\left. \begin{aligned} (\omega_k t + \varphi_k)T &= \chi_1 \pi, \\ (\omega_m t + \varphi_m)T &= \chi_2 \pi, \end{aligned} \right\} \quad (10)$$

где $(\chi_1, \chi_2) = 1, 2, 3 \dots$

Из (10) определяют частоты, удовлетворяющие условиям ортогональности:

$$\left. \begin{aligned} \omega_k &= \frac{(\chi_1 + \chi_2)\pi}{2T}, \\ \omega_m &= \frac{(\chi_1 - \chi_2)\pi}{2T}. \end{aligned} \right\} \quad (11)$$

Условие 2 – начальные фазы $\varphi_k = \varphi_m$, ортогональность сигналов достигается подбором фаз [3]:

$$\sin[(\omega_k + \omega_m)T + \varphi_k + \varphi_m] - \sin(\varphi_k + \varphi_m) = 0, \quad (12)$$

при $\omega_k = \omega_m$.

При одинаковых частотах, определяемых из условия $\omega_k = \chi_1 \pi / T$, сигналы являются ортогональными, если сдвиг фаз между ними равен $\pm \pi / 2$.

Условие 3 – сигналы $S_k(t)$ и $S_m(t)$ противофазны: $\varphi_k = -\varphi_m$ [3].

Двоичные системы передают одну двоичную единицу (или один бит) информации. Алфавит, содержащий n символов, передает $\log_2 n$ двоичной единицы информации на каждый символ. Посредством многопозиционной манипуляции несущего колебания по амплитуде, частоте и фазе формируют m -ичные сигналы, используя $m = 2^k$, где k – количество двоичных единиц в символе. Применение сигналов с многопозиционной АМн возможно в каналах с низким уровнем аддитивного шума.

Различия между системами передачи информации определяются объемом алфавита источника n и сигнала m . В зависимости от m системы передачи информации разделяют на двоичные ($m = 2$) и m -ичные ($m > 2$).

На рисунке 3 [4; 5] показана зависимость вероятности ошибочного приема символа от нормированного отношения сигнал/шум, определяемого как отношение сигнал/шум, приходящееся на двоичную единицу информации [4]:

$$h_2 = \frac{h_k}{k} = h_m \log_2 m. \quad (13)$$

Кривые построены для сигналов с многопозиционной фазовой манипуляцией при $m = 2, 4, 8, 16$ и 32 .

Из рисунка 3 следует, что при малом отношении сигнал/шум h_2 (дБ) < 0 для $m = 2$ вероятность ошибочного приема символа меньше, чем для $m > 2$. В этой связи сигнал с двухпозиционной ФМн имеет

предпочтение перед сигналами с $m > 2$, так как, не обнаружив двоичный ФМн сигнал в шумах при отношении сигнал/шум < 0 (дБ), можно гарантировать защищенность сигналов многопозиционного сигнала.

При некогерентном приеме начальная фаза неизвестна и является случайной величиной [2]. Наибольшая помехоустойчивость имеет место при передаче двоичной информации ортогональными сигналами.

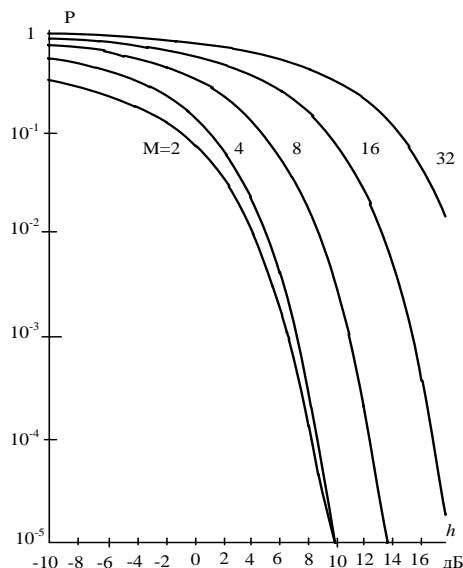


Рис. 3. Зависимость вероятности ошибки P от отношения мощности сигнала к мощности шума h для когерентного ФМн сигнала

Вероятность ошибки при некогерентном приеме двух ортогональных сигналов определяется из формулы (4). Помехоустойчивость когерентного и некогерентного приема двух ортогональных сигналов отличается незначительно.

Широкое применение на практике получили сигналы КАМ. При одинаковом числе сигналов в ансамбле сигналы КАМ обеспечивают более высокую энергетическую эффективность по сравнению с сигналами ФМн, АМн, ЧМн. Сигналы КАМ обладают большими значениями минимального межточечного расстояния d_{\min} , чем АМн и ФМн, а следовательно, и большим значением помехоустойчивости [6].

Средняя вероятность ошибки для КАМ-М при $m = 2^k$ (k – четно) на m -арный символ равна [6]

$$P_{\text{ош}} = 4 \left(1 - \frac{1}{\sqrt{m}}\right) Q \cdot \left(\frac{d}{\sqrt{2N_0}}\right) \left[1 - \left(1 - \frac{1}{\sqrt{m}}\right) Q \cdot \left(\frac{d}{\sqrt{2N_0}}\right)\right], \quad (14)$$

где $\frac{d}{\sqrt{2N_0}} = \sqrt{\frac{h_m^2}{(\sqrt{n}-1)^2}} = \sqrt{\frac{3h_c^2}{m-1}}$, $h_m^2 = \frac{E_m}{N_0}$, $h_c^2 = \frac{E_c}{N_0}$.

Сравним характеристики качества КАМ и ФМн для заданного объема сигналов m . Оба типа сигналов являются двухмерными.

Аппроксимация вероятности ошибки на символ m -позиционной ФМн выглядит следующим образом [6]:

$$P_{\sqrt{m_{\text{ФМн}}}} \approx 2Q \left(\sqrt{2\gamma_s} \sin \frac{\pi}{m} \right). \quad (15)$$

Аппроксимация вероятности ошибки на символ m -позиционной КАМ представлена в виде [6]

$$P_{m_{\text{КАМ}}} = 2 \left(1 - \frac{1}{\sqrt{m}}\right) Q \left(\sqrt{\frac{3}{m-1} \frac{E_{\text{ср}}}{N_0}} \right). \quad (16)$$

Отношение двух аргументов Q -функции [6]

$$K_m = \frac{P_{m_{\text{КАМ}}}}{P_{\sqrt{m_{\text{ФМн}}}}} = \frac{3/(m-1)}{2 \sin^2(\pi/m)}. \quad (17)$$

Из ряда работ следует, что наилучшей помехоустойчивостью обладает двоичная система сигналов с фазовой манипуляцией ФМн при когерентном приеме, отношение сигнал/шум которой [7]

$$h = \sqrt{\frac{2E}{N_0}}, \quad (18)$$

где N_0 – спектральная плотность мощности шума.

При $m > 2$ преимущества ФМн снижаются из-за взаимного влияния, обусловленного взаимной корреляцией между сигналами.

В многопозиционной системе ЧМн символы передаются с различными значениями частоты. Некогерентный прием ортогональных сигналов ЧМн (рис. 4) в области нормированного отношения сиг-

нал/шум на двоичную единицу h_2 (дБ) меньше 0 дБ имеют ту же зависимость от вероятности ошибочного приема P , т.е. двоичная система ($m = 2$) имеет преимущество по сравнению с системами $m > 2$.

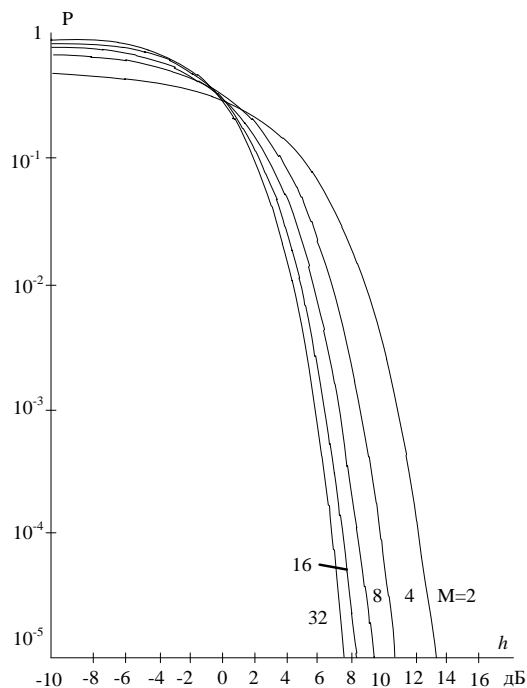


Рис. 4. Зависимость вероятности ошибки P от отношения мощности сигнала к мощности шума h для некогерентного ЧМн сигнала [4]

Отношение сигнал/шум для m -арного символа связано с отношением сигнал/шум на двоичную единицу [4]

$$h_2 = \frac{h}{k}. \quad (19)$$

Вероятность ошибки [4]

$$P_i < m \cdot \exp\left(-k \frac{h_2}{2}\right) = \exp\left[-k \left(\frac{h_2}{2} - \ln 2\right)\right].$$

Вероятность ошибочного приема экспоненциально стремится к нулю при условии, что отношение сигнал/шум на двоичную единицу удовлетворяет неравенству [4]

$$h_2 > \ln 2. \quad (20)$$

При пороговых значениях, если энергия сигнала на двоичный символ превышает спектральную плотность мощности шума в 1,39 раза, можно говорить об использовании двоичного ортогонального сигнала в качестве измерительного в канале утечки информации, учитывая его преимущества в большей чувствительности и несинхронности системы.

Формирование двоичного ортогонального сигнала рассмотрено в литературе [8]. Спектр ЧМн сигнала представлен на рисунке 5.

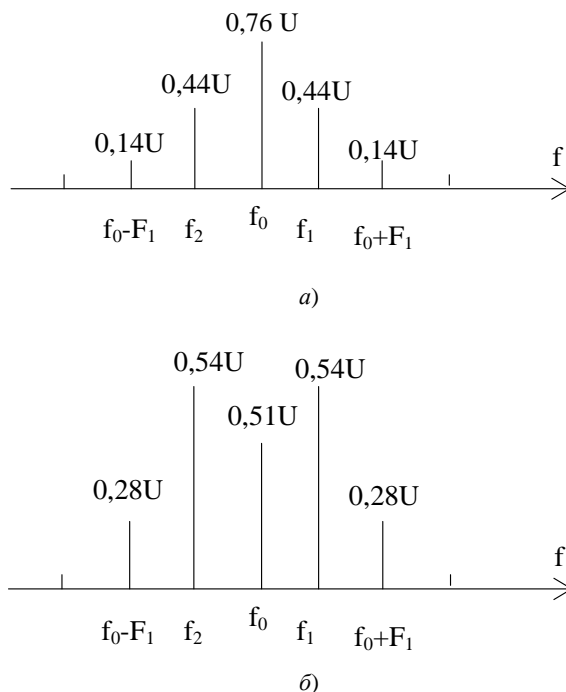


Рис. 5. Спектр ЧМн сигнала:
а – для $k_{\text{ЧМн}} = 0,8$; б – для $k_{\text{ЧМн}} = 1,2$ [8]

Из рисунка 5 следует, что при малых значениях коэффициента модуляции $k_{\text{ЧМн}}$ энергия колебания сосредоточена в полосе частот вблизи несущей частоты f_0 . При этом ширина спектра колебания ЧМн сиг-

нала без разрыва фазы равна ширине спектра колебания АМн. По мере увеличения коэффициента $k_{\text{ЧМн}}$ энергия ЧМн сигнала концентрируется в области боковых частот.

Заключение. В результате проведенной работы предложено для оценки защищенности цифровых каналов утечки информации использовать в качестве измерительных двоичные ортогональные некогерентные сигналы без разрыва фазы, зависимости вероятности ошибки от отношения мощности сигнала к мощности шума которых находятся вблизи границы Шеннона.

ЛИТЕРАТУРА

1. Скляр, Б. Цифровая связь. Теоретические основы и практическое применение / Б. Скляр. – 2-е изд. пер. с англ. – М.: Издат. дом «Вильямс», 2007. – 1104 с.
2. Дядюнов, А.Н. / Адаптивные системы сбора и передачи аналоговой информации. Основа теории / А.Н. Дядюнов, Ю.А. Онищенко, А.И. Сенин. – М.: Машиностроение, 1988. – 288 с.
3. Клюев, Н.И. Информационные основы передачи сообщений / Н.И. Клюев. – М.: «Связь», 1966. – 360 с.
4. Стейн, С. Принципы современной теории связи и их применение к передаче дискретных сообщений / С. Стейн, Дж. Джонс; отв. ред. Л.М. Финк. – М.: «Связь», 1971. – 376 с.
5. Витерби, Э.Д. Принципы когерентной связи / Э.Д. Витерби; пер. с англ. под ред. Б.Р. Левина. – М.: «Советское радио», 1966. – 392 с.
6. Савищенко, Н.В. Многомерные сигнальные конструкции: их частотная эффективность и потенциальная помехоустойчивость приема / Н.В. Савищенко; под ред. Д.Л. Бураченко. – СПб.: Изд-во Политехн. ун-та, 2005. – 420 с.
7. Варакин, Л.Е. Теория систем сигналов / Л.Е. Варакин. – М.: «Сов. радио», 1978. – 304 с.
8. Железняк, В.К. Основы теории модулированных колебаний: учебное пособие / В.К. Железняк, С.В. Дворников. – СПб.: ГУАП, 2006. – 160 с.

Поступила 19.04.2013

STUDY OF AN OPTIMAL SIGNAL FOR THE ESTIMATION OF SAFETY OF DIGITAL CHANNELS OF INFORMATION LEAKAGE

V. ZHELEZNYAK, D. RYABENKO

An optimal signal for the estimation of safety of digital channels of information leakage is studied. The technology of technical protection of information forms generic requirements to the theory and technique of transmission of systems of signals. The system of signals is considered, which is used for transmission of information as a total of signals, united by an integrated rule of construction. It is known, that noise immunity is one of the main requirements to the system of transmission. An optimal system of signals, providing maximal noise immunity at minimal ratio of bit energy to the spectral density of noise power in the channels of information leakage, methods of estimation of safety of discrete signal systems in the channels of information leakage under the influence of high level noises like Gaussian white noise and the choice and explanation of an optimal signal, which will allow to estimate the safety of digital channels of information leakage are proposed.

УДК 519.1:519.8:004.6

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И АЛГОРИТМ ПОСТРОЕНИЯ РАСПИСАНИЯ
УЧЕБНЫХ ЗАНЯТИЙ УНИВЕРСИТЕТА****М.В. ДЕКАНОВА***(Полоцкий государственный университет)*

Представлена математическая модель составления расписания учебных занятий для университета на основе гиперграфа. В модели учтены обязательные ограничения: занятия, которые проводятся с одной и той же группой (одним и тем же преподавателем, в одной и той же аудитории), должны быть назначены на различные интервалы времени; выполнение всех аудиторных работ в течение всего планового периода; общее число всех видов занятий в выделенный промежуток времени не должно превышать имеющийся аудиторный фонд; отсутствие «форточек» для студентов; обеспечение возможности сдвоенности пар; обеспечение возможности проведения определенных пар в различные дни. Показаны критерии, по которым производится оптимизация. Представлена свертка всех критериев оптимизации в одну целевую функцию. Предложен алгоритм решения задачи методом раскраски гипервершин гиперграфа.

Введение. Направление в науке, получившее название «теория расписаний», берет свое начало с известной работы Генри Гантта 1903 года [1], предложившего то, что сегодня называют диаграммами Гантта, которые встречаются во многих работах по теории расписаний [2; 3].

Первые исследования в области автоматизации составления расписания работ относятся к производственным системам и появились в 50 – 60-е годы XX века в связи с внедрением автоматизированных систем управления производством [4].

На рубеже XX и XXI веков актуальным стало создание систем автоматизированного управления учебным процессом в образовательных системах массового обслуживания [5]. Связано это с усилением требований к качеству обучения, появлением разнообразных форм обучения, развитием форм дистанционного обучения, необходимостью повышения экономической эффективности обучения и др.

Формулировка и решение задачи составления расписания занятий с помощью аппарата классической теории расписаний связана с рядом сложностей и требует модификации традиционной постановки задачи составления расписания для учета специфических особенностей организации процесса обучения в образовательных системах. Так, в работе [6] предлагается для решения задачи составления расписания учебных занятий использовать теоретический аппарат составления «производственных» расписаний.

Область применения известных методов составления «производственных» расписаний для составления расписания учебных занятий ограничена малыми образовательными системами с малым числом ограничений, накладываемым на расписание. Все это привело к появлению нового направления решения задачи составления расписания учебных занятий, основанного на непосредственном использовании методов целочисленного программирования [7 – 10].

Для решения задачи построения учебных расписаний разработано множество алгоритмов (как правило, эвристических), учитывающих те или иные особенности постановки задачи. Существуют коммерческие пакеты программ составления учебных расписаний – как отечественные [11], так и зарубежные [12; 13]. Все известные работы в этом направлении, посвященные автоматизации процедуры составления расписания занятий [14 – 26], можно условно разделить на две группы: к первой группе относятся работы, использующие классические методы решения задач целочисленного программирования (методы полного перебора, ветвей и границ, перебора в глубину, метод Гомори и др.) [14 – 24]; вторая группа работ основана на методах решения задач целочисленного программирования, использующих интеллектуальные алгоритмы решения данных задач [25; 26].

С математической точки зрения задача построения оптимального расписания учебных занятий является достаточно сложной, поскольку она принадлежит классу так называемых NP-трудных задач [27].

Используемая терминология по теории расписаний соответствует монографии [28], а по теории графов – монографии [29].

Несмотря на большую проработанность рассматриваемой проблемы, до сих пор в математических моделях не учтен ряд ограничений и критериев оптимизации (например, равномерность распределения занятий в течение дня, минимизация количества дней, занятых учебной нагрузкой, выполнение всех аудиторных работ в течение всего планового периода), не предложено эффективных алгоритмов их решения, особенно для задач большой размерности.

1. Математическая модель составления расписания

Задачу составления расписания учебных занятий университета, как и в ряде работ [16 – 18; 21; 24], предлагается решать по этапам. На *первом этапе* на основе документов, связанных с организацией

учебного процесса (как правило, учебных поручений преподавателей, нагрузкой), формируется непротиворечивое начальное расписание, соответствующее критериям оптимальности и ограничениям. *Второй этап* – назначение аудиторий для занятий, расписание которых составлено на первом этапе. *Третий этап* связан с улучшением расписания, связанным с устранением конфликтов, возникших после выполнения первых двух этапов. На каждом из этапов возможно применение различного математического аппарата и, соответственно, появление различных математических моделей и алгоритмов их решения.

В работе представлены гиперграфовая модель и алгоритм формирования рационального расписания учебных занятий университета.

Теоретико-графовая математическая модель расписания представляется следующим образом. В процессе решения задачи построения расписания необходимо назначить каждое занятие на конкретный промежуток времени рабочего дня недели. Поскольку длительность одной пары фиксирована, то заранее известны все возможные интервалы времени, на которые могут быть назначены пары. Введем обозначения: T_d – множество дней недели (понедельник, вторник, ..., суббота); T_{II} – множество регламентированных «пар» по сменам (пары от 1 до 8); T_q – множество признаков недели (расписание по четным и нечетным неделям может быть различным). Также введем множество таймслотов $T = T_d \times T_{II} \times T_q$, элементы которого однозначно определяют день недели, «пару»/смену и признак недели. Для каждого занятия m известно множество таймслотов, когда оно свободно, то есть в это время занятие может быть проведено $T_m^+ \subset T$, и множество недоступных таймслотов $T_m^- \subset T$, причем $T_m^+ \cup T_m^- = T$; $T_m^+ \cap T_m^- = \emptyset$.

Таким образом, расписание занятий вуза может быть представлено гиперграфом расписания:

$$H_{РАСП} = (M, T, E_{расп}),$$

где $E_{РАСП}$ – бинарное отношение множества учебных событий («пар»):

$$S = \{(m, t) | E_{РАСП}(m, t), m \in M, t \in T \}.$$

Расписания занятий в университете являются периодическими, поэтому достаточно составить расписание учебных занятий на один плановый период. Обычно периодом является одна неделя или же две последовательные недели (если расписания для них существенно отличаются одно от другого). Нарушения периодичности, возникающие в процессе реализации расписания, можно учитывать в режиме корректировки основного расписания, составленного на один плановый период.

Для проведения занятий студенты объединяются в академические группы $g \in G$. Обозначим через n_g количество занятий группы g за две недели. Множество групп G включает подмножества групп факультета G_U .

Лекционные занятия проводятся в потоках $r \in R$. Пусть R_U – множество потоков, включающих хотя бы одну группу, принадлежащую факультету U , а R_U – множество потоков, включающих только группы факультета U .

Для проведения некоторых занятий группы могут делиться на подгруппы $d \in D$. Обозначим D_U – множество подгрупп групп факультета U .

Введем гиперграф структуры групп, подгрупп и потоков:

$$H_{СТР} = (G, D, R, E_{стр}),$$

где $E_{СТР}$ – бинарное отношение множества пар:

$$M = \{(g, z) | E_{СТР}(g, z), g \in G, z \in D \cup R \cup G \}.$$

Множество преподавателей P также может разбиваться на подмножества P_U преподавателей различных кафедр университета. Введем гиперграф учебной нагрузки:

$$H_{УН} = (M, P, E_{УН}),$$

где $E_{УН}$ – бинарное отношение, порождающее множество пар:

$$L_B = \{(m, p) | E_{УН}(m, p), m \in M, p \in P \cup \{\emptyset\} \}.$$

Инцидентность m и p означает, что занятие $l_b \in L_B$ с контингентом студентов m (группа/подгруппа/поток) проводит преподаватель p ; $B = \{b_1, b_2, b_3\}$ (b_1 – идентификатор дисциплины; b_2 – вид занятия; b_3 – количество подгрупп, на которые подразделяется группа для проведения данного занятия ($z \in D$)).

Исходная информация для составления расписания включает аудиторный фонд. Обозначим $A_{мм}$ – множество мультимедийных аудиторий, $A_{к/к}$ – множество компьютерных аудиторий, $A_{снец}$ – множество

специализированных аудиторий, $A_{общ}$ – аудитории общего назначения. Будем считать эти множества непересекающимися, тогда $|A_{мм}| + |A_{к/к}| + |A_{спец}| + |A_{общ}|$ – количество аудиторий университета.

Введем гиперграф учебных поручений:

$$H_{УП} = (L_B, A_{мм}, A_{к/к}, A_{спец}, A_{общ}, E_{УП}),$$

где $E_{УП}$ – бинарное отношение, порождающее множество пар:

$$V = \{(b, a) | E_{УП}(l_b, a), l_b \in L_b, a \in A_{мм} \cup A_{к/к} \cup A_{спец} \cup A_{общ} \cup \{\emptyset\}\}.$$

Инцидентность между l_b и a означает, что учебное поручение $v \in V$ для занятия l_b содержит требуемую аудиторию a или «пустую» аудиторию (для ее более позднего определения).

При формировании расписания занятий должен учитываться ряд обязательных ограничений. Для их формирования введем булево обозначение учебного события $s \in S$.

$$s = s(g, z, p, a, t, l_b) = \begin{cases} 1, & \text{если с группой } g \text{ в таймслоте } t \text{ преподавателем } p \\ & \text{проводится занятие } l_b \text{ в аудитории;} \\ 0 & \text{в противном случае.} \end{cases}$$

Можно выделить следующие *обязательные ограничения*:

1) в любом таймслоте t для каждой группы g может проводиться не более одного занятия (таким образом, если два занятия проводятся с одной и той же группой студентов, то они должны быть назначены на различные интервалы времени):

$$\sum_{z:z \in R \cup G} s + \sum_{z:z \in D} \frac{s}{b_3} \leq 1, \forall t \in T, \forall g \in G; \quad (1)$$

2) в любом таймслоте t любой преподаватель p может вести не более одного занятия (таким образом, если два занятия проводятся одним преподавателем, то они должны быть назначены на различные интервалы времени):

$$\sum_z s \leq 1, \forall t \in T, \forall p \in P; \quad (2)$$

3) в любом таймслоте t в любой аудитории a не может проходить более одного занятия (таким образом, если два занятия проводятся в одной и той же аудитории, то они должны быть назначены на различные интервалы времени):

$$\sum_{z:z \in R \cup G} s + \sum_{z:z \in D} \left(\frac{s}{b_3} \right) \leq 1, \forall t \in T, \forall g \in A; \quad (3)$$

4) общее число всех видов занятий в каждом таймслоте t не должно превышать имеющийся в университете аудиторный фонд:

$$\sum_{z:z \in R \cup G} s + \sum_{z:z \in D} \left(\frac{s}{b_3} \right) \leq |A_{мм}| + |A_{к/к}| + |A_{спец}| + |A_{общ}|, \forall t \in T; \quad (4)$$

5) расписание занятий каждой группы в любой день не должно содержать «окон» (т.е. для студентов в построенном расписании не должно быть так называемых «форточек»).

Обозначим через j_{\max} номер последнего занятия группы g в день t_d недели t_q :

$$j_{\max} = \max\{t_n : s = 1\}, \forall g \in G, \forall t_d \in T, \forall t_q \in T.$$

Аналогично определим номер первого занятия j_{\min} :

$$j_{\min} = \min\{t_n : s = 1\}, \forall g \in G, \forall t_d \in T, \forall t_q \in T.$$

Тогда количество пар у группы g в день t_d :

$$j_{col} = \sum_{T_{II}} \{s : z \in R \cup G\} + \sum_{T_{II}} \left\{ \frac{s}{b_3} : z \in R \right\}.$$

Тогда искомое ограничение примет вид:

$$j_{\max} - j_{\min} - j_{col} = 0, \forall g \in G, \forall t_d \in T, \forall t_q \in T; \quad (5)$$

б) для каждой группы g должны выполняться все виды аудиторной работы в течение двух недель:

$$\sum_T \{s : z \in R \cup G\} + \sum_T \left\{ \frac{s}{b_3} : z \in D \right\} = n_g, \forall g \in G; \quad (6)$$

7) необходимо обеспечить возможность проводить некоторые занятия непосредственно одно за другим (например, лабораторные занятия, требующие две последовательные пары).

Пусть задано множество предметов L_z , для которых необходимо проведение сдвоенных пар. Тогда для некоторого занятия $l_z \in L_z$ определяется основное занятие j' , непосредственно назначаемое на определенный интервал времени t в день t_d недели t_q в аудиторию a' для группы g , и зависимое занятие j'' , следующее непосредственно за основным занятием.

Тогда искомое ограничение примет вид:

$$j'' - j' = 1; a' = a'', \forall g \in G, \forall t_d \in T, \forall t_q \in T; \forall l_z \in L_z; \quad (7)$$

8) необходимо обеспечить возможность проводить какие-либо занятия в разные дни (например, три занятия по высшей математике не должны проводиться в один день с одной группой учащихся).

Такая необходимость обеспечивается с помощью определения подмножества занятий L_w , никакие три из которых не должны проводиться в один и тот же день.

Тогда количество пар у группы g в день t_d недели t_q , принадлежащих подмножеству занятий L_w , удовлетворяет ограничению:

$$\sum_T \{s : z \in R \cup G; l_w \in L_w\} + \forall g \in G + \forall t_d \in T, \forall t_q \in T. \quad (8)$$

Таким образом, задача формирования расписания заключается в формировании гиперграфа $H_{РАСП}$, т.е. определении для каждого занятия таймслота и аудитории с учетом выполнения вышеприведенных обязательных ограничений (1) – (8).

В последующем может проводиться оптимизация всего сформированного расписания и ряд локальных оптимизаций расписаний отдельных подразделений университета по определенным желаемым критериям. В качестве частных критериев оптимизации расписания занятий могут быть предложены следующие:

1) равномерность распределения занятий по количеству «пар» в день:

$$K_1 = \left\{ \sum_{T_d} s \rightarrow \min, s \mid \forall g \in G, \forall t_d \in T, \forall t_q \in T \right\}; \quad (9)$$

2) возможность проведения занятий в определенную «смену» (временной промежуток до или после «пары» t'_q) для двух недель:

$$K_2 = \left\{ (j_{\max} \leq t'_q) \vee (j_{\min} \leq t'_q + 1), s \mid \forall g \in G, \forall t_d \in T, \forall t_q \in T \right\}. \quad (10)$$

Для оптимизации расписания преподавателей критерии применяются избирательно, в зависимости от предпочтений определенного преподавателя $p \in P$;

3) равномерность распределения занятий по количеству «пар» в день для двух недель:

$$K_3 = \left\{ \sum_{T_d} s \rightarrow \min, s \mid p \in P, \forall t_d \in T, \forall t_q \in T \right\}; \quad (11)$$

4) отсутствие занятий в определенные дни $T_d^p \subset T_d$, т.е. необходимо предусмотреть возможность запрещать назначение пар для преподавателя на тот или иной интервал времени (например, в связи с методическими днями или в связи с личными пожеланиями преподавателя):

$$K_4 = \left\{ \sum_{T_d} s = 0, s \mid p \in P, \forall t_d \in T, \forall t_q \in T_d^p \right\}; \quad (12)$$

5) минимизация количества окон:

$$K_5 = \left\{ \sum_{T_d \times T_d} (j_{\max} - j_{\min}) - \sum_T s \rightarrow \min, s \mid p \in P \right\}; \quad (13)$$

б) минимизация количества дней, занятых учебной нагрузкой:

$$T_{УН}^p \subseteq T_u \times T_d, (t_u, t_d) \in T_{УН}^p \Leftrightarrow \exists t_n, \exists z, \exists a s(p, z, a, (t_u, t_d, t_n)) = 1:$$

$$K_6 = \{T_{УН}^p \rightarrow \min, s | p \in P\}. \quad (14)$$

Эти критерии рассматриваем как частные критерии.

Предлагается построить обобщенный аддитивный критерий, который объединяет (свертывает) все частные критерии в одну целевую функцию, представляющую собой взвешенную сумму частных критериев:

$$F = \sum_{i=1}^6 w_i K_i,$$

где w_i – весовой коэффициент критерия K_i ; $0 \leq w_i \leq 1$; $F = \sum_{i=1}^6 w_i = 1$.

Коэффициенты w_i определяются на основе экспертных оценок.

2. Раскраска гипервершин гиперграфа

Для получения оптимального расписания занятий на основе описанной выше гиперграфовой модели могут использоваться методы, применяемые для раскраски гипервершин гиперграфа [21].

Следовательно, задача составления расписания сводится к формированию гиперграфа учебных поручений $H_{УП}$ гиперграфа $H_{РАСП}$, имеющего оптимальное значение критерия F (9) – (14) при соблюдении обязательных ограничений (1) – (8).

В задаче оптимальной раскраски гиперграфа требуется построить раскраску φ : гипервершин гиперграфа $H_{УП} = (V, E_{УП})$, которая содержит минимальное число цветов. Покажем, каким образом можно представить задачу построения расписания в терминах раскраски гиперграфа $H_{УП} = (V, E_{УП})$, который изображен на рисунке и строится согласно приведенным далее правилам.

Предполагается, что каждая нефиктивная гипервершина гиперграфа $H_{УП} = (V, E_{УП})$ соответствует определенному занятию, которое проводится конкретным преподавателем с конкретной группой учащихся в конкретной аудитории (если аудитория задана заранее) или в аудитории конкретного типа (если аудитория заранее не задана). Раскраска $\varphi: V \rightarrow N$ интерпретируется следующим образом: если нефиктивная гипервершина $v_i \in V$ окрашена в цвет $\varphi(v_i) \in N = \{1, 2, \dots, n\}$, то занятие, определенное гипервершиной $v_i \in V$, должно быть проведено в $\varphi(v_i)$ -й по порядку следования интервал времени из упорядоченного множества W интервалов, предназначенных для проведения пар в плановом периоде. Гиперграф $H_{УП} = (V, E_{УП})$ может содержать фиктивные гипервершины, включение которых во множество V определяется критерием K_4 оптимизации расписания преподавателей в терминах раскраски гиперграфа.

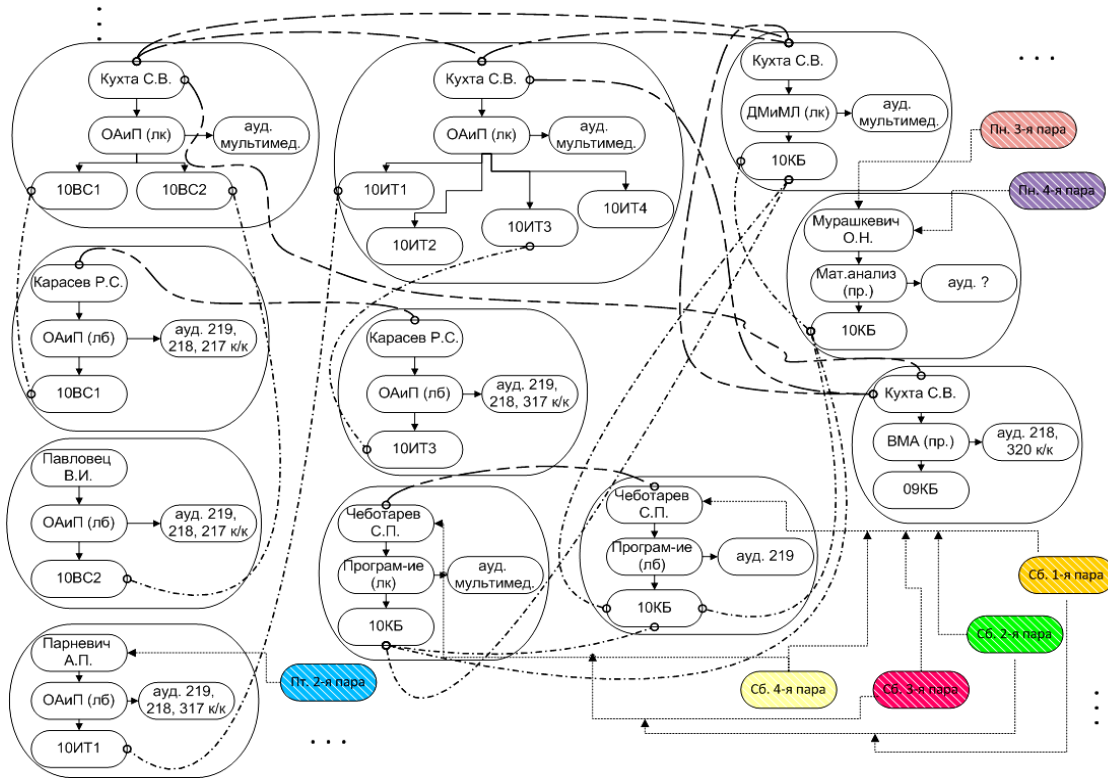
Гиперребра множества $E_{УП}$ гиперграфа $H_{УП} = (V, E_{УП})$ используются для определения обязательных ограничений (1) – (8) и критериев локальной оптимизации расписания занятий (9) – (14). В частности, гиперребро $[v_i, v_j] \in E_{УП}$, инцидентное нефиктивным гипервершинам $v_i \in V$ и $v_j \in V$, может обозначать, что занятия, соответствующие данным гипервершинам, проводятся у одной и той же группы (обязательное ограничение (1)). Поскольку в раскраске φ смежные гипервершины гиперграфа $H_{УП}$ должны быть окрашены в различные цвета, то занятия с одной и той же группой учеников должны быть проведены в различные интервалы времени из множества W . Точно так же следует ввести в гиперграф $H_{УП}$ гиперребра $[v_i, v_j] \in E_{УП}$ в связи с обязательным ограничением (2) для преподавателей и в связи с обязательным ограничением (3) для аудиторий.

Если в процессе построения гиперграфа возникают кратные гиперребра, то их следует удалить, оставив в гиперграфе $H_{УП}$ по одному гиперребру для каждого множества кратных гиперребер, инцидентных одной и той же паре гипервершин. Гипервершины, представляющие занятия одного и того же преподавателя (или одной и той же группы, или проводимые в одной и той же аудитории), должны быть попарно соединены между собой гиперребрами. В результате множество таких гипервершин определяет полный подграф гиперграфа $H_{УП}$. Следовательно, в раскраске φ гиперграфа $H_{УП}$ все гипервершины полного подграфа будут окрашены в различные цвета.

Таким образом, раскраска $\varphi: V \rightarrow N$ построенного гиперграфа $H_{УП} = (V, E_{УП})$ определяет расписание занятий, для которого выполнены обязательные ограничения (1) – (3).

Следующее ограничение, которое необходимо учитывать при составлении расписания, связано с количеством имеющихся в университете аудиторий каждого типа (обязательное ограничение (4)). Это ограничение учитывается путем введения дополнительного ограничения на использование одного и того же цвета для ограниченного числа гипервершин из заданного подмножества множества V . Конкретный цвет

гипервершин означает, что занятия, представленные гипервершинами одного цвета, проводятся в один и тот же интервал времени из множества W . Следовательно, учет при построении раскраски $\varphi: V \rightarrow N$ заданной границы на число гипервершин одного цвета позволяет ограничить количество аудиторий соответствующего типа, которые могут быть использованы в одном и том же интервале времени из множества W . Такая интерпретация обязательного ограничения (4) предполагает применение нетрадиционного понятия раскраски гипервершин гиперграфа. Для соответствующего обобщения раскраски (оптимальной раскраски) будем использовать термин допустимая раскраска $\varphi: V \rightarrow N$ (оптимальная допустимая раскраска).



Пример гиперграфа

Если в процессе построения гиперграфа возникают кратные гиперребра, то их следует удалить, оставив в гиперграфе $H_{уп}$ по одному гиперребру для каждого множества кратных гиперребер, инцидентных одной и той же паре гипервершин. Гипервершины, представляющие занятия одного и того же преподавателя (или одной и той же группы, или проводимые в одной и той же аудитории), должны быть попарно соединены между собой гиперребрами. В результате множество таких гипервершин определяет полный подграф гиперграфа $H_{уп}$. Следовательно, в раскраске φ гиперграфа $H_{уп}$ все гипервершины полного подграфа будут окрашены в различные цвета.

Таким образом, раскраска $\varphi: V \rightarrow N$ построенного гиперграфа $H_{уп} = (V, E_{уп})$ определяет расписание занятий, для которого выполнены обязательные ограничения (1) – (3).

Следующее ограничение, которое необходимо учитывать при составлении расписания, связано с количеством имеющихся в университете аудиторий каждого типа (обязательное ограничение (4)). Это ограничение учитывается путем введения дополнительного ограничения на использование одного и того же цвета для ограниченного числа гипервершин из заданного подмножества множества V . Конкретный цвет гипервершин означает, что занятия, представленные гипервершинами одного цвета, проводятся в один и тот же интервал времени из множества W . Следовательно, учет при построении раскраски $\varphi: V \rightarrow N$ заданной границы на число гипервершин одного цвета позволяет ограничить количество аудиторий соответствующего типа, которые могут быть использованы в одном и том же интервале времени из множества W . Такая интерпретация обязательного ограничения (4) предполагает применение нетрадиционного понятия раскраски гипервершин гиперграфа. Для соответствующего обобщения раскраски (оптимальной раскраски) будем использовать термин допустимая раскраска $\varphi: V \rightarrow N$ (оптимальная допустимая раскраска).

При составлении расписания следует учитывать необходимость или пожелания преподавателей, не назначать для них занятия на какие-то конкретные промежутки времени из множества W (или не проводить занятия в каких-то аудиториях в конкретные промежутки времени). Для задания такого требования или пожелания требуется сопоставить преподавателю (или аудиториям определенного типа) конкретные цвета (номера интервалов времени из множества W), запрещенные для них по тем или иным причинам (критерий оптимизации K_4 (12)). В терминах сетевой модели такое требование может быть реализовано следующим образом.

При построении допустимой раскраски необходимо раскрасить гипервершины гиперграфа $H_{\text{УП}} = (V, E_{\text{УП}})$ так, чтобы определенные его гипервершины не получили запрещенные для них цвета.

Для представления такого условия в гиперграф $H_{\text{УП}} = (V, E_{\text{УП}})$ вводится соответствующее множество фиктивных гипервершин (по одной фиктивной гипервершине для каждого интервала времени из множества W , который может оказаться запрещенным).

Например, если составляется расписание на шесть рабочих дней недели по шесть занятий каждый день, то требуется рассмотреть 48 интервалов времени для проведения занятий: $|W| = n = 48$. Если каждый из этих интервалов может быть запрещенным (для преподавателя или для какого-то типа аудиторий), то необходимо добавить 48 фиктивных вершин в построенный гиперграф, заранее окрасив их последовательно в различные цвета из множества $N = \{1, 2, \dots, 48\}$. Затем следует добавить в полученный гиперграф гиперребра, инцидентные фиктивным гипервершинам и тем гипервершинам гиперграфа $H_{\text{УП}}$, для которых следует запретить соответствующие цвета в допустимой раскраске. Например, фиктивная гипервершина v_i , заранее окрашенная в цвет $\varphi(v_i) \in N = \{1, 2, \dots, 24\}$, должна быть соединена гиперребром с каждой гипервершиной, определяющей занятие, которое нельзя проводить в $\varphi(v_i)$ -м по порядку следования интервале времени из упорядоченного множества W . В процессе последующей раскраски гиперграфа $H_{\text{УП}}$ соответствующие нефиктивные гипервершины не могут окрашиваться в запрещенные для них цвета, и, следовательно, соответствующие занятия не будут проводиться в запрещенные для них интервалы времени из множества W .

В отличие от обязательного ограничения (4), приведенная выше интерпретация критерия оптимизации K_4 не требует обобщения традиционного понятия раскраски гиперграфа. Для учета критерия оптимизации K_4 достаточно лишь дополнить множество гипервершин гиперграфа соответствующим множеством фиктивных гипервершин и соединить их гиперребрами с гипервершинами, для которых требуется запретить использование соответствующих цветов. Следует заметить, однако, что построение допустимой раскраски гиперграфа $H_{\text{УП}}$ следует начинать при условии, что части его гипервершин (а именно всем фиктивным гипервершинам) заранее приписаны конкретные цвета, причем эти цвета не могут быть изменены в дальнейшем, даже если это исключает возможность получения оптимальной раскраски гиперграфа $H_{\text{УП}}$.

Ограничение (7) учитывается в процессе построения допустимой раскраски $\varphi: V \rightarrow N$ следующим образом. При окрашивании подмножества гипервершин, связанных ограничением (7), цвет необходимо выбирать только для основной гипервершины из этого подмножества. При этом неосновным гипервершинам подмножества цвета назначаются исходя из их расположения относительно окрашенной основной гипервершины. Аналогично обязательному ограничению (4) такая интерпретация ограничения (7) также приводит к нетрадиционному понятию раскраски гиперграфа, состоящему в принудительном окрашивании неосновных гипервершин гиперграфа $H_{\text{УП}}$ согласно цветам, которые назначены его основным вершинам. Для учета этого обобщения понятия раскраски будем использовать тот же термин «допустимая раскраска» гиперграфа $H_{\text{УП}}$.

Обязательные ограничения (5), (6), (8) отражаются непосредственно в целевой функции, которая используется при построении оптимальной допустимой раскраски гиперграфа $H_{\text{УП}}$, и учитываются как тенденция к запрещению цветов, нарушающих эти условия уже в самом алгоритме построения допустимой раскраски.

Опишем алгоритм, который реализуется перед началом раскраски гипервершин гиперграфа. На этом этапе определяются гипервершины, цвета которым назначаются заранее, а также множества допустимых цветов, которые могут быть назначены остальным гипервершинам. Если явно не оговорено, какие гипервершины просматриваются, то предполагается, что просматриваются нефиктивные гипервершины гиперграфа $H_{\text{УП}}$.

1. Фиктивные гипервершины окрашиваются в соответствующие цвета из множества N .
2. Для всех нефиктивных гипервершин производится анализ, с какой гипервершины необходимо начинать.
3. Начинаем искать гипервершину v_i для совместителей (занятия у совместителей обычно проводятся либо по субботам, либо во вторую смену). Такую гипервершину красим в единственно возможный для нее цвет и просматриваем все смежные с ней гипервершины, запрещая для них цвет, в который окрашена гипервершина v_i .

4. Далее выбирается гипервершина, у которой наибольшее количество окрашенных «соседей». Такую гипервершину красим в единственно возможный для нее цвет и просматриваем все смежные с ней гипервершины, запрещая для них цвет, в который окрашена гипервершина v_i .

5. Находим гипервершину с аудиторией, которая является самой «востребованной» (необходимо определить индекс востребованности). Такую гипервершину красим в единственно возможный для нее цвет и просматриваем все смежные с ней гипервершины, запрещая для них цвет, в который окрашена гипервершина v_i .

6. Находим гипервершину с преподавателем, который является самым «востребованным». Такую гипервершину красим в единственно возможный для нее цвет и просматриваем все смежные с ней гипервершины, запрещая для них цвет, в который окрашена гипервершина v_i .

7. Просматриваем все гипервершины до тех пор, пока будем находить неокрашенные гипервершины, которые можно окрасить только в один цвет из множества N . Если находим такую гипервершину v_i , то красим ее в единственно возможный для нее цвет из множества N . Просматриваем все смежные с ней гипервершины и запрещаем для них цвет, в который окрашена гипервершина v_i . Если для какой-либо гипервершины остается только один возможный цвет, то после просмотра всех гипервершин шаг 7 повторяется снова.

Таким образом, находим гипервершины, которые можно окрасить только в один цвет из множества N . Если найдена такая гипервершина v_i , то гипервершину v_i помечаем как просмотренную. Для непомеченных гипервершин находим все допустимые цвета. Среди непросмотренных гипервершин находим группы идентичных гипервершин: они соответствуют одной и той же группе учащихся, одному и тому же преподавателю и требуют для своего проведения либо один и тот же тип аудитории, либо одну и ту же аудиторию. Идентичные гипервершины помечаются как просмотренные, и для них находятся цвета, в которые они могут быть окрашены. Затем выполняются действия, аналогичные шагу 7. Определяются списки допустимых цветов для всех гипервершин, которые еще не окрашены.

Алгоритм раскраски гипервершин гиперграфа представляет собой выполнение следующих шагов:

- шаг 1. Назначаются последовательно цвета из множества N всем гипервершинам, которые необходимо раскрасить. Назначаются цвета гипервершинам блока в соответствии с цветом, который получила вершина из массива цветов;

- шаг 2. Вычисляются оценки качества полученного назначения цветов. Для этого формируется массив конфликтных вершин, в который помещаются смежные гипервершины с одинаковыми цветами. Для каждой гипервершины v_i просматриваются смежные с ней гипервершины. Если какая-либо из смежных гипервершин имеет такой же цвет, как и гипервершина v_i , то гипервершина v_i помещается в массив конфликтных вершин, а штраф за назначение цветов увеличивается на величину соответствующего штрафа;

- шаг 3. Для каждого интервала времени и каждого ресурса проверяется, превышено ли число имеющихся ресурсов. Если оно превышено, все гипервершины, которые запрашивают данный ресурс в данный интервал времени, помечаются как конфликтные. Штраф за полученное назначение цветов увеличивается на произведение штрафа за данное нарушение на число таких нарушений. Если нарушено, например, правило – соответствующие занятия не следуют непосредственно одно за другим, то обе гипервершины помечаются как конфликтные, а штраф за данное назначение цветов увеличивается на величину штрафа за это нарушение. Если нарушено правило, например, соответствующие занятия проводятся в один день, то одна из гипервершин помечается как конфликтная, а цена решения увеличивается на величину штрафа за это нарушение. Определяются конфликты, приводящие к появлению форточек для студентов. Для каждого дня проверяется, есть ли у группы студентов в этот день форточки. Если есть, то подсчитывается их количество, а штраф за данное назначение цветов увеличивается на произведение штрафа за такое нарушение на количество форточек. Если форточка одна, то конфликтной считается либо первая гипервершина v_j , либо последняя гипервершина v_j , окрашенная в цвет из множества цветов для данного дня; если имеется больше одной форточки, то обе гипервершины v_i и v_j считаются конфликтными;

- шаг 4. Если штраф за полученное назначение цветов равняется нулю, то имеем допустимую раскраску, которая определяет расписание учебных занятий. В противном случае выполняется шаг 3;

- шаг 5. Разрешение конфликтов. Выделяются блоки гипервершин, которые содержат конфликтные гипервершины. Все гипервершины такого блока помечаются как конфликтные. Если блок помечен как имеющий фиксированный цвет, то такой блок не может содержать конфликтных гипервершин. Все конфликтные гипервершины считаются неокрашенными, и соответствующие ресурсы освобождаются. Случайным образом выбираем блок с конфликтными гипервершинами. Вычисляются возможные цвета для определяющей гипервершины выбранного блока с учетом уже окрашенных гипервершин. Если количество допустимых цветов отлично от нуля, то случайным образом выбирается один из них. Если нет допустимых цветов во множестве N , то выбирается один из возможных цветов без учета уже окрашенных гипервершин. Этот цвет назначается определяющей гипервершине блока, а также остальным его гипервершинам. Ресурсы, запрашиваемые гипервершинами блока, определяются как используемые.

Среди конфликтных гипервершин выбирается случайным образом гипервершина v_i . Просматриваются все гипервершины, смежные с v_i , и если им приписан цвет, то он помечается как запрещенный для гипервершины v_i . Для каждого цвета из множества N , который остался допустимым, определяем, можно ли назначить его гипервершине v_i с учетом использованных ресурсов. Если нельзя, помечаем этот цвет как недопустимый. Затем удаляем форточки из списка допустимых цветов. Для каждой группы студентов, связанной с гипервершиной v_i , находим количество еще не окрашенных гипервершин, а также количество форточек в имеющемся на данный момент расписании для рассматриваемой группы студентов. Если количество форточек равно количеству неокрашенных гипервершин, то среди всех допустимых цветов из множества N разрешаем те цвета, которые соответствуют форточкам. Если имеются допустимые цвета, то выбираем один из них. Если нет допустимых цветов, то выбираем один из первоначально допустимых цветов. Окрашиваем гипервершину v_i в выбранный цвет и помечаем ресурсы, запрашиваемые гипервершиной v_i , как используемые.

Среди оставшихся конфликтных гипервершин случайным образом выбирается гипервершина v_j , для которой повторяются все действия предыдущего шага. Затем выполняется шаг 2.

Заключение. В результате проведенного исследования разработана математическая модель решения задачи составления расписания учебных занятий, которая учитывает ряд обязательных ограничений и критериев оптимизации. Данная модель имеет ряд отличий:

- учитывает равномерность распределения занятий в течение дня; возможность проведения занятий в определенную смену;

- минимизирует количество дней, занятых учебной нагрузкой;

- обеспечивает выполнение всех аудиторных работ в течение всего планового периода и др.

Предложен алгоритм решения задачи составления расписания учебных занятий методом раскраски вершин гиперграфа. Разработанные математическая модель и алгоритм решения задачи могут быть положены в основу автоматизированной системы составления расписания учебного заведения.

ЛИТЕРАТУРА

1. Лазарев, А.А. Методы и алгоритмы решения задач теории расписаний для одного и нескольких приборов и их применение для задач комбинаторной оптимизации: дис. ... д-ра физ.-мат. наук / А.А. Лазарев. – М.: Ин-т проблем управления им. В.А. Трапезникова РАН, 1997. – 413 с.
2. Танаев, В.С. Теория расписаний. Многостадийные системы / В.С. Танаев, Ю.Н. Сотсков, В.А. Струевич. – М.: Наука, 1989. – 328 с.
3. Гафаров, Е.Р. Задачи теории расписаний. Алгоритмы и применение / Е.Р. Гафаров // Современные проблемы фундаментальных и прикладных наук: тр. 49-й науч. конф. МФТИ: Управление и прикладная математика. – Москва-Долгопрудный, 2006. – С. 82 – 83.
4. Баронов, В.В. Автоматизация управления предприятием / В.В. Баронов. – М.: Инфра-М, 2000.
5. Логоша, Б.А. Комплекс моделей и методов оптимизации расписания занятий в вузе / Б.А. Логоша, А.В. Петропавловская // Экономика и математические методы. – 1993. – Т. 29, № 4.
6. Строкина, Ю.Г. Алгоритмические процедуры формирования гетерогенных расписаний для производственных систем: дис. ... канд. техн. наук / Ю.Г. Строкина. – Уфа: УГАТУ, 1997. – 150 с.
7. Волков, И.К. Исследование операций: учебник для вузов / И.К. Волков, Е.А. Загоруйко; под ред. В.С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2000. – 436 с.
8. Ху, Т. Целочисленное программирование и потоки в сетях / Т. Ху; пер. с англ. – М.: Мир, 1974. – 520 с.
9. Таха, Хэмди А. Введение в исследование операций / Хэмди А. Таха. – 6-е изд.; пер. с англ. – М.: Издат. дом «Вильямс», 2001.
10. Сигал, И.Х. Введение в прикладное дискретное программирование: модели и вычислительные алгоритмы: учеб. пособие / И.Х. Сигал, А.П. Иванова. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 240 с.
11. Расписание Про 2.3 [Electronic resource]. – Mode of access: <http://www.softsearch.ru/programs/28-705-raspisanie-pro-download.shtml>.
12. Magic Timetable Master [Electronic resource]. – Mode of access: <http://www.imagictimetablessoftware.com/>
13. TimeTabler [Electronic resource]. – Mode of access: <http://www.timetabler.com/>
14. Burke, E.K. A University Timetabling System Based on Graph Coloring and Constraint Manipulation / E.K. Burke, D.G. Elliman, R.F. Weare // Journal Of Research on Computing in Education. – 1993.
15. Бабкин, Э.А. Проектирование и реализация алгоритмов составления учебного расписания на основе многоагентных технологий / Э.А. Бабкин, И.М. Ретинский // Технические, программные и математические аспекты управления сложными распределенными системами: материалы науч.-техн. конф. – Нижний Новгород, 2003. – С. 10–12.
16. Безгинов, А.Н. Обзор существующих методов составления расписания / А.Н. Безгинов, С.Ю. Трегубов // Информационные технологии и программирование: межвуз. сб. ст. – Вып. 2(14). – М.: МГИУ, 2005.

17. Ерунов, В.П. Формирование оптимального расписания учебных занятий в вузе / В.П. Ерунов, И.И. Морковин // Вестн. ОГУ. – 2001. – № 3.
18. Каширина, И.Л. Генетический алгоритм решения квадратичной задачи о назначениях специального вида / И.Л. Каширина // Вестн. ВГУ. – 2003. – № 1.
19. Костенко, В.А. Локально-оптимальные алгоритмы построения расписаний, основанные на использовании сетей Хопфилда / В.А. Костенко, А.В. Винокуров // Программирование. – 2003. – № 4. – С. 27 – 40.
20. Костин, С.А. Модели и алгоритмы глобальной оптимизации первоначального расписания вуза / С.А. Костин, Н.Н. Клеванский // Информационные технологии в образовании: XIV междунар. конф.-выст.: материалы конф. – М.: МИФИ, 2004.
21. Маслов, М.Г. Эвристический алгоритм решения задачи составления расписания учебных занятий в вузе / М.Г. Маслов // Математические методы в технике и технологиях: сб. тр. XV междунар. науч. конф., Тамбов, 2 – 4 июня 2002 г.: в 10-ти т. – Тамбов, 2002. – Т. 9. – С. 86 – 88.
22. Система моделей и методов рационального планирования и организации учебного процесса в вузе / под ред. В.В. Гусева, Н.Я. Краснера. – Воронеж: ВГУ, 1984. – 152 с.
23. Ханов, Г.В. Автоматизация составления расписаний с учетом неопределенности / Г.В. Ханов, Е.В. Алабужев // Информационные технологии в образовании, технике и медицине: материалы междунар. конф.: в 3-х т. Т. 1 / ВолгГТУ. – Волгоград, 2004.
24. Сотсков, Ю.Н. Построение расписаний учебных занятий на основе раскраски вершин графа / Ю.Н. Сотсков, С.В. Балтак // Информатика. – 2006. – С. 58 – 69.
25. Льюггер, Джордж, Ф. Искусственный интеллект: стратегии и методы решения сложных проблем / Джордж, Ф. Льюггер. – 4-е изд.; пер. с англ. – М.: Издат. дом «Вильямс», 2003. – 864 с.
26. Кузин, Л.Т. Основы кибернетики / Л.Т. Кузин: в 2 т. Т. 1. Математические основы кибернетики: учеб. пособие для вузов. – 2-е изд., перераб. и доп. – М.: Энергоатомиздат, 1994. – 576 с.
27. Гэри, М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М.: Мир, 1982. – 416 с.
28. Харари, Ф. Теория графов / Ф. Харари. – М.: Мир, 1973. – 300 с.
29. Клеванский, Н.Н. Формирование расписания с использованием динамических критериев загруженности / Н.Н. Клеванский, Е.А. Макарцова // Информационные технологии в образовании: XI междунар. конф.-выст. Ч. IV. – М.: МИФИ, 2001. – С. 139 – 140.
30. Клеванский, Н.Н. Моделирование стратегии формирования расписания занятий вуза средствами реляционной алгебры / Н.Н. Клеванский, Е.А. Макарцова, С.А. Костин // Прикладные проблемы образовательной деятельности: межвуз. сб. науч. тр. – Воронеж: ВГПУ, 2003. – С. 71 – 74.
31. Кристофидес, Н. Теория графов. Алгоритмический подход / Н. Кристофидес. – М.: Мир, 1978. – 432 с.
32. Leighton, F.T. A graph coloring algorithm for large scheduling problems / F.T. Leighton // Journal of Research of the National Bureau of Standards. – 1979. – Vol. 84. – P. 489 – 506.
33. Werra D. de. An introduction to timetabling / D. de Werra // European Journal of Operations Research. – 1985. – Vol. 19. – P. 151 – 162.
34. Tuza Z. Graph colorings with local constraints: a survey / Z. Tuza // Discussions Mathematical Graph Theory. – 1997. – Vol. 17. – P. 101 – 228.

Поступила 20.09.2013

MATHEMATICAL MODEL AND ALGORITHM OF TIMETABLING AT UNIVERSITY

M. DEKANOVA

The mathematical model of drawing up the schedule of university studies on the basis of the hyper graph is presented. In the mathematical model obligatory restrictions are considered: classes, which are conducted with the same group (by the same teacher, in the same audience) must be appointed at various time intervals, all classroom assignments must be done during the planned period, the total number of all types of classes in the allocated period mustn't exceed the available classroom fundre. The mustn't be "free periods" for students, the must be possibility to have dual classes and to conduct certain classes in various days. Optimization carried out according to the following criteria: the uniformity of the distribution of classes during the day, the possibility of carrying out classes in a certain change, the minimization of the number of the days with an academic load, the minimization of "free periods" for teachers, lack of classes in certain days for teachers. The convolution of all optimization criteria is presented one criterion function. The algorithm of the task solution by method of a coloring of hyper tops of the hyper graph is offered.

УДК 621.37.037

СВЕРХРАЗРЕШЕНИЕ ИСТОЧНИКОВ ИЗЛУЧЕНИЯ НА ОСНОВЕ МЕТОДА КЕЙПОНА

канд. техн. наук, доц. **В.Ф. ЯНУШКЕВИЧ, А.О. БЕЗДЕЛЬ**
(Полоцкий государственный университет)

Представлены результаты теоретического анализа сверхразрешения источников излучения с помощью антенных решеток с использованием метода Кейпона. Применены методы цифровой обработки, обеспечивающие эффективные характеристики подавления помех и высокие качества обработки радиолокационных сигналов. Проанализированы функции углового разрешения объектов. Приведены результаты исследований для одного и двух источников излучения. Даны рекомендации по уменьшению вычислительной сложности алгоритмов обработки. Модифицирован метод Кейпона. Результаты исследований могут быть использованы в радиопеленгации и радиолокации.

Введение. Одной из важнейших задач радиолокации является измерение угловых координат источника излучения полезного сигнала, которое основано на определении направления прихода электромагнитных волн, излученных или отраженных целью.

Актуальность исследования обусловлена тем, что к современным радиолокационным станциям предъявляют жесткие ограничения по массе, габаритным размерам и энергопотреблению, которые исключают применение мощных передающих устройств и ограничивают размеры полотна антенной решетки.

Теоретический анализ. Метод Кейпона был предложен в 1969 году для разрешения спектральных компонент дискретного спектра [1]. Используя аналогию частотного и пространственного спектров, метод был предложен также для оценивания углов прихода сигналов с помощью антенной решетки.

Модель сигналов, принимаемых антенной решеткой из N элементов, может быть представлена следующим образом:

$$Z = \sum_{k=1}^J a_k S(\varphi_k) + X, \quad (1)$$

где J – число дискретных источников; a_k , φ_k – комплексная амплитуда и угол прихода волны, соответствующей индексу k ; $S(\varphi_k)$ – N -мерный вектор сигналов, принимаемых решеткой от источника с номером k ; X – N -мерный вектор собственного шума элементов антенной решетки; Z – N -мерный вектор принятого сигнала.

Модель предполагает, что комплексные амплитуды волн различных источников статистически независимы, то есть

$$\langle a_k a_m^* \rangle = \begin{cases} \sigma_k^2 & (k = m); \\ 0 & (k \neq m). \end{cases} \quad (2)$$

Поскольку источники дискретные, то сигналы в элементах решетки для каждого источника предполагаются коррелированными (коэффициент корреляции равен единице). Это значит, что вектор сигнала $S(\varphi_k)$ каждого источника является детерминированным и описывает одновременно характер волнового фронта и конфигурацию антенной решетки. Шумы в её элементах предполагаются некоррелированными и равной мощности. Задача ставится следующим образом: необходимо найти весовой вектор W , который минимизирует среднюю мощность на выходе антенной решетки при условии, что для некоторого угла прихода φ коэффициент передачи решетки фиксирован и равен, например, единице. Математически эту задачу запишем следующим образом:

$$\min_w \langle |W^H Z|^2 \rangle \text{ при условии } W^H S(\varphi) = 1. \quad (3)$$

Это задача на условный экстремум.

Для ее решения мы должны составить функционал Лагранжа в виде

$$\Phi(W) = \langle |W^H Z|^2 \rangle - \chi (W^H S(\varphi) - 1). \quad (4)$$

Здесь χ – неопределенный множитель Лагранжа.

Первое слагаемое в (4) запишем так

$$\langle |W^H Z|^2 \rangle = W^H \langle ZZ^H \rangle W = W^H M W, \quad (5)$$

где $M = \langle ZZ^H \rangle$ – корреляционная матрица сигналов на входе антенной решетки.

Для принятой модели сигналов (1) нетрудно вычислить корреляционную матрицу при условии (2). В результате получим

$$M = \sum_{k=1}^J \sigma_k^2 S(\varphi_k) S^H(\varphi_k) + \sigma^2 E. \quad (6)$$

Здесь σ^2 – средняя мощность шума в одном элементе антенной решетки.

С учетом (5) выражение (4) преобразуется к виду

$$\Phi(W) = W^H M W - \chi (W^H S(\varphi) - 1). \quad (7)$$

Градиент этого функционала приравняем к нулю и получим следующее равенство:

$$M W = \lambda S(\varphi). \quad (8)$$

Отсюда находим весовой вектор

$$W = \chi M^{-1} S(\varphi), \quad (9)$$

где M^{-1} – обратная корреляционная матрица входных сигналов.

Теперь весовой вектор (9) следует подставить в требуемое условие (3), и тогда неопределенный множитель χ можно найти в виде

$$\chi = [S^H(\varphi) M^{-1} S(\varphi)]^{-1}. \quad (10)$$

Окончательное решение задачи получается, если (10) подставить в (1). В результате весовой вектор, минимизирующий выходную среднюю мощность по критерию Кейпона, будет иметь вид

$$W = \frac{1}{S^H(\varphi) M^{-1} S(\varphi)} M^{-1} S(\varphi). \quad (11)$$

В оптимальном состоянии, когда весовые коэффициенты антенной решетки установлены в соответствии с (11), средняя выходная мощность находится подстановкой (11) в (5). Эта величина является разрешающей функцией, которую мы обозначим $\eta_1(\varphi)$.

Таким образом, для метода Кейпона разрешающая функция получается равной

$$\eta_1(\varphi) = \frac{1}{S^H(\varphi) M^{-1} S(\varphi)}. \quad (12)$$

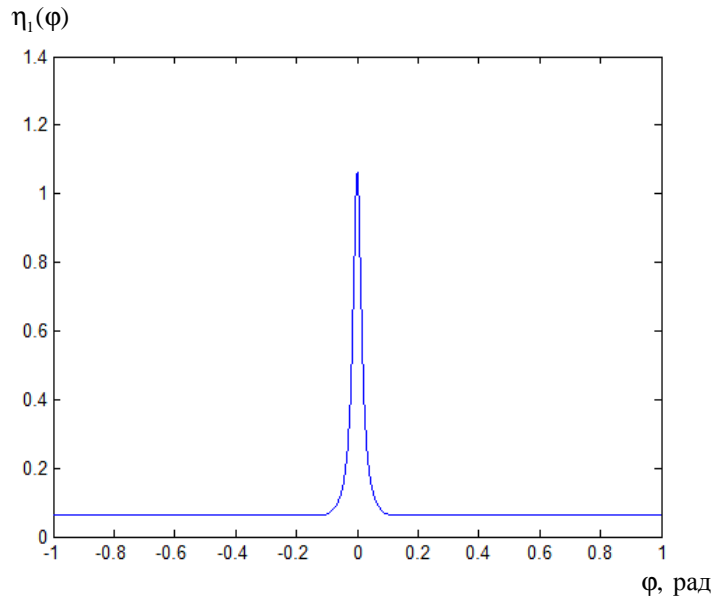
Средняя выходная мощность может быть измерена. Поэтому эта величина представляет интерес с точки зрения углового разрешения источников.

1. Допустим, что в пространстве имеется один источник. Этот пример полезно рассмотреть, хотя в случае одного источника вопрос об угловом разрешении не имеет смысла. В данном случае корреляционная матрица (6) принимает вид

$$M = \sigma_1^2 S(\varphi_1) S^H(\varphi_1) + \sigma^2 E. \quad (13)$$

На рисунке 1 изображена функция (12). Расчеты выполнялись для линейной эквидистантной антенной решетки с числом элементов $N = 16$ и полуволновым расстоянием между элементами.

Предполагалось, что плоская волна от источника приходит по направлению нормали к антенной решетке, т.е. $\varphi_1 = 0$. Средняя мощность полезного сигнала и собственного шума в каждом элементе принимались соответственно $\sigma_1^2 = 1$, $\sigma^2 = 1$.

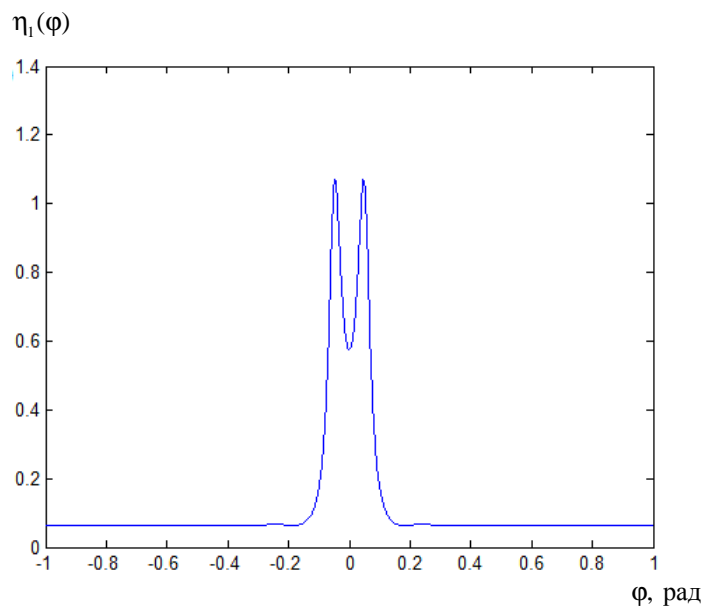
Рис. 1. Средняя мощность на выходе антенной решетки от угла φ

2. Допустим теперь, что в пространстве имеется два источника. Тогда корреляционная матрица принятых сигналов (6) имеет вид

$$M = \sigma_1^2 S(\varphi_1) S^h(\varphi_1) + \sigma_2^2 S(\varphi_2) S^h(\varphi_2) + \sigma^2 E. \quad (14)$$

Предположим, что оба источника имеют одинаковые средние мощности в каждом элементе антенной решетки. Мы также полагаем, что отношение мощности сигнала одного источника к мощности собственного шума в каждом элементе равно единице, т.е. $\sigma_1^2 = \sigma_2^2 = \sigma^2 = 1$; углы прихода волн полагаем равными: $\varphi_1 = \pi/64$, $\varphi_2 = -\pi/64$.

Как прежде, предположим, что измерение углов прихода волн выполняется с помощью линейной эквидистантной антенной решетки из 16 элементов ($N=16$) и полуволновым межэлементным расстоянием ($d/\lambda = 0,5$). Попробуем применить метод Кейпона (12), который не требует априорного знания числа источников, а опирается только на знание корреляционной матрицы входных сигналов. Результаты приведены на рисунке 2.

Рис. 2. Средняя мощность на выходе антенной решетки в зависимости от угла φ при $\sigma_1^2 = \sigma_2^2 = \sigma^2 = 1$

Метод Кейпона имеет следующие преимущества:

- не требует априорного знания количества источников излучения;
- меньшая вычислительная сложность алгоритма, чем у метода максимального правдоподобия;
- имеет большую разрешающую способность, чем метод максимального правдоподобия;
- меньшее влияние случайных амплитудно-фазовых ошибок.

Однако методу Кейпона присущи и недостатки:

- худшие точностные показатели, чем у метода «теплового шума»;
- данный метод не позволяет оценить количество источников излучения.

Для уменьшения вычислительной сложности данного алгоритма можно воспользоваться симметричностью корреляционной матрицы, представленной на рисунке 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617	-0.2457	-0.5472	-0.8358	-1.1046	-1.3473
2	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617	-0.2457	-0.5472	-0.8358	-1.1046
3	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617	-0.2457	-0.5472	-0.8358
4	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617	-0.2457	-0.5472
5	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617	-0.2457
6	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676	0.0617
7	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648	0.3676
8	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463	0.6648
9	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053	0.9463
10	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359	1.2053
11	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324	1.4359
12	-0.2457	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903	1.6324
13	-0.5472	-0.2457	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059	1.7903
14	-0.8358	-0.5472	-0.2457	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763	1.9059
15	-1.1046	-0.8358	-0.5472	-0.2457	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3	1.9763
16	-1.3473	-1.1046	-0.8358	-0.5472	-0.2457	0.0617	0.3676	0.6648	0.9463	1.2053	1.4359	1.6324	1.7903	1.9059	1.9763	3

Рис. 3. Пример корреляционной матрицы

При вычислении корреляционной матрицы размерности N необходимо вычислить N^2 элементов, при использовании же симметричности нам необходимо $2N - 1$ вычислений.

Также можно использовать симметричность корреляционной матрицы при вычислении обратной корреляционной матрицы, которая также симметрична (рис. 4). Ее симметричность можно использовать при вычислении разрешающей функции для метода Кейпона, что позволит уменьшить количество вычислений в 2 раза.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.8549	-0.1400	-0.1315	-0.1200	-0.1056	-0.0887	-0.0697	-0.0491	-0.0273	-0.0049	0.0177	0.0398	0.0610	0.0808	0.0986	0.1141
2	-0.1400	0.8626	-0.1316	-0.1227	-0.1109	-0.0964	-0.0797	-0.0610	-0.0410	-0.0199	0.0016	0.0230	0.0439	0.0638	0.0822	0.0986
3	-0.1315	-0.1316	0.8714	-0.1225	-0.1135	-0.1018	-0.0877	-0.0716	-0.0537	-0.0346	-0.0146	0.0057	0.0258	0.0454	0.0638	0.0808
4	-0.1200	-0.1227	-0.1225	0.8807	-0.1134	-0.1048	-0.0937	-0.0804	-0.0651	-0.0484	-0.0305	-0.0118	0.0071	0.0258	0.0439	0.0610
5	-0.1056	-0.1109	-0.1135	-0.1134	0.8893	-0.1053	-0.0974	-0.0873	-0.0750	-0.0610	-0.0456	-0.0291	-0.0118	0.0057	0.0230	0.0398
6	-0.0887	-0.0964	-0.1018	-0.1048	-0.1053	0.8967	-0.0989	-0.0921	-0.0832	-0.0722	-0.0596	-0.0456	-0.0305	-0.0146	0.0016	0.0177
7	-0.0697	-0.0797	-0.0877	-0.0937	-0.0974	-0.0989	0.9020	-0.0948	-0.0893	-0.0818	-0.0722	-0.0610	-0.0484	-0.0346	-0.0199	-0.0049
8	-0.0491	-0.0610	-0.0716	-0.0804	-0.0873	-0.0921	-0.0948	0.9048	-0.0934	-0.0893	-0.0832	-0.0750	-0.0651	-0.0537	-0.0410	-0.0273
9	-0.0273	-0.0410	-0.0537	-0.0651	-0.0750	-0.0832	-0.0893	-0.0934	0.9048	-0.0948	-0.0921	-0.0873	-0.0804	-0.0716	-0.0610	-0.0491
10	-0.0049	-0.0199	-0.0346	-0.0484	-0.0610	-0.0722	-0.0818	-0.0893	-0.0948	0.9020	-0.0989	-0.0974	-0.0937	-0.0877	-0.0797	-0.0697
11	0.0177	0.0016	-0.0146	-0.0305	-0.0456	-0.0596	-0.0722	-0.0832	-0.0921	-0.0989	0.8967	-0.1053	-0.1048	-0.1018	-0.0964	-0.0887
12	0.0398	0.0230	0.0057	-0.0118	-0.0291	-0.0456	-0.0610	-0.0750	-0.0873	-0.0974	-0.1053	0.8893	-0.1134	-0.1135	-0.1109	-0.1056
13	0.0610	0.0439	0.0258	0.0071	-0.0118	-0.0305	-0.0484	-0.0651	-0.0804	-0.0937	-0.1048	-0.1134	0.8807	-0.1225	-0.1227	-0.1200
14	0.0808	0.0638	0.0454	0.0258	0.0057	-0.0146	-0.0346	-0.0537	-0.0716	-0.0877	-0.1018	-0.1135	-0.1225	0.8714	-0.1316	-0.1315
15	0.0986	0.0822	0.0638	0.0439	0.0230	0.0016	-0.0199	-0.0410	-0.0610	-0.0797	-0.0964	-0.1109	-0.1227	-0.1316	0.8626	-0.1400
16	0.1141	0.0986	0.0808	0.0610	0.0398	0.0177	-0.0049	-0.0273	-0.0491	-0.0697	-0.0887	-0.1056	-0.1200	-0.1315	-0.1400	0.8549

Рис. 4. Пример обратной корреляционной матрицы

Также для увеличения быстродействия можно предложить модификацию метода Кейпона, в которой антенная решетка разделяется на две подрешетки. Корреляционная матрица размерностью N разделяется на две размерностью $N/2$, что позволяет уменьшить количество элементов корреляционной матрицы до $N^2/2$ и количество вычислений в 2 раза. Однако применение данной модификации приводит к незначительному ухудшению разрешающей способности, тем не менее данный метод обладает лучшей разрешающей способностью, чем оригинальный метод Кейпона примерно с такой же вычислительной сложностью, что можно увидеть из рисунка 5.

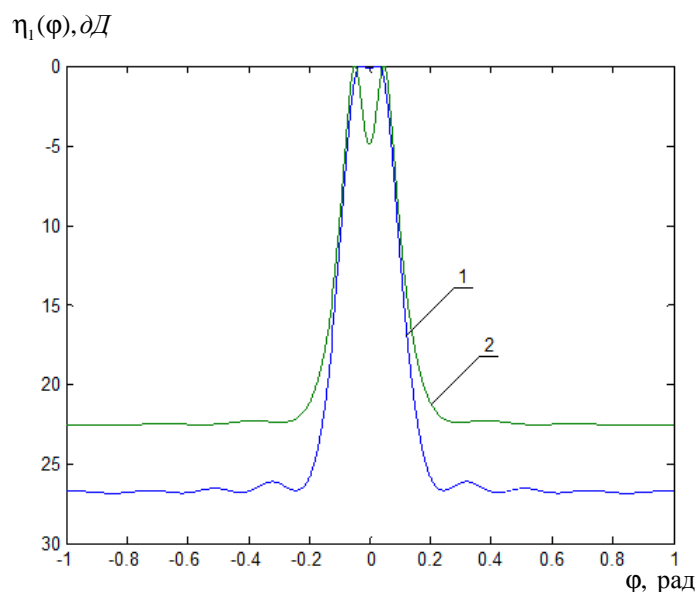


Рис. 5. Зависимости $\eta_1(\varphi) = \psi(\varphi)$ для оригинального (кривая 1) и модифицированного (кривая 2) метода Кейпона

Заключение. Проведен анализ метода Кейпона для получения сверхразрешения. Разработаны рекомендации, основанные на симметричности матриц, позволяющие уменьшить вычислительную сложность алгоритмов, что позволяет упростить цифровую реализацию данных алгоритмов. Представлена модификация для метода Кейпона, позволяющая уменьшить вычислительную сложность данного алгоритма. Результаты исследований могут быть использованы в радиопеленгации и радиолокации.

ЛИТЕРАТУРА

1. Кейпон, Дж. Пространственно-временной спектральный анализ с высоким разрешением / Дж. Кейпон // ТИИЭР. – 1969. – Т. 57, № 8. – С. 59 – 69.
2. Ермолаев, В.Т. Оценивание параметров сигналов, принимаемых антенной решеткой / В.Т. Ермолаев, А.Г. Флакман, А.А. Анурын // Изв. вузов. Радиофизика. – 1996. – Т. 39, № 9. – С. 1144 – 1160.

Поступила 05.09.2013

SUPERRESOLUTION OF LIGHT SOURCES ON THE BASIS OF KEYPON'S METHOD

V. YANUSHKEVICH, A. BEZDEL

The results of theoretical analysis of superresolution of light sources using antenna arrays with the help of Keypon's method have been presented. Digital techniques to ensure effective noise reduction features and high-quality processing of radar signals have been applied. The function of the angular resolution of authorization objects is analyzed. The results of studies for one or two light sources are given. Recommendations are given to reduce the computational complexity of algorithms. The Keypon's method has been modified. The results of the research can be used in radio position finding and radiolocation.

УДК 621.397

КОМПЕНСАЦИЯ ДВИЖЕНИЯ В ВИДЕОКОДЕКАХ НА ОСНОВЕ ОПТИМАЛЬНЫХ РАЗБИЕНИЙ КОДИРУЕМЫХ БЛОКОВ

д-р техн. наук, доц. С.В. ДВОРНИКОВ

(Военная академия связи им. Маршала Советского Союза С.М. Буденного, Санкт-Петербург);

В.В. ЦВЕТКОВ, А.А. УСТИНОВ

(ЗАО «НПФ «ТИРС», Санкт-Петербург)

Предлагается метод компенсации движения в видеокодеках на основе оптимальных разбиений кодируемых блоков. Сформулирована задача разбиения кодируемых трехмерных фрагментов подвижных изображений на непересекающиеся группы. Доказывается, что снижение межкадровых различий внутри каждой группы приводит к уменьшению числа значимых коэффициентов трехмерного преобразования. Обосновывается, что уменьшение числа значимых коэффициентов преобразования приводит к увеличению коэффициента сжатия.

Теоретические основы сжатия подвижных изображений на основе трехмерных ортогональных преобразований

Прямое трехмерное дискретное косинусное преобразование (ДКП-3) для куба размером $N \times N \times N$ задается следующим образом [1]:

$$F(i, j, k) = \sqrt{\frac{8}{N^3}} c(i)c(j)c(k) \sum_{z=0}^{N-1} \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y, z) \cos \left[\frac{(2x+1)\pi i}{2N} \right] \cos \left[\frac{(2y+1)\pi j}{2N} \right] \cos \left[\frac{(2z+1)\pi k}{2N} \right],$$

где $f(x, y, z)$ – значение яркостной или цветоразностной компонент пикселя с координатами $x, y, z \in [0, \dots, N-1]$; $F(i, j, k)$ – коэффициент преобразования с координатами $i, j, k \in [0, \dots, N-1]$, функция $c(k)$ определяется как

$$c(k) = \begin{cases} 1/2, & k = 0; \\ 1, & k \neq 0. \end{cases}$$

Обратное дискретное косинусное преобразование вычисляется следующим образом:

$$f(x, y, z) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sqrt{\frac{8}{N^3}} c(i)c(j)c(k) F(i, j, k) \cos \left[\frac{(2y+1)\pi j}{2N} \right] \cos \left[\frac{(2z+1)\pi k}{2N} \right].$$

Свойства трехмерного дискретного косинусного преобразования

К основным свойствам ДКП-3 следует отнести такие, как:

- локализация большей части энергии сигнала в небольшом числе коэффициентов преобразования. Данное свойство позволяет исключить наименее значимые коэффициенты из рассмотрения при кодировании с потерями [1];

- ДКП-3 представляет собой ортогональное преобразование, которое может быть реализовано путем последовательного выполнения одномерных ДКП по строкам, столбцам и оси времени [1].

Данные свойства определили широкое применение ДКП-3 при сжатии подвижных изображений. Типовая схема кодека на основе ДКП-3 предполагает выполнение следующих основных операций: входная обработка, ДКП-3, квантование коэффициентов преобразования, энтропийное кодирование. На приеме данные процедуры выполняются в обратной последовательности [1].

Однако в условиях кодирования высокодинамичных сцен ДКП-3 может быть не оптимальным в смысле концентрации энергии в незначительном числе коэффициентов, расположенных на временной оси (в различных кадрах кодируемого блока).

Для примера на рисунке показаны квантованные коэффициенты ДКП-3 для кодируемого блока размером $8 \times 8 \times 8$. Данный блок был намеренно выбран из высокодинамичной области кадров тестового подвижного изображения «foreman».

Анализ представленных коэффициентов ДКП-3 показывает наличие ненулевых коэффициентов в каждой плоскости куба $8 \times 8 \times 8$.

Наиболее распространенным подходом к устранению данного недостатка является изменение размеров кодируемых трехмерных блоков [2 – 8]. В основе данного подхода лежит оценка степени подвижности каждого трехмерного кодируемого куба по нескольким фиксированным градациям. Далее для каждой градации определяется оптимальный размер кодируемого куба, например, $16 \times 16 \times 1$, $16 \times 16 \times 8$

или $8 \times 8 \times 8$. Недостатки такого подхода очевидны. Во-первых, градация на заданное число уровней подвижности не вполне обоснованна и носит эмпирический характер. Во-вторых, фиксированное число уровней градации не может описать всего многообразия степени изменения сцен в трехмерных кодируемых фрагментах.

<p>a)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>183</td><td>-3</td><td>0</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-12</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	183	-3	0	-1	0	0	0	0	-12	-1	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<p>б)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>3</td><td>5</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-11</td><td>-1</td><td>0</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	3	5	0	1	0	0	0	0	-11	-1	0	-1	0	0	0	0	-1	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
183	-3	0	-1	0	0	0	0																																																																																																																																		
-12	-1	0	0	0	0	0	0																																																																																																																																		
1	0	0	-1	0	0	0	0																																																																																																																																		
-1	-1	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
3	5	0	1	0	0	0	0																																																																																																																																		
-11	-1	0	-1	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
-1	-1	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
<p>в)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>32</td><td>2</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-10</td><td>-2</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-2</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>1</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	32	2	-1	0	0	0	0	0	-10	-2	1	0	0	0	0	0	-2	-1	0	0	0	0	0	0	-1	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	<p>г)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>17</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>-2</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	17	2	0	0	0	0	0	0	-1	2	0	0	0	0	0	0	0	-2	-1	0	0	0	0	0	-1	1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0								
32	2	-1	0	0	0	0	0																																																																																																																																		
-10	-2	1	0	0	0	0	0																																																																																																																																		
-2	-1	0	0	0	0	0	0																																																																																																																																		
-1	1	-1	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
17	2	0	0	0	0	0	0																																																																																																																																		
-1	2	0	0	0	0	0	0																																																																																																																																		
0	-2	-1	0	0	0	0	0																																																																																																																																		
-1	1	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
<p>д)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>16</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>1</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>-1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	16	0	0	0	0	0	0	0	3	1	-1	0	0	0	0	0	-1	1	0	0	0	0	0	0	-1	-1	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<p>е)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>2</td><td>-1</td><td>2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>4</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-3</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	2	-1	2	0	0	0	0	0	4	0	0	1	0	0	0	0	-3	1	1	0	0	0	0	0	-1	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0								
16	0	0	0	0	0	0	0																																																																																																																																		
3	1	-1	0	0	0	0	0																																																																																																																																		
-1	1	0	0	0	0	0	0																																																																																																																																		
-1	-1	1	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
2	-1	2	0	0	0	0	0																																																																																																																																		
4	0	0	1	0	0	0	0																																																																																																																																		
-3	1	1	0	0	0	0	0																																																																																																																																		
-1	0	1	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
<p>ж)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>4</td><td>-2</td><td>-1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>4</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-2</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	4	-2	-1	1	0	0	0	0	4	1	1	0	0	0	0	0	-1	1	1	0	0	0	0	0	-2	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<p>з)</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td>1</td><td>-3</td><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>-1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	1	-3	-1	0	0	0	0	0	2	1	0	1	0	0	0	0	-1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0								
4	-2	-1	1	0	0	0	0																																																																																																																																		
4	1	1	0	0	0	0	0																																																																																																																																		
-1	1	1	0	0	0	0	0																																																																																																																																		
-2	-1	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
1	-3	-1	0	0	0	0	0																																																																																																																																		
2	1	0	1	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
-1	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		
0	0	0	0	0	0	0	0																																																																																																																																		

Пример коэффициентов ДКП-3 для кодируемого блока размером $8 \times 8 \times 8$:
 а – 1-го слоя куба; б – 2-го слоя куба; в – 3-го слоя куба; г) 4-го слоя куба;
 д) 5-го слоя куба; е – 6-го слоя куба; ж – 7-го слоя куба ; з – 8-го слоя куба

В следующем пункте рассмотрим развитие данной идеи на основе оптимального разбиения кодируемого блока на непересекающиеся наборы кадров.

Формализация задачи поиска оптимального разбиения трехмерных кубов на непересекающиеся области

Пусть на вход кодера поступают L кадров подвижного изображения размером $P_1 \times N_1$ пикселей. Разделим видеоданные на непересекающиеся области (кубы) размером $P \times N \times L$ пикселей в виде трехмерных матриц $A_{P \times N \times L}$. С целью упрощения дальнейших рассуждений трехмерную матрицу $A_{P \times N \times L}$ представим множеством из L одномерных векторов размерностью $P \times N$ элементов: $A_{P \times N \times L} \rightarrow \{\vec{A}_1, \vec{A}_2, \dots, \vec{A}_L\}$. Отметим, что любой i -й вектор сформированного множества из L векторов получен путем последовательной N -кратной канкатенации P -мерных столбцов i -го слоя матрицы $A_{P \times N \times L}$ в единый вектор размерности $P \times N$. Используя терминологию линейной алгебры, множество $\{\vec{A}\} = \{\vec{A}_1, \vec{A}_2, \dots, \vec{A}_L\}$ будем рассматривать как множество из L точек в евклидовом пространстве размерности $P \times N$ ($E^{P \times N}$).

Разделим исходное множество $\{\vec{A}\}$ на R подмножеств $\{\vec{A}\}_1, \{\vec{A}\}_2, \dots, \{\vec{A}\}_R$ так, чтобы выполнялись условия:

$$\{\vec{A}\}_1 \cup \{\vec{A}\}_2 \cup \dots \cup \{\vec{A}\}_R = \{\vec{A}\};$$

$$\{\vec{A}\}_1 \cap \{\vec{A}\}_2 \cap \dots \cap \{\vec{A}\}_R = \{0\}.$$

Первое условие обеспечивает объединение R подмножеств в единое множество $\{\vec{A}\}$. Второе условие обеспечивает то, что R образованных подмножеств являются непересекающимися.

Исходя из необходимости достижения максимального сжатия, будем считать некоторое разбиение множества $\{\vec{A}\}$ оптимальным, если суммарное число ненулевых коэффициентов преобразования после выполнения ДКП-3 над каждым образованным подмножеством является оптимальным.

Введем функцию $f_{\text{ДКП-3}}(\{\vec{A}\}_i)$, которая определяет число ненулевых (значимых) коэффициентов преобразования, полученных после выполнения ДКП-3 над i -м подмножеством. Тогда формально задачу поиска оптимального разбиения можно записать следующим образом:

$$\sum_{k=1}^R f_{\text{ДКП-3}}(\{\vec{A}\}_k) \rightarrow \min_{\{\vec{A}\}_k, \forall k=1, R}. \quad (1)$$

Задача (1) предполагает поиск некоторого разбиения на конечном множестве возможных разбиений, для которого суммарное число ненулевых коэффициентов преобразования по каждому подмножеству разбиения минимально. В этом смысле данная задача является целочисленной. Кроме того, в ходе решения задачи необходимо для каждого разбиения выполнять ДКП-3 по всем сформированным подмножествам. В связи с этим метод решения данной задачи, основанный на переборе возможных разбиений с последующим выполнением ДКП-3 для найденных подмножеств, является сложным.

Поэтому рассмотрим постановку задачи поиска оптимального разбиения, которая не требует вычисления ДКП-3 для оценки количества значимых коэффициентов.

Известно, что ДКП-3 является наиболее эффективным, если корреляция между соседними элементами внутри обрабатываемого куба является относительно высокой [9]. Коэффициент корреляции между соседними элементами можно сопоставить со степенью их похожести или близости. Выберем в качестве меры близости средний квадрат ошибки (СКО). Тогда можно утверждать, что чем меньше СКО между элементами, тем выше коэффициент корреляции между ними. Следовательно, для решения задачи (1) без выполнения ДКП-3 необходимо найти такое разбиение исходного множества на R непересекающихся подмножеств, при котором суммарный СКО между элементами каждого из подмножеств был бы минимальным.

Для формальной постановки такой задачи необходимо математически описать следующее:

- способ разбиения исходного множества $\{\vec{A}\}$ на R непересекающихся подмножеств;
- способ вычисления СКО между элементами каждого из подмножеств.

Разбиение исходного множества $\{\vec{A}\}$ на R непересекающихся подмножеств зададим в виде матрицы X размером $R \times L$ элементов. На элементы матрицы X наложим ограничения:

$$X(i, j) \in \{1, 0\}; \quad (2)$$

$$\sum_{i=1}^R X(i, j) = 1, \quad \forall j = \overline{1, L}. \quad (3)$$

Ограничение (2) формально описывает способ разбиения в следующих случаях:

$$\text{если } X(i, j) = 1, \text{ то } \vec{A}_j \in \{\vec{A}\}_i;$$

$$\text{если } X(i, j) = 0, \text{ то } \vec{A}_j \notin \{\vec{A}\}_i.$$

Ограничение (3) обеспечивает выполнение условия непересекаемости подмножеств: каждый элемент $\vec{A}_j, j = 1, 2, \dots, L$ может принадлежать только одному подмножеству.

Для вычисления СКО между элементами определим средний элемент i -го подмножества в виде [2]

$$\hat{A}_i = \frac{[A]\bar{X}_i^T}{\bar{X}_i \bar{1}},$$

где $[A]$ – матрица размером $P \times L$ элементов; \bar{X}_i – i -я строка матрицы X . Отметим, что каждый j -й столбец матрицы $[A]$ (\bar{A}_j), является j -м элементом исходного множества $\{\bar{A}\}$.

Если j -й элемент принадлежит i -му подмножеству, то данный элемент определим в виде $\bar{A}_{j \in \{\bar{A}\}_i}$. Тогда СКО в i -й группе определим следующим образом:

$$CKO_i = \sum_{j=1}^{n_i} \left(\bar{A}_{j \in \{\bar{A}\}_i} - \frac{[A]\bar{X}_i^T}{\bar{X}_i \bar{1}} \right)^T \left(\bar{A}_{j \in \{\bar{A}\}_i} - \frac{[A]\bar{X}_i^T}{\bar{X}_i \bar{1}} \right), \quad (4)$$

где n_i – число элементов множества $\{\bar{A}\}$, принадлежащих i -му подмножеству; $\bar{1}$ – единичный вектор-столбец размерности $L \times 1$.

Исходя из физического смысла матрицы X , принадлежность j -го элемента i -му множеству можно представить в виде произведения элемента на $X(i, j)$. В этом случае суммарное значение СКО по всем R подмножествам запишем в виде

$$CKO = \sum_{i=1}^R \sum_{j=1}^L \left(\bar{A}_j X(i, j) - X(i, j) \frac{[A]\bar{X}_i^T}{\bar{X}_i \bar{1}} \right)^T \left(\bar{A}_j X(i, j) - X(i, j) \frac{[A]\bar{X}_i^T}{\bar{X}_i \bar{1}} \right). \quad (5)$$

Раскрыв скобки в выражении (5), получим:

$$CKO = \sum_{i=1}^R \sum_{j=1}^L \left(\bar{A}_j^T \bar{A}_j X^2(i, j) - 2X^2(i, j) \frac{\bar{A}_j^T [A]\bar{X}_i^T}{\bar{X}_i \bar{1}} + X^2(i, j) \frac{\bar{X}_i [A]^T [A]\bar{X}_i^T}{(\bar{X}_i \bar{1})^2} \right). \quad (6)$$

Проанализируем каждое слагаемое в выражении (6) с учетом суммирования:

- первое слагаемое

$$\sum_{i=1}^R \sum_{j=1}^L \bar{A}_j^T \bar{A}_j X^2(i, j) = \sum_{j=1}^L \bar{A}_j^T \bar{A}_j \sum_{i=1}^R X^2(i, j) = \sum_{j=1}^L \bar{A}_j^T \bar{A}_j,$$

поскольку $\sum_{i=1}^R X^2(i, j) = \sum_{i=1}^R X(i, j) = 1$ в соответствии с ограничениями (2) и (3);

- второе слагаемое

$$2 \sum_{i=1}^R \sum_{j=1}^L X^2(i, j) \frac{\bar{A}_j^T [A]\bar{X}_i^T}{\bar{X}_i \bar{1}} = 2 \sum_{j=1}^L \bar{A}_j^T \sum_{i=1}^R \frac{X^2(i, j) [A]\bar{X}_i^T}{\bar{X}_i \bar{1}} = 2 \sum_{j=1}^L \frac{\bar{X}_j [A]^T [A]\bar{X}_j^T}{\bar{X}_j \bar{1}};$$

- третье слагаемое

$$\sum_{i=1}^R \sum_{j=1}^L X^2(i, j) \frac{\bar{X}_i [A]^T [A]\bar{X}_i^T}{(\bar{X}_i \bar{1})^2} = \sum_{j=1}^L X^2(i, j) \sum_{i=1}^R \frac{\bar{X}_i [A]^T [A]\bar{X}_i^T}{(\bar{X}_i \bar{1})^2} = \sum_{i=1}^R \frac{\bar{X}_i [A]^T [A]\bar{X}_i^T}{(\bar{X}_i \bar{1})},$$

поскольку $\sum_{j=1}^L X^2(i, j) = \bar{X}_i \bar{1}$, при выполнении ограничения (2).

С учетом полученных выражений для 1 – 3 слагаемых выражение (6) окончательно запишем в виде

$$CKO = \sum_{j=1}^L \bar{A}_j^T \bar{A}_j - \sum_{i=1}^R \frac{\bar{X}_i [A]^T [A]\bar{X}_i^T}{(\bar{X}_i \bar{1})}. \quad (7)$$

Как было отмечено ранее, оптимальным разбиением является разбиение, при котором минимизируется величина СКО. Следовательно, задача поиска оптимального разбиения сводится к минимизации вы-

ражения (7) по всем возможным матрицам разбиений. Поскольку первое слагаемое в выражении (7) от искомой переменной не зависит, то окончательно задачу поиска оптимального разбиения можно записать как

$$\sum_{i=1}^R \frac{\vec{X}_i [A]^T [A] \vec{X}_i^T}{(\vec{X}_i \vec{1})} \rightarrow \max_{\vec{X}_i, \forall i=1, R} \quad (8)$$

при ограничениях (2) и (3) на искомые значения элементов матрицы X .

Решением задачи (8) является матрица X , задающая оптимальное разбиение исходного трехмерного куба размером $P \times N \times L$ пикселей на R непересекающихся подмножеств.

Заключение. Для практической реализации алгоритма сжатия подвижных изображений на основе ДКП-3 и поиска оптимального разбиения трехмерных кубов на подмножества необходимо осуществить решение оптимизационной задачи (8). Данная задача относится к классу задач нелинейного целочисленного программирования. Её решение и анализ эффективности предложенного решения компенсации движения требует дальнейших исследований.

ЛИТЕРАТУРА

1. Беляев, Е.А. Сжатие видеoinформации на основе трехмерного дискретного псевдо-косинусного преобразования для энергоэффективных систем видеонаблюдения / Е.А. Беляев, Т.М. Сухов, Н.Н. Шостацкий // Компьютерная оптика. – Т. 34, № 2. – С. 260 – 272, 210.
2. Use of an Adaptive 3D-DCT Scheme for Coding Multiview Stereo Images / N.P. Sgouros [et al.] // IEEE Proceedings of ISSPIT. – 2005. – P. 180 – 185.
3. Chan, Y.L. Variable Temporal-length 3-D Discrete Cosine Transform Coding / Y.L. Chan, W.C. Siu // IEEE Transactions on Image Processing. – 1997. – Vol 6, №. 5. – P. 758 – 763.
4. Koivusaari, J.J. Simplified Three-Dimensional Discrete Cosine Transform Based Video Codec / J.J. Koivusaari, J.H. Takala // SPIE Proc. of Multimedia on Mobile Devices. – 2005. – P. 11 – 21.
5. Dugad, R. A Fast Scheme for Image Size Change in The Compressed Domain / R. Dugad, N. Ahuja // IEEE Transactions on Circuits and Systems for Video Technology. – 2001. – P. 461 – 474.
6. Koivusaari, J.J. Image Coding Using Adaptive Resizing in The Block-DCT Domain / J.J. Koivusaari, J.H. Takala, M. Gabbouj // SPIE Proc. of Multimedia on Mobile Devices II. – 2006. – P. 1 – 9.
7. Use of Adaptive Resizing in 3-D DCT Domain for Video Coding / J. Li [et al.] // Picture Coding Symp., Lisbon, Portugal. – 2007. – P. 7 – 9.
8. Dai, Q. Fast Algorithms for Multidimensional DCT-to-DCT Computation Between a Block and Its Associated Subblocks / Q. Dai, X. Chen, C. Lin // IEEE Trans. Circuits and Syst. for Video Tech. – 2003. – Vol. 13, № 7. – P. 717 – 725.
9. Ахмед, Н. Ортогональные преобразования при обработке цифровых сигналов / Н. Ахмед, К.Р. Рао. – Л.: Связь, 1980. – 248 с.

Поступила 16.05.2013

COMPENSATION OF MOTION IN VIDEO CODECS ON THE BASIS OF OPTIMAL PARTITION OF ENCODED BLOCKS

S. DVORNIKOV, V. TSVETKOV, A. USTINOV

A method for compensation of motion in video codecs on the basis of optimal partition of encoded blocks is proposed. The aim to divide encoded three-dimensional fragments of movable images into non-intersecting groups is formulated. It is proven, that reduction of interframe differences inside of each group leads to reduction of the number of significant coefficients of three-dimensional transformation. It is grounded, that reduction of the number of significant coefficients of transformation leads to increase of the coefficient of compression.

УДК 621.3.072.6

ОПТИМИЗАЦИЯ БАЛАНСОВОЙ НАДЕЖНОСТИ ДЕФИЦИТНОЙ ЭНЕРГОСИСТЕМЫ С УЧЕТОМ МЕЖСИСТЕМНЫХ СВЯЗЕЙ

*Н.В. РАДОМАН, канд. техн. наук, доц. О.И. АЛЕКСАНДРОВ;
канд. техн. наук, доц. Д.Н. СВИРСКИЙ
(Белорусский государственный технологический университет, Минск)*

Рассматривается проблема оптимизации балансовой надежности дефицитной энергосистемы с учетом межсистемных связей, которая описывается с помощью многокритериальной целевой функции, включающей в себя минимум отклонения значений перетоков от запланированных значений, минимум суммарного расхода топлива, минимум потерь мощности и энергии как в энергосистеме, так и в межсистемных линиях электропередач. Оптимизация выполняется с учетом приведенной системы линейных двухсторонних ограничений-неравенств на переменные режима работы электроэнергетической системы. Для экспресс-оценки баланса между производством и потреблением электроэнергии предлагается метод минимизации топливной составляющей издержек электростанциями электроэнергетической системы с одновременной минимизацией потерь активной мощности в сети. Особое внимание уделяется минимизации потерь электроэнергии на транспорт в электрических сетях.

Введение. В настоящее время важной *проблемой* является определение рациональных режимов работы электроэнергетической системы (ЭЭС), т.е. наиболее целесообразных значений внутри- и межсистемных перетоков мощности и энергии в энергообъединении (ЭО), которая подразумевает комплекс оптимальных условий существования режима при соблюдении ряда ограничений.

В реализации транзитного потенциала Республика Беларусь исходит из положений и принципов Европейской Энергетической Хартии, целью которой является создание общего недискриминационного рынка электроэнергии (ЭЭ) на евразийском континенте посредством организации параллельной работы крупных энергетических объединений – UCTE, CENTREL, NORDEL – и объединения энергосистем стран СНГ и ОЭС Балтии (ЭК БРЭЛЛ). Этому способствует участие Беларуси в разработке крупных международных проектов «Восток – Запад в условиях функционирования Балтийского кольца», «Параллельная работа стран СНГ с ОЭС Европы» [1].

Объединенная энергосистема (ОЭС) Республики Беларусь граничит с энергосистемами пяти сопредельных государств (Россия, Литва, Латвия, Украина, Польша), имея с ними межгосударственные перетоки ЭЭ по 38 линиям. В связи с острой проблемой обеспеченности ОЭС энергоресурсами основной задачей является рациональное их использование с учетом получения ЭЭ из смежных, избыточных по мощности энергосистем (ЭС) (номинальный импорт ЭЭ оценивается величиной около 5,0 млрд. кВт·ч/год). Потому взаимодействие ЭС с учетом высоковольтных межсистемных транзитов (МТ) определяет надежность электроснабжения дефицитных регионов и приобретает не только режимный, но и экономический смысл. Поскольку современные питающие и межсистемные линии электропередачи (МЛЭП) представляют собой многоконтурную и сложно-замкнутую сетевую структуру, выбор рациональных режимов эксплуатации энергосистем с учетом МТ представляет довольно трудную задачу, которая имеет ряд принципиальных особенностей.

Постановка задачи. Ведение оптимального режима в ЭЭС подразумевает решение комплекса задач, обеспечивающих минимальные издержки. В число этих задач входят такие, как выбор состава работающего оборудования, распределение нагрузок между станциями с минимизацией расхода топлива и с наименьшими потерями в электрических сетях, а также рациональная работа всего ЭО с учетом МЛЭП. Формулировку данной проблемы можно описать с помощью *многокритериальной целевой функции*, которая включает в себя минимум отклонения значений перетоков от запланированных значений, минимум суммарного расхода топлива, минимум потерь мощности и энергии как в ЭС, так и в МЛЭП. Основные показатели, необходимые для принятия оптимальных решений, были разработаны ранее на основе системного анализа с охватом всех определяющих факторов, т.е. были созданы документы для оперативного использования диспетчерским персоналом смежных ЭС [2; 3]. В нашем случае речь идет о многолетнем оперативно-диспетчерском взаимодействии Системного Оператора Единой энергетической системы России и Объединенного диспетчерского управления Беларуси. Важным элементом такого управления является оптимизация баланса мощностей для межсистемных перетоков в ОЭС, т.е. оптимизация текущего режима. В этом случае задача управления ОЭС в течение определенного интервала (например, суток) распадается на ряд последовательных задач, результаты решения которых в агрегированном виде дают искомый суточный график для ведения режима ОЭС. Если при этом были выполнены ограничивающие условия по изменяющимся параметрам и достигнут минимум затратных средств, то режим ОЭС будет оптимальным.

В данной постановке под задачей оптимизации баланса мощности понимается определение рациональных значений как внутренних, так и межсистемных потоков мощности и энергии для дефицитной ЭС. При дефиците мощности в ОЭС, а в ряде случаев и при нехватке топлива выполняется оптимизация баланса мощностей и выработки ЭЭ в условиях взаимодействия со смежными ЭС, включая рынок зарубежных поставщиков. Для разных часовых интервалов (особенно во время сезона пиковых нагрузок) выполняется закупка ЭЭ в соседних ЭС. В этом случае минимальный уровень заявляемой мощности определяется разностью между системным максимумом потребления и обеспеченной резервом мощностью собственных электростанций системы [4]. Экономически целесообразные величины закупаемой мощности и энергии определялись на основе технико-экономических и режимных расчетов с учетом технических, режимных, директивных и ценовых ограничений. На первом этапе определяется оптимальная загрузка электростанций ЭС для различных часовых интервалов. Получаемое при этом рациональное значение дефицита мощности для разных нагрузок ЭС позволяет определить оптимальное количество покупной ЭЭ на планируемый расчетный период.

Математическая модель. Исходными данными для анализа режимов совместной ОЭС являются расчетные балансы мощности по ЭС, принимаемые на основе прогнозов электропотребления и электрических нагрузок на перспективный период. Исходя из интересов дефицитной ЭС целевая функция в детерминированной постановке для суточного режима может быть представлена в следующем виде:

$$Z = \left\{ \sum_{t=1}^{24} \alpha_{1t} Y_{1t}(p) + \alpha_{2t} Y_{2t}(B) + \alpha_{3t} Y_{3t}(\pi) + \alpha_{4t} Y_{4t}(R) + \alpha_{5t} Y_{5t}(H) + \alpha_{6t} Y_{6t}(3p) + \alpha_{7t} Y_{7t}(P_{omcl}) \right\} \rightarrow \min, \quad (1)$$

где $\alpha_{1t}, \dots, \alpha_{7t}$ – весовые корректирующие коэффициенты стоимости, определяемые на основании экспертных оценок специалистов диспетчерских и экономических служб; $Y_{1t}(p)$ – суммарное отклонение величин перетоков мощности от запланированных значений по контролируемым линиям связи, где организован учет часовых отклонений потребления ЭЭ областными ЭС от заявленных величин; $Y_{2t}(B)$ – суммарный расход топлива на электростанциях; $Y_{3t}(\pi)$ – суммарные потери в ЭО.

При составлении эквивалентной схемы замещения ОЭС основные перетоки мощности от шин передающих подстанций (П/СТ) избыточной ЭС представлены в виде генерирующих узлов, а шины принимающих П/СТ дефицитной ЭС – в виде потребительских узлов. Тогда суммарные потери активной мощности в общем виде могут оцениваться матричным выражением:

$$\pi_p = [P, Q, p, q, U_0] \cdot \mathbf{B}[P, Q, p, q, U_0] \cdot [P, Q, p, q, U_0]^*$$

Здесь P, p – векторы-строки активных мощностей в генерирующих и потребляющих узлах эквивалентной схемы замещения ОЭС; Q, q – векторы-строки реактивных мощностей в генерирующих и потребляющих узлах; U_0 – номинальное напряжение базисного (балансирующего) узла; $\mathbf{B}(P, Q, p, q, U_0)$ – матрица коэффициентов потерь. Функция минимизации потерь

$$\pi_p \rightarrow \min$$

при условии соблюдения двусторонних ограничений для узловых мощностей.

Наименьший на протяжении суток фактический резерв активной мощности в ОЭС $Y_{4t}(R)$ может быть представлен как

$$R = \min [R_0(t) = P(t) - p(t)] \rightarrow \max,$$

где $R_0(t) = P(t) - p(t), t \in T$ – график резерва активной мощности в ОЭС; $P(t)$ – планируемый график изменения располагаемой мощности ОЭС в течение рассматриваемого интервала времени T ; $p(t) = p_{max}^{c.n.}(t) + p_{max}^{обм}(t), t \in T$ – прогнозируемый график суточных максимумов нагрузки ОЭС с учетом потерь в сети на период T ; $p_{max}^{c.n.}(t)$ – прогнозируемый график суточных максимумов собственного потребления с учетом потерь в сети; $p_{max}^{обм}(t)$ – прогнозируемый график выдачи мощности в ОЭС в часы максимума нагрузки данной ЭС.

Тогда с учетом изложенного $Y_{4t} = 1/R$.

Показатель системной надежности $Y_{5t}(H)$, который в самом общем виде может быть записан как показатель эффективности работы ОЭС:

$$F = \sum_{l=1}^N A_l(t) - \Delta A_{\Sigma}(t) / \sum_{l=1}^N A_l(t), l \in N, t \in T,$$

где $\sum_{l=1}^N A_l(t)$ – суммарная энергия, потребляемая элементами системы, причем в зависимости от типа рассматриваемого элемента она может быть генерируемая, передаваемая и потребляемая; $\Delta A_{\Sigma}(t)$ – суммар-

ный недоотпуск ЭЭ по всей ОЭС; T – полное время, за которое рассчитывается недоотпущенная ЭЭ; l – текущий индекс элемента системы.

Для приближенной оценки этого показателя можно принять, что $Y_{5l} = 1/F$.

Индекс надежности δ представим как отношение отпущенной ЭЭ потребителям к энергии спроса:

$$\delta = \frac{A_{отп}}{A_{отп} + \Delta A_{\Sigma}}.$$

Здесь $A_{отп}$ – отпущенная потребителям ЭЭ из передающей ЭЭС; ΔA_{Σ} – недоотпущенная ЭЭ из-за различных нарушений в системе, включая потерю устойчивости и отказы противоаварийной системной автоматики.

Показатель $Y_{6l}(3p)$ – приведенные затраты на поддержание системы управления мощностью потребителей. В этом случае рассматриваются наиболее мощные и энергоемкие промышленные предприятия, которые могут фактически повлиять на совмещенный график нагрузки ОЭС. В качестве совмещенного графика нагрузки может выступать график любой режимной энергетической иерархии – от группового графика крупного промышленного предприятия до эквивалентного графика нагрузки ЭЭС в составе ОЭС. Образуется деформированный график нагрузки, который одновременно выравнивается, сглаживается и уплотняется, выполяняя известную процедуру горизонтально-вертикального маневрирования электропотреблением и их комбинаций. В этом случае оптимальный график нагрузки потребителей выполняется при условии $Y_{6l} \rightarrow \min$.

Показатель $Y_{7l}(P_{откл})$ – суммарный ущерб промышленных потребителей при ограничении их мощности на величину $\Delta P_{откл}$ в результате различных нарушений электроснабжения. Количественная величина ущерба обычно определяется с помощью интегральных характеристик ущербов для узлов электропотребителей:

$$3 \left(yT \sum_{i=1}^N Y_i P_{откл} \right) \rightarrow \min,$$

где Y_i , руб./кВт·ч, – ущерб i -го потребителя при его отключении (ограничении); $P_{откл}$ – отключаемая мощность; T – время отключения; y – число отключений за период T .

Основные ограничивающие условия. В общем случае проблема оптимальной коррекции плана формулируется как задача нелинейного программирования:

$$\mathbf{F}(\bar{\mathbf{J}} + \Delta \bar{\mathbf{J}}, \mathbf{G}) \rightarrow \min; \quad (2)$$

$$\mathbf{W}(\mathbf{X}(\mathbf{G}, \bar{\mathbf{J}} + \Delta \bar{\mathbf{J}}), \bar{\mathbf{J}} + \Delta \bar{\mathbf{J}}, \mathbf{G}) = 0. \quad (3)$$

Здесь уравнение (2) – целевая функция управления при отклонении параметров на $\Delta \bar{\mathbf{J}}$ от вектора исходных данных $\bar{\mathbf{J}}$; (3) – уравнение установившегося режима ЭС (\mathbf{X} – зависимые параметры; \mathbf{G} – управляющие воздействия, формирующиеся в виде двусторонних ограничений-неравенств, наложенных на параметры режима).

Основные значения перетоков обменной мощности по МЛЭП определяются в рамках двусторонних нестрогих ограничений:

$$S_{\xi}^{\min} \leq S_{\xi} \leq S_{\xi}^{\max}, \quad \xi \in \Lambda,$$

где Λ – множество контролируемых линий; S_{ξ}^{\min} и S_{ξ}^{\max} – нижняя и верхняя границы допустимых значений перетоков по ξ -й МЛЭП соответственно; S_{ξ} – поток мощности в ξ -й линии.

После проверки ЛЭП по пропускной способности выделяется множество линий $v \in V$ с нарушенными режимными ограничениями. Соответственно, формируется корректирующий вектор $\Delta \dot{S}$:

$$\Delta \dot{S} = colon [\Delta \dot{S}_1, \Delta \dot{S}_2, \dots, \Delta \dot{S}_v, \dots, \Delta \dot{S}_v],$$

компоненты которого равны

$$\Delta \dot{S}_v = \begin{cases} v^{\max} - v, & \text{если } v^{\max} \leq v; \\ 0, & \text{если } v^{\max} \geq v. \end{cases}$$

Здесь v^{\max} – максимально допустимый поток мощности в v -й линии; v – поток мощности в v -й линии в рассматриваемом режиме; V – множество линий с нарушенными режимными ограничениями.

На первом этапе решения задачи коррекции режима ЭЭС ограничиваемся возможностью регулирования перетоков ЛЭП с помощью средств генерации активной и реактивной мощности (электростанции системы и источники реактивной мощности). В этом случае можно записать:

$$\Delta \dot{S} = \underline{C} \Delta s,$$

где \underline{C} – матрица коэффициентов распределения токов; Δs – вектор-столбец задающих мощностей в узлах.

На этом этапе критерием оптимальности считаются минимальные изменения узловых мощностей. Вектор $\Delta \dot{S}$, найденный из последнего соотношения, будет удовлетворять этому критерию.

На следующем этапе в коррекцию параметров включаются ветви, имеющие трансформаторы с РПН. Принимая в первом приближении допущение о том, что приведенные напряжения в узлах расчетной схемы одинаковы по величине и фазе и равны U_{cp} , можно записать:

$$\Delta \dot{S} = \underline{C} \Delta \dot{s} + \text{diag} \underline{Y} (\underline{M} \cdot \underline{C} + 1) (e^{(m)} - \Delta \underline{k}) U_{cp}^2.$$

Здесь $\text{diag} \underline{Y}$ – диагональная матрица проводимостей ветвей; \underline{M} – первая матрица инцидентий; $\Delta \underline{k}$ – вектор-столбец отклонений относительных коэффициентов трансформации по отношению к исходному базисному режиму; $e^{(m)}$ – вектор-столбец, состоящий из m единиц.

Минимизация дефицита мощности. Рассматривалась ОЭС, основная сеть которой включает: генерирующие источники; узлы привязки к односторонним межсистемным связям, постоянно импортирующих мощность извне; узлы привязки к межсистемным связям с реверсивными перетоками мощности; начальные узлы экспортных линий с межсистемными перетоками мощности; потребительские узлы энергоемких промышленных предприятий, выполняющие роль потребителей-регуляторов; узлы-приемники с фиксированным потреблением мощности в течение рассматриваемого временного интервала. Дефицит мощности в детерминированной постановке можно описать в виде оценки по поставке мощности на территорию ОЭС и её потребления [5]:

$$Def(P) = \left(\sum_{i=1}^n P_i^G + \sum_{j=1}^N P_j^{имп} + P_{рез}^G - \Delta P_{рем}^G \right) \pm \sum_{k=1}^K \Delta P_k^{имп(эксп)} - \left(\sum_{l=1}^L \Delta P_l^{эксп} + \sum_{v=1}^V P_v^{pp} + \sum_{\mu=1}^M P_{\mu}^n \right), \quad (4)$$

где $\sum_{i=1}^n P_i^G$ – располагаемая активная мощность генераторов ОЭС; $\sum_{j=1}^N P_j^{имп}$ – импорт электрической мощности в ОЭС; $P_{рез}^G$ – резервная мощность генераторов в ОЭС; $\Delta P_{рем}^G$ – выведенная мощность ремонтируемых агрегатов; $\sum_{k=1}^K \Delta P_k^{имп(эксп)}$ – электрическая мощность реверсивных перетоков в МЛЭП; $\sum_{l=1}^L \Delta P_l^{эксп}$ – экспорт мощности; $\sum_{v=1}^V P_v^{pp}$ – регулируемая мощность энергоемких промышленных предприятий; $\sum_{\mu=1}^M P_{\mu}^n$ – фиксированная мощность крупных нагрузочных узлов.

Таким образом, требуется минимизировать суммарный дефицит мощности в ОЭС (1) при условии выполнения системы линейных двусторонних ограничений-неравенств на переменные режима:

$$P_i^{\min} \leq P_i^G \leq P_i^{\max}, \quad i \in \overline{1, n}; \quad (5)$$

$$P_j^{\min} \leq P_j^{имп} \leq P_j^{\max}, \quad j \in \overline{1, N}; \quad (6)$$

$$P_k^{\min} \leq P_k^{имп(эксп)} \leq P_k^{\max}, \quad \gamma \in \overline{1, K}; \quad (7)$$

$$P_l^{\min} \leq P_l^{эксп} \leq P_l^{\max}, \quad l \in \overline{1, L}; \quad (8)$$

$$P_v^{\min} \leq P_v^{pp} \leq P_v^{\max}, \quad v \in \overline{1, V}. \quad (9)$$

Все ограничения формируются с учетом двусторонних соглашений между дефицитной и избыточной энергосистемами, а также с учетом заявок промышленных потребителей, формирующих совмещенный суточный график нагрузки дефицитной ЭС. В этом случае условие минимизации для прогнозируемого временного интервала t можно записать как

$$Def(P_t) \rightarrow \min \quad (10)$$

в области пространства переменных (5) – (9) и баланса мощности. Значения мощности в разных узлах определяются в течение расчетного временного интервала, равного одному часу (получасу) совмещенного суточного графика нагрузки ОЭС. Значения мощностей некоторых узлов могут иметь и постоянные величины. Тогда они фиксируются и выводятся из состава ограничений (5) – (9).

Для экспресс-оценки баланса между производством и потреблением ЭЭ предлагается *упрощенный метод минимизации топливной составляющей издержек электростанциями ЭЭС с одновременной минимизацией потерь активной мощности в сети*. В качестве независимых переменных принимались мощности генерирующих узлов и узлов, связанных с транзитными перетоками поступающей ЭЭ в сеть ЭЭС. Рассматривалась задача суточной оптимизации режима ЭС с усредненным часовым интер-

валом. Целевая функция включает в себя такие показатели, как минимум расхода топлива в ЭС, минимум потерь мощности, минимум отклонения величин сальдо-перетоков от договорных обязательств.

Расход топлива на электростанциях ЭЭС определяется с помощью расходных характеристик:

$$T_i = f_i(P_i), i \in \{n\} \quad (11)$$

где T_i – расход топлива на i -й электростанции; P_i – мощность i -й электростанции; n – число генерирующих узлов.

Характеристики могут быть представлены в виде полиномов второй степени:

$$T_i = K'_i + K''_i P_i + K'''_i P_i^2, i \in \{n\}. \quad (12)$$

Здесь K'_i – расход топлива при минимальной нагрузке (т у.т.); K''_i и K'''_i – коэффициенты полинома второй степени в уравнении (12).

Во многих случаях характеристики современных блочных электростанций можно приближенно представить линейной зависимостью

$$T_i = K'_i + K''_i P_i, i \in \{n\},$$

тогда сумма расхода топлива на всех электростанциях за каждый временной интервал выразится как

$$T_{\Sigma i} = \sum_{i=1}^n (K'_i + K''_i P_i), i \in \{n\},$$

или с учетом (8), отбрасывая постоянную составляющую, получим:

$$T_{\Sigma i} = \sum_{i=1}^n \beta_i \theta_{it}(P) P_{it}, i \in \{n\}, t = 1, 2, \dots, 24,$$

где θ_{it} – значение удельного расхода топлива (т у. т. / МВт) для i -й электростанции в усредненном интервале t -го часа; P_{it} – усредненная прогнозируемая мощность i -й электростанции за t -й час суток, МВт; β_i – корректирующий коэффициент стоимости топлива (учитывает расходы на добычу, транспортные расходы, качественные показатели).

В этом случае потери активной мощности, обусловленные технологическим расходом ЭЭ на транспорт в электрических сетях, могут быть приближенно выражены следующим образом:

$$\pi(P) = \sum_{\mu, \nu=1}^N b_{\mu\nu} P_{\mu\nu}^2.$$

Здесь $b_{\mu\nu}$ – величины, определяемые узловыми собственными и взаимными комплексными сопротивлениями исследуемой сети (« b -коэффициенты»); μ, ν – текущие индексы генераторных или нагрузочных узлов ($\mu, \nu = 1, 2, \dots, N$; $\mu \neq \nu$; N – общее число узлов ЭЭС).

Суммарные издержки на генерацию и распределение активной мощности в ЭЭС можно приближенно представить с помощью выражения

$$H_{\Sigma t} = \sum_{i=1}^n \beta_i \theta_{it}(P) P_{it} + \gamma \sum_{\mu, \nu=1}^N b_{\mu\nu} P_{\mu\nu}^2, \quad (13)$$

где γ – топливный эквивалент стоимости потерь в сети. Тогда задачу оптимизации можно сформулировать как задачу минимизации функционала (13):

$$\left(\sum_{i=1}^n \beta_i \theta_{it}(P) P_{it} + \gamma \sum_{\mu, \nu=1}^N b_{\mu\nu} P_{\mu\nu}^2 \right) \rightarrow \min.$$

Допустимая область существования целевой функции определяется системой ограничений, основные из которых следующие:

1) условия баланса активных мощностей для каждого t -го интервала времени:

$$\sum_{i=1}^n P_{it} - \sum_{j=1}^m P_{jt} - \pi(P) = 0,$$

где P_{it} – усредненная мощность i -го объекта генераторной группы в течение временного интервала t ; P_{jt} – суммарная активная нагрузка потребителей j -й группы; m – общее число условных нагрузочных узлов (групп потребителей);

2) ограничивающие условия, которые накладываются на расход топлива по i -й электростанции:

$$\sum_{i=1}^{24} T_{it} \leq T_{it}^*$$

Здесь T_{it} – часовой расход условного топлива на i -й электростанции, для которого задан полный расход топлива T_{it}^* за сутки;

3) регулировочный диапазон изменения активной нагрузки v -го объекта:

$$P_{vt} \leq P_v \leq \bar{P}_v, v \in \{n\},$$

где $\underline{P}_v, \bar{P}_v$ – нижняя и верхняя предельные границы изменения мощностей энергоузлов.

4) сумма мощностей генерирующих узлов для каждого часа суток:

$$\sum_{i=1}^n P_{it} = \Gamma;$$

5) допустимые пределы изменения активной мощности элементов сети (линии, группы линий, трансформаторы):

$$P_{it} \leq \bar{P}_i, i \in \{L\},$$

где \bar{P}_i – верхняя предельная граница передаваемой мощности по элементам сети; L – число контролируемых элементов сети.

В матричных обозначениях сформулированную задачу можно представить в следующем виде:

$$H_{\Sigma t}(P) = \mathbf{T}^* \mathbf{P} + \mathbf{P}^* \mathbf{B} \mathbf{P}; \tag{14}$$

$$\mathbf{D}^* \mathbf{P} \leq \mathbf{G}; \mathbf{P} \geq 0, \tag{15}$$

где $\mathbf{P}^* = [P_G | P_G^M | P_H^M | P_H]$ – столбцовая матрица активных узловых мощностей; $P_G = P_1, P_2, \dots, P_n$ – генераторы; $P_G^M = [P_{G1}^M, P_{G2}^M, \dots, P_{Gk}^M]$ – импортируемые потоки мощности (внешние генераторы); $P_H^M = [P_{H1}^M, P_{H2}^M, \dots, P_{Hk}^M]$ – экспортируемые потоки мощности (внешние нагрузки); $P_H = [P_1, P_2, \dots, P_m]$ – нагрузки; $\mathbf{B} = \gamma \frac{1}{2U^2} \mathbf{Z}'$ – матрица коэффициентов потерь; \mathbf{Z}' – вещественная составляющая матрицы узловых собственных и взаимных проводимостей сети; U – среднее значение напряжения в сети; \mathbf{D} – матрица, составленная из коэффициентов ограничений; \mathbf{G} – столбцовая матрица ограничивающих констант. Сформулированная проблема хорошо вписывается в классическую модель квадратичного выпуклого программирования, которая может быть решена за конечное число шагов.

При условии выполнения ограничений (5) – (9), варьируя значения P_i и Q_i для электростанций (а в идеальном случае и для потребителей-регуляторов), определяется минимум нагрузочных потерь для ОЭС. Снижение технологического расхода ЭЭ позволяет частично разгрузить генераторы электростанций с пологими расходными характеристиками, сократив при этом суммарный расход топлива в ОЭС.

Минимизация потерь ЭЭ на транспорт. При взаимодействии ЭС с учетом передачи мощности по высоковольтным МЛЭП важным аспектом является минимизация технологического расхода ЭЭ на транспорт в электрических сетях [6]. В настоящее время в ОЭС потери ЭЭ постепенно сокращаются за счет повышения эффективности технической эксплуатации основного электрооборудования и усовершенствования методики учета технологических потерь.

На рисунке 1 представлена динамика изменения значений потерь ЭЭ (%) с 2000 по 2010 год по данным US Energy Information Administration для нескольких государств [7]. Среди представленных стран Республика Беларусь имеет средние показатели по потерям, и потому снижение их является актуальной задачей.

При поиске оптимального режима работы ОЭС представляет интерес переменная составляющая потерь, влияя на которую можно добиться их снижения. Условие минимума технологического расхода ЭЭ на транспорт в электрических сетях можно представить следующим образом:

$$\sum_{i=1}^{24(48)} \alpha_1 \cdot Y_1(\Delta P) \rightarrow \min, \tag{16}$$

где α_1 – весовой корректирующий коэффициент, определяемый на основании экспертных оценок; $Y_1(\Delta P)$ – функция стоимости суммарного технологического расхода ЭЭ.

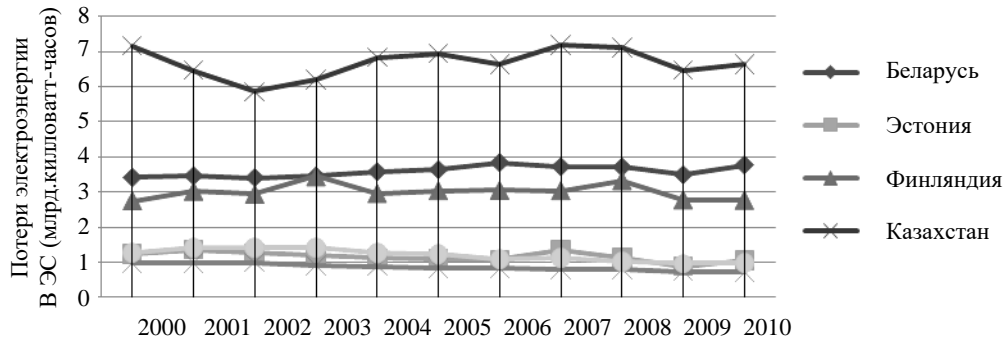


Рис. 1. Изменение потерь электроэнергии в 2000 – 2010 годах

Рассматриваются автономный режим работы ОЭС и параллельная работа ОЭС (наличие межсистемных и межгосударственных потоков мощности). Для случая автономного режима работы учитываются только нагрузочные потери активной мощности, суммарное значение которых может быть представлено в виде:

$$\Delta P_{\Sigma} = \sum_{i=1}^n \frac{P_i^2 + Q_i^2}{U_i^2} \cdot R_i.$$

Здесь U_i – напряжение ЛЭП; R_i – активное сопротивление ЛЭП; P_i – активная мощность в ЛЭП; Q_i – реактивная мощность в ЛЭП; $i = 1, \dots, n$ – текущий индекс ЛЭП, для которой находят величину потерь.

Для рассмотрения возможности снижения нагрузочных потерь следует варьировать, в основном, величины P_i и Q_i , так как изменять напряжения в сети в достаточно больших пределах нежелательно. Поэтому в качестве значения напряжения на данном этапе можно принять среднее его значение $U_{\text{ср}}$. Тогда при условии выполнения режимных ограничений, варьируя значения P_i и Q_i для электростанций, находим минимум нагрузочных потерь для ОЭС. Все расчеты проводились в денежном выражении, поэтому функция стоимости суммарного технологического расхода ЭЭ на транспорт в электрических сетях была приведена к виду:

$$Y_1 = C \sum_{l=1}^{24(48)} \sum_{i=1}^n \frac{P_i^2 + Q_i^2}{U_i^2} \cdot R_i \cdot t, \quad (17)$$

где C – стоимость ЭЭ в рассматриваемой ЭС; t – промежуток времени, в течение которого все параметры сети считаются условно-постоянными (в данном случае – час либо полчаса); $l = 1, \dots, 24$ (48) – номер интервала времени (часовой, получасовой), $i = 1, \dots, n$ – номер ЛЭП.

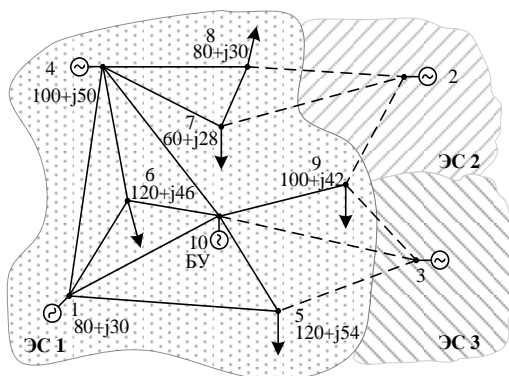


Рис. 2. Схема сложно-замкнутой ЭЭС для случая автономной работы

Таким образом, для каждого часа при кратковременном планировании режимов работы ЭС необходимо найти оптимальное распределение суммарной нагрузки ЭС между основными электростанциями, чтобы выражение (16) было минимальным. Для начального приближения коэффициент α_1 принимается равным единице.

На рисунке 2 приведен пример схемы сложно-замкнутой сети, состоящей из 10 узлов и 18 ветвей.

Параметры ЛЭП для расчетной схемы представлены в таблице 1.

Узлы 1, 4...10 входят в состав ЭС 1, а узлы 2 и 3 представляют собой узлы из ЭС 2 и 3, соответственно. Следовательно, линии 2...7, 2...8, 2...9, 3...5, 3...9, 3...10 являются межсистемными, для случая автономной работы будут считаться отключенными. Базисное напряжение в ОЭС – 110 кВ. Расчет режима и поиск значений мощности генераторов, чтобы значения технологического

расхода ЭЭ при транспорте в электрических сетях ЭС 1 при автономной работе были минимальны, производились с помощью программы в среде MATLAB. Полученные значения были проверены с помощью эталонной программы «Rastr».

Таким образом, задача поиска оптимальных значений мощностей генераторов ЭС, при которых значения потерь ЭЭ в сети будут минимальны, сводится к задаче нелинейного программирования с целевой функцией, представленной формулой (18).

Таблица 1

Параметры ЛЭП для расчетной схемы

№ линии	R, Ом	X, Ом	№ линии	R, Ом	X, Ом
1...4	16	28	3...10	2	10
1...5	9	22	4...6	9	14
1...6	6	21	4...7	3	9
1...10	10	26	4...8	7	16
2...7	12	20	4...10	2	8
2...8	6	14	5...10	8	24
2...9	5	16	6...10	7	14
3...5	4	12	7...8	9	15
3...9	6	15	9...10	5	16

$$F \left[\frac{1}{U_{cp}^2} \cdot C \sum_{i=1}^n (P_{vi}^2 + Q_{vi}^2) \cdot R_i \right] \rightarrow \min, \quad (18)$$

где P_{vi}^2, Q_{vi}^2 – активная и реактивная мощности i -й ЛЭП, соответственно.

Таким образом, необходимо найти минимум функции многих переменных при соблюдении двусторонних ограничений:

$$P_{\min_i}^{\text{ген}} \leq P_i^{\text{ген}} \leq P_{\max_i}^{\text{ген}}, \quad (19)$$

При решении задачи оптимизации используется один из методов нелинейного программирования – метод барьерных функций, при котором вводится запрет на выход оптимизируемых величин за границы их допустимых значений. Алгоритм программного расчета состоит из следующих этапов:

- 1) запускается файл чтения данных, в котором содержатся все параметры рассчитываемой схемы;
- 2) задаются минимальные и максимальные величины мощности генераторов, участвующих в процессе оптимизации, и вектор их заданных значений в качестве начального приближения;
- 3) с помощью метода барьерных функций проводится поиск оптимальных компонентов вектора $P_{\text{ген}}$, и найденные значения используются для уточненного расчета режима ЭС:
 - первоначальные значения мощностей узлов корректируются с помощью учета зарядных мощностей ЛЭП:

$$Q_i = \frac{1}{2} U_{cp}^2 \sum_{j=1}^m b_{ij}, \quad (20)$$

где Q_i – реактивная мощность i -го узла, которая обусловлена зарядной мощностью инцидентных этому узлу линий, имеющих емкостные поперечные проводимости; U_{cp} – среднее напряжение в сети; b_{ij} – емкостная проводимость линии $i - j$; m – число инцидентных этому узлу ветвей;

- выполняется операция разнесения потерь мощности в линии по ее концам:

$$\Delta S_i = \frac{1}{2} \Delta S_k; \quad \Delta S_j = \frac{1}{2} \Delta S_k, \quad (21)$$

где i и j – номера начального и конечного узлов k -й линии соответственно;

- производится коррекция узловых мощностей в узлах i и j :

$$S'_i = S_i + \Delta S_i; \quad S'_j = S_j + \Delta S_j; \quad (22)$$

- уточняется значение зарядной мощности в узлах:

$$Q'_i = \frac{1}{2} U_{yi}^2 \sum_{j=1}^m b_{ij}, \quad (23)$$

где U_{yi} – вычисленное значение напряжения в i -м узле;

- уточняется значение узловой мощности в i -м узле:

$$S''_i = S'_i + jQ_i - jQ'_i, \quad (24)$$

где Q_i – первоначальное значение узловой зарядной мощности; Q'_i – ее уточненное значение;

4) вычисляются значения параметров режима ЭС, величина технологического расхода ЭЭ на транспорт в сети с помощью выражения (16) и соответствующее значение функции стоимости потерь электроэнергии согласно выражению (17).

В таблице 2 приведены значения потерь в сети за 1 час, в течение которого параметры ЭС приняты условно постоянными, для заданных значений мощности (№ расчета – 1) и для рассчитанных оптимальных значений (№ расчета – 2), стоимость ЭЭ принята равной для ЭС 1 – 0,035 у. е. Мощность первой электростанции может изменяться в пределах 50 – 160 МВт, а второй – от 50 до 200 МВт.

Таблица 2

Изменение потерь мощности в ЛЭП ЭС 1 при автономной работе

№ расчета	P_1 , МВт	P_4 , МВт	ΔP_{Σ} , МВт	Y_1 , у. е.
1	80	100	18,04	0,63
2	105,5	164	17,68	0,62

Таким образом, для данной схемы выигрыш при перераспределении нагрузок генераторов электростанций 1 и 4 в течение рассматриваемого интервала времени (1 час) составляет 0,36 МВт или 1,27 у. е. После расчета получают оптимальные значения мощностей генераторов ЭС 1.

Для рассмотрения случая межсистемных перетоков мощности при покупке/продаже ЭЭ в формулу (16) вносятся следующие изменения:

$$Y_1 = \sum_{i=1}^{24(48)} \sum_{i=1}^n \frac{P_i^2 + Q_i^2}{U_{cp}^2} \cdot R_i \cdot t \cdot C_i, \quad (25)$$

так как стоимость ЭЭ в разных ЭЭС различна и больше не является постоянной величиной.

При отнесении потерь ЭЭ на принимающую сторону возможны следующие варианты:

- 1) при полном учете потерь принимающей стороной для R_i берется полное сопротивление МЛЭП;
- 2) при полном отнесении потерь на передающую сторону R_i будет равно нулю;
- 3) при «пропорциональном» распределении затрат на потери ЭЭ выражение (12) преобразуется:

$$Y_1 = \sum_{i=1}^{24(48)} \sum_{i=1}^n \frac{P_i^2 + Q_i^2}{U_{cp}^2} \cdot k_i \cdot R_i \cdot t \cdot C_i, \quad (26)$$

где $-k_i = \frac{L_{прин_i}}{L_{лэп_i}}$ коэффициент учета потерь ЭЭ в i -й ЛЭП для принимающей ЭС; $L_{лэп_i}$ – общая длина

i -й ЛЭП (км); $L_{прин_i}$ – длина ЛЭП, которая относится к принимающей ЭЭС (км); $k_i = 1 - \frac{L_{прин_i}}{L_{лэп_i}}$ – коэффициент учета потерь ЭЭ в i -й ЛЭП для передающей ЭС.

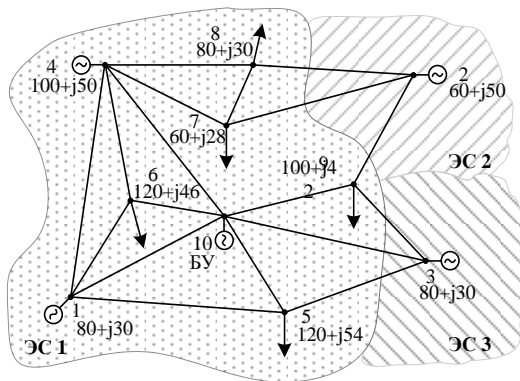


Рис. 3. Схема сложно-замкнутой ЭЭС для случая параллельной работы

В таблице 3 приведены значения потерь для заданных значений мощности (расчет № 1) и для рассчитанных оптимальных значений (расчет № 2) при стоимости ЭЭ, равной для ЭС1 0,035 у. е., для ЭС2 – 0,05 у. е., для ЭС3 – 0,04 у. е.; заданные и оптимальные значения потерь, когда стоимость ЭЭ в ЭС1 – 0,035 у. е., ЭС2 – 0,04 у. е., ЭС3 – 0,05 у. е. (расчет № 3 и 4 соответственно); соответствующие значения функции стоимости потерь ЭЭ в электрических сетях при указанных ценах ЭЭ в данных энергосистемах.

После расчета мы получаем оптимальные значения мощностей генераторов, относящихся к энергосистемам 1, 2 и 3. Таким образом, для данной схемы выигрыш при перераспределении генерации электростанций 1 и 4 и межсистемных перетоков мощности от энергосистем 2 и 3 в течение рассматриваемого

В случае учета межсистемного транзита мощности в результате изменения потокораспределения потери могут как увеличиваться, так и уменьшаться. Для анализа ситуации выполнялся расчет нагрузочных потерь сначала для режима без учета транзита ЭЭ, а затем с учетом добавления транзитных перетоков. После чего проводилось сравнение этих значений, и была дана оценка влияния МТ мощности на ОЭС.

Рассмотрим случай пропорционального отнесения потерь мощности в межсистемных ЛЭП для ЭС 1.

Рамки изменения мощности электростанций № 1 и 4 при проведении оптимизации остаются такие же, как в предыдущем случае, а для электростанций № 2 и № 3 мощность генераторов может изменяться в пределах 50 – 150 МВт. На рисунке 3 представлена ЭС1, но с подключенными к ней с помощью МЛЭП энергосистемами 2 и 3.

го часа составляет для первого случая – 0,8 у. е., а для второго – 1,39 у. е. Для проведения расчетов была написана программа в среде MatLab для поиска такого распределения мощностей электростанций, чтобы потери мощности в сети были минимальными.

Таблица 3

Изменение потерь мощности в ЛЭП ЭС 1 при параллельной работе

№ расчета	P1, МВт (ЭС 1)	P4, МВт (ЭС 1)	P2, МВт (ЭС 2)	P3, МВт (ЭС 3)	ΔPΣ, МВт (ЭС 1)	Y1, у. е. (ЭС 1)
1. C1 – 0,035 у. е., C2 – 0,05 у. е., C3 – 0,04 у. е.	80	100	60	80	11,47	42,82
2. C1 – 0,035 у. е., C2 – 0,05 у. е., C3 – 0,04 у. е.	72,8	120,4	72,7	80,4	11,15	42,02
3. C1 – 0,035 у. е., C2 – 0,04 у. е., C3 – 0,05 у. е.	80	100	60	80	11,47	43,65
4. C1 – 0,035 у. е., C2 – 0,04 у. е., C3 – 0,05 у. е.	78,4	112,2	86,4	67,2	11,12	42,26

Заключение. Выполненное исследование позволяет более широко и всесторонне решать важную триединую задачу – бесперебойное электроснабжение промышленных и бытовых потребителей, обеспечение максимально возможной экономичности работы собственных генерирующих мощностей, а также одновременное выполнение договорных графиков сальдо-перетоков ЭЭ.

ЛИТЕРАТУРА

1. ЭСКО – Электронный журнал энергосервисной компании «Экологические Системы» // Электроэнергетика Беларуси. – 2012. – № 5, май.
2. Концепция регулирования частоты и перетоков в энергообъединении стран СНГ и Балтии: утв. решением ЭЭС СНГ от 27.10.2007.
3. Основные технические требования к параллельно работающим энергосистемам стран СНГ и Балтии. Правила и рекомендации по регулированию частоты и перетоков: утв. решением ЭЭС СНГ от 12.10.2007.
4. Александров, О.И. Математическая модель оптимизации электроснабжения дефицитных регионов по межсистемным линиям связи / О.И. Александров, М.Ш. Мисриханов, Н.В. Радоман // Методические вопросы исследования надежности больших систем энергетики. Вып. 62; отв. ред. Н.И. Воропай, В.А. Савельев. – Иваново: ПресСто, 2011. – С. 468 – 476.
5. Радоман, Н.В. Минимизация дефицита мощности в объединенной энергосистеме / Н.В. Радоман, О.И. Александров // Автоматический контроль и автоматизация производственных процессов: материалы междунар. науч.-техн. конф., Минск, 17 – 18 мая 2012 г. – Минск: БГТУ, 2012. – С. 127 – 129.
6. Радоман, Н.В. Минимизация технологического расхода электроэнергии на транспорт в электрических сетях энергосистемы / Н.В. Радоман, О.И. Александров // Труды БГТУ. Научный журнал. – 2012. – № 6(153). – С. 107 – 112.
7. US Energy Information Administration, International Energy Statistics, 2000 – 2010 [Electronic resource]. – Mode of access: <http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=2&pid=2&aid=9&cid=AR,AM,AU,AJ,BO,BE,BR,BU,CA,CH,EZ,DA,EN,FI,FR,GM,HU,IC,EL,IS,IT,JA,KZ,KG,LG,LH,MD,NL,NO,PL,RS,SW,TI,UP,UK,US,UZ,VE,&syid=1996&eyid=2008&unit=BKWH>.

Поступила 10.06.2013

OPTIMIZATION OF THE BALANCE RELIABILITY OF THE SCARCE POWER SUPPLY SYSTEM TAKING INTO ACCOUNT INTERSYSTEM POWER LINES

N. RADOMAN, O. ALEKSANDROV, D. SVIRSKIJ

A problem of optimization of balance reliability of a scarce power supply system taking into account intersystem power lines is considered. This problem is described by multicriteria objective function which includes the minimum of deviation values from the planned values of power flows, the minimum total fuel consumption, the minimum of power and energy losses in supply system and intersystem power lines. Optimization is performed taking into account the given system of linear bilateral inequality constraints on variables of an operating mode of electrical power system. The method of minimization of a fuel component of expenses of power plants with simultaneous minimization of losses of active power in electric networks is offered for the rapid assessment of the balance between electricity production and consumption. Special attention is paid to the minimization of losses of the electric power on transport in electric networks.

УДК 658.29-049.5

**КОМПЛЕКСНЫЙ МЕТОДИЧЕСКИЙ ПОДХОД
ОПЕРАТИВНОЙ ОЦЕНКИ УРОВНЯ КИБЕРПРЕСТУПЛЕНИЙ***канд. техн. наук В.В. МАЛИКОВ**(Центр повышения квалификации руководящих работников
и специалистов Департамента охраны МВД Республики Беларусь, Минск);**канд. техн. наук Е.А. КРИШТОПОВА**(Белорусский государственный университет информатики и радиоэлектроники, Минск)*

Предложен комплексный методический подход по оценке уровня киберпреступлений, учитывающий нормативно-правовые, технологические, экономические критерии и параметры. Рассматриваются принципы, реализующие оценку уровня киберпреступлений. Показана совокупная экономическая оценка прибыли от реализации атак, учитывающая величину уровня доходности по критерию близости риска совершения киберпреступлений, основанному на оценках технико-экономического потенциала рынка и уровня наказания в соответствии с нормативно-правовыми актами. Классифицируется рынок киберпреступности.

Введение. В настоящее время использование информационно-коммуникационных технологий со стороны организаций различных форм собственности, а также физических лиц приобрело трансграничный характер. Государства, на территории (или через территории) которых осуществляется активное использование таких технологий, должны принимать меры для нормативно-правового, организационно-технического и технического регулирования в данной сфере. Обеспечение внутригосударственного регулирования невозможно без учета аспектов международного законодательства и сотрудничества между ведущими индустриально развитыми странами.

Глобальный рынок киберпреступности активно развивается и совершенствуется в соответствии с передовыми направлениями информатизации общества, внедрением электронных систем коммуникаций, электронных платежных систем [1].

В целях организации действенной системы противостояния современным вызовам и угрозам, реализуемым для противоправной деятельности со стороны преступного сообщества, необходима разработка методической базы, способствующей пресечению и раскрытию киберпреступлений.

Методика исследований. Для оценки состояния рынка киберпреступности, устанавливающего уровень его соответствия определенным критериям и показателям, необходимо проведение совокупного экономического анализа прибыли от реализации атак с учетом величины уровня доходности по критерию близости риска совершения киберпреступлений, который базируется на оценках технико-экономического потенциала рынка и уровня наказания в соответствии с нормативно-правовыми актами. Существующие в настоящее время методы оценки, а также программные комплексы, разработанные на их основе, имеют определенный перечень недостатков и уязвимостей, а также значительную (во многих случаях избыточную) стоимость и низкую функциональность [2].

Для достижения цели, заключающейся в разработке методического подхода по технической и экономической оценке уровня киберпреступлений, необходимо задаться перечнем принципов, его реализующих. В качестве *основополагающих принципов* выделим следующие: цели, согласования критериев, равнопрочности оценки.

Под *принципом цели* будем понимать следующее: все решения при проведении оценки ориентированы на главную цель – обеспечение требуемого уровня достоверности информации для принятия управленческих решений.

Под *принципом согласования критериев* будем понимать то, что оценки состояния киберпреступности на всех уровнях иерархии анализируемой структуры должны быть согласованы и базироваться на оценках состояния элементов нижнего уровня иерархии.

Под *принципом равнопрочности оценки* будем понимать следующее: оценка риска совершения комплексной кибератаки в целом и отдельных технологических атак в частности равна величине максимального риска совершения преступления, т.е. оценка риска производится по наиболее уязвимому элементу.

Будем считать, что логически законченные подсистемы структуры рынка киберпреступности P_m являются независимыми модулями, т.е. вероятность реализации атаки подсистемой O_A не влияет на вероятность реализации атаки подсистемой O_B , т.е. $P_{O_A}(O_B) = P_B$.

Однако известно, что каждая подсистема множества O определяется множеством $R = \{R_1, R_2, \dots, R_n\}$ признаков, которые будем называть *показателями атакующей системы*.

Таким образом, для рынка киберпреступности требуется определить множества $P = \{P_1, P_2, \dots, P_m\}$, общий риск для каждой атакующей подсистемы множества O с учетом множества показателей атакующей системы $R = \{R_1, R_2, \dots, R_n\}$ как по уровню структуры киберпреступлений, так и всей структуры киберпреступлений в целом [3].

Результаты и обсуждение. Оценка технического и экономического уровня киберпреступлений была проведена через определенные действия.

1. *Декомпозиция рынка киберпреступности на законченные функциональные уровни и модули* (рис. 1). Предлагается использование варианта декомпозиции, который учитывает полный технологический цикл осуществления атак, включающий как разработку вредоносного программного обеспечения, так и непосредственно его использование в преступных целях. В качестве базовых уровней предлагаются следующие: интернет-мошенничество; спам; DDoS-атаки; рынок криминальных средств систем и услуг [2; 4; 5].



Рис. 1. Классификация рынка киберпреступности

2. Следующий этап – *проведение уровневой технико-экономической оценки потенциала модулей рынка киберпреступности.*

Техническую оценку будем проводить на основе параметра технологичности атаки – $K_{техн}$, показатели которого будут иметь следующие возможные значения:

- высокотехнологичная атака/услуга (H): не описана в базах знаний по уязвимостям и сборниках эксплойтов – является таргетированной (0-day, Watering Hole, social engineering и др.) [4; 5];
- низко технологичная атака/услуга (L): описана в базах знаний по уязвимостям и сборниках эксплойтов – не является таргетированной.

Экономическую оценку будем проводить на основе параметра потенциального объема рынка $P_{пот}$, показатели которого будут иметь следующие возможные значения:

- высокий объем рынка (H): потенциальный объем ≥ 1 млн. долларов США;
- низкий объем рынка (L): потенциальный объем менее 1 млн. долларов США.

Таким образом, представление уровневой технико-экономической оценки потенциала рынка киберпреступности $V_{кибер}$ с учетом описанных параметров оценки будет иметь вид:

$$V_{кибер} = \{K_{техн}, P_{пот}\}. \tag{1}$$

В качестве базовых уровней технико-экономической оценки потенциала рынка киберпреступности $V_{кибер}$ предлагаются такие: $A1 (H, H)$, $B1 (H, L)$, $C1 (L, H)$, $D1 (L, L)$.

3. Далее проведем *уровневую оценку величины наказания за совершенные киберпреступления по действующим нормативно-правовым актам*.

Оценку длительности срока лишения свободы осуществим на основе параметра величины срока наказания H_{cp} , показатели которого будут иметь следующие возможные значения:

- высокий уровень наказания (H): срок лишения/ограничения свободы ≥ 3 лет;
- низкий уровень наказания (L): срок лишения/ограничения свободы менее 3 лет.

Финансовая оценка наказания осуществляется на основе параметра величины срока наказания $H_{фин}$, показатели которого будут иметь следующие возможные значения:

- высокий уровень (H): сумма штрафа ≥ 100 тыс. долларов США;
- низкий уровень (L): сумма штрафа менее 100 тыс. долларов США.

Таким образом, представление *уровневой оценки величины наказания за совершенные киберпреступления по действующим нормативно-правовым актам* $У_{НПА}$ с учетом описанных параметров оценки будет иметь вид:

$$У_{НПА} = \{H_{cp}, H_{фин}\}. \quad (2)$$

В качестве *базовых уровней оценки величины наказания за совершенные киберпреступления по действующим нормативно-правовым актам* $У_{НПА}$ предлагаются уровни: $A2 (H, H)$, $B2 (H, L)$, $C2 (L, H)$, $D2 (L, L)$.

4. Затем необходимо провести *экспертные оценки средней стоимости организации атаки /услуги на рынке киберпреступности* $C_{орг}$, а также *потенциально возможного среднего дохода от организации атаки /услуги* $C_{дох}$.

5. На следующем этапе необходимо рассчитать *уровень доходности атаки* $У_{дох}$ с учетом *коэффициента риска* $K_{риска}$ по формуле

$$У_{дох} = K_{риска} \cdot \left(\frac{C_{дох}}{C_{орг}} \right). \quad (3)$$

Коэффициент риска $K_{риска}$ будем рассчитывать по формуле

$$K_{риска} = K_{кибер} \cdot K_{НПА}. \quad (4)$$

Соответствующие значения $K_{кибер}$ и $K_{НПА}$ на основе экспертной оценки будем определять из соответствующих параметров оценки $У_{кибер}$ и $У_{НПА}$, указанных в таблицах 1 и 2 соответственно.

Таблица 1

Значения коэффициента $K_{кибер}$

Показатель	Базовый уровень $У_{кибер}$			
	A1	B1	C1	D1
Значение $K_{кибер}$	0,9	0,6	0,3	0,1

Таблица 2

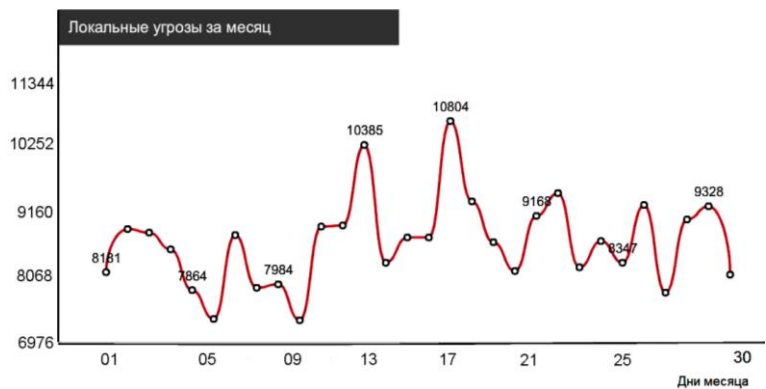
Значения коэффициента $K_{НПА}$

Показатель	Базовый уровень $У_{НПА}$			
	A2	B2	C2	D2
Значение $K_{НПА}$	0,9	0,7	0,5	0,3

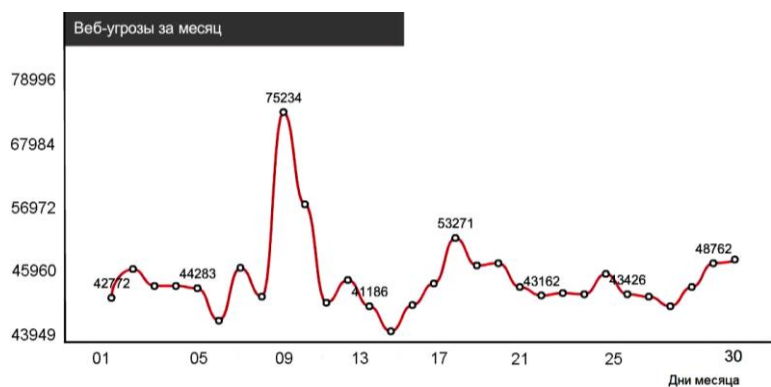
При этом государственным регуляторам и организациям по борьбе с киберпреступлениями в первую очередь необходимо обратить внимание на преступления с $У_{дох} \geq 0,4$, так как в данных случаях высокий уровень дохода при низком уровне ответственности будет способствовать увеличению такого рода атак.

6. Далее определим *потенциально возможное количество атак /услуг* $Ч_{общ}$, на основе экспертных оценок и/или результатов практического сканирования уязвимостей.

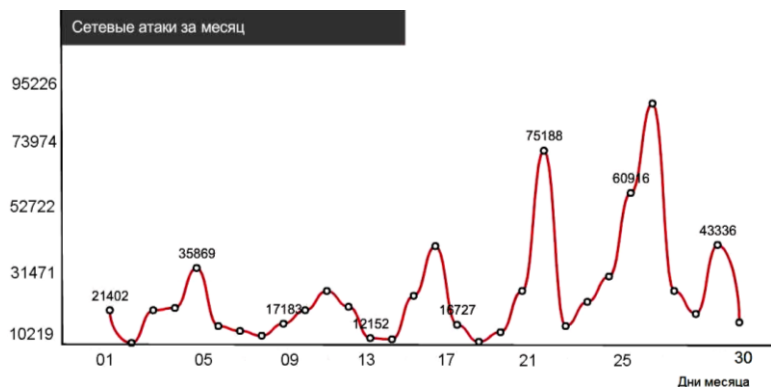
Например, в Республике Беларусь только количество угроз, определенное результатами on-line сканирования в одном из месяцев 2013 года (рис. 2) [6], в день может составить (по среднему числу угроз в месяц) более 50 тысяч.



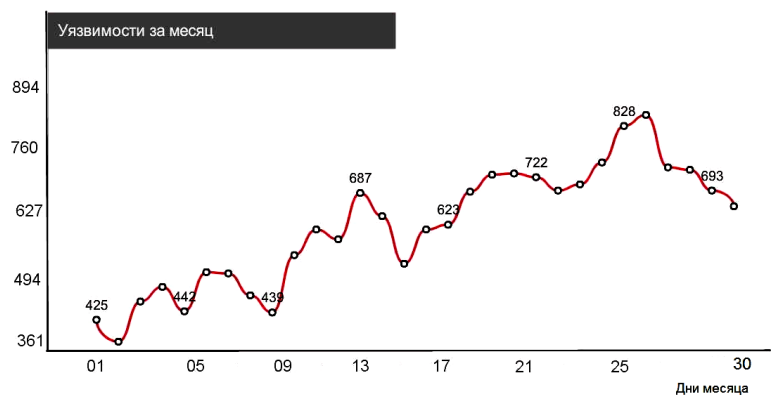
а)



б)



в)



г)

Рис. 2. Результаты on-line сканирования уровня угроз, сетевых атак и уязвимостей в Республике Беларусь за один из месяцев 2013 года: количество локальных угроз, шт. (а); количество веб-угроз, шт. (б); количество сетевых атак, шт. (в); количество уязвимостей, шт. (г)

7. На следующем этапе определяем *общий средний доход от реализации атак / услуг* по формуле:

$$D_{\text{общ}} = Ч_{\text{общ}} \cdot C_{\text{дох}} \quad (5)$$

Например, по оценкам, суммарная величина доходов рынка киберпреступности в Российской Федерации за 2012 год составила 1,938 млрд. долларов США [2].

Заключение. Предложенный методический подход по оценке уровня киберпреступлений позволяет проводить совокупную экономическую оценку прибыли от реализации атак с учетом величины уровня доходности по критерию близости риска совершения киберпреступлений, основанному на оценках технико-экономического потенциала рынка и уровня наказания в соответствии с нормативно-правовыми актами.

Результаты проведения практической оценки рынка киберпреступлений позволят государственным регуляторам и организациям по борьбе с киберпреступлениями оперативно изменять положения нормативно-правовых актов, методику оперативно-розыскных мероприятий и общей информированности граждан.

ЛИТЕРАТУРА

1. Вехов, В.Б. Криминалистическое учение о компьютерной информации и средствах ее обработки: автореф. дис. ... д-ра юрид. наук: 12.00.09 / В.Б. Вехов; ФГОУ ВПО «Волгогр. акад. МВД Рос. Федерации». – Волгоград, 2008. – 59 с.
2. Рынок преступлений в области высоких технологий: состояние и тенденции 2013 года // group-ib.ru [Электронный ресурс]. – 2003–2013. – Режим доступа: <http://www.group-ib.ru/list/1008-analytics/?view=article&id=1155>.
3. Маликов, В.В. Повышение эффективности информационных и инженерно-технических систем защиты критически важных объектов: дис. ... канд. техн. наук: 05.13.19 / В.В. Маликов; БГУИР. – Минск, 2010. – 174 л.
4. DDoS-атаки первого полугодия 2013 года // securelist.com [Электронный ресурс]. – 2013. – Режим доступа: http://www.securelist.com/ru/analysis/208050810/DDoS_ataki_pervogo_polugodiya_2013_goda. Спам в июле 2013 // securelist.com [Электронный ресурс]. – 2013. – Режим доступа: http://www.securelist.com/ru/analysis/208050808/Spam_v_iyule_2013.
5. Статистика 2013 // securelist.com [Электронный ресурс]. – 2013. – Режим доступа: <http://www.securelist.com/ru/statistics#/ru/map/vul/month/Belarus>.

Поступила 26.09.2013

METHODICAL APPROACH OF OPERATIVE ASSESSMENT OF THE LEVEL OF CYBER CRIME

V. MALIKOV, E. KRYSHTOPOVA

The comprehensive methodological approach of assessing the level of cyber crime, which takes into account regulatory, technological, economic criteria and parameters is proposed. The article covers the principles of implementing the assessment of the level of cybercrime. The total economic valuation of the profit gained from the realization of attacks, taking into account the value of the profitability level by proximity of the risk of committing cybercrime, based on estimates of technical and economic potential of the market and the level of punishment in accordance with the legal regulations is shown. The classification of the cyber crime market is given.

УДК 004.896:613.62

**НЕЧЕТКО-МНОЖЕСТВЕННЫЙ ПОДХОД
К ЭКСПЕРТНОЙ ОЦЕНКЕ ПРОФЕССИОНАЛЬНЫХ РИСКОВ
НА ПРИМЕРЕ УСЛОВИЙ ТРУДА РАБОТНИКОВ НЕФТЕПЕРЕРАБАТЫВАЮЩЕГО ЗАВОДА**

Ю.А. БУЛАВКА
(Полоцкий государственный университет)

Обоснована необходимость применения нечетко-множественного подхода к экспертной оценке профессионального риска для здоровья работников НПЗ, предложена нечеткая модель. С целью совершенствования механизма оценки профессионального риска целесообразно использовать аппарат теории нечетких множеств и нечеткого вывода, основное преимущество которого заключается в возможности создания количественных оценок для лингвистических переменных и эффективного отображения зависимости между этими переменными в виде нечетких правил. Показано применение метода экспертных оценок профессионального риска для здоровья работников на основе нечетко-множественного подхода на примере условий труда для работников нефтеперерабатывающих предприятий.

Введение. На современном этапе развития социотехнических систем удобным средством моделирования сложных динамических процессов в условиях неопределенности и многокритериальности является применение методов теории нечетких множеств и нечеткого вывода [1 – 3]. Теория нечеткой логики (или теория нечетких множеств, Fuzzy Logic) – новый подход к описанию процессов, в которых присутствует неопределенность, затрудняющая и даже исключающая применение точных количественных методов и подходов [4 – 5]. *Преимущества использования метода нечетких множеств:* возможность включения в анализ качественных переменных и оперирования нечеткими входными данными и лингвистическими критериями; быстрое моделирование сложных динамических систем и сравнение их с заданной степенью точности и др.

Процесс воздействия производственных факторов на работника труднопрогнозируем, зависит от многих обстоятельств и условий, поэтому процедура принятия решения по определению уровня профессионального риска представляет собой совокупность переменных различной природы, в связи с этим применительно к данному процессу целесообразно использовать модель нечеткого вывода. Таким образом, цель данной работы – разработка модели нечеткого логического вывода для оценки профессионального риска для здоровья работников.

Результаты и их обсуждение. В качестве нечеткой модели приняты три системы нечеткого вывода: СНВ₁, СНВ₂ и СНВ₃ (рис. 1).

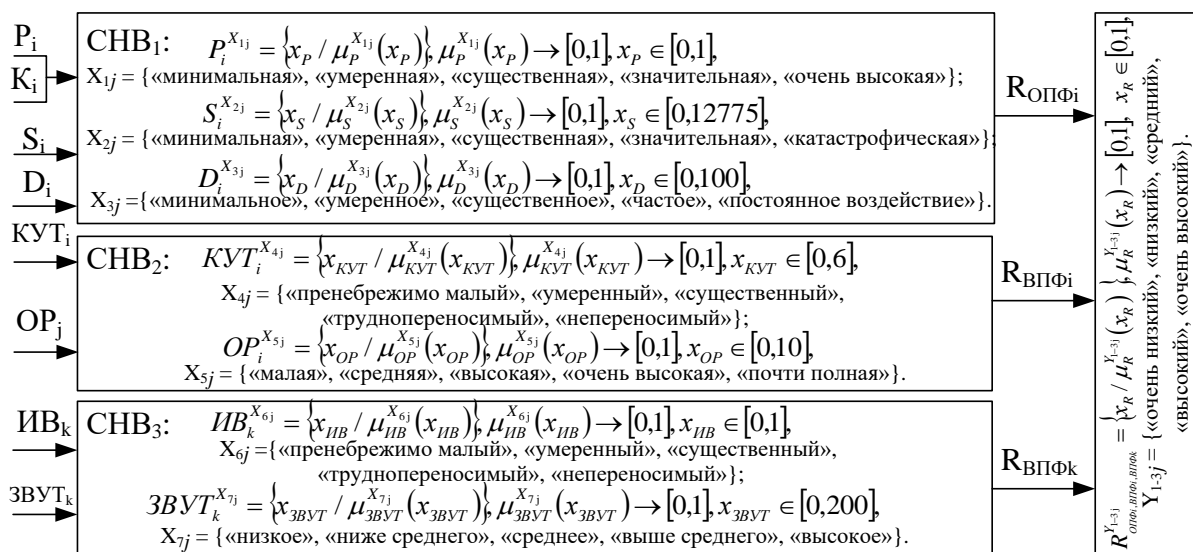


Рис. 1. Системы нечеткого вывода в модели оценки профессионального риска:
X_j, Y_j – значения лингвистических переменных (термы).
Выражение A = {x / μ_A(x)} – совокупность упорядоченных пар нечеткого подмножества A, где μ(x) – функция принадлежности значения базовой переменной x к подмножеству A

Входными переменными первой системы нечеткого вывода являются составляющие риска для i -й опасности: вероятность (частота) проявления опасности (P_i) с учетом давности происшествий (K_i), серьезность (тяжесть) последствий воздействия опасности (S_i) и длительность воздействия опасности (D_i). Коэффициент давности происшествий предложено учитывать при определении вероятности (частоты) проявления опасности, увеличивая его величину от 10 до 40 %. В качестве выходной переменной первой системы нечеткого вывода используется уровень профессионального риска воздействия опасного производственного фактора (ОПФ) по i -й опасности ($R_{ОПФ_i}$), который является основой для принятия решения о необходимости мер по управлению риском.

Для второй системы нечеткого вывода приняты две переменные: *класс условий труда* – $KУТ_i$ (оптимальные – 1; допустимые – 2; вредные – 3.1, 3.2, 3.3, 3.4; опасные – 4) по i -й опасности, полученный в результате аттестации рабочих мест по условиям труда (АРМ) и характеризующий меру вероятности вреда для здоровья при воздействии вредного производственного фактора (ВПФ), и *относительный риск* ($ОР_j$) для j -го класса болезней, получаемый в результате анализа ЗВУТ и выступающий в качестве критерия тяжести. Результатом нечеткого вывода второй системы является лингвистическая переменная – «профессиональный риск воздействия ВПФ» по i -й опасности ($R_{ВПФ_i}$), с помощью которой принимается решение о необходимости разработки эффективных превентивных управленческих решений.

Первой переменной третьей системы нечеткого вывода является предложенный показатель – индекс вредности ($ИВ_k$) для k -й профессии (должности) либо для структурного подразделения (установки, цеха, производства), характеризующий комплексное воздействие факторов производственной среды, определяемых в результате аттестации рабочих мест по условиям труда, и выступающий в роли вероятности воздействия ВПФ; второй переменной $СНВ_3$ принят показатель, характеризующий тяжесть вреда для всех систем организма, – число случаев временной нетрудоспособности по всем болезням, на 100 работающих ($ЗВУТ_k$) для k -й профессии (должности). На выходе третьей системы нечеткого вывода будет лингвистическая переменная – «профессиональный риск комплексного воздействия ВПФ» по i -й опасности ($R_{ВПФ_k}$), характеризующая опасность производственной среды для профессий (должностей) либо всего структурного подразделения (установки, цеха, производства).

Для адекватного определения суммарной вредности условий труда, учитывающей влияние всех производственных факторов, предлагается ввести дополнительный показатель для постоянных рабочих мест – *индекс вредности (ИВ)* [6 – 9], определяемый с округлением до пятого знака по формуле:

$$ИВ = \frac{0,00001 \cdot \sum_{i=1}^n O_i^{\text{класс } 1}}{1 + \sum_{i=1}^n O_i^{\text{класс } 1}} + \frac{0,0001 \cdot \sum_{i=1}^n D_i^{\text{класс } 2}}{1 + \sum_{i=1}^n D_i^{\text{класс } 2}} + \frac{0,001 \cdot \sum_{i=1}^n B_i^{\text{класс } 3.1}}{3 + \sum_{i=1}^n B_i^{\text{класс } 3.1}} + \frac{0,1 \cdot \sum_{i=1}^n B_i^{\text{класс } 3.3}}{2 + \sum_{i=1}^n B_i^{\text{класс } 3.3}} + \frac{1 \cdot \sum_{i=1}^n B_i^{\text{класс } 3.4}}{1 + \sum_{i=1}^n B_i^{\text{класс } 3.4}}, \quad (1)$$

где $O_i^{\text{класс } 1}$, $D_i^{\text{класс } 2}$ – количество i -тых производственных факторов, по которым установлены *оптимальные* условия труда класса 1 и *допустимые* условия труда класса 2 соответственно; $B_i^{\text{класс } 3.1}$, $B_i^{\text{класс } 3.2}$, $B_i^{\text{класс } 3.3}$, $B_i^{\text{класс } 3.4}$ – количество i -тых производственных факторов, по которым установлены *вредные* условия труда класса 3.1, 3.2, 3.3, 3.4 соответственно; 0,00001; 0,0001; 0,001; 0,01; 0,1; 1 и числа в знаменателе дробей 1,2,3 – коэффициенты, нивелирующие неравнозначность опасности условий труда при возрастании их класса и необходимые для выполнения условия не применения каскадного суммирования в итоговой оценке; n – общее количество определяемых производственных факторов на рабочем месте.

Индекс вредности не предполагает учет *опасных (экстремальных)* условий труда (класса 4), так как работа в таких условиях не допустима за исключением ликвидации аварий, проведения экстренных работ для предупреждения аварийных ситуаций, при этом работы проводятся в соответствующих средствах индивидуальной защиты и при соблюдении регламентированных режимов.

Значение ИВ изменяется в диапазоне от 0 до 1 и характеризует степень соответствия общего уровня производственных факторов действующим гигиеническим нормативам и стандартам с учетом неравнозначности проявления этих факторов. Величина ИВ может выступать в качестве основного показателя при априорной оценке категории подозреваемого профессионального риска в соответствии с данными таблицы 1.

В настоящее время существует несколько алгоритмов нечеткого вывода, наиболее известными из которых являются алгоритмы Мамдани, Цукамото, Сугено и Ларсена [10]. Принят алгоритм Мамдани, который один из первых нашел применение в системах нечетких множеств. Основными этапами нечеткого вывода по данному алгоритму являются [4]: формирование базы правил системы нечеткого вывода; фаззификация входных параметров; агрегирование; активизация подусловий в нечетких правилах продукций; аккумуляция заключений; дефаззификация.

Таблица 1

Критерии оценки подозреваемого профессионального риска по ИВ

Границы ИВ	Категория подозреваемого профессионального риска
$0 < ИВ < 0,00005$	Риск для здоровья отсутствует
$0,00005 \leq ИВ < 0,00025$	Пренебрежимо малый (переносимый, допустимый) риск для здоровья
$0,00025 \leq ИВ < 0,0005$	Малый (умеренный) риск для здоровья, возможны восстанавливающиеся функциональные изменения в организме
$0,0005 \leq ИВ < 0,005$	Средний (существенный) риск для здоровья, возможен рост производственно обусловленной заболеваемости, развитие начальных признаков либо легких форм профессиональных заболеваний
$0,005 \leq ИВ < 0,05$	Высокий (труднопереносимый) риск для здоровья, возможен рост производственно обусловленных хронических патологий и появление профессиональных болезней легкой и средней степеней тяжести
$1 < ИВ \leq 0,05$	Очень высокий (непереносимый) риск для здоровья, возможно развитие тяжелых форм профессиональных заболеваний, высокие уровни заболеваемости с временной утратой трудоспособности и значительный рост числа хронических патологий

Этап формирования базы правил систем нечеткого вывода заключается в представлении эмпирических либо экспертных знаний в виде конечного множества правил нечетких продукций. При этом в каждом из нечетких высказываний должны быть определены функции принадлежности значений термножества для каждой лингвистической переменной. Введем лингвистические переменные для реализации алгоритма нечеткого вывода Мамдани.

В качестве термножества $СНВ_1$ входной переменной «Вероятность проявления опасности» будем использовать множество $X_1 = \{\text{«минимальная», «умеренная», «существенная», «значительная», «очень высокая»}\}$. Числовое значение переменной $[0, 1]$ и определяется по коэффициенту частоты производственного травматизма. Коэффициент частоты производственного травматизма K_q показывает количество несчастных случаев на производстве, происходящих на 1000 среднесписочного числа работающих за определенный период (месяц, квартал, год), рассчитывается по формуле [11]:

$$K_q = \frac{N_{НС}}{N_p} \cdot 1000, \tag{2}$$

где $N_{НС}$ – количество пострадавших от НС на производстве с нетрудоспособностью на один рабочий день и более, в том числе с летальным исходом в отчетный период; N_p – среднесписочная численность работников.

В качестве термножества $СНВ_1$ входной переменной «Серьезность последствий воздействия опасности» будем использовать множество $X_2 = \{\text{«минимальная», «умеренная», «существенная», «значительная», «катастрофическая»}\}$. Числовое значение переменной $[0, 12775]$. Граница интервала (значение в днях) эквивалентна потере 35 лет трудоспособности вследствие смертельной травмы.

В качестве термножества $СНВ_1$ входной переменной «Длительность воздействия опасности» используется множество $X_3 = \{\text{«минимальная», «умеренная», «существенная», «частое», «постоянное воздействие»}\}$. Числовое значение переменной принадлежит интервалу $[0, 100]$ и определяется в процентах продолжительности воздействия опасности за смену или рабочий день.

В качестве термножества $СНВ_2$ входной переменной «Класс условий труда» будем использовать множество $X_4 = \{\text{«пренебрежимо малый», «умеренный», «существенный», «труднопереносимый», «непереносимый»}\}$. Числовое значение переменной $[0, 6]$ и определяется аппроксимированием классов условий труда в виде балльной оценки по следующей шкале [12]: оптимальные – 1 балл; допустимые – 2 балла; вредные класса 3.1 – 3 балла; вредные класса 3.2 – 4 балла; вредные класса 3.3 – 5 баллов; вредные класса 3.4 – 6 баллов (опасные условия труда класса 4 не рассматриваются).

В качестве термножества $СНВ_2$ входной переменной «Относительный риск» будем использовать множество по степени производственной обусловленности заболеваний $X_5 = \{\text{«малая», «средняя», «высокая», «очень высокая», «почти полная»}\}$. Числовое значение переменной $[0, 10]$ и определяется по принятой в медицине труда шкале [13].

В качестве термножества $СНВ_3$ входной переменной «Число случаев временной нетрудоспособности по всем болезням, на 100 работающих» будем использовать множество $X_6 = \{\text{«низкое», «ниже среднего», «среднее», «выше среднего», «высокое»}\}$. Числовое значение переменной $[0, 200]$ и определяется по шкале Е.Л. Ноткина [13].

В качестве термножества $СНВ_3$ входной переменной «Индекс вредности» будем использовать множество $X_7 = \{\text{«пренебрежимо малый», «умеренный», «существенный», «труднопереносимый», «непереносимый»}\}$. Числовое значение переменной $[0, 1]$ и определяется по формуле (1).

В качестве терм-множества выходной переменной $СНВ_1 - СНВ_3$ «Профессиональный риск» используется множество $Y_{1-3} = \{«очень низкий», «низкий», «средний», «высокий», «очень высокий»\}$. Числовое значение выходной переменной $[0, 1]$.

Сформулированы базы правил систем нечеткого вывода в виде совокупности нечетких предикатных правил вида:

$$P_1: \text{если } x_1 \text{ есть } X_1 \text{ и } x_2 \text{ есть } X_2, \text{ то } y \text{ есть } Y_1,$$

$$P_i: \text{если } x_1 \text{ есть } X_i \text{ и } x_2 \text{ есть } X_i, \text{ то } y \text{ есть } Y_i, i = 1, k, \quad (3)$$

где x_i – входные переменные; y – выходная переменная; X_i, Y_i – значения лингвистических переменных (термы).

Правила для $СНВ_1$ с учетом всех возможных комбинаций:

P_1 : если «Вероятность проявления опасности» есть «минимальная» и «Серьезность последствий воздействия опасности» есть «минимальная» и «Длительность воздействия опасности» есть «минимальная», то «Профессиональный риск воздействия ОПФ» есть «очень низкий»;

P_2 : если «Вероятность проявления опасности» есть «минимальная» и «Серьезность последствий воздействия опасности» есть «минимальная» и «Длительность воздействия опасности» есть «умеренная», то «Профессиональный риск воздействия ОПФ» есть «очень низкий»;

P_3 : если «Вероятность проявления опасности» есть «минимальная» и «Серьезность последствий воздействия опасности» есть «минимальная» и «Длительность воздействия опасности» есть «существенный», то «Профессиональный риск воздействия ОПФ» есть «очень низкий»;

P_4 : если «Вероятность проявления опасности» есть «минимальная» и «Серьезность последствий воздействия опасности» есть «минимальная» и «Длительность воздействия опасности» есть «частое», то «Профессиональный риск воздействия ОПФ» есть «низкий»;

P_5 : если «Вероятность проявления опасности» есть «минимальная» и «Серьезность последствий воздействия опасности» есть «минимальная» и «Длительность воздействия опасности» есть «постоянное воздействие», то «Профессиональный риск воздействия ОПФ» есть «низкий»;

...

P_{125} : если «Вероятность проявления опасности» есть «очень высокая» и «Серьезность последствий воздействия опасности» есть «катастрофическая» и «Длительность воздействия опасности» есть «постоянное воздействие», то «Профессиональный риск воздействия ОПФ» есть «очень высокий».

Базы правил системы нечеткого вывода $СНВ_2$ и $СНВ_3$ представлены в матричном виде в таблицах 2 и 3 соответственно.

Таблица 2

База правил системы нечеткого вывода $СНВ_2$

Относительный риск	Степень производственной обусловленности заболеваний				
	малая	средняя	высокая	очень высокая	почти полная
Класс условий труда	Категория профессионального риска				
Пренебрежимо малый	очень низкий	низкий	низкий	средний	средний
Умеренный	низкий	низкий	средний	средний	высокий
Существенный	низкий	средний	средний	высокий	высокий
Труднопереносимый	средний	средний	высокий	высокий	очень высокий
Непереносимый	средний	высокий	высокий	очень высокий	очень высокий

Таблица 3

База правил системы нечеткого вывода $СНВ_3$

Заболееваемость с УВТ	Число случаев временной нетрудоспособности по всем болезням на 100 работающих				
	низкое	ниже среднего	среднее	выше среднего	высокое
Индекс вредности	Категория профессионального риска				
Пренебрежимо малый	очень низкий	низкий	низкий	средний	средний
Умеренный	низкий	низкий	средний	средний	высокий
Существенный	низкий	средний	средний	высокий	высокий
Труднопереносимый	средний	средний	высокий	высокий	очень высокий
Непереносимый	средний	высокий	высокий	очень высокий	очень высокий

На основе экспертных оценок и принципа лингвистического распознавания образов установлено, что наиболее полно описывают изменение входных переменных термы, распределенные по треугольной функции принадлежности (кроме входных переменных OP_j и IB_k и выходных переменных $R_{ОПФ_i}$, $R_{ВПФ_i}$, $R_{ВПФ_k}$, для которых характерны трапециевидные функции принадлежности).

Графики треугольных функций принадлежности для термов входных и выходных лингвистических переменных первой системы нечеткого вывода представлены на рисунках 2 – 5.

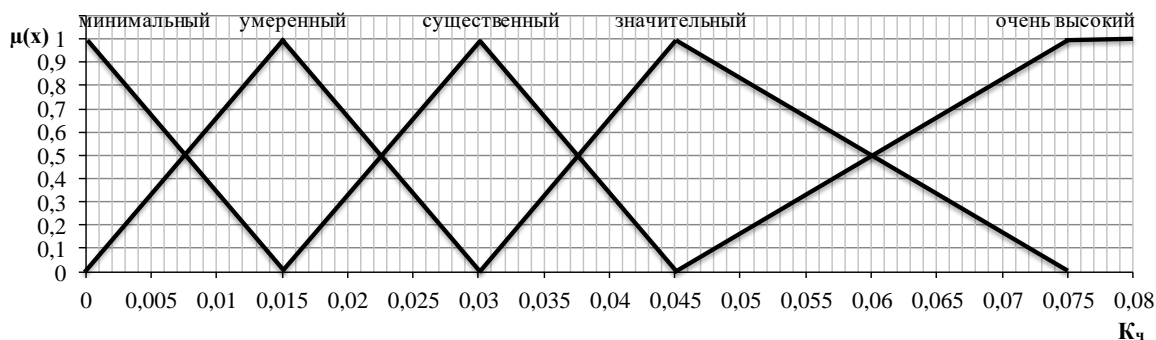


Рис. 2. Графики функций принадлежности для термов входной лингвистической переменной $СНВ_4$ «Вероятность проявления опасности»

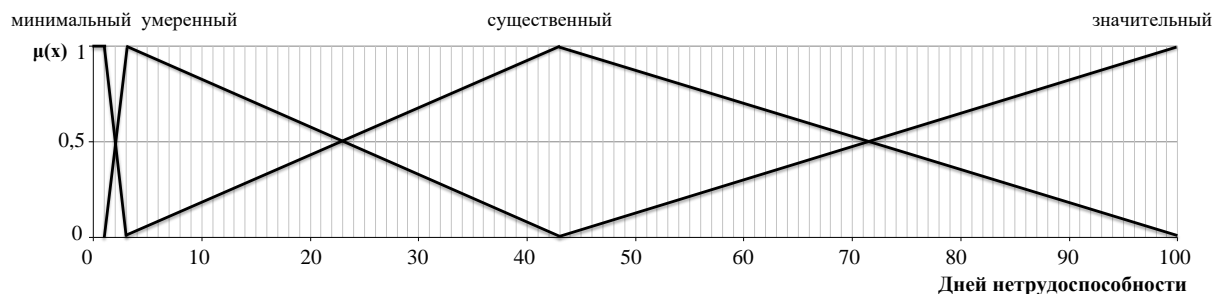


Рис. 3. Графики функций принадлежности для термов входной лингвистической переменной $СНВ_1$ «Серьезность последствий воздействия опасности»

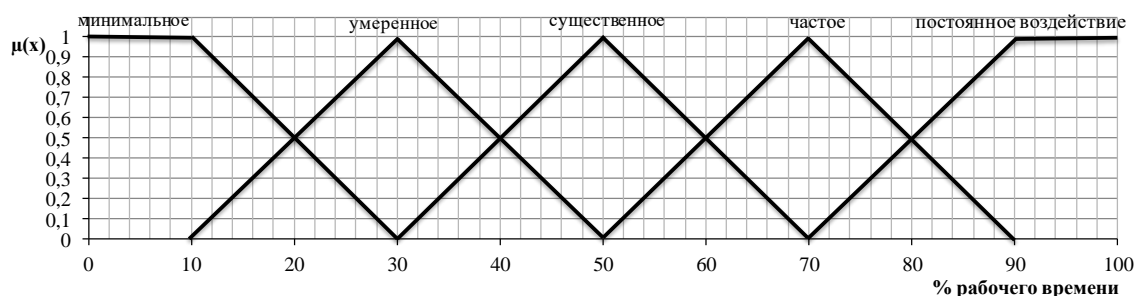


Рис. 4. Графики функций принадлежности для термов входной лингвистической переменной $СНВ_1$ «Длительность воздействия опасности»

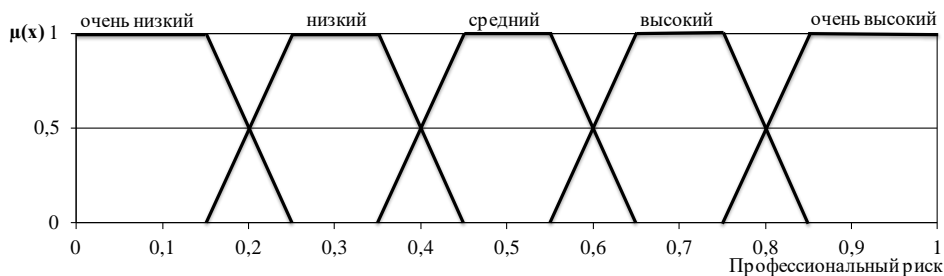


Рис. 5. Графики функций принадлежности для термов выходных лингвистических переменных $СНВ_1 - СНВ_3$ «Профессиональный риск»

Этап агрегирования представляет собой процедуру определения степени истинности условий по каждому из правил системы нечеткого вывода. Если условие правила составное высказывание, определяется степень истинности условия на основе известных степеней истинности подусловий. Для определения результата нечеткой конъюнкции (связки «И») и для определения результата нечеткой дизъюнкции (связка «ИЛИ») существуют различные формулы [4]. Активизация в системах нечеткого вывода представляет собой процесс нахождения степени истинности каждого из подзаклучений правил нечетких продукций.

В нашем случае для соединения высказываний необходима \min -конъюнкция:

$$\mu'(y) = \min\{c, \mu(y)\}, \quad (4)$$

где c – произведение степени истинности условия на весовой коэффициент правила (принимается равным единице); $\mu(y)$ – степень истинности нечеткого высказывания о результате в правиле.

Находим уровни «отсечения» для предпосылок каждого из правил (с использованием операции \min). На этапе аккумуляции находим функции принадлежности для каждой из выходных лингвистических переменных множества, т.е. объединяем все степени истинности заключений для получения функции принадлежности каждой из выходных переменных с использованием операции \max [5].

Для выполнения дефазификации, т.е. нахождения обычного (четкого) значения для каждой из выходных лингвистических переменных множества, предложено применение метода центра тяжести (центроид площади) для дискретного множества значений функции принадлежности по формуле:

$$y = \frac{\int_{\min}^{\max} x \cdot \mu(x) dx}{\int_{\min}^{\max} \mu(x) dx}, \quad (5)$$

где y – результат дефазификации; x – переменная, соответствующая выходной лингвистической переменной; $\mu(x)$ – функция принадлежности нечеткого множества, соответствующего выходной переменной после этапа аккумуляции; \min и \max – левая и правая точки интервала носителя нечеткого множества рассматриваемой выходной переменной.

При дефазификации методом центра тяжести обычное (не нечеткое) значение выходной переменной равно абсциссе центра тяжести площади, ограниченной графиком кривой функции принадлежности соответствующей выходной переменной [4].

Для выполнения расчетов согласно предлагаемой модели разработан программный продукт* на языке программирования C# в среде разработки Microsoft Visual Studio 2010 Express Edition, упрощающий процедуру оценки профессионального риска.

Применение программного средства для существующих условий при оценке профессионального риска воздействия опасного производственного фактора по $СНВ_1$ (рассматривалась опасность получения химического и термического ожога, отравления вредными веществами) на оператора технологических установок производства смазочных масел и битумов НПЗ показало, что при остановке и подготовке оборудования к ремонту риск получения химического ожога ($R_{ОПФ} = 0,50$) «средний» со степенью уверенности 100 %, отравления вредными веществами ($R_{ОПФ} = 0,33$) – «низкий» с 100 %-й уверенностью, термического ожога ($R_{ОПФ} = 0,58$) – «средний» с 71 %-й уверенностью.

Аналогичные расчеты могут производиться для всех видов выполняемых работ.

Расчет риска снижения остроты слуха по $СНВ_2$ для операторов технологических установок различных производств показал, что на основных производствах (топлив и масел) ($R_{ВЛФ} = 0,35$) риск соответствует «низкому» со 100 %-й уверенностью, а для некоторых вспомогательных цехов, например товарно-сырьевого, ($R_{ВЛФ} = 0,75$) – «высокому» уровню со степенью уверенности 98 %.

В таблице 4 приведены данные по расчету профессионального риска комплексного воздействия вредных факторов производственной среды для работников различных производственных подразделений, полученные по модели $СНВ_3$ при помощи разработанной программы.

В соответствии с принятыми критериями приемлемости на НПЗ предложена и апробирована шкала оценки рисков (R) по нечеткому 01-классификатору, в соответствии с которой в диапазоне $0 \leq R < 0,6$ риски приемлемые, при $R \geq 0,6$ неприемлемые и требуют разработки мер по управлению ими, а при $R \geq 0,8$ необходимы срочные мероприятия по снижению либо устранению опасности.

* Совместная разработка с магистрантом Е.В. Сташевич.

Высокому риску комплексного воздействия факторов производственной среды подвержены работники товарно-сырьевого цеха, базы оборудования и ремонтного производства (см. табл. 4), что требует разработки управляющих превентивных воздействий по снижению уровня риска.

Таблица 4

Профессиональный риск комплексного воздействия факторов производственной среды для работников НПЗ

Структурное подразделение	Уровень риска	Степень уверенности, %	Категория риска
Нефтяных топлив и ароматики	0,47	100	Средний
Производство масел смазочных и битумы	0,51	100	Средний
Ремонтное производство	0,70	100	Высокий
Производство энергоснабжения и очистных сооружений	0,54	100	Средний
Товарно-сырьевой цех	0,70	100	Высокий
Цех электроснабжения	0,50	100	Средний
Цех контрольно-измерительных приборов и автоматики	0,56	94	Средний
Центральная лаборатория	0,50	100	Средний
Автотранспортный цех	0,51	100	Средний
Цех базы оборудования	0,70	100	Высокий
Цех приготовления и расфасовки органических растворителей	0,48	100	Средний
Производство присадок	0,54	100	Средний
НПЗ в целом	0,59	56	Средний

Для выполнения расчетов по предложенному методу также могут найти применение специализированные программные инструменты MatLab (пакет программ Fuzzy LogicToolbox), fuzzyTECH и др. Преимуществом специализированных программных инструментов является возможность визуального отображения полученных результатов.

На рисунке 6 показана трехмерная поверхность нечеткого вывода зависимости уровня риска снижения остроты слуха для операторов технологических установок производства смазочных масел и битумов НПЗ от входных переменных, построенная в программе fuzzyTECH 5.81d.

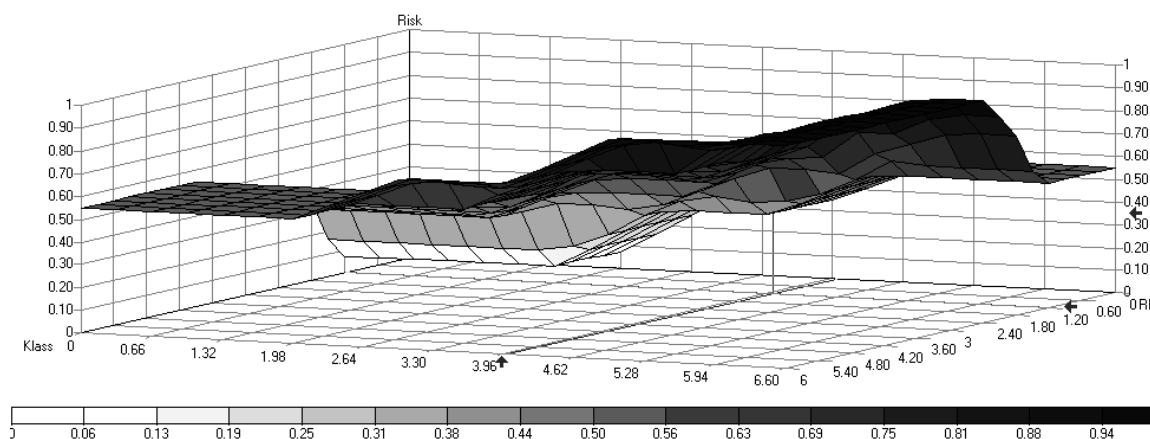


Рис. 8. Поверхность нечеткого вывода, построенная в программе fuzzyTECH 5.81d (klass – класс условий труда; RR – Relative Risk (относительный риск); Risk – риск)

Заключение. Применение предложенного концептуального подхода, реализованного в программном продукте для оценки профессионального риска для здоровья работников, позволит предопределять адекватные управленческие решения по устранению либо ограничению воздействия производственных факторов в условиях неопределенности, что повысит качество функционирования систем управления охраной труда в организациях.

ЛИТЕРАТУРА

1. Mahant, Narendra. Risk Assessment is Fuzzy Business – Fuzzy Logic Provides the Way to Assess Off-site Risk from Industrial Installations // Risk 2004. – 2004. – No. 206. – P. 39 – 44.

2. Tah, H.M. A proposal for construction project risk assessment using fuzzy logic / H.M. Tah, V.J. Carr // Construction Management & Economics. – 2000. – Vol. 18, № 4. – P. 491 – 500.
3. Алтунин, А.Е. Модели и алгоритмы принятия решений в нечетких условиях: моногр. / А.Е. Алтунин, М.В. Семухин. – Тюмень: Изд-во Тюмен. гос. ун-та, 2000. – 352 с.
4. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А. Леоненков. – СПб.: БХВ-Петербург, 2003. – 736 с.
5. Згуровский, М.З. Модели и методы принятия решений в нечетких условиях / М.З. Згуровский. – Киев: Наук. думка, 2011. – 275 с.
6. Булавка, Ю.А. Развитие комплексной оценки профессионального риска путем учета суммарной вредности условий труда / Ю.А. Булавка // Гигиена и санитария. – 2013. – № 4. – С. 47 – 54.
7. Булавка, Ю.А. Совершенствование методики оценки профессионального риска путем использования индекса вредности / Ю.А. Булавка // Сахаровские чтения 2013 года: экологические проблемы XXI века: материалы 13-й междунар. науч. конф., Минск, 16 – 17 мая 2013 г.; под ред. С.П. Кундаса, С.С. Позняка, Н.А. Лысухо. – Минск: МГЭУ им. А.Д. Сахарова, 2013. – С. 94.
8. Булавка, Ю.А. Аттестация рабочих мест по условиям труда как основная составляющая системы оценки профессиональных рисков / Ю.А. Булавка // Культура и безопасность в современном мире: материалы междисциплинар. науч.-практ. конф. с междунар. участием. – М.: Акад. ГПС МЧС России, 2013. – С. 10 – 13.
9. Булавка, Ю.А. Совершенствование априорной оценки профессиональных рисков путем учета суммарной вредности условий труда / Ю.А. Булавка // Проблемы техносферной безопасности-2013: материалы 2-й междунар. науч.-практ. конф. молодых ученых и спец. – М.: Акад. ГПС МЧС России, 2013. – С. 184 – 186.
10. Ярушкина, Н.Г. Основы теории нечетких и гибридных систем: учеб. пособие / Н.Г. Ярушкина. – М.: Финансы и статистика, 2004. – 320 с.
11. Булавка, Ю.А. Анализ производственного травматизма на нефтеперерабатывающем предприятии / Булавка Ю.А. // Вестн. Полоц. гос. ун-та. Сер. В. Промышленность. Прикладные науки. – 2011. – № 3. – С. 130 – 137.
12. Минько, В.М. Математическое моделирование в управлении охраной труда / В.М. Минько. – Калининград: ФГУИПП «Янтарный сказ», 2002. – 184 с.
13. Критерии оценки и показатели производственно обусловленной заболеваемости для комплексного анализа влияния условий труда на состояние здоровья работников, оценки профессионального риска: (инструкция) / Р.Д. Клебанов [и др.]. – Минск: ГУ РНМБ, 2009. – 33 с.

Поступила 23.09.2013

FUZZY-SET APPROACH FOR EXPERT OCCUPATIONAL RISK ASSESSMENT (AN EXAMPLE OF EMPLOYEES AT REFINERY)

Y. BULAUKA

Fuzzy model and the necessity of applying fuzzy-set approach to expert assessment of occupational risks to workers' health are presented. In order to improve the mechanism of occupational risk assessment one should use the tools of the theory of fuzzy inference. Its advantage is the ability to create a quantitative assessment of linguistic variables and effective display of the relationship between these two variables in the form of fuzzy rules. Application of the method of expert assessment of occupational risk for health of workers based on fuzzy-set approach by the example of the working conditions for employees of refineries is shown.

УДК 004.02

ОБЗОР АЛЬТЕРНАТИВНЫХ МЕТРИК ПРОГРАММНОГО КОДА ДЛЯ ПОИСКА ПЛАГИАТА

М.В. МАТЮШ

(Полоцкий государственный университет)

Представлен обзор альтернативных метрик программного кода. Анализ программного кода используется для решения актуальных задач определения качества, поиска плагиата и определения авторства исходного кода. Для каждой конкретной задачи могут подходить только определённые методы и алгоритмы анализа исходного текста программ. В свою очередь, для каждого метода и алгоритма требуется собственный, уникальный набор характеристик программного кода. На текущий момент существует множество методов и алгоритмов для анализа и поиска уникальных характеристик программного кода, часть из которых будет кратко представлена и обобщена в работе.

Актуальность анализа исходного кода. При создании программного обеспечения все затраты приходятся на создание первого образца, а дальнейшее его тиражирование стоит несравненно меньше, т.е. себестоимость программного обеспечения составляет стоимость исходного кода этого программного обеспечения. При модификациях программного обеспечения все изменения ведутся непосредственно над программным кодом. Анализ исходных кодов программ может использоваться для решения актуальных задач: определения качества, поиска плагиата, определения авторства.

Качество кода в свою очередь включает определение метрик, «читаемости» кода, документирования кода, надлежащего форматирования, количества ошибок проектирования. Качество кода в большинстве случаев определяет стоимость исходного кода. Качественный код легче поддается дальнейшим модификациям, содержит в себе меньше ошибок, может поддерживаться и модифицироваться программистами, не относящимися к числу авторов первоначального варианта.

Хищение программного продукта – преступление, связанное с использованием копии программного обеспечения без соблюдения лицензионного соглашения, по которому оно распространяется. Но хищение исходных кодов программного продукта – более тяжкое преступление. Поэтому поиск плагиата именно в исходных текстах программ является актуальной задачей. Качество поиска плагиата определяет возврат затрат на создание первой копии программного продукта.

В случае исходного кода вредоносного программного обеспечения задача правоохранительных органов состоит в определении авторства для дальнейшего привлечения к ответственности. Также определение первоначального авторства может быть актуальным при определении авторства исходного кода при двух и более лиц. Задачи поиска плагиата и определения авторства различные по реализации, хотя могут преследовать одну и ту же цель. Так, задача поиска плагиата подразумевает сравнение двух исходных текстов программ и поиск одинаковых блоков, а задача определения авторства – поиск общих черт у различных по назначению программ одного автора.

Цель данного исследования – обобщить известные и предложить теоретическое описание собственных, альтернативных метрик программного кода, которые используют совокупность стандартных характеристик и смогут охватить большой спектр решения задач анализа программного кода.

Определение качества программного кода. К стандартным характеристикам программного кода можно отнести качество. Качественные характеристики носят субъективный характер и для различных ситуаций определяются различными способами. Характеристики программного кода определяются на основе метрик – численных показателей, определяемых на основе программного кода. Однако зачастую разработчики программного обеспечения не в состоянии адекватно оценить значение той или иной метрики, поскольку не до конца понимают её природу [1].

Из основных видов метрик программного кода можно выделить [2]:

1) *количественные метрики.* Самой элементарной количественной метрикой является количество строк кода. Различают физические и логические строки, при этом к последним относят количество команд в программе. Также к количественным метрикам относят: количество пустых строк; количество комментариев; процент комментариев; среднее число строк для функций (классов, файлов); среднее число строк, содержащих исходный код для функций (классов, файлов); среднее число строк для модулей и многие другие характеристики, основанные на подсчете некоторых единиц в коде программы;

2) *метрики сложности потока управления программой.* Суть данного типа метрик заключается в строительстве ориентированного графа, содержащего лишь один вход и один выход. При этом вершины графа соотносят с теми участками кода программы, в которых имеются лишь последовательные вычисления и отсутствуют операторы ветвления и цикла, а дуги соотносят с переходами от блока к блоку и ветвями выполнения программы. Условие при построении данного графа: каждая вершина достижима из начальной и конечная – из любой другой вершины. Данная метрика определяет сложность программного кода;

3) *метрики сложности потока управления данными*. Необходимы для расчёта информационной прочности использования различных типов данных, к примеру, анализа характера использования переменных из списка ввода-вывода;

4) *метрики сложности потока управления и данных программы*. Данный класс метрик устанавливает сложность структуры программы как на основе количественных подсчетов, так и на основе анализа управляющих структур;

5) *объектно-ориентированные метрики*. Эти метрики основаны на анализе методов класса, дерева наследования и т.д. В этом классе можно выделить следующие метрики: центростремительное сцепление, т.е. количество классов вне этой категории, зависящих от классов внутри этой категории; центробежное сцепление, т.е. количество классов внутри этой категории, зависящих от классов вне этой категории; нестабильность; абстрактность; суммарная сложность всех методов класса; глубина дерева наследования; количество потомков; сцепление между классами, показывающее количество классов, с которыми связан исходный класс; множество методов, которые могут быть потенциально вызваны методом класса в ответ на данные; недостаток сцепления методов;

6) *метрики надежности*. Метрики, близкие к количественным, но основанные на количестве ошибок и дефектов в программе: количество структурных изменений, произведенных с момента прошлой проверки; количество ошибок, выявленных в ходе просмотра кода; количество ошибок, выявленных при тестировании программы, и количество необходимых структурных изменений, необходимых для корректной работы программы;

7) *гибридные метрики*. Метрики данного класса основываются на более простых метриках и представляют их взвешенную сумму.

Как правило, все перечисленные метрики являются показателями качества программного кода, но плохо подходят для поиска плагиата и определения авторства. В то же самое время совокупность метрик может определять другие нестандартные характеристики программного кода.

Стандартная задача поиска плагиата. Если рассматривать задачу плагиата исходных текстов программ, то здесь нужно четко разделять *плагиат дизайна программного обеспечения* и *плагиат исходных кодов*. Сходства программного продукта – основная причина анализа исходных кодов. Однако это не дает 100 % гарантии использования тех же самых исходных кодов. В данном случае похищен может быть только дизайн приложения. Следующей причиной анализа на плагиат может послужить общая или схожая функциональность двух приложений. Но схожая функциональность – это больше чем причина анализа, это неотъемлемое свойство плагиата, причем тут нужно понимать и видеть разницу между общим и модульным функционалом, так как могут быть различные назначения двух программных продуктов, т.е. различная общая функциональность, но при детальном анализе и делении на составные компоненты, каждый из которых имеет собственную функциональность.

Эти отдельные компоненты или группы могут совпадать по функциональности с другим программным обеспечением, и здесь может иметь место плагиат. Компоненты также могут делиться на более мелкие части-модули, и так до эмпирической константы, которая в случае исходного кода может быть функцией, процедурой или методом.

Более детальное рассмотрение до уровня переменных не имеет смысла, так как особенность исходных кодов в том, что используется конечный алгоритм, поэтому при анализе плагиата важен не столько лексический набор инструкций, сколько синтаксический.

Когда мы определили эмпирический минимум в виде функции, мы имеем в виду реализацию этой функции. Функции являются идентичными, если они идентичны по семантической составляющей. В качестве подобного минимума допустимо использовать отдельную строку кода в исходном файле. Это объясняется, с одной стороны, тем, что все основные операции во время кодирования, включая копирование и вставку, выполняются над одной или несколькими строками, а с другой – позволяет упростить предварительную обработку кода. Однако использование строк негативно сказывается на качестве обнаруживаемых клонов, а также существенно затрудняет определение синтаксических характеристик кода [2]. Наиболее предпочтительным и целесообразным является использование токенов в качестве единицы. Такой подход уменьшает количество ложных срабатываний, но приводит к существенному увеличению количества единиц кода в сравнении со строками кода и потребует реализации лексического анализатора для каждого поддерживаемого языка программирования. Введение единиц кода позволяет абстрагироваться от конкретных особенностей работы с исходным кодом. Наиболее предпочтительным и целесообразным является использование токенов в качестве единицы. Такой подход уменьшает количество ложных срабатываний, но приводит к существенному увеличению количества единиц кода в сравнении со строками кода и потребует реализации лексического анализатора для каждого поддерживаемого языка программирования. Введение единиц кода позволяет абстрагироваться от конкретных особенностей работы с исходным кодом. Но методы и алгоритмы, применяемые в поиске плагиата, не позволяют определить качество программного кода, а проверять качественный и не качественный программный код на плагиат нецелесообразно, так как присутствие плагиата в разных программах подразумевает, что эти программы схожи по своему качеству.

Особенности анализа программного кода. Исходные коды программ отличаются от обычного текста, им присущи специфические характеристики: структурированность, зависимость от входных данных, ограниченный набор лексем, часто повторяющиеся синтаксические конструкции и т.д. Применение стандартных алгоритмов поиска плагиата будет давать неверные результаты, так как в текстах программ всегда будут встречаться одинаковые конструкции, операторы, переменные, литералы и т.д. [3; 4]. Также стандартные алгоритмы поиска плагиата не учитывают паттерны проектирования, это такая повторяемая архитектурная особенность, представляющая собой решение проблемы проектирования в рамках некоторого часто возникающего контекста [5]. Очевидно, что использование одних и тех же правил проектирования исходного кода программ приводит к большой их схожести, затрудняя анализ исходных текстов на наличие плагиата. Далее предложим собственные характеристики с учетом ранее рассмотренной специфики программного кода.

Почерк программиста. В исходном тексте программ, помимо стандартных характеристик, сохраняется больше характеристик, свойственных конкретному автору (наименование переменных, стиль написания текста, лингвистические особенности комментариев, количество разных типов данных, стабильный уровень определенных метрик и др.) [3; 6 – 8], т. е. уместно будет употреблять понятие *почерк программиста*.

Задача определения почерка программиста заключается в поиске необязательных характеристик кода, которые присущи конкретному программисту, но это не является тривиальной задачей, как в случае с рукописным почерком. Почерк программиста может проявиться в архитектурном решении, например, в злоупотреблении каким-либо шаблоном проектирования, или при кодировании в правилах именования переменных, методов и других структурных единиц языка программирования, таких как:

- именование классов, методов и переменных, локальных переменных метода, параметров метода;
- место инициализации переменных класса; создание неиспользуемых методов класса (т.е. избыточность, пришедшая из проектирования, но не используемая в текущей реализации);
- количество комментариев, вид комментариев, место комментариев;
- длина имени методов; использование типов;
- использование перечислений;
- объявление интерфейсов, но не использование полиморфизма;
- употребление определённого шаблона во всех программах автора, определённого эталонного набора метрик программного кода, которые имеют одинаковое значение во всех программах конкретного программиста;
- и множество других.

Почерк программиста в первую очередь должен однозначно определять автора исходного кода, а также в задачах поиска плагиата находить почерк автора. Так как под почерком программиста мы также подразумеваем определённый поддерживаемый уровень метрик исходного кода, то качество почерка имеет эквивалентную зависимость с качеством программного кода. Таким образом, данная характеристика охватывает одновременно все три вышеперечисленные задачи: расчет качества кода, поиск плагиата и определение авторства.

Акцент программиста. Как и в случае с людьми, владеющими различными языками, у программистов также есть такое понятие, как «родной язык программирования». В отличие от языка общения, «родным» языком программирования является не тот, что изучался первым, а тот, на котором программист пишет постоянно, тот, на котором создавать программный продукт ему комфортно. И если случается, что программисту следует написать программу на не «родном» для него языке программирования, в исходном тексте этой программы могут появиться особенности, присущие только его «родному» языку программирования, т.е. проявится *акцент программиста*. Акцент программиста полностью не отделен от почерка программиста и является составной частью этой характеристики. Но в то же самое время выделение акцента программиста позволяет помимо большей точности в задачах поиска плагиата и определения авторства еще и проводить статистические исследования по миграции программистов с того или иного языка программирования. А также акцент позволяет определить отрицательный коэффициент соответствия стандартов программирования для исследуемого программного кода, что в свою очередь является качеством кода.

Эвристические методы для выделения акцента программиста должны выделить априорное множество обязательных и необязательных характеристик, присущих определённому языку программирования, сведенное в общий класс особенностей, и с последующим поиском этих особенностей в исходном коде с отличным языком программирования данного класса. Данная характеристика также может предполагать неопределенность в том случае, когда исходный код не соответствует ни одному классу акцентов из проверяемого множества, причем даже классу соответствующего языка программирования исходного кода, над которым будет осуществляться анализ. В этом случае считается, что это не акцент, а простая безграмотность программиста.

Грамотность программиста. Очевидным является то, что при поиске акцента программиста этот процесс становится частью процесса определения уникального почерка программиста. Поэтому обязательные стилистические особенности программирования не панацея к действию, так как изначально человек может придерживаться правил, и это из задачи поиска акцента перейдет в поиск почерка программиста. Если сравнивать с языком как словесным выражением мысли, то такие программисты являются неграмотными. Последнее предположение выделяет еще одну метрику – *грамотность программиста*.

Грамотность программиста – это соответствие исходного кода программиста стандартам, предложенным и повсеместно применяемым для конкретного языка. Эта задача актуальна, к примеру, при приеме на работу и при отслеживании качества программного кода. Решение задачи сводится к определению критериев грамотности, правил оформления и далее – последовательная проверка кода программиста на соответствие. К правилам может относиться именование переменных, функций или методов, далее принятая синтаксическая последовательность лексем языка. Также следует сразу отметить, что правила оформления для одного языка не соответствуют правилам оформления для другого. Из-за этого и появляется акцент при переходе от одного языка программирования к другому.

Грамотность программиста, с одной стороны, является противоположной характеристикой к акценту программирования, с другой – неотъемлемой частью почерка программиста, определяющего его качество. Но помимо этого грамотность определяет корректность инструкций и правил, которые характерны для большей части языков программирования.

Уникальный стиль программиста. Почерк программиста охватывает все остальные характеристики, но однозначно не определяет ту уникальную особенность, которая характерна именно для этого программиста. Эту особенность назовем *уникальным стилем программиста*, который будет рассчитываться путем поиска уникальных показателей ранее рассматриваемых характеристик и иметь как положительную, так и отрицательную тенденцию к изменению качества исходного кода. Типичными метриками, отражающими уникальный стиль программиста, могут быть нестандартное комментирование исходного кода, добавление собственных приставок и суффиксов к именам классов, методов и переменных.

Отличительной особенностью данной характеристики является то, что ее может намеренно изменять сам программист. Данная характеристика может быть специально введена программистом для защиты своего исходного кода. Также можно разделить стиль программирования на плохой и хороший по отношению к правилам оформления программного кода, а соответственно, и к качеству программного кода. Причем в этом контексте плохой может означать только, что автор исходного текста специально не следует общим правилам языка программирования, придерживается собственных по причине безграмотности или с целью оставить код уникальным. Хорошим стилем считается стиль, который не противоречит правилам оформления, но имеет избыточные комментарии или инструкции, являющиеся отличительными чертами автора.

Чистота программного кода. Положительная динамика данной характеристики является пропорционально зависимой от низких показателей акцента программиста, нулевых показателей уникальности стиля программиста, высоких показателей грамотности программиста и остальных хороших показателей почерка программиста. Предлагается считать программный код «чистым», если он идеален для конкретно решенной задачи. Основное затруднение в поиске данного критерия – это заранее определённые идеальные решения, что является утопичным. Поэтому предлагается применять данную характеристику только к определённым инструкциям и соответствию шаблонам проектирования.

Особенности командного программирования. Современные высоконагруженные проекты – результат работы не одного, а группы программистов. Разрабатываться командой могут как отдельные несвязанные модули, так и один класс (как синтаксическая единица языка). Эта особенность может потребовать применять вышеперечисленные метрики не к одному программисту, а к команде. В этом случае это будет общая усредненная характеристика всех участников команды. Также в случае командного программирования, если будут известны работы каждого программиста и при предварительном наполнении эталонных характеристик для каждого программиста, станет возможным проводить анализ на процент участия в написании исходного кода каждого программиста в команде.

Еще одной особенностью командного программирования является разный уровень навыков у каждого программиста, что выражается в различных характеристиках, а именно показателях почерка, грамотности, акцента и т.д. Данная особенность позволяет определить максимум и минимум, а также измерить динамику изменения характеристик в программном коде. Данный показатель позволит более качественно выделить общую динамику изменения показателей всей команды при изменении ее участников.

Также стоит отметить, что именно особенность командного программирования в совокупности с миграцией разработчиков между компаниями-производителями программного обеспечения и порождает

прецеденты плагиата. Подразумевается, что предлагаемые выше альтернативные метрики программного кода после их реализации смогут выявлять и такие нарушения авторского права.

В заключение проведенного исследования можно сделать следующие **выводы**:

- вопросы анализа программных кодов в виде решения трех задач: расчет качества, поиск плагиата и определение автора исходного кода – являются актуальными;
- приведены стандартные метрики исходного кода. Показана стандартная задача поиска плагиата в исходных текстах программ. Рассмотрены особенности анализа программного кода в сравнении с аналогичным литературным текстом;
- предложены собственные альтернативные метрики: почерк программиста, акцент программиста, грамотность программиста, уникальный стиль программиста и чистота исходного кода. Также показана ситуация анализа при командном программировании;
- представлен обобщенный обзор предлагаемых альтернативных метрик для дальнейшего продолжения исследования в данной области анализа исходного текста программных кодов, а именно планируется подготовка алгоритмического обеспечения по выделению каждой из предложенных характеристик с целью решения задач расчета качества, поиска плагиата и определения авторства.

ЛИТЕРАТУРА

1. Проблемы автоматизированного исследования качества программного кода // Журнал научных публикаций [Электронный ресурс]. – 2010. – Режим доступа: <http://jurnal.org/articles/2010/inf26.html>. – Дата доступа: 26.07.2013.
2. Метрики кода программного обеспечения // PVS-Studio [Электронный ресурс]. – 2008. – Режим доступа: <http://www.viva64.com/ru/a/0045/>. – Дата доступа: 01.08.2013.
3. Анализ алгоритмов выявления плагиата в кодах программ, написанных на языках высокого уровня // Современная техника и технологии: XVIII междунар. науч.-практ. конф. [Электронный ресурс]. – 2012. – Режим доступа: http://www.lib.tpu.ru/fulltext/v/Conferences/2012/C2/V2/v2_212.pdf. – Дата доступа: 10.07.2013.
4. Использование статистических характеристик программного кода в задачах дедупликации и оценки качества // Науч.-попул. интернет-журнал [Электронный ресурс]. – 2012. – Режим доступа: <http://novainfo.ru/archive/9/ispolzovanie-statisticheskikh-harakteristik-programmnogo-koda-v-zadachah-deduplikacii-i-ocenki-kachestva>. – Дата доступа: 25.07.2013.
5. Шаблон проектирования // Википедия: свободная энцикл. [Электронный ресурс]. – 2002. – Режим доступа: http://ru.wikipedia.org/wiki/Шаблон_проектирования. – Дата доступа: 20.09.2013.
6. Идентификация дублирования и плагиата в исходном тексте прикладных программ // Институт проблем управления РАН [Электронный ресурс]. – 2005. – Режим доступа: <http://lab18.ipu.ru/projects/conf2006/1/15.htm>. – Дата доступа: 22.08.2013.
7. Поиск плагиата в исходных текстах программ // Белорус. гос. ун-т [Электронный ресурс]. – 2011. – Режим доступа: <http://www.fpmi.bsu.by/ImgFpmi/Cache/36173.pdf>. – Дата доступа: 03.09.2013.
8. Метрики сложности программного обеспечения как критерии оценки плагиата в исходных кодах программ // Материалы междунар. науч.-практ. конф. [Электронный ресурс]. – 2012. – Режим доступа: <http://www.apriori-nauka.ru/uploads/files/RIBANOVK6.pdf>. – Дата доступа: 21.09.2013.

Поступила 23.09.2013

REVIEW OF ALTERNATIVE METRICS PROGRAM CODE FOR SEARCH OF PLAGIARISM

M. MATSIUSH

Analysis of the alternative metrics of program code is presented, which is used to solve the actual problems of determining the quality, searching plagiarism and authorship of source code. For each specific target can be approached only certain methods and algorithms of analyzing the source code of programs. In turn, each method and algorithm requires its own unique set of characteristics of program code. At the moment there are a lot of methods and algorithms for analyzing and searching for unique characteristics of the program code, part of which will be briefly presented and summarized in the article. The purpose of this paper is to summarize the well-known and provide a theoretical description of own, irregular specifications of the software code, which use a set of standard features and will be able to cover a wide range of solutions the problems of analyzing the program code.

УДК 519.71; 612.82

**PREDICTIVE AGENTS FOR CONTROL
OF INTELLIGENT MANUFACTURING ENTERPRISES****A. LINKEVICH, F. PANTELEENKO**
(*Polotsk State University*)

A general mathematical framework is put forward to model modern intelligent manufacturing enterprises that constitute distributed complex systems. The problem of control of such an object is reduced, to a considerable extent, to the task of forecasting future states of elements of the system, which amounts to the standard issue of prediction of time series. To cope with the latter problem, it is suggested to use frames in Hilbert space.

1. Introduction

A number of economical, social, biological, and other dynamical systems constitute ensembles of interacting agents that exhibit a reach variety of collective phenomena including deterministic chaos, fractals, self-organization, etc. Agents can have different beliefs, expectations, desires, intentions, strategies, and objectives. (The term *agent* is also used in another sense in computer science – see Section 2).

An important particular class is distributed dynamical systems that have little centralized control and communication between their agents, but still should behave in a prescribed manner (a variety of systems of such a kind are considered, e.g., in [1] and references available at

<http://www.ge.infn.it/econophysics> and

<http://www.unifr.ch/econophysics>).

The seminal paper [2] has introduced a simple model of adaptive economic agents and posed the so-called El Farol bar problem. In a generalized form the question is when elements (agents) of a distributed system do not operate “at cross purposes” and can achieve a prescribed global goal avoiding frustration. Treatment of actions in terms of the minority game [3] provides a fruitful approach to model market mechanisms and investigate self-organization and other collective phenomena in social and biological systems. Other theoretical tools are also successfully applied such as methods of quantum field theory [4] and gauge fields geometry [5], a self-similarity analysis [6] and so on.

The main aim of our work is to advance a framework to model similarly modern *manufacturing enterprises* (MEs) and to broaden thereby the area of application of methods of mathematical physics in addition to social and biological systems considered, e.g., in papers [1 – 8]. In a sense, our treatment of a ME is close in spirit to the minority game. So, in our model, like in markets, agents simultaneously and adaptively compete for limited resources of the ME, have heterogeneous strategies, expectations, knowledge, objectives, etc. Our second objective is to outline basic points of software applications developed together with our collaborators for real industrial employment.

The key idea of our approach is to provide each node (element, constituent) of a ME with an intelligent agent that should fulfil control of operation of the node. Such a controlling agent has to process available information to forecast the state of the ME, which appears efficient to elaborate decisions for the node.

A modern ME should meet a stiff competition, produce rapid responses to market changes, incorporate new technologies and other handy innovations (see [9; 10] and references therein). To cope with cost, quality and timeliness of products, MEs have to ensure integration and distributed organization of heterogeneous constituents, open and dynamic structure, integration of humans with software and hardware, interoperability, agility, scalability, fault tolerance, etc. To accomplish these and other objectives, *autonomous agents and multi-agent systems* (MASs) can be useful (for an introduction see, e.g., [10]). So, in distributed intelligent MEs, they can be used to encapsulate legacy software applications and integrate MEs’ activities with those of their suppliers, customers and partners into an open, distributed intelligent system via networking, represent manufacturing resources, products, operations, etc. to facilitate planning, scheduling and execution control, provide administration services and information management, facilitate communication, cooperation and coordination between agents, incorporate a whole scheduler or planner into manufacturing planning and scheduling systems.

An interesting particular concept of intelligent MEs is the *holonic manufacturing system* suggested in 1994 by the HMS consortium [11]. That is a ME whose elements (materials, machines, products, etc.) have autonomous and cooperative properties. In contrast to a centralized setup, activity of any element is determined by cooperation with other constituents. Intelligent software and hardware agents are used to represent such components of the ME.

Our paper is organized as follows. In the next section we sketch a standpoint on autonomous agents and multi-agent systems and overview two developed tools to handle uncertain information in the course of inter-agent communication. Section 3 contains a concise presentation of a software platform designed so as to support operation of controlling agents of a ME. In Section 4 we put forward a mathematical framework to model modern MEs composed of interacting constituents. Main properties of frames defined in Hilbert spaces are summarized in Section 5, while in Section 6 frames are used as a means to perform least squares approximation of experimental data and prediction of time series. The paper relies significantly on our previous works [10; 12 – 17] and some papers of our colleagues (see, e.g., [18]).

2. Autonomous Agents and Multi-Agent Systems

Agent-based computing is a promising paradigm of producing complex application software. Despite a great activity there is still much debate and little consensus concerning what agents are, though there appears some convergence of points of view. We adopt the following outlook on agents (for more detail see [10]). An agent can exercise control over its actions and operate independently of humans or others. An agent is provided with certain objectives and carries out flexible, content-dependent activity to achieve its objectives.

An agent may be characterized by mental attitudes. An agent acquires information about itself, other agents and its environment that constitutes its beliefs. The agent can experience different desires that compel to pursue different objectives, elaborate intentions, make and implement decisions.

An agent can interact with other agents (and humans) through communication by means of exchange of messages expressed in a knowledge language. The content of communication is meaningful statements about properties of agents and their environment, knowledge of agents, and their mental (propositional) attitudes in respect to the information.

A multi-agent system is designed so as to obtain an ensemble of interconnected agents that can operate beyond the capabilities of any individual. Complexity of tasks assigned to achieve together with the autonomy (self-control) of agents prevent forecasting and planning interrelations and foster the agents to make and implement flexible, content-dependent decisions on their interactions. Cooperation and competition among interacting agents enable self-organization of such a system. To endow controlling agents of a ME with ability to communicate with one another and with humans in accord with requirements of the international association FIPA (for more detail see, e.g., [10] and references therein), the model of FIPA compliant agents has been generalized in spirit of the fuzzy logic, and an interpreter of basic Prolog with an extended functionality has been offered [19; 10]. This program facilitate handling uncertain information by agents and should increase effectiveness of performance of the ME because information to be processed may naturally occur incomplete and imperfect due to a number of reasons.

Yet another means to cope with uncertain information and facilitate interaction of humans with agents, is the so-called “Regularized” English-Like Language (RELL) suggested in [10] for semantic treatment of information. To implement this way of communication, a first version of an interpreter has been developed [20] which allows translation (through a dialog) from the natural English to the RELL and vice versa.

3. Software Platforms to Support Operation of Agents

To provide an effective setup, a software platform has been suggested on the basis of economic principles modelled in such papers as [1 – 8]. This should enable, due to cooperation and competition between agents, self-organization of the system and thereby improvement of its performance [10]. The user of such a system is offered to suggest tasks from a certain problem domain and assign some points (money) to each of the tasks, while agents have mechanisms to generate the desire to gain points. A Task Manager Agent (TMA) of the system acts on behalf of the user and offer and allocate jobs between the agents on the basis of negotiations with them, and evaluate the results produced by the agents.

If an agent encounters difficulties in achieving a task it can ask other agents to help to tackle the job. This cooperation of agents becomes feasible through communication and negotiations, while competition between agents is caused by their attempts to gain points out of a limited total amount. This perfectly suits to models of a ME in which agents have to compete for limited resources of the system.

If the TMA regards that some task has successfully been achieved by an agent then the agent gets a certain amount of points. Each agent should pay for any resource and any service. An agent can obtain free of charge only information about resources and services provided on the platform. Certain points is automatically debited to support common services of the platform. If an agent has spent all its points, it can be eliminated from the game. In this way the Darwinist selection of most efficient agents is implemented similar to models of markets (see [1 – 8] for more detail).

In order to ensure proper operation of a complex ME, it is essential to carry out relevant measurements and adjustments. Such a system can include a number of elements so that parallel multi-channel signal processing is required. Moreover, extraordinary situations cannot be prevented when there are no regular routine ways to wrestle with emerging problems. Therefore development of systems for parallel signal processing and on-line decision-making is of great importance. Here we sketch a MAS intended to achieve such tasks [14].

Generally, data to be processed can fall into the following categories. *Calm data* pertain to cases when everything appears going well and no control influence is necessary, whereas *provoking data* signify appearance of tokens of abnormal states so that certain control actions are required. In particular, *extraordinary data* indicate a state of emergency when special actions are essential. *Monitoring data* should be processed so as to watch provoking signals. *Recording data* must be memorized. Provoking data may entail either a *routine* reaction or *intelligent* activity of the controller.

A meta-agent Manager (Master) governs interaction between all the other (working) agents (Slaves) and provides facilities for human control of the performance of the MAS.

Communication between agents is implemented through a common file that seems like an “ocean of information” simultaneously accessible to any agent.

4. A Model of an Intelligent Manufacturing Enterprise

Let us represent a manufacturing enterprise (ME) as a directed graph consisting of N uniformized generalized nodes that operate in discrete time $t=0,1,\dots$. It is convenient to enumerate all relevant utilities (products, services, workers, materials, information, energy, and other resources) that can, in principle, appear as an input or an output of a node and collect them in a set $U = \{u_\mu, \mu \in M\}$ where (multi) index μ may take on values from a set M .

Each node gets some utilities from other nodes, transforms them (or, in particular, keeps intact), and gives some available utilities to other nodes. Thus we suppose that a generalized node i at each time step t ($t=0,1,\dots$) has $m+1$ inputs $x_{i0\mu_0}^t, x_{i1\mu_1}^t, \dots, x_{im\mu_m}^t$, and $n+1$ outputs $y_{i0\nu_0}^t, y_{i1\nu_1}^t, \dots, y_{in\nu_n}^t$, where $x_{ij\mu}^t \in U$ is the utility of kind μ obtained at time t by node i from node j , while $y_{ik\nu}^t \in U$ is the utility of kind ν transferred at time t by node i to node k . Notice that input $x_{i0\mu}^t$ serves to denote utilities produced by node i by itself, and output $y_{i0\nu}^t$ correspond to utilities qualified as ultimate outputs of the ME as a whole.

Let us suppose further that control of operation of a node is performed by an intelligent, in a sense, agent endowed with abilities to gather, collect, and process relevant information so as to elaborate decisions for the node. Then it can appear advantageous if such an agent is able to predict values of so-called switching variables for future time moments given values of these quantities in the past. These variables are defined so as to indicate decisions made by node i .

Namely, the value $s_{ij\mu}^t = +1$ corresponds to the decision made by node i at time moment t to give (starting with the next time moment $t+1$) utility μ to node j , while $s_{ij\mu}^t = -1$ means the decision to try to get utility μ from node j (though this intention, can appear not successful).

We introduce also the probability $P_{ij\mu}(T)$ that the decision to transfer utility μ from node i to node j will be accomplished within time interval T .

It is obvious that not all of the switching variables are relevant for operation of a given node. It is reasonable therefore to determine, for any node i , vector $\sigma_i = (\sigma_{\alpha_1\beta_1\rho_1}, \sigma_{\alpha_2\beta_2\rho_2}, \dots, \sigma_{\alpha_{t_i}\beta_{t_i}\rho_{t_i}})$ whose components constitute an ordered subset of all quantities $s_{ij\mu}$ arranged according to attainability and importance for the agent (node) i to make and exploit forecasting. Then the task to be achieved by agent i reduces to the following: given values $\sigma_i^{t'}$ for $t'=t, t-1, \dots, t-T_p$, find values $\sigma_i^{t''}$ for $t''=t+1, \dots, t+T_f$.

It seems promising to treat any σ_i as an integer number, or even as an approximation of a rational number. In the latter case, a natural way to improve (if necessary) predictions is to enhance the accuracy of the approximation of these rational numbers by taking into account the states of more nodes.

Thus the problem of control of operation of nodes is reduced, in part, to the standard task of forecasting time series, and controlling agents should be provided with ability to make such predictions. It turns out that frames in Hilbert spaces appear a promising tool to accomplish forecasting relevant events.

5. Frames in Hilbert Spaces

As is known, a set of vectors $\{|h_a\rangle \in H, a \in A\}$ can constitute a basis in a vector space H if these vectors are linearly independent. This property allows to obtain the formula of the inverse transform of vectors. It is remarkable however that the above-mentioned condition can be weakened so that so-called *frames* may be used instead of *bases* for transforms of vectors, i.e. *frame* is a generalization of *basis*.

We accept the following definition [21; 15 – 17]. A *frame* in a Hilbert space H is a set of vectors $\{|h_a\rangle \in H, a \in A\}$ such that there exists a set of vectors $\{|h^a\rangle \in H, a \in A\}$ constituting a *reciprocal (dual) frame* with respect to the first one in the sense that jointly these sets provide resolution of unity in the form

$$I = \int_A d\mu(a) |h^a\rangle \langle h_a| = \int_A d\mu(a) |h_a\rangle \langle h^a|. \tag{1}$$

Here $\mu(a)$ is a Borel measure supported on a set A . These representations provide the inverse transforms for the expansion of vectors. So, if we expand a vector $|u\rangle \in H$ over the reciprocal frame $\{|h^a\rangle, a \in A\}$ so that $|u\rangle = \int_A d\mu(a) u_a |h^a\rangle$ then the components $u_a = \langle h_a | u \rangle$ are expressed through the vectors of the original frame $\{|h^a\rangle, a \in A\}$. And vice versa, if we take the representation $|u\rangle = \int_A d\mu(a) u^a |h^a\rangle$ then the coefficients are $u^a = \langle h^a | u \rangle$.

It is the main point how to construct sets of vectors $|h_a\rangle$ and $|h^a\rangle$ so as to provide frames in Hilbert space H that are reciprocal (dual) to each other. Suppose that there exists an operator T such that $|h^a\rangle = T |h_a\rangle, a \in A$.

Then the above problem amounts to finding this operator given a set of vectors $|h_a\rangle$. Let us introduce the so-called metric operator $G = \int_A d\mu(a) |h_a\rangle \langle h_a|$.

Then the required resolution of unity (1) is achieved if the operator T satisfies the equation $TG = I$. Therefore $T = G^{-1}$, and the problem is reduced to finding the inverse matrix G^{-1} if the label variable a can take on only discrete values.

6. Least Squares Approximation of Experimental Data and Forecasting Time Series Using Frames

Suppose that it is required to find such a function $y(x)$ that describes how an observable y depends on another observable x , i.e. this function should provide the best approximation of experimental data

$(x_1, y_1), \dots, (x_N, y_N)$ in the sense that the functional $E = \frac{1}{2} \sum_{i=1}^N \left(\frac{y_i - y(x_i)}{\sigma_i} \right)^2$ takes the minimum. Here σ_i is the standard (mean square) deviation of y_i . Then from the so-called normal equations one can find the coefficients

$$\text{of the expansion: } y^a = \langle h^a | \bar{y} \rangle = \sum_{i=1}^N \bar{h}^a(x_i) y_i / \sigma_i^2.$$

Suppose now that a scalar quantity $x(t)$ is measured at discrete times $t_0, t_1, \dots, t_k = k\tau, k = 0, 1, \dots$, which gives a time series x_0, x_1, \dots to be processed. The common approach to forecasting time series is to use vectors in the embedding space (for an introduction see, e.g., [22]). Then the problem can be formulated as follows.

Let $y \in \mathbf{R}^d$ be a d – dimensional vector composed of time delays of an observable x , i.e. $y_i \equiv y(t_i) \equiv y(i) = (x(i), x(i-1), \dots, x(i-d+1))$ and $x_i \equiv x(t_i) \equiv x(i)$. It is required to find a map $F(\cdot)$ that describes time evolution of y : $y_{i+1} = F(y_i), i = 0, 1, \dots$

Let us expand the function $F(y)$ over a frame $\{|h_a(y), a \in A\}$ so as $F(y) = \int_A d\mu(a) F^a h^a(y)$. Using

$$\text{the resolution of unity one can find } y_{i+1} = \sum_{j=1}^{N_i} \Delta(y_i, y_j) y_{j+1} / \sigma_j^2 \text{ where } \Delta(y, y') = \int_A d\mu(a) h_a(y) \bar{h}^a(y').$$

A similar form for the map $F(y)$ with a function $\Delta_M(y, y')$ of a particular form is suggested by the kernel density estimation approach (see [22] and references therein). Our framework yields an expression for the probability density that looks similarly but we observe also how this representation arises and what the origin of the kernel $\Delta_M(y, y')$ is.

7. Conclusions

In this paper we put forward a general mathematical framework to model MEs that constitute complex distributed systems composed of interacting elements (nodes). To ensure a proper performance of such a system, it seems promising to provide each sufficiently autonomous element (node) of the system with a controlling intelligent agent. Then the effective control of the ME as a whole reduces, to a considerable degree, to predictions of states of nodes interconnected with a given element. The latter issue is nothing, but the standard problem of forecasting time series and different methods can be borrowed (see, e.g., [22]).

In particular, frames in Hilbert spaces can turn out to be an effective means for making such predictions. This conclusion is based on our preliminary computer simulations. We employed yet a simple model specified by the widely used logistic map $ax(1-x)$ to describe behavior of adaptive, predictive and competitive agents in an evolving chaotic environment along the lines of the paper [23]. A more detailed presentation of our numeric results as well as description of more realistic modeling MEs will be given elsewhere.

It is worth noting that effectiveness of operation of a ME as a whole depends heavily on the extent to which constituents (agents) of the ME compete with one another, because it is just presence of competition along with cooperation that makes self-organization of a system possible. Also, it seems extremely interesting to investigate how agents can be learnt (trained) to be efficient and what statistical properties a system of such agents can exhibit in its performance after the learning. One can expect that theoretical findings presented in [1; 3; 16; 17] can be useful to this end. Now this work is in progress.

The authors are thankful to A.B. Antonevich, S.A. Astapchick, A.M. Krot, G.G. Krylov, S.P. Kundas, V.I. Kuvshinov, S.K. Rakhmanov, P.A. Vityaz for useful discussions and interest to the work.

REFERENCES

1. Wolpert D., K. Tumer An Introduction to Collective Intelligence. – Tech Report NASA-ARC-IC-99-63 (A shorter version of this paper is: J.M. Bradshaw (ed.) Handbook of Agent Technology (AAAI Press/MIT Press, 1999)).
2. Arthur W.B. Am. Econ. Assoc. Papers and Proc., 1994, p. 406 – 411.
3. Challet D., Zhang Y.-C. Physica A, 1997, vol. 246, no. 3, p. 407 – 418; Physica A, 1998, vol. 256, no. 4, p. 514 – 532; Caldarelli G., Marsili M., Zhang Y.-C. Europhys. Lett., 1997, vol. 40, no. 5, p. 479 – 484. Zhang Y.-C. Europhys. News, 1998, vol. 29, p. 51; Physica A, 1999, vol. 269, p. 30 – 44; Challet D., Marsili M., Zecchina R. Phys. Rev. Lett., 2000, vol. 84, no. 8, p. 1824 – 1827.
4. Barkema G.T., Howard M.J., Cardy J.L. Phys. Rev. E, 1996, vol. 92, p. 2017; Bak P., Paczuski M., Shubik M. Physica A, 1997, vol. 246, p. 430 – 453.
5. Ilinski K.J. Phys A, 2000, vol. 33, no. 1, P. L5 – L14.
6. Yukalov V.I. J. Math. Phys., 1991, vol. 32, p. 1235 – 1239; Gluzman S., Yukalov V.I. Phys. Rev. E, 1997, vol. 55, p. 3983 – 3999; Yukalov V.I., Gluzman S. Phys. Rev. E, 1998, vol. 58, p. 1359 – 1382; Andersen J.V., Gluzman S., Sornette D. Eur. Phys. J. B, 2000, vol. 14, p. 579 – 601.
7. Amaral L.A.N., Cizeau P., Gopikrishnan P., Liu Y., Meyer M., Peng C.-K., Stanley H.E. Computer Physics Commun., 1999, vol. 121 – 122, p. 145 – 152; Marsili M. Physica A, 1999, vol. 269, p. 9 – 15; Paczuski M., Bassler K.E., Corral A. Phys. Rev. Lett., 2000, vol. 84, no. 14, p. 3185 – 3188; Palmer R.G., Arthur W.B., Holland J.H., Le Baron B., Tayler P. Physica D, 1994, vol. 75, pp. 264-274; Savit R., Manuca R., Riolo R. Phys. Rev. Lett., 1999, vol. 82, no. 10, p. 2203 – 2206; Solomon S. Comp. Physics Commun., 1999, vol. 121 – 122, p. 161 – 167; Stanley H.E., Amaral L.A.N., Canning D., Gopikrishnan P., Lee Y, Liu Y. Physica A, 1999, vol. 269, p. 156 – 169.
8. Anderson P.W., Arrow K.J., Pines D. The Economy as an Evolving Complex System (Addison Wesley, Reading, MA, 1988).
9. Van Leeuwen E.H., Norrie D.H. Manufacturing Engineer, 1997, vol. 76, no. 2, p. 86 – 88; Shen W., Norrie D.H. Knowledge and Information Systems, 1999, vol. 1, no. 2, p. 129 – 156.

10. Linkevich A.D. Nonlinear Phenomena in Complex Systems, 2001, vol. 4, no. 1, p. 18 – 46.
Hayashi H. In: Proc. of 24th ISIR, Japan Industrial Robot Association [Electronic resource]. – Mode of access: <http://www.ims.org/>.
11. Konstantinov V.M., Linkevich A.D., Panteleenko F.I. An intelligent WWW-based multi-agent system for synthesis of resistant surfacing layers. – In: S.A. Astapchick and P.A. Vityaz (eds.). Theoretical and Technological Foundations of Hardening and Restoration of Machine-Building Articles (Technoprint, PSU, Minsk, 2001), p. 126 – 129.
12. Kundas S.P., Tsyrel'chuck I.N. Computational material science: A modern state and prospects of its development. – In: S.A. Astapchick and P.A. Vityaz (eds.). Theoretical and Technological Foundations of Hardening and Restoration of Machine-Building Articles (Technoprint, PSU, Minsk, 2001), pp. 41 – 47 (in Russian).
13. Kogan D.B., Linkevich A.D. In: Proc. Intern. Seminar “Modern Problems of Design and Production of Radioelectronic Means”, Novopolotsk, May 29 – 31, 2000 (Polotsk State University, Novopolotsk, 2000), p. 241 – 246.
14. Linkevich A.D. Nonlinear Phenomena in Complex Systems, 2000, vol. 3, no. 2, p. 135 – 152.
15. Linkevich A.D. Nonlinear Phenomena in Complex Systems, 2000, vol. 3, no. 3, p. 253 – 259; Frames in Hilbert Spaces: A Tool for Artificial Intelligence (to be published); Intelligent agents in the Internet: Using frames for learning, forecasting, and understanding (to be published).
16. Linkevich A.D. Mathematical Methods and Models for Investigation of Neurodynamical Mechanisms of Cognitive Processes (IEC NAS RB / PSU, Minsk / Novopolotsk, 2001).
17. Panteleyencko F., J. Siemanski. Some ways of creation of materials designated for tribotechnical processes. – In: Proc. of the Intern. Sci. Conf. “MECHANICS 2000”, June, 2000, Pzeszow, p. 273 – 283.
18. Gorokhov V.Yu., Linkevich A.D. (to be published).
19. Linkevich A.D., Putrov V.S. (to be published).
20. Daubechies I. Ten Lectures on Wavelets (SIAM, Philadelphia, 1992). Kaiser G. A Friendly Guide to Wavelets (Birkhauser, Boston, 1994).
21. Abarbanel H.D.I., Brown R., Sidorowich J.J., Tsimring L.S. Rev. Mod. Phys., 1993, vol. 65, p. 1331; Casdagli M., Eubank S. Nonlinear Modeling and Forecasting (Addison-Wesley, Redwood City, CA, 1992).
22. Pierre D., Hubler A. Physica D, 1994, vol. 75, p. 343 – 360.

Послупила 04.06.2013

PREDICTIVE AGENTS FOR CONTROL OF INTELLIGENT MANUFACTURING ENTERPRISES

A. LINKEVICH, F. PANTELEENKO

A general mathematical framework is put forward to model modern intelligent manufacturing enterprises that constitute distributed complex systems. The problem of control of such an object is reduced, to a considerable extent, to the task of forecasting future states of elements of the system, which amounts to the standard issue of prediction of time series. To cope with the latter problem, it is suggested to use frames in Hilbert space.

ФИЗИКА

УДК 621.365.46:621.396.6

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ИНФРАКРАСНОГО НАГРЕВА ДЛЯ МОНТАЖА ЭЛЕКТРОННЫХ КОМПОНЕНТОВ

д-р техн. наук, проф. В.Л. ЛАНИН, Т.Э. ЛАВОР, В.Т. ХОТЬКИН
(Белорусский государственный университет информатики и радиоэлектроники, Минск)

Рассматривается применение современных электронных компонентов в микроминиатюрном исполнении, значительно повышающих плотность монтажных соединений в электронных модулях при монтаже компонентов на печатных платах, что заставляет применять бесконтактные методы нагрева, способные повысить качество соединений и устранить перегрев компонентов. Локальный нагрев инфракрасным излучением позволяет достаточно точно дозировать энергию нагрева в зоне пайки при условии оптимизации его параметров. Моделированием параметров инфракрасного нагрева в системе SolidWorks показана эффективность совместного применения кварцевого концентратора и защитной маски, что уменьшает потери излучаемой энергии и способствует ее концентрации в зоне монтажа электронных компонентов на плате.

Введение. В связи с прогрессирующей миниатюризацией электронных компонентов и модулей все большее значение приобретает качество монтажа компонентов на печатных платах. Несмотря на то, что технологии пайки постоянно совершенствуются, требования к качеству электромонтажных соединений растут еще быстрее. Большие трудности возникают для электронных модулей с плотным размещением поверхностно-монтажных компонентов SMD (Surface Mounted Devices) на печатной плате, что заставляет искать новые технологии, способные повысить качество пайки, увеличить ее скорость и снизить стоимость. Это особенно важно для корпусов BGA и печатных плат, которые нагреваются первыми, а затем тепло передается контактными площадкам и сферическим BGA-выводам для формирования паяных соединений (рис. 1). Во избежание теплового удара компоненты не должны перегреваться, а контактные площадки и плавкие выводы BGA должны быть нагреты до температуры, достаточной для образования паяного соединения [1].

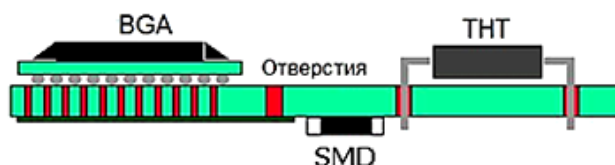


Рис. 1. Монтаж компонентов BGA, SMD и THT на плате

Одна из совершенных в настоящее время технологий пайки – локальная инфракрасная (ИК), когда нагрев производится сфокусированным пучком ИК-излучения только в местах пайки. Установки локальной ИК-пайки состоят из двух нагревателей, один из которых подогревает плату снизу до сравнительно невысокой температуры, а другой, верхний, осуществляет в нужный момент быстрый локальный нагрев требуемой области платы до температуры плавления припоя. Применение ИК-пайки является перспективным направлением в технологии поверхностного монтажа, так как обеспечивает уменьшение затрат на эксплуатацию оборудования при одновременном повышении качества паяных соединений [2].

Из всего спектра ИК-излучения с длинами волн в диапазоне 0,72...1000 мкм для пайки используется только малая часть этого диапазона: близкое излучение 0,7...1,5 мкм, среднее излучение 1,5...5,6 мкм и дальнее – 5,6...10,0 мкм. Длина волны ИК-излучения зависит от температуры источника излучения (при высоких температурах нагрева излучение имеет меньшую длину волны). Температура нагрева тела будет в общем случае зависеть от длины волны излучения, степени черноты или его отражающей способности, длительности облучения и от массы тела. Чем короче длина волны излучения, тем глубже оно проникает в тело, поэтому близкое ИК-излучение будет обеспечивать более глубокое его проникновение по сравнению со средним и дальним [3].

При практическом применении ИК-излучения возникли определенные проблемы:

- неравномерность нагрева различных компонентов на поверхности платы из-за разной степени поглощения энергии;

- невоспроизводимость режимов нагрева вследствие рассогласования спектра излучения источника и спектров поглощения компонентов и платы;
- наличие затененных участков платы, где нет прямого ИК-излучения;
- необходимость защиты ИК-излучателей от испаряющихся флюсов.

Первые разработки в области применения ИК-нагрева для пайки основывались на использовании в качестве излучателей ИК-ламп с температурой излучающего тела, превышающей 800...1100 °С. Так как температура пайки обычно находится в диапазоне 220...240 °С, то режимы нагрева значительно отличались от равновесных, что приводило к возникновению перегрева участков на поверхности платы, связанных с различной степенью черноты, массой компонентов и другими причинами [3]. Чаще всего ИК-установки работают в условиях воздушной атмосферы, однако присутствие кислорода в зоне пайки не всегда желательно, так как может происходить окисление припоя и разрушение органических материалов. Поэтому при пайке в воздушной атмосфере ограничивается длительность цикла нагрева, которая не должна превышать 100...150 с [4]. Перспективно применение при ИК-пайке азота с малым содержанием кислорода и водородно-азотной смеси. При этом улучшается качество пайки за счет исключения окисления припоя, повышения активности флюса, возрастает также краевой угол смачиваемости флюса, что уменьшает количество флюса, затекающего под корпуса элементов, остаточные ионные загрязнения на платах, упрощает очистку печатных плат.

Моделирование параметров инфракрасного нагрева. Экспериментальная проверка работоспособности ИК-устройств на реальных объектах довольно дорога, а зачастую невозможна не только по экономическим, но и по конструктивным соображениям. В таких случаях моделируют теплообмен с помощью прикладных пакетов. В модуле Flow Simulation, интегрированном в систему SolidWorks, можно моделировать геометрию ИК-нагревателей и выполнять тепловые расчеты и анализы «в одном окне». Это значительно снижает вероятность возникновения ошибок импорта/экспорта через промежуточный формат данных (SAT, IGES и т. д.).

Обычно ИК-излучатели имеют ограниченную площадь, при этом зона наибольшего нагрева находится непосредственно под нагревателем. Теплопередача излучением является эффективным и мощным механизмом передачи тепла:

$$T(K)E = bT^4, \tag{1}$$

где E – тепловая энергия излучения, пропорциональна четвертой степени абсолютной температуры; b – постоянная Стефана – Больцмана; T – температура источника излучения.

Длину волны ИК-излучения источника нагрева обычно выбирают среднюю из диапазона, для которой рассчитывают интенсивность излучения:

$$I = \frac{C_1}{\lambda_5} \left(\frac{C_2}{e^{\lambda T} - 1} \right)^{-1}, \tag{2}$$

где C_1, C_2 – постоянные Планка; λ – длина волны излучения.

Методика анализа ИК-нагрева в SolidWorks включает: создание геометрической модели, задание свойств материалов и граничных условий, настройку опций расчета, настройку сетки и просмотр результатов. Для решения поставленной задачи в программном пакете SolidWorks используют метод конечных объемов [5]:

$$\frac{\partial \rho \varphi}{\partial t} + \nabla(\rho U_\varphi) - \nabla(D_\varphi \varphi) = S_\varphi, \tag{3}$$

где $\frac{\partial \rho \varphi}{\partial t}$ – изменение физической величины φ ; $\nabla(\rho U_\varphi)$ – конвективное слагаемое в законе сохранения величины φ ; $\nabla(D_\varphi \varphi)$ – диффузное слагаемое в законе сохранения величины φ ; S_φ – источниковое слагаемое в законе сохранения величины φ .

Первым этапом алгоритма моделирования нагрева является создание трехмерной модели ИК-нагрева печатной платы с установленными на ней электронными компонентами (рис. 2, а). Для моделирования выбраны керамические ИК-панели (рис. 2, б), которые обладают возможностью отдавать высокую мощность при хорошей равномерности нагрева. Для оптимизации работы расчетного модуля Flow Simulation пакета SolidWorks Premiumx64 Edition [6] необходимо построить геометрическую модель, состоящую из печатной платы, паяемого компонента, верхнего и нижнего ИК-нагревателей.

Для достижения более эффективного использования излучаемой энергии применен кварцевый концентратор энергии, который локализует ИК-излучение от нагревателя на печатной плате.

На этапе подготовки моделирования процесса ИК-пайки устанавливают параметры расчетного домена (Computational Domain), указывают область, заполненную воздухом (Fluid Subdomain), выбирают материалы каждой детали сборки с необходимыми для инженерного расчета параметрами. Источниками

тепла (Heat Sources) являются нижний и верхний нагреватели, для которых выбирают значения мощности из диапазона ее варьирования. Далее указываются поверхности, получающие ИК-излучение от нагревателей (Radiative Surfaces), типы источников излучения (Radiative Sources), интенсивность их излучения и мощность. В месте пайки электронного компонента на плату указывается контактное тепловое сопротивление (Contact Resistances) для адекватности анализа. Для ускорения моделирования устанавливают начальную температуру печатной платы, равную предполагаемой температуре предварительного нагрева (Initial Conditions).

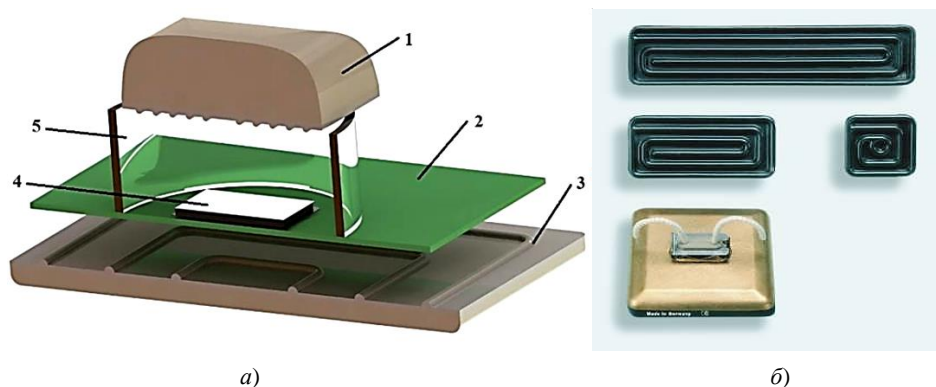


Рис. 2. Геометрическая модель ИК-нагрева (а) и керамические ИК-нагреватели (б):
1 – верхний нагреватель; 2 – печатная плата; 3 – нижний нагреватель;
4 – паяемый компонент; 5 – кварцевый концентратор энергии

Целевыми функциями оптимизации параметров ИК-нагревателей (при значениях мощности 500, 750 и 1000 Вт и на расстоянии от нагревателя до печатной платы 20 и 30 мм) являются более эффективное использование ИК-излучения и минимизация нагрева электронных компонентов для снижения теплового удара. Распределение температур на поверхности платы и на поверхности паяемого электронного компонента (на расстоянии до платы 20 мм) в обычном режиме приведено на рисунке 3 (а); при использовании кварцевого концентратора и маскирующей пластины – на рисунке 3, б.

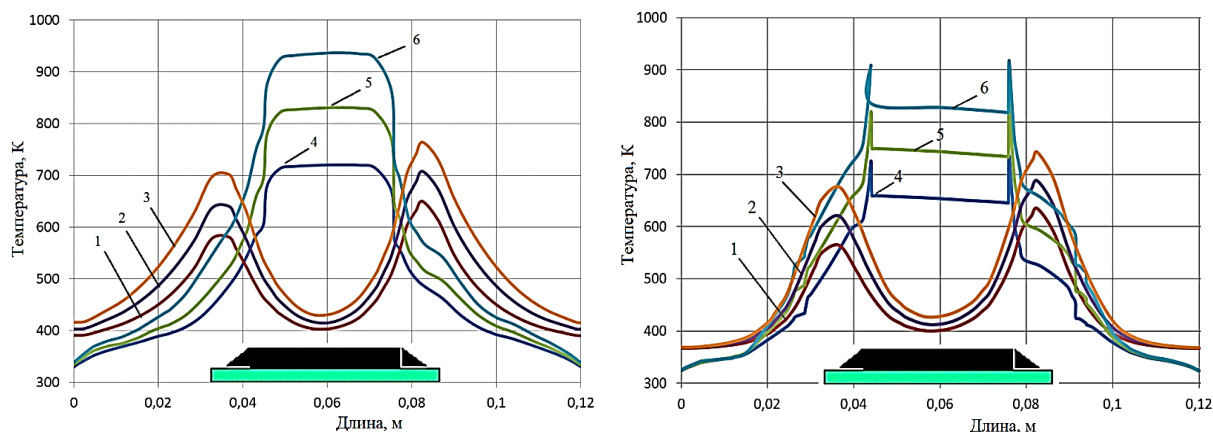


Рис. 3. Распределение температуры печатной платы (1, 2, 3) и электронного компонента (4, 5, 6) при мощности нагревателя 500, 750, 1000 Вт соответственно:
а – в обычном режиме; б – при использовании кварцевого концентратора и маскирующей пластины

Задача оптимизации температуры нагрева электронного компонента решается с учетом комбинаций влияющих факторов: x_1 – расстояния до платы, x_2 – мощности нагревателя методом наименьших квадратов в MathCAD.

Влияние комбинаций факторов на нагрев электронного компонента

x_1 , мм	25,50	26,00	26,50	27,00	27,50	28,00	28,50	29,00	29,50
x_2 , Вт	445,76	477,29	513,77	556,09	605,47	663,70	733,52	819,65	931,58
y , К	659,26	664,91	671,45	679,08	688,10	698,92	712,22	729,17	752,17

Полученную таблицу комбинаций факторов решением функции в заданном диапазоне аргументов с заданным шагом подставляют в функцию нагрева электронных компонентов и определяют ее поведение при воздействии различных комбинаций влияющих факторов.

Уравнения регрессии для нагрева печатной платы, полученные методом наименьших квадратов в пакете MathCAD: в обычном режиме (а), с кварцевым концентратором (b) и с кварцевым концентратором и маской (с), имеют вид:

$$a) \quad a(x_1, x_2) = 36,97x_1 + 0,602x_2 - 0,015x_1x_2 - 0,813x_1^2 - 2,702 \cdot 10^{-5} x_2^2;$$

$$b) \quad b(x_1, x_2) = 35,198x_1 + 0,454x_2 - 0,011x_1x_2 - 0,673x_1^2 - 1,366 \cdot 10^{-6} x_2^2;$$

$$c) \quad c(x_1, x_2) = 40,624x_1 + 0,216x_2 - 8,094 \cdot 10^{-3} x_1x_2 - 0,811x_1^2 - 1,364 \cdot 10^{-4} x_2^2.$$

Получены уравнения равного выхода для нагрева печатной платы на уровне 523 К для тех же условий:

$$f_a(x_1) = 11143,61 - 18507,83\sqrt{1,33 \cdot 10^{-4} x_1^2 - 0,014x_1 + 0,307} - 275,28x_1;$$

$$f_b(x_1) = 166124,341 - 366139,43\sqrt{1,12 \cdot 10^{-4} x_1^2 - 0,0096x_1 - 0,203} - 394,97x_1;$$

$$f_c(x_1) = 29,67 + 3665,59\sqrt{5,08 \cdot 10^{-4} x_1^2 - 0,026x_1 + 0,328} - 791,1.$$

Минимум нагрева паяемого электронного компонента достигнут при использовании кварцевого концентратора и алюминиевой маскирующей пластины (рис. 4).

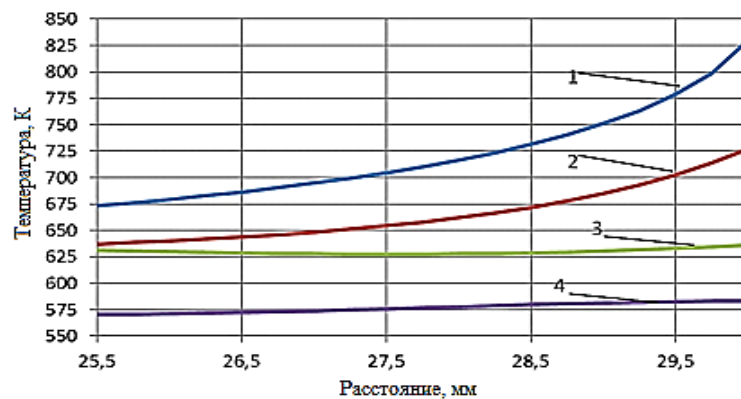


Рис. 4. Температура электронного компонента в зависимости от комбинаций влияющих факторов: 1 – обычный режим; 2 – с маской; 3 – с кварцевым концентратором; 4 – с кварцевым концентратором и маской

В диапазоне мощности ИК-нагревателя 500...1000 Вт оптимальное расстояние между нагревателем и платой равно 28,75 мм. Этому соответствует распределение температур, приведенное на рисунке 5.

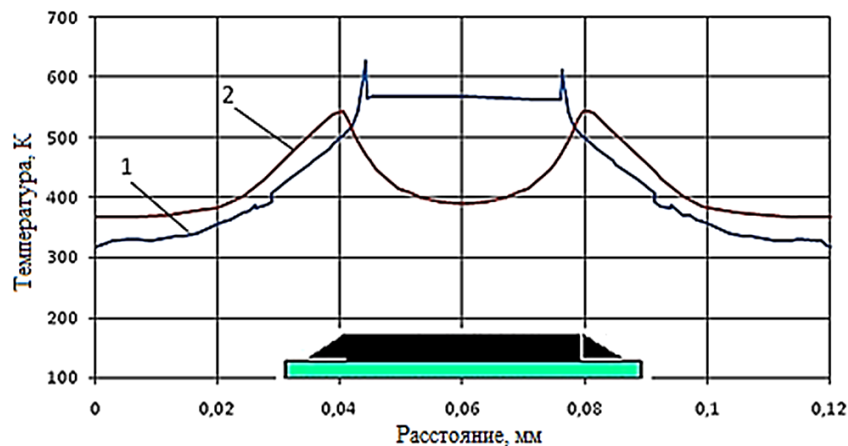


Рис. 5. Распределение температуры электронного компонента (1) и печатной платы (2) в оптимальном режиме

Таким образом, можно утверждать, что использование кварцевого концентратора уменьшает потери излучаемой ИК-энергии и способствует ее концентрированию на площадках для пайки. В результате температура нагрева в зоне монтажа компонента увеличивается на 13...15 %, а нагрев печатной платы снижается на 15...17 %.

Выводы

1. Бесконтактный нагрев инфракрасным излучением позволяет достаточно точно дозировать энергию нагрева в зоне пайки электронных компонентов при условии оптимизации его параметров. Применение SolidWorks для моделирования процесса нагрева ИК-излучением позволяет адекватно описывать реально происходящие процессы при пайке электронных компонентов на печатных платах, включая процессы нагрева компонентов, соединений и печатной платы.

2. В результате моделирования параметров инфракрасного нагрева установлено: применение кварцевого концентратора повышает эффективность процесса пайки, так как снижаются потери излучаемой ИК-энергии и происходит ее концентрация в рабочей зоне; в результате температура нагрева в зоне монтажа компонента увеличивается на 13...15 %, а нагрев платы снижается на 15...17 %; минимальный нагрев многовыводного электронного компонента в диапазоне мощности инфракрасного нагревателя 500...1000 Вт при оптимальном расстоянии от нагревателя до платы 28,75 мм и применении защитной маски.

ЛИТЕРАТУРА

1. Wassink, K.R. J. Soldering in Electronics / K.R.J. Wassink. – Ayr, Scotland. Electrochem. Publ. 2002. – 285 p.
2. Джюд, М., Пайка при сборке электронных модулей / М. Джюд, К. Бридли; пер. с англ. – М.: Издат. Дом «Технологии», 2006. – 416 с.
3. Ланин, В.Л. Электромонтажные соединения в электронике. Технология, оборудование, контроль качества / В.Л. Ланин, В.А. Емельянов. – Минск: Интегралполиграф, 2013. – 406 с.
4. Ланин, В.Л. Инфракрасный нагрев в технологии пайки изделий электроники / В.Л. Ланин // Электронная обработка материалов. – 2007. – № 5. – С. 91 – 96.
5. Патанкар, С.В. Численное решение задач теплопроводности и конвективного теплообмена при течи в каналах / С.В. Патанкар; пер. с англ. – М.: Изд-во МЭИ, 2003. – 312 с.
6. SolidWorks 2007-2008. Компьютерное моделирование в инженерной практике / А.А. Алямовский [и др.]. – СПб.: БХВ-Петербург, 2008. – 1040 с.

Поступила 13.09.2013

OPTIMIZATION OF INFRA-RED HEATING PARAMETERS FOR INSTALLATION OF ELECTRONIC COMPONENTS

V. LANIN, T. LAVOR, V. KHOTSKIN

Application of modern electronic components in micro miniature execution considerably raises density of assembly connections in electronic modules at installation of components on printed-circuit boards that forces to apply contactless methods of the heating, capable to raise quality of connections and to eliminate an overheat of components. Local heating by infra-red radiation allows to dose out precisely enough energy of heating to a zone of the soldering under condition of optimization of its parameters. By simulation of parameters of infra-red heating in SolidWorks efficiency of application of the quartz concentrator which reduces losses of radiated energy is shown and promotes its concentration in a zone of installation of electronic components on printed circuit board.

УДК 546.28:621.315.592

ИССЛЕДОВАНИЕ ПОВЕРХНОСТИ ПОЛИМЕРОВ МОДИФИЦИРОВАННОЙ РАДИАЦИОННОЙ ОБРАБОТКОЙ

*А.А. ХАРЧЕНКО, канд. физ.-мат. наук Д.И. БРИНКЕВИЧ,
канд. хим. наук С.Д. БРИНКЕВИЧ*

*(Белорусский государственный университет, Минск);
канд. физ.-мат. наук, доц. С.А. ВАБИЩЕВИЧ, Н.В. ВАБИЩЕВИЧ
(Полоцкий государственный университет);*

*канд. физ.-мат. наук, доц. М.Г. ЛУКАШЕВИЧ, канд. физ.-мат. наук В.С. ПРОСОЛОВИЧ,
д-р физ.-мат. наук, проф. В.Б. ОДЖАЕВ
(Белорусский государственный университет, Минск)*

Методом атомно-силовой микроскопии исследована модификация поверхности пленок полиимида, полиэтилентерефталата и полиэфирэфиркетона при γ -облучении и имплантации ионами Ni^+ , Ag^+ , Ar^+ , Fe^+ и V^+ с энергиями 30...100 кэВ в интервале доз $1 \cdot 10^{15}$... $1,5 \cdot 10^{17}$ см⁻². Обнаружено формирование при низких дозах облучения конусообразных структур высотой до 80 нм и диаметром в основании до 400 нм, хаотично расположенных на поверхности полимерных пленок. Их образование обусловлено преимущественной локализацией радиационно-химических процессов вблизи дефектов структуры (границы раздела фаз) полимера. Формирование указанных дефектов за пределами зоны имплантации обусловлено переносом возбуждения и/или миграцией дырок из области трека иона.

Введение. Среди разнообразных способов получения полимерных нанокомпозитов (золь-гель химический синтез, вакуумное осаждение, совместная плазменная полимеризация и др.) метод ионно-лучевой имплантации выделяется возможностью формирования наночастиц имплантируемой примеси на различных глубинах под поверхностью облучаемой матрицы при практически неограниченной концентрации вводимой примеси. При этом исходная молекулярная структура и элементный состав облучаемого полимера, а также режимы ионного облучения (тип и энергия иона, доза имплантации и др.) определяют микроструктуру, фазовый состав и, соответственно, физико-химические свойства формируемой нанокомпозитной пленки [1; 2]. Отметим, что интенсивное воздействие высокоэнергетических ионов на полимер приводит к появлению принципиально новых свойств у облученного приповерхностного слоя полимерной пленки.

В свете вышеизложенного особый интерес представляют исследования изменения морфологии поверхности полимеров в процессе облучения высокоэнергетичными частицами.

Цель настоящей работы – выявление особенностей модификации поверхности различных полимеров при ионной имплантации и γ -облучении.

Методика эксперимента. Промышленные пленки полиимида (ПИ), полиэтилентерефталата (ПЭТФ) и полиэфирэфиркетона (ПЭЭК) были имплантированы ионами Ni, Ag, Ar, Fe и V с энергиями 30...100 кэВ в интервале доз $1 \cdot 10^{15}$... $1,5 \cdot 10^{17}$ см⁻² в остаточном вакууме 10^{-5} Па на имплантаторе ИЛУ-3 при плотности ионного тока $j = 4$ мкА/см². Во избежание перегрева и деструкции образца в процессе имплантации использовалась кассета, обеспечивающая эффективный сток ионного заряда с поверхности полимера и плотный контакт пленок с металлическим основанием, охлаждаемым водой. При таких условиях в процессе имплантации температура образцов не превышала 350 К. Облучение полимерных пленок γ -квантами проводилось на установке МРХ- γ -25М с источником ⁶⁰Со. Мощность поглощенной дозы составляла $0,36 \pm 0,008$ Гр/с. Интервал поглощенных доз 6...2900 кГр.

Морфология поверхности модифицированной имплантацией полимерной пленки исследовалась методом атомно-силовой микроскопии (АСМ) при комнатной температуре в полуконтактном резонансном режиме на частоте 145 кГц на приборе Solver P-47. Одновременно измерялась среднеарифметическая шероховатость (R_a) поверхности. Использовались кантилеверы серии NSG 01 с радиусом закругления 10 нм.

Результаты и их обсуждение. При низкой поглощенной дозе наблюдалось «залечивание» каверн, существовавших в исходных образцах ПЭЭК и ПИ, что приводило к небольшому снижению шероховатости. Это характерно в основном для γ -облученных образцов. Вероятнее всего, это обусловлено тем, что энергия, поглощаемая при облучении, превышает потенциальный барьер конформационной перестройки молекул полимера. Молекулы становятся подвижными, поверхностная энергия минимизируется за счет уменьшения площади поверхности, что стимулирует «залечивание» каверн и, соответственно, снижение шероховатости. Возможны также рекристаллизация поверхностного слоя и движение участков поверхности в масштабе сферолитов [6].

Облучение γ -квантами ориентированных пленок полиимида приводило к формированию волнистой поверхности (рис. 1). Это можно объяснить следующим: при выравнивании поверхности валик давит

на выступающие части пленки, при этом происходит частичная кристаллизация и ориентация полимера. Отметим, что радиационная деструкция полимеров приводит к значительному разбуханию полимера [3]. С другой стороны, при γ -облучении различных полимеров имеет место высокая локальная неоднородность процессов радиационно-индуцированной деструкции и сшивания полимеров [4]. Радиационные явления наиболее интенсивно протекают в аморфных областях.

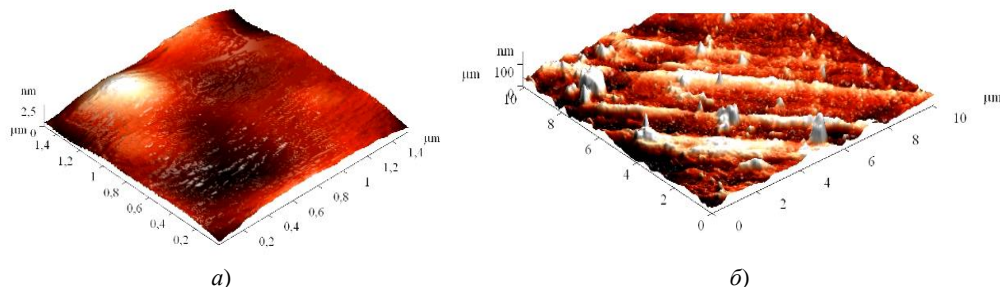


Рис. 1. Трехмерное АСМ-изображение поверхности пленки исходного (а) и γ -облученного (б) полиимида (поглощенная доза 123 кГр)

Учитывая вышесказанное, можем предположить, что в результате ионной имплантации процессу разбухания в первую очередь подвержены аморфные области полимера, что при облучении частично ориентированных полимеров приводит к формированию «волн» на поверхности.

Облучение γ -квантами приводит также к появлению на поверхности полимерной пленки конусообразных структур с высотой до 80 нм и диаметром в основании до 400 нм (см. рис. 1).

Аналогичная модификация поверхности полимера наблюдалась и при низких дозах имплантации ($\Phi < 1 \cdot 10^{16} \text{ см}^{-2}$) ионов металлов (рис. 2...5). Плотность таких образований достигает 2 мкм^{-2} . Шероховатость при этом изменяется незначительно. Конусообразные структуры распределены по поверхности хаотично. Плотность и размеры этих образований зависели от вида полимерного материала и условий облучения. Конусообразные структуры наиболее отчетливо проявлялись при имплантации ПЭТФ ионами Ni^+ (рис. 3, б). Их высота составляла $\sim 40 \text{ нм}$, а диаметр в основании – до 100 нм, при этом плотность их распределения по поверхности достигала 5 мкм^{-2} в отдельных местах. При имплантации ионов Ag^+ (рис. 4) их размеры и плотность были существенно ниже.

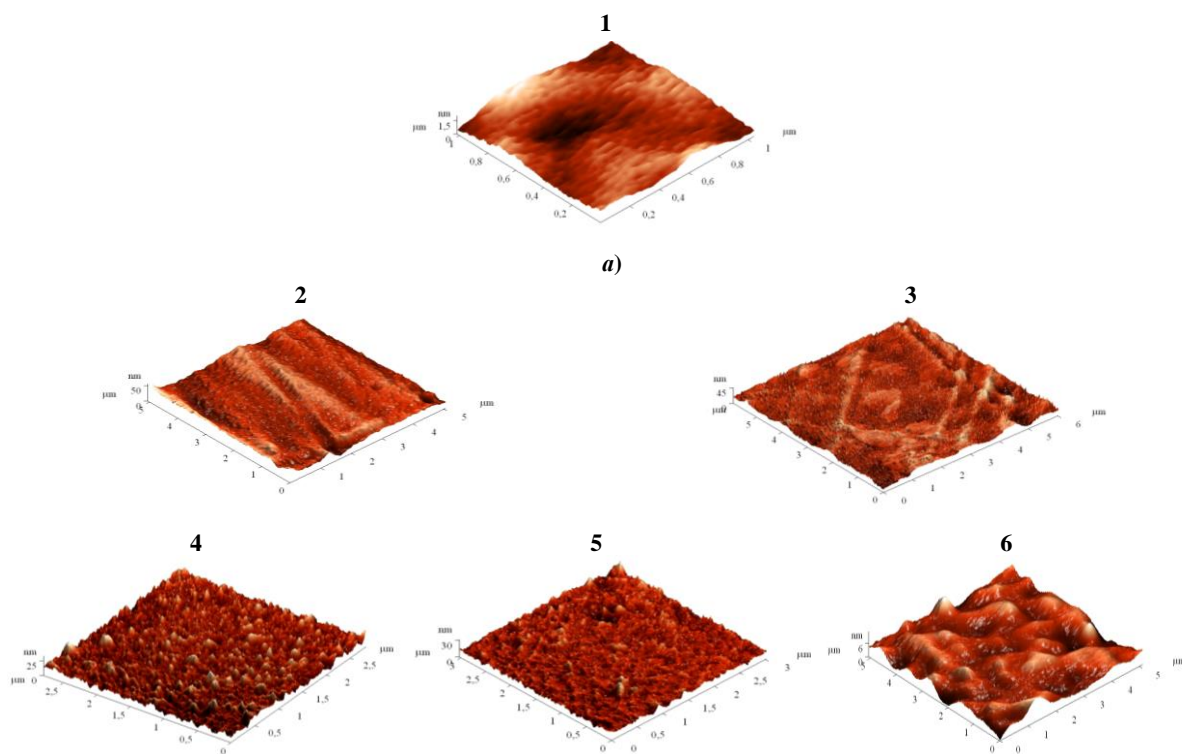


Рис. 2. Трехмерное изображение АСМ-поверхностей исходной (1) и имплантированных ионами никеля (а, 2...6) и серебра (б, 7...11) пленок полиимида. Флюенс, см^{-2} : (2, 7) – $0,25 \cdot 10^{17}$; (3, 8) – $0,5 \cdot 10^{17}$; (4, 9) – $0,75 \cdot 10^{17}$; (5, 10) – $1 \cdot 10^{17}$; (7, 11) – $1,5 \cdot 10^{17}$

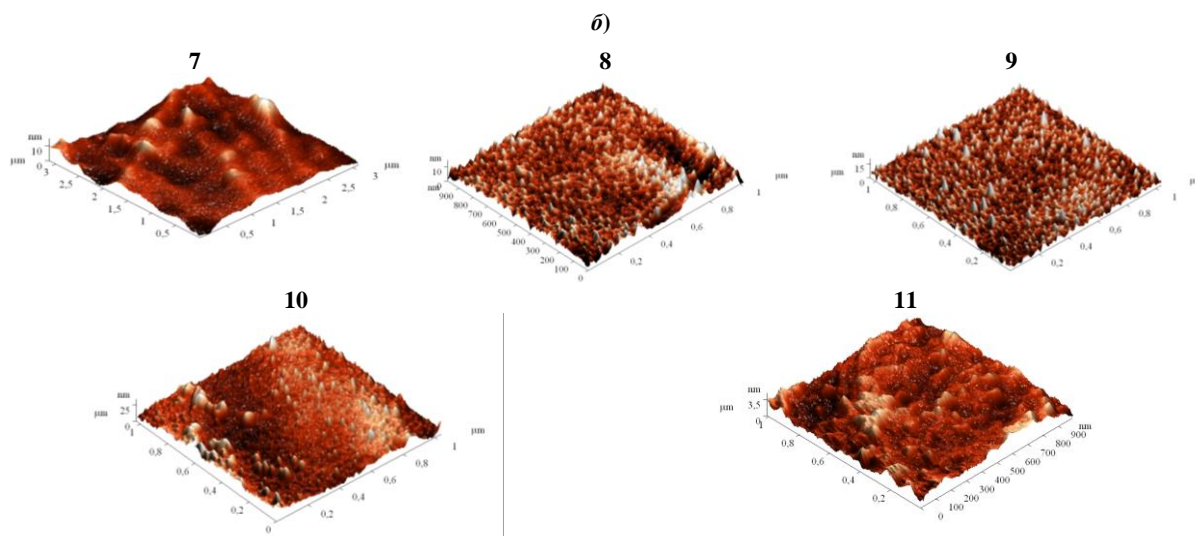


Рис. 2 – Окончание (начало см. с. 84). Трёхмерное изображение АСМ-поверхностей исходной (1) и имплантированных ионами никеля (а, 2...6) и серебра (б, 7...11) плёнок полиимида. Флюенс, см⁻²: (2, 7) – $0,25 \cdot 10^{17}$; (3, 8) – $0,5 \cdot 10^{17}$; (4, 9) – $0,75 \cdot 10^{17}$; (5, 10) – $1 \cdot 10^{17}$; (7, 11) – $1,5 \cdot 10^{17}$

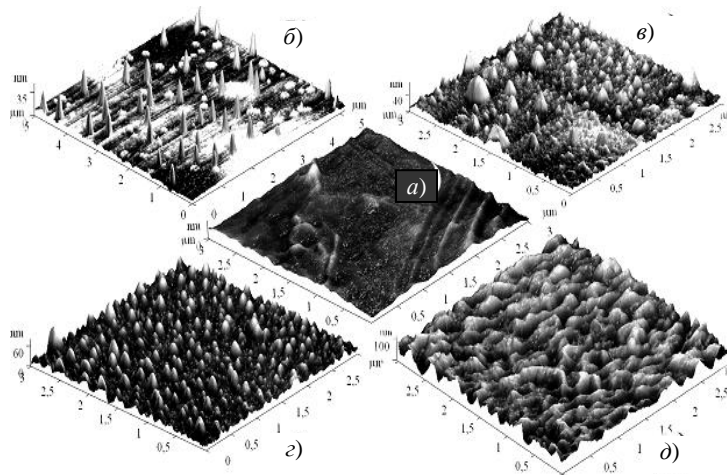


Рис. 3. АСМ-трёхмерное изображение поверхности плёнки ПЭТФ, имплантированной ионами Ni⁺. Флюенс, см⁻²: а – 0; б – $2,5 \cdot 10^{16}$; в – $1 \cdot 10^{17}$; г – $1,25 \cdot 10^{17}$; д – $1,5 \cdot 10^{17}$

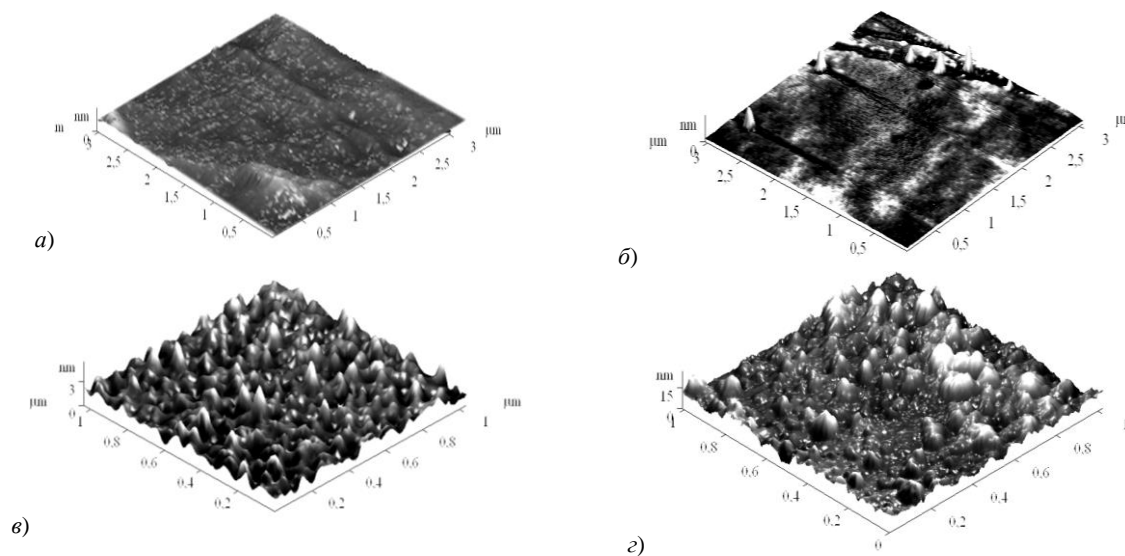


Рис. 4. АСМ-изображение поверхности плёнки ПЭТФ, имплантированной ионами Ag⁺. Флюенс, см⁻²: а – исходный образец; б – $2,5 \cdot 10^{16}$; в – $1 \cdot 10^{17}$; г – $1,5 \cdot 10^{17}$

При имплантации ПИ (см. рис. 2) и ПЭЭК (рис. 5) плотность таких конусообразных структур была существенно ниже, а диаметр в основании достигал 200...400 нм, что больше, чем в ПЭТФ. В полиимиде они наблюдались лишь при $\Phi = 5 \cdot 10^{16} \text{ см}^{-2}$ и распределены неравномерно: большая часть поверхности свободна от таких структур, имеются области их локализации с плотностью до 1...2 шт. на $\mu\text{м}^2$. Более отчетливо они наблюдались при γ -облучении.

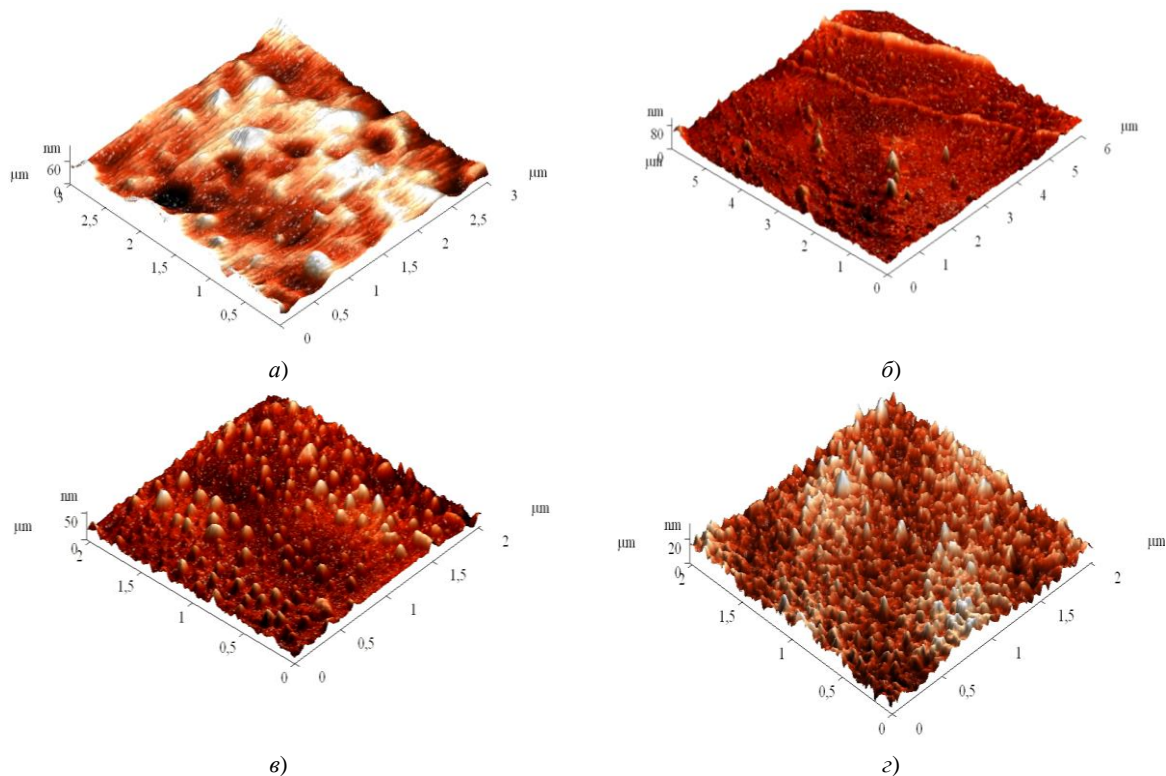


Рис. 5. АСМ-изображение поверхности полиэфирэфиркетона, исходного (а) и имплантированного ионами Ni^+ (б...г).
Флюенс, см^{-2} : б – $1 \cdot 10^{16}$; в – $5 \cdot 10^{16}$; г – $1 \cdot 10^{17}$

При увеличении дозы имплантированных ионов высота подобных конусообразных структур уменьшается, а диаметр их основания увеличивается до 800 нм (см. рис. 2, в, г и 3, в). При этом увеличивается как количество таких образований на единице площади, так и средний диаметр в основании, а также начинается их перекрытие. Шероховатость поверхности растет как для полиимида (рис. 6), так и для полиэтилентерефталата (рис. 7).

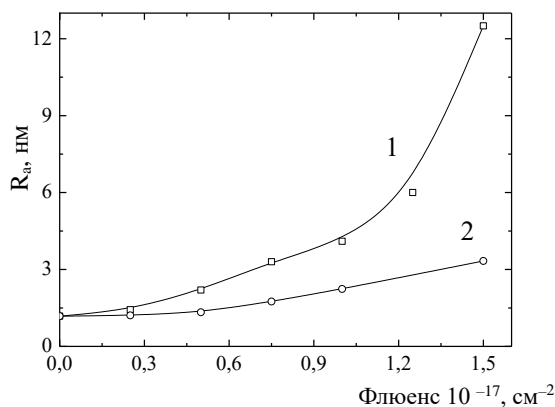


Рис. 6. Зависимость шероховатости поверхности пленок полиимида, имплантированных ионами никеля (1) и серебра (2), от дозы имплантации при плотности ионного тока $j = 4 \text{ мкА/см}^2$

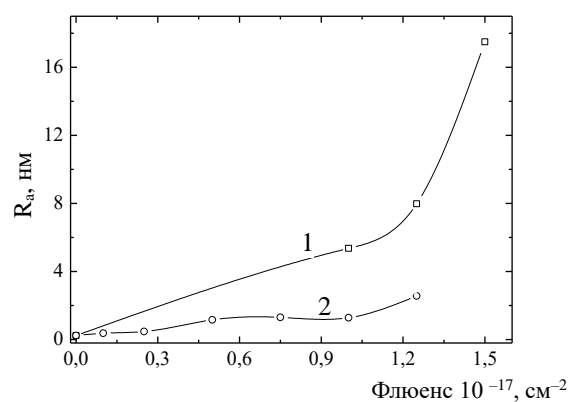


Рис. 7. Зависимость шероховатости поверхности пленок ПЭТФ, имплантированных ионами никеля (1) и серебра (2), от дозы имплантации при плотности ионного тока $j = 4 \text{ мкА/см}^2$

При дозе имплантации $1,5 \cdot 10^{17} \text{ см}^{-2}$ вышеописанные неоднородности поверхности в виде отдельных конусообразных островков полностью сливаются. Морфология поверхности качественно меняется, превращаясь в сплошную лабиринтоподобную структуру (см. рис. 1, д). При этом наблюдается резкий рост шероховатости до $R_a = 17 \text{ нм}$ при имплантации ионов никеля (рис. 6 и 7 кривые 1).

Отметим, что для тяжелых ионов серебра доза начала формирования лабиринтоподобной структуры и роста шероховатости поверхности несколько выше, чем для более легких ионов Ni^+ . Такая закономерность наблюдалась как для полиимида (см. рис. 6), так и для полиэтилентерефталата (см. рис. 7).

Изменение морфологии поверхности имплантированного полимера может быть обусловлено действием различных факторов:

- 1) образованием углеродно-металлических нанокластеров в модифицированной имплантацией приповерхностной области полимера;
- 2) процессами деструкции и распыления приповерхностного слоя полимера;
- 3) локальным разогревом матрицы в области ионного трека;
- 4) деформацией полимерной основы под действием давления, оказываемого выходом газообразных продуктов деструкции полимера.

Наблюдающаяся при больших дозах имплантации лабиринтоподобная структура поверхности полимера обусловлена, на наш взгляд, действием преимущественно первых двух факторов: 1) деструкцией и распылением приповерхностного слоя полимера и 2) образованием углеродно-металлических нанокластеров.

Распыление приповерхностного слоя в процессе ионной имплантации может приводить к выходу на поверхность полимера углеродно-металлических нанокластеров, сформировавшихся в имплантированном слое на начальной стадии имплантации. Для проверки данного утверждения нами были проведены исследования модификации поверхности пленки полиимида сфокусированными (диаметр $\sim 1 \text{ мкм}$) пучками Fe и Au с энергией 100 кэВ в интервале доз $1,0 \cdot 10^{16} \dots 1,0 \cdot 10^{18} \text{ см}^{-2}$ при плотности ионного тока 4 мкА/ см^2 .

Полученное методом атомно-силовой микроскопии трехмерное изображение поверхности полиимида, имплантированного пучком ионов Fe^+ дозой $1,0 \cdot 10^{17} \text{ см}^{-2}$, представлено на рисунке 8.

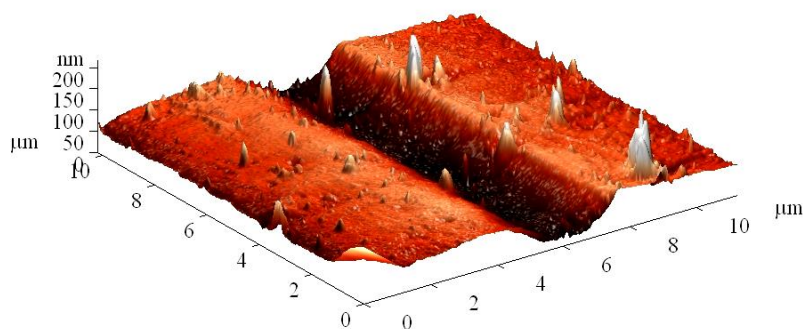


Рис. 8. Трехмерное АСМ изображение поверхности пленки полиимида, имплантированной сфокусированным пучком ионов Fe^+ дозой $1,0 \cdot 10^{17} \text{ см}^{-2}$

Аналогичные изображения были получены и для ионов Au^+ . Видно, что сфокусированный ионный пучок формирует на поверхности полиимида «канавку», ширина которой при увеличении дозы ионов слабо возрастала в пределах $2 \dots 4 \text{ мкм}$ [5], что несколько больше диаметра ионного пучка. Глубина этих «канавок» слабо возрастала с ростом флюенса ионов, для ионов Fe^+ составила $\sim 80 \text{ нм}$ при $\Phi = 5 \cdot 10^{16} \text{ см}^{-2}$ и около 110 нм при $\Phi = 1 \cdot 10^{18} \text{ см}^{-2}$. Для ионов Au^+ она была несколько ниже и в этом же диапазоне доз ионов возрастала с 55 до 70 нм [5]. Отметим, что рассчитанные с помощью программы SRIM [6] пробеги (R_p) ионов Fe^+ и Au^+ в полиимиде составили 110 и 60 нм со среднеквадратичным отклонением $\Delta R_p = 30$ и 9 нм соответственно.

С учетом вышесказанного можно утверждать, что при дозах свыше $5 \cdot 10^{16} \text{ см}^{-2}$ наблюдается выход имплантированного слоя на поверхность полиимида. При больших дозах имплантации происходит распыление имплантированного слоя. При этом скорость распыления металлических нанокластеров ниже, чем скорость деструкции и распыления окружающих их остатков полимерной матрицы. Поэтому при этих дозах на поверхности полимера формируется лабиринтоподобная структура, коррелирующая со структурой металлических нанокластеров. Аналогичные выводы справедливы, на наш взгляд, также для полиэтилентерефталата и полиэфирэфиркетона.

Конусообразные структуры имеют другую природу и воздействием четырех перечисленных выше факторов объяснены быть не могут. Отметим, что «конусы» наблюдаются не только при низких дозах имплантации, но и при других видах воздействия: при γ -облучении; при воздействии лазерными наносекундными импульсами с энергией ниже пороговой величины энергии испарения [7]. Кроме того, они могут формироваться за пределами области высокоэнергетического воздействия. Так, в работе [8] при облучении пленок ПЭТФ толщиной 3,8 мкм вакуумно-ультрафиолетовым излучением в спектральном диапазоне 115...135 нм наблюдалось формирование подобных «конусов» на обратной (не облучаемой) стороне пленки, которое не было каким-то образом объяснено.

Формирование подобных «конусов» за пределами области воздействия фокусированного ионного пучка представлено на рисунке 8. Причем наблюдались они на достаточно большом удалении (до 5 мкм и более) от «канавок», сформированных фокусированным ионным пучком Fe^+ и Au^+ с энергией 100 кэВ в интервале доз $1 \cdot 10^{17} \dots 1 \cdot 10^{18} \text{ см}^{-2}$. То есть наблюдался определенный эффект «дальнодействия» – радиационно-стимулированная модификация полимера происходила на расстояниях в несколько микрометров от области высокоэнергетического воздействия.

Объяснить формирование «конусов» можно, принимая во внимание особенности молекулярной структуры и специфику радиационного воздействия на полимеры. При высокоэнергетичном воздействии происходит процесс ионизации, заключающийся в удалении электрона с определенной молекулярной орбитали и формировании так называемой «дырки». В макромолекулах следует учитывать возможность образования делокализованных ионизированных состояний с эффективным размером, существенно превышающим размер элементарного звена, и возможность быстрой (недиффузионной) миграции «дырки» по цепи макромолекулы на значительные в молекулярном масштабе расстояния.

Такое же рассмотрение применимо и к возбужденным состояниям. Первичные физические процессы (ионизация или возбуждение) и следующие за ними химические изменения (разрыв связи, деструкция) могут быть разделены существенным расстоянием вследствие эффективной миграции «дырок» и переноса возбуждения [9].

Формирование «конусов» вне пределов зоны высокоэнергетического воздействия обусловлено переносом возбуждения и/или миграции «дырки». Так, по данным авторов [10], энергия возбуждения может передаваться в полимере на расстояние более 1000 мономерных звеньев.

С другой стороны, следует учитывать микрофазную неоднородность (микрорегетерогенность) макроструктуры полимера. При этом возможен перенос электрона или «дырки» через границу раздела фаз, что может привести к локализации радиационных повреждений в определенных микрообластях системы или вблизи раздела фаз [9].

Конусообразные структуры на поверхности полимера и являются, на наш взгляд, проявлением такой локализации радиационно-индуцированной модификации полимеров, приводящей к локальному хаотичному вспучиванию поверхности полимера. В пользу такого предположения говорит также формирование сплошной линии «конусов» вдоль структурного дефекта – царапины на поверхности γ -облученного полиимида (рис. 9).

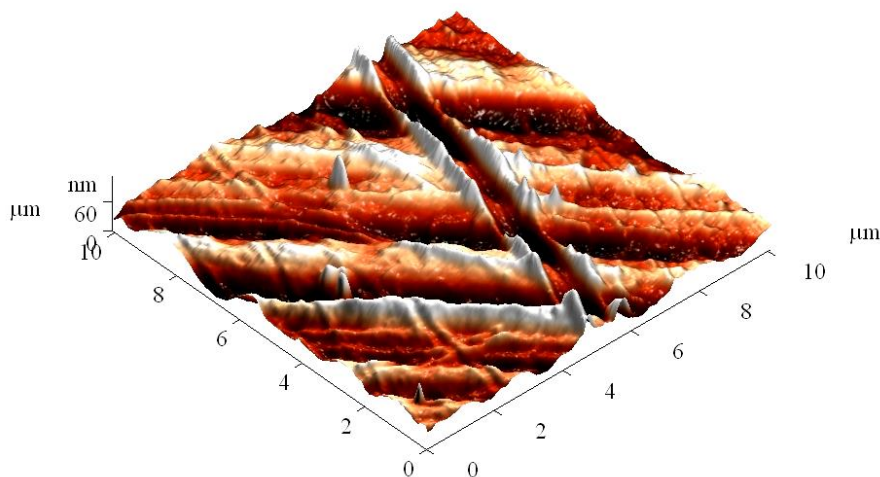


Рис. 9. Трехмерное АСМ изображение поверхности (с царапиной) пленки γ -облученного полиимида. Поглощенная доза 100 кГр

Возможен также другой «термодинамический» подход к объяснению возникновения «конусов» на поверхности облученных полимеров, который дополняет изложенное выше и позволяет оценить размер «конуса». Этот подход предполагает наличие локальных упругих напряжений сжатия в полимере, например, у границы раздела микрофаз. Плотность упругой энергии в сжатом материале [11]

$$f = p^2/2E,$$

где p – напряжение; E – модуль упругости.

При релаксации напряжений сжатия может приводить к формированию «конуса». Будем считать упругие напряжения сферическими с радиусом r , при образовании «конуса» освобождается упругая энергия в объеме $\sim r^3$, равная

$$-\Delta F_{\text{упр}} = f \cdot r^3 = (p^2/2E) r^3,$$

и требуется затрата работы на создание новой поверхности площадью $\sim l^2$:

$$\Delta F_{\text{нов}} = l^2 \sigma.$$

Здесь σ – поверхностная энергия; l – радиус конуса в основании.

Изменение энергии системы

$$\Delta F = \Delta F_{\text{упр}} + \Delta F_{\text{нов}} = l^2 \sigma - (p^2/2E) r^3.$$

Предполагая полную релаксацию $\Delta F = 0$, можно оценить размеры конусов в основании:

$$l = p (r^3/2E\sigma)^{1/2}.$$

Таким образом, размеры формирующихся «конусов» прямо пропорциональны величине напряжений сжатия и размерам напряженной области в степени $3/2$.

Отметим, что высокоэнергетичное облучение выступает в данном случае в качестве лишь «спускового крючка» для высвобождения энергии, запасенной в деформированных областях полимера. Поэтому такие процессы могут протекать далеко за пределами области высокоэнергетического воздействия.

Такой подход с высокой степенью достоверности способен описать формирование «конусов» за пределами области высокоэнергетического воздействия.

В области имплантации картина усложняется. Вследствие высокой локальной неоднородности процессов радиационно-индуцированной модификации полимеров возможно формирование полей упругих напряжений в процессе облучения. То есть одновременно протекают два процесса:

- 1) релаксация существующих упругих напряжений;
- 2) формирование новых деформационных полей.

Заключение. В результате проведенного исследования получены следующие результаты:

- обнаружено формирование при γ -облучении и низких дозах имплантации ионов конусообразных структур высотой до 80 нм и диаметром в основании до 400 нм, хаотично расположенных на поверхности полимерных пленок полиимида, полиэтилентерефталата и полиэфирэфиркетона;
- образование конусообразных структур обусловлено преимущественной локализацией радиационно-химических процессов вблизи дефектов структуры (границы раздела фаз) полимера;
- формирование указанных дефектов за пределами зоны имплантации обусловлено переносом возбуждения и/или миграцией дырок из области трека иона.

ЛИТЕРАТУРА

1. Степанов, А.Л. Оптические свойства металлических наночастиц, синтезированных в полимере методом ионной имплантации (Обзор) / А.Л. Степанов // Журнал технической физики. – 2004. – Т. 74, № 2. – С. 1 – 12.
2. Попок, V.N. Compositional and structural alterations of polymers under low-to-medium-energy ion implantation / V.N. Popok // Surface Science Research; ed. C.P. Norris. – Amsterdam: Nova Science Publishers Inc, 2005. – P. 147 – 193.
3. Грасси, Н. Деструкция и стабилизация полимеров / Н. Грасси, Дж. Скотт. – М.: Мир, 1988. – 246 с.

4. Песчанская, Н.Н. Скачкообразная деформация γ -облученного полиметилметакрилата / Н.Н. Песчанская, А.С. Смолянский, А.С. Шведов // Физика твердого тела. – 2009. – Т. 51, № 6. – С. 1218 – 1222.
5. Формирование низкоразмерных структур на полимерной пленке фокусированным ионным пучком / А.А. Харченко [и др.] // Вестн. БГУ. Сер. 1. – 2012. – № 2. – С. 29 – 31.
6. Ziegler, J.F. The stopping and range of ions in solids / J.F. Ziegler, J.P. Biersack, U. Littmark. – New York, 1985 (SRIM-2010 software at <http://www.srim.org>).
7. Малышев, А.Ю. Модель лазерного свеллинга полимеров при воздействии наносекундных импульсов / А.Ю. Малышев, М.Ю. Битюрин // Квантовая электроника. – 2009. – Т. 35, № 9. – С. 825 – 830.
8. Исследование поверхности пленок из полиэтилентерефталата, модифицированных вакуумно-ультрафиолетовым облучением на воздухе / А.В. Митрофанов [и др.] // Поверхность. – 2009. – № 7. – С. 30 – 38.
9. Экспериментальные методы химии высоких энергий / под общ. ред. М.Я. Мельникова. – М.: МГУ, 2009. – С. 169 – 178.
10. Светочувствительные полимерные материалы / под ред. А.В. Ельцова. – Л.: Химия, 1985. – 296 с.
11. Щукин, Е.Д. Влияние активной среды на механическую устойчивость и повреждаемость поверхности твердого тела / Е.Д. Щукин // Вестн. МГУ. Сер. 2. Химия. – 2012. – Т. 53, № 1. – С. 50 – 72.

Поступила 19.09.2013

INVESTIGATION OF POLIMERS SURFACE MODIFIED BY IRRADIATION

**A. KHARCHENKO, D. BRINKEVICH, S. BRINKEVICH,
S. VABISHCHEVICH, N. VABISHCHEVICH,
M. LUKASHEVICH, V. PROSOLOVICH, V. ODZHAEV**

The modification of polyimide, polyethyleneterephthalate and polyetheretherketone films surface at the γ -irradiation and Ni^+ , Ag^+ , Au^+ , Fe^+ and B^+ 30...100 keV ions implantation in the fluence range $1 \cdot 10^{15}$... $1,5 \cdot 10^{17}$ cm^{-2} have been investigated by the atomic force microscopy. The formation of single cone off with height up to 80 nm and diameters up to 400 nm on the polymer films surface was observed at low implantation doze. Their formation is due to the predominant localization of radiation-chemical processes near the defect structure (the interface) polymer. The formation of these defects outside the implantation range is caused by the transfer of excitation and/or migration of holes from the field ion track.

УДК 539.216.2

ОПТИЧЕСКОЕ ПОГЛОЩЕНИЕ В ПЛЕНКАХ $Cu_xIn_xZn_{2-2x}Se_2$ **В.В. ХОРОШКО, канд. техн. наук И.Н. ЦЫРЕЛЬЧУК***(Белорусский государственный университет информатики и радиоэлектроники, Минск);***д-р физ.-мат. наук В.Ф. ГРЕМЕНОК***(ГНПО «Научно-практический центр НАН Беларуси по материаловедению», Минск)*

Представлены результаты исследований изменения коэффициента поглощения тонких пленок $Cu_xIn_xZn_{2-2x}Se_2$, полученных путем селенизации базовых слоев $ZnSe/(Cu+In)$, в зависимости от концентрации атомов цинка. На основании представленных результатов получены значения ширины запрещенной зоны в тонких пленках $Cu_xIn_xZn_{2-2x}Se_2$. Полученные данные свидетельствуют о том, что тонкие пленки CIZS являются перспективными материалами для тонкопленочных солнечных элементов, при этом стоимость пленок будет заметно ниже, так как в них происходит замена дорогостоящих материалов галлий и индий на менее дорогостоящий цинк.

Введение. Тонкопленочные солнечные элементы (ТСЭ) в последнее время прочно заняли долю рынка в производстве солнечных элементов (СЭ), а проводимые исследования позволяют говорить о том, что в будущем данный класс СЭ сможет лидировать при производстве СЭ. Так, если на 2001 год суммарный процент производимых всех видов ТСЭ был чуть более 1 % от общей массы [1], то к 2011 году такой процент уже составляет около 12 %, а общая мощность установленных батарей – 100 ГВт, что в 100 раз больше, чем в 2001 году [2].

Одним из перспективных направлений развития ТСЭ являются многокомпонентные полупроводниковые соединения со структурой халькопирита (в особенности $CuInSe_2$ (CIS)). Производимые промышленно ТСЭ с использованием $Cu(In,Ga)Se_2$ (CIGS) в качестве поглощающего слоя позволяют получать СЭ со значением КПД 12...14 % при стоимости около 0,8 US\$/Вт пиковой мощности [3]. Слои CIGS содержат такие дорогостоящие материалы, как галлий и индий, соответственно поиск материалов, сходных по физическим свойствам, является актуальной задачей. Замещение $CuInSe_2$ на его четверной аналог $Cu_xIn_xZn_{2-2x}Se_2$ (CIZS) позволяет получить материал с физическими характеристиками, плавно изменяемыми в зависимости от химического состава. Ширина запрещенной зоны CIZS в зависимости от состава ($0 < x < 1$) меняется в пределах от 1,04 до 2,67 эВ [4], что позволяет получить материал, согласованный по своим характеристикам с максимумом солнечной энергии. При этом также важнейшим параметром материала является коэффициент поглощения – характеристика отдельной полосы поглощения бездефектного кристалла (твердого раствора) или среды с заданным содержанием поглощающих дефектов. Получение материалов с большими значениями коэффициента поглощения $\alpha > 10^4 \text{ см}^{-1}$ позволяет выполнять в промышленных ТСЭ толщину поглощающего слоя 1...2,5 мкм, что приводит к удешевлению продукции [5].

Методика эксперимента. Для получения поликристаллических пленок $Cu_xIn_xZn_{2-2x}Se_2$ применялся метод двухстадийной селенизации, особенностью которого является то, что процесс синтеза протекает в атмосфере одного из наиболее летучих компонент, образующих выбранную систему, тогда как остальные компоненты наносятся в виде тонких пленок на твердый носитель, в качестве которого применяются стеклянные пластины. При повышении температуры компоненты пленок вступают в реакцию между собой и сосуществующей паровой фазой и образуют гомогенную пленку требуемого состава. В качестве паровой фазы использовался селен, поэтому в целом такой процесс получил название «селенизация» [6].

Для получения базовых слоев Cu-In-Zn использовался метод термического испарения. Осаждение слоев Cu-In-Zn методом термического испарения осуществлялось на подложки из боросиликатного стекла. Подложки предварительно очищались кипячением в перекисно-аммиачной среде и промывались деионизованной водой. Нанесение Cu-In-Zn осуществлялось на вакуумной установке УВН-71П-3. Давление остаточных газов в процессе осаждения металлов составляло $8,7 \cdot 10^{-4}$ Па. Температура подложек поддерживалась на уровне 100...120 °С. Навески металлов чистоты ВЗ распылялись из молибденовой лодочки одновременно со средней скоростью 0,5 мкм. Для получения пленок различного элементного состава соотношение навесок Cu/In/Zn варьировалось в пределах (0,1...0,5)/1/1. Толщина осажденных слоев составляла 600...750 нм. Введение цинка в осаждаемые металлические пленки этим способом приводит к сильному загрязнению вакуумной установки. В связи с этим получение базового слоя, содержащего цинк, осуществлялось методом испарения селенида цинка ZnSe, который испаряется возгонкой с сохранением стехиометрии. Навеска селенида цинка рассчитывалась исходя из задаваемого состава получаемой пленки. Селенид цинка испарялся из отдельного танталового испарителя. Первая стадия селенизации осуществлялась при температуре 240...270 °С в течение 10...30 минут, вторая стадия – при температуре 510...540 °С в течение 10...50 минут.

Спектры пропускания экспериментальных пленочных образцов (одиночных пленочных покрытий на аморфных стеклянных подложках) $T(\lambda) = I/I_0$ (I_0 – интенсивность потока света, входящего в пленоч-

ное покрытие; I – интенсивность потока света, вышедшего из образца) исследовались в диапазоне длин волн 0,19...3,00 мкм на спектрофотометре Cary-500 Scan (*UV-Vis-NiRSpectrometer, Varian, США*) и использовались как первичные экспериментальные данные для оценки спектров линейного поглощения (α) и значений края поглощения (E_g) пленкообразующего материала. В случае широкозонных прямозонных твердых растворов стехиометрического состава поглощение связано с прямыми зона-зонными переходами, а в случае сред, содержащих собственные (*центры окраски*) и примесные дефекты (*ионы внедрения или замещения*) – еще и с переходами между уровнями поглощающего дефекта в твердом растворе. Поглощаемость среды A в общем случае определяется выражением $A = 1 - T - R - S$, где R и S – коэффициенты отражения и рассеяния, а T – коэффициент пропускания. В случае пленочных образцов определение точных значений коэффициентов поглощения α или оптической плотности D осложняется отличием микроструктуры пленочных покрытий от ламинарной (сплошной) (что может приводить к существенному вкладу R и S) и наличием подложки. Практическое определение точных значений коэффициентов поглощения α исследованных пленочных покрытий осложнялось существенным вкладом материала подложки в спектральные характеристики системы «пленка – подложка» в случае малопоглощающих тонких покрытий, формируемых на стеклянных подложках ($E_g \sim 330...350$ нм), и малыми значениями пропускания ($T(\lambda) = I/I_0$) в случае сильно поглощающих покрытий с металлическим блеском ($T_{\text{макс}}$ в области края $\sim 1...10$ %).

В связи с вышеизложенными особенностями исследованных образцов, приводящими к неконтрастной картине краевого поглощения и отсутствию возможности оценки коэффициента преломления n и геометрической толщины покрытий h (исходя из спектрального положения интерференционных пиков в спектрах пропускания), коэффициент поглощения α определялся с учетом показательной (экспоненциальной) зависимости поглощения от толщины (*закон Бугера – Ламберта*) и независимых измерений толщины методами сканирующей микроскопии скола покрытий. Оцененные спектры поглощения использовались для установления края поглощения (λ_g) исследуемых образцов. Коэффициент поглощения α в случае преимущественного вклада прямых зона-зонных переходов связан с энергией краевого поглощения E_g соотношением: $\alpha = (h\nu - E_g)^{1/2}/h\nu$.

Зависимость $(\alpha h\nu)^2$ от $h\nu$ позволяет оценить значения энергии фундаментального края поглощения путем линейной экстраполяции $h\nu - E_g$ в область $(\alpha h\nu)^2 = 0$.

Спектры возбуждения и фотолюминесценции измерялись при комнатной температуре (300 К) в диапазоне 200...800 нм с использованием в качестве источника возбуждающего излучения дуговой ксеноновой лампы. Измерения спектров люминесценции осуществлялись на спектрофлуориметре SFL-1211 А (*SOLAR, Беларусь*) в стандартной 90° геометрии возбуждение – регистрация.

Результаты и их обсуждение. На рисунке 1 представлены спектральные зависимости оптической плотности D и величины $(\alpha h\nu)^2$ для пленок CIZS с высоким ($\approx 19,6$ ат. %) содержанием Zn.

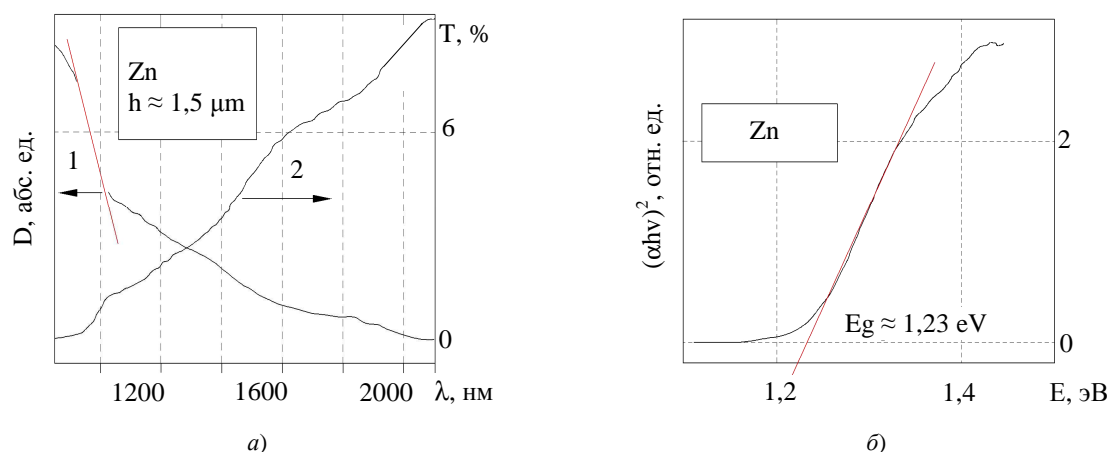


Рис. 1. Спектральные зависимости оптической плотности D (а) и величины $(\alpha h\nu)^2$ (рисунк б) для пленок CIZS с содержанием Zn $\approx 19,6$ ат. %

На рисунке 2 представлены такие же зависимости для CIZS с низким ($\approx 5,5$ ат. %) содержанием цинка. Как видно, коэффициент оптического пропускания T (кривая 2) для этих образцов является небольшим. Низкие значения коэффициента оптического пропускания могут быть обусловлены использованием диафрагмы для детектирования сигнала в связи с малым размером образцов и наличием в них проколов. Вследствие этого амплитуда пропускания существенно уменьшалась, а точное определение ширины запрещенной зоны по краю пропускания оказалось затруднительным.

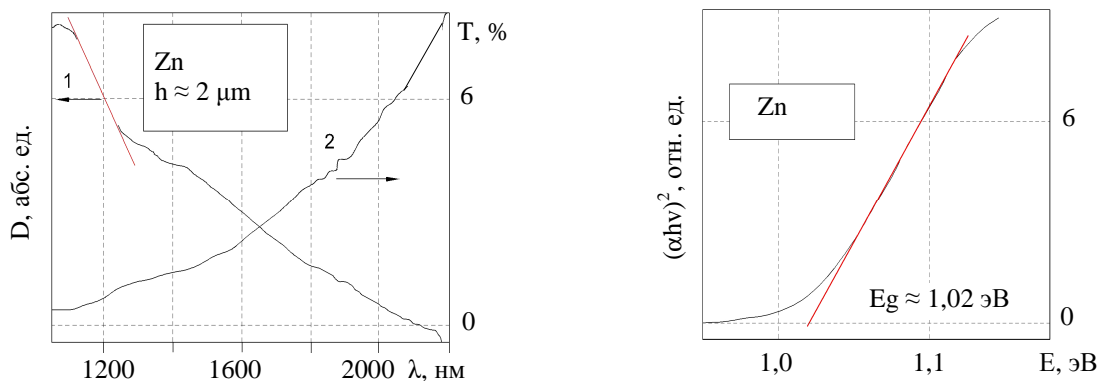


Рис. 2. Спектральные зависимости оптической плотности D (рисунок a) и величины $(\alpha h\nu)^2$ (рисунок b) для пленок CIZS с содержанием Zn $\approx 5,5$ ат. %

В таблице приведены составы исследуемых образцов.

Составы исследуемых образцов

№ образца	Элементы, ат. %			
	Cu	In	Zn	Se
1	25,15	22,94	–	51,01
2	21,11	20,74	5,4	52,74
3	20,24	17,36	9,98	52,52
4	18,11	16,23	18,11	51,4
5	15,44	14,12	19,59	50,84

Оценка ширины запрещенной зоны для твердого раствора CIZS с содержанием цинка 5,5 ат. % дает значение $E_g \sim 1,02$ eV (1215 nm), что соответствует ширине запрещенной зоны исходного соединения CuInSe₂. Данный факт, по-видимому, свидетельствует о том, что при небольших концентрациях цинка он играет роль легирующей примеси с небольшой энергией ионизации. Для CIZS с содержанием цинка 19,5 ат. % оценка ширины запрещенной зоны дает значение $E_g \sim 1,23$ eV (1007 nm). Исследование поглощения пленок CIZS различного состава, полученных при одинаковых режимах путем селенизации базовых слоев ZnSe / (Cu + In) с содержанием ZnSe от 0,05 до 0,25 мол. %, позволило установить зависимость ширины запрещенной зоны E_g от состава.

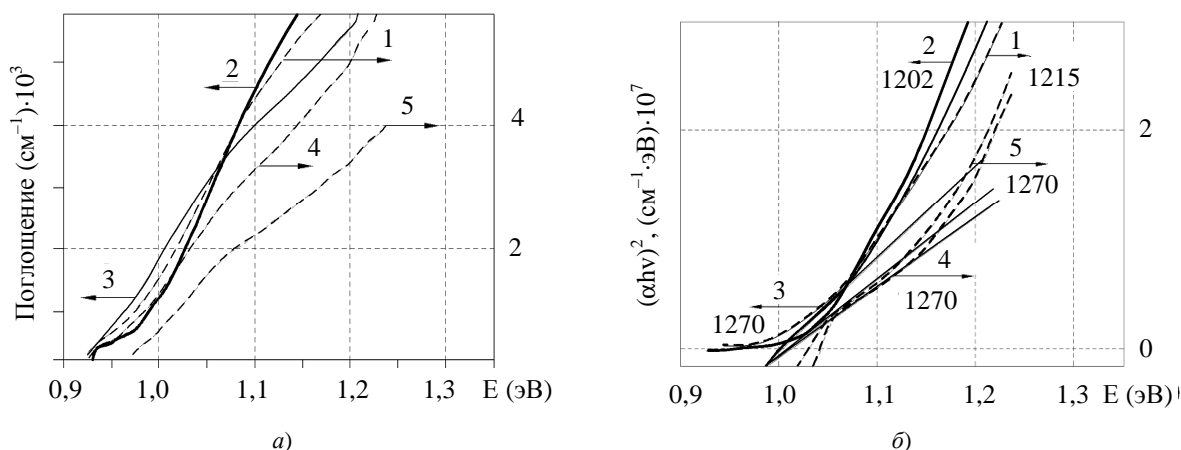


Рис. 3. Зависимости поглощения от концентрации цинка (a), зависимости $(\alpha h\nu)^2$ от энергии фотона E (b)

Как видно из рисунка 3, a , с увеличением номера образца (соответственно с увеличением содержания Zn) уменьшается крутизна края поглощения и абсолютная величина поглощения в коротковолновой области спектра (т.е. в области, где поглощение обусловлено свободными носителями). Величина поглощения в спектральной области 1000...1032 нм уменьшается от $\sim 10^4$ см⁻¹ (образцы 1, 2) до $\sim 4 \cdot 10^3$ см⁻¹ (образец 5). На рисунке 3, b приведены рассчитанные зависимости $(\alpha h\nu)^2$ от E , позволяющие оценить положение края E_g (оптическую ширину запрещенной зоны) в предположении, что краевое поглощение обусловлено преимущественно прямыми межзонными переходами (в этом случае $\alpha \sim (h\nu - E_g)^{1/2}/h\nu$).

Уменьшение абсолютной величины поглощения коррелирует с увеличением пропускания T (%) в длинноволновой области спектра (~ 1300 нм) от 0,8...0,4 % (образцы 1 и 2) до $\sim 1,5...3,0$ % (образцы 3...5). При этом величина прикраевого диффузного отражения R в образцах 1,2,4,5 составляет 11...12 %. С учетом того, что поглощение A описывается выражением $A = 1 - R - T$ (где T и R – оптическое пропускание и отражение соответственно), уменьшение поглощения обусловлено увеличением пропускания, отмеченного на рисунке 3.

Очевидно, что повышение концентрации цинка в пленках CIZS сопровождается увеличением прикраевого поглощения (уменьшения пропускания T), повышением крутизны и небольшим длинноволновым сдвигом края поглощения при сохранении величины диффузного отражения.

Таким образом, проведенные исследования оптического поглощения пленок CIZS позволяют заключить, что эти материалы характеризуются прямыми межзонными переходами и имеют коэффициент оптического поглощения $\alpha > 10^4$ см⁻¹ в области края фундаментального поглощения. Оптическая ширина запрещенной зоны пленок CIZS с увеличением содержания цинка смещается в длинноволновую область и для ряда исследованных образцов меняется в пределах 1,02...1,23 эВ, что соответствует значениям данного параметра для солнечных элементов на основе CIGS, производимых промышленно [7].

Полученные результаты позволяют говорить о том, что тонкие пленки CIZS являются перспективными материалами для ТСЭ, при этом стоимость пленок будет заметно ниже, так как в них происходит замена дорогостоящих материалов галлий и индий на менее дорогостоящий цинк. При таком большом коэффициенте поглощения более 99 % излучения будет поглощаться в пленке толщиной порядка 1 мкм, а более 96 % излучения – в пленке толщиной 0,5...0,6 мкм.

Заключение. В результате проведенных исследований было установлено, что введение в исходную пленку CuInSe₂ атомов цинка приводит к образованию собственных дефектов замещения типа Zn_{Cu} и Zn_{In}, создающих в запрещенной зоне соответственно донорные и акцепторные энергетические уровни. Полученные материалы характеризуются высоким значением коэффициента оптического $\alpha > 10^4$ см⁻¹ поглощения, что вполне достаточно для использования в качестве поглощающего слоя при толщине всего 1 мкм. Добавление к исходному соединению CIS атомов цинка приводит к увеличению ширины запрещенной зоны и коэффициента оптического поглощения, что в целом приводит к повышению эффективности создаваемых фотоэлектрических преобразователей.

ЛИТЕРАТУРА

1. Photovoltaic materials, history, status and outlook / A. Goetzberger [et al.] // Material Science and Engineering. – 2003. – Vol. 40. – P. 1 – 46.
2. Free Solar Panel Price Survey//CivicSolarCorpotation [Electronic resource]. – 2013. – Mode of access: http://www.civicsolar.com/products/solar-panels/?solrsort=fs_price_watt_break_1_0%20asc. – Date of access: 02.09.2013.
3. Фотоэлектрические свойства поверхностно-барьерных структур на основе пленок Zn_{2-2x}Cu_xIn_xSe₂ / В.Ю. Рудь [и др.] // Физика и техника полупроводников. – 2005. – Т. 39. Вып 9. – С. 1070 – 1074.
4. Formation and Investigation of Photosensitive Structures Based on Laser-Deposited CuInSe₂-2ZnSe Films / Rud Yu.V. [et al.] // Phys. Stat. Solidi (a). – 2001. – V. 118, № 3. – P. 1077 – 1085.
5. Goodrich, A. CIGS Road Map / A. Goodrich, M. Woodhouse, R. Noufi // NREL technical report. – 2011. – P. 60.
6. Rau, U. Cu(In,Ga)Se₂ Solar Cells / U. Rau, H.W. Schock // Series of Photoconversion of Solar Energy. – 2001. – Vol. 1. – P. 277 – 345.
7. 19.9 %-efficient ZnO/CdS/Cu(In,Ga)Se₂ solar cell with 81.2 % fill factor / I. Repins [et al.] // Progress in Photovoltaics. – 2008. – V. 16. – P. 235 – 239.

Поступила 09.09.2013

OPTICAL ABSORPTION IN Cu_xIn_xZn_{2-2x}Se₂ FILMS

V. KHOROSHKO, I. TSYRELCHUK, V. GREMENOK

The results of studies of change of absorption coefficient of Cu_xIn_xZn_{2-2x}Se₂ thin films are shown, obtained by selenium saturation of basic layers ZnSe/(Cu+In) depending on the concentration of zinc atoms. On the basis of the results shown values of the width of forbidden zone in Cu_xIn_xZn_{2-2x}Se₂ thin films are obtained. The data obtained show that thin films CIZS are perspective materials for thin film solar elements, and in addition the price of thin films will be much lower, because replacement of expensive materials as Ga and In for less expensive zinc takes place.

УДК 539.3

**ЧЕТЫРЕХФУНКЦИОНАЛЬНЫЙ МЕТОД РАСЧЕТА
НАПРЯЖЕННО-ДЕФОРМИРОВАННОГО СОСТОЯНИЯ
В СИСТЕМЕ ПАРАЛЛЕЛЬНЫХ ОСТАТОЧНЫХ ЛИНЗОВИДНЫХ
ДЕФОРМАЦИОННЫХ ДВОЙНИКОВ В МАРТЕНСИТНОЙ ФАЗЕ Ni_2MnGa**

*канд. физ.-мат. наук О.М. ОСТРИКОВ, Е.В. ШМАТОК
(Гомельский государственный технический университет им. П. О. Сухого)*

Разработан метод расчета напряженно-деформированного состояния, обусловленного остаточными параллельными линзовидными двойниками, находящимися вдали от поверхности. Метод пригоден для расчета смещений, деформаций и напряжений у двойников со сложной формой границ. Установлено, что максимальные напряжения локализуются у вершин двойников и на их границах.

Введение. Двойникование кристаллов представляет собой распространенное явление, связанное с поворотом кристаллической решетки [1 – 4]. Образование двойников происходит по разным причинам, например, в результате термического либо механического воздействия на кристалл [1]. Двойникование играет важную роль в процессах разрушения материалов [3], влияет на закономерности их деформирования и, как следствие, – на их физико-механические свойства [4]. Ввиду этого развитие теории двойникования является важной научной проблемой и шагом на пути разработки методов прогнозирования влияния двойникования на свойства используемых в технике материалов [4]. Магнитные материалы с памятью формы (сплавы Гейслера) в настоящее время находят все более широкое практическое применение в технических системах нового поколения, позволяя миниатюризировать изделия из этих материалов [5 – 7]. Физико-механические свойства магнитных сплавов Гейслера в настоящее время недостаточно изучены. Необходимость таких исследований обусловлена возрастающими требованиями к уровню технологии магнитных материалов с памятью формы. Применение магнитных сплавов с памятью формы в технических системах позволяет вывести их на новый, самый современный уровень. Несмотря на большой интерес в мире к данной проблематике, имеется ряд нерешенных задач, связанных с исследованием взаимодействия механического двойникования и бездиффузионных фазовых превращений [8; 9]. Слабо развита теория бездиффузионных фазовых превращений в ферромагнитных материалах с эффектом запоминания формы.

Цель работы – разработка метода расчета напряженно-деформированного состояния в часто встречающейся на практике системе параллельных линзовидных двойников неправильной формы.

Постановка задачи. Условия формирования деформационных остаточных двойников в монокристаллах Ni_2MnGa изучены в [8; 9]. На рисунке 1 показано схематическое изображение таких двойников для реализации четырехфункционального способа расчёта, использование которого целесообразно в случае более сложной формы двойниковых границ и при более существенных отклонениях двойника от линзовидной формы.

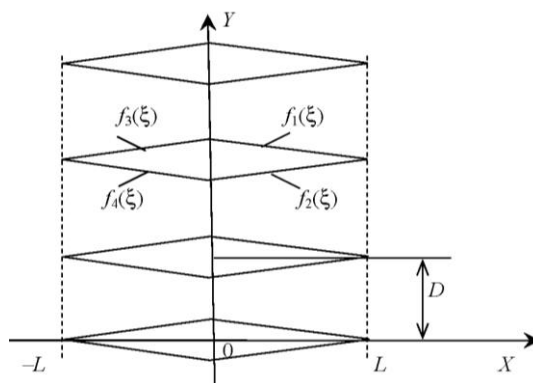


Рис. 1. Схематическое изображение системы параллельных линзовидных двойников в рамках четырехфункциональной модели

Название метода расчета связано с необходимостью использования четырех функций $f_1(\xi)$, $f_2(\xi)$, $f_3(\xi)$ и $f_4(\xi)$ для описания формы двойников. Меньшее количество функций целесообразно использовать в случае формы двойников, близкой к линзовидной форме [10].

Удобно приближение ромбовидных двойников. В этом случае форма их границ может быть описана линейными функциями [4]:

$$\begin{aligned} f_1(\xi) &= -\frac{H}{2}\left(\frac{\xi}{L}-1\right), \quad f_2(\xi) = \frac{H}{2}\left(\frac{\xi}{L}-1\right), \\ f_3(\xi) &= \frac{H}{2}\left(\frac{\xi}{L}+1\right), \quad f_4(\xi) = -\frac{H}{2}\left(\frac{\xi}{L}+1\right), \end{aligned} \quad (1)$$

где L – половина длины двойника; H – наибольшая ширина двойника.

С применением приближения о непрерывном распределении двойниующих дислокаций на двойниковых границах, использованного в [4], смещения u_i и напряжения σ_{ij} у такой системы двойников можно рассчитать по формулам:

$$u_i(x, y) = u_i^{(1)}(x, y) + u_i^{(2)}(x, y) + u_i^{(3)}(x, y) + u_i^{(4)}(x, y); \quad (2)$$

$$\sigma_{ij}(x, y) = \sigma_{ij}^{(1)}(x, y) + \sigma_{ij}^{(2)}(x, y) + \sigma_{ij}^{(3)}(x, y) + \sigma_{ij}^{(4)}(x, y), \quad (3)$$

где i, j принимают значения x, y или z ; $u_i^{(1)}(x, y)$, $u_i^{(2)}(x, y)$, $u_i^{(3)}(x, y)$, $u_i^{(4)}(x, y)$ и $\sigma_{ij}^{(1)}(x, y)$, $\sigma_{ij}^{(2)}(x, y)$, $\sigma_{ij}^{(3)}(x, y)$, $\sigma_{ij}^{(4)}(x, y)$ – смещения и напряжения, созданные участками границ двойника.

В соответствии с [4] данные смещения и напряжения можно определить по формулам:

$$u_i^{(k)}(x, y) = \sum_{n=0}^{M-1} \int_0^L \sqrt{1 + (f_k'(\xi))^2} \rho_k(\xi) u_i^{(k,0)}(x, y - nD, \xi) d\xi; \quad (4)$$

$$\sigma_{ij}^{(k)}(x, y) = \sum_{n=0}^{M-1} \int_0^L \sqrt{1 + (f_k'(\xi))^2} \rho_k(\xi) \sigma_{ij}^{(k,0)}(x, y - nD, \xi) d\xi. \quad (5)$$

Здесь k принимает значения 1, 2, 3 или 4; n – индекс суммирования; M – число двойников в системе; ξ – параметр интегрирования; $\rho_k(\xi)$ – плотность двойниующих дислокаций на k -м участке двойниковой границы; D – расстояние между двойниками (см. рис. 1).

В (4) и (5) принято, что

$$\begin{aligned} u_x^{(k,0)}(x, y - nD, \xi) &= \frac{b_{sp}}{2\pi} \left[\operatorname{arctg} \frac{y - f_k(\xi) - nD}{x - \xi} + \frac{(x - \xi)(y - f_k(\xi) - nD)}{2(1 - \nu)((x - \xi)^2 + (y - f_k(\xi) - nD)^2)} \right], \\ u_y^{(k,0)}(x, y - nD, \xi) &= -\frac{b_{sp}}{2\pi} \left[\frac{1 - 2\nu}{4(1 - \nu)} \ln((x - \xi)^2 + (y - f_k(\xi) - nD)^2) + \frac{(x - \xi)^2 - (y - f_k(\xi) - nD)^2}{4(1 - \nu)((x - \xi)^2 + (y - f_k(\xi) - nD)^2)} \right], \end{aligned} \quad (6)$$

$$u_z^{(k,0)}(x, y - nD, \xi) = \frac{b_e}{2\pi} \operatorname{arctg} \frac{y - f_k(\xi) - nD}{x - \xi};$$

$$\begin{aligned} \sigma_{xx}^{(k,0)}(x, y - nD, \xi) &= -\frac{\mu b_{sp}}{2\pi(1 - \nu)} \frac{(y - f_k(\xi) - nD)[3(x - \xi)^2 + (y - f_k(\xi) - nD)^2]}{[(x - \xi)^2 + (y - f_k(\xi) - nD)^2]^2}, \\ \sigma_{yy}^{(k,0)}(x, y - nD, \xi) &= \frac{\mu b_{sp}}{2\pi(1 - \nu)} \frac{(y - f_k(\xi) - nD)[(x - \xi)^2 - (y - f_k(\xi) - nD)^2]}{[(x - \xi)^2 + (y - f_k(\xi) - nD)^2]^2}, \\ \sigma_{xy}^{(k,0)}(x, y - nD, \xi) &= \frac{\mu b_{sp}}{2\pi(1 - \nu)} \frac{(x - \xi)[(x - \xi)^2 - (y - f_k(\xi) - nD)^2]}{[(x - \xi)^2 + (y - f_k(\xi) - nD)^2]^2}, \end{aligned} \quad (7)$$

$$\sigma_{zz}^{(k,0)}(x, y - nD, \xi) = -\frac{\mu b_{sp} \nu}{2\pi(1 - \nu)} \frac{y - f_k(\xi) - nD}{(x - \xi)^2 + (y - f_k(\xi) - nD)^2},$$

$$\sigma_{zx}^{(k,0)}(x, y - nD, \xi) = -\frac{\mu b_e}{2\pi} \frac{y - f_k(\xi) - nD}{(x - \xi)^2 + (y - f_k(\xi) - nD)^2},$$

$$\sigma_{zy}^{(k,0)}(x, y - nD, \xi) = \frac{\mu b_e}{2\pi} \frac{x - \xi}{(x - \xi)^2 + (y - f_k(\xi) - nD)^2}.$$

Здесь b_{sp} и b_e – соответственно модули краевой и винтовой составляющей вектора Бюргера частичной двойниующей дислокации; μ – модуль сдвига; ν – коэффициент Пуассона.

Результаты расчета и их обсуждение. В расчетах для мартенситной фазы Ni_2MnGa принимались следующие данные: $L = 60$ мкм; $H = 20$ мкм; $D = 45$ мкм; $M = 5$; $\rho_1(\xi) = \rho_2(\xi) = const = \rho = 10^6$; $\nu = 0,3$; $\mu = 19,231$ ГПа; $b_{xp} = 2,91 \text{ \AA}$; $b_s = 2,74 \text{ \AA}$ [11] – [15]. Результаты расчётов отражены на рисунках 2 и 3.

Расчет смещений u_i представлен на рисунке 2. Общий вид линий равных величин смещений для u_x и u_z идентичен (рис. 2, а и в), а отличие заключается в величине численного значения смещений в заданных точках. Рассматриваемые смещения симметричны относительно начала декартовой системы координат (рис. 2, а, в). Результат расчёта полей смещений u_y (рис. 2, б) существенно отличается от результатов, полученных при расчете смещений u_x и u_z . Наибольшие смещения наблюдаются в области зарождения двойников.

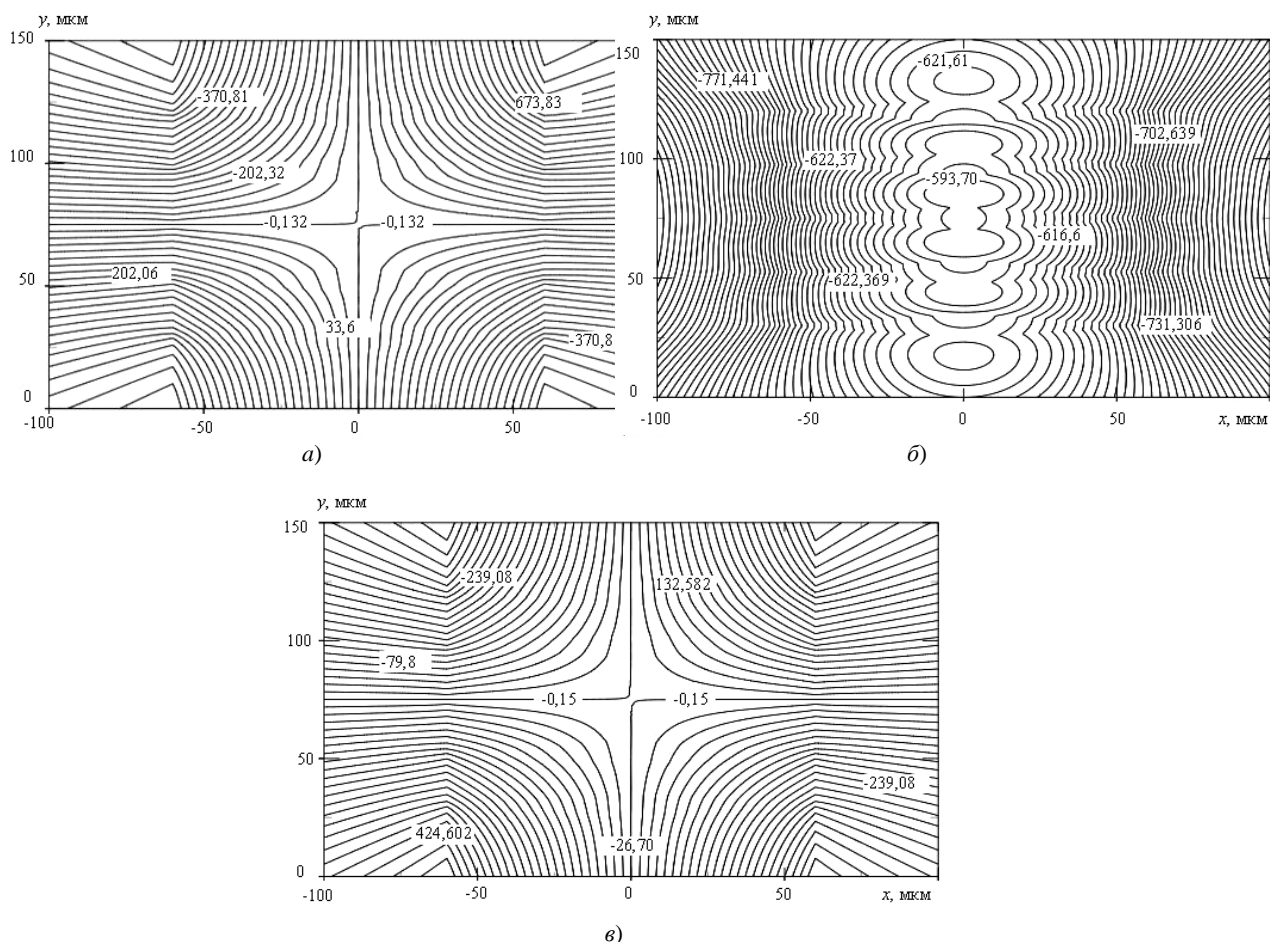


Рис. 2. Результаты расчетов на основании четырехфункциональной модели полей смещений u_i в Ni_2MnGa , обусловленных системой параллельных линзовидных двойников (нм): а – u_x ; б – u_y ; в – u_z

Расчет компонент тензора напряжений представлен на рисунке 3. Во всех случаях расчета компонент тензора напряжений чётко прослеживаются двойниковые границы, являющиеся концентраторами напряжений. Наибольшая концентрация напряжений наблюдается у вершин двойников.

На рисунке 3, а представлен результат расчета *нормальной компоненты тензора напряжений* σ_{xx} . В данном случае максимальные по модулю значения напряжений σ_{xx} наблюдаются у вершин двойников группы параллельных двойников. Напряжения в центре у верхней и нижней границ двойников значительно ниже, чем у вершин и в окрестностях двойников. Уровень напряжений в окружающем двойник материале возрастает по мере продвижения вдоль оси Ox к началу координат. Графический результат расчёта *сдвиговой компоненты тензора напряжений* σ_{xy} показан на рисунке 3, б. Здесь наблюдается несколько иная картина: общий окрестный фронт напряжений постепенно убывает по мере продвижения вдоль оси Ox в положительном направлении. *Распределение сдвиговой компоненты* σ_{xz} представлено на рисунке 3, в. Из полученного результата видно, что напряжения σ_{xz} у верхних и нижних границ двойников существенно различаются по модулю своих значений. Напряжения у верхних границ на порядок вы-

ше напряжений у нижних границ двойников. Более того, напряжения возрастают по мере продвижения вдоль оси OY от крайних двойников к двойникам, находящимся в центре системы параллельных двойников. Также заметно, что линии равных напряжений сходятся к вершинам двойников.

Картина напряжений, приведенная на рисунке 3, *з*, отображает результат расчета *нормальной компоненты тензора напряжений* σ_{yy} . В данном случае наблюдаются минимальные значения напряжений у центральной части системы параллельных двойников, максимальные модульные значения – у крайних двойников. Результат расчета *сдвиговой компоненты* σ_{yz} тензора напряжений представлен на рисунке 3, *д*. Данная компонента имеет симметричные модульные значения напряжений относительно оси OY , а по мере продвижения в положительном направлении вдоль оси OX значения компоненты возрастают. На рисунке 3, *е* показан результат расчета *нормальной компоненты* σ_{zz} тензора напряжений, созданных рассматриваемой системой двойников. Полученные данные имеют схожесть с данными, представленными на рисунке 3, *в*, для компоненты σ_{xz} . Отличие заключается лишь в численных значениях.

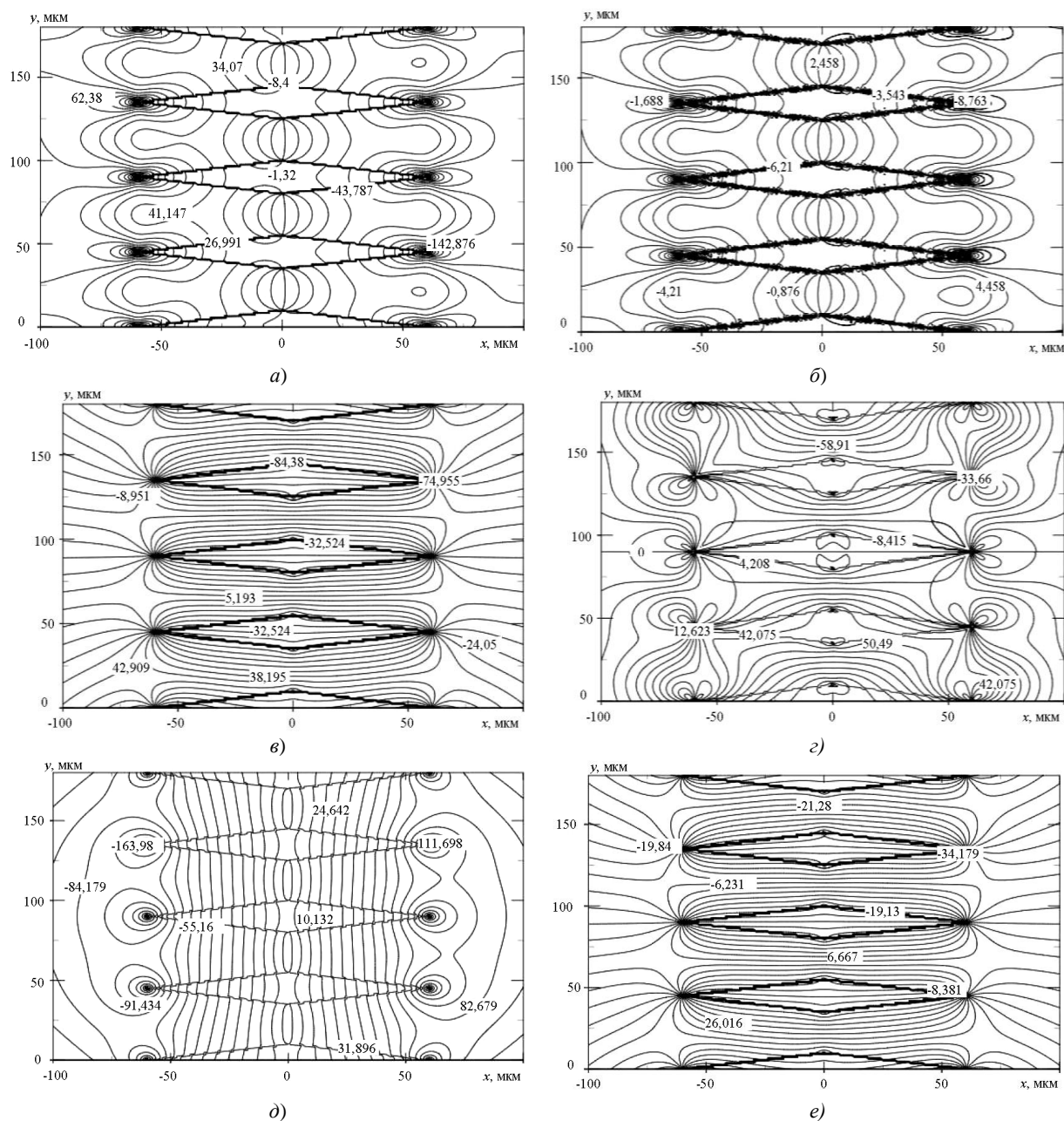


Рис. 3. Поля напряжений σ_{ij} в Ni_2MnGa , созданные параллельными четырёхфункциональными двойниками (МПа):

$$a - \sigma_{xx}; \quad б - \sigma_{xy}; \quad в - \sigma_{xz}; \quad г - \sigma_{yy}; \quad д - \sigma_{yz}; \quad е - \sigma_{zz}$$

Исходя из соотношений Коши [16]

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (8)$$

или формул [16]

$$\left. \begin{aligned} \varepsilon_{xx} &= \frac{1}{E} [\sigma_{xx} - \nu(\sigma_{yy} + \sigma_{zz})], \\ \varepsilon_{yy} &= \frac{1}{E} [\sigma_{yy} - \nu(\sigma_{xx} + \sigma_{zz})], \\ \varepsilon_{zz} &= \frac{1}{E} [\sigma_{zz} - \nu(\sigma_{xx} + \sigma_{yy})], \\ \varepsilon_{xy} &= \frac{1}{2\mu} \sigma_{xy}, \\ \varepsilon_{xz} &= \frac{1}{2\mu} \sigma_{xz}, \\ \varepsilon_{yz} &= \frac{1}{2\mu} \sigma_{yz}, \end{aligned} \right\} \quad (9)$$

где ε_{ij} – тензор деформаций; E – модуль Юнга.

С использованием формул (2) – (7) нетрудно рассчитать и компоненты тензора деформаций.

Заключение. На основании принципа суперпозиции разработан метод расчета напряженно-деформированного состояния в системе параллельных остаточных механических двойников, находящихся вдали от поверхности. Метод применен для расчета напряжений в мартенситной фазе Ni_2MnGa , в которой присутствует такая система двойников. Определены области локализации напряжений, изучена конфигурация полей смещений и напряжений в системе параллельных двойников, возникающих в монокристалле Ni_2MnGa при его деформировании.

ЛИТЕРАТУРА

1. Классен-Неклюдова, М.В. Механическое двойникование кристаллов / М.В. Классен-Неклюдова. – М.: АН СССР, 1960. – 262 с.
2. Косевич, А.М. Дислокации в теории упругости / А.М. Косевич. – Киев: Наук. думка, 1978. – 220 с.
3. Финкель, В.М. Разрушение кристаллов при механическом двойниковании / В.М. Финкель, В.А. Федоров, А.П. Королев. – Ростов н/Д. – 1990. – 172 с.
4. Остриков, О.М. Механика двойникования твердых тел: монография / О.М. Остриков. – Гомель: Учреждение образования «Гомельский государственный технический университет им. П.О. Сухого», 2008. – 301 с.
5. Васильев, А.Н. Ферромагнетики с памятью формы / А.Н. Васильев, В.Д. Бучельников, Т. Такаги // Успехи физических наук – 2003 – Т. 173, № 6. – С. 577 – 608.
6. Бучельников, В.Д. Магнитные сплавы с памятью формы: фазовые переходы и функциональные свойства / В.Д. Бучельников, А.Н. Васильев, В.В. Коледов // Успехи физических наук. – 2006. – Т. 176, № 8. – С. 900 – 906.
7. Planes, A. Ferromagnetic shape memory alloys / A. Planes, L. Manosa // Materials Science Forum. – 2006. – Vol. 512. – P. 145 – 152.
8. Остриков, О.М. Особенности механического двойникования и локального разрушения монокристаллического сплава Ni_2MnGa при индентировании его поверхности / О.М. Остриков, Е.В. Шматок // Вестн. Тамбов. ун-та. Серия. Естественные и технические науки. – 2013. – Т. 18, № 4 (ч. 2). – С. 1924 – 1925.

9. Остриков, О.М. Исследование пластической деформации поверхности монокристалла Ni₂MnGa методом индентирования / О.М. Остриков, А.Л. Созинов, А.В. Сорока // Инженерно-физический журнал. – 2012. – Т. 85, № 5. – С. 1132 – 1141.
10. Остриков, О.М. Метод расчета полей напряжений в системе параллельных остаточных линзовидных двойников в сплаве Гейслера Ni₂MnGa / О.М. Остриков, Е.В. Шматок // Вестн. Тамбов. ун-та. Серия: Естественные и технические науки. – 2013. – Т. 18, № 4 (ч. 2). – С. 1920 – 1921.
11. Sozinov, A. Giant magnetic-field-induced strain in NiMnGa seven-layered martensitic phase / A. Sozinov [et al.]. – Helsinki, University of Technology, Laboratory of Physical Metallurgy and Materials Science; App.phys.letters. – 2002. – Vol. 80. – P. 1746 – 1748.
12. Heczko, O. Temperature dependence and temperature limits of magnetic shape memory effect / O. Heczko, L. Straka // Journal of Applied Physics. – 2003. – Vol. 94, № 11. – P. 7139 – 7143.
13. Heczko, O. Magnetic properties and domain structure of magnetic shape memory Ni-Mn-Ga alloy / O. Heczko, K. Jurek, K. Ullakko // Journal of Magnetism and Magnetic Materials – 2001. – Vol. 226 – 230. – P. 996 – 998.
14. Muellner, P. Large cyclic deformation of a Ni-Mn-Ga shape memory alloy induced by magnetic fields / P. Mullner [et al.]. – Journal of Applied Physics. – 2002. – Vol. 92, № 11. – P. 6708 – 6713.
15. Hirsinger, L. Modelling detwinning of martensite platelets under magnetic and stress actions on Ni-Mn-Ga alloys / L. Hirsinger, C. Lexcellent // Journal of Magnetism and Magnetic Materials – 2003. – V. 254 – 255. – P. 275 – 277.
16. Хирт, Дж. Теория дислокаций / Дж. Хирт, И. Лоте. – М.: Атомиздат, 1972. – 600 с.

Поступила 04.09.2013

**FOUR-FUNCTIONAL METHOD OF CALCULATION
OF THE INTENSE-DEFORMED CONDITION
IN THE SYSTEM OF PARALLEL RESIDUAL KIND LENS DEFORMATION TWINS
IN MARTENSIT PHASE Ni₂MNGA**

O. OSTRIKOV, E. SHMATOK

The method of calculation of the intense-deformed condition caused by residual parallel kind lens twins, being far from a surface is developed. The method is suitable for calculation of displacement, deformations and pressure at twins with the difficult form of borders. It is established that the maximum pressure is localized at tops of twins and on their borders.

УДК 550.837.75:553.98

ВЗАИМОДЕЙСТВИЕ АМПЛИТУДНО-МОДУЛИРОВАННОГО СИГНАЛА С ПОДСТИЛАЮЩЕЙ СРЕДОЙ НАД АНИЗОТРОПНЫМИ ОБРАЗОВАНИЯМИ

*канд. техн. наук, доц. В.Ф. ЯНУШКЕВИЧ
(Полоцкий государственный университет)*

Представлены результаты теоретического анализа взаимодействия амплитудно-модулированного сигнала с анизотропной средой. Исследованы электродинамические параметры поверхностей над залежами углеводородов. Для расчетов используются экспериментально полученные данные характеристик среды реальных месторождений. Проанализировано поведение тензора диэлектрической проницаемости в зависимости от вариации свойств подстилающих поверхностей. Исследования выполнены на основе квазигидродинамического метода исследования параметров сред. Даны рекомендации по оптимизации характеристик электромагнитных волн для повышения точности и достоверности оконтуривания залежей нефти и газа. Результаты исследований могут быть использованы для разработки новых методов поиска и идентификации углеводородов.

Введение. Актуальность рассматриваемых в настоящей работе задач заключается в анализе влияния характеристик поверхностей на электродинамические параметры среды над углеводородными залежами (УВЗ) в режиме амплитудно-модулированных (АМ) сигналов. Электромагнитные методы (ЭММ) разведки залежей нефти и газа (углеводородов) основаны на дифференциации земной поверхности по электромагнитным свойствам [1]. Характеристики электромагнитных волн (ЭМВ) определяются геоэлектрическим строением изучаемого профиля местности. Выделение искомого объекта осуществляется по выявленным электромагнитным аномалиям на данной местности. Анализ существующих методов показывает, что при представлении среды над УВЗ в виде анизотропной неоднородности (АН) плазмopodobного типа использование для поиска и выделения месторождений нефти и газа АМ-сигналов ведет к повышению точности и достоверности георазведки углеводородов [2; 3].

Теоретический анализ. В данной работе проведен анализ влияния подстилающих поверхностей на электродинамические характеристики сред над УВЗ в режиме АМ-сигналов вида:

$$e(t) = E(1 + k_m \cos \Omega t) \cos \omega t, \tag{1}$$

где E – амплитуда несущего колебания частоты $\omega = 2\pi f$; k_m – коэффициент амплитудной модуляции; $\Omega = 2\pi F$ – частота модуляции.

Тензор диэлектрической проницаемости может быть представлен следующим образом [1]:

$$\hat{\epsilon} = \begin{vmatrix} \dot{\epsilon}_1 & -j\dot{\epsilon}_2 & 0 \\ j\dot{\epsilon}_2 & \dot{\epsilon}_1 & 0 \\ 0 & 0 & \dot{\epsilon}_3 \end{vmatrix}. \tag{2}$$

Компоненты тензора диэлектрической проницаемости АН для такого режима имеют вид [2]:

$$\begin{aligned} \dot{\epsilon}_1 &= \epsilon_r + \sum_{i=1}^2 \omega_{pi}^2 \frac{\omega_{Gi}^2 - \omega^2 - \nu_i^2}{(\nu_i^2 + \omega_{Gi}^2 - \omega^2) + 4\omega^2 \nu_i^2} + j \left[\frac{\epsilon_r \cdot k_m \cdot \Omega \cdot \sin \Omega t}{\omega + (1 + k_m \cdot \cos \Omega t)} - \frac{\sigma_r}{\omega \epsilon_0} - \frac{\omega_{pi}^2 \nu_i}{\omega} \cdot \frac{\omega^2 + \nu_i^2 + \omega_{Gi}^2}{(\nu_i^2 + \omega_{Gi}^2 - \omega^2) + 4\omega^2 \nu_i^2} \right], \\ \dot{\epsilon}_2 &= \sum_{i=1}^2 \left\{ \frac{\omega_{pi} \omega_{Gi}}{\omega} \cdot \frac{\omega_{Gi}^2 - \omega^2 + \nu_i^2}{(\nu_i^2 + \omega_{Gi}^2 - \omega^2)^2 + 4\omega^2 \nu_i^2} - \frac{2j \nu_i \omega_{pi} \omega_{Gi}}{(\nu_i^2 + \omega_{Gi}^2 - \omega^2)^2 + 4\omega^2 \nu_i^2} \right\}, \\ \dot{\epsilon}_3 &= \epsilon_r + \sum_{i=1}^2 \left\{ \omega_{pi}^2 \cdot \frac{1}{\nu_i^2 + \omega^2} + j \left[\frac{\epsilon_r k_m \Omega \sin \Omega t}{\omega(1 + k_m \cos \Omega t)} - \frac{\sigma_r}{\omega \epsilon_0} - \frac{\omega_{pi}^2 \nu_i^2}{\omega} \cdot \frac{1}{\nu_i^2 + \omega^2} \right] \right\}, \end{aligned} \tag{3}$$

где ϵ_r – диэлектрическая проницаемость среды над УВЗ; ω_{pi}, ω_{Gi} – плазменная и гиротропная частоты; ν_i – частота столкновений электрона с тяжелыми частицами; σ_r – удельная электрическая проводимость. При разработке новых ЭММ георазведки большое внимание уделяется определению глубины залегания УВЗ, поэтому анализ составляющей тензора $\dot{\epsilon}_3$ для геологических объектов, отличающихся по диэлектрической, магнитной проницаемости, по электропроводности, весьма актуален.

Результаты и их обсуждение. График частотной зависимости вещественной части диэлектрической проницаемости приведен на рисунке 1.

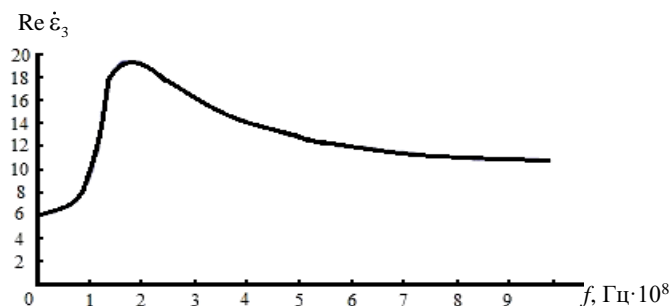


Рис. 1. Частотная зависимость $\text{Re } \dot{\epsilon}_3$

Практически все вещества, кроме чистых металлов, могут быть отнесены к классу диэлектриков с конечной проводимостью. В связи с этим применяют понятие комплексной диэлектрической проницаемости:

$$\epsilon = \dot{\epsilon} - j\dot{\epsilon}', \quad (4)$$

где $\dot{\epsilon}$ — действительная часть, связанная с поляризацией диэлектрика под действием приложенного поля; $\dot{\epsilon}'$ — мнимая часть, связанная с конечной проводимостью диэлектрика, $\text{Ом} \cdot \text{м}^{-1}$.

Величина диэлектрической проницаемости нелинейно увеличивается от 6 на частотах $0 \dots 10$ Гц до 19 при $f = 200$ МГц. С дальнейшим увеличением частоты зондирующей ЭМВ значение $\text{Re } \dot{\epsilon}_3$ нелинейно уменьшается до величины 10,5 на частоте $f = 1000$ МГц.

Приведенные расчеты показывают, что наибольшее влияние на модуляцию вещественной части составляющей тензора диэлектрической проницаемости оказывают частоты в диапазоне от 100 до 300 МГц. При этом вариации коэффициента амплитудной модуляции и частоты модуляции не оказывают воздействия на величину $\text{Re } \dot{\epsilon}_3$.

На рисунке 2 представлена зависимость модуля вещественной части диэлектрической проницаемости от удельной проводимости среды над УВЗ.

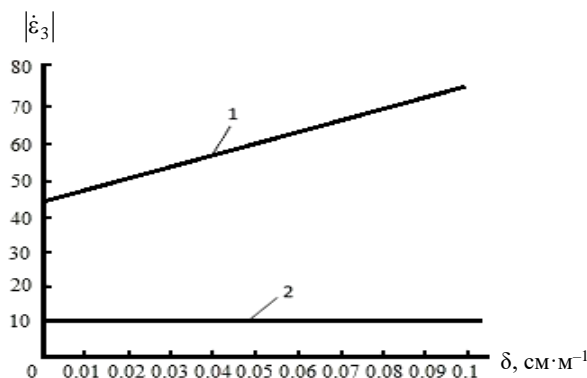


Рис. 2. Влияние удельной проводимости среды δ на модуль тензора диэлектрической проницаемости:
1 — для $f = 100$ МГц; 2 — для $f = 1000$ МГц

Наиболее стабильно величина $|\dot{\epsilon}_3|$ ведет себя для частоты ЭМВ $f = 1000$ МГц, когда при изменении удельной проводимости в пределах от 0 до $0,1 \text{ см} \cdot \text{м}^{-1}$ модуль вещественной части диэлектрической проницаемости лежит в диапазоне $10,78 \dots 10,94$. При уменьшении частоты зондирующего сигнала модуль тензора изменяется по линейному закону, увеличиваясь с ростом δ . Так, для $f = 100$ МГц вариация величины $|\dot{\epsilon}_3|$ лежит в пределах $47 \dots 72$. Для частоты ЭМВ $f = 1$ МГц диапазон изменения модуля $\dot{\epsilon}_3$ составляет $5100 \dots 8600$.

Поведение фазы диэлектрической проницаемости показано на рисунке 3. Фаза составляющей тензора в исследуемом диапазоне частот положительна. Она уменьшается с ростом частоты. Наибольшее влияние на вариацию фазовой компоненты оказывают более низкие частоты несущего колебания. Так,

для частоты $f = 0 \dots 10$) Гц $\arg \dot{\epsilon} \approx 42^\circ$, а на частоте 200 МГц фаза 20° . В более высокочастотном диапазоне наблюдается плавное уменьшение фазовой компоненты. На частоте $f = 600$ МГц величина $\arg \dot{\epsilon} \approx 12^\circ$, а затем она медленно уменьшается до 10 на частоте ЭМВ $f = 1000$ МГц.

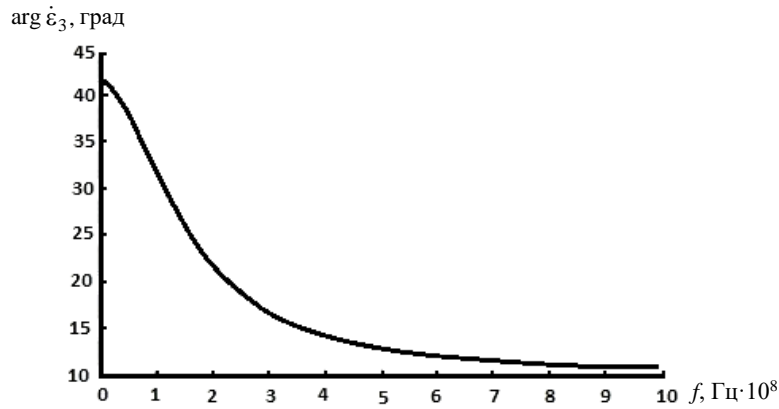


Рис. 3. Частотная зависимость фазовой составляющей тензора диэлектрической проницаемости $\arg \dot{\epsilon}_3$

На рисунке 4 представлен анализ поведения модуля диэлектрической проницаемости в зависимости от величины ϵ_r на частоте $f = 100$ МГц.

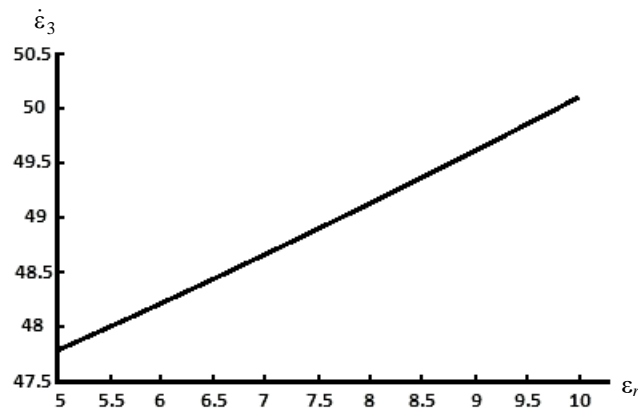


Рис. 4. Зависимость $|\dot{\epsilon}_3 = \psi(\epsilon_r)|$

Расчеты показали, что величина модуля диэлектрической проницаемости линейно возрастает с ростом ϵ_r , увеличиваясь от 47,6 до 50,1. Аналогичный характер прослеживается в диапазоне несущей частоты (100...1000) МГц, за исключением численных значений, которые уменьшаются по сравнению со значениями, представленными на рисунке 4.

На отрезке частот 1...100 МГц величина $|\dot{\epsilon}_3|$ изменяется также по линейному закону, однако уменьшается с ростом ϵ_r .

Концентрация частиц N_i показывает влияние на изменение плазменной частоты:

$$\omega_{pi} = q_i \left(\frac{N_i}{m_i \epsilon_0} \right)^{\frac{1}{2}}, \tag{5}$$

где q_i, m_i – заряд и масса частиц, $\epsilon_0 = 8,85 \cdot 10^{-12}$ ф/м – диэлектрическая проницаемость вакуума.

При расчетах использовались данные для равновесной концентрации электронов и ионов, т.е. $N_e = N_u$.

Влияние концентрации частиц отражено на рисунке 5.

Расчеты представлены на частоте $f = 1000$ МГц.

Для исследуемого диапазона концентраций частиц вариация модуля составляет примерно октаву, изменяясь от 10,5 до 19,5. Наблюдается линейная зависимость указанной характеристики. Аналогичный характер прослеживается для других несущих частот ЭМВ, за исключением численных значений.

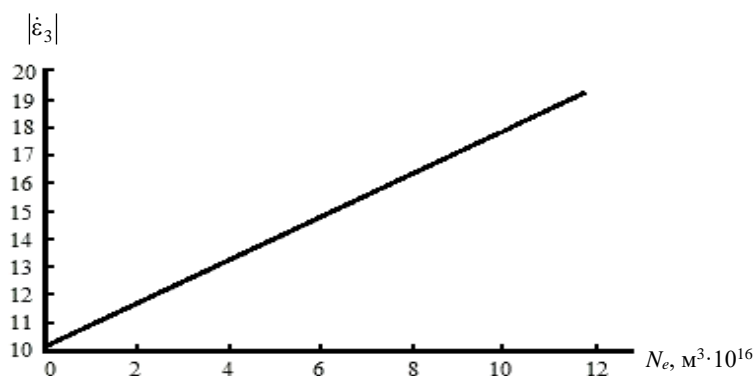


Рис. 5. Зависимость $|\dot{\epsilon}_3| = \psi(N_e)$

На рисунке 6 представлены зависимости модуля вещественной составляющей от частоты столкновений частиц. Режим исследования выбран следующий: на частоте зондирующей ЭМВ $f = 100$ МГц, коэффициент амплитудной модуляции $K_m = 1$. С ростом частоты модуль линейно уменьшается. При частоте столкновений $\nu = 2 \dots 4 \cdot 10^8 \text{ с}^{-1}$ величина $|\dot{\epsilon}_3|$ лежит в пределах 85...75. При высоких частотах столкновений наблюдается плавное уменьшение модуля диэлектрической проницаемости. Аналогичный характер зависимостей прослеживается и для других несущих частот зондирующего сигнала, за исключением численных значений. На частотах более 1000 МГц характеристики приобретают вид с выпуклостью вверх.

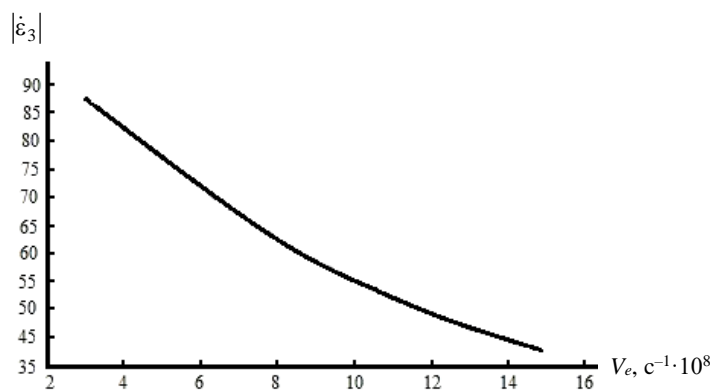


Рис. 6. Зависимость $|\dot{\epsilon}_3| = \psi(\nu)$

Поскольку для исследований применяется АМ-сигнал, поэтому представляет интерес влияние модулирующих факторов на модуль диэлектрической проницаемости (рис. 7).

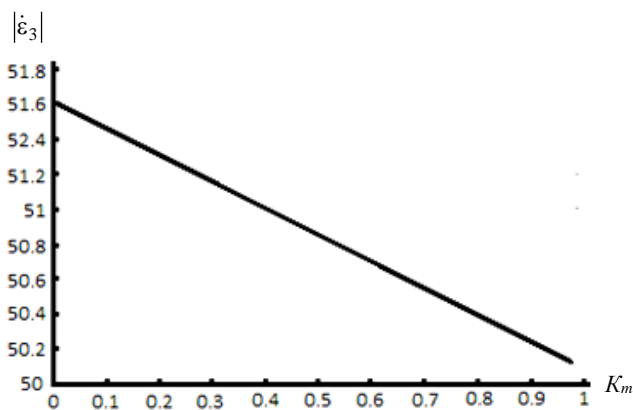


Рис. 7. Зависимость $|\dot{\epsilon}_3| = \psi(K_m)$

Зависимость рассчитана для $f = 100$ МГц. Следует отметить, что вариация характеристики несущественна, что объясняется незначительным вкладом составляющей тензора (3), включающей величину K_m . С ростом коэффициента АМ значение $|\dot{\epsilon}_3|$ уменьшается по линейному закону. В данном случае приращение $\Delta \dot{\epsilon}_3 = 1,5$.

В диапазоне частот, больших 1000 МГц, происходит нелинейное незначительное увеличение $|\dot{\epsilon}_3|$ с ростом величины K_m .

Заключение. Проведенный анализ влияния параметров подстилающей среды и зондирующего сигнала на электродинамические характеристики среды над УВЗ показал, что результаты исследований могут быть использованы для определения глубины залегания углеводородов.

Данные, полученные на основе проведенного исследования, могут быть использованы при построении радиотехнических систем поиска и обнаружения месторождений нефти и газа.

ЛИТЕРАТУРА

1. Москвичев, В.Н. Исследования взаимодействия электромагнитных волн с углеводородной залежью / В.Н. Москвичев // Радиотехника и электроника. – Минск: Выш. шк., 1989. – Вып. 18. – С. 91 – 96.
2. Гололобов, Д.В. Влияние параметров амплитудно-модулированного воздействия на фазовые составляющие поверхностного импеданса анизотропной среды типа / Д.В. Гололобов, О.В. Орлова, В.Ф. Янушкевич // Вестн. Полоц. гос. ун-та. Серия С. Фундаментальные науки. – 2010. – № 9. – С. 98 – 101.
3. Гололобов, Д.В. Импедансные граничные условия анизотропной среды для амплитудно-модулированного сигнала / Д.В. Гололобов, С.В. Калинин, В.Ф. Янушкевич // Докл. БГУИР. – Минск, 2010. – № 6(52). – С. 13 – 17.

Поступила 01.04.2013

INTERACTION OF THE AMPLITUDE-MODULATED SIGNAL WITH THE SPREADING ENVIRONMENT OVER ANISOTROPIC FORMATIONS

V. YANUSHKEVICH

Results of the theoretical analysis of interaction of the amplitude-modulated signal with the anisotropic environment are presented. Electrodynamics parameters of surfaces over deposits of hydrocarbons are investigated. For calculations experimentally obtained data of characteristics of the environment of real fields are used. The behavior of a tensor of dielectric permeability depending on a variation of properties of spreading surfaces is analysed. Researches are executed on the basis of a quasihydrodynamic method of research of parameters of environments. Recommendations about optimization of characteristics of electromagnetic waves for increase of accuracy and reliability of contouring of deposits of oil and gas are made. Results of researches can be used for development of new methods of search and identification of hydrocarbons.

УДК 621.354.6

ЗАРЯДНО-ПИТАЮЩИЕ УСТРОЙСТВА НА ОСНОВЕ ЕМКОСТНЫХ ЭЛЕМЕНТОВ

*канд. техн. наук, доц. С.Э. ЗАВИСТОВСКИЙ, П.П. РЕДЬКО
(Полоцкий государственный университет)*

Анализируются методика и способы заряда аккумуляторных батарей. На основании проведенного анализа разработана схемотехника зарядно-питающих устройств с использованием емкостных элементов, предусматривающая цель долговечности, надежности и расширения диапазона действия (универсальность), а также электронного упрощения при сохранении соответствующих параметров. Приведены необходимые данные и наименования элементной базы для реального практического применения и соответствующие примеры расчета. Показано, что емкостно-преобразовательный подход к решению рассматриваемого вопроса позволяет совмещать простоту схемотехнических решений с автоматической стабилизацией зарядного тока в силу превращения источника напряжения в источник тока.

В настоящее время существует множество методик и практических рекомендаций, а также многообразие реализованных и проектируемых устройств для заряда (доподзарядки) аккумуляторов и аккумуляторных батарей различного назначения: стационарных, передвижных (автотранспортных), малогабаритных, миниатюрного исполнения и т.п., по химическому составу: кислотных, щелочных, оксидных и иных. Однако всех их можно подразделить на две основные группы по виду электрического заряда, т.е. при использовании в этих целях 1) источника постоянного по величине напряжения и 2) источника постоянного по величине тока. К отдельной группе можно отнести зарядные устройства, работающие по усложненным принципам с соответствующими выходными характеристиками, имеющие цель лечебно-профилактических и иных специальных мер.

Источник постоянного напряжения при использовании его в режиме подзаряда аккумуляторных батарей порождает непостоянный по величине ток на протяжении цикла заряда. В исходном разряженном состоянии аккумулятор потребляет наибольший ток, обладая пониженным собственным напряжением на его клеммах. По мере зарядки повышается напряжение самого аккумулятора, и разница между напряжением зарядного устройства и аккумулятора уменьшается, что приводит к результирующему снижению действующего зарядного напряжения и, следовательно, тока. Этот принцип является предпочтительным в случае подзаряда в нестационарном состоянии (например, движущиеся транспортные средства), когда аккумулятор загерметизирован, и основной дозаряд происходит без активного газовыделения [1; 2]. Во-вторых, данный метод оправдан в необслуживаемом постоянно пункте, когда имеет место сверхнормативное время заряда, что может привести при постоянном по величине зарядном токе к перезаряду, создающему угрозу осыпания пластин и разрушения их активной массы и даже к их короткому замыканию.

Для полной реабилитации аккумулятора в режиме дозаряда реализуют в стационарных условиях в основном ток постоянный и оптимальный по величине на протяжении всего цикла заряда, так как превышение приводит к короблению пластин и преждевременному осыпанию (разрушению) их активной массы, а уменьшение тока (ниже оптимального) не позволяет восстановить активный состав по всей толщине пластин, порождая тем самым необратимый процесс их сульфатации и, следовательно, сокращение срока службы аккумулятора.

Исключением являются малогабаритные и миниатюрные аккумуляторы, так называемые «батарейки», дозаряд которых рекомендуется проводить иногда уменьшенным током, даже до 1/4 от оптимального, с соответствующим сокращением времени дозаряда, реализованные на ином химическом принципе (веществе), так называемые «сухие», перезаряд которых можно контролировать, например, по повышенному нагреву корпуса, а также те, для которых завод-изготовитель предусматривает специально рекомендованный им оптимальный режим подзаряда.

Для основных же типов аккумуляторов оптимальным током является ток, составляющий 0,1/h от его номинальной емкости, приведенной к так называемому десятичасовому режиму заряд-разряда C_{10} , Ah. Для автотранспортных аккумуляторов принята двадцатичасовая емкость, т.е. C_{20} , Ah (ампер·час). Однако оптимальным током стационарного дозаряда является десятичасовой режим, т.е. C_{10} , Ah. Например, для аккумулятора $C_{20} = 60$ Ah дозаряд осуществляют исходя из ранее указанных рекомендаций током 0,1/h от C_{20} , т.е. 6 А, за исключением первоначального заряда, осуществляемого, как правило, в условиях завода-изготовителя током 0,05/h от C_{20} в течение всего цикла заряда, т.е. 20 часов, и «загерметизированного» с ограниченным газовыделением (в стационарных условиях следует открывать воздушные окна) в течение

0,5 времени цикла дозаряда током 0,1/h от C_{20} и оставшееся время током 0,05/h от C_{20} во избежание активного газовыделения. Максимальный ускоренный режим, осуществляемый в экстремальных и исключительных случаях, составляет не более 0,15/h от C_{20} первой половины цикла дозаряда и оставшееся время не более 0,1/h от C_{20} . Под первой половиной цикла дозаряда понимают время, в течение которого аккумулятор набирает примерно 75 % (3/4) его номинальной емкости. Напряжение при этом достигает 2,4 В на элемент (для 12 В имеем $2,4 \cdot 6 = 14,4$ В; для 24 В – 28,8 В), плотность электролита 1,25...1,26 г/см³. Окончание заряда можно определить по установившейся постоянной плотности электролита в пределах 1,27...1,29 г/см³, постоянству напряжения на клеммах и активному газовыделению [1 – 2].

Величина емкости аккумулятора обычно указывается приведенная к 10-ти часовому режиму заряд-разряда. При разряде током 0,1/h аккумулятор расходует свой ресурс в соответствии с указанным временем цикла, т.е. в данном случае 10 часов. При разряде меньшим по величине током его адекватный цикл увеличивается, т.е. при разряде меньшим током полученная от аккумулятора энергетическая емкость окажется больше указанной в паспорте.

В настоящее время существует огромное разнообразие всевозможных зарядных устройств аккумуляторных батарей, реализованных по усложненным электронным и трансформаторным упрощенным схемам. Однако существенное количество их не выдерживают токовых перегрузок и приходят в негодность, в частности при всевозможных неисправностях заряжаемых аккумуляторов. Кроме того, они, как правило, предусмотрены для заряда аккумуляторов определенных номиналов по напряжению и токовым нагрузкам. Зарядные устройства, обеспечивающие постоянство тока заряда на протяжении цикла, являются более сложными.

На рисунке 1 приведена схемотехника зарядного устройства, реализованного на емкостных элементах.

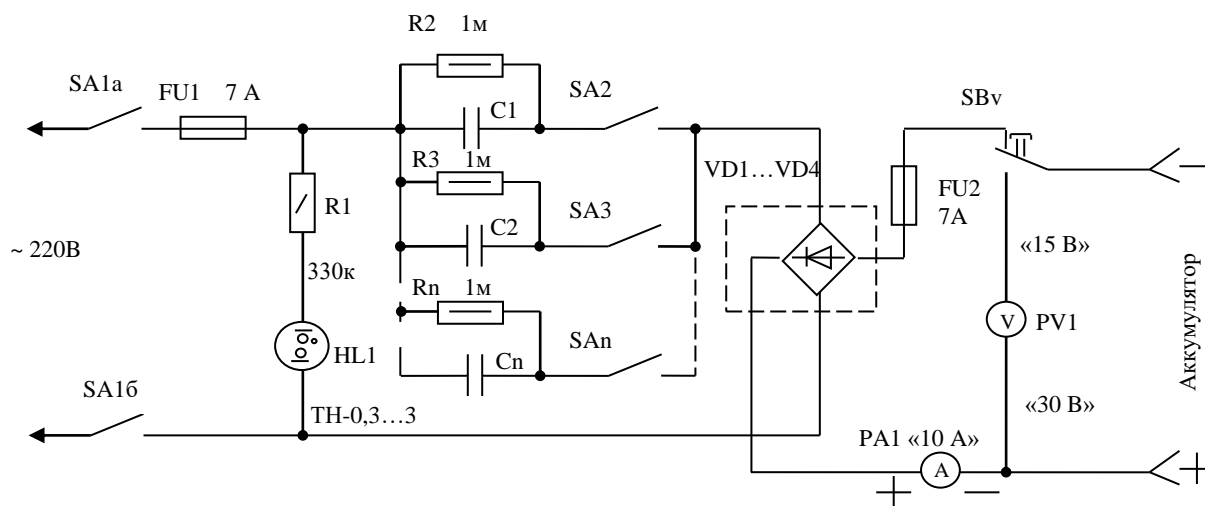


Рис. 1. Зарядное бестрансформаторное устройство аккумуляторов

Помимо простоты конструкции, надежности и долговечности, в силу $x_c \gg r_{ак}$ представленное зарядное устройство является источником постоянного по величине тока, зависящего от емкости набора конденсаторов и питающего сетевого напряжения и не зависящего практически от состояния аккумуляторов, их неисправностей, а также от короткого замыкания на выходных клеммах.

Величина емкости набора конденсаторов определяется диапазоном требуемого зарядного тока в зависимости от параметров аккумулятора.

Например, для аккумулятора в 55 Ah емкость конденсаторов составит при заряде током 0,1/h

$$C_{[мкФ]} = \frac{I_{зар} \cdot 10^6}{2\pi f (U - U_{зар.аккумулятор} - U_{VD})} = \frac{5,5 \cdot 10^6}{2 \cdot 3,14 \cdot 50 \cdot (220 - 16 - 4)} = 88.$$

При заряде током 0,05/h, $C = 44$ мкФ для аккумуляторов в 70 Ah емкость конденсаторов составит 111 и 55 мкФ соответственно. Для аккумулятора 24 В или двух последовательно соединенных 12 В, т.е. $2U_{зар}$ емкости будут соответственно для 55 Ah 95 (48) мкФ, 70 Ah 121 (60) мкФ, т.е. данное зарядное устройство без каких-либо изменений пригодно и для зарядки номиналов аккумуляторных батарей больше-

го или меньшего значений, а также можно использовать их последовательное соединение со степенью точности, зависящей от измененных номиналов.

Набор конденсаторов $C_1...C_n$ с коммутационными переключателями SA2...SA n (например, 48 мкФ, 24 мкФ, 16 мкФ) обеспечат возможную токовую комбинацию в зарядном устройстве 1,5 А и 1 А, являясь, таким образом, ступенчатой регулировкой.

При изменении номинала переменного питающего напряжения ~ 220 В (например, ~ 380 В) ток зарядки будет изменен в той же пропорции (в соответствии с этим необходимо скорректировать и допустимые значения напряжений на конденсаторах и диодах).

Конденсаторы используются неэлектролитические, предназначенные для работы в электроустановках на допустимое рабочее напряжение не менее 250...300 В (для питающего напряжения ~ 220 В). Приоритетом является взрывобезопасное их исполнение. В противном случае при испытаниях, а также в окончательном варианте предусматривается защитная конструкция корпуса.

Следует внимательно относиться к использованию конденсаторов длительного хранения и бывших в употреблении. Входной предохранитель должен быть из легкоплавкого металла в целях быстрого разъединения и незначительного превышения номинала предельного рабочего тока.

Переключатели должны выдерживать максимальные рабочие токи (иногда используют запараллеливание контактов). В качестве входного SA1 можно применить сдвоенный автоматический типа ВА47-39-С8 и аналогично одинарные в качестве SA1... SA n . Учитывая длительность срабатывания по времени их тепловой защиты, наличие входного легкоплавкого предохранителя в данной конструкции обязательно.

Резисторы R2...R n выполняют функцию повышенной утечки, т.е. разряда конденсаторов при их отключении.

Диоды выбирают на допустимое напряжение не менее 300 В и ток не менее $0,5I_{зар.мах}$. Например, для наибольшего значения тока в 10 А подойдут диоды типов: Д231, Д245, Д246, Д247, Д248 – с любыми буквами: КД202 Ж, К, Н, Р и иные [4]. С точки зрения рациональности пользуются справочными данными. Монтируются на радиаторы с охлаждаемой поверхностью в соответствии с токопотерями.

Предохранитель FU₂ необходим как защита от короткого замыкания аккумулятора при возможном пробое диодов.

В качестве вольтметра может быть использован прибор электронного типа при постоянном неотключаемом режиме на предельное напряжение 300 В.

Время дозаряда рассчитывается как разность заводской емкости (с учетом процента потерь в зависимости от длительности и интенсивности эксплуатации аккумулятора) и существующей на момент заряда, разделенной на номинальный ток:

для 55 А·ч (7 % потерь)

$$\frac{51 \text{ А} \cdot \text{ч} - 32 \text{ А} \cdot \text{ч} (\text{емкость разряженного})}{5,5 \text{ А}} \approx 35 \text{ ч.}$$

Величину емкости аккумулятора на момент дозаряда можно оценить опытным путем по результатам эксплуатации, а также замерам плотности электролита и напряжения на клеммах в соответствии с представленной таблицей [3].

Состояние степени заряженности аккумуляторной батареи

Процент заряженности аккумулятора	Плотность электролита, г/см ³	Напряжение на клеммах, В
100	1,27...1,29	12,7...12,8
75	1,24...1,26	12,5...12,6
50	1,20...1,22	12,2...12,3
20	1,15...1,17	11,95...12,05
→0	≤ 1,00...1,03	≤ 11,6...11,7

Напряжение на клеммах аккумулятора приведено для 12 В модели (для 24 В необходимо приведенные в таблице данные удвоить).

В целях автоматизации отключения по окончании расчетного времени заряда возможно использование электронного таймера (реле времени). Обобщенная в этом случае схема показана на рисунке 2.

Представленная схема может быть скорректирована аналогичным образом и для автоматического переключения на 0,5I от основного тока заряда в соответствии с рекомендациями [1 – 3].

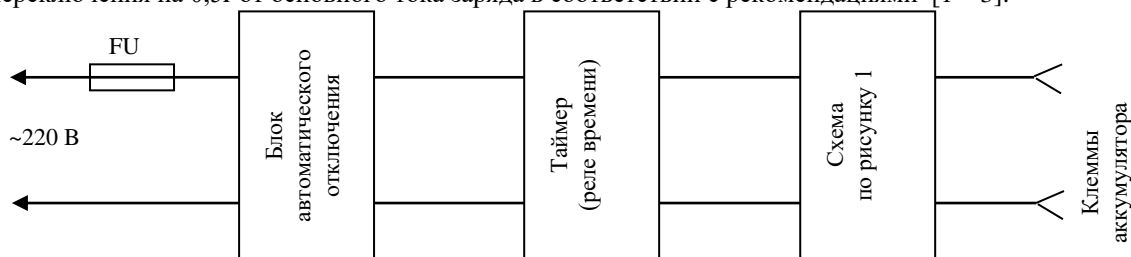


Рис. 2. Схема автоматизации зарядки

С точки зрения энергопотребления зарядное устройство представляет собой активно-емкостную нагрузку. Общий ток при одновременной работе с активной нагрузкой (активное освещение и т.п.) будет определяться векторным сложением. Совместная работа с индуктивностью приводит к ее компенсации и в общем случае к компенсации индуктивной составляющей нагрузки энергосети, что является положительным в системе энергоснабжения (рис. 3, а, б, в).

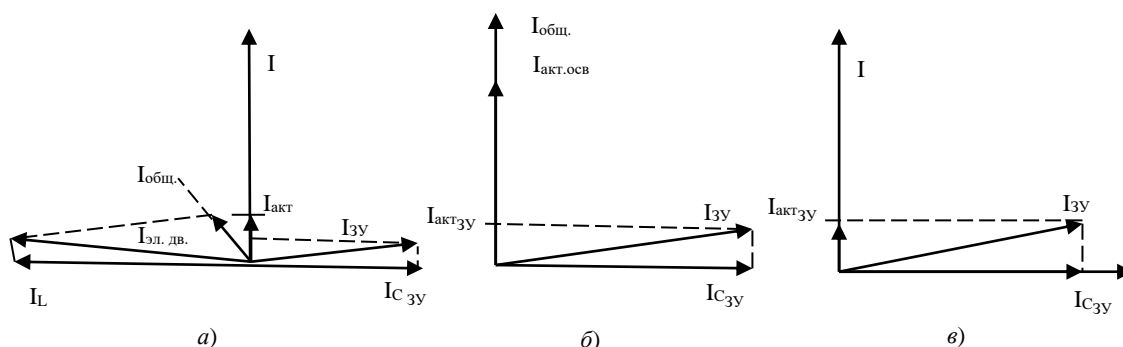


Рис. 3. Векторные диаграммы активно-индуктивно-емкостных нагрузок с учетом зарядного устройства

Активная токовая составляющая зарядного устройства порождает активную потребляемую мощность. Его коммутационные и разъединительные, а также соединительные элементы должны выдерживать полный потребляемый ток с учетом работы. Полный ток зарядного устройства с активным освещением может быть рассчитан примерно так: $\sqrt{I_{акт. осв.}^2 + I_{ЗУ}^2}$ (рис. 3, б).

Положительным фактором в схеме, представленной на рисунке 1, является исключение трансформатора с экономией цветного металла (меди).

В целях электробезопасности (наличие сетевого потенциала непосредственно на аккумуляторных клеммах) при изготовлении и эксплуатации зарядного устройства необходимо учитывать и соблюдать меры техники электробезопасности (безопасные и защищенные от прикосновения присоединительные клеммы, предупреждающие надписи, соответствующая расцветка, ограждения от посторонних лиц и т.д.) [5; 6].

Использование рекомендуется на предприятиях государственного и коммерческого типов с наличием электротехнического персонала, а также индивидуальными владельцами, знающими и выполняющими требования электробезопасности.

Данное зарядное устройство может быть рекомендовано и в гальванике.

При нанесении покрытий ввиду сложного рельефа поверхности, на которую они будут нанесены, и ионных хаотических процессов возникают токовые флуктуации, влияющие на качество исходя из неравномерности нанесения. По рекомендованной схеме они устраняются конденсаторами, выполняющими в данном случае функцию фильтра верхних частот. В этом случае также следует использовать элементы блокировок и защиты с точки зрения электробезопасности [5; 6].

Эту схему в упрощенном варианте (рис. 4) можно применять и для заряда аккумуляторов электроинструментов (дрели, шуруповерты и т.п.). Схема обеспечивает ток зарядки в 150...200 мА. Блок согласования и индикации возможен типовой. Без переделок его рекомендуется применять (во избежание перегрузки цепи индикации напряжения при отсоединенном аккумуляторе или разомкнутой его электрической цепи) в строгой подаче питающего напряжения только при подсоединенном и безобрывном

аккумуляторе. Предпочтительной является модернизация цепи индикации напряжения с добавлением кнопки на переключение, аналогичной SB2 (см. рис. 1).

Так как ток на протяжении всего цикла заряда остается неизменным, индикация окончания заряда возможна по напряжению, т.е. по достижению на аккумуляторе величины 22,5...23 В (номинал 18 В) или 16,5...17 В (номинал 12 В). В этом случае рекомендуется некоторая коррекция параметров резисторов в цепях светодиодной индикации. При нагреве корпуса аккумулятора заряд следует прекратить.

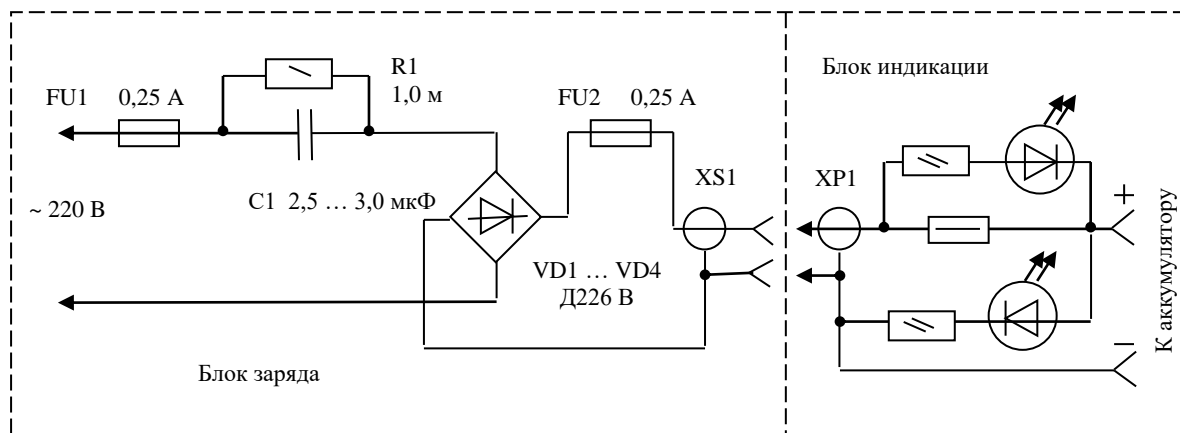


Рис. 4. Схема заряда аккумуляторов электроинструмента

Величину питающего сетевого напряжения можно использовать и меньших номиналов (например, ~127 В, ~110 В), увеличив пропорционально емкость конденсатора C1, что несколько повышает электробезопасность, но понижает возможное пробивное напряжение внутриаккумуляторных элементов.

Аналогичным образом может быть реализована и непосредственная схема сетевого питания некоторых иных потребителей (рис. 5).



Рис. 5. Зарядное устройство в режиме непосредственного электропитания с использованием ограничительных нелинейных элементов

Эта схема по своим функциям аналогична схеме стабилизатора с параллельно включенным регулирующим элементом. Полярный конденсатор (в случае необходимости) сглаживает пульсации выпрямленного напряжения. Второй стабилитрон (дублирующий) выбирают по напряжению, несколько большему, и он выполняет функцию гарантированного (дополнительного) ограничения напряжения.

Возможно использование и трансформаторно-емкостного варианта [7] (рис. 6).

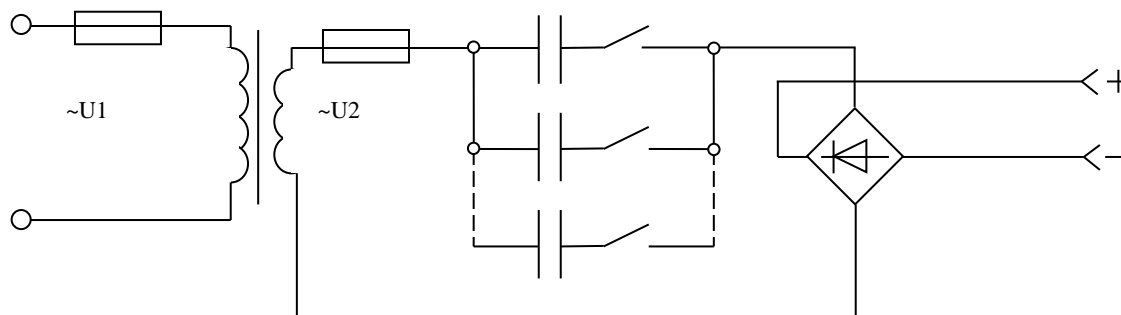


Рис. 6. Трансформаторно-емкостная зарядная схема

Трансформаторно-емкостной зарядный способ рекомендуется для заряда маломощных (миниакумуляторов) аккумуляторов. Величина трансформированного напряжения выбирается в пределах 6...8-кратной относительно заряжаемых аккумуляторов. Принцип работы и расчетные соотношения соответствуют рассмотренным ранее. При напряжении $\sim U_2 \leq 42$ В устройство является электробезопасным.

Заключение. Емкостно-преобразовательный подход к решению рассматриваемой проблемы позволяет совмещать простоту схемотехнических решений с автоматической стабилизацией зарядного тока в силу превращения источника напряжения в источник тока. Изложенный анализ методики заряда позволит потребителям квалифицированно и рационально использовать те или иные зарядные устройства. Разработка, изготовление и использование зарядно-питающих устройств на основе емкостных элементов могут способствовать снижению дефицита цветных металлов, а также трансформаторных и иных узлов, оказать некоторое воздействие на снижение индуктивной составляющей токовой нагрузки энергосети. Однако в этом случае требуется более внимательное и строгое выполнение техники электробезопасности при работе и обслуживании такого рода зарядно-питающих устройств.

ЛИТЕРАТУРА

1. Чижков, Ю.П. Электрооборудование автомобилей и тракторов: учебник / Ю.П. Чижков. – М.: Машиностроение, 2007. – 656 с.
2. Набоких, В.А. Эксплуатация и ремонт автомобилей и тракторов: учебник для студ. высш. учеб. заведений / В.А. Набоких. – 3-е изд. стер. – М.: Издат. центр «Академия». 2006. – 240 с.
3. Справочник по полупроводниковым диодам, транзисторам и интегральным схемам / под. общ. ред. Н.Н. Горюнова. – М.: Энергия, 1976.
4. Белявин, К.Е. Электробезопасность при эксплуатации электроустановок / К.Е. Белявин. – М., 2004. – 186 с.
5. Правила технической эксплуатации электроустановок потребителей и Правила техники безопасности при эксплуатации электроустановок потребителей. – 4-е изд. перераб. и доп. с изм. – М.: Энергоатомиздат, 1989. – 427 с.
6. Редько, П.П. Заряд микроаккумуляторов / П.П. Редько // Радиомир. – 2001. – № 8. – С. 14.

Поступила 10.07.2013

CHARGING-FEEDING DEVICES ON THE BASIS OF CAPACITIVE ELEMENTS

S. ZAVISTOVSKIY, P. REDKO

Methods and means of charging accumulators are analyzed. On the basis of the research a circuit technology of charging-feeding devices based on capacitive elements is developed, which provides for the aim of longevity, reliability and expansion of the range (universality), as well as electronic simplification preserving the corresponding parameters. Necessary data and names of elements base for real practical application and corresponding examples of calculation are given. It is shown, that capacitive-transformative approach to the solution of the given problem allows to combine the simplicity of circuit technological solutions with automatic stabilization of charging current because of transformation of voltage source into current source.

СОДЕРЖАНИЕ

ИНФОРМАТИКА

<i>Железняк В.К., Барков А.В.</i> Маскирование RGB-видеокадров синхронным и адаптивным шумовым RGB-видеокадром.....	2
<i>Бездель А.О., Янушкевич В.Ф., Чертков В.М.</i> Компьютерное моделирование сверхразрешения источников излучения на основе проекционного метода	7
<i>Дворников С.В., Дьяконов В.В., Дворников С.С., Железняк А.В.</i> Синтез вейвлет-манипулированных сигналов	14
<i>Железняк В.К., Рябенко Д.С.</i> Обоснование оптимального сигнала для оценки защищенности цифровых каналов утечки информации.....	18
<i>Деканова М.В.</i> Математическая модель и алгоритм построения расписания учебных занятий университета.....	24
<i>Янушкевич В.Ф., Бездель А.О.</i> Сверхразрешение источников излучения на основе метода Кейпона	34
<i>Дворников С.В., Цветков В.В., Устинов А.А.</i> Компенсация движения в видеокодеках на основе оптимальных разбиений кодируемых блоков.....	39
<i>Радоман Н.В., Александров О.И., Свирский Д.Н.</i> Оптимизация балансовой надежности дефицитной энергосистемы с учетом межсистемных связей	44
<i>Маликов В.В., Криштопова Е.А.</i> Комплексный методический подход оперативной оценки уровня киберпреступлений.....	54
<i>Булавка Ю.А.</i> Нечетко-множественный подход к экспертной оценке профессиональных рисков на примере условий труда работников нефтеперерабатывающего завода	59
<i>Матюш М.В.</i> Обзор альтернативных метрик программного кода для поиска плагиата	67
<i>Linkovich A., Panteleenko F.</i> Predictive Agents for Control of Intelligent Manufacturing Enterprises	72

ФИЗИКА

<i>Ланин В.Л., Лавор Т.Э., Хотькин В.Т.</i> Оптимизация параметров инфракрасного нагрева для монтажа электронных компонентов.....	78
<i>Харченко А.А., Бринкевич Д.И., Бринкевич С.Д., Вабищевич С.А., Вабищевич Н.В., Лукашевич М.Г., Просолович В.С., Оджаев В.Б.</i> Исследование поверхности полимеров модифицированной радиационной обработкой	83
<i>Хорошко В.В., Цырельчук И.Н., Гременок В.Ф.</i> Оптическое поглощение в пленках $Cu_xIn_xZn_{2-2x}Se_2$	91
<i>Остриков О.М., Шматок Е.В.</i> Четырехфункциональный метод расчета напряженно-деформированного состояния в системе параллельных остаточных линзовидных деформационных дойников в мартенситной фазе Ni_2MnGa	95
<i>Янушкевич В.Ф.</i> Взаимодействие амплитудно-модулированного сигнала с подстилающей средой над анизотропными образованиями	101
<i>Завистовский С.Э., Редько П.П.</i> Зарядно-питающие устройства на основе емкостных элементов.....	106