МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Одеський національний політехнічний університет

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Том 6, № 3

Volume 6, No. 3

Одеса – 2016
Odesa – 2016

Журнал внесений до переліку наукових фахових видань України
(технічні науки)
згідно наказу Міністерства освіти і науки України № 463 від 25.04.2013 р.

Виходить 4 рази на рік

Заснований Одеським національним політехнічним університетом у 2011 році

Свідоцтво про державну реєстрацію КВ № 17610 - 6460Р від 04.04.2011р.

Published 4 times a year

Founded by Odessa National Polytechnic University in 2011

Certificate of State Registration KB № 17610 - 6460P of 04.04.2011

Друкується за рішенням редакційної колегії та Вченої ради Одеського національного політехнічного університету

Оригінал-макет виготовлено редакцією журналу

Адреса редакції: просп. Шевченка, 1, Одеса, 65044, Україна
Телефон: +38 048 705 8506
Web: http://immm.opu.ua E-mail: immm.ukraine@gmail.com
Editorial address: 1 Shevchenko Ave., Odessa, 65044, Ukraine
Tel.: +38 048 705 8506
Web: http://immm.opu.ua E-mail: immm.ukraine@gmail.com

ЗМІСТ / CONTENTS

ІНФОРМАТИЗАЦІЯ ПРОГНОЗУВАННЯ РИЗИКУ СТРУКТУРНО СКЛАДНИХ ТЕХНІЧНИХ СИСТЕМ ЗА ДОПОМОГОЮ МОДЕЛЕЙ БАЙЕСОВСКИХ МЕРЕЖ ДОВІРИ В.В. Вичужанін, Н.О. Шибаєва	205	INFORMATIZATION OF PROGNOSTICATION OF RISK STRUCTURALLY OF THE DIFFICULT TECHNICAL SYSTEMS BY MEANS OF MODELS OF THE BAYES NETWORKS OF TRUST Vychuzhanin V., Shibaeva N.
ОПТИМАЛЬНІСТЬ НЕУСІЧЕНОЇ ПОСЛІДОВНОЇ ПРОЦЕДУРИ ВАЛЬДА В ЗАДАЧАХ ПЕРЕВІРКИ ДВОХ ПРОСТИХ ПРОГНОЗІВ НЕСАНКЦІОНОВАНОГО ДОСТУПУ В ІНФОРМАЦІЙНИХ МЕРЕЖАХ ДЕРЖАВИ В.Б. Дудикевич, І.Р. Опірський, П.І. Гаранюк, О.А. Ваврічен	215	OPTIMALITY IS NOT TRUNCATED CONSISTENT PROCEDURES WALD IN SCAN TASKS OF TWO SIMPLE BETS UA IN INFORMATION NETWORKS STATE Dudykevich V., Opirsky I., Garanyuk P., Vavrichen O.
ТРАНСФОРМАЦІЯ ПАРАДИГМ ЗАХИСТУ ІНФОРМАЦІЇ, ІНФОРМАЦІЙНОЇ ТА СОЦІАЛЬНО-ПСИХОЛОГІЧНОЇ БЕЗПЕКИ (Частина 1) С.О. Гнатюк, В.О. Гнатюк, В.Г. Кононович, І.В. Кононович	227	TRANSFORMATION OF INFORMATION AND SOCIAL-PSYCHOLOGICAL SECURITY PARADIGMS (Part 1) Gnatyuk S., Gnatyuk V., Kononovich V., Kononovich I.
ОБГРУНТУВАННЯ ЗАСТОСУВАННЯ ФРАКТАЛЬНОГО ПІДХОДУ ДЛЯ СТВОРЕННЯ КОМПЛЕКСУ АПАРАТУРИ КОНТРОЛЯ СПРАВЖНОСТІ ФОНОГРАМ ПРИ ЕКСПЕРТИЗІ МАТЕРІАЛІВ ТА ЗАСОБІВ ЦИФРОВОГО ЗВУКОЗАПИСУ О.В. Рибальський, В.І. Соловйов, В.В. Журавель	240	JUSTIFICATION OF FRACTAL APPROACH USING TO CREATE COMPLEX EQUIPMENT FOR CONTROL OF AUTHENTIC DIGITAL PHONOGRAMS AT THE EXAMINATION OF DIGITAL AUDIO MATERIALS AND TOOLS Rybalsky O., Solovyov V., Zhuravel V.
МЕТОД КОДУВАННЯ НА ОСНОВІ ФІБОНАЧЧІЄВОЇ Q-МАТРИЦІ А.В. Свірідов, Т.І. Петрушина	249	THE FIBONACCI Q-MATRIX CODING METHOD Sviridov A., Petrushina T.

ОЦІНКА ЯКОСТІ ВИМОВИ МЕТОДОМ ПОРІВНЯННЯ З ЕТАЛОНОМ
Г.А. Добровольський, О.О. Тодоріко, Н.Г. Кеберле
259
PRONUNCIATION QUALITY ASSESSMENT BY COMPARISON WITH SAMPLE
Dobrovolsky G., Todoriko O., Keberle N.

ВИБІР ЕФЕКТИВНОЇ БАЗОВОЇ ОСНОВИ МОДУЛЯ ПРИ БАГАТОРАЗОВОМУ ПРОРІДЖУВАННІ ПРОБНИХ ЗНАЧЕНЬ В МЕТОДІ ФАКТОРИЗАЦИИ ФЕРМА З НЕРІВНОМІРНИМ КРОКОМ
Є.В. Максименко
270
SELECTION OF EFFECTIVE BASIC BASIS OF MODULE WITH MULTIPLE THINNING TRIAL VALUE IN THE FACTORIZATION FERMAT'S METHOD WITH IRREGULAR PITCH
Maksymenko Ye.

ОПТИМІЗАЦІЯ ПАРАМЕТРІВ В СИСТЕМАХ ОПТИМИЗАЦІЇ З САМООРГАНІЗАЦІЄЮ
О.Д. Франжева
280
OPTIMIZATION OF PARAMETERS IN SELF-ORGANIZATING SYSTEMS
Franzheva E.

ВПЛИВ НЕЛІНІЙНОСТІ НЕЙРОНА НА ЦИКЛІЧНУ СИСТЕМУ УПРАВЛІННЯ
В.Г. Кононович, О.Ю. Козлова, О.Ю. Кунянский
290
INFLUENCE OF NON-LINEARITY OF NEURON ON CYCLIC SYSTEM MANAGEMENTS
Kononovich V., Kozlova O., Kunjanskij O.

РОЗРОБКА ЕФЕКТИВНИХ СТРУКТУР СЛОВНИКА ДЛЯ ЗАДАЧ РОЗПІЗНАВАННЯ МОВЛЕННЯ
Д.В. Заганич, І.Є. Мазурок
297
DEVELOPMENT OF EFFECTIVE VOCABULARY STRUCTURES FOR THE SPEECH RECOGNITION TASKS
Zahanich D., Mazurok I.

ПОРІВНЯННЯ ВЛАСТИВОСТЕЙ НОМІНАЛЬНОГО ТИПУ ОБ'ЄКТІВ РІЗНИХ ПРЕДМЕТНИХ ПІДОБЛАСТЕЙ В РЕЛЯЦІЙНИХ БАЗАХ ДАНИХ
М.Г. Глава
302
COMPARISON OF THE NOMINAL TYPE PROPERTIES OF OBJECTS OF DIFFERENT SUBJECT SUBDOMAINS IN RELATIONAL DATABASES
Glava M.

УДК 004.732.629.5-52

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 205-214

ИНФОРМАТИЗАЦИЯ ПРОГНОЗИРОВАНИЯ РИСКА СТРУКТУРНО СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ С ПОМОЩЬЮ МОДЕЛЕЙ БАЙЕСОВСКИХ СЕТЕЙ ДОВЕРИЯ

В.В. Вычужанин, Н.О. Шибаева

Одесский национальный морской университет,
Мечникова 34, Одесса,65404, Украина, e-mail: vint532@yandex.ru

Предлагается методика оценки риска структурно сложных технических систем, основывающаяся на математическом аппарате динамических байесовских сетей доверия. Проведено исследование разработанных моделей оценки риска функционально-взаимосвязанных и взаимодействующих подсистем и их элементов, входящих в сложные технические системы по текущей информации об их вероятностях отказа. Полученные результаты оценок риска позволяют прогнозировать значения вероятности отказа, риск и поддерживать принятие решений при поиске дефектов в отказавших элементах системы.

Ключевые слова: оценка риска, сложная техническая система, прогнозирование, динамическая байесовская сеть доверия.

Введение

В настоящее время эксплуатация и обслуживание стремительно растущих по размерностям структурно сложных технических систем (СТС) связаны с необходимостью обеспечения более жестких требований по повышению эффективности их использования [1,2]. В этой связи возрастает роль методов, базирующихся на современном программном обеспечении диагностики и прогнозирования СТС. Необходима разработка новых моделей диагностики и прогнозирования СТС в целях оперативного обнаружения нештатных ситуаций, оценивания риска СТС, а также предоставления информации с возможностью ее реализации в пределах ресурса допустимого риска [3].

При эксплуатации СТС не исключены ситуации, связанные с внезапными, непрогнозируемыми отказами функционально-взаимосвязанных и взаимодействующих подсистем и их элементов, входящих в СТС. В целях продления срока эксплуатации СТС необходимо учитывать, что наиболее существенный период времени за пределами гарантийного срока обслуживания характеризуется медленным увеличением вероятности отказа подсистем и их элементов, усилением процессов их старения. При этом необходимо оценивать и прогнозировать фактическое техническое состояние с учетом реальной специфики конкретного типа СТС за эксплуатационный период. Т.е. состояние, характеризующее систему в конкретный момент времени, при неопределенных условиях окружающей внешней и внутренней среды и с учетом регламентируемых значений эксплуатационных параметров СТС [4-6].

Для успешного решения задачи обеспечения надежности и комплексной оценки рисков СТС необходимо снять ряд неопределенностей, каждая из которых является достаточно сложной и значимой. К таким неопределенностям относятся: неполнота

информации о внешних и внутренних воздействиях на систему, о состоянии таких систем; неопределенность в поведении систем. Снятие перечисленных неопределенностей должно основываться на решении таких задач для СТС как оценка риска и его прогнозирование, разработка стратегии периодичности технического обслуживания.

Таким образом, все вышеперечисленные проблемы, связанные с эксплуатацией СТС, а именно прогнозирование вероятности отказа, риска, заставляют искать новые методы решения подобных задач. Такие задачи могут быть успешно решены с использованием математического моделирования, разработкой эффективных методов оценки риска систем и их алгоритмов, реализованных в виде комплексов проблемно-ориентированных программ.

Учитывая специфику СТС, существующие проблемы по обеспечению надежности их эксплуатации задача информатизации прогнозирования оценки риска функционально-взаимосвязанных и взаимодействующих подсистем и их элементов структурно сложных технических систем актуальна.

Цель статьи и постановка задачи исследования

Целью статьи является повышение надежности эксплуатации функционально-взаимосвязанных и взаимодействующих подсистем и их элементов структурно сложных технических систем.

Задача исследования – информатизация и прогнозирование оценки риска функционально-взаимосвязанных и взаимодействующих подсистем и их элементов структурно сложных технических систем.

Основная часть

Для оценки и прогнозирования значения вероятности отказа, риска функционально-взаимосвязанных и взаимодействующих подсистем и их элементов СТС по априорным и апостериорным данным, а также поддержки принятия решений при поиске дефектов в отказавших элементах системы разработана модель, основанная на динамических байесовских сетях доверия (ДБСД) [7.8].

Использование ДБСД позволяет с большой точностью определять элементы СТС наиболее приближенные к критическому состоянию и выходу их из строя. Поставленная задача решается использованием постоянной системы опроса всех элементов системы на различных ее уровнях за конкретный период времени. Это позволяет с помощью ДБСД моделировать аварийные ситуации и точно определять критические значения риска выхода из строя элементов СТС. Появляется возможность устанавливать период времени связанного с разработанным алгоритмом опроса элементов системы, за который проводятся динамические оценки и прогнозирование значения вероятности отказа, риска СТС.

Построение и исследование ДБСД проведено с использованием программного продукта GiNIe [9]. Применение среды GiNIe позволяет:

- осуществить динамическую диагностику каждого элемента ДБСД;
- произвести регрессионный анализ влияния каждого родительского элемента сети на ее соответствующий дочерний элемент;
- осуществить графическое отображение результатов прогностического анализа поведения групп элементов СТС;
- произвести расчеты значения вероятности отказа, риска, а также наступления критического состояния для функционально-взаимосвязанных и взаимодействующих подсистем и их элементов структурно сложных технических систем.

В качестве примера при разработке и исследовании моделей оценки риска по текущей информации о вероятностях отказа функционально-взаимосвязанных и взаимодействующих подсистем и их элементов, входящих в СТС, выбрана судовая энергетическая установка (СЭУ). Структура ДБСД СЭУ представлена на рис. 1. Используемые условные обозначения на рис. 1 приведены в табл. 1.

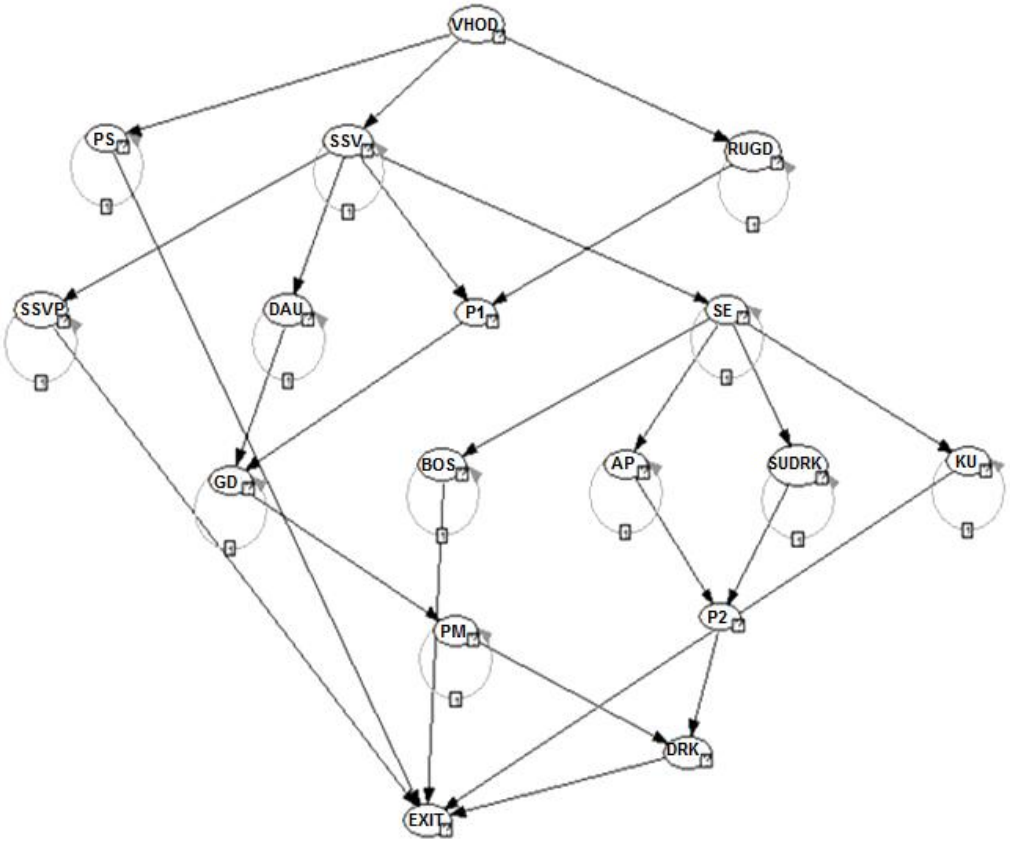


Рис. 1. Структура ДБСД СЭУ

Таблица 1.

Условные обозначения элементов СЭУ в ДБСД

Наименование элемента	Условное обозначение
Входной элемент	ВХОД, VHOD
Ручное управление главным двигателем	РУГД, RUGD
Система сжатого воздуха	ССВ, SSV
Система управления движетельно-рулевым комплексом (ДРК)	СУДРК, SUDRK
Котельная установка	КУ, KU
Судовая электростанция	СЭ, SE
Противопожарная система	ПС, PS
Главный двигатель	ГД, GD
Система дистанционного автоматизированного управления (ДАУ) главного двигателя	ДАУ, DAU
Балластно-осушительная система	БОС, BOS
Передача мощности от главного двигателя к движетелю	ПМ, PM
Аварийный привод ДРК	АП, AP
Движительно-рулевой комплекс	ДРК, DRK
Система санитарной водоподготовки	ССВП, SSVP

Структура ДБСД СЭУ – это многоуровневая система расположения элементов, состоящая из 14 элементов системы, 7 уровней с добавлением специализированных промежуточных узлов Р1 и Р2, обеспечивающих реализацию многоуровневой структуры сети. Для элементов структуры ДБСД СЭУ верхнего уровня задаются условные вероятности отказа с учетом влияния элементов системы более низкого иерархического уровня на элементы более высокого иерархического уровня.

Целевая функция оценки работоспособности элементов СЭУ посредством ДБСД имеет вид

$$F(P_b) = \{G, M\},$$

где G – ациклический направленный граф сети; M – множество элементов СЭУ, составляющих сеть.

Вершинами графа являются функционально-взаимосвязанные и взаимодействующие элементы СЭУ, которые, с учетом иерархии сети и в соответствии с [7,10], определяются

$$v = \left\{ v_i^j \middle| \overline{1,n}, j = \overline{1,m} \right\},$$

где v – наименование элемента СЭУ;

i – номер блока в сети;

n – число блоков в сети;

j – номер уровня в сети;

m – число уровней в сети.

Вероятности отказа при прогнозировании надежности (работоспособности) элементов СЭУ изменяются в соответствии с экспоненциальным законом распределения и использованием логистической регрессии, способствующей предсказанию вероятности возникновения некоторого события по значениям множества признаков. Для этого вводится зависимая переменная, принимающая лишь одно из двух значений. Как правило, это числа 0 (событие не произошло) и 1 (событие произошло), и множество независимых переменных (также называемых признаками, предикторами или регрессорами), на основе значений которых вычисляется вероятность принятия того или иного значения зависимой переменной.

При моделировании ДБСД СЭУ (рис. 2) для различных значений вероятности (риска) отказа входного элемента определены значения вероятности (риска) отказа функционально-взаимосвязанных и взаимодействующих подсистем и их элементов за 20000 часов эксплуатации СЭУ (рис. 3, 4). Из результатов проведенных исследований следует, что при росте риска отказа входного элемента от 0.09 к 0.2 растут значения рисков отказа всех дочерних, нижестоящих элементов ДБСД СЭУ в соответствии с данными, приведенными в табл. 2.

Из полученных результатов оценок вероятностей и рисков отказов функционально-взаимосвязанных и взаимодействующих подсистем и их элементов СЭУ определено время их отказа (табл. 3).

Целевым назначением применения ДБСД при оценках, как вероятности отказа, так и риска отказа является апостериорный вывод. Он заключается в том, что при поступлении новой информации об отказах приравниваются к нулю несовместимые со свидетельством априорные вероятности (риски) отказа.

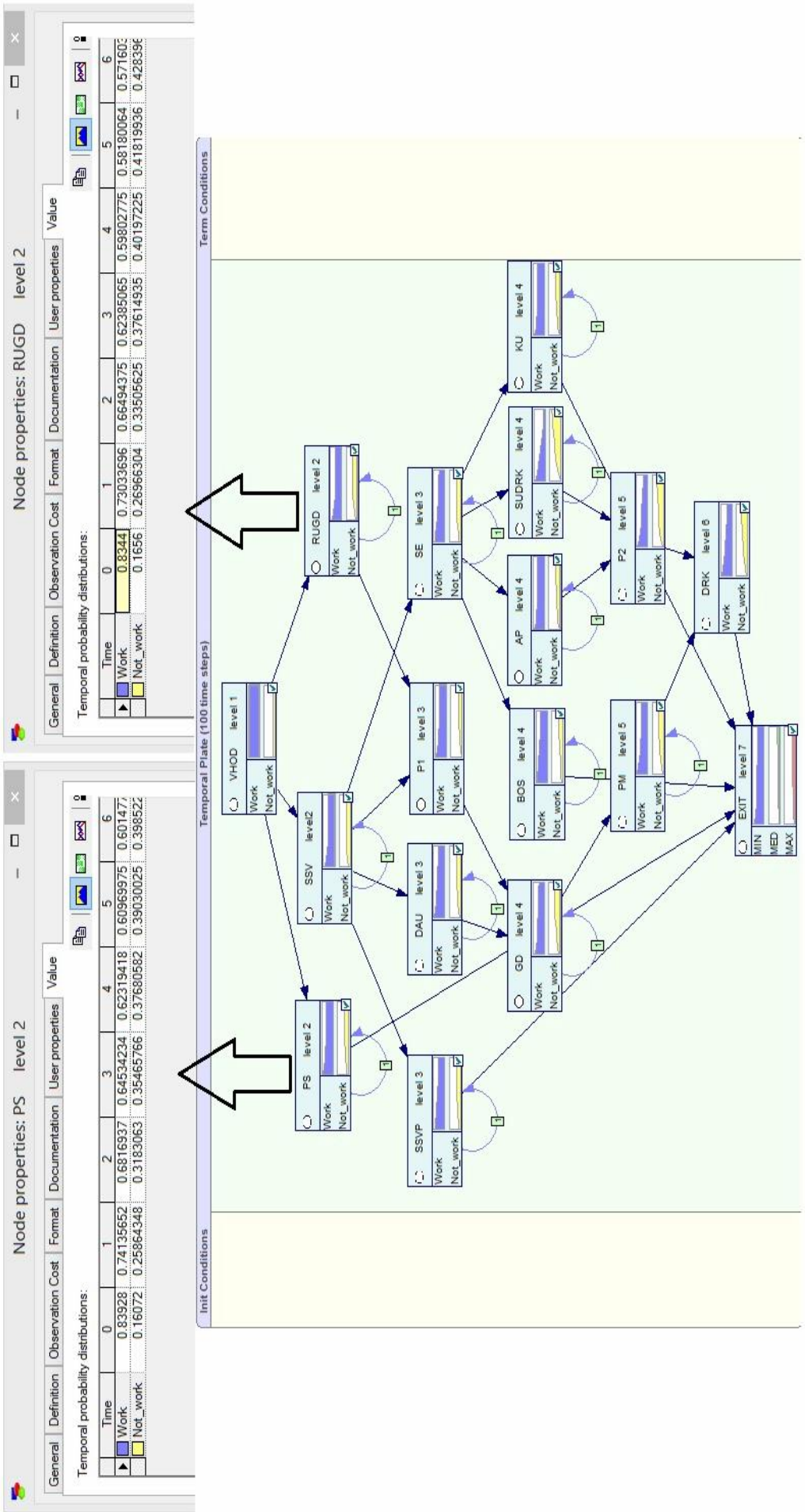


Рис. 2. Динамическая байесовская сеть доверия СЭУ в среде GeNIe при риске отказа входного элемента 0.2

Таблица 2.

Значения рисков отказа дочерних элементов ДБСД СЭУ

№	Наименование элемента	Риск отказа на входном элементе 0.09	Риск отказа на входном элементе 0.2
1	ПС	0.07	0.16
2	ССВ	0.08	0.17
3	РУГД	0.07	0.16
4	ССВ	0.07	0.14
5	ДАУ	0.07	0.13
6	СЭ	0.08	0.15
7	ГД	0.08	0.14
8	БОС	0.07	0.12
9	АП	0.07	0.12
10	СУДРК	0.07	0.13
11	КУ	0.07	0.13
12	ПМ	0.06	0.1
13	ДРК	0.08	0.13

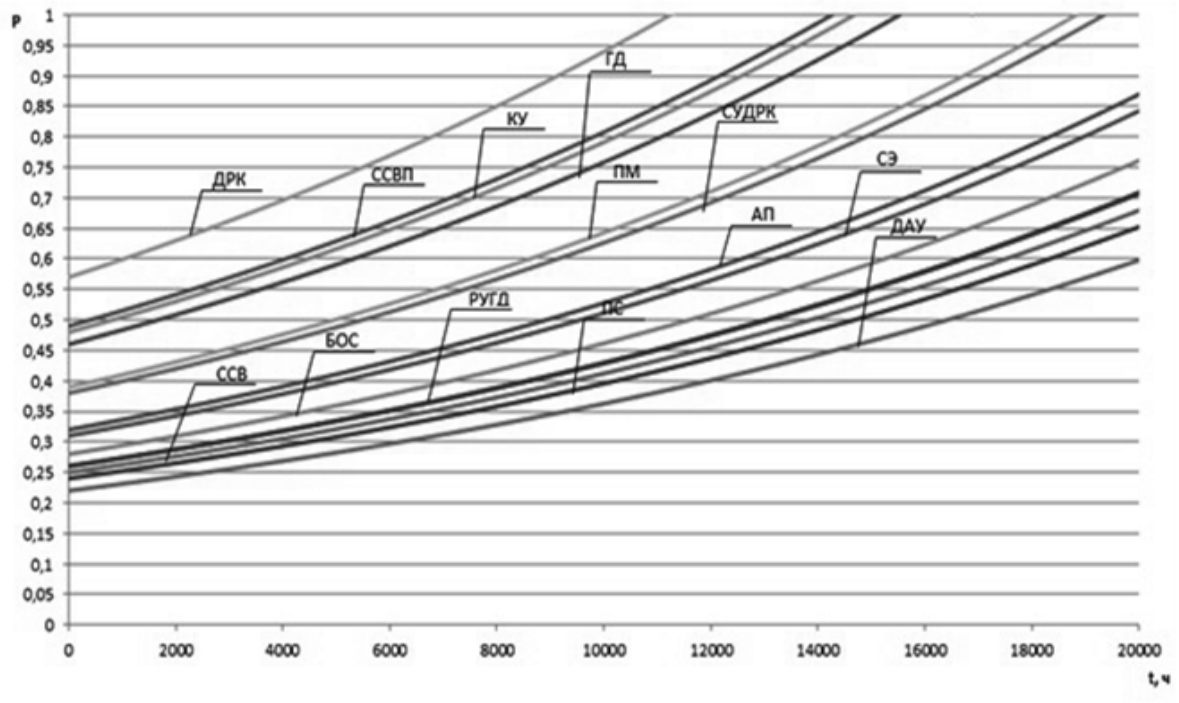


Рис. 3. Вероятность отказа элементов СЭУ при вероятности отказа на входе системы 0.26

Таблица 3.

Время отказа элементов СЭУ

Элемент отказа	Время отказа, ч
ДРК	10898
ССВП	14270
КУ	14683
ГД	15534
СУДРК	19355

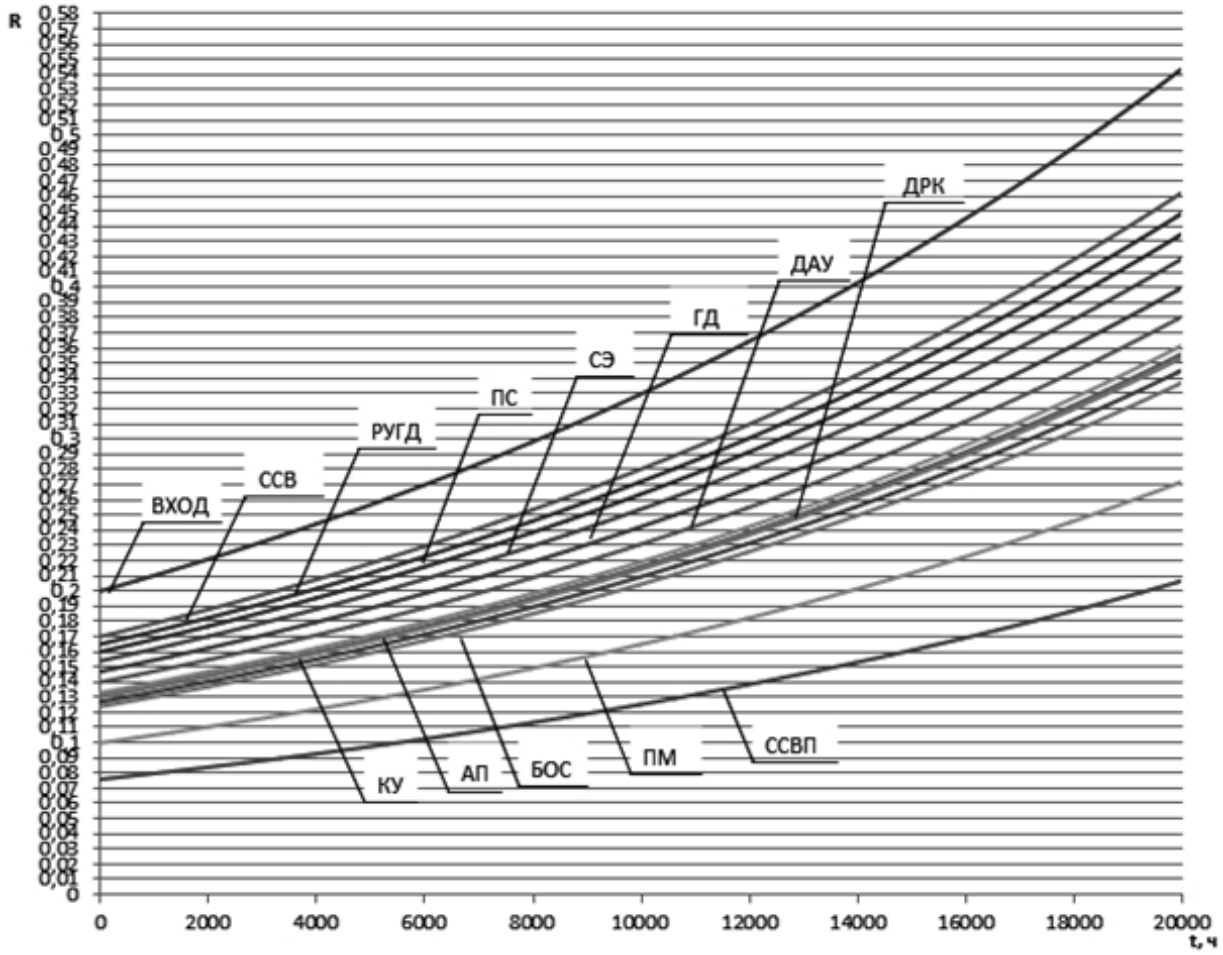


Рис. 4. Риск отказа элементов СЭУ при риске отказа на входе системы 0.2

Затем нормируются оставшиеся вероятности (риски) отказа с учетом априорных вероятностей (рисков) для того, чтобы их сумма равнялась единице [7]. Т.е. априорные данные динамически пересчитываются и формируют апостериорную оценку вероятности (риска), которая является априорной информацией, для обработки новой информации. Апостериорный вывод основан на процедурах анализа данных, получаемых вследствие использования ДБСД. При реализации такого подхода в исследованиях, моделированием по априорным и апостериорным данным определены подсистемы СЭУ, наибольшим образом влияющие на работоспособность всей системы за различные промежутки времени. Установлено, что к таким подсистемам относятся: ССВ, СУДРК, КУ, СЭ, ГД, ПМ, ДРК, для которых некоторые результаты исследований приведены на рис. 5, 6.

Из проведенных исследований следует, что ДРК, зависящий от СЭ, имеет наибольший риск выхода из строя и может перейти в статус критически уязвимого элемента. Если ДРК проверен через 100 часов работы и имеет риск отказа 0.08845, то работоспособность подсистемы не нарушена, а значит, работоспособна вся СТС.

Использование результатов исследований разработанных моделей в целях ретроспективного анализа аварийных ситуаций на СТС позволяет решить задачу определения их причин. Особую актуальность это принимает, когда анализ аварии, выявление ее первопричины и принятие соответствующих мер позволяет исключить или уменьшить вероятность повторения неблагоприятных событий, а значит выполнить поставленную задачу, повысить надежность эксплуатации функционально-

взаимосвязанных и взаимодействующих подсистем и их элементов структурно сложных технических систем.

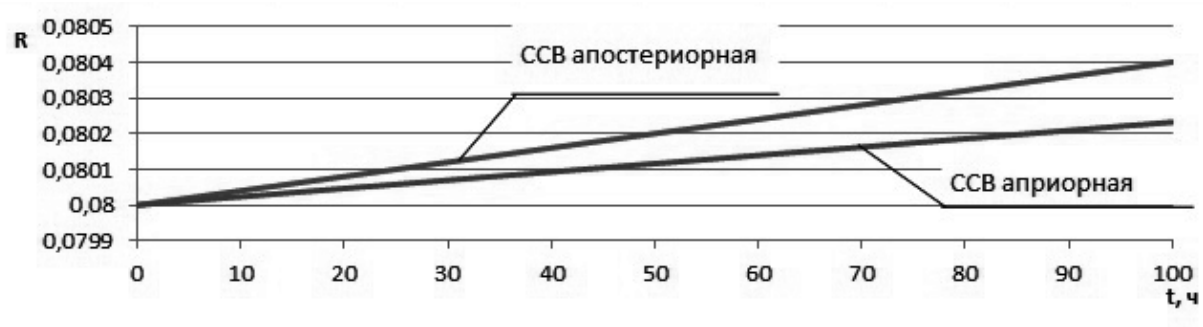


Рис. 5. Априорная и апостериорная оценки риска отказа ССВ СЭУ при поступлении информации об отказах

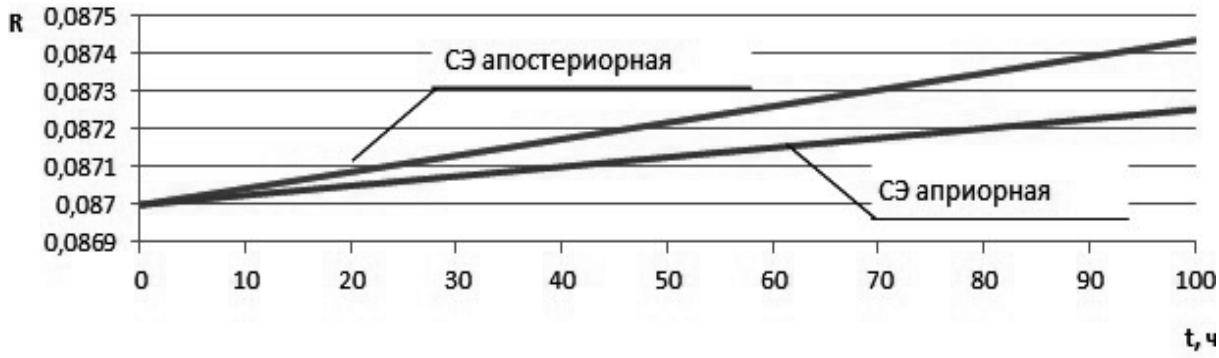


Рис. 6. Априорная и апостериорная оценки риска отказа СЭ СЭУ при поступлении информации об отказах

Выводы

Впервые разработаны методологические основы информационного обеспечения прогнозирования технических рисков отказа судовых структурно сложных технических систем, базирующиеся на моделях байесовских сетей доверия с учетом взаимосвязанности и взаимодействия подсистем и их элементов.

Использование разработанной методики позволяет прогнозировать тенденции изменения вероятности и риска отказов СТС во времени с учетом изменения вероятности и риска отказов отдельных ее элементов.

Предлагаемая методика оценки риска отказа судовых СТС позволяет определить значимость действующих в системе взаимосвязей и взаимодействий элементов систем.

Полученные результаты оценок вероятности и риск отказа позволяют прогнозировать значения вероятности и риск отказа, поддерживать принятие решений при поиске дефектов в отказавших элементах системы.

Список литературы

1. Дорожко, И.В. Методика синтеза оптимальных стратегий диагностирования автоматизированных систем управления сложными техническими объектами с использованием априорной информации / И.В. Дорожко, Н.А. Осипов // Тр. СПИИРАН, 2012. – Вып. 1 (20). – С. 165-185.
2. Половко, А.М. Основы теории надежности / А.М. Половко, С.В. Гуров. – СПб.: БХВ-Петербург, 2006. – 704 с.
3. Панкратова, Н. Д. Системный анализ в динамике диагностирования сложных технических систем / Н.Д. Панкратова // Системні досліжениня та інформаційні технології. – 2008. – № 1. – С. 33–49.
4. Вычужанин, В.В. Информатизация дистанционного диагностирования состояния сложных технических систем / В.В. Вычужанин, С.Н. Коновалов // Информатика и математические методы в моделировании. – 2016. – Том 6, №1. – С. 303-311.
5. Вычужанин, В.В. Технические риски сложных комплексов функционально взаимосвязанных структурных компонентов судовых энергетических установок / В.В. Вычужанин, Н.Д. Рудниченко // Вісник Одеського національного морського університету, збірник наукових праць. – 2014. – Випуск 2(40) . – С. 68 –77.
6. Вычужанин, В.В. Проблемы дистанционного мониторинга, диагностики и прогнозирования состояния судовых технических систем / В.В. Вычужанин, С.Н. Коновалов, Н.О. Шибаева // Вісник Одеського національного морського університету, збірник наукових праць. – 2015. – Випуск 2(44). – С. 97 –107.
7. Тулупьев, А.Л. Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах / А.Л. Тулупьев, А.В. Сироткин, С.И. Николенко. – СПб.: Изд-во С.-Петерб. ун-та, 2009. – 400 с.
8. Jensen, F.V. Bayesian Networks and Decision Graphs / F.V. Jensen, T.D. Nielsen. - Berlin, Springer, 2007. - 457 p.
9. Дорожко, И.В. Оценка надежности структурно сложных технических комплексов с помощью моделей байесовских сетей доверия в среде GeNIe / И.В. Дорожко, А.Г. Тарасов, А.М. Барановский // Intellectual Technologies on Transport. – 2015. – № 3. – С. 36-45.
10. Бидюк, П.И. Построение и методы обучения байесовских сетей / П.И. Бидюк, А.Н. Терентьев // «Таврійський вісник інформатики і математики». – 2004. - №2. - С. 139-154.

ІНФОРМАТИЗАЦІЯ ПРОГНОЗУВАННЯ РИЗИКУ СТРУКТУРНО СКЛАДНИХ ТЕХНІЧНИХ СИСТЕМ ЗА ДОПОМОГОЮ МОДЕЛЕЙ БАЙЕСОВСКИХ МЕРЕЖ ДОВІРИ

В.В. Вичужанін, Н.О. Шибаєва

Одеський національний морський університет,
Мечникова 34, Одеса, 65404, Україна, e-mail: vint532@yandex.ru

Пропонується методика оцінки ризику структурно складних технічних систем, яка ґрунтується на математичному апараті динамічних байесовских мереж довіри. Проведено дослідження розроблених моделей оцінки ризику функціонально-взаємопов'язаних і взаємодіючих підсистем і їх елементів, що входять в складні технічні системи за поточною інформацією про їх ймовірності відмови. Отримані результати оцінок ризику дозволяють прогнозувати значення ймовірності відмови, ризик і підтримувати прийняття рішень при пошуку дефектів в відмовили елементах системи.
Ключові слова: оцінка ризику, складна технічна система, прогнозування, динамічна баєсова мережа довіри.

INFORMATIZATION OF PROGNOSTICATION OF RISK STRUCTURALLY OF THE DIFFICULT TECHNICAL SYSTEMS BY MEANS OF MODELS OF THE BAYES NETWORKS OF TRUST

V.V. Vychuzhanin, N.O. Shibaeva

Odessa National Maritime University,
Mechnikov 34, Odessa, 65404, Ukraine, e-mail: vint532@yandex.ru

Methodology of estimation of risk structurally of the difficult technical systems, being base on mathematical vehicle of the dynamic Bayes networks of trust, is expounded in the article. A study of the worked out models of estimation of risk of the functionally-associate and interactive subsystems and their elements, included in the difficult technical systems on current information about their probabilities of refuse, is undertaken. The got results of risk estimations allow to forecast the values of probability of refuse, risk and to support making decision at the search of defects in the saying no elements of the system.
Keywords: risk estimation, difficult technical system, prognostication, dynamic Bayes network of trust.

УДК 004.056:061.68; 004.3.75:061.68

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 215-226

ОПТИМАЛЬНІСТЬ НЕУСІЧЕНОЇ ПОСЛІДОВНОЇ ПРОЦЕДУРИ ВАЛЬДА В ЗАДАЧАХ ПЕРЕВІРКИ ДВОХ ПРОСТИХ ПРОГНОЗІВ НЕСАНКЦІОНОВАНОГО ДОСТУПУ В ІНФОРМАЦІЙНИХ МЕРЕЖАХ ДЕРЖАВИ

В.Б. Дудикевич[1], І.Р. Опірський[1], П.І. Гаранюк[1], О.А. Вавpічен[2]

[1] Національний університет «Львівська політехніка»,
вул. Ст.Бандери,12, Львів,79000,Україна; e-mail: iopirsky@gmail.com
[2] Національна академія Державної прикордонної служби України ім. Б.Хмельницького,
вул. Шевченка, 46, м. Хмельницький, 29003, Україна

В роботі математично обґрунтовано та виведено вирази, що точно описують послідовне правило Вальда, і в рамках зроблених припущень та досліджень виведено вирази оптимальної перевірки двох простих прогнозів, що в свою чергу дозволяє отримати подальший розвиток вирішенню проблеми прогнозування несанкціонованого доступу. На основі дослідження оптимальності не усіченого послідовного правила Вальда отримано рівності, що зв'язують ймовірності помилок з заданими порогами і пороги з заданими ймовірностями помилок які справедливі для загального випадку залежних неоднорідних спостережень і для незалежних однорідних спостережень.

Ключові слова: несанкціонований доступ, інформаційні системи держави, процедура Вальда, прогноз, прогнозування, поріг, оптимальне послідовне правило, спостереження

Вступ

Створення прогнозів – дуже специфічний процес, тому їх перевірка представляє собою важкий і потрібний етап прогнозування. В загальному випадку знайти конструктивне рішення не завжди вдається навіть в двох альтернативних прогнозах. Тому основну увагу приділимо випадкам, коли побудова конструктивних рішень можлива. Розглянемо не тільки двох альтернативні, але і багато альтернативні процедури. В багатьох практичних задачах закони розподілу ймовірностей досліджуваних випадкових процесів і не спостережуваних ситуацій відомі не повністю, а в кращому випадку – з точністю до сукупності заважаючи параметрів. Задачі прогнозування несанкціонованого доступу (НСД) тісно пов'язані з задачами контролю поточного стану інформаційних мереж держави (ІМД), як вирішуються автоматично і характеризуються по-перше, великим об'ємом оброблюваної інформації і, по-друге, значною одноманітністю порівняно невеликого числа основних операцій. Як показує досвід, програми контролю і прогнозування мають в окремих випадках тисячі команд, а їх виконання зводиться до багаторазових звернень до пам'яті і підпрограмою контролю і прогнозу. Перечисленні особливості дають вирішальний вплив на питання раціонального розміщення вихідних даних, програм контролю та прогнозу в пам'яті ЕОМ і на організацію структури програм.

Таким чином, проблема визначення основних алгоритмів при перевірці прогнозів НСД в ІМД, а також розвитку теорії прогнозування НСД в ІМД на базі математичного

апарату теорії ймовірностей є актуальним і потребує детального і подальшого наукового дослідження. В нашій роботі ми продовжуємо поглиблюватись у проблему прогнозування НСД в ІМД, використовуючи, конкретно в цій статті, сучасний математичний апарат теорії ймовірності, а саме використовуючи процедури Вальда.

Метою даної роботи є дослідження та аналіз використання двопорогової послідовної неусіченої процедури Вальда для перевірки двох простих прогнозів несанкціонованого доступу до інформаційних мереж держави.

Основна частина

Розглянемо не усічену послідовну перевірку двох альтернативних прогнозів при незалежних, можливо, неоднорідних спостереженнях при дотримані умови $R_n^N(T_n) = \min\left\{R_n^0(T_n), R_{nП}^V(T_n)\right\}, n = \overline{1, N-1},$ для виконання якого в даному випадку достатньо виконати умову $g_{ij}(n)P_i(\tau^0 > n) \to 0, n \to \infty,\ N \to \infty$, відповідно з теоремою 1 [1] оптимальне не усічене правило може бути отримано з усіченого шляхом граничного переходу $N \to \infty$. Тому оптимальна неусічена процедура має вигляд:

$$u_n^0(\Lambda_n) = \begin{cases} 1,\ \Lambda_n \geq B_n, \\ 0, \Lambda_n \leq A_n, \\ u_П, \Lambda_n \in (A_n, B_n), n \geq 1, \end{cases} \tag{1}$$

де пороги A_n, B_n знаходяться з рівнянь $G_{n0}(\Lambda_n) = G_{nП}(\Lambda_n), G_{n1}(\Lambda_n) = G_{nП}(\Lambda_n),$ в яких $G_{nП}(\Lambda_n) = \lim\limits_{N\to\infty} G^N{}_{nП}(\Lambda_n)$.

Таким чином, при довільній залежності втрат від номера кроку спостереження і неоднаково розподілених спостережень оптимальна не усічена процедура перевірки двох простих прогнозів полягає в порівнянні відношення правдоподібності (ВП) з двома змінними (що залежать від n) порогами.

Припустимо тепер, що спостереження однорідні $(p_{in}(x_n) = p_i(x_n)),$ а функція втрат має вигляд:

$$g_{ij}(n) = \varphi_{ij} + c_{ij}n, i, j = 0,1, \tag{2}$$

де c_i - вартість затримки у винесенні рішення на один крок при $\theta = i$; φ_{ij} - втрати при прийнятті j-го рішення в i-й ситуації $(\theta = i)$, що не залежить від n.

Для виконання умови $\rho^0 \equiv \rho(\delta_0) = \lim\limits_{N\to\infty} \rho_N(u_0^N),$ достатньо виконання умови $nP_i(\tau^0 > n) \to 0, n \to \infty, i = 0,1.$ Останні виконуються у випадку кінцевого середнього ризику (СР) $\rho^0 = \rho(\tau^0).$

Припустимо, що оптимальні пороги не залежать від n $(A_n = A, B_n = B).$ Тоді, підставляючи (2) в $G_{nj}(\Lambda_n) = \chi\Lambda_n g_{1j}(n) + g_{0j}(n)$ і

$$G_{nП}^N(\Lambda_n) = \sum_{\nu=1}^{N-n}\sum_{j=0}^{1}\{\chi\Lambda_n g_{1j}(n+\nu)P_{1j}^{(\nu)}(\Lambda_n, n, N) + g_{0j}(n+\nu)P_{0j}^{(\nu)}(\Lambda_n, n, N)\}, n = \overline{1, N-1};$$

отримаємо:

$$G_{n\Pi}(\Lambda_n) = \widetilde{G}_{\Pi}(\Lambda_n) + n(c_0 + c_1\chi\Lambda_n);$$
$$G_{nj}(\Lambda_n) = \widetilde{G}_{j}(\Lambda_n) + n(c_0 + c_1\chi\Lambda_n),$$

де

$$\widetilde{G}_j(\Lambda_n) = \chi\Lambda_n\varphi_{1j} + \varphi_{0j};$$

$$\widetilde{G}_{\Pi}(\Lambda_n) = \sum_{\nu=1}^{\infty}\sum_{j=0}^{1}\ [\ \chi\Lambda_n(\varphi_{1j} + c_1\nu)P_{1j}^{(\nu)}(\Lambda_n) + (\varphi_{0j} + c_0\nu)P_{0j}^{(\nu)}(\Lambda_n)\ ],n = 1,2,...; \tag{3}$$

$$P_{ij}^{(\nu)}(\Lambda_n) = \int\limits_{\{X_{n+1}^{\Pi}\}} P_{ij}^{(\nu-1)}(\Lambda_{n+1})p_i(x_{n+1})dx_{n+1}, \nu \geq 2;\ \ P_{ij}^{(1)}(\Lambda_n) = \int\limits_{\{X_{n+1}^{j}\}} p_i(x_{n+1})dx_{n+1};$$

$$X_{n+1}^{\Pi} = \{x_{n+1} : \Lambda`(x_{n+1}) \in (A/\Lambda_n, B/\Lambda_n)\};$$

$$X_{n+1}^{0} = \{x_{n+1} : \Lambda`(x_{n+1}) \leq (A/\Lambda_n)\};$$

$$X_{n+1}^{1} = \{x_{n+1} : \Lambda`(x_{n+1}) \geq B/\Lambda_n\};$$

не залежать від n. Оптимальні пороги A і B при цьому знаходяться з рівнянь

$$G_j(\Lambda) = \widetilde{G}_{\Pi}(\Lambda), j = 0,1, \tag{4}$$

і виявляються постійними. Відповідно, при лінійній залежності втрат від номера кроку спостереження і однаково розподілених спостережень оптимальна послідовна процедура базується на порівнянні ВП з двома постійними порогами. Вперше ця процедура була запропонована і досліджена А. Вальдом [2]. Тому в подальшому будемо називати двопорогову послідовну неусічену процедуру

$$u_n^*(\Lambda_n) = \begin{cases} 1, \Lambda_n \geq B, \\ 0, \Lambda_n \leq A, \\ u_{\Pi}, \Lambda_n \in (A,B), n \geq 1, \end{cases} \tag{5}$$

процедурою Вальда.

Пороги A, B, що входять в (5), залежать від відношень між коефіцієнтами $\varphi_{ij}, c_i, i, j = 0,1,$ апріорної ймовірності P_1 і степені різноманітності прогнозів (щільностей $p_1(x), p_0(x)$). Знайдені пороги з рівняння (4) в реальному вигляді представляють собою проблему.

Вважаючи в (4) $P_{ij}^{(\nu)}(\Lambda_n) = 0, \quad \nu \geq 2;\ \ P_{ij}^{(1)}(\Lambda_n) = 0, \quad i \neq j;\ \ P_{11}^{(1)}(\Lambda_n) = P_{00}^{(1)}(\Lambda_n) = 1,$ $\Lambda_n \in (A,B),$ отримаємо лінійну апроксимацію для функції $\widetilde{G}_{\Pi}(\Lambda_n)$. Очевидно, що відповідні цій апроксимації пороги (див. (4))

$$\underline{A} = \frac{c_0}{\chi(\varphi_{10} - \varphi_{11} - c_1)},\ \overline{B} = \frac{\varphi_{01} - \varphi_{00} - c_0}{\chi^{c_1}}$$

є нижньою та верхньою границями відповідно для оптимальних значень A, B, при чому A і B тим ближчі до $\underline{A}$ і $\overline{B}$, чим більше відрізняються розподілення $p_1(x), p_0(x)$.

Розглянемо тепер умовно екстремальну задачу, в якій необхідно знайти оптимальне послідовне правило $u_0(x)$, що мінімізує середній час спостереження $\overline{\tau}(u(x)) = M\tau(u) = P_1\overline{\tau_1}(u) + (1-P)\overline{\tau}_0(u)(\overline{\tau_1} = M_i\tau)$ в класі послідовних правил, що мають кінцеві середні довжини $\overline{\tau}_i < \infty, i = 0,1,$ і задовольняють обмеження

$$\alpha_i(u(x)) \le \overline{\alpha}_i, i = 0,1, \tag{6}$$

де $\alpha_i(u) = P_i(u_\tau(x^{\tau_1}) = j),\ i \ne j = 0,1,$ – ймовірні помилки рішень; $\overline{\alpha}_i$ – задані ймовірності помилкових рішень. Позначимо клас таких правил через $\Delta(\overline{\alpha}_0, \overline{\alpha}_1)$.

Припустимо, що існує функція $u(x) \in \Delta(\overline{\alpha}_0, \overline{\alpha}_1)$, така, що $\alpha_i u(x) < \overline{\alpha}_1, i = 0,1$ (умова Слейтера). Тоді умовно екстремальна задача може бути зведена до безумовної [1]: $L(\lambda_0, \lambda_1, u_0(x)) = \min\limits_{u(x)} L(\lambda_0, \lambda_1, u_0(x)),$ де λ_0, λ_1 - невизначені множники;

$$L(\lambda_0, \lambda_1, u_0(x)) = \overline{\tau}(u(x)) + \sum_{i=0}^{1} \lambda_1(\alpha_i(u(x)) - \overline{\alpha}_i).$$

Допустимо також, що $\alpha_i(u_0(x)) = \overline{\alpha}_i, i = 0,1.$ Тоді $\lambda_0 > 0, \lambda_1 > 0.$ Слідуючи міркуваннями [1], отримаємо, що апостеріорний ризик (АР)

$$R_{nj}(\Lambda_n) = (1 + \chi\Lambda_n)^{-1}\{\widetilde{\varphi}_{0j} - \overline{\alpha}_0 + [(1-P_1)/\lambda_0 + \chi\Lambda_n P_1/\lambda_1]n + \chi\Lambda_n(\widetilde{\varphi}_{1j} - \overline{\alpha}_1)\}, \tag{7}$$

де $\widetilde{\varphi}_{ij} = \begin{cases} 1, i \ne j, \\ 0, i = j. \end{cases}$

Звідси витікає, що розглянута умовно екстремальна задача еквівалентна байєсівській при функції втрат $g_{ij}(n) = \varphi_{ij} + c_i n,\ i, j = 0,1,$ де $\varphi_{ij} = \begin{cases} 1 - \overline{\alpha}_i, i \ne j, \\ -\overline{\alpha}_i, i = j. \end{cases}$ $c_i = \pi_{0i}/\lambda_i,\ i = 0,1,$ $(\pi_{00} = 1 - \pi_{01}),$ і, відповідно, при незалежних однорідних спостереженнях оптимальне правило (5) з порогами $A(\overline{\alpha}_0, \overline{\alpha}_1, \pi_{01}),\ B(\overline{\alpha}_0, \overline{\alpha}_1, \pi_{01}),$ що залежать від обмежень і апріорної ймовірності π_{01}.

Однак виявляється, що правило Вальда мінімізується в класі $\Delta(\overline{\alpha}_0, \overline{\alpha}_1)$ не тільки безумовну середню довжину спостереження, але і дві умовні середні довжини $\overline{\tau}_i$. Сам по собі цей факт не є тривіальним, і, очевидно, якщо він має місце, то оптимальні пороги $A = A(\overline{\alpha}_0, \overline{\alpha}_1),\ B = B(\overline{\alpha}_0, \overline{\alpha}_1),$ не залежать від апріорної ймовірності π_{01}. Позначимо через $\Delta^0(\overline{\alpha}_0, \overline{\alpha}_1)$ клас послідовних не усічених правил, для яких в (6) справедливі рівності

$$\Delta^0(\overline{\alpha}_0, \overline{\alpha}_1) = \{u(x) : \alpha_i(u(x)) = \overline{\alpha}_i, \quad \overline{\tau}_i(u(x)) < \infty, i = 0,1\ \},\ \overline{\alpha}_1 + \overline{\alpha}_0 < 1. \tag{8}$$

Припустимо, що для будь-якого значення $\pi_{01} \in (0,1)$ знайдуться $g_1(\pi_{01}) > 0,$ $g_0(\pi_{01}) > 0,$ такі, що байєсівське правило

$u_0(x) = \arg \inf_{u(x)\in\Delta} \sum_{i=0}^{1} \pi_{0i}\left[g_i(\pi_{01})\alpha_i(u(x)) + \bar{\tau}_i(u(x))\right]$, де Δ - клас різних нерандомізованих правил, належить класу $\Delta^0(\bar{\alpha}_0, \bar{\alpha}_1)$. Тоді у відповідності з теоремою 1 [1]. правило $u_0(x)$ є оптимальним в умовно екстремальній задачі:

$$\bar{\tau}_i u_0(x) = \inf_{u(x)\in\Delta(\bar{\alpha}_0,\bar{\alpha}_1)} \bar{\tau}\ (u(x)), i = 0,1. \tag{9}$$

Оскільки, як було показано вище, оптимальним байєсівським правилом є (5), при виконанні вказаного припущення воно оптимальне в сенсі (8), тобто мінімізує дві умовні середні довжини в класі $\Delta^0(\bar{\alpha}_0, \bar{\alpha}_1)$. З існування необхідних констант $g_i(\pi_{01}), i = 0,1$, витікає із леми 6 гл. 3 монографії [3], котра тут не відтворюється.

Отримані результати можна сумувати і вигляді наступної теореми.

Теорема 2. Якщо спостереження $x_1, x_2, ..., x_n, ...$ незалежні і однаково розподілені як при прогнозі H_0 так і при H_1, то: при функції втрат (2) оптимальним байєсівським правилом є правило Вальда $u^*(\Lambda) = \{u^*{}_n(\Lambda_n), n \geq 1\}$, що визначене співвідношенням (6); верхня і нижня границя для порогів визначаються рівністю (7); якщо послідовне правило Вальда $u^*(\Lambda) \in \Delta^0(\bar{\alpha}_0, \bar{\alpha}_1)(\bar{\alpha}_0 + \bar{\alpha}_1 > 1)$, то воно мінімізує умовні середні тривалості спостереження $\bar{\tau}_i, i = 0,1$, в класі (послідовних і непослідовних) правил $\Delta(\bar{\alpha}_0, \bar{\alpha}_1)$, що мають ймовірності помилок не більше заданих $\bar{\alpha}_0, \bar{\alpha}_1$ і кінцеві $\bar{\tau}_i, i = 0,1$.

Відзначимо наступні обставини. Випадок $\tau_i = \infty$ не представляє інтересу. Однак можна показати, що правило (6) оптимальне і в цьому випадку. Якщо правило Вальда не належить класу $\Delta^0(\bar{\alpha}_0, \bar{\alpha}_1)$, але належить $\Delta(\bar{\alpha}_0, \bar{\alpha}_1)$, то, в класі $\Delta(\bar{\alpha}_0, \bar{\alpha}_1)$ може знайтись найкраще (9) правило. Приклад такого роду приведений в [4]. Теорема справедлива і для загального випадку спостереження процесу x_t в дискретному $(t \in \{0,1,2,...\})$ або неперервному $(t \in [0,\infty])$ часі, якщо логарифми ВП (ЛВП) $z_t = \ln \Lambda_t$ має незалежні однорідні прирощення.

В [3] показано, що в задачі перевірки прогнозів про середнє значення вінерівського процесу правило Вальда оптимальне в класі

$$\tilde{\Delta}(\bar{\alpha}_0, \bar{\alpha}_1) = u(x) : \beta(\alpha_0(u), \alpha_1(u)) \geq \beta(\bar{\alpha}_0, \bar{\alpha}_1),$$
$$\beta(\alpha_1(u), \alpha_0(u)) \geq \beta(\bar{\alpha}_1, \bar{\alpha}_0)$$

більш широким, ніж $\Delta(\bar{\alpha}_0, \bar{\alpha}_1)$. Тут

$$\beta(x, y) = (1-x)\ln[(1-x)/y] - x\ln[(1-y)/x], \tag{10}$$

причому $M_1\tau(u^*) = 2\beta(\bar{\alpha}_1, \bar{\alpha}_0), M_0\tau(u^*) = 2\beta(\bar{\alpha}_0, \bar{\alpha}_1)$.

Визначимо зв'язок порогів A, B і середніх тривалостей $\bar{\tau}_i{}^* = M_1\tau^*$ з заданими ймовірностями помилок $\bar{\alpha}_0, \bar{\alpha}_1$, де

$$\tau^* = \inf\{n : z_n \notin (\bar{\alpha}_0, \bar{\alpha}_1)\} = \inf\{n : \Lambda_n \notin (A, B)\}$$

- тривалість Вальда $(z_n \ln \Lambda_n)$. Позначимо через $\delta_1 = z_{\tau^*} - a_1, \delta_0 = z_{\tau^*} - a_0$ перескоки порогів $a_1 \ln B$, $a_0 \ln A$ ЛВП в момент зупинки τ^*. Для ймовірностей помилок правила Вальда маємо

$$\alpha_0^* = \int\limits_{\{u^*_{\tau^*}=1\}} p_0(x_1^{\tau^*})dx_1^{\tau^*} = \int\limits_{\{u^*_{\tau^*}=1\}} e^{-Z_{\tau^*}} p_1(x_1^{\tau^*})dx_1^{\tau^*} = B^{-1} \int\limits_{\{u^*_{\tau^*}=1\}} e^{-\delta_1} p_1(x_1^{\tau^*})dx_1^{\tau^*} = B^{-1}(1-\alpha_1^* - \varepsilon_1)e_1;$$

$$\alpha_0^* = \int\limits_{\{u^*_{\tau^*}=0\}} p_1(x_1^{\tau^*})dx_1^{\tau^*} = \int\limits_{\{u^*_{\tau^*}=0\}} e^{Z_{\tau^*}} p_0(x_1^{\tau^*})dx_1^{\tau^*} = A \int\limits_{\{u^*_{\tau^*}=1\}} e^{\delta_0} p_0(x_1^{\tau^*})dx_1^{\tau^*} = A(1-\alpha_0^* - \varepsilon_0)e_0; \quad (11)$$

де $e_0 = M_0(e^{\delta_0} \big| u^*{}_{\tau^*} = 0)$ і $e_1 = M_1(e^{\delta_1} \big| u^*{}_{\tau^*} = 1)$, $\varepsilon_i = P_i(\tau^* = \infty)$ – ймовірності не закінчення спостережень. Допустимо, що правило Вальда закінчується з ймовірністю $1: P_i(\tau^* < \infty) = 1,\ i = 0,1$ (ці умови слабкіші, чим постульовані в теоремі 2 умови кінцевості $\bar{\tau}^*{}_i$). Тоді $\varepsilon_i = 0$ з (11) отримуємо формули, що зв'язують ймовірності помилок з заданими порогами і пороги з заданими ймовірностями помилок:

$$\alpha_0^*(A,B) = \frac{e_1(1-e_0A)}{B-e_0e_1A};\ \ \alpha_0^*(A,B) = \frac{e_0A(B-e_1)}{B-e_0e_1A}; \quad (12)$$

$$A(\bar{\alpha}_0, \bar{\alpha}_1) = \frac{\bar{\alpha}_1}{(1-\bar{\alpha}_0)e_0},\ \ B(\bar{\alpha}_0, \bar{\alpha}_1) = \frac{(1-\bar{\alpha}_1)e_1}{\bar{\alpha}_0}. \quad (13)$$

Рівність (13) забезпечує строгі рівності $\alpha^*{}_i = \bar{\alpha}_i$ у випадку, коли розподілення z_n не решітчасте. В іншому випадку можуть існувати такі значення $\bar{\alpha}_i$ при яких $\alpha^*{}_i \neq \bar{\alpha}_i$.

Позначимо через $S_\tau = \sum_{n=1}^{\tau} s'_n$ суму незалежних однаково розподілених величин s'_n, де τ – деякий момент зупинки. Якщо $M\big|s'\big| < \infty,\ \bar{\tau} = M_\tau < \infty,$ то справедлива рівність [2,4]:

$$MS_\tau = \bar{\tau}Ms', \quad (14)$$

що носить назву тотожності Вальда. Якщо до того ж $M(s')^2 < \infty,$ то [2,4]

$$M(S_\tau - \bar{\tau}Ms')^2 = \bar{\tau}Ds',$$

де $Ds' = M(s' - Ms')^2$.

Нехай $z' = \ln\big[p_1(x)/p_0(x)\big]$ і $I_1 = M_1 z'(x) \neq 0, I_0 = -M_0 z'(x) \neq 0,$ що справедливо, коли міра точок x, в яких $p_0(x) = p_1(x),$ рівна 0. Величини I_1, I_0 називають інформаційними кількостями (числами) Кульбака-Лейблера [5]. Вони характеризують кількість інформації , що міститься в спостереженнях x на користь прогнозу H_1 проти H_0 і на користь H_0 проти H_1 відповідно. Нехай також

$$M_i\big[\ z'(x)\ \big]^2 < \infty, i = 0,1, \quad (15)$$

що забезпечує кінцеве l_i серед перескоків $M_i\delta_i$ та e_i. Тоді для любого кінцевого моменту τ у відповідності з тотожністю (14) маємо

$$M_i z_\tau = I_1 M_1 \tau,\ M_0 z_\tau = I_0 M_0 \tau. \tag{16}$$

Для моменту τ^*, очевидно,

$$M_i z_{\tau^*} = M_1(a_1 + \delta_1 \big| u^*_{\tau^*} = 1) P_i(u^*_{\tau^*} = 1) + M_1(a_0 + \delta_0 \big| u^*_{\tau^*} = 0) P_i(u^*_{\tau^*} = 0). \tag{17}$$

З (16), (17) отримуємо

$$\bar\tau_1^* = \left[\alpha_1^*(a_0 + \bar\delta_{01}) + (1-\alpha_1^*)(a_1 + \bar\delta_{11})\right]/ I_1;\ \ \bar\tau_1^* = -\left[\alpha_0^*(a_1 + \bar\delta_{10}) + (1-\alpha_0^*)(a_0 + \bar\delta_{00})\right]/ I_0, \tag{18}$$

де $\bar\delta_{ij} = M_j(\delta_i \big| u^*{}_{\tau^*} = i), i, j = 0,1,$ - умовні середні перескоки. Використовуючи (13), з (18) отримуємо зв'язок умовних середніх тривалостей правила Вальда з заданими ймовірностями помилок:

$$\begin{aligned} \bar\tau_1^* &= (\bar\alpha_0, \bar\alpha_1) = \left[\beta((\bar\alpha_1, \bar\alpha_0) + (1-\bar\alpha_1)(\bar\delta_{11} + \ln e_1) + \bar\alpha_1(\bar\delta_{01} - \ln e_0)\right]/ I_1; \\ \bar\tau_0^* &= (\bar\alpha_0, \bar\alpha_1) = \left[\beta((\bar\alpha_0, \bar\alpha_1) - (1-\bar\alpha_0)(\bar\delta_{00} - \ln e_0) - \bar\alpha_0(\bar\delta_{10} + \ln e_1)\right]/ I_0, \end{aligned} \tag{19}$$

де $\beta(x, y)$ - функція, визначена в (10).

Для використання (12), (13), (18), (19) в практичних розрахунках необхідне визначення величин $\delta_{ij}, e_i,$ що в загальному випадку представляє собою важку задачу. Тому доцільно отримати асимптотичні і наближені формули. В якості першого наближення можуть служити формули Вальда [2], що ігнорують ефект перескоку порогів. Припускаючи в (13) і (19)) $e_i = 1, \delta_{ij} = 0,$ отримуємо

$$A(\bar\alpha_0, \bar\alpha_1) \approx \frac{\bar\alpha_1}{1-\bar\alpha_0},\ B(\bar\alpha_0, \bar\alpha_1) \approx \frac{\overline{1-\alpha_1}}{\bar\alpha_0}; \tag{20}$$

$$\bar\tau_1^*(\bar\alpha_0, \bar\alpha_1) \approx \beta(\bar\alpha_1, \bar\alpha_0)/ I_1,\ \bar\tau_0^*(\bar\alpha_0, \bar\alpha_1) \approx \beta(\bar\alpha_0, \bar\alpha_1)/ I_0. \tag{21}$$

Точність формул (20), (21) зростає при наближенні прогнозу (при зменшенні в середньому збільшень ЛВП за крок) і при зменшенні заданих ймовірностей помилок. Справді, якщо $\bar\alpha_i \to 0,$ то при виконанні умови (15) $\bar\delta_{ij} < \infty, e_i < \infty$ і з (14), (19) витікає, що вибір порогів по формулі Вальда (20) забезпечує асимптотичні рівності

$$\alpha_0^* = e_1\bar\alpha_0(1 + O(\bar\alpha_1)),\ \alpha_1^* = e_0\bar\alpha_1(1 + O(\bar\alpha_0)), \tag{22}$$

$$\bar\tau_1^* = \frac{|\ln\bar\alpha_0|}{I_1}(1 + o(1)),\ \ \bar\tau_0^* = \frac{|\ln\bar\alpha_1|}{I_0}(1 + o(1)). \tag{23}$$

Головні члени співвідношення (22) при малих $\bar\alpha_i$ співпадають з правими частинами (21). Якщо, до цього зближаються прогнози, то $e_i \to 1$ і, відповідно, головні члени (22) дорівнюють $\bar\alpha_0$ і $\bar\alpha_1$. Таким чином, точність формул (21) може видатись

задовільними навіть в випадку достатньо інформативних спостережень при малих $\overline{\alpha}_i$, так як $\overline{\delta}_{ij}$ і $\ln e_i$ входять в (13) мультиплікативно.

З сказаного слідує, що формули (20) можуть видатись досить грубими, забезпечуючи ймовірності помилок значно менше аніж необхідно (див. (22)). Їх уточнення, зв'язане з оцінкою e_i, може бути проведено з допомогою теорії відновлення [6]. Результати робіт [3, 6] дозволяють записати наступні асимптотичні точні формули:

$$\begin{aligned} \alpha_0^* &= (A,B) = \frac{\beta_1(1-\beta_0 A)}{B-\beta_0\beta_1 A} + o(\frac{A}{B}); \\ \alpha_1^* &= (A,B) = \frac{\beta_0 A(B-\beta_1)}{B-\beta_0\beta_1 A} + o(\frac{A}{B}), A \to 0, B \to \infty; \end{aligned} \tag{24}$$

$$\begin{aligned} A(\overline{\alpha}_0,\overline{\alpha}_1) &= \frac{\overline{\alpha}_1}{(1-\overline{\alpha}_0)\beta_0}(1+o(\overline{\alpha}_0^{\gamma_0})+o(\overline{\alpha}_1^{1+\gamma_1}))^{-1}; \\ B(\overline{\alpha}_0,\overline{\alpha}_1) &= \frac{\beta_1(1-\overline{\alpha}_1)}{\overline{\alpha}_0}(1+o(\overline{\alpha}_0^{\gamma_1})+o(\overline{\alpha}_1^{1+\gamma_0}))^{-1}, \overline{\alpha}_i \to 0, \end{aligned} \tag{25}$$

де константи β_i, γ_i залежать від розподілу спостереження і визначаються співвідношенням

$$\beta_0 = \frac{\varphi_-(1)}{\varphi'_-(0)};\ \beta_1 = \frac{\varphi_+(0)}{\varphi'_+(1)}, \tag{26}$$

в яких $\varphi\pm(\lambda)$ – компоненти канонічної факторизації [6] функції $\varphi(\lambda) = 1 - M_0 \exp\{\lambda z'(x)\}(\varphi(\lambda) = \varphi+(\lambda)\varphi-(\lambda))$. В деяких випадках в співвідношеннях (24), (25) символ o замінюється на O [3].

З (24) і (25) слідує, що якщо пороги вибрати по формулам

$$B(\overline{\alpha}_0,\overline{\alpha}_1) = \frac{\beta_1(1-\overline{\alpha}_1)}{\overline{\alpha}_0};\ \ A(\overline{\alpha}_0,\overline{\alpha}_1) = \frac{\overline{\alpha}_1}{(1-\overline{\alpha}_0)\beta_1}, \tag{27}$$

то $\alpha^*{}_i \approx \overline{\alpha}_i$, причому точність зростає експоненціально швидко при збільшенні $|\alpha_i|$. В той же час якщо пороги вибрати по формулам Вальда (20), то $\alpha^*{}_i \approx \beta_j\ \overline{\alpha}_i, \overline{\alpha}_i \to 0 (i,j = 0,1, i \neq j)$. Таким чином, Вальдівська апроксимація дає хороші результати тільки в тому випадку, коли мала абсолютна величина перескоку $(\beta_i \approx 1)$, тобто фактично при $\overline{\alpha}_i << 1$. Для середніх тривалостей з (18), (24), (27) при достатньо малих $\overline{\alpha}_i$ витікає наступна апроксимація:

$$\overline{\tau}_1^*(\overline{\alpha}_0,\overline{\alpha}_1) \approx \left[\beta(\overline{\alpha}_1,\overline{\alpha}_0) + \overline{\delta}_1 + \ln\beta_1\right]/I_1;\ \ \overline{\tau}_0^*(\overline{\alpha}_0,\overline{\alpha}_1) \approx \left[\beta(\overline{\alpha}_0,\overline{\alpha}_1) + \overline{\delta}_0 + \ln\beta_0\right]/I_0, \tag{28}$$

де $\overline{\delta}_i = M_i\delta_i$ - середні перескоки $\delta_i + z_{\pi} - a_i$ в односторонніх правилах $\tau_1 = \inf\{n : z_n \geq a_1\};\ \ \tau_0 = \inf\{n : z_n \leq a_0\}$, які можуть бути вичислені з використанням теорії відновлення [6].

Покажемо, що для умовних середніх тривалостей $\overline{\tau}_i(u) = M_{i\tau}(u)$ довільного правила $u(x) \in \Delta(\overline{\alpha}_0,\overline{\alpha}_1)$ при виконанні умови (15) справедливі рівності

$$\bar{\tau}_1(u) \geq \beta(\bar{\alpha}_1, \bar{\alpha}_0)/I_0;\ \bar{\tau}_0(u) \geq \beta(\bar{\alpha}_0, \bar{\alpha}_1)/I_0. \tag{29}$$

Для цього знадобиться наступна лема, доказ якої елементарно можна знайти [5].

Лема. Нехай $(X, \Im)$ - вимірюваний простір, в якому визначені взаємно абсолютно неперервні σ-кінцеві міри P і $\mu, z(x) = \ln\frac{dP}{d\mu}(x)$. Для довільних непересічних множин $Y_i \in F$, $i \geq 0$, таких, що $\bigcup\limits_{i\geq 0} Y_i = X$, має місце нерівність

$$\int\limits_X z(x)d\mu(x) \geq \sum_{i\geq 0} P(Y_i)\ln\frac{P(Y_i)}{\mu(Y_i)}, \tag{30}$$

рівність в якому досягається тільки в тому випадку , коли $P(Y_i) = \mu(Y_i), i \geq 0$.

Припускаючи спочатку $P = P_1, \mu = P_0$, а потім $P = P_0, \mu = P_1$ і $Y_i = \{u_\tau = i\}$, з (30) отримуємо

$$M_1 z_{\tau(u)} \geq \sum_{i=0}^{1} P_1(u_\tau = i)\ln\frac{P_1(u_\tau = i)}{P_0(u_\tau = 0)} = \alpha_1(u)\ln\frac{\alpha_1(u)}{1-\alpha_0(u)} + (1-\alpha(u))\ln\frac{1-\alpha_1(u)}{\alpha_1(u)} =$$

$$= \beta(\alpha_1(u), \alpha_0(u));$$

$$M_0 z_{\tau(u)} \leq -\sum_{i=0}^{1} P_0(u_\tau = i)\ln\frac{P_0(u_\tau = i)}{P_1(u_\tau = i)} = -\left[\alpha_0(u)\ln\frac{\alpha_0(u)}{1-\alpha_1(u)} + (1-\alpha_0(u))\ln\frac{1-\alpha_0(u)}{\alpha_1(u)}\right] =$$

$$= -\beta(\alpha_0(u), \alpha_1(u)).$$

Справедливість невірності (29) слідує тепер з (30), (16) і обставини, що функція $\beta(x, y)$ є спадною в області $x + y \leq 1$.

З теореми 2 оптимальність правила Вальда витікає тільки в тому випадку, коли справедливі точні рівності $a^*_i = \bar{a}_i$. Як вказувалось вище, ці рівності виконуються для не решітчастих розподілень при виборі порогів по формулам (13), в яких точно вичислити величини e_i зазвичай не вдається. З співвідношення (21) витікає, що нижні оцінки (29) для умовних середніх тривалостей наближено досягаються для правил Вальда, коли можна знехтувати перескоками порогів $\delta_{i.}$. Останні, як вже відмічалось, справедливі при розрізненні «близьких» прогнозів і при малих $\bar{a}_i$ (див. також (23)). Таким чином, правило Вальда з порогами (20) в цьому випадку близьке до оптимального, причому степінь наближення тим вище, чим «ближче» щільність $p_0(x), p_1(x)$ і менше $\bar{a}_i$. Найкращі результати, однак, дають формули (27), що включаються при виборі порогів степінь схожості щільностей $p_0(x), p_1(x)$.

Підкреслимо, що рівності (20) стають точними, коли стають відсутні перескоки порогів статистикою ЛВП. Це справедливо, наприклад, коли час спостереження і траєкторії ЛВП неперервні. Якщо ЛВП є гаусовським процесом з незалежними однорідними приростами, то точними є рівності (21). Рівність (12), (13), (20) справедливі для загального випадку залежних неоднорідних спостережень, в той час як (18), (19), (21), (23)–(28) – для незалежних однорідних спостережень.

Розглянемо випадок залежних спостережень $x_1,...,x_n,...,$ таких, що існують взаємно однозначні декорелятивні перетворення $F^{(n)}(x^n{}_1)=\{F_k(x^k{}_1),k=\overline{1,n}\}$, що не залежать від номера прогнозу i. Тоді згідно (18), (19)

$$\Lambda_n(x_1^n)=\prod_{k=1}^{n}\frac{\widetilde{p}_{1k}(\widetilde{x}_k)}{\widetilde{p}_{0k}(\widetilde{x}_k)}=\widetilde{\Lambda}_n(\widetilde{x}_1^n),n\geq 1, \tag{31}$$

де $\bar{x}_k=F_k(x^k{}_1)$; щільності $p_{ik}(\bar{x}_k)$ залежать від k, так як $\bar{x}_k,k\geq 1,$ не однаково розподілені.

Якщо до того ж виконана умова $R_n^0(x_1^n)=R_n^N(T_n),\ n=\overline{1,N}$, то оптимальною (в байєсівському сенсі) здається процедура (1) з змінними порогами. Структурна схема оптимальної системи розпізнання представлена на рис. 1. Для строгої оптимальності правила Вальда (як в байєсівській постановці при втратах виду (1), так і в умовно екстремальній постановці – в класі $\Delta(\overline{\alpha}_0,\overline{\alpha}_1)$) необхідно, щоб ЛВП $\bar{z}_n=\ln\Lambda_n=\sum_1^n\widetilde{z}_k'$ $(\widetilde{z}_k'=\ln[\widetilde{p}_{1k}(\widetilde{x}_k)/\widetilde{p}_{0k}(\widetilde{x}_k)])$ представляв процес з незалежними однорідними приростами. Справді, момент зупинки, що відповідає правилу Вальда, в силу (31) можна записати у вигляді $\tau(u(\Lambda))=\inf\{n\geq 1:\widetilde{z}_n\notin(\ln A,\ln B)\}$.

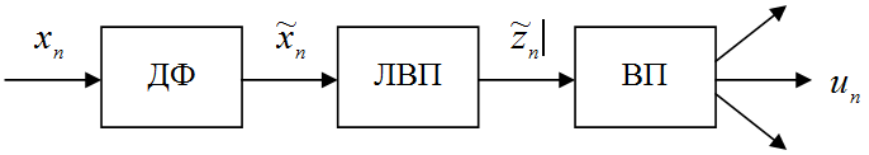


Рис. 1. Структурна схема оптимальної системи розпізнання двох прогнозів при залежних спостереженнях: ДФ – декорелятивний фільтр; ЛВП – блок формування ЛВП; ВП – двохпороговий вирішальний пристрій з змінними порогами.

Якщо величини $\widetilde{z}_k',k\geq 1,$ однаково розподілені (це означає однорідність $\{\widetilde{z}_n,n\geq 1,\}$, застосовується теорема 2, в відповідності до якої правило (5) оптимальне в класі $\Delta(\overline{\alpha}_0,\overline{\alpha}_1)$. В деяких задачах значення $\widetilde{z}_k'$ однаково розподілені при $k\geq M\geq 2$. Тому з збільшенням часу спостереження процес $\{\widetilde{z}_n\}$ стає однорідним і в випадку великої середньої тривалості правило Вальда оптимальне. При цьому справедлива рівність (29), в яких $I_1=M_1\widetilde{z}',I_0=-M_0\widetilde{z}',(\widetilde{z}'=\widetilde{z}_k,k\geq M),$ а також при «близьких» прогнозах і малих $\overline{\alpha}_0,\overline{\alpha}_1$ наближення рівності (21).

Щоб проілюструвати сказане, розглянемо задачу повірки двох прогнозів $H_0:\theta=0,H_1:\theta=1$ про середнє значення марковської послідовності $\{x_n,n\geq 1\}:x_n+S_\theta+\eta_n,\theta=0,1,$ де $S_\theta,$ S_1 - дві константи; $\{\eta_n,n\geq 1\}$ задовольняють співвідношенню $\eta_{n+1}=R\eta_n+\xi_{n+1},n\geq 1,\eta_0=0,$ де $\{\xi_n,n\geq 1\}$ - послідовність незалежних однаково розподілених величин з нульовим середнім, що має розподіл з щільністю $f(\xi)$. Декорельовані перетворення мають вигляд $F_k(x^k{}_1)=x_k-Rx_{k-1},k\geq 2,F_1(x_1)=1;$ при цьому $\widetilde{x}_k=S_\theta(1-R)+\xi_k,k\geq 2;x_1=\widetilde{x}_1.$

Відповідно, $\widetilde{z}_n=\sum_1^m\widetilde{z}_k';z_k'=\ln\frac{f(\widetilde{x}_k-S_1(1-R))}{f(\widetilde{x}_k-S_0(1-R))},k\geq 2;\ z_1'=\ln\frac{f(\widetilde{x}_1-S_1)}{f(\widetilde{x}_1-S_0)}.$

Значення $\tilde{z}'_k, k \geq 2,$ незалежні і однаково розподілені, величина $\tilde{z}'_1$ незалежна від значень $\tilde{z}'_k, k \geq 2,$ і відрізняється від них по розподіленню. Процес $\{\tilde{z}_n, n \geq 1\},$ є марковським неоднорідним, а $\{\varepsilon_n, n \geq 2\}$ – однорідним $\left\{\varepsilon_n, \sum_2^n \tilde{z}'_k\right\}$. Тому, якщо перше спостереження носить новий вклад, що виконано при великому середньому числі кроків спостереження, то правило Вальда близьке до оптимального і можна користуватись всіма отриманими результатами.

Висновки

На основі математичного обґрунтування та виведення виразів, що точно описують послідовне правило Вальда і в рамках зроблених припущень та досліджень виведено вирази оптимальної перевірки двох простих прогнозів, що в свою чергу дозволило отримати подальший розвиток вирішенню проблеми прогнозування НСД. На основі дослідження оптимальності не усіченого послідовного правила Вальда отримано рівності, що зв'язують ймовірності помилок з заданими порогами і пороги з заданими ймовірностями помилок справедливі для загального випадку залежних неоднорідних спостережень і для незалежних однорідних спостережень.

З отриманих співвідношень витікає, що нижні оцінки для умовних середніх тривалостей, наближено досягаються для правил Вальда, коли можна знехтувати перескоками порогів $\delta_{i.}$. Таким чином, отримане і досліджене правило Вальда з порогами в нашому випадку близьке до оптимального, причому степінь наближення тим вище, чим «ближче» щільність $p_0(x), p_1(x)$ і менше $\overline{a}_i$.

При деяких умовах, а саме коли величини $\tilde{z}'_k, k \geq 1,$ однаково розподілені, застосовується запропонована нами теорема 2, у відповідності до якої правило Вальда (двопорогова послідовна не усічена процедура) оптимальне в класі $\Delta(\overline{\alpha}_0, \overline{\alpha}_1)$. В деяких задачах значення $\tilde{z}'_k$ однаково розподілені при $k \geq M \geq 2.$ Тому з збільшенням часу спостереження процес $\{\tilde{z}_n\}$ стає однорідним і у випадку великої середньої тривалості правило Вальда оптимальне. Отже, правило Вальда близьке до оптимального і можна користуватись всіма отриманими результатами.

Список літератури

1. Опірський, І.Р. Оптимізація послідовних процесів прийняття рішень при умовно екстремальній постановці задачі / І.Р. Опірський // СНУ ім. В.Даля: Інформаційна безпека. – 2014. - №4(16). –С.120-127.
2. Орлов, А. И. Теория принятия решений: учебник. — М.: Экзамен, 2006. — 573 с.
3. Вероятность и математическая статистика: Энциклопедия / Под ред. Ю.В.Прохорова. — М.: Большая российская энциклопедия, 2003. — 912 с.
4. Ширяев, А.Н. Вероятность. В 2-х кн. Кн.2 – 3-е изд. – М: МЦНМО, 2004. – 408 с.
5. Кудряшов, Б.Д. Теория информации.– СПб: Питер, 2009.–320 с.
6. Siegmund, D. Sequential Analysis. Tests and confidence intervals.–N.Y.: Springer veriage, 2005.–270 p.

ОПТИМАЛЬНОСТЬ НЕ УСЕЧЕННОЙ ПОСЛЕДОВАТЕЛЬНОЙ ПРОЦЕДУРЫ ВАЛЬДА В ЗАДАЧАХ ПРОВЕРКИ ДВУХ ПРОСТЫХ ПРОГНОЗОВ НСД В ИНФОРМАЦИОННЫХ СЕТЯХ ГОСУДАРСТВА

В.Б. Дудыкевич[1], И.Р. Опирский[1], П.И. Гаранюк[1], О.А.Вавричен[2]

[1]Национальный университет «Львовская политехника»,
ул. Бандеры, 12, Львов, 79000, Украина; e-mail: iopirsky@gmail.com
[2]Национальная академия Государственной пограничной службы Украины им. Б.Хмельницкого,
ул. Шевченко, 46, г. Хмельницкий, 29003, Украина

В работе математически обоснованы и выведены выражения, которые точно описывают последовательное правило Вальда. В рамках сделанных предположений и исследований выведено выражения оптимальной проверки двух простых прогнозов, что в свою очередь позволяет получить дальнейшее развитие решению проблемы прогнозирования НСД. На основе исследования оптимальности не усеченного последовательного правила Вальда получено равенства, связывающие вероятности ошибок с заданными порогами и пороги с заданными вероятностями ошибок которые справедливы для общего случая зависимых неоднородных наблюдений и для независимых однородных наблюдений.

Ключевые слова: несанкционированный доступ, информационные системы государства, процедура Вальда, прогноз, прогнозирование, порог, оптимальное последовательное правило, наблюдения

OPTIMALITY IS NOT TRUNCATED CONSISTENT PROCEDURES WALD IN SCAN TASKS OF TWO SIMPLE BETS UA IN INFORMATION NETWORKS STATE

V.B. Dudykevich[1], I.R. Opirsky[1], P.I. Garanyuk[1], O.A.Vavrichen[2]

[1] National University "Lviv Polytechnic",
Str. Bandera, 12, Lviv, 79000, Ukraine; e-mail: iopirsky@gmail.com
[2] National Academy of State Border Service of Ukraine. B. Khmelnitsky,
Str. Shevchenko, 46, Khmelnytskyi, 29003, Ukraine

The paper is mathematically valid and withdrawn expression that describes sequential rule Wald. As part of the assumptions and research derived the optimal expression of testing two simple prediction, which in turn allows you to further development addressing unauthorized access prediction. Based on a study of optimality is not truncated serial entitled to receive equity Wald linking the probability of errors with the set thresholds and thresholds with given probabilities of error which are valid for the general case of dependent observations of inhomogeneous and homogeneous for independent observations.

Keywords: unauthorized access, information system of the state, the procedure Wald, forecast, forecasting, threshold, optimum sequence typically, observation.

UDC 004.056:619.8:351.007

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 227-239

TRANSFORMATION OF INFORMATION AND SOCIAL-PSYCHOLOGICAL SECURITY PARADIGMS (Part 1)

S. Gnatyuk[1], V. Gnatyuk[1], V. Kononovich[2], I. Kononovich[3]

[1] National Aviation University,
Kosmonavta Komarova Ave, 1, Kyiv, 03680, Ukraine; E-mail: s.gnatyuk@nau.edu.ua
[2] Odessa National Polytechnic University,
Shevchenko Ave, 1, Odessa, 65044, Ukraine; E-mail: vl_kononovich@ukr.net
[3] Odessa National Academy of Food Technologies,
Kanatna Str, 112, Odessa, 65039, Ukraine; E-mail: kononovich@mail.ru

The paper presents the results of a retrospective analysis of transitional paradigm of information security – from the data and information security to the minds and behavior security. Formulated modern paradigms of information security. Described new problems in critical information infrastructures security. Based on the analysis offered partial solutions to a number of partial problems. The implementation of the identity determination and management system will allow to lock the fullness of security mechanisms. Same way is achieved the ability to track each transaction online. Information and psychological security from the destructive impact of information requires the use of social psychological methods. Offered the simple model of forming group consciousness around the idea of cybercrime fighting. The received systematization and solving problems results allows to increase the work efficiency of information, cyber social and psychological security systems and formalize directions for further researches in developing effective security systems.

Keywords: information security, cybersecurity, information & communication system, individual & group mind, social-psychological security, legal framework, paradigm.

Introduction

The evolutionary process of development of methods and technologies of information security characterized by drama and high rates of acceleration. According to the International Telecommunication Union the growth of cybercrime has exponential nature [1; 2]. Number of information security incidents is growing exponentially. Volatility of terminology also shows that problems with information security are far from complete. The problem of this study constitute systematization of representations in content and directions of transformation paradigms of information security, the direction of the transformation of many paradigms of security in the field of information security.

Review of the problems and achievements needs analysis of large volume publications. Three stages in the development of security of state secrets systems of independent Ukraine considered in [3]. Judging by the materials of the foreign press the problems of information security in data processing systems were a big surprise. Since then (from 1960), when these problems were considered as separate, approaches to their solution also went through the initial three stages [4]. Known as periodization phases of information security of Kievan Rus, Lithuanian-Russian state Zaporizhzhya Sich, Cossack Hetmanate and further several stages to this day [5]. It has historical interest. When the information security problems in the Soviet Union went out of the shadows of secrecy, the materials were published in the open press, which became widely available for scientific and technical experts. In 1980, was translated and published the book by L. Hoffmann [6]. An important addition was the book by D. Xiao,

D. Kerr and S. Mednick [7]. In 1989 «Foreign electronics» journal has placed a number of materials and a special edition of «Information Security» [8]. Later, experts learned about the work of software security [9]. The principles presented in these publications remain almost intact to this day. Further transformation of information security forms out of preservation and expansion of the basic principles of all previous stages. In 1992, A. Timonin came to conclusion that «in general, the problem of security in automated information systems refers to not algorithmically solvable problems» [10]. However, a number of problems remain unresolved: still remains the closed functional completeness of the set security services, insufficient front of speakers security processes researches, needs to be expanded the modeling of social and psychological security. The urgency of these problems stems from the Cybersecurity Strategy of Ukraine.

The main *aim of this paper* is to identify the characteristics of the gradual transformation paradigms from data and information security to behavior security, individual and group mind, changes systematization of the information security paradigms; production of further transformation of information security paradigms; reasoning the implementation of identity definition and defined identity management.

Stages differentiation of information security transformation

Further stages of information security will be divided on the basis of a paradigm shift in information security. The concept of paradigm considered as a set of values, methods, approaches, technical skills and resources, problem solving methods accepted in the scientific community of experts within the established scientific tradition in a certain period of time. The paradigm is undergoing changes depending on accumulated practical experience and research results. Terms of stages separation determine the facts of the appearance of the relevant national legal documents. We consider the following stages of transformation information security technologies:

Stage 1 (1992 – 1996) – the establishment of its own system for the security of classified information in Ukraine and creation of technical security of information (TSI) in the fields;

Stage 2 (1996 – 2000) – the creation and development of the legal framework and integrated systems for information security in automated systems;

Stage 3 (2000 – 2004) – the development of the legal framework of security of state information resources and harmonization of national and international regulatory framework;

Stage 4 (2004 – 2008) – further development of information resources security in all types of public, commercial and personal data in information and telecommunication systems and information space of Ukraine. Expanding the scope of information security on public information;

Stage 5 (2008 – 2012) – development of information security against the backdrop of a foreign network-centric paradigm of information and influence;

Stage 6 (2012 – 2016) – the development of cyber cyberspace state. Expanding the scope of information security for commercial and social spheres;

Stage 7 (current) – IS update tasks as components in the plane of the information confrontation and information warfare. The establishment of security system behavior, individual and group mind.

The transformation is not complete, on the contrary we have the transformation accelerating, the development of methodologies and interdisciplinary approaches. Transformations and changes of paradigms are summarized in the Table 1.

Historically, information security paradigms successively change one another, maintaining, improving and complementing the previous ones. Some paradigm enacted into life simultaneously - for example, some in the public sector, and others simultaneously in the private sector. Some paradigm implemented into action and improved for several stages.

Historically, the first standard that has made a huge impact on the security of information networks has become the standard US Department of Defense «Evaluation Criteria Trusted Computer Systems» (first published in August 1983). Thus, in the mid 80 of the last century laid the foundation of information security strategy.

Classical information security paradigm based on access control

The classic paradigm of information security based on access control was introduced on 1 phase (1992 – 1996). This is a phase of development of its own system for the security of classified information in Ukraine and creation of technical security of information (TSI) in the fields. Was first applied to the automated system of Class 1 – one machine and one user complex, which processes the information to one or more categories of confidentiality [11].

Table 1.

Stages and factors of transformation in information security

№	Paradigm	Scope of security	Security tool and technologies	Basic terminology
1	Classical, based on access control	Information security in technical information processing systems	Ensuring the confidentiality, integrity, availability (CIA), observability, guarantees	TSI, CSI, access control
2	Frontier security (perimeter defense)	Information security in automated systems (computers)	CIA + access control	CSM, security policy, security
3	Multi-layered information security system	Information security in automated systems (computer networks)	CIA + access control + firewall + privacy (P)	Government Information Resources
4	Network-centric (I)	Information Security of information resources (IS IR) of technology and major communications	CIA + access control + firewall + System of detection prevention and mitigation for IS incidents	ICS, information space, personal information
5	Network-centric (II)	Information security of critical infrastructures	CIA + access control + firewall + System of detection prevention and mitigation for IS incidents + search for vulnerabilities	Critical information infrastructure
6	Cyberspace	Cybersecurity environment, resources, social capital, information production	CIA + access control + firewall + System of detection prevention and mitigation for IS incidents + search for vulnerabilities + Audit, monitoring and insurance	Cyber security, state cyberspace
7	Sociocentric, security behavior and consciousness	National security, social and psychological security	… + Cybersecurity + Psychological security	Information impact, information war

Control access to information organized so that only authorized person or process are entitled to read, write, create, or delete information. The information system was autonomous and did not extend beyond the agency or organization. An example of such a system is autonomous personal electronic computer (the PC). The main goal of information security is to prevent information security threats, preventing information theft and computer facilities, disclosure, loss, destruction and distortion of information, ensure the normal production of all departments of the facility information. The main tasks of the security were considered: restricting who is allowed to access; creating a system of passwords and access for users to information by categories. The theoretical basis of security systems was the theory guaranteed secure systems [12]. The general principle of the information security is a maximum efficiency of risk accepted not below the fixed risk when operational risk is minimal. Organizational data security is divided into technical information security (TSI) and cryptographic security of information (CPI).

Technically, the goal of information security is to implement the rules, measures and action to prevent damage and / or loss in the case of attacks and threats to information. Security carried out comprehensive information security system (CISS), which consists of legal, organizational, methodological, technical, software, information and mathematical provisions that prevent the realization of the threat or significantly impede the realization of the attacks. The complex remedies considered as a set of functional services that combine to create the necessary functionality profile security. Each service is a set of features that allow you to withstand a certain set of threats. Security policy can be implemented using a variety of services and mechanisms, alone or in combination, depending on the objects of policy. In general arrangements belong to one of three classes, which may overlap: prevention, registration, renewal. To provide services using security mechanisms. According to international recommendations ITU-T [13] security network built in a hierarchical multi-modular, security - security services - functional security services - security mechanisms. [14].

The classic paradigm of information security was extended to automated Class 2 - localized multimachine multiuser systems that process information for one or more categories of confidentiality. An example of such a system is a local area network (LAN) connection between computers which do not go beyond the controlled area or local communication system. CISS were based on the theory guaranteed secure networking model which includes guaranteed by components and is guaranteed by channels that connect the components together. In communication systems to counter potential threats to international recommendations ITU-T [15] defined tasks such security networks: confidentiality of information that is stored or transferred; the integrity of the data, i.e. information that is stored or transferred; the integrity of the system, in particular, the problem the operating system security; reporting (this includes the problem of observability), where each object / subject should be responsible for any action which he initiated; readiness (availability) is the property of the information environment in which all legitimate objects must get correct access to information system.

Its highest classical paradigm of information security in automated systems reached Class 3 - Distributed multi machine and multi user complexes which process different categories of information confidentiality and where is the need to transfer information via unsecured environment. At this stage, the terminology has changed: instead of the term automated systems were used correct term: Information and Communication System (ICS) and a short – Infocommunication system. Add another important characteristic of these systems, - Class 3 assign ICS belonging to one operator (one owner), allowing it to implement a uniform security policy.

Examples of Class 3 are a special system of confidential communication or public communication networks of one operator. CISS distributes its functions for network elements. In general, the problems of CISS are to: prevent, detect, respond and scare. An adequate level

of information security can be achieved only through an integrated approach that involves the systematic use of physical, software and technical and organizational measures and means. Security is required for all components of information and telecommunication systems, lines, channels, transmission systems, hardware, software, information and personnel. The ultimate aim is the selection of effective means to counter threats in the implementation of information security, the cost of which, in any case, should not exceed the value of losses expected from the sale of threats.

Paradigm of information activity objects border security

Frontier security paradigm of information objects (perimeter defense system) operated in phase 2 (1996 - 2000) - the stage of the creation and development of the legal framework and integrated systems for information security in automated systems. Since the mid 90^{th} years widely spread personal computers, local area networks and the Internet. To access the distributed database implemented technology client - server. The situation with the information security began to deteriorate. In 1995, the open print publications appearing on the concept of information confrontation and information war. In response developed and implemented this paradigm. Frontier security paradigm based on guaranteed secure core idea around which provided several lines of circular defense and demilitarized zones [16].

Communication centers, switching centers, final system focused on a relatively small area and security system built on a «circular defense» (defense barrier or perimeter security). All items are secured, are located in a secure physical environment of the area that is secured. Station equipment switching units placed on the secured object, where the full cycle of organizational and technical measures for comprehensive information security qualified certain level. From the theoretical positions of technical security of information known as the weakest link, which is not blocked by institutional and / or organizational and technical or cryptographic means, determines the resulting level security system.

When ensuring confidentiality must allocate tasks source message non-repudiation consumer messages non-repudiation network that receives messages from source to destination delivery. The same considerations affecting the distribution of tasks and authentication of information interaction network in the transfer of information between network operators, we are different.

Telecommunication networks secured «distributed by». The concept of «distributed» systems of security applies in particular in IP-networks. The concept of channel transmission in such networks blurred. Path cannot be secured in one message. Message divided into packets, each of which can be transmitted to an arbitrary route. They can create virtual channels of communication. The security level of the network route determined by the weakest security s of all possible routes. A security route is determined by its weakest link.

To secure computer networks from unauthorized access foreign organized criminals filtering security perimeter network using a network between screens - firewall (firewall). Between the firewall – a set of hardware and software that monitor and filter network packets that pass through it, in accordance with desired configuration rules. Complex of multifrontier security provided by operating systems with multi-level security from unauthorized access, encrypted transaction methods of cryptography, start to block access cards.

But the effectiveness of the information security of computer networks proved insufficient. Up to 60% of all incidents of information security in networks account for internal attackers and the human factor: errors and not enough staff qualification, malicious actions, lack of control and so on.

Paradigm of layered multilevel information security system

The paradigm of multi-layered information security system of information resources and technologies appeared on stage 3 (2000 - 2004). This stage is characteristic by development of the legal framework of state information resources security and harmonization of national and international legal framework of information security. The paradigm has evolved with further expansion of the use of information technology, global distribution of Internet, mass introduction of remote access to distributed client - server database technology. Paradigm prerequisites are the emergence of multi-operating system secured from non-authorized access, usage of cryptography to encrypt transactions, implementation of locking connecting devices means.

This paradigm layered, multi-level information security based on the idea of guaranteed secure core around which provided several lines of circular defense and demilitarized zones. The system of circular defense by multiple threats provided CISS, understood as a set of measures and means for preventing information leaks by technical, acoustic, vibroacoustic, electromagnetic (and aiming), laser, infrared, radiation, chemical channels that created the main and additional (auxiliary) technical means of processing information and security against unauthorized access to information and means of processing in automated systems. Continuous security is provided as time - at all stages of the life cycle of information security, the decision on security, development of technical specifications, design, creation of security, usage and disposal after its decommissioning. From the theoretical positions of information security is known that the weakest link, which is not blocked by means of security, determine the resulting level of security.

A characteristic feature of this paradigm for foreign countries is the transition from concepts of information security to the concept of information security technologies and information resources. In addition to the main goal of information security is added to ensure stable operation of information and communication systems, security of the legitimate interests of enterprises from illegal encroachments, prevent theft of funds, improve service quality and security guarantees property rights and interests of customers. Conceptual engineering model of multi-layered information security system represented by a group of international ISO / IEC 15408 standards [17], which determine the development of technology security profiles and security projects. [18] In Ukraine, he is introduced as the industry standard, for example, in the banking sector [19, 20]. Conceptual model of information security system additionally includes a set of security services and security mechanisms that implement services that provide monitoring functions, security and adaptation of information resources to prevent the possibility of gradual penetration offender detect the fact of penetration, object localization invasion and attack and neutralize expulsion of the offender, restore lost functions of the system. New in the conceptual model is widely used filters, firewalls that secure the perimeter. Considering that about half of information security problems associated with the human factor, the security against internal and external malicious intrusion detection systems, recognition of abnormal behavior, adaptive algorithms recovery systems and facilities security are implementing.

Network-centric paradigm of information security

Network-centric paradigm of information security of information resources launched on 4 phase (2004 - 2008). It marked a further development of information resources security in all types of public, commercial and personal data in information and telecommunication systems and information space of Ukraine. It was expanded by the scope of information security in open information. The paradigm stems from the wealth of modern international experience and scientific advances in this field, with the rapid development of infocommunication fact, the development, the complexity and the increasing role of

communications networks as a critical public resource. Since the beginning of 2000 «information security has become directly tied to the security infrastructure of the country and welfare of the nation» [21]. The network-centric (network-centric) paradigm of information security began to develop. In Ukraine, this paradigm has been recognized as inappropriate. The inertia of the previous paradigm of multi-layered information security system is not allowed to fully appreciate the importance of the new paradigm. Proposed by the concept of information security of telecommunication networks [22] found no significant response nor the scientific community nor the official agencies. However, network-centric paradigm has had a significant impact on information security business and telecommunications. For Ukraine, this paradigm can be formulated as follows:

For Phase 4 - implementation of network-centric paradigm of information security of the state critical infrastructure, including information and communication networks as the most critical public resource.

For the next 5 stage - the paradigm of increasing requirements for survivability of information systems as part of critical infrastructure and are characterized by a high degree of resource allocation and decentralization of management to enhance the role of technical operation in terms of requirements to preserve a minimal set of critical functions to the survivability of information systems, security factor to the action of destabilizing factors in the environment, including information influence. The concept of this paradigm is the main problem of increased requirements for information systems that were characterized by dispersed resources and decentralized management. Characteristic features of this paradigm are: more interlocking processes of information security management utilities – energy, transport, telecommunications, pipeline networks; access to forefront properties availability and integrity as indicators of sustainable and effective functioning of the systems; transition to the next stage of information security technology cybersecurity cyberspace to businesses, organizations and users.

For example, the Law of Ukraine «About Telecommunications» are requirements for preparedness and survivability, providing support for such properties as the reliability of the telecommunications system, its sustainability, resource availability, integrity and recoverability system structure. Information security of telecommunications networks should be ensured in the integration of information and communication technologies, various types of networks and telecommunications services, quantity and quality are constantly increasing, and activities on the networks of different ownership forms. It is necessary to harmonize methods of information security for the various components of information and telecommunication systems and networks, including information resources, applications, and telecommunication protocols. An integrated approach means a need for a network infrastructure for information security vulnerabilities as any network link can cause problems for all involved, both for the providers and operators and consumers of services. Information security in telecommunication systems are difficult complex task. Secure telecommunications network should be secured from malicious and unintentional attacks, be reliable, scalable, provide a guaranteed response time, availability of services and information, integrity of data and equipment, accurate billing information. The security components of the telecommunications network is critical to the security of the entire network, including applications and services.

However, because the network combines a large number of elements, making progress determines their ability to interact or lack of such capacity. Information security should be implemented threats not only from each item or service, and should be provided in collaboration tools and security features in a multimedia environment with the full implementation of the overall security of information transfer from one end [14].

The complexity of modern telecommunication networks and information and communication systems (ICS), management and interaction of networks leads to the necessity and usefulness allocation separately ICS Class 4 - Global Distributed multicomputer, Multi,

multidomain complex which processes information of various categories of confidentiality and has different owners domains. Domain 3 includes class X, owned by one owner and has the CISS, securing the perimeter domain, its information security management system, its system of prevention, detection, treatment and elimination of incidents of information security, a single domain security policy. Domains of different owners may have different security policies.

In telecommunications information security is intertwined with: management of the quality of communication services where security and preparedness information resources are an integral part of assessing the quality of services; management of economic efficiency, which is the relationship between information and economic risks; tasks in technical operation of software requirements to preserve a minimal set of features critical to the survivability of information systems to the security factor by the action of destabilizing factors of the environment.

The hazard level threats target information influence is directly proportional to the level of technological development and scale networks use computers in a network management system, the industry and the state as a whole. For the growing importance of telecommunication networks requirements to ensure the integrity and reliability of information transmission, security violations routing accuracy and timeliness of information delivery (minimum delay messages), and secure against unauthorized access to information resources, networking and physical security infrastructure. But the fifth stage in Ukraine did not happen.

Network-centric principle in national security system of Russia

At stage 5 (2008 - 2012) meant to be the development of information security against the backdrop of a foreign network-centric paradigm of information and influence. Expansion of the information security sphere was scheduled to the commercial and social sectors. Ignoring international experience, including neglect of network-centric paradigm has led to serious consequences for Ukraine. On the contrary, Russia reacted to carefully develop in the US new concept of network-centric war. Its foundation is the «network» and the basic principle is the principle of «network-centrism». The principle of network-centrism, in regard to information system management of Russian national security lies «in addition vertical administrative-command network structures of government by horizontal informal, self-organized network structures of civil society in all their diversity, organized by the network principle in historical time, geographical space, subordination and problem targeting national security» [23]. With the principle of «network-centrism» follows the use of «organizational weapons». «Organizational weapon - a combination of national and transnational, sympathizing network of structures that combine small, but very influential group of politicians, senior government officials, military, law enforcement officials, the media, big business, political parties, etc., who are willing on whether other reasons contribute to the promotion of Russian national interests» [23]. The use of organizational weapons can pursue constructive and destructive purposes. The design objective is to create the necessary conditions to create in their environment state «affective-cognitive-volitional consonance - unity understanding and experience of voluntary departures in the majority of its members». On the contrary, destructive goals – a state of «affective-cognitive-dissonance willful», that is a hard conflict in the area of understanding, experience and network will in member institutions whose activities are disgusted with the national interests of Russia. Selector strategic and operational decisions in the field of national and military security is the use of the totality of interacting distributed in historical time, geographical area, the subordination and problem orientation multilateral, multi-business strategic computer games, which are models of national security Russia.

Consider what we did not realize at the time until 2013. That is, consider the problem of the emergence of unfriendly and hostile information influence on Ukraine. Expanded information war (and now too hot) with Russia based on serious scientific research, particularly on solid mathematical training processes information influence and mathematical modeling. The first publications in the media on the concept of information confrontation and information war there in 1995. Published monographs [24-25] and many other publications on this topic. Attention is drawn to the mathematical side of research. In the book [25] presented a mathematical apparatus for research capacity systems self-studying, in terms of targeting. In 2001, D. Chernavskii published detailed results of research in the dynamic information theory. Described model generating valuable information «can be used in various fields of biology, linguistics, sociology and history» [26]. There are research results risks modernization, management of regional industrial complexes, of socio-historical development ... and the language of war. Objectivity model tested and refined on the historic under-lays, in our time tested in practice, realized in an aggressive policy of Russia. Here is a brief description of this model.

Use mathematical apparatus of the theory of dynamical systems. The system is composed of several types of *i* objects, which belong to the same set of power system N. Each element has a type of information. Information may be language, culture characteristics, production capacity, psychological parameters and more. Element system can occur and disappear. The life of each *i* -then element of τ_j is less life expectancy on the same system. Each *i*-th element contributes to the objects of the same type. «We can offer in information distributed in the space of a dynamic system type» [26]

$$\frac{\partial u_i}{\partial t} = \frac{1}{\tau_i} u_i - \sum_{j \neq i} b_{ij} u_i u_j - a_i u_i^2 + D_i \Delta u_i,\ a_i < b_i, \qquad (1)$$

where u_i – the concentration of i - type elements and each element has i -type information. $(1/\tau_i)u_i$ describes autocatalytic (mutually supported) by reproducing the characteristic time of autoreproduction. $b_{ij}u_i u_{ij}$ describes the interaction of elements. Sign «minus» means that the interaction is antagonistic (or competitive) nature. At a meeting of two different elements, each of them or seeking to impose their information second or «destroy» him. $a_i u_i^2$ describes the effect of «struggle» or destruction at the meeting of the two identical items due to competition for resources environment. This significant member when the concentration of the same elements becomes too large. The latter term describes the diffusion element when the elements are moved and mixed in space. D_i – elements diffusion coefficient.

Equation (1) used to describe the emergence of a single genetic code, description of the violation of punitive symmetry to describe historical events when added to the equation (1) members describing the geography of a territory that is modeled. The model was tested on the description of the main events of European history, from the Middle Ages to the present day. During the events understood the formation of new powers and disappearance of others. Simulation shows satisfactory resemblance to the real facts of history, if choose appropriate coefficients model. The model allows global forecast more or less distant future.

Unfortunately, this model can be used for planning the language of war. For example, when modeling the dynamics of ethno received the following conclusion: «With increasing migration length (which is inevitable with the development of techniques and technologies), physical obstacles and cease to play the role of space can be considered homogeneous. Then all clusters are driven one, which is - it is impossible to present. In other words, the world will be global and will be managed by another state. There will be one state language and common rules of conduct that is – clearly cannot provide ...» [26].

In the face of all this kind of dangers and threats to Ukraine was disarmed physically and psychologically information. Only decisive action conscious of society turned then lost its political independence.

Paradigm of enterprises and organizations cyberspace security

The paradigm of cyber environment cybersecurity of companies, organizations and users within 6 phase (2012 - 2016) is still in its infancy. To be implemented cyber cyberspace development of the state and expand the scope of information security to the commercial and social sectors. Meanwhile, the problem of cyber security has become topical in the world. Tools cyberattacks were used to obtain advantages in information exposure and cyberwarfare. In many countries «formed special units that have a purpose: conducting exploration work in networks, securing their own networks, blocking and» collapse «of enemy structures using cyberspace capabilities» [27].

Among the starting legal documents approved «Cybersecurity Strategy of Ukraine». Terminology and regulatory framework moves to the development stage. We know that cybersecurity (or rather «the sphere of cyber security» in terms of specialists ICS - Ed.) - A set of tools, strategies, security principles, security guarantees, guidelines, approaches to risk management, performance, training, practical experience, insurance and technology that can be used to secure cyber environment, resources, organization and person. «Resources include user or organization connected computer devices, personnel, infrastructure, applications, services, telecommunications systems, and the totality of transmitted and / or stored information in cyber environment. Cybersecurity is to try to achieve and preserve the properties of the security resources of the organization or user directed against the relevant threats cyber environment. General security tasks include: accessibility; integrity, which may include authenticity and non-recovery; confidentiality. Cyber environment includes users, network devices; all software processes are stored or transit information, applications, services and systems that can be directly or indirectly connected with the networks» [28]. In other words, should reflect the paradigm shift from information security technologies and important communications to various kinds of cyber security and resources cyber environment all businesses, organizations and users. Especially emphasizes critical information security (especially information) infrastructures. «Cybersecurity includes a social capital, information production. In today's business environment concept perimeter disappears. The boundaries between internal and external networks become more blurred» [28]. Security is ensured at all levels of telecommunications networks, network access, network, and transport levels, levels of network management and provision of services. Among the strategic aspects of cybersecurity Ukraine formulated the problem of «Building effective mechanisms to secure national interests and the need to develop a single vision of cybersecurity as state bodies and business structures» [29].

Conclusions

In this paper classified transformation stages and directions of information security paradigms, shows change relative importance of types of information security. In the second part will be supplemented by a list of actual problems in information security, proposed application of the definition of identity and identity management definition. The results will improve management information systems, information-psychological and cyber security.

References

1. Обисо, М. Развитие международного сотрудничества в области кибербезопасности. Глобальный ответ на глобальный вызов / Марко Обисо // Межрегиональный семинар для стран Европы, Азиатско-Тихоокеанского содружества независимых государств (Европа-АТР-СНГ) «Современные методы борьбы с киберпреступностью». – Одеса, Украина, 28-30 марта 2012.
2. Бурячок, В.Л. Інформаційна та кібербезпека: соціотехнічний аспект: підручник / В.Л. Бурячок, В.Б. Толубко, В.О. Хорошко, С.В. Толюпа; за заг. ред. д-ра техн. наук, професора В.Б. Толубка. – К.: ДУТ, 2015. – 288 с.
3. Корченко, О.Г. Оцінювання шкоди національній безпеці України у разі витоку державної таємниці: Монографія / О.Г. Корченко, О.С. Архипов, Ю.О. Дрейс. – К.: наук.-вид. центр НА СБ України, 2014. – 332 с.
4. Герасименко, В.А. Проблемы защиты данных в системах их обработки / В.А. Герасименко // Зарубежная радиоэлектроника. Защита информации. – 1989. – Специальный выпуск. № 12. – С. 5-21.
5. Гуз, А.М. Історія захисту інформації в Україні та провідних країнах світу: Навчальний посібник / А.М. Гуз. – К.: КНТ, 2007. – 260 с.
6. Гофман, Л.Дж. Современные методы защиты информации / Л.Дж. Гофман. – М.: Сов. радио, 1980. – 264 с.
7. Сяо, Д. Защита ЭВМ / Д. Сяо, Д. Керр, С. Мэдник. – М.: Мир, 1982. – 203 с.
8. Защита информации. Специальный выпуск / Зарубежная радиоэлектроника. – № 12/1989. – 112 с.
9. Защита программного обеспечения, Пер. с англ. / Д. Гроувер, Р. Сатер, Дж. Фипс и др. // Под редакцией Д. Гроувера. – М.: Мир, 1992. – 288 с.
10. Грушо, А.А. Теоретические основы защиты информации / А.А. Грушо, Е.Е. Тимонина. – М.: Издательство «Яхтсмен», 1996. – 192 с.
11. НД ТЗІ 2.5-005-99. Класифікація автоматизованих систем і стандартні функціональні профілі захищеності оброблюваної інформації від несанкціонованого доступу [Електронний ресурс]. – К.: ДСТСЗИ СБУ, 1999. – 20 с. – Режим доступу: http://www.dsszzi.gov.ua/dstszi/control/uk/publish/article?art_id=403818&cat_id=38835
12. Зегжда, Д.П. Основы безопасности информационных систем / Д.П. Зегжда, А.Г. Ивашко. – Г.: Горячая линия - Телеком, 2000. – 452 с.
13. Recommendation ITU-T X.800. Security architecture for Open Systems Interconnection for CCITT applications [Електронний ресурс]. – Geneva, 1991. – 48 с. – Режим доступу: http://www.itu.int/rec/T-REC-X.800-199103-I
14. Кононович, В.Г. Технічна експлуатація систем захисту інформації телекомунікаційних мереж загального користування. Частина 3. Архітектура безпеки Концепція захисту інформації: [навч. посібник для вузів, затверджено Міністерством транспорту та зв'язку України] / Кононович В.Г. – Одеса, ОНАЗ, 2009. – 194 с.
15. Рекомендация МСЭ-Т Е.408. Требования к безопасности сетей электросвязи [Электронный ресурс]. – Женева, 2004. – 21 с. – Режим доступу: https://www.google.com.ua/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiw6pmYjZvQAhVBEywKHTkKDrgQFggcMAA&url=https%3A%2F%2Fwww.itu.int%2Frec%2Fdologin_pub.asp%3Flang%3De%26id%3DT-REC-E.408-200405-I!!PDF-R%26type%3Ditems&usg=AFQjCNFq_wLxyafcx223VwO6vGv0bzGbhA&sig2=TtlAbZuD4XmAX9XuYOBCeA&bvm=bv.138169073,d.bGg
16. Кононович, В.Г. Аналіз проблеми розподілу витрат на інформаційну безпеку інформаційно-телекомунікаційних систем / В.Г. Кононович, М.Ф. Тардаскін, Т.М. Тардаскіна // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні : науково-технічний збірник. – 2004. – Вип. 8. – С. 62-68.
17. Standart ISO/IEC 15408:2000. Information technology - Security techniques - Evaluation criteria for IT security. - Part 1: Introduction and general model. - Part 2: Security functional requirements. - Part 3: Security assurance requirements.
18. Бондаренко, Г. Перспективы применения международного стандарта ISO/IEC в Украине / Г. Бондаренко, Л. Скрыпник, А. Потий // Правовое, нормативное и метрологическое обеспечение системы защиты информации в Украине. – 2001. – Вып. 3. – С. 7-26.
19. ГСТУ СУІБ 1.0/ISO/IEC 27001:2010 (ISO/IEC 27001:2005, MOD). Інформаційні технології. Методи захисту. Система управління інформаційною безпекою. Вимоги [Галузевий стандарт України]. – К.: Національний банк України. 2010. – 49 с.
20. ГСТУ СУІБ 2.0/ISO/IEC 27002:2010 (ISO/IEC 27002:2005, MOD). Інформаційні технології. Методи захисту. Звід правил управління інформаційною безпекою [Галузевий стандарт України]. – К.: Національний банк України. 2010. – 163 с.

21. Леваков, А. Анатомия информационной безопасности США. Информационная безопасность [Электронный ресурс] / А. Леваков // Информационный бюллетень. – М.: Jet Info online. – № 6 (109), 2002. – 40 с. – Режим доступа: http://www.jetinfo.ru/Sites/info/Uploads/2002_6.DF9C812FFBD9496BAE9694E27F2D9D1D.pdf
22. Кононович, В.Г. Основні положення концепції інформаційної безпеки телекомунікаційних мереж загального користування / В.Г. Кононович, М.Ф. Тардаскін // Захист інформації. – 2006. – № 1(28). – С. 18-30.
23. Никитенко, Е.Г. Облик перспективной информационно-управляющей системы обеспечения национальной безопасности России / Е.Г. Никитенко, Н.А. Сергеев // Оборонно-промышленный комплекс России. – 2012. – Т. 8. – С. 491-506.
24. Почепцов, Г.Г. Информационные войны / Г.Г. Почепцов. – М.: Рефл-бук, –К.: Ваклер, 2000. – 576 с.
25. Расторгуев, С.П. Информационная война / С.П. Расторгуев. – М: Радио и связь, 1999. – 416 с.
26. Чернавский, Д.С. Синергетика и информатика: Динамическая теория информации / Д.С. Чернавский // Предисл. и послесловие Г.Г. Малинецкого. – М.: Книжный дом «ЛИБРОКОМ», 2001, 2009. – 304 с.
27. Дубов, Д.В. Кібербезпека: світові тенденції [Електронний ресурс] / Д.В. Дубов, М.А. Ожеван // Доповідь на Міжнародній конференції 26 травня 2011 р. – К.: НІСТ, 2011. – 30 с. – Режим доступу: http://www.niss.gov.ua/articles/510.
28. Recommendation ITU-T X.1205. Telecommunication security. Overview of cybersecurity. – Geneva: 2008. – 56 p.
29. Дубов, Д.В. Стратегічні аспекти кібербезпеки [Текст] / Д.В. Дубов // Стратегічні пріоритети. – 2013. – № 4 (29). – С. 119-126.

ТРАНСФОРМАЦІЯ ПАРАДИГМ ЗАХИСТУ ІНФОРМАЦІЇ, ІНФОРМАЦІЙНОЇ ТА СОЦІАЛЬНО-ПСИХОЛОГІЧНОЇ БЕЗПЕКИ (Частина 1)

С.О. Гнатюк[1], В.О. Гнатюк[1], В.Г. Кононович[2], І.В. Кононович[3]

[1] Національний авіаційний університет,
просп. Космонавта Комарова, 1, м. Київ, 03680, Україна; e-mail: s.gnatyuk@nau.edu.ua
[2] Одеський національний політехнічний університет,
просп. Шевченко, 1, Одеса, 65044, Україна; e-mail: vl_kononovich@ukr.net
[3] Одеська національна академія харчових технологій,
вул. Канатна, 112, м. Одеса, 65039, Україна; e-mail: kononovich@mail.ru

У роботі представлені результати ретроспективного аналізу етапів трансформації парадигми сфери інформаційної безпеки – від захисту інформації та інформаційної безпеки до захисту свідомості та поведінки людей. Сформульовані сучасні парадигми інформаційного захисту. Описані нові проблеми забезпечення захисту критичних інформаційних інфраструктур. В результаті аналізу запропоновано способи вирішення низки часткових задач. Впровадження системи визначення ідентичності та управління визначенням ідентичності дасть можливість замкнути повноту механізмів захисту. Тим самим досягається можливість відслідковувати кожну транзакцію в мережі. Інформаційно-психологічний захист від деструктивного інформаційного впливу вимагає застосування соціально-психологічних методів. Пропонується проста модель формування групової свідомості навколо ідеї боротьби з кіберзлочинністю. Отримана систематизація та результати вирішення задач дозволяє підвищити ефективність роботи систем забезпечення інформаційної, кібернетичної та соціально-психологічної безпеки й формалізувати напрямки подальших досліджень щодо розробки ефективних систем безпеки.

Ключові слова: захист інформації, інформаційна безпека, кібербезпека, інформаційно-комунікаційні системи, індивідуальна та групова свідомість, соціально-психологічний захист, правова система.

ТРАНСФОРМАЦИЯ ПАРАДИГМ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННОЙ И СОЦИАЛЬНО-ПСИХОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ (Часть 1)

С.А. Гнатюк[1], В.О. Гнатюк[1], В.Г. Кононович[2], И.В. Кононович[3]

[1] Национальный авиационный университет,
просп. Космонавта Комарова, 1, г. Киев, 03680, Украина; e-mail: s.gnatyuk@nau.edu.ua
[2] Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: vl_kononovich@ukr.net
[3] Одесская национальная академия пищевых технологий,
ул. Канатная, 112, м. Одесса, 65039, Украина; e-mail: kononovich@mail.ru

В работе представлены результаты ретроспективного анализа этапов трансформации парадигмы в сфере информационной безопасности – от защиты информации и информационной безопасности до защиты сознания и поведения людей. Сформулированные современные парадигмы информационной защиты. Описаны новые проблемы обеспечения защиты критических информационных инфраструктур. В результате анализа предложены способы решения ряда частных задач. Внедрение системы определения идентичности и управления определением идентичности даст возможность замкнуть полноту механизмов защиты. Тем самым достигается возможность отследить каждую транзакцию в сети. Информационно-психологическая защита от деструктивного информационного влияния требует использования социально-психологических методов. Предлагается простая модель формирования группового сознания вокруг идеи борьбы с киберпреступностью. получена систематизация и результаты решения задач позволяет повысить эффективность работы систем обеспечения информационной, кибернетической и социально-психологической безопасности и формализовать направления дальнейших исследований и разработки эффективных систем безопасности.

Ключевые слова: защита информации, информационная безопасность, кибербезопасность, информационно-коммуникационные системы, индивидуальное и групповое сознание, социально-психологическая защита, правовая система.

УДК 621.317.799 297 + 681.849

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 240-248

ОБОСНОВАНИЕ ПРИМЕНЕНИЯ ФРАКТАЛЬНОГО ПОДХОДА ДЛЯ СОЗДАНИЯ КОМПЛЕКСА АППАРАТУРЫ КОНТРОЛЯ ПОДЛИННОСТИ ЦИФРОВЫХ ФОНОГРАММ ПРИ ЭКСПЕРТИЗЕ МАТЕРИАЛОВ И СРЕДСТВ ЦИФРОВОЙ ЗВУКОЗАПИСИ

О.В. Рыбальский, В.И. Соловьев, В.В. Журавель

Национальная академия внутренних дел,
пл. Соломянская, 1, Киев, 03056, Украина; e-mail: rybalsky_ol@mail.ru

Показан фрактальный характер оцифрованного аналогового сигнала при записи информации на аппаратуре цифровой звукозаписи и фрактальный характер проявлений паразитных параметров такой аппаратуры, фиксируемых в информационных сигналах и сигналах шумов, записываемых на цифровых носителях. При этом показано, что малая величина этих проявлений требует применения чувствительных методов анализа. Показано, что, учитывая мультифрактальность этих проявлений, такой подход обеспечивает необходимую точность анализа и является наилучшим решением задачи создания требуемого инструментария проведения экспертиз материалов и средств цифровой звукозаписи.

Ключевые слова: аппаратура цифровой звукозаписи, аналого-цифровое преобразование, паразитные параметры аппаратуры записи, цифровая фонограмма, фрактальность

Введение

При разработке комплекса аппаратуры для контроля оригинальности и подлинности цифровых фонограмм (ЦФ), предназначенной для проведения экспертиз материалов и средств цифровой звукозаписи, после проведенного анализа концептуально был выбран фрактальный подход к созданию такого экспертного инструментария. *Цель* статьи – показать теоретическое обоснование выбора такого подхода к созданию необходимой аппаратуры.

Основная часть

Известно, что аналоговый сигнал (АС) представленный, например, как $s(t) = A_m \cos \omega_0 t$, после прохождения через систему аналого-цифро-аналогового преобразования (АЦАП) записывается в виде соотношения

$$s\ (t) = \sum_{n_2=-\infty}^{\infty} rect\left(\frac{t-nT}{T}\right) A_m \cos \omega_0 (nT), \qquad (1)$$

где n – номер выборки, T – период дискретизации входного аналогового сигнала [1].

Этот сигнал, по своей сути является сигналом, присутствующим на выходе цифроаналогового преобразователя (ЦАП) аппаратуры цифровой звукозаписи (АЦЗЗ) после прохождения тракта записи-воспроизведения информации при прямой импульсно-кодовой модуляции (ИКМ) исходного АС. Не рассматривая процессы записи и воспроизведения сигнала на носитель (считая их идеальными и не вносящими дополнительных искажений в исходный сигнал), рассмотрим процесс прохождения сигнала в системе АЦАП более подробно. Цель такого рассмотрения – определение характера искажений, вносимых в сигнал на выходе АЦЗЗ при прохождении в системе АЦАЦ.

Известно, что в процессе преобразования аналогового сигнала в цифровую форму он подвергается процессу дискретизации во времени и квантования по уровню. Для дискретизации звуковых сигналов применяется устройство выборки и хранения (УВХ) [1].

Таким образом, если рассматривать сигнал на выходе системы АЦАП с точностью до $\pm 1/2$ младшего разряда квантования (МР), то саму систему АЦАП можно представить (согласно с известным методом [2–4]), как линейную систему, к которой прибавляются шумы квантования, как это представлено на рис. 1, где $\eta(t)$ – шумы квантования по уровню, $x^*(t)$ – сигнал на выходе УВХ, $y(t)$ – сигнал на выходе квантователя.

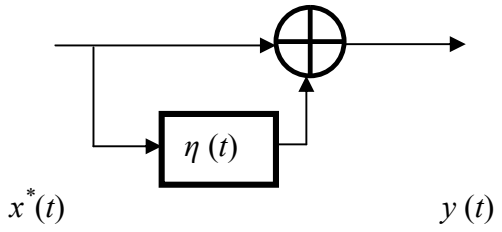


Рис. 1. Представление АЦАП, как линейной системы

Предложенный подход позволяет рассматривать систему АЦАП как линейное устройство, подчиняющееся принципу «суперпозиции».

Для определения искажений АС, происходящих в системе АЦАП в работе [3] был предложен подход к аналитическому описанию происходящих в такой системе процессов. Этот подход позволил получить необходимые соотношения, описывающие процессы в рассматриваемой системе. Воспользуемся этими соотношениями для описания искажений АС, определяемых процессами дискретизации и квантования.

Оператор действия УВХ

$$x^*(t) = L\{x(t)\} = \sum_{k=-\infty}^{\infty} x(nT)\varphi(t-nT) = \sum_{n=-\infty}^{\infty} \varphi(t-nT) \times \int_{-\infty}^{\infty} x(t')\delta(t'-nT)dt', \qquad (2)$$

где

$$\varphi(t) = rect\left[\frac{t-T/2}{T}\right],$$

а

$$\varphi(t-nT) = rect\left[\frac{t-\frac{(2n+1)T}{2}}{T}\right],$$

где n – номер выборки дискретизированного АС, T – период дискретизации [3].

Из (2) следует, что УВХ – линейное устройство и может быть рассмотрено в частотной области.

Найдем передаточную функцию УВХ в частотной области [3]. Сигнал $x^*(t)$ на его выходе определится выражением

$$x^*(t) = L\{x(t)\} = \sum_{n=-\infty}^{\infty} x(nT)\varphi(t-nT), \tag{3}$$

а передаточная функция определится прямым преобразованием Фурье входного и выходного процессов УВХ

$$X^*(j\omega) = Sa\left(\frac{\omega T}{2}\right)e^{-j\frac{\omega T}{2}}\sum_{l=-\infty}^{\infty} X\left(\omega - l\frac{2\pi}{T}\right), \tag{4}$$

где $X^*(j\omega)$ и $X(j\omega)$ – спектры выходного и входного процессов соответственно, l – номер отсчета при дискретизации, $T = 2\pi/\omega_s$ – период дискретизации.

Для обобщенного определения спектра сигнала на выходе системы, показанной на рис. 1 необходимо также определить спектр погрешности квантования.

Известно, что сигнал на выходе квантователя является суммой входного сигнала и шумов квантования, т.е.

$$x_q(t) = x(t) + z(t), \tag{5}$$

где $x(t)$ – сигнал на входе квантователя, $z(t)$ – шум квантования.

При этом ошибка квантования функционально связана со значением преобразуемого сигнала $x(t)$, т.е. она определяется как

$$z(t) = \psi[x(t)] = kq - x \text{ при } kq - 0.5q \le x \le kq + 0.5q, \tag{6}$$

где k – номер интервала квантования, q – шаг квантования [4].

Используя периодичность $\psi[x(t)]$, эту функцию можно разложить в ряд Фурье

$$\psi(x) = \sum_{i=1}^{\infty} b_i \sin\left(i\frac{2\pi}{q}x\right), \tag{7}$$

где i – номер гармоники, а

$$b_i = -\frac{2}{q}\int_{-0,5q}^{0,5q} x\sin\left(i\frac{2\pi}{q}x\right)dx = -\frac{4}{q}\int_{0}^{0,5q} x\sin\left(i\frac{2\pi}{q}x\right)dx = \frac{q}{i\pi}(-1)^i. \tag{8}$$

Исходя из результатов разложения, ошибка квантования

$$z(t) = \psi[x(t)] = \frac{q}{\pi}\sum_{i=1}^{\infty}\frac{(-1)^i}{i}\sin\left(i\frac{2\pi}{q}x(t)\right). \tag{9}$$

При подаче на вход квантователя уровня (КУ) гармонического воздействия $x(t) = A_m\cos(\omega_0 t)$ погрешность квантования на его выходе будет

$$z(t)=\psi[\tilde{x}(t)]=\frac{q}{\pi}\sum_{i=1}^{\infty}\frac{(-1)^i}{i}\sin\left(i\frac{2\pi}{q}\tilde{x}(t)\right)=\frac{q}{\pi}\sum_{i=1}^{\infty}\frac{(-1)^i}{i}\sin\left(\frac{2\pi iA_m\cos\omega_0 t}{q}\right). \quad (10)$$

Воспользуемся известным соотношением [5]:

$$\sin(x\cos\varphi)=2\sum_{m=1}^{\infty}(-1)^m J_{2m-1}(x)\cos(2m-1)\varphi, \quad (11)$$

где $x=\frac{2\pi iA_m}{q}$, $\varphi=\omega_0 t$.

Тогда шум квантования

$$\begin{aligned} z(t)&=\frac{2q}{\pi}\sum_{i=1}^{\infty}\frac{(-1)^i}{i}\sum_{m=1}^{\infty}(-1)^{m-1}J_{2m-1}\left(\frac{2\pi iA_m}{q}\right)\cos[(2m-1)\omega_0 t]= \\ &=\frac{2q}{\pi}\sum_{m=-\infty}^{\infty}F_m\cos(2m-1)\omega_0 t, \end{aligned} \quad (12)$$

где

$$F_m=\sum_{i=1}^{\infty}\frac{(-1)^{i+m-1}}{i}J_{2m-1}\left(\frac{2\pi iA_m}{q}\right), \quad (13)$$

где $m=1,2,3,\ldots$ – порядок функции Бесселя.

Спектр шума квантования при гармоническом воздействии на входе квантователя получается путем воздействия на шум квантования гармонического сигнала оператором прямого преобразования Фурье, т.е.

$$S_z(j\omega)=2q\sum_{m=1}^{\infty}F_m\{\delta[\omega-(2m-1)\omega_0]+\delta[\omega+(2m-1)\omega_0]\}. \quad (14)$$

Учитывая, что сигнал на входе КУ $x(t)=A_m\sin\omega_0 t$, в соответствии с [3] получаем в результате прохождения входного сигнала через квантователь в выходном сигнале системы новые гармонические составляющие, т.е.

$$z(t)=\sum_{i=-\infty}^{\infty}b_i\sin\omega_i t, \quad (15)$$

где $\omega_i=i\omega$, т.е. частоты гармоник шума квантования кратны частоте входного сигнала.

Тогда (см. рис. 1) сигнал на выходе квантователя

$$x_q^*(t)=A_m\sin\omega t+\sum_{i=-\infty}^{\infty}b_i\sin\omega_i t=A_m\sin\omega_0 t+\sum_{i=-\infty}^{\infty}b_i\sin i\omega_0 t=\sum_{n=1}^{\infty}b_n'\sin n\omega_0 t, \quad (16)$$

где

$$b'=\begin{cases} b' \text{ при } i\geq 2 \\ b_1+A_m \text{ при } i=1 \end{cases} \text{ и } b_{-i}'=-b_i'.$$

Воздействуя на обе части равенства (16) оператором прямого преобразования Фурье, получаем спектр сигнала на выходе КУ

$$X_q^*(j\omega) = \sum_{i=-\infty}^{\infty} b_i' \int_{-\infty}^{\infty} \sin \omega_0 it \cdot e^{-j\omega t} dt = \frac{\pi}{j} \sum_{i=1}^{\infty} b_i' \left[\delta(\omega - i\omega_0) + \delta(\omega + i\omega_0) \right]. \quad (17)$$

Учитывая (4) и (16) в случае гармонического воздействия $x(t)$ на вход системы АЦАП, спектр отклика определяется соотношением

$$\begin{aligned} Y(j\omega) &= Sa\left(\frac{\omega T}{2}\right) e^{-j\frac{\omega T}{2}} \sum_{l=-\infty}^{\infty} X_q\left(\omega - l\frac{2\pi}{T}\right) = \\ &= Sa\left(\frac{\omega T}{2}\right) e^{-j\frac{\omega T}{2}} \left\{ \sum_{l=-\infty}^{\infty} A_m \pi \left[\delta\left(\omega - l\frac{2\pi}{T} - \omega_0\right) + \delta\left(\omega - l\frac{2\pi}{T} + \omega_0\right) \right] + \right. \\ &\left. + 2q \sum_{m=1}^{\infty} F_m \sum_{l=-\infty}^{\infty} \delta\left[\omega - l\frac{2\pi}{T} - (2m-1)\omega_0\right] + \delta\left[\omega - l\frac{2\pi}{T} + (2m-1)\omega_0\right] \right\}, \end{aligned} \quad (18)$$

где l – номер отсчета при дискретизации, m – порядок функции Бесселя, q – шаг квантования.

Таким образом, при проведении операций дискретизации во времени и квантования по уровню, неизбежных в процессе преобразования АС в цифровую форму, возникают дополнительные частотные составляющие в спектре исходного сигнала и, следовательно, происходят искажения его формы.

Сигнал с частотой дискретизации в низкочастотной части спектра выходного дискретизованного и квантованного сигнала отсутствует. Вместе с тем, в его высокочастотной части будут присутствовать комбинационные суммарные и разностные частотные составляющие, определяемые частотой входного сигнала, частотой дискретизации во времени и частотой квантования по уровню. Таким образом, в спектре выходного сигнала будут присутствовать дополнительные частотные составляющие. При этом комбинационные частоты будут повторяться во всей полосе частот, занимаемой сигналом, а их уровень снижаться с ростом частоты. Это свидетельствует о мультифрактальном характере спектра сигнала, подвергнутого аналого-цифровому преобразованию, что вполне поясняется характером спектральной плотности для последовательности прямоугольных импульсов, которой описывается выборка сигнала при дискретизации [6].

Из соотношения (18) вытекает фрактальность оцифрованного аналогового сигнала, представленного во временной области. Достаточно просто посмотреть на статическую характеристику КУ (см. рис. 2), чтобы представить ступенчатость сигнала на выходе ЦАП, что и определяет фрактальность его характера.

Мультифрактальность структуры оцифрованного гармонического сигнала можно проиллюстрировать его вейвлет портретами, выполненными с разной степенью детализации, что и показано на рис. 3. На этом рисунке видны ветвления сигнала при разных уровнях детализации, что говорит о его мультифрактальности после оцифровки при записи на АЦЗЗ.

Но сам факт мультифрактальности оцифрованного гармонического сигнала еще не является основанием для применения фрактального подхода для идентификационных и диагностических экспертных исследований АЦЗЗ и ЦФ, поскольку для их проведения необходимы идентификационные признаки аппаратуры записи, обладающие рядом свойств. К ним относятся строгая индивидуальность, фиксируемость в ЦФ, повторяемость и достаточность (в количественном смысле с точки зрения статистической представительности).

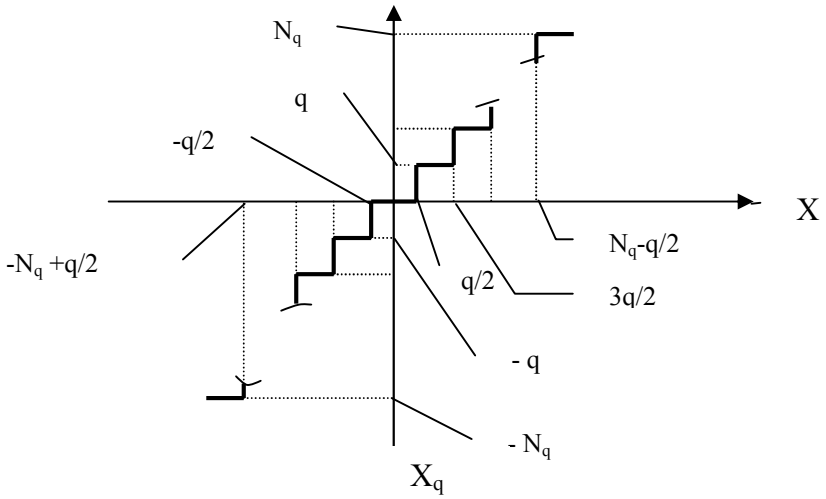
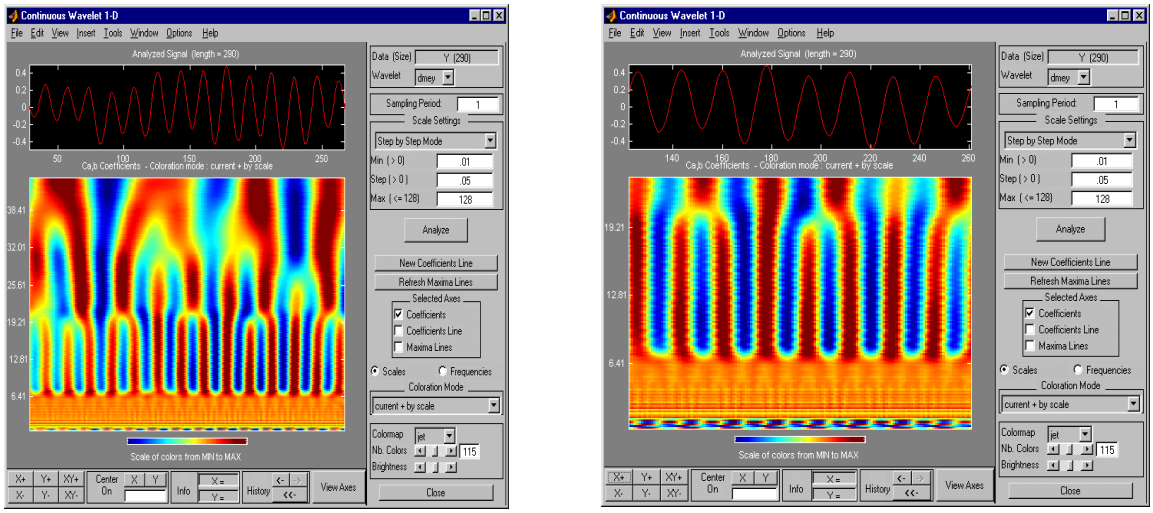


Рис. 2. Статическая характеристика квантователя по уровню



а б

Рис. 3. Вейвлет-портрет модуля сигнала в паузе между речевыми сигналами, записанного и воспроизведенного при ЧД 8 кГц с 16-разрядной оцифровкой, полученный в программе MatLab: а – при малой детализации; б – при увеличении детализации

Такими свойствами обладают некоторые паразитные параметры АЦЗЗ и, в частности, паразитные параметры отдельных узлов аналого-цифровых преобразователей (АЦП) и ЦАП.

К таким узлам относятся матрицы КУ, источники опорного напряжения и операционные усилители, входящие, как в АЦП любого типа, так и в ЦАП.

Общая погрешность АЦП, в которую обязательно входит КУ, можно определить, как

$$\overline{\delta}^2_{АЦП} = \overline{\delta}^2_0 + \overline{\delta}^2_{дин} + \overline{\delta}^2_q , \qquad (19)$$

где $\overline{\delta}^2_0 = \dfrac{\Delta^2_0}{U^2_{оп}}$ – средний квадрат статической погрешности АЦП, определяемой нестабильностью во времени и температуре его элементов, собственными шумами

аналоговых узлов, неточностью технологического изготовления отдельных узлов, в частности квантователя;

$\overline{\delta}^2_{дин} = \frac{\Delta^2_{дин}}{U^2_{ол}}$ – дисперсия динамической погрешности АЦП;

$\overline{\delta}^2_q = \frac{q^2}{12}$ – дисперсия погрешности квантования по уровню (для равномерно распределенной случайной величины);

$U_{оп}$ – значение опорного напряжения АЦП [7].

Проанализируем эти погрешности, с точки зрения возможности их использования в экспертизе при проведении идентификационных исследований АЦЗЗ и проверки подлинности ЦФ.

Динамическая погрешность не представляет для нас интереса по двум причинам:

- во-первых, в АЦП, использующем УВХ, такая погрешность практически отсутствует (а именно такие АЦП и применяются в АЦЗЗ и звуковых картах компьютеров);
- во-вторых, ее невозможно определить в проверяемой фонограмме, ибо неизвестна исходная форма входного сигнала до его преобразования.

Не интересен для нас и средний квадрат погрешности квантования по уровню, поскольку он определяется только значением МР преобразования.

Но статическая погрешность АЦП, да и ЦАП, включающая в себя технологические неточности изготовления КУ, представляет интерес для экспертных исследований [1; 2]. Поэтому ее следует подробно проанализировать. Среди погрешностей АЦП, относящихся к СХ КУ, которые потенциально могут быть использованные при экспертизе, относятся лишь две, а именно:

- *дифференциальная нелинейность* $\delta_{ЛД}$ статической характеристики (ДНСХ) т.е. отклонение разности двух аналоговых сигналов, которые отвечают соседним значениям кода, от значения единицы МР. Измеряется в процентном отношении от максимального значения преобразуемого сигнала, т.е. от значения $U_{оп}$. Превышение $\delta_{ЛД}$ значения ± 1 МР приводит к немонотонности статической характеристики преобразования;
- *немонотонность статической характеристики* (НСХ) $\delta_{НМ}$ преобразования – т.е. неидентичность знака приращения мгновенных значений изменения входного и выходного сигналов КУ хотя бы на одном из уровней квантования [7].

Эти неточности изготовления показаны на рис. 4.

С точки зрения возможности выявления при экспертизе интерес представляют и НСХ, и ДНСХ КУ [1; 2].

Рассмотрим искажения сигнала, происходящие из-за НСХ, при квантовании сигналов по уровню в АЦП и обратном преобразовании в ЦАП на примере наиболее характерных схем АЦП и ЦАП, стандартно используемых в АЦЗЗ и компьютерах для ввода/вывода АС.

СХ КУ образуется за счет применения резистивной матрицы типа R–2R. На эту матрицу подается опорное напряжение от стабильного источника, в результате чего на ней происходит деление величины этого напряжения пропорционально суммарной величине сопротивлений ее резисторов. Эта величина зависит от состояния ключей, к которым подключены резисторы с сопротивлением 2R, определяемым значением двоичного кода величины преобразуемого сигнала [1; 2; 7]. Именно разброс величин сопротивлений резисторов матрицы R–2R и разброс величин сопротивлений открытых каналов ключей, управляемых двоичным кодом, приводит к возникновению ДНСХ и НСХ преобразователя.

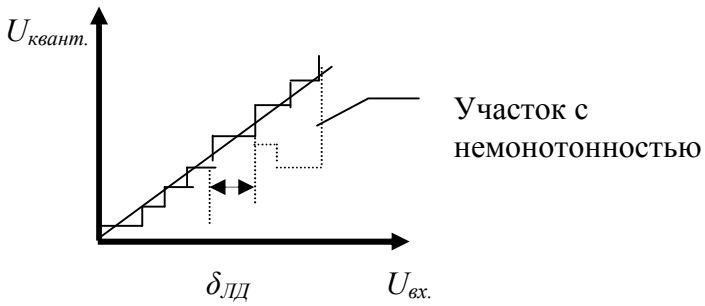


Рис. 4. Участки статической характеристики преобразования КУ с дифференциальной нелинейностью и немонотонностью

Этот разброс значений сопротивлений имеет индивидуальный характер и присутствует в каждом экземпляре АЦП и ЦАП. Поэтому величина интегральной предельной нелинейности характеристики преобразования (отклонения от прямой линии) всегда нормируется в АЦП и ЦАП и является их паспортной характеристикой. Она измеряется в процентах от диапазона входного (для АЦП) или выходного (для ЦАП) сигналов или в единицах МР и не должна превышать значения, определяемого классом преобразователя [1; 2; 7].

Проявление влияния НСХ и ДНСХ в выходных сигналах системы АЦАП выражается в возникновении в выходных сигналах искажений формы в виде всплесков или провалов. Они возникают на всех уровнях квантования, имеющих эти технологические отклонения, и поэтому могут присутствовать во всем диапазоне значений амплитуд преобразуемых сигналов, записываемых на ЦФ [2]. Поскольку таких уровней в любом КУ достаточно много, а их размещение на СХ преобразователя строго индивидуально, то создаваемые ими искажения могут служить идентификационными признаками АЦЗЗ. Проявление этих технологических дефектов во всем диапазоне величин уровней сигналов придает возникающим из-за них искажениям формы (а, следовательно, и спектра) сигналов мультифрактальный характер. К этим искажениям сигналов следует добавить флуктуационные изменения величины опорного напряжения и колебания нулевого уровня операционных усилителей, также носящих индивидуальный характер. Также на процесс передачи сигналов в АЦЗЗ влияют индивидуальные особенности работы усилителей, систем автоматической регулировки усиления, систем управления и т.п. вспомогательных узлов аппаратуры записи. Учитывая их малый уровень (для 16-разрядного АЦП прямой ИКМ они теоретически лежат в пределах минус 90-96 дБ), применение классических методов анализа (например, спектрального Фурье-анализа) неэффективны [2]. Поэтому выявление технологических неточностей изготовления и функционирования преобразователей (т.е. паразитных параметров АЦЗЗ), фиксируемых в ЦФ, требуют более чувствительных методов анализа. К таким методам относится фрактальный анализ, а характер проявления влияния паразитных параметров АЦЗЗ в выходных сигналах дает основания для применения фрактального подхода к построению аппаратуры для идентификационных и диагностических экспертных исследований материалов и средств цифровой звукозаписи.

Выводы

Показан фрактальный характер оцифрованного аналогового сигнала при записи звуковой информации на АЦЗЗ и фрактальный характер проявлений паразитных параметров аппаратуры записи, фиксируемых в информационных сигналах и сигналах шумов, записываемых на цифровых носителях. При этом показано, что малая величина этих проявлений требует применения чувствительных методов анализа. Учитывая

мультифрактальность этих проявлений, фрактальный подход является наилучшим решением задачи создания требуемого инструментария.

Список литературы

1. Рыбальский, О.В. Современные методы проверки аутентичности магнитных фонограмм в судебно-акустической экспертизе / О.В. Рыбальский, Ю.Ф. Жариков. – К. : Нац. акад. внутр. справ України, 2003. – 300 с.
2. Рибальський, О. В. Застосування вейвлет-аналізу для виявлення слідів цифрової обробки аналогових і цифрових фонограм у судово-акустичній експертизі / О. В. Рибальський. – К. : Нац. акад. внутр. справ України, 2004. – 167 с.
3. Семенов, О.Б. О нелинейных искажениях при аналого-цифро-аналоговом преобразовании сигналов / О.Б. Семенов // Техника средств связи. – Серия Техника радиовещательного приема и акустика. – 1981. – Вып. 1. – С. 77–86.
4. Баранов, Л.А. Квантование по уровню и временная дискретизация в цифровых системах управления / Л.А. Баранов. – М.: Энергоатомиздат, 1990. – 304 с.
5. Воронцов, А.А. Специальные функции задач теории рассеяния: Справочник / А.А. Воронцов, С.Д. Мировицкая. – М.: Радио и связь, 1991. – 200 с.
6. Денисенко, А.Н. Теоретическая радиотехника: Справочное пособие Ч. 1: Детерминированные сигналы (методы анализа) / А.Н. Денисенко, О.А. Стеценко. – М.: Издательство стандартов, 1993. – 215 с.
7. Федорков, Б.Г. Микроэлектронные цифро-аналоговые и аналого-цифровые преобразователи / Б.Г. Федорков, В.А. Телец, В.П. Дегтяренко. – М.: Радио и связь, 1984. – 120 с.

ОБГРУНТУВАННЯ ЗАСТОСУВАННЯ ФРАКТАЛЬНОГО ПІДХОДУ ДЛЯ СТВОРЕННЯ КОМПЛЕКСУ АПАРАТУРИ КОНТРОЛЯ СПРАВЖНОСТІ ФОНОГРАМ ПРИ ЕКСПЕРТИЗІ МАТЕРІАЛІВ ТА ЗАСОБІВ ЦИФРОВОГО ЗВУКОЗАПИСУ

О.В. Рибальський, В.І. Соловйов, В.В. Журавель

Національна академія внутрішніх справ,
пл. Солом'янська, 1, Київ, 03056, Україна; e-mail: rybalsky_ol@mail.ru

Показано фрактальний характер оцифрованого аналогового сигналу при запису інформації на апаратурі цифрового звукозапису та фрактальний характер проявів паразитних параметрів такої апаратури, що фіксуються в інформаційних сигналах і сигналах шумів, які записуються на цифрових носіях. При цьому показано, що мала величина цих проявів вимагає застосування чутливих методів аналізу. Показано, що, враховуючи мультифрактальність цих проявів, такий підхід забезпечує необхідну точність аналізу та є найкращим рішенням задачі створення необхідного інструментарію проведення експертиз матеріалів та засобів цифрового звукозапису.
Ключові слова: апаратура цифрового звукозапису, аналого-цифрове перетворення, паразитні параметри апаратури запису, цифрова фонограма, фрактальність

JUSTIFICATION OF FRACTAL APPROACH USING TO CREATE COMPLEX EQUIPMENT FOR CONTROL OF AUTHENTIC DIGITAL PHONOGRAMS AT THE EXAMINATION OF DIGITAL AUDIO MATERIALS AND TOOLS

O.V. Rybalsky, V.I. Solovyov, V.V. Zhuravel

National Academy of Internal Affairs,
1, Solomenskaya Sq., Kiev, 03056, Ukraine; e-mail: rybalsky_ol@mail.ru

Fractal character of the digitised analog signal at the record of information on the apparatus of digital record of sound and fractal character of displays of parasite parameters of such apparatus is shown, fixed in informative signals and signals noises recordable on digital carriers. It is thus shown that the small size of these displays requires application of sensible methods of analysis. It is shown that, taking into account the multifractal of these displays, such approach provides necessary exactness of analysis and is the best decision of task of creation of the required tool.
Keywords: apparatus of digital record of sound, digital phonogram, analog-digital transformation, parasite parameters of apparatus of record, fractal

UDC 519.725

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 249-258

THE FIBONACCI Q-MATRIX CODING METHOD

A.V. Sviridov, T.I. Petrushina

Odessa I.I. Mechnikov National University,
Dvoryanskaya st., 2, Odessa, 65026; e-mail: laestr.path@gmail.com

This paper presents the results of research, formalization and mathematical justification of the Fibonacci Q-matrix coding method. This method allows finding errors in the encoded message with high probability and correcting them in certain cases. The notes on algorithm implementation are given. The developed "block Q-matrix" method based on the standard method is described. The comparative analysis of the algorithms is presented.

Keywords: coding methods, Fibonacci numbers, error detection and correction, Q-matrix.

Introduction

The question of effective encoding and protection of the data in communication channels is rather important in the modern IT sphere.

Most of the known error detection and correction codes make it possible to restore single bits or combinations of bits [1-4], which is surely useful for many fields of application. However, the presented "Fibonacci Q-matix" coding method uses an entirely different approach: it allows restoring one of the predefined parts of the message – no matter how big is – given the condition that the damage affected only that part. The flaw is that errors, even small ones, in other parts of the message make the whole message unreadable.

However, there is an opportunity to develop new methods based on the standard Q-matrix method. The developed "block Q-matrix method" presented in the paper divides the message into fixed-length segments and applies the standard algorithm to them. That allows correcting errors scattered throughout the whole message. Also, in case method fails to restore some damaged segments, only that segments becomes unreadable.

The aim of the research is to study and evolve the coding methods based on the Fibonacci numbers.

The task of the research is to formalize, justify mathematically and analyze the Fibonacci Q-matrix coding method, analyze the ways to improve the algorithm, develop a program library which implements the method and make the characteristic of its work.

«Fibonacci Q-matrix» properties

Fibonacci Q-matrix is a following square 2×2 matrix [5]:

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \qquad (1)$$

Property 1. There is a property which connects the Q-matrix with the Fibonacci numbers:

$$Q^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix}, \tag{2}$$

where F_i is the Fibonacci number *i*.

This can be proved by the induction method.

$$n = 2: \; Q^2 = \begin{bmatrix} 1+1 & 1+0 \\ 1+0 & 1+0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

The elements of the Q^2 matrix are the corresponding Fibonacci numbers: $F_1 = F_2 = 1$, $F_3 = 2$.

Assume that $Q^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix}$ and calculate Q^{n+1}.

$$Q^{n+1} = Q^n \times Q = \begin{bmatrix} F_{n+1} + F_n & F_{n+1} + 0 \\ F_{n+1} + 0 & F_n + 0 \end{bmatrix} = \begin{bmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{bmatrix}.$$

The property is proved.

Property 2. $\det(Q^n) = (-1)$.

To prove it a determinant of Q-matrix can be calculated:

$$\det(Q) = 1 \times 0 - 1 \times 1 = -1.$$

Using a property stating that a determinant of the product of two square matrixes of the same size equals the product of their determinants [8]:

$$\det(Q^n) = (-1)^n. \tag{3}$$

Consequence. As $\det(Q^n) = F_{n+1}F_{n-1} - F_n^2 = (-1)^n$, then for each three consecutive Fibonacci numbers the following is true:

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n. \tag{4}$$

Property 3. With the help of the cofactor method [8] the following presentation of the Q^{-n} matrix can be acquired:

$$Q^{-n} = (-1)^n \begin{bmatrix} F_{n-1} & -F_n \\ -F_n & F_{n+1} \end{bmatrix} = \begin{bmatrix} (-1)^n F_{n-1} & (-1)^{n+1} F_n \\ (-1)^{n+1} F_n & (-1)^n F_{n+1} \end{bmatrix}. \tag{5}$$

The standard Fibonacci Q-matrix coding method

Let's look at the standard Fibonacci Q-matrix coding method [5].

Assume there is a message M, which can be divided into 4 parts, represented in numeric values as a 2×2 matrix. Each element m_i is a nonnegative integer number

$$M = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix}. \tag{6}$$

Encoding. An encoded message M' can be fetched in the following way.

$$M' = M \times Q^n = \begin{bmatrix} F_{n+1}m_1 + F_n m_2 & F_n m_1 + F_{n-1} m_2 \\ F_{n+1}m_3 + F_n m_4 & F_n m_3 + F_{n-1} m_4 \end{bmatrix} = \begin{bmatrix} m_1' & m_2' \\ m_3' & m_4' \end{bmatrix}. \tag{7}$$

As F_i and m_i are nonnegative integer numbers, then the elements m_i' of the encoded matrix M' are nonnegative as well.

The power of the Q-matrix can be a random positive number and it serves as an encryption key of this method.

Decoding. The primal message M'' can be fetched from the decoded message M' in the following way.

$$M'' = M' \times Q^{-n} = \begin{bmatrix} m_1'' & m_2'' \\ m_3'' & m_4'' \end{bmatrix} = \begin{bmatrix} (-1)^n F_{n-1}m_1' + (-1)^{n+1} F_n m_2' & (-1)^{n+1} F_n m_1' + (-1)^n F_{n+1} m_2' \\ (-1)^n F_{n-1}m_3' + (-1)^{n+1} F_n m_4' & (-1)^{n+1} F_n m_3' + (-1)^n F_{n+1} m_4' \end{bmatrix}. \tag{8}$$

We need proof that $M'' = M$. For this we must prove that each m_i'' equals m_i:

$$m_1'' = (-1)^n F_{n-1}m_1' + (-1)^{n+1} F_n m_2' =$$

$$= (-1)^n F_{n-1}F_{n+1}m_1 + (-1)^n F_{n-1}F_n m_2 + (-1)^{n+1} F_n^2 m_1 + (-1)^{n+1} F_n F_{n-1} m_2 =$$

$$= (-1)^n F_{n-1}F_{n+1}m_1 - (-1)^n F_n^2 m_1 + (-1)^n F_n F_{n-1} m_2 - (-1)^n F_n F_{n-1} m_2 =$$

$$= (-1)^n m_1 (F_{n-1}F_{n+1} - F_n^2) = (-1)^n \times (-1)^n m_1 = m_1;$$

$$m_2'' = (-1)^{n+1} F_n m_1' + (-1)^n F_{n+1} m_2' =$$

$$= (-1)^{n+1} F_n F_{n+1} m_1 + (-1)^{n+1} F_n^2 m_2 + (-1)^n F_{n+1} F_n m_1 + (-1)^n F_{n+1} F_{n-1} m_2 =$$

$$= (-1)^n F_{n+1} F_n m_1 - (-1)^n F_{n+1} F_n m_1 + (-1)^n F_{n+1} F_{n-1} m_2 - (-1)^n F_n^2 m_2 =$$

$$= (-1)^n m_2 (F_{n+1}F_{n-1} - F_n^2) = (-1)^n \times (-1)^n m_2 = m_2;$$

$$m_3'' = (-1)^n F_{n-1}m_3' + (-1)^{n+1} F_n m_4' =$$

$$= (-1)^n F_{n-1}F_{n+1}m_3 + (-1)^n F_{n-1}F_n m_4 + (-1)^{n+1} F_n^2 m_3 + (-1)^{n+1} F_n F_{n-1} m_4 =$$

$$= (-1)^n F_{n-1}F_{n+1}m_3 - (-1)^n F_n^2 m_4 + (-1)^n F_n F_{n-1} m_3 - (-1)^n F_n F_{n-1} m_4 =$$

$$= (-1)^n m_3 (F_{n-1}F_{n+1} - F_n^2) = (-1)^n \times (-1)^n m_3 = m_3;$$

$$m_4'' = (-1)^{n+1} F_n m_3' + (-1)^n F_{n+1} m_4' =$$

$$= (-1)^{n+1} F_n F_{n+1} m_3 + (-1)^{n+1} F_n^2 m_4 + (-1)^n F_{n+1} F_n m_3 + (-1)^n F_{n+1} F_{n-1} m_4 =$$

$$= (-1)^n F_{n+1}F_n m_3 - (-1)^n F_{n+1}F_n m_4 + (-1)^n F_{n+1}F_{n-1} m_3 - (-1)^n F_n^2 m_4 =$$

$$= (-1)^n m_4 (F_{n+1}F_{n-1} - F_n^2) = (-1)^n \times (-1)^n m_4 = m_4.$$

Error detection and correction

One of the features of this coding method is the possibility of detecting and correcting the errors. From (3) it can be known that:

$$\det(M') = (-1)^n \det(M).$$

By passing the value of the determinant with the message, we can allow the receiver to check whether it matches the determinant of the received matrix before starting to decode. If the message is corrupted, one or more of the elements of the matrix will differ, and the determinants won't match.

As $\det(M') = m_1'm_4' - m_2'm_3'$, we can restore the corrupted part of the message, if the determinant and the other three parts are unharmed.

$$\begin{aligned} m_1' &= \frac{m_2'm_3' + \det(M')}{m_4'},\ m_2' = \frac{m_1'm_4' + \det(M')}{m_3'}, \\ m_3' &= \frac{m_1'm_4' + \det(M')}{m_2'},\ m_4' = \frac{m_2'm_3' + \det(M')}{m_1'}. \end{aligned} \tag{9}$$

If it is not obvious which part of the message was corrupted, the correction for each part can be calculated. The corrected part must be integer, so with high probability there will be a single matching result.

If no correction provides integer results, then two or more parts of the message were damaged, and the restoration is likely impossible.

Example:

$$M = \begin{bmatrix} 65 & 115 \\ 104 & 97 \end{bmatrix},\ Q^6 = \begin{bmatrix} 13 & 8 \\ 8 & 5 \end{bmatrix},$$

$$M' = M \times Q^6 = \begin{bmatrix} 1765 & 1095 \\ 2128 & 1317 \end{bmatrix},$$

$$\det(M') = -5655.$$

Let's add a «corruption» into the second part of the message, changing it to 1112:

$$M^{err} = \begin{bmatrix} 1765 & 1112 \\ 2128 & 1317 \end{bmatrix},$$

$$\det(M^{err}) = -41831 \neq -5655.$$

Let's assume that the first part is damaged, and try to correct it:

$$m_1' = \frac{m_2' m_4' + \det(M')}{m_4'} = \frac{1112 \times 2128 - 5655}{1317} = 1792.47 .$$

The result is non-integer. Let's assume the second part was damaged:

$$m_2' = \frac{m_1' m_4' + \det(M')}{m_3'} = \frac{1765 \times 1317 - 5655}{2128} = 1095 .$$

Let's assume the correction is true:

$$M'' = M' \times Q^{-n} = \begin{bmatrix} 1765 & 1095 \\ 2128 & 1317 \end{bmatrix} \times \begin{bmatrix} 5 & -8 \\ -8 & 13 \end{bmatrix} = \begin{bmatrix} 65 & 115 \\ 104 & 97 \end{bmatrix}.$$

So, we have restored the message, where one of the parts was corrupted.

Notes on algorithm implementation

The following notes are given for the implementation of the Fibonacci Q-matrix coding method.

1. The encrypted parts of the message along with the determinant are translated into the Fibonacci code [7]. Because each Fibonacci code contains only two consecutive 1-digits, which are located at the end of the code, it can help to determine which parts of the message were corrupted. Also, if a binary data transmission is used, we won't have to translate the encrypted values into the binary numeral system, because the Fibonacci code consists only of 0- and 1-digits. In this way we avoid unnecessary calculations while transforming numbers from one numeral system into another.
2. The encrypted parts are brought to the common length by adding 0-bits to the lesser parts. The message is increased by 1-2 bytes in most cases this way. In computer implementation developed for this method the common length of each part is increased till it becomes divisible by 8, so that each byte has only one corresponding part of the message.
3. The writing of encoded data is done in the following way. The size of the determinant in bytes and its sign are written. The determinant is written. The decoded parts are written.
4. The reading of encoded data is done in the following way. The size of the determinant and its sign are read. The determinant is acquired by reading the number of bytes equal to its size. The rest of the message is divided into 4 equal-sized parts, each one has its 0-bits deleted from the end until the first 1-bit is met.
5. The data read is checked for the correct Fibonacci codes. If the determinant is damaged, the correcting of other errors will be useless. If the Fibonacci code of the determinant is correct, and one part of the message is damaged, that part is restored. If the Fibonacci codes of the determinant and the encoded parts are correct, but the received determinant doesn't match the calculated determinant, an attempt to restore each part is made, and an integer result is treated as a correct one. If no attempt gave the integer result, the message is decoded with errors.

Block Q-matrix algorithm

The following algorithm has been developed on the basis of studied properties of Fibonacci Q-matrix and the Fibonacci code.

Assume there is a message M, which can be divided into segments 4 bytes each:

$$M = \{b_{11}b_{12}b_{13}b_{14}, b_{21}b_{22}b_{23}b_{24}, ..., b_{n1}b_{n2}b_{n3}b_{n4}\}, 0 \le b_{ij} \le 255. \quad (10)$$

Let's encode the message M into M' by applying the standard Fibonacci Q-matrix method for each quad of bytes

$$M' = \{B'_1, B'_2, ..., B'_n\}, \quad (11)$$

where B'_i is a group of bytes $b_{i1}b_{i2}b_{i3}b_{i4}$ encoded with the help of the standard Fibonacci Q-matrix method.

The power of Q-matrix $n = 5$ is chosen for the implementation of this method. The implementation of the standard Fibonacci Q-matrix method is done as described previously.

To decode the message, the standard Q-matrix method needs to be applied for each encoded segments B'_i. However, in order to separate B'_i correctly in the encoded message, the length of these segments must be a fixed length of k bytes. To find it, we'll calculate the largest possible values of determinant and the elements of the encoded matrix.

The biggest value of determinant in case $n = 5$ is achieved when $\det(M') = m'_1 m'_4 - m'_2 m'_3 = 255 \times 255 - 0 \times 0 = 65025$. Translated into Fibonacci code it will be written as 100001000101010000001011; the length of this presentation is 24 bits, i.e. 3 bytes.

The biggest value of the element of the encoded matrix is achieved when $m'_i = F_{n+1} m_i + F_n m_{i+1} = 8 \times 255 + 5 \times 255 = 3315$. Translated into Fibonacci code it will be written as 001010100100010011; the length of this presentation is 18 bits. All four elements of the matrix will have their collective length no more than 9 bytes.

Therefore, considering that at least 1 more byte is needed for the sign and the size of determinant, we can say that each encoded segment B'_i won't exceed $k = 13$ bytes. If the segment is smaller, the missing bits are filled with the values which can be ignored by the decoding algorithm (for example, 0-bits).

So, the optimal block Q-matrix method produces 104 bits (13 bytes) of encoded message from each 32 bits (4 bytes) on the initial message.

It is worth noting that the computer implementation, the work of which is described later in the paper, uses the implementation of the standard Q-matrix method with more massive amount of output data (providing divisibility by 8 for encoded parts), and each 4 bytes are encoded into 18 bytes.

Interleaving modification

The described algorithm can handle the errors in different parts of the message, restoring each 4-byte segment if the damage affected only one byte of this segment. However, this means that the errors must affect only single bytes, not consecutive ones. In any case, the undamaged segments of the message will remain readable.

To increase the effectiveness, the algorithm is modified with the help of interleaving [9]: the encoded bytes are shuffled in certain order, and the consecutive damaged bytes will affect different 4-byte segments in the end.

In the computer implementation, the first bytes of each segment are placed at the beginning, then the second ones, and so on, till the last bytes. This process allocates bytes of each segment equally, and the errors will need to "guess" several bytes of each segment to prevent it from restoration. Even if it happens, only the 4 bytes of this segment will become unreadable in the decoded message.

Results of the numerical experiments

A C++ library has been developed, which performs encoding and decoding with the help of the standard Fibonacci Q-matrix method and the block method, applying the notes listed above.

The ability to correct errors has been tested for each algorithm. The given number of byte damages has been generated for the message: in one case the single damaged bytes were randomly scattered throughout the message, in another case the consecutive bytes were damaged in the random position. For each number of errors 100 tests were handled, which calculated the number of completely restored message cases.

The message was 576 bytes long. For the standard Fibonacci Q-matrix method ($n = 5$) the encoded message was 1250 bytes. For the block Q-matrix method it was 2592 bytes. The test results are shown on the diagrams on fig. 1 and fig. 2. The comparative analysis of the work of the algorithms is presented in table 1.

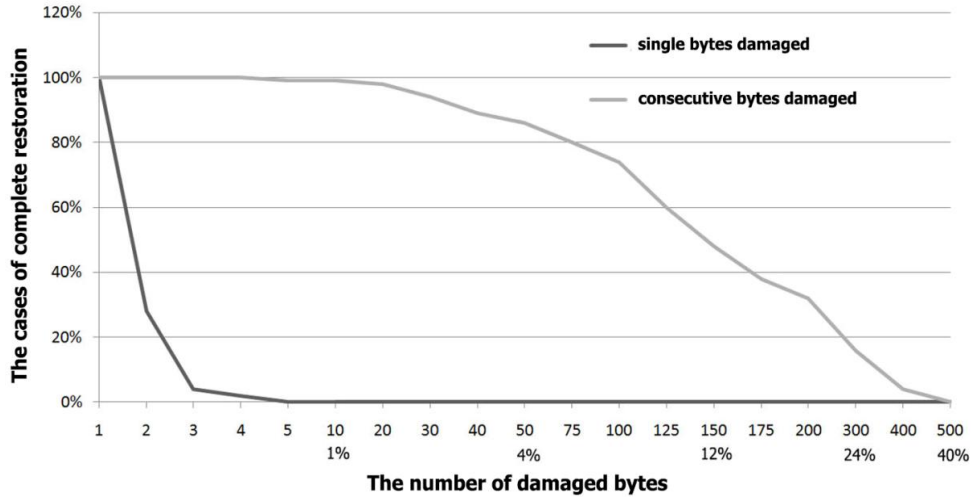


Fig. 1. The cases of complete restoration of the initial message in the standard Fibonacci Q-matrix method

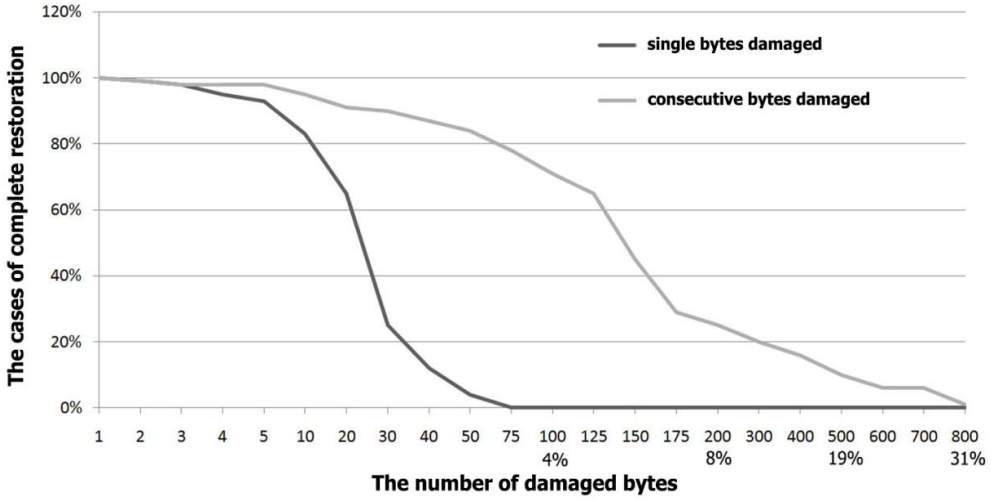


Fig. 2. The cases of complete restoration of the initial message in the block Q-matrix method

The standard Fibonacci Q-matrix method survives the considerable amount of consecutive damaged bytes (being completely ineffective in case more than 40% of the message is damaged), although it is useless in case of single damaged bytes in different parts of the message. In that case the half or the whole message becomes unreadable.

The block Q-matrix method handles the consecutive damaged bytes a little worse (being completely ineffective in case more than 31% of the message is damaged), and allows the restoration of the scattered single error bytes (up to 4% if the message). Also, even if it fails to correct, only the failed damaged groups remain unreadable.

Table 1.

Comparative analysis of the methods

	Standard Ficonacci Q-matrix ($n = 5$)	Block Q-matrix method
Ratio of the encoded message size to the initial message size	2.1622	4.5
Code rate	0.4625	0.22
Complete restoration of consecutive error bytes	Up to 40% damage of the message	Up to 31% damage of the message
Complete restoration of single error bytes	Less than 1% damage of the message	Up to 4% damage of the message
If failed to correct all errors	Half or all of the message is unreadable	Only the failed 4-byte segments remain unreadable
Time to encode 4000 bytes	1.13 sec	1.1 sec
Time to decode 4000 bytes	4.65 sec	0.75 sec

The standard Fibonacci Q-matrix method doesn't belong to linear block or convolutional codes [2], which are widespread in coding theory for error detection and correction, or any other category known to the authors. In terms of this it is difficult to compare it with the work of other codes.

While the widespread codes make it possible to correct single bits and their combinations, the Fibonacci Q-matrix allows restoring the elements of the matrix, the size of which is theoretically unlimited. The implementation of this method, however, may require arbitrary-precision arithmetic.

This code also allows encrypting the data, using the power of the Q-matrix as an encryption key. This gives the code an additional advantage in cryptography.

The use of this method might prove useful for the **digital signature** technology. If both the power of the Q-matrix and the determinant are present, the validation of the document can be verified by comparing the calculated determinant with the known one.

The results of numerical experiments made it possible to calculate the rate of the code. The **code rate** [3] is a relation of bits of «useful» information (primal message) to the number of bits of redundant information (encoded message).

Using the developed library it was experimentally proved that for $n = 5$, while the size of the message tends to infinity, the rate of the standard Fibonacci Q-matrix code approaches approximately 0.4625. Increasing the power n for Q-matrix will decrease the code rate, therefore increasing the size of the encoded message compared to the primal.

For comparison, the rate of Hamming (3.1) - code is 0.333, and Hamming (7.4) - code is 0.571 [3].

The standard Fibonacci Q-matrix coding method is quite useful if the damage took place in only one part of the message, even if the whole part was corrupted. However, the damage in more parts will make half of the message, or the whole message, unreadable and uncorrectable. Interleaving isn't effective for this method, unless we can predict which bytes will be damaged beforehand.

Nevertheless, given that $\lim_{n\to\infty}\frac{F_{n+1}}{F_n}=\varphi$, where F_n is a Fibonacci number n, and φ is the Golden ratio, there are ways to correct errors even in two or three elements in the matrix [5,6]. However, these methods require additional study.

The developed block Q-matrix method relates to the block codes [10]. The code rate is $\frac{4}{18}=0.22$ in computer implementation, and the optimized algorithm's code rate can reach $\frac{4}{13}=0.307$. In both cases the number of redundant bits exceeds the standard method.

With the help of interleaving this code can correct the consecutive damaged bytes a little worse than the standard Fibonacci Q-matrix method. However, at first, it leaves the unharmed code readable, and, at second, is able to restore single damaged bytes in different parts of the message even without interleaving.

Both methods are inferior to most of the modern codes in speed and amount of calculations, but their correcting ability can be high for specific types of damage.

Conclusion

This paper presents the Fibonacci Q-matrix coding method, which allows detecting and correcting data errors. The authors have formalized, systematized and justified the existing researches in this field. The "block Q-matrix" method has been developed on the basis of the standard algorithm. The comparative analysis of these methods has been made.

The advantage of these methods is that they allow correcting considerably large information units, the size of which is theoretically unlimited, instead of single bits and their combinations. A matrix element that can be an integer of unlimited value is a minimal information unit for the Fibonacci Q-matrix coding method.

However, both methods are inferior to most of the modern codes in speed and amount of calculations. The standard method also cannot restore the message, if the damage is out of limits of predefined area. This disadvantage has been amended in the block Q-matrix method.

References

1. Сидельников, В.М. Теория кодирования. Справочник по принципам и методам кодирования / В.М. Сидельников. - Мех-мат, МГУ, 2006 г. – 289 с.
2. Никитин, Г.И. Сверточные коды: Учеб. пособие / Г.И. Никитин. - СПбГУАП. СПб., 2001. – 80 с.
3. W. Cary Huffman. Fundamentals of Error-Correcting Codes. / W. Cary Huffman, Vera Pless. - Cambrige University Press, 2003. – 665 p.
4. Цымбал, В.П. Теория информации и кодирование : Учебник. – 4-е изд., перераб. и доп. / В.П. Цымбал. – К. : Вища шк., 1992. – 263 с.: ил.
5. Stakhov, A.P. Fibonacci matrices, a generalization of the "Cassini formula", and a new coding theory / A.P. Stakhov // Chaos, Solitons and Fractals. – 2006. – No. 30. - pp. 56-66.
6. Стахов, А.П. Тьюринг, филлотаксис, математика гармонии и «золотая» информационная технология. Часть 2. «Золотая» информационная технология [Электронный ресурс] / А.П. Стахов // Электронное периодическое издание «Академия Тринитаризма». – с. 46-51: – Режим доступа: http://www.trinitas.ru/rus/doc/0232/004a/02321090.pdf

7. Aviezri S. Fraenkel. Robust Universal Complete Codes for Transmission and Compression / Aviezri S. Fraenkel, Shmuel T. Klein // Discrete Applied Mathematics. – 1996. – Volume 64, Issue 1. - Pages 31-55.
8. Мальцев, А.И. Основы линейной алгебры / А.И. Мальцев. – М.: Наука, 1975. – 400 с.
9. What is Interleaving? [Electronic resource] // Techopedia – Режим доступа: https://www.techopedia.com/definition/5683/interleaving
10. Intuitive Guide to Principles of Communications [Electronic resource] // Complex to Real – Режим доступа: http://complextoreal.com/wp-content/uploads/2013/01/block.pdf

МЕТОД КОДУВАННЯ НА ОСНОВІ ФІБОНАЧЧІЄВОЇ Q-МАТРИЦІ

А.В. Свірідов, Т.І. Петрушина

Одеський національний університет ім. І.І. Мечникова,
вул. Дворянська, 2, Одеса, 65026; e-mail: laestr.path@gmail.com

В статті викладаються результати дослідження, формалізації та математичного обґрунтування метода кодування на основі «фібоначчієвої Q-матриці». Даний метод кодування дозволяє з високою ймовірністю виявляти помилки в закодованому повідомленні і виправляти їх у певних випадках. Наведені зауваження щодо реалізації алгоритму. Розроблений «блоковий» алгоритм Q-матриці на основі базового. Наведений порівняльний аналіз роботи алгоритмів.
Ключові слова: методи кодування, числа Фібоначчі, виявлення та виправлення помилок, Q-матриця.

МЕТОД КОДИРОВАНИЯ НА ОСНОВЕ ФИБОНАЧЧИЕВОЙ Q-МАТРИЦЫ

А.В. Свиридов, Т.И. Петрушина

Одесский национальный университет им. И.И. Мечникова,
ул. Дворянская, 2, Одесса, 65026; e-mail: laestr.path@gmail.com

В статье излагаются результаты исследования, формализации и математического обоснования метода кодирования на основе «фибоначчиевой Q-матрицы». Данный метод кодирования позволяет с высокой вероятностью обнаруживать ошибки в закодированном сообщении и исправлять их в определённых случаях. Приведены замечания к реализации алгоритма. Разработан «блочный» алгоритм Q-матрицы на основе базового метода. Приведен сравнительный анализ работы алгоритмов.
Ключевые слова: методы кодирования, числа Фибоначчи, обнаружение и исправление ошибок, Q-матрица.

UDC 004.934.2

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 259-269

PRONUNCIATION QUALITY ASSESSMENT BY COMPARISON WITH SAMPLE

G.A. Dobrovolsky, O.A. Todoriko, N.G. Keberle

Zaporizhzhya National University,
66, Zhukovskogo str., Zaporizhzhya, 69600, Ukraine; e-mail: gen.dobr@gmail.com

The task of pronunciation quality assessment by comparison with a reference example usually requires large training set of such examples. Unfortunately, such sets even for widely used human languages are rare. Most annotated speech corpora contain examples of mispronunciation, without reference utterance examples. In this paper we propose an approach to assess pronunciation quality by comparison with a reference example given small set of reference utterance examples. Dynamic time warping with silence model allows to compare reference utterance by teacher/native speaker with student's utterance and to obtain feature sets describing mispronunciation at word and phone level. Student's utterance is then classified as correct or mispronounced using bagging method.

Keywords: computer-aided pronunciation training, language learning, mispronunciation detection, dynamic time warping, bagging.

Problem statement

Computer-Aided Language Learning (CALL) systems [1-2] have gained new attention nowadays, as speech recognition technologies (SRT) widely used in human-computer interaction with search engines can be adapted to distant language learning. Computer-Aided Pronunciation Training (CAPT) systems respond to the demand of SRT client to be understood. There are various technologies to teach reading, listening, and grammar, to improve and expand vocabulary. At the same time, oral speech and correct pronunciation training are harder to automate, and are more to the research, than to the technology, however several pronunciation assessment services already exist [3-4].

The straightforward way to assess pronunciation is to use automatic speech recognition (ASR) system. Current ASR systems are based on supervised machine learning techniques. Training of ASR system requires a large corpus of annotated (manually/automatically) reference data – audio files storing sound of a phoneme/word/phrase/text utterance of a person in a given language. Such a prerequisite causes a bottleneck of direct adoption of ASR system to pronunciation assessment – necessary datasets are only available for the most used languages [5-6], whereas there are 7102 languages spoken in the world [7]. One more bottleneck of ASR system adoption is the vocabulary used. Sufficient datasets are available only for the most common, everyday topics (e.g. British English corpus WJSCAM0 [8] for news). Specific terminology words, professional slang, rare vocabulary words will be substituted by similarly sounding words.

Therefore, there is a need of exploring alternative approaches that do not require large reference data, and do not perform extra operations, e.g. do not perform full ASR.

Related Work

At the early stages, pronunciation quality assessment was performed for the whole phrase with the help of hidden Markov model. Obtained results did not depend on a teacher,

but did not point to the error type [10-12]. To overcome this difficulty the researchers focused on various ways of detection of "problematic" phonemes extracted from utterance examples, and their classification as pronounced correctly or mispronounced [13-16]. The results of such approach have shown increased precision of pronunciation assessment. Approaches to extend ASR system with typical pronunciation errors [17] lead to increased quality of assessment. However, they require a-priori sets of typical pronunciation errors, inherent to language learners of different nationalities. As a result, only those typical errors could be assessed, i.e. person-specific utterances within the same nationality are not taken into account.

Recently, comparison-based approaches to mispronunciation detection [9], [18] appear, attempting to avoid usage of a full ASR system. They differ in the way how classification is performed, and how feature sets of utterances are obtained. In [9] SVMs are used for classification, and utterance feature sets extracted with Gaussian posteriograms (GP) and Mel frequency cepstral coefficients (MFCC) are compared. In [18] classification is done with Gaussian mixed models (GMMs), and deep neural networks (DNNs) are used for extraction of feature sets.

Aim of the paper

In this paper, we propose to use bootstrap aggregating (bagging) algorithm to improve classification of example utterances, taming the problem of small reference datasets. The approach is inspired by the previous success in application of dynamic time warping (DTW) with silence model [9] to mispronunciation detection. However, in [9] support vector machines (SVM) are used in classification of example utterances, which require a large reference dataset for classifier training. Bagging algorithm allows starting pronunciation assessment with a small reference dataset, incrementally adding new references. Such an environment is inherent to a socially-oriented on-line language learning system, where teachers/native speakers can add their utterances of sample phrases, and the system reclassifies students' pronunciation accordingly.

Results

Mispronunciation quality assessment simplified method is based on the following assumptions:

- if a phrase uttered by a student similar to a phrase uttered by a teacher, then the student has a good pronunciation;
- similarity criterion is a distance function between correspondent features' values of conditional phonemes utterance by teacher and student;
- uttered phrase is split into conditional phonemes in assumption that features of the sound change essentially between different conditional phonemes, rather than inside one conditional phoneme;
- silence and pauses between words are not taken into account.

Claimed that student's pronunciation is well-trained if his/her phrase is similar to a teacher phrase. This allows at the beginning only a small set of teacher sample utterances. The benefit of such an approach is its simplicity, incremental pronunciation quality assessment improvement as more correctly pronounced samples (e.g. by students) are put into a sample set.

Sound file preparation stage

Sound file preparation stage is traditional for speech recognition (see Fig.1). First, low-frequency component is removed as not important for speech recognition by means of signal smoothing:

$$x_k = \alpha \cdot x_k + (1-\alpha) \cdot x_{k\text{-}1}, \ |\alpha| < 1, \tag{1}$$

where α - is a parameter, regulating the level of smoothing.

Then, the signal amplitude is mapped to the segment [-1, 1] and the signal is split into frames, F (see Fig. 1). Frames are overlapping fragments of the sound file, having length depending on the frequency of the sound. In our case, as sound was recorded at 22 kHz, and fast Fourier transformation requires 2^n discrete signal values in a frame, frame length was 23 ms (512 values), and overlapping window was 11 ms (256 values).

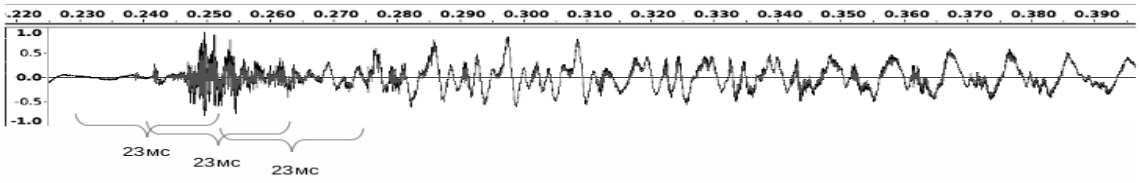


Fig. 1. Dependency of a sound signal on time and frame size explanation

For each frame t of F we calculated MFCC [19] feature set, and additionally energy, entropy, and their first and second derivatives, resulting in a feature set f_t of 42 features.

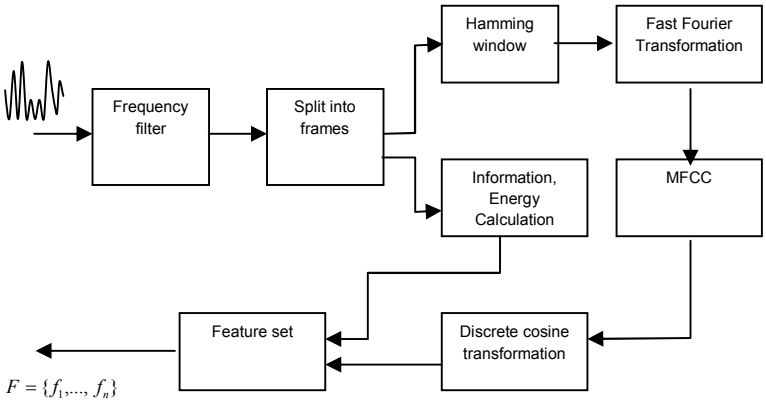


Fig. 2. Sound file preparation steps

For a frame t its energy is evaluated as biased estimate of the variance of the input signal:

$$E_t = \frac{1}{N}\sum_{k=0}^{N-1}\left(x_k - \bar{x}_t\right)^2, t = \overline{1,T}, \tag{2}$$

where $\bar{x}_t$ - is an average value of signal in a frame t, N – quantity of amplitude values in a frame, t – frame number, T – total quantity of frames.

To evaluate entropy we obtain amplitude sweep $[a_{\min}, a_{\max}]$, the resulted segment is split into R parts $[a_0, a_1]$, $[a_1, a_2]$, … , $[a_{R-1}, a_R]$, where $a_0 = a_{\min}$ and $a_R = a_{\max}$, and for each

frame we calculate the quantity of amplitudes, belonging to the segment and obtain frequency histogram. Then, using Shannon's definition of information entropy, we obtain:

$$I = -\sum_{i=1}^{R} p_i \ln(p_i), \tag{3}$$

where p_i – is a signal amplitude share, belonging to the segment $[a_{i-1}, a_i]$.

Usage of Mel Frequency Cepstral Coefficients (MFCC) is one of the standard techniques to obtain features of a sound in ASR systems [19]. MFCC features are obtained with the help of a set of frequency filters, taking into account the peculiarity of a human ear to have different sensibility in different parts of the audio spectrum – almost linear for frequencies below 1 kHz and logarithmic for higher frequencies.

At the first step we calculate signal energy logarithm upon application of each filter

$$S(t,m) = \ln\left(\sum_{n=0}^{N-1} \left|X(t,n)\right|^2 H(m,n)\right), t = \overline{1,T}, m = \overline{0,M-1}, \tag{4}$$

where $X(t,n)$ – is a n-th component of Fourier image in the frame t, $H(m,n)$ – is a n-th component of m-th Mel-Frequency filter, N – window size, M – predefined quantity of Mel filters, T – quantity of frames. Usually in ASR systems $M = 20$, but $M = 12$ is also acceptable.

At the second step we perform discrete cosine transformation of $S(t,m)$ values:

$$c(t,m) = \sum_{m_1=0}^{M-1} S(t,m_1) \cos\left(\frac{m(m_1 - 0{,}5)\pi}{M}\right), t = \overline{1,T}, m = \overline{0,M-1}, \tag{5}$$

We also calculate first and second derivatives to take into account human ear reaction to the spectrum changes in time:

$$\begin{aligned} dc(t,m) &= c(t+2,m) - c(t-2,m), \\ d^2c(t,m) &= c(t+1,m) - c(t-1,m). \end{aligned} \tag{6}$$

The same derivatives are calculated for energy E and information entropy I as well.

Values (2), (3), (5), (6) form a feature set f_t for each frame t, resulting in a feature set of 42 features

$$f_t = \left\langle c(t,m), dc(t,m), d^2c(t,m), E, dE, d^2E, I, dI, d^2I \right\rangle. \tag{7}$$

Preparation of samples

To detect silence we seek frames with minimal information entropy values [20] that are considered as noise. Frames contain informative speech, if its Mahalanobis distance to any of frames considered as noise exceeds a given threshold [21].

Sequence of frames, *F*, is then separated into conditional phonemes, by pair wise comparison of Euclidean distances between correspondent MFCC values of each two neighbor frames $f_t, f_{t=1}$. We assume that sound characteristics change essentially between two different conditional phonemes, rather than within the same phoneme. To calculate Euclidean distances we use MFCC features of the same nature (energies, frequencies etc). Conditional

phonemes set may not coincide with the traditional sound set of the language, and for each specific phrase may differ.

Separation of a sample phrase into words may be performed manually or with the help of some ASR system.

Comparison with sample

After sample and student utterances are prepared as shown in Fig.3, DTW algorithm is used to align two frame sets (see Fig.3).

Given sample $FT = \{ft_1,...,ft_n\}$ and student's $FS = \{fs_1,...fs_m\}$ frame sets, DTW distance matrix Φ is constructed as

$$\Phi(i,j) = D(ft_i, fs_j), i = 1..n, j = 1..m,$$

where D – is Euclidean distance between sample/student frames.

As student utterance is uncertain, with pauses, we use DTW with modified distance function, taking silence frames into account, as in [9].

Silence vector ϕ_{sil} keeps average distances from each frame of FS to each frame of FT, marked as silence,

$$\phi_{\text{sil}}(j) = \frac{1}{r}\sum_{k=1}^{r}\Phi(k,j),$$

where r – is a quantity of frames in FT marked as silence.

Modified distance matrix is then obtained as

$$\Phi'(i,j) = \begin{cases} \min(\Phi(i,j), \phi_{sil}(j)), & i \in B \\ \Phi(i,j), & i \notin B \end{cases},$$

where $\phi_{sil}(j)$ – average distance between *j*-th frame of *FS* and frames of *FT*, marked as silence, *i* – sample frame index, *j* – student frame index, *B* – set of sample frames, where student can (or is allowed to) make a pause.

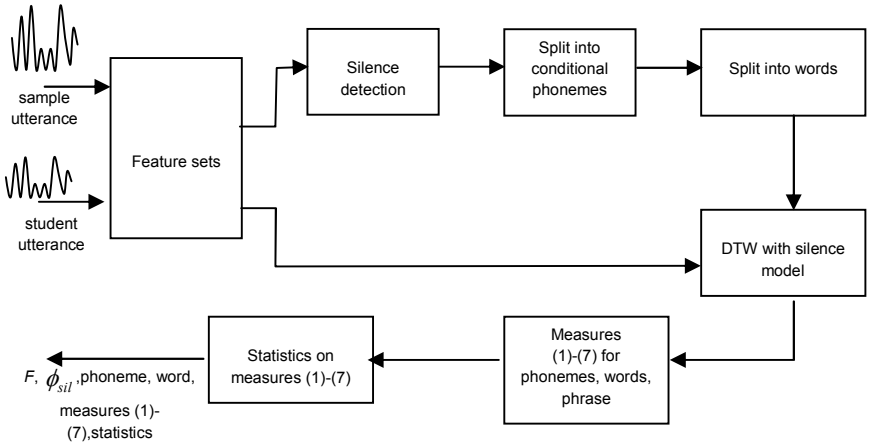


Fig. 3. Comparison steps

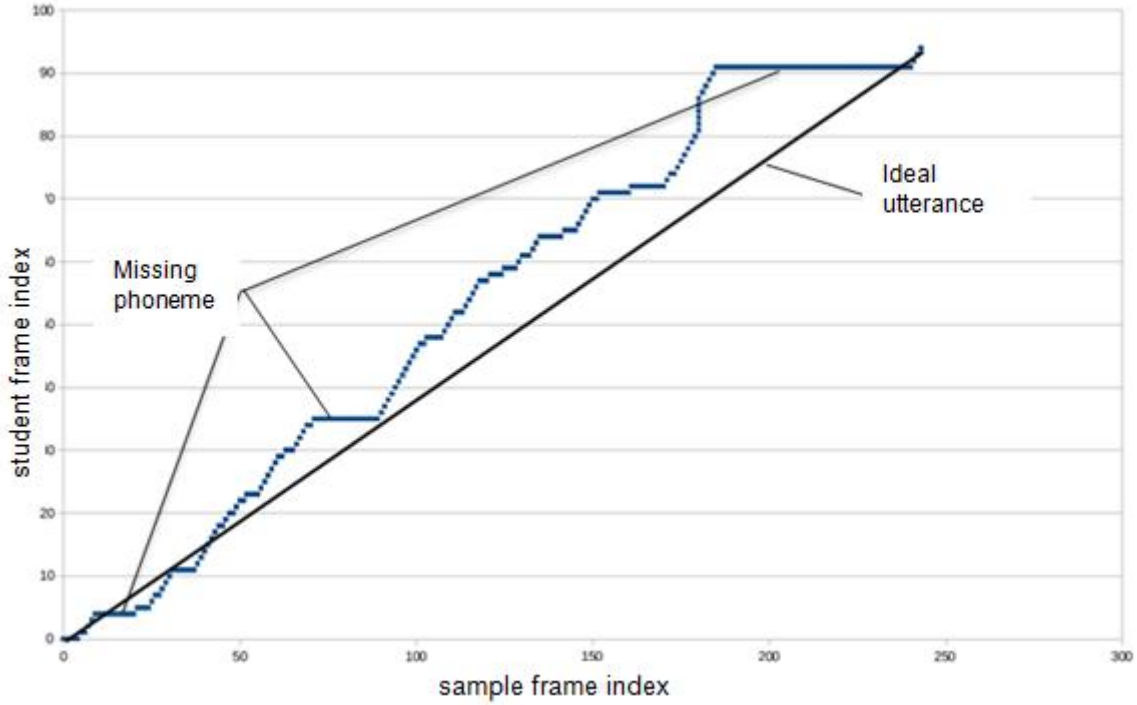


Fig. 4. Sample (ideal) and student utterances of a phrase “This woman has got a good dress”

Given $t_{t\min}, t_{t\max}$ – begin/end indexes of frames of a particular conditional phoneme of sample utterance, $t_{s\min}, t_{s\max}$ – of student utterance, we obtain the following set of measures:

– max/min indexes of student frames t_s given the index of sample frame t_t

$$\begin{aligned} s_0(t_t) &= \min(t_s \mid t_t), \\ s_1(t_t) &= \max(t_s \mid t_t); \end{aligned} \tag{8}$$

– max/min indexes of sample frames t_t given the index of student frame t_s

$$\begin{aligned} t_0(t_s) &= \min(t_t \mid t_s), \\ t_1(t_s) &= \max(t_t \mid t_s); \end{aligned} \tag{9}$$

– average angle of a slope of the graph of a linear function (see Fig. 4)

$$K = \frac{t_{s\max} - t_{s\min}}{t_{t\max} - t_{t\min}}; \tag{10}$$

– deviation from the graph of a linear function

$$C = \sum_{t=t_{t\min}}^{t_{t\max}} \max\left(\left|s_0(t) - t \cdot K\right|, \left|s_1(t) - t \cdot K\right|\right), \tag{11}$$

– maximal deviation from the graph of a linear function

$$D = \max_{t_{t\min} \le t \le t_{t\max}} \left(\max\left(\left|s_0(t) - t \cdot K\right|, \left|s_1(t) - t \cdot K\right|\right)\right), \tag{12}$$

– maximal quantity of student frames correspondent to one sample frame

$$S = \max(s_1(t) - s_0(t)), t_{t_{\min}} \le t \le t_{t_{\max}}; \quad (13)$$

– maximal quantity of sample frames correspondent to one student frame

$$S = \max(t_1(s) - t_0(s)), t_{s_{\min}} \le s \le t_{s_{\max}}. \quad (14)$$

Enlisted measures aim at evaluation of the pronunciation speed and duration of a phoneme utterance.

To measure similarity of utterance of two phonemes of the same length (in frames) we used Euclidean distance between each pair of phonemes $R_1 = \sum_{t=t_{t_{\min}}}^{t_{t_{\max}}} \sum_{s=s_0(t)}^{s_1(t)} \Phi'(t,s)$, and $R_2 = \sum_{t=t_{t_{\min}}}^{t_{t_{\max}}} \Phi'(t, t \cdot K)$.

Classification

Given a feature set f_t and measures set (8)-(14), it is possible to classify student utterance as correct or mispronounced. Classification task was formulated as follows: given a small set of sample utterances and an example utterance, obtain pronunciation quality as similarity measure. Each sound file is presented as two-dimensional array: $F = \{f_t\}, t = \overline{1,T}$, where t – is a frame number, f_t – set of 42 features (7), calculated for the frame t.

To compare different recordings, their durations were equalized with DTW, hence all the sounds were presented as two-dimensional arrays of the same size.

Let classifier be presented with an unknown function $h : F \to \{-1,+1\}$, where "-1" and "+1" are classes correspondent to "mispronounced" and "pronounced correctly". Function h is selected such that if $h(F) < 0$, an example utterance is considered as mispronounced, and if $h(F) > 0$ – as pronounced correctly.

The main problem for mispronunciation detection task was the small set of samples. Most machine learning techniques (Bayesian classifiers, neural networks, hidden Markov's models) require large training sets. Usage of small training sets leads to overfitting problem – classifier simply stores the whole training set, without learning and generalizing, so even slight modification of the sample leads to errors. However, modifications are unavoidable, because sample utterances can be recorded by different people, having different recording devices.

Therefore, only simple classifiers, such as support vector machines (SVM) dealing with small training sets, can be applied. SVM classification technique seeks for hyper plane separating two clusters in multidimensional space, where the most important points are the closest to the borders of clusters. However, in our task it is unknown if such a hyper plane exists, because it is possible, that there is not a plane, but a complex surface (parabolic etc.). We conducted a set of experiments with SVM, as in [9], but on a small training set, and obtained unsatisfactory results.

Hence, we concluded at selection of machine learning ensemble meta-algorithm Bootstrap Aggregating (Bagging) [22].

Bagging

As the sample set is relatively small, classifiers like SVM cannot be used due to large training set required, we propose to use bootstrap aggregating, or bagging algorithm [22] to generate training set for classifier.

The main idea of bagging is to create an ensemble of simple classifiers, each of which is trained on a randomly selected training subset

$$h_q(F, z_q), q = \overline{1, Q},$$

where Q – number of classifiers, z_q – some adjustable parameters, F – audio file feature set. Q is either predefined or adjusted depending on the training results.

After training, we obtain a set of h_q, on average behaving as a $h(F)$ we seek for, and the resulting classifier is averaging all the h_q:

$$h(F) = \frac{1}{Q}\sum_{q=1}^{Q} sign(h_q(F, z_q)),$$

that is a value of comparison between a sample and an example.

Training set construction for bagging

To create training sets three consequent random generators were used. First generator selected a feature index m_q from the feature set f_m (integer from 1 to 42), second – a moment of time t_q. Third generator worked several times – it selected indexes of elements from the set of all utterances, both sample and students', $\{l_{qi}\}, i = \overline{1, I}$.

Training subset is a set of pairs

$$(F_{t_q, m_q}[l_{qi}], class[l_{qi}])$$

where $class[l_{qi}] = 1$, if $F_{t_q, m_q}[l_{qi}]$ is correctly pronounced, $class[l_{qi}] = -1$, if $F_{t_q, m_q}[l_{qi}]$ is mispronounced.

As functions $h_q(F[l_{qi}], z_q)$ we selected linear functions

$$h_q(F[l_{qi}], z_q) = F_{t_q, m_q}[l_{qi}] + z_q, \quad (15)$$

where $F_{t_q, m_q}[l_{qi}]$ is a real number – the value of m_q for feature set f_{t_q}, for l_{qi}-th utterance, z_q - some real number.

To train each classifier $h_q(F, z_q)$ it is necessary to find z_q, minimizing the error

$$ERR_q = \sum_{i=1}^{I} \left| class[l_{qi}] - sign\left(h_q\left(F_{t_q, m_q}[l_{qi}], z_q\right)\right)\right|$$

Classifiers (8) are simple, easy to create and to train. For each classifier $h_q(F, z_q)$ calculated is the frequency of errors, and the most precise classifiers remain, others are

removed. The selection assumes each classifier decided its dominant class, "-1" or "+1", and then class number is averaged.

The benefits of bagging are: there is no overfitting, adding "noise" is a step of classifiers creation; best features are selected automatically at the classifiers selection stage; rather complex surfaces, not just planes in the feature space, can be dealt with bagging.

Experiments

To assess pronunciation quality, calibration of results is needed. To calibrate the system, we use small additional set of utterances by students with good pronunciation grades, confirmed by a teacher. This additional set was used to obtain minimum and maximum permissible values of each feature.

Phoneme or word considered as mispronounced if any of measures (8)-(14) go beyond permissible values. A phrase is considered as mispronounced if any of phonemes or words was mispronounced.

Examples of pronunciation quality assessment are shown in Table 1, where "-f" – female student, "-m" – male student.

Table 1.

Grades of pronunciation quality assessment for the phrase "This woman has got a good dress"

Example	Worst word grade	Worst conditional phoneme	Expert grade
students with good pronunciation grades			
2-f	0.250	0.250	good
3-f	0.250	0.250	good
4-f	0.250	0.389	good
5-f	0.250	0.250	good
6-f	0.250	0.250	good
7-m	0.250	0.250	good
8-m	0.250	0.250	good
9-m	0.250	0.250	good
students			
06-f	1.499	1.033	weak accent
00-m	1.887	3.764	strong accent
05-f	3.949	6.386	strong accent
07-f	3.748	2.499	strong accent
08-f	5.936	6.440	strong accent
08-f	4.343	6.814	strong accent
09-m	2.017	3.976	strong accent
01-f	2.191	9.836	strong accent
03-f	5.247	3.582	missed word
04-f	3.106	7.746	other phrase
10-f	25.487	25.487	one word instead of phrase
11-m	2.666	2.666	one word instead of phrase

Concluding remarks

The paper discusses the possibility to adopt known algorithms, used in ASR systems, to a comparison-based CAPT system. The proposed combination is MFCC-based sound feature set, DTW with silence model and bagging for creation/training pronunciation classifiers given a small sample set. Training is performed for each sample utterance separately, and allows for a small sample set. Adding a new sample does not require the whole system rebuilding, hence the solution is scalable.

Proposed approach evaluates both correctness of pronunciation and duration/number of phonemes. To define proper pronunciation a small training set is enough – nearly 10 samples of each phrase, uttered by different voices and at different rate of speech.

Directions of future work are seen as follows. First, to compare the quality of results on other corpora possessing both sample and student utterances. Second, to apply other classifier types that are tolerant to small sample sets.

References

1. Witt, S.M. Automatic Error Detection in Pronunciation Training: Where we are and where we need to go / S.M. Witt // Proc. IS ADEPT. – 2012. – Vol. 1. – PP. 1-8.
2. Lohiya, S.V. Survey on Computer Aided Language Learning using automatic accent assessment techniques / S.V. Lohiya, M.V. Kamble // Proc. ICPC. – 2015. – PP. 1-4.
3. Duolingo [Electronic resource]. – Access to resource: https://www.duolingo.com
4. Englishtown [Electronic resource]. – Access to resource: http://www.englishtown.com
5. Nuance Recognizer [Electronic resource]. – Access to resource: http://www.nuance.com/for-business/by-solution/customer-service-solutions/solutions-services/inbound-solutions/self-service-automation/recognizer/recognizer-languages/index.htm
6. Google voice search [Electronic resource]. – Access to resource: https://en.wikipedia.org/wiki/Google_Voice_Search
7. Ethnologue: Languages of the World [Electronic resource]. – Access to resource: http://www.ethnologue.com/
8. Robinson, T. WSJCAM0: A British English speech corpus for large vocabulary continuous speech recognition / T. Robinson, J. Fransen, D. Pye, J. Foote, S. Renals // Proc. ICASSP 1995. – IEEE Computer Society. – 1995. – PP. 81-84.
9. Lee, A.A. Comparison-based Approach to Mispronunciation Detection / A. Lee, J. Glass // Proc. SLT Workshop 2012. – IEEE. – 2012. – PP. 382-387.
10. Eskenazi, M. An overview of spoken language technology for education / M. Eskenazi // Speech Communication. – 2009. – 51(10). – PP. 832-844.
11. Delmonte, R. Exploring Speech Technologies for Language Learning, Speech and Language Technologies / R. Delmonte // InTech. – 2011. – PP.71-105.
12. Levis, J. Computer technology in teaching and researching pronunciation / J. Levis // Annual Review of Applied Linguistics. – 2008. – Volume 27. – PP.184-202.
13. Franco, H. Automatic detection of phone-level mispronunciation for language learning / H. Franco, L. Neumeyer, M. Ramos, H. Bratt // Proc. Eurospeech 99. – ICSA. – 1999. – PP. 851-854.
14. Witt, S.M. Phone-level pronunciation scoring and assessment for interactive language learning / S.M. Witt, S. Young // Speech Communication. – 2000. – 30(2-3). – PP. 95-108.
15. Yoon, S.-Y. Landmark-based Automated Pronunciation Error Detection / S.-Y. Yoon, M. Hasegawa-Johnson, R. Sproat // Proc. Interspeech. – ISCA. – 2010. – PP. 614-617.
16. Ai, R. Automatic Pronunciation Error Detection and Feedback Generation for CALL Applications / R. Ai // Lecture Notes in Computer Science. – 2015. – Volume 9192. – PP. 175-186.
17. Harrison, A.M. Improving mispronunciation detection and diagnosis of learners' speech with context-sensitive phonological rules based on language transfer / A.M. Harrison, W.Y. Lau, H.M. Meng, L. Wang // Proc. Interspeech. – ISCA. – 2008. – PP. 2787-2790.
18. Nicolao, M. Automatic assessment of English learner pronunciation using discriminative classifiers / M. Nicolao, A.V. Beeston, T. Hain // Proc. ICASSP 2015. – IEEE. – 2015. – PP. 5351-5355.
19. Mida, L. Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques / L. Muda, M. Begam, I. Elamvazuthi // Journal of Computing. – 2010. – 2(3). – PP. 138-143.

20. Asgari, M. Voice Activity Detection Using Entropy in Spectrum Domain / M. Asgari, A. Sayadian, M. Farhadloo // Proc. Telecommunication Networks and Applications. – IEEE. – 2008. – PP. 407-410.
21. Dobrovolsky, G.A. Application of Shannon entropy for voice activity detection in noisy sound recordings (in Russian) / G.A. Dobrovolsky, O.O. Todoriko // Herald of Kherson National Technical University. – 3(58). – 2016 (in press).
22. Breiman, L. Bagging Predictors / L. Breiman // Machine Learning. – 1996. – 24(2). – PP. 123-140.

ОЦІНКА ЯКОСТІ ВИМОВИ МЕТОДОМ ПОРІВНЯННЯ З ЕТАЛОНОМ

Г.А. Добровольський, О.О. Тодоріко, Н.Г. Кеберле

Запорізький національний університет
вул. Жуковського, 66, м. Запоріжжя, 69600, Україна; e-mail: gen.dobr@gmail.com

Задача оцінки якості вимови за допомогою порівняння з еталонною вимовою зазвичай потребує великої кількості еталонів. На жаль, підібрати необхідну кількість еталонів навіть для розповсюдженої мови важко, оскільки переважна більшість анотованих корпусів містить лише набори прикладів некоректної вимови, без еталонних прикладів. У даній статті запропоновано один підхід до оцінки якості вимови методом порівняння з еталоном в умовах невеликої кількості еталонних вимов. Метод DTW з урахуванням тиші дозволяє співставити еталонну фразу, яку вимовив вчитель/носій мови, із фразою учня, та отримати набір властивостей вимови рівня слова і фонеми. На цьому наборі властивостей виконується класифікація фрази як коректно/некоректно вимовленої за допомогою методу bagging, який не потребує великої кількості еталонів для навчання.

Ключові слова: комп'ютеризоване навчання вимові, вивчення мови, визначення помилок у вимові, dynamic time warping, bagging

ОЦЕНКА КАЧЕСТВА ПРОИЗНОШЕНИЯ МЕТОДОМ СРАВНЕНИЯ С ЭТАЛОНОМ

Г.А. Добровольский, О.А. Тодорико, Н.Г. Кеберле

Запорожский национальный университет
ул. Жуковского, 66, г. Запорожье, 69600, Украина; e-mail: gen.dobr@gmail.com

Задача оценки качества произношения путем сравнения с эталонным произношением обычно требует большого количества эталонов. К сожалению, подобрать нужное количество эталонов даже для широко распространенных языков трудно, подавляющее большинство аннотированных корпусов содержат лишь наборы примеров неправильного произношения, но не эталонные примеры. В данной статье предлагается один подход к оценке качества произношения методом сравнения с эталоном при условии небольшого количества эталонов. Dynamic Time Warping с учетом тишины позволяет сопоставить эталонную фразу, произнесенную учителем/носителем языка, с фразой ученика, и получить набор свойств произношения уровня слова и фонемы. На основании этого набора свойств выполняется классификация фразы как правильно/неправильно произнесенной с помощью метода bagging, который не требует большого количества эталонов для обучения.

Ключевые слова: компьютеризированное обучение произношению, изучение языка, определение ошибок в произношении, dynamic time warping, bagging

УДК 511:003.26.09

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 270-279

ВЫБОР ЭФФЕКТИВНОГО БАЗОВОГО ОСНОВАНИЯ МОДУЛЯ ПРИ МНОГОКРАТНОМ ПРОРЕЖИВАНИИ ПРОБНЫХ ЗНАЧЕНИЙ В МЕТОДЕ ФАКТОРИЗАЦИИ ФЕРМА С НЕРАВНОМЕРНЫМ ШАГОМ

Е.В. Максименко

Институт специальной связи и защиты информации Национального технического университета Украины «Киевский политехнический институт»,
ул. Верхнеключевая, 4, Киев, 03056, Украина; e-mail: iszzi@i.ua

Рассмотрена задача поиска базового основания модуля (*bb*) при многократном прореживании пробных значений в методе Ферма с неравномерным шагом. Для достижения максимального коэффициента ускорения при известном ограничении на объем памяти ЭВМ, используемой для хранения допустимых пробных значений *x*, сформулирована математическая постановка такой задачи и предложен способ ее решения на основании установленного соотношения для минимальных значений коэффициентов ускорения. Показано, что последовательность приращений пробных *x* периодически повторяется и сумма элементов такой периодической части может быть значительно меньшим чем *bb*, либо, если *bb* делится на 4 без остатка, меньшим чем *bb*/2. Обсуждаются вопросы решения задачи поиска эффективного *bb* в случае фиксированного *N*, когда значение *N*mod*bb* может меняться при изменении *bb*.

Ключевые слова: факторизация, метод Ферма, прореживание, ускорение.

Введение

В информационно-телекоммуникационных системах для целей криптографической защиты информации применяется RSA алгоритм, что вызывает интерес к его криптоанализу. В работе [1] показано, что известные примеры компрометации RSA алгоритма не являются эффективней задачи факторизации. Поэтому на современном этапе актуальным является решение задачи факторизации. Основные используемые методы решения задачи факторизации представлены в работах [2-4]. В работе [5] для решения задачи факторизации чисел предложен подход, связанный с прореживанием пробных значений с неравномерным шагом при использовании базового и множества других оснований модуля. Там же были приведены примеры выбора разных вариантов базового основания модуля и на основании численных экспериментов установлено, что время решения задачи факторизации методом прореживания существенно зависит от выбора базового основания, что требует специального исследования.

Постановка задачи

Пусть задано составное нечетное число $N = pq$, которое следует разложить на множители, где p и q некоторые нечетные числа, не обязательно являющиеся простыми. Согласно исходному варианту метода Ферма для определения p и q решают уравнение

$$X^2 = N + Y^2, \qquad (1)$$

где X и Y – целые положительные числа.

Если в (1) неизвестную X представить в виде $X = \left(\left[\sqrt{N}\right]+1\right)+x = x_0 + x$, то решение уравнения (1) получают перебором пробных значений $x = 0,1,2,\ldots,$ до тех пор, пока остаток $X^2 - N$ не окажется полным квадратом целого числа.

Назовем допустимыми те пробные значения x по модулю некоторого основания b, при которых разность $(X^2 - N) \bmod b$ является квадратичным остатком по модулю b, и недопустимыми остальные. Исключение недопустимых пробных значений x будем называть их прореживанием. Как показано в работе [5], при решении задачи факторизации целесообразно использовать множество различных оснований модуля, но при этом особая роль отводится базовому основанию модуля (в дальнейшем bb), которое играет роль первичного просеивания до дальнейшего анализа на допустимость значений x при других основаниях модуля.

В работе [5] на основании численных экспериментов при согласованных условиях для ряда вариантов факторизуемых чисел было определено время расчета при разных значениях базового основания модуля. Оказалось, что по сравнению с базовым основанием $bb = 277200$ в случае $bb = 25200$ $(25200 = 277200/11)$ время расчета увеличилось в $1.74 \div 1.78$ раза (в среднем в 1.77 раза). Для обеспечения сравнимости результатов суммарное число простых чисел в дополнительных основаниях модулей увеличилось на единицу за счет замены основания $b_{12} = 109$ на $b'_{12} = 1199 = 109 \cdot 11$. При использовании значения основания модуля $bb = 3600$ $(3600 = 27720/11/7)$ время расчета увеличилось в $3.06 \div 3.17$ раза (в среднем в 3.11 раза).

Следовательно, для метода Ферма с прореживанием пробных значений базовое основания модуля bb является существенно влияющим параметром, и способам выбора эффективного bb посвящено настоящее исследование.

Характеристики, определяющие эффективность основания модуля

Пусть $KN(b,N)$ - число допустимых значений x для основания b, для которых остаток $\left(x^2 - N\right) \bmod b$ является квадратичным остатком по модулю b, $MK(b)$ – множество чисел k от 1 до b-1, для которых НОД$(k,b) = 1$, а n_{MK} – число элементов множества $MK(b)$. Пусть также как и в [5]:

$|r(b)|_{\min}$ – минимальное число элементов множества допустимых x среди всех значений $N \bmod b$, не имеющих общих делителей с b: $|r|_{\min} = \min\limits_{k \in MK(b)} (KN(b,k))$,

$|r(b)|_{\max}$ – максимальное такое число элементов: $|r|_{\max} = \max\limits_{k \in MK(b)} (KN(b,k))$,

$r(b)_{cp}$ – среднее значение числа элементов: $|r|_{cp} = \sum\limits_{k \in MK(b)} KN(b,k) / n_{MK}$.

Определим коэффициенты ускорения:

$z_{\min}(b) = b / |r(b)|_{\max}$ – коэффициент минимального ускорения,

$z_{\max}(b) = b / |r(b)|_{\min}$ – коэффициент максимального ускорения,

$z_{cp}(b) = b / |r(b)|_{cp}$ – усредненное значение коэффициента ускорения.

В работах [6, 7] для коэффициента $z_{cp}(b)$ экспериментально установлено:

1. $z_{\min}(2^4) = z_{cp}(2^4) = z_{\max}(2^4) = 4$,

2. $z_{cp}(2^k) < 6$ при $k > 4$,

3. $z_{cp}(cb) = z_{cp}(c) \cdot z_{cp}(b)$, если c и b взаимно простые,

4. $z_{cp}(b) = 2$, если b простое число.

Свойства 2 и 4 не выполняются для $z_{\min}(b)$, а выполнение (или невыполнение) свойства 3 требует обоснования. На основании численных экспериментов в работе [6] определено, что для случая простых b $z_{\min}(b) = 2b/(b+1)$. В работе [5] на основании такого равенства получена оценка для $z_{\min}(b)$, когда b является произведением двух простых чисел. Коэффициент $z_{\min}(b)$ определяет условия максимально сложного варианта для метода многократного прореживания, который для случая метода факторизации Ферма будет при тех $N \bmod b$, при которых коэффициент ускорения будет минимальным. Поэтому при выборе b важно иметь оценку для $z_{\min}(b)$.

Исходя из соотношения $z_{\min}(b) = 2b/(b+1)$, для случая простых b можно предположить, что увеличения $z_{\min}(bb)$ можно добиться за счет умножении исходного основания bb на простые числа, которые еще не являются его множителями. Элементарные вычисления показывают, что число таких простых множителей не может быть большим, поскольку очень быстро растет их произведение и, как следствие, требуемая оперативная память компьютера для хранения допустимых x. Поэтому при выборе базового основания модуля целесообразно использовать и степени простых чисел. В табл. 1 приведены результаты численного определения $z_{\min}(b)$ для степеней $m = 1,2,3$ простых чисел $p < 24$, а также числа $4 = 2^2$.

Таблица 1.

Значения коэффициентов ускорения

Z	$Z_{\min}$			$Z_{\max}$			Z_{cp}		
m	1	2	3	1	2	3	1	2	3
$p=3$	1.5	3	3	3	4.5	6.75	2	3.6	4.154
$p=4$	2	4	4	2	4	8	2	4	5.333
$p=5$	1.667	2.5	2.5	2.5	3.571	4.032	2	2.941	3.086
$p=7$	1.75	2.333	2.333	2.333	3.0625	3.236	2	2.649	2.711
$p=11$	1.833	2.2	2.2	2.2	2.630	2.683	2	2.396	2.418
$p=13$	1.857	2.167	2.167	2.167	2.522	2.558	2	2.331	2.346
$p=17$	1.889	2.125	2.125	2.125	2.388	2.407	2	2.249	2.257

Как следует из данных табл. 1, для простых $p > 2$

$$z_{\min}\left(p^m\right) = \begin{cases} 2p/(p+1), m=1 \\ 2p/(p-1), m>1 \end{cases}. \qquad (2)$$

А в случае составного числа $p = 4$ при $m > 1$ $z_{\min}\left(4^m\right) = 4$.

Поскольку для произвольного bb существует $k = N \bmod bb$, для которого $|r|_{\max} = KN(b,k)$, то задача построения эффективного bb, гарантирующего заданное ускорение независимо от $N \bmod bb$, сводится к задаче построения наименьшего bb, обеспечивающего заданное минимальное ускорение. Для решения такой задачи определим общую структуру bb.

Общая структура bb и оценка значения $z_{\min}(bb)$

Согласно основной теореме арифметики произвольное целое число можно единственным способом представить в виде произведения простых чисел (с учетом перестановки множителей и их знаков). Поэтому общая структура базового основания модуля может быть записана в виде

$$bb = \prod_{k=1}^{n(bb)} p_k^{m_k}, \tag{3}$$

где $n(bb)$ – количество простых чисел – множителей bb, p_k $(k = 1 \div n)$ – простые числа – множители bb, m_k $(k = 1 \div n)$ – показатели степеней c_k. При этом, согласно данным табл. 1, имеет смысл использовать степени простых чисел не больше второй, а для числа 2 – четвертой. Если для произведения взаимно простых чисел $p1$ и $p2$ верно соотношение

$$z_{\min}(p1 \cdot p2) \geq z_{\min}(p1) \cdot z_{\min}(p2), \tag{4}$$

то для $z_{\min}(bb)$ будет достоверной оценка

$$z_{\min}(bb) \geq \prod_{k=1}^{n(bb)} z_{\min}(p_k^{m_k}). \tag{5}$$

Покажем, что соотношение (4) справедливо для произвольных взаимно простых чисел $p1$ и $p2$. Для этого первоначально докажем, что существует N такое, что количество чисел y в диапазоне от 0 до $p1 \cdot p2 - 1$, для которых $(y^2 - N) \bmod p1$ и $(y^2 - N) \bmod p2$ одновременно являются квадратичными остатками, равно произведению $z_{\min}(p1) \cdot z_{\min}(p2)$.

Действительно, для произвольных $p1$ и $p2$ существуют $z_{\min}(p1)$ и $z_{\min}(p2)$. Тогда существуют числа $k1\,(0 < k1 < p1)$ и $k2\,(0 < k2 < p2)$ такие, что $|r(p1)|_{\max} = KN(p1, k1)$ и $|r(p2)|_{\max} = KN(p2, k2)$.

Очевидно, что если для некоторого $x1\,(0 < x1 < p1)$ $(x^2 - k1) \bmod p1$ является квадратичным остатком по модулю $p1$, то таким же квадратичным остатком по модулю $p1$ будут числа $x_i = x1 + p1 \cdot i$ для $i \geq 0$.

Пусть $MPK(p1, p2, t)$ – множество чисел $\{(t + p1 \cdot i) \bmod p2\}_{i=0}^{p2-1}$. Множество $MPK(p1, p2, t)$ содержит все числа от 0 до $p2 - 1$. Если бы это было не так, то нашлись два равных значения $(t + p1 \cdot i1) \bmod p2 = (t + p1 \cdot i2) \bmod p2$, где $i1 < i2$. Следовательно, разность $(t + p1 \cdot i2) - (t + p1 \cdot i1) = p1 \cdot (i2 - i1)$ делилась бы на $p2$. Но $i2 - i1 < p2$, откуда следовало бы, что $p1$ и $p2$ имеют общий делитель больший единицы, т.е. не являются взаимно простыми.

Пусть $|r(p2)|_{\max} = \max_{k \in MK(p2)} (KN(p2, k)) = KN(p2, k2)$, где $0 \leq k2 \leq p2$. Поскольку множество $MPK(p1, p2, t)$ содержит все числа от 0 до $p2 - 1$, то среди элементов множества $\{t + p1 \cdot i\}_{i=0}^{p2-1}$ будет ровно $KN(p2, k2)$ значений $\{t_j = t + p1 \cdot i_j\}_{j=0}^{KN(p2,k2)}$, для которых $(t_j^2 - k2) \bmod p2$ будут квадратичными остатками по модулю $p2$. Так как это

верно для произвольного t в пределах от 0 до $p1-1$, то верно и для всех x, для которых $\left(x^2-k1\right)\bmod p1$ является квадратичным остатком.

Согласно китайской теореме об остатках существует единственное значение N такое, что $N \bmod p1 = k1 \bmod p1$, $N \bmod p2 = k2 \bmod p2$ и $0 < N < p1 \cdot p2$. Поэтому для произвольного $x1\,(0 < x1 < p1)$, для которого $\left(x^2-k1\right)\bmod p1$ является квадратичным остатком по модулю $p1$, существует $KN(p2,k2)$ чисел $\left\{x_j = x1 + p1 \cdot i_j\right\}_{j=0}^{KN(p2,k2)}$, для которых $\left(x_j^2-N\right)\bmod p1$ является квадратичным остатком по модулю $p1$, а $\left(x_j^2-N\right)\bmod p2$ – квадратичным остатком по модулю $p2$. А с учетом того, что при разных значениях $x1$ ($x1_1$ и $x1_2$) множества $MPK(p1,p2,x1_1)$ и $MPK(p1,p2,x1_2)$ не пересекаются, то суммарное количество чисел в диапазоне от 0 до $p1 \cdot p2 - 1$, для которых $\left(y^2-N\right)\bmod p1$ и $\left(y^2-N\right)\bmod p2$ одновременно являются квадратичными остатками, равно произведению $z_{\min}(p1) \cdot z_{\min}(p2)$, что и требовалось доказать.

Пусть теперь для взаимно простого с $p1$ и $p2$ числа $N\,(0 < N < p1 \cdot p2)$ $\left|r(p1 \cdot p2)\right|_{\max} = \max\limits_{k \in MK(p1 \cdot p2)} (KN(p1 \cdot p2,k)) = KN(p1 \cdot p2,N)$. Покажем, что для каждого $x\,(0 < x < p1 \cdot p2)$, для которого $\left(x^2-N\right)\bmod(p1 \cdot p2)$ является квадратичным остатком по модулю $p1 \cdot p2$, выполнены условия:

- $\left(x^2-N\right)\bmod p1$ является квадратичным остатком по модулю $p1$,
- $\left(x^2-N\right)\bmod p2$ является квадратичным остатком по модулю $p2$.

Поскольку $\left(x^2-N\right)\bmod(p1 \cdot p2)$ является квадратичным остатком по модулю $p1 \cdot p2$, то существует $y\,(0 < y < p1 \cdot p2)$, что $\left(x^2-N\right)\bmod(p1 \cdot p2) = y^2 \bmod(p1 \cdot p2)$.

Тогда $y^2 = c \cdot p1 \cdot p2 + r0$, где c – некоторое натуральное число или 0, а $r0 = y^2 \bmod(p1 \cdot p2)$. Если $r0$ представить в виде $r0 = c1\,p1 + r1$, где $c1$ – некоторое натуральное число или 0, а $r0$ – натуральное число, то $r1 = r0 \bmod p1$. Следовательно, $y^2 \bmod p1 = (c \cdot p1 \cdot p2 + c1 \cdot p1 + r0) \bmod p1 = r0$, а для x значение $\left(x^2-N\right)\bmod p1$ является квадратичным остатком по модулю $p1$.

Аналогично доказывается, что $\left(x^2-N\right)\bmod p2$ является квадратичным остатком по модулю $p2$.

Из доказанного следует, что в случаях N, когда максимальным будет количество значений $x\,(0 < x < p1 \cdot p2)$, для которых $\left(x^2-N\right)\bmod(p1 \cdot p2)$ является квадратичным остатком по модулю $p1 \cdot p2$, такое количество не превосходит числа тех $x\,(0 < x < p1 \cdot p2)$, для которых $\left(x^2-N\right)\bmod p1$ является квадратичным остатком по модулю $p1$, а $\left(x^2-N\right)\bmod p2$ – квадратичным остатком по модулю $p2$. Следовательно, $\left|r(p1 \cdot p2)\right|_{\max} \le \left|r(p1)\right|_{\max} \cdot \left|r(p2)\right|_{\max}$, а

$$z_{\min}(p1 \cdot p2) = \frac{p1 \cdot p2}{\left|r(p1 \cdot p2)\right|_{\max}} \ge z_{\min}(p1) \cdot z_{\min}(p2) = \frac{p1}{\left|r(p1)\right|_{\max}} \cdot \frac{p2}{\left|r(p2)\right|_{\max}},$$

что доказывает справедливость оценок (4).

Для оценки соотношения между $z_{\min}(p1 \cdot p2)$ и $z_{\min}(p1) \cdot z_{\min}(p2)$ были проведены ряд численных экспериментов и некоторые из их результатов приведены в табл. 2.

Таблица 2.

Сравнительные данные о коэффициентах ускорения

№ п/п	bb	p	2	3	5	7	11	13	$\prod z_{\min}$	$z_{\min}(bb)$
1	3600	m	4	2	2				30	30
		$z_{\min}(p^m)$	4	3	5/2					
2	25200	m	4	2	2	1			52.5	52.5
		$z_{\min}(p^m)$	4	3	5/2	7/4				
3	277200	m	4	2	2	1	1		96.25	96.25
		$z_{\min}(p^m)$	4	3	5/2	7/4	11/6			
4	3603600	m	4	2	2	1	1	1	178.75	178.75
		$z_{\min}(p^m)$	4	3	5/2	7/4	11/6	13/7		

В табл. 2 символом p обозначены простые числа, символом m – показатели их степеней, а $\prod z_{\min}$ – это произведение $z_{\min}(p^m)$.

В данных табл. 2 значение $z_{\min}(bb)$ в соотношении (5) в точности равно $\prod_{k=1}^{n(bb)} z_{\min}(c_k^{m_k})$. Поэтому для оценки эффективности базового основания модуля вместо $z_{\min}(bb)$ допустимо использовать произведение $\prod_{k=1}^{n(bb)} z_{\min}(c_k^{m_k})$.

Значение $z_{\min}(bb)$ позволяет определить размер оперативной памяти, необходимой для хранения максимально возможного количества допустимых x, равного $|r(b)|_{\max} = bb / z_{\min}(bb)$. Такие данные для вариантов bb, представленных в табл. 2, приведены в табл. 3. Следует отметить, что bb увеличивается значительно быстрее, чем коэффициент ускорения. Следовательно, быстро растет объем данных о допустимых x, что является ограничением и что следует учитывать при выборе bb.

Таблица 3.

Данные о количестве допустимых x для вариантов bb, представленных в табл. 2

Вариант bb	bb	$z_{\min}(bb)$	$\|r(b)\|_{\max} = bb / z_{\min}(bb)$
1	3600	30	120
2	25200	52.5	480
3	277200	96.25	2880
4	3603600	178.75	20160

Математическая постановка задачи выбора эффективного bb

Базовое основание модуля будем считать эффективным, если оно позволяет обеспечить максимальное ускорение при заданном ограничении на объем требуемой оперативной памяти ЭВМ для хранения информации о допустимых пробных значениях x для произвольных $N \bmod bb$, полагая при этом выполнение равенства в соотношении (5).

Согласно определению эффективности для bb можно сформулировать следующую задачу нелинейного программирования:

$$\prod_{k=1}^{n(bb)} z_{\min}(c_k^{m_k}) \to \max, \tag{6}$$

$$bb / \prod_{k=1}^{n(bb)} z_{\min}(c_k^{m_k}) \le S, \tag{7}$$

где S – некоторое заданное значение.

При решении задачи (6)-(7) целесообразно учесть некоторые соотношения и оценки для отношения $c_k^{m_k} / z_{\min}(c_k^{m_k})$, являющегося составной частью $bb / \prod_{k=1}^{n(bb)} z_{\min}(c_k^{m_k})$.

Для произвольного простого числа $p > 2$ определим значение отношения $c_k^{m_k} / z_{\min}(c_k^{m_k})$ для $m_k = 1$ и $m_k = 2$, являющегося множителем для левой части неравенства (7):

$$\begin{cases} p / z_{\min}(p) = p / (2p / (p+1)) = (p+1)/2 \\ p^2 / z_{\min}(p^2) = p^2 / (2p / (p-1)) = p(p-1)/2 \end{cases}.$$

Если при этом $p1 > p2 > 2$, то

$$\begin{cases} p1 / z_{\min}(p1) = (p1+1)/2 > (p2+1)/2 = p2 / z_{\min}(p2), \\ p1^2 / z_{\min}(p1^2) = p1(p1-1)/2 > p2(p2-1)/2 = p2^2 / z_{\min}(p2^2). \end{cases} \tag{8}$$

Следовательно, отношение $c_k^{m_k} / z_{\min}(c_k^{m_k})$ будет тем меньшим, чем меньшим будет простое число c_k.

Кроме того, при $p1 > p2 > 2$, для $t1 = p1^2 \cdot p2$ и $t2 = p1 \cdot p2^2$

$$\begin{aligned} t1 / z_{\min}(t1) - t2 / z_{\min}(t2) &= p1(p1-1)/2 \cdot (p2+1)/2 - (p1+1)/2 \cdot p2(p2-1)/2 = \\ &= (p1^2 p2 + p1^2 - p1p2 - p1) - (p2^2 p1 + p2^2 - p1p2 - p2) = \\ &= (p1 - p2)(p1p2 - p1 - p2 - 1) > (p1 - p2)(2p1 - p1 - p2 - 1) = \\ &= (p1 - p2)(p1 - p2 - 1) > 0. \end{aligned} \tag{9}$$

Согласно соотношению (9) $p1^2 \cdot p2 / z_{\min}(p1^2 \cdot p2) < p1 \cdot p2^2 / z_{\min}(p1 \cdot p2^2)$, если $p1 > p2 > 2$. Следовательно, в решении задачи (6)-(7) показатели степени простых чисел не могут увеличиваться при росте значения простого числа в произведении (3).

Из соотношений (8) и (9) следует, что если индекс k для простых чисел c_k в формуле (3) означает порядковый номер простого числа, причем (как исключение) $c_1 = 4$, то в результате решения задачи (6)-(7) получится bb следующего вида:

$$bb = \prod_{k=1}^{n(bb)} c_k \cdot \prod_{k=1}^{n2(bb)} c_k, \tag{10}$$

где $n2(bb)$ – некоторое число, не превышающее $n(bb)$.

Исходя из структуры bb, представленной формулой (10), решение задачи (6)-(7) будет состоять в переборе вариантов значений $n(bb)$ и $n2(bb)$ и выборе лучшего из них.

Рассмотрим пример задачи (6)-(7).

Пусть располагаемый объем оперативной памяти ЭВМ для хранения информации о допустимых значениях x ограничен 10^6 числами. Необходимо определить bb, обеспечивающее максимальное значение минимального ускорения, при котором количество допустимых x $|r(bb)|_{\max}$ не превышает 10^6.

Результаты расчетов для разных вариантов bb в зависимости от $n(bb)$ и $n2(bb)$, величины $z_{\min}(bb)$ и $|r(bb)|_{\max}$ представлены в табл. 4 для случая $|r(bb)|_{\max} \le 10^6$.

Таблица 4.

Результаты расчетов для примера задачи (6)-(7)

$n2$	bb, $z_{\min}(bb)$ $\lvert r(bb)\rvert_{\max}$	n				
		4	5	6	7	8
0	bb	420	4620	60060	1021020	19399380
	z	8.75	16.0417	29.7917	56.2731	106.919
	r	48	288	2016	18144	181440
1	bb	1680	18480	240240	4084080	77597520
	z	17.5	32.0833	59.5833	112.5463	213.838
	r	96	576	4032	36288	362880
2	bb	5040	55440	720720	12252240	232792560
	z	35	64.1667	119.167	225.0926	427.6759
	r	144	864	6048	54432	544320
3	bb	25200	277200	3603600	61261200	
	z	52.5	96.25	178.75	337.6389	
	r	480	2880	20160	181440	
4	bb	176400	1940400	25225200	*428828400*	
	z	70	128.33333	238.33333	*450.18519*	
	r	2520	15120	105840	*952560*	
5	bb		21344400	277477200		
	z		154	286		
	r		138600	970200		

Полученное решение задачи (6)-(7) выделено в табл. 4 курсивом, где $bb = 428828400 = 2^4 \cdot 3^2 \cdot 5^2 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17$.

Как следует из данных табл. 3 и 4, быстрый рост bb сопровождается относительно медленным ростом $z_{\min}(bb)$ и, следовательно, быстрым ростом требуемой памяти ЭВМ для хранения информации о допустимых x. Поэтому представляют интерес способы сокращения объема требуемой памяти ЭВМ без потери информации о допустимых x либо увеличения коэффициента ускорения без увеличения объема требуемой памяти ЭВМ. Эти вопросы рассматриваются ниже.

Сокращение объема требуемой памяти ЭВМ для эффективного bb

Согласно формуле (10), определяющей структуру эффективного базового основания модуля, bb всегда нацело делится на 4. Поэтому в случае допустимого $x < bb/2$, допустимыми окажутся также значения: $bb/2 - x$, $bb/2 + x$, $bb - x$. Действительно,

$$(bb/2-x)^2 \bmod bb = (bb^2/4 - bb\cdot x + x^2) \bmod bb = (bb(b/4-x)+x^2) \bmod bb = x^2 \bmod bb,$$

$$(bb/2+x)^2 \bmod bb = (bb^2/4 + bb\cdot x + x^2) \bmod bb = (bb(b/4+x)+x^2) \bmod bb = x^2 \bmod bb,$$

$$(bb-x)^2 \bmod bb = (bb^2 - 2bb\cdot x + x^2) \bmod bb = (bb(b-2x)+x^2) \bmod bb = x^2 \bmod bb.$$

Следовательно, если определить некоторое начальное допустимое значение x и приращения для определения следующих допустимых x, то последовательность приращений будет повторяться через $KN(bb,N)/2$ значений, где сумма приращений равна $bb/2$. Поэтому в памяти ЭВМ достаточно хранить данные только о приращениях для допустимых x, сумма которых равна $bb/2$.

В дальнейшем для периодической последовательности приращений, сумма периодической части элементов которой равна некоторому числу P, будем говорить, что такая последовательность обладает периодом длины P вне зависимости от числа элементов ее периодической части. При этом для произвольного bb независимо от N длина периода $P = bb$. В случае же, когда bb делится без остатка на 4, $P = bb/2$.

Следует отметить, что при незначительном усложнении программного кода можно сократить объем хранимой информации еще в два раза, если воспользоваться условием, что вместе с x допустимым является $bb/2 - x$.

Дальнейшее сокращение объема требуемой памяти ЭВМ возможно для отдельных случаев чисел N, когда длина периода P может быть меньшей чем $bb/2$, что требует более глубокого анализа структуры N и bb.

Выводы

Проведенные исследования показали, что задачу выбора эффективного базового основания модуля следует рассматривать в двух постановках:

1. как задачу поиска bb такого, что независимо от $N \bmod bb$ будет выполнено ограничение на максимально возможный объем памяти ЭВМ, необходимой для хранения информации о допустимых x, и тогда математическая постановка задачи сводится к задаче динамического программирования (6)-(7), которую, с учетом полученного общего вида bb (формула (10)), можно решить перебором небольшого числа вариантов;
2. как задачу поиска эффективного bb для разложения на множители известного N, когда при изменении bb может меняться $N \bmod bb$. Тогда общая постановка задачи поиска базового основания модуля, обеспечивающего максимальное ускорение при выполнении ограничений на объем требуемой памяти ЭВМ формально может соответствовать задаче (6)-(7). Но при этом не будут выполняться соотношения (4), (8), (9), а ее решение возможно на основе перебора вариантов показателей степеней простых чисел, которые могут превышать значение 2.

Список литературы

1. Brown, Daniel R.L. Breaking RSA May Be As Difficult As Factoring [Электронный ресурс] / Daniel R.L. Brown // Cryptology ePrint Archive. – Report 2005/380. –. Режим доступа: http://eprint.iacr.org/2005/380
2. Василенко, О.Н. Теоретико-числовые алгоритмы в криптографии / О.Н. Василенко. – М.: МЦНМО, 2003. – 328с.
3. Song, Y. Yan Primality Testing and Integer Factorization in Public-Key Cryptography (Advances in Information Security) / Y. Yan Song. – Springer, 2009. – 372 pp.

4. Ишмухаметов, Ш.Т. Методы факторизации натуральных чисел: учебное пособие / Ш.Т. Ишмухаметов. – Казань: Казан. ун., 2011. – 190 с.
5. Винничук, С.Д. Многократное прореживание для ускорения метода факторизации ферма при неравномерных шагах для неизвестной / С.Д. Винничук, Е.В. Максименко // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – 2016. – № 64. – С. 13-24.
6. Винничук, С.Д. Алгоритм Ферма факторизации чисел вида N=pq методом прореживания / С.Д. Винничук, А.В. Жилин, В.Н. Мисько // Электронное моделирование. – 2014. – №2. – Т. 36. – С. 3-14.
7. Місько, В.М. Прискорення методу Ферма методом проріджування з використання декількох баз / В.М. Місько // Безпека інформації. – 2015. – №1. – Т. 21. – С. 64-68.

ВИБІР ЕФЕКТИВНОЇ БАЗОВОЇ ОСНОВИ МОДУЛЯ ПРИ БАГАТОРАЗОВОМУ ПРОРІДЖУВАННІ ПРОБНИХ ЗНАЧЕНЬ В МЕТОДІ ФАКТОРИЗАЦИИ ФЕРМА З НЕРІВНОМІРНИМ КРОКОМ

Є.В. Максименко

Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут»
вул. Верхньоключова, 4, Київ, 03056, Україна; e-mail: iszzi@i.ua

Розглянуто задачу пошуку базової основи модуля (*bb*) при багаторазовому проріджуванні пробних значень в методі Ферма з нерівномірним кроком. Для досягнення максимального коефіцієнта прискорення при відомому обмеження на обсяг пам'яті ЕОМ, яка використовується для зберігання допустимих пробних значень *x*, сформульована математична постановка такого завдання і запропонований спосіб її вирішення на підставі встановленого співвідношення для мінімальних значень коефіцієнтів прискорення. Показано, що послідовність збільшень пробних *x* періодично повторюється і сума елементів такої періодичної частини може бути значно меншим за *bb*, або, якщо *bb* ділиться на 4 без залишку, меншим за *bb*/2. Обговорюються питання вирішення завдання пошуку ефективного *bb* в разі фіксованого *N*, коли значення *N*mod*bb* може змінюватися при зміні *bb*.
Ключові слова: факторизація, метод Ферма, проріджування, прискорення.

SELECTION OF EFFECTIVE BASIC BASIS OF MODULE WITH MULTIPLE THINNING TRIAL VALUE IN THE FACTORIZATION FERMAT'S METHOD WITH IRREGULAR PITCH

Ye. Maksymenko

Institute of Special Communication and Information Protection of National Technical University of Ukraine "Kyiv Polytechnic Institute",
4, Verhnyoklyuchova st., Kyiv, 03056, Ukraine; e-mail: iszzi@i.ua

The task of searching the basic module basis (*bb*) with repeated thinning test values in the Fermat's method with irregular pitch was considered. To achieve the maximum acceleration rate during a known restriction on the limited amount of computer memory used to store the valid test values of *x*, the mathematical formulation of this problem is formulated and proposed a way to solve it on the basis of the established relations for the minimum values of the acceleration coefficient. It is shown that the sequence of test increments of *x* periodically repeated and the sum of elements of the periodic part can be significantly smaller for *bb*, or if divided by 4 *bb* without a residue for less *bb/2*. The issues with the problem of search effective *bb* in the case of fixed *N*, when *N*mod*bb* value can change with *bb* are discussed.
Keywords: factorization, method of Fermat, thinning, acceleration.

UDC 519.8

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 280-289

OPTIMIZATION OF PARAMETERS IN SELF-ORGANIZATING SYSTEMS

E.D. Franzheva

Odessa National Polytechnic University,
Shevchenko Ave., Odessa, 65044, Ukraine; e-mail: franjeva.lena@gmail.com

In article there are reviewed models of systems with self-organization, which use system's conditions in previous moments of time. The conditions are included in system as linear combinations with coefficients, which are determined. Such linear combinations could be reviewed as stabilizing controls according to the principle of feedback. As the controls could be chosen not in one way, it is necessary to give additional conditions on controls' properties as criteria of optimization. There is offered a new criteria for parametric optimization of modeling of self-organizing nonlinear systems. There is shown an algorithm for constructing such parametric controls in the low-order systems. There are given examples.

Keywords: nonlinear dynamic systems, optimal stabilization, self-organization, modeling

Introduction

One of the fundamental properties of self-organization [1] (in technical, biological, economic, social systems, etc.) is the ability to copy or reproduce completed structures in the process of evolution. Such versatility suggests the copy process is caused by relatively simple properties of systems which generate it. It also gives hope to building a relatively simple model of the studied phenomenon by using nonlinear dynamical systems of a special type. Error can accumulate if templates are reproduced many times, this error leads to partial or complete loss of templates due to nonlinearity of dynamical system. We can achieve that the template will start to recover, if enter accounting of prehistory of system structure. Prehistory is taken into account as a linear combination or mixing with the given parameters in a special way. The template can be considered as the position of equilibrium of the system, and prehistory – as control according to principle of feedback which aims to stabilize the unknown position of equilibrium. One of the most efficient methods among many ones of local stabilization of unknown cycles or positions of equilibrium (e. g. the review [2]) is the Pyragas method [3]. However it has significant disadvantages [4]. A modification of the Pyragas method is proposed in [5]. Note that the set of admissible stabilizing controls generally consist in more than one element. That is why it is advisable to add criteria in the form of additional requirements for controls.

The goal of this paper is a constructing a mathematical model of systems with self-organization, in which disturbed (for whatever reason) process of reproducing stationary forms that would be restored with the greatest speed.

The task. Construct an algorithm for computing the model parameters for which the value of the recovery rate of original template adopted as the optimality criterion would be the maximum under given constraints. It is assumed that the evaluation of the rate of template loss is known.

Main part

We assume that the copy of template in the system without self-organization is carried out due to dynamic system of the form:

$$x_{n+1} = f(x_n), \tag{1}$$

where $x_n \in R^m$, $f : A \to A \subset R^m$, is a continuous vector function given on the invariant set A. It is known [6] that a set is called invariant one for system (1) if $x_0 \in A$ implies $f(x_0) \in A$.

Exact reproducing of template means that if $x_0 = X$ then $x_1 = f(x_0) = X$, $x_{i+1} = f(x_i) = X$ for all $i = 1,2,\ldots$. Thus, $x_n \equiv X$ is a position of equilibrium of system (1) and the correct copying means local asymptotic steadiness of this position of equilibrium. In this turn, local asymptotic steadiness of under review of position of equilibrium means that all of eigenvalues of the Jakobi matrix $f'(X)$, called as multipliers, are contained in the central unit circle. If this condition breaks down then any arbitrarily small error in determination of intermediate copies x_n after a few steps will lead to the fact that the system cannot reproduce the template. Moreover, this template can be lost totally.

To ensure the local asymptotic steadiness of the position of equilibrium in original system we can use the structure with self-organization which recovers the template by principle of feedback. For example:

$$x_{n+1} = f(\sum_{j=1}^{N} a_j x_{n-j+1}), a_j \geq 0,\ j = 1,\ldots,N, \sum_{j=1}^{N} a_j = 1. \tag{2}$$

If $x_{n-j+1} = X$, $j = 1,\ldots,N$, then $\sum_{j=1}^{N} a_j x_{n-j+1} = X$, as convex combination with weights a_j, $j = 1,\ldots,N$, и $f(X) = X$.

The values of parameters a_j and the value N are determined ambiguously and depend on system's multipliers. In [7] it is suggested an algorithm of this values' choice which minimize the number N. The article suggests an improvement of this algorithm such that maximizes the rate of process of template's recover. This algorithm is presented for case $N = 2$. I.e. when system (2) has form:

$$x_{n+1} = f(a_1 x_n + a_2 x_{n-1}),\ a_1 \geq 0, a_2 \geq 0,\ a_1 + a_2 = 1. \tag{3}$$

To illustrate the algorithm idea let us at first consider simple linear system of the form:

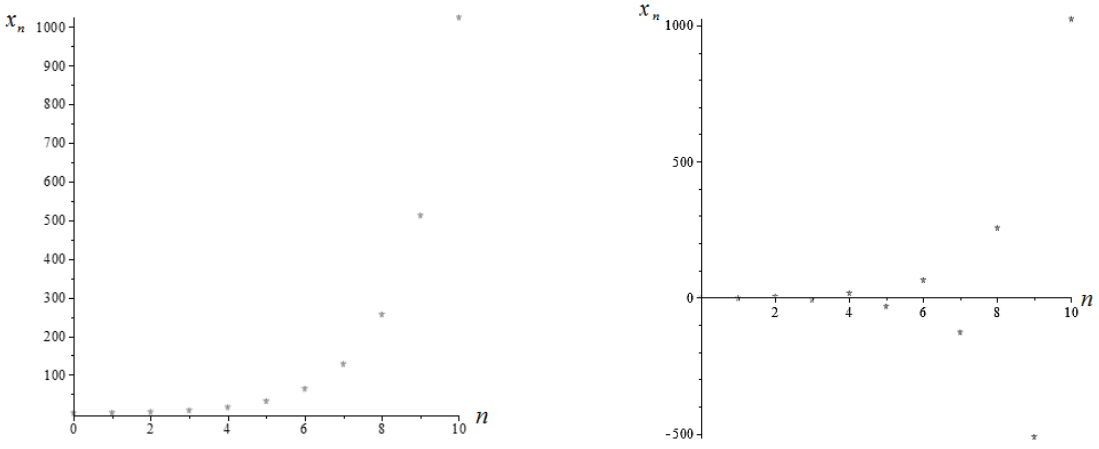
$$x_{n+1} = 2x_n. \tag{4}$$

System (4) has the position of equilibrium $x_n \equiv 0$.

Without using of control solutions of this system form a divergent sequence $\{2^n x_0\}$ with rate of increasing is 2^n. (Fig. 1, a).

If we enter a control with parameters $a_1 = \frac{1}{2}, a_2 = \frac{1}{2}$, then the system takes the form

$$x_{n+1} = x_n + x_{n-1}. \tag{5}$$



a b

Fig. 1. Behavior of dynamical system: a - $x_{n+1} = 2x_n$; b - $x_{n+1} = -2x_n$

I.e. solutions (5) form famous Fibonacci sequence which diverges with lower rate: $\left(\frac{\sqrt{5}-1}{2}\right)^n$, because $\frac{\sqrt{5}-1}{2} < 2$.

Choosing different values for a_1 and a_2 we can minimize the rate of divergence of process but we can't make it convergent.

The situation will change principally if system (1) has the form

$$x_{n+1} = -2x_n . \tag{6}$$

Let's find a common solution of equation (6): $x_n = (-2)^n x_0$. The sequence $\{(-2)^n x_0\}$ is illustrated in Fig. 1,(b).

In this case, if we chose parameters $a_1 = \frac{1}{3}, a_2 = \frac{2}{3}$ then the sequence of solutions of equation $x_{n+1} = -2\left(\frac{1}{3}x_n + \frac{2}{3}x_{n-1}\right)$ will diverge, although more slowly, then original (Fig. 2, a).

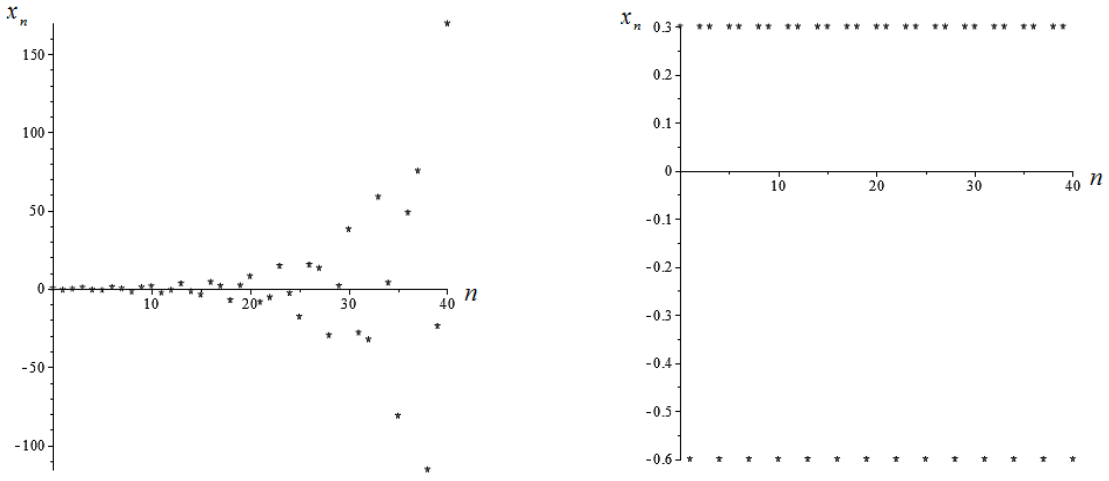
If we enter control with parameters $a_1 = \frac{1}{2}, a_2 = \frac{1}{2}$, the sequence of solutions of the equation $x_{n+1} = -x_n - x_{n-1}$ will be bounded (Fig. 2, b). By choosing parameters $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$, we will get a sequence of solutions of the equation $x_{n+1} = -2\left(\frac{2}{3}x_n + \frac{1}{3}x_{n-1}\right)$ which will be convergent (Fig. 3).

Thus, the use of averaging over time of the linear system (path) reduces the rate of divergence and for (6) makes a divergent process from vergent.

For realization of suggested idea of averaging over system's path, in common case of a nonlinear system, let us apply a method of linearization [6]. For research of stability of position of equilibrium $x_n \equiv X$ of system (2) let us consider the linearized system

$$x_{n+1} = f'(x)\sum_{j=1}^{N} a_j x_{n-j+1} , \tag{7}$$

where $f'(X)$ is the Jakobi matrix, which is calculated at position of equilibrium X.



a b

Fig. 2. Using of control with parameters for system (6): a - $a_1 = \frac{1}{3}, a_2 = \frac{2}{3}$; b - $a_1 = \frac{1}{2}, a_2 = \frac{1}{2}$

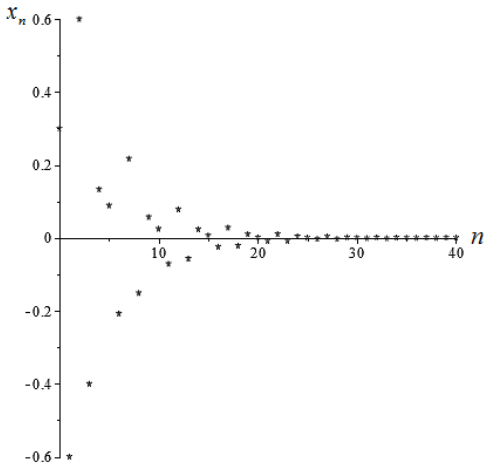


Fig. 3. Using of control with parameters $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$ for system (6)

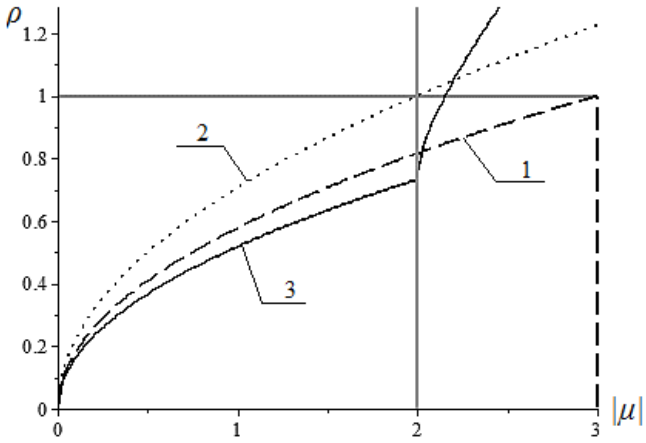


Fig. 4. Comparison of dependency ρ on $|\mu|$ with different values a_1, a_2: 1 – with $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$, 2 – with $a_1 = a_2 = \frac{1}{2}$, 3 – with $a_1 = \sqrt{3} - 1, a_2 = 2 - \sqrt{3}$.

The characteristic equation for linearized system in a neighborhood of the equilibrium position X of system (3) can be written as [8]:

$$\det\left[\lambda^N E - f'(X)\sum_{j=1}^{N} a_j \lambda^{N-j}\right] = 0. \qquad (8)$$

For our case, $N = 2$, we get

$$\prod_{j=1}^{m}\left[\lambda^2 - \mu_j(a_1\lambda + a_2)\right] = 0. \qquad (9)$$

where $\mu_j, j = 1,\ldots,m$ are the eigenvalues of matrix $f'(X)$.

Let any of the multipliers $f'(X)$ be more than 1. Let us chose it as μ_1. In this case the equation

$$\lambda^2 - \mu_1(a_1\lambda + a_2) = 0 \tag{10}$$

has necessarily a root greater than 1, because the left side of equation at $\lambda = 1$ that is less than 0, and at $\lambda = 2\mu_1$ we can estimate the left side in the following way: $4\mu_1{}^2 - \mu_1(a_1 2\mu_1 + a_2) \geq 2\mu_1^2 - \mu_1$, because $a_1 2\mu_1 + a_2 < 2\mu_1 + 1$.

Then $2\mu_1^2 - \mu_1 > 0$ at $\mu_1 > 1$ and, by Rolle's theorem, we get that on the interval $(1.2\mu_1)$ the equation (10) has a root greater than 1. That's why it is inexpedient for this task to consider the case $\mu_j > 1$. If all of $\mu_j \in (-1,1)$, then the position of equilibrium is stable without the use of control. If $\mu_j \in (-3,-1)$, $j = 1,\ldots,m$, then there exist [6] a_1, a_2, such that the equation (3) at these a_1, a_2 has all of its roots in the central unit circle, i.e. there exist a_1, a_2 such that if $\lambda^2 - \mu_j(a_1\lambda + a_2) = 0$, $\mu_j \in (-3,-1)$, then $|\lambda| < 1$. It will, for example, at $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$.

Let μ^* be the largest among all absolute values of eigenvalues of Jakobi matrix $f'(X)$, which is calculated in position of equilibrium X. μ^* is called the spectral radius of matrix $f'(X)$. The value μ^* defines the template loss rate.

The convergence rate of perturbed solutions to position of equilibrium is defined by the maximum value of the modules of the roots of characteristic equation (9). If $\mu^* \approx -3$ then the coefficients are $a_1 \approx \frac{2}{3}, a_2 \approx \frac{1}{3}$. Then the maximum root of characteristic equation (9) at module, for example, at $\mu^* = -2.99$ equals 0.998, and the convergence rate therefore equals $\frac{1}{0.998} \approx 1.02$.

Let us define ρ as distance from the origin to the most distant root of characteristic equation (9) then let us accept that the rate of template recover equals ρ^{-1}.

Let us define a dependence of this value on value of the template loss rate μ^* and parameters of the system a_1, a_2 ($a_1 + a_2 = 1$). In case $N = 2$ we will get a quadratic equation which we can solve explicitly and find a function $\rho(\mu)$. Considering that $\mu < 0$ we will use the function $\rho(|\mu|)$.

We have

$$\rho(|\mu|) = \begin{cases} \sqrt{|\mu|(1-a_1)}, & |\mu| < \dfrac{4(1-a_1)}{a_1^2} \\ \dfrac{1}{2}\left(a_1|\mu| + \sqrt{a_1^2\mu^2 - 4(1-a_1)|\mu|}\right), & |\mu| \geq \dfrac{4(1-a_1)}{a_1^2}. \end{cases} \tag{11}$$

Let us note that for N larger than 4, it is impossible to write out the dependency $\rho(|\mu|)$ explicitly.

Figure 6 shows a chart $\rho(|\mu|)$ at $a_1=\frac{2}{3}, a_2=\frac{1}{3}$ by a dotted line, at $a_1=a_2=\frac{1}{2}$ by points and at $a_1=\sqrt{3}-1, a_2=2-\sqrt{3}$ by a solid line.

Let us note that if $|\mu|>3$ then at $N=2$ it will necessarily need $\rho(|\mu|)>1$, that's why in this case in a system with self-organization N must be larger than 2. The coefficients $a_1=\frac{2}{3}, a_2=\frac{1}{3}$ provide the template recover for any of $|\mu^*|<3$ but not always with maximum rate (Fig. 4).

Let us formulate the task in the following way: the evaluation of template loss rate $\mu\in(-\mu^*,-1]$ is given and desired recovery rate ρ^{-1} is given too. Find the minimum N and coefficients $a_1,\cdots,a_N$ which provide the template recovery with the recovery rate ρ^{-1} at known loss rate of a template μ^* and for any of $\mu\in(-\mu^*,-1]$.

We can give an alternative formulation of the problem. Let μ^* and N be given. It is necessary to maximize the template recovery rate ρ^{-1} through coefficients $a_1,\cdots,a_N$ choosing them from the set $\left\{a_j\ge 0, j=1,\cdots,N, \sum_{j=1}^{N} a_j=1\right\}$.

Let us proceed to solving problems.

So, let $\rho(|\mu|)$ be modulo the maximal root of the equation

$$\lambda^2-\mu(a_1\lambda+a_2)=0,\ \mu<0. \tag{12}$$

Let us define the value

$$\left\|\rho(\mu^*)\right\|=\max_{\mu\in[0,\mu^*]}\{\rho(|\mu|)\}.$$

It is required to minimize $\left\|\rho(\mu^*)\right\|$ by parameters a_1,a_2 choosing them from the set $\{(a_1,a_2): a_1+a_2=1, a_1\ge 0, a_2\ge 0\}$.

At $N=2$ $\left\|\rho(\mu^*)\right\|=\rho(\mu^*)$ we get the minimum of the value $\rho(\mu^*)$ at condition of equality of equation roots (12), i.e. when the correlation $\mu^*=\frac{4(1-a_1)}{a_1^2}$ is performed.

Then

$$\rho^*=\min_{a_1,a_2}\rho(\mu^*)=\sqrt{\frac{4(1-a_1)^2}{a_1^2}}=\frac{2(1-a_1)}{a_1}, \tag{13}$$

where from

$$a_1=\frac{2}{\rho^*+2}. \tag{14}$$

Since

$$a_2=1-a_1,\ \text{then} \tag{15}$$

$$a_2 = \frac{\rho^*}{\rho^* + 2}. \tag{16}$$

Put (14) in (11):

$$\rho^* = \sqrt{\mu^*(1 - \frac{2}{\rho^* + 2})}, \text{ where from } \rho^{*2} = |\mu^*| \frac{\rho^*}{\rho^* + 2},$$

then:

$$\mu^* = \rho^{*2} + 2\rho^*. \tag{17}$$

From (17) follow that

$$\rho^* = -1 + \sqrt{1 + \mu^*}. \tag{18}$$

Thus we explicitly get dependency of the maximal template recovery rate $\frac{1}{\rho^*}$ on loss rate of template μ^*

$$\frac{1}{\rho^*} = \frac{1}{-1 + \sqrt{1 + \mu^*}} = \frac{1}{\mu^*}\left(1 + \sqrt{1 + \mu^*}\right), \tag{19}$$

and formulas for optimal coefficients too

$$a_1^* = \frac{2(-1 + \sqrt{1 + \mu^*})}{\mu^*},\ a_2^* = \frac{\mu^* - 2(-1 + \sqrt{1 + \mu^*})}{\mu^*}. \tag{20}$$

Examples

As an example of applying the algorithm with the use of self-organized system let us consider the system with logistic function $f(x) = x \to 4x(1-x)$ which keeps point $X = 0.75$ on any step of the copying (Fig. 5). Herein $\mu^* = |f'(0.75)| = |-2| = 2$. Let us consider that the value of point step i was defined with on error, i.e. for example $x_i = 0.7500001$. Then last copies will be much different than the template.

The system of self-organization must recover the lost template. If we accept parameters' values as $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$ then from (11) it follows that the rate of template recovery equals $\frac{1}{\rho} \approx 1.225$. Let us find values of parameters a_1, a_2, in which the template is recovering with maximal rate. From (18), (19) in $\mu^* = 2$ we will get $a_1 = \sqrt{3} - 1 \approx 0.732,\ a_2 = 2 - \sqrt{3} \approx 0.238,$ $\rho^* = 0.732$, $\frac{1}{\rho^*} = \frac{1}{2}(1 + \sqrt{3}) \approx 1.366 > 1.225$.

The exact recovery of template occurs on the interval $n \in [1,10]$. It is expected that the value of x_{10} is defined with an error of 10^{-7}. Thus on interval $n \in [10, 40]$ the process of full

template loss is demonstrated. The control is activated at the moment $n = 40$ in form of system of self-organization. On interval $n \in [40, 60]$ the process of template recovery occurs. For self-organization the following parameters are used $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$ (process is shown by *asterisk*) and $a_1 = \sqrt{3} - 1$, $a_2 = 2 - \sqrt{3}$ (process is shown by *circle*). It is seen that parameters' values $a_1 = \sqrt{3} - 1$, $a_2 = 2 - \sqrt{3}$ are more efficient than values $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$ in the sense of the rate of convergence.

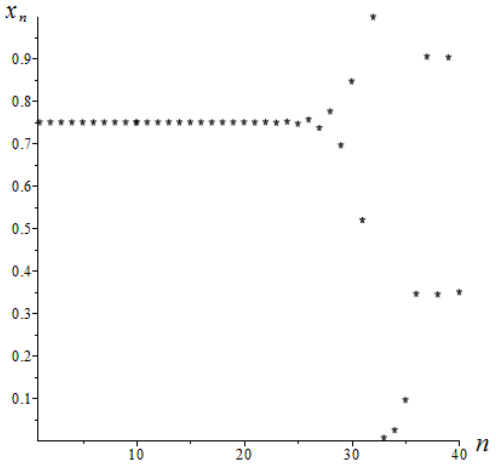


Fig. 5. Result of the copying without errors and with the error 10^{-7} which occurred on 10 step of the copying

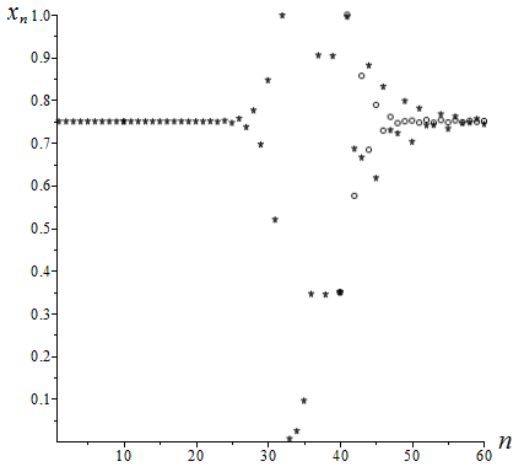


Fig. 6. The dynamic of copy process which described by the equations (1) and (2)

In the next example let us consider the tent map [9] $f(x) = x \to -2\left|x - \frac{1}{2}\right| + 1$ in which the point $X = \frac{2}{3}$ is kept (Fig. 7). Wherein $\mu = -2$, the point's value is defined with an error e.g. $x_i = X + 0.0000001$.

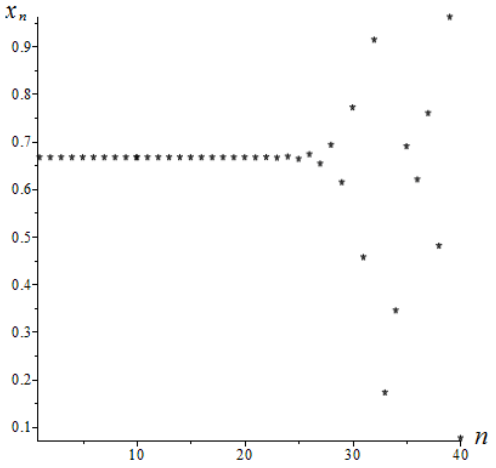


Fig. 7. The result of the copying without errors and with the error 10^{-7} which occurred on 10 step of the copying for the tent map

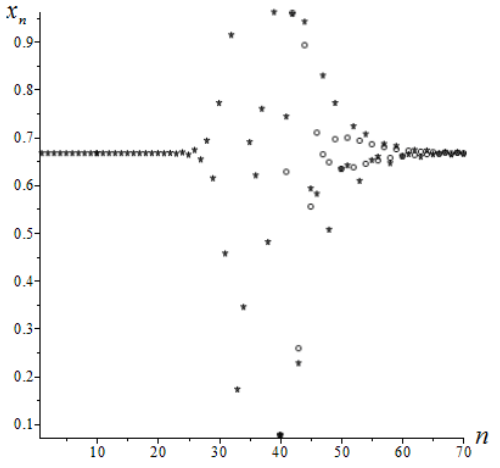


Fig. 8. The dynamic of copy process which described by the equations (1) and (2) for the tent map

Let's consider the vector case model of Markov-Ferhulst [10]

$$\begin{cases} x_{n+1} = c_{11} f(x_n) + c_{12} f(y_n), \\ y_{n+1} = c_{21} f(x_n) + c_{22} f(y_n), \end{cases} \tag{21}$$

where $f(x) = 4x(1-x)$, c_{ij} are elements of matrix $C = \begin{pmatrix} 0.3 & 0.7 \\ 0.7 & 0.3 \end{pmatrix}$. The vector $\begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix}$ is the position of equilibrium of the system (21). Wherein $\mu^* = |f'(0.75)| = |-2| = 2$. We use the control with parameters for self-organization $a_1 = \sqrt{3} - 1, a_2 = 2 - \sqrt{3}$ (Fig. 9) – process is shown by symbols ● and the control with parameters for self-organization $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$ – process is shown by symbols *.

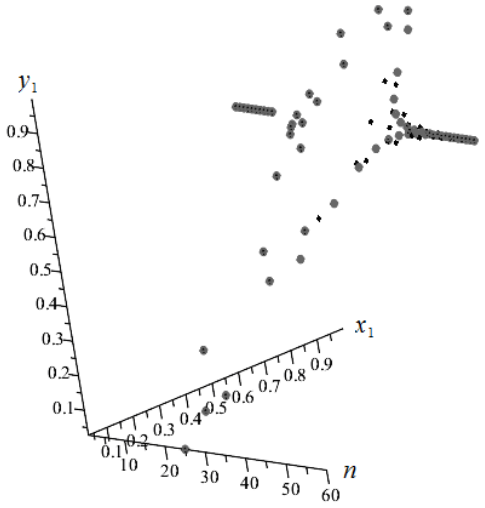


Fig. 9. The dynamic of the copy process for a vector case: ● - with parameters $a_1 = \sqrt{3} - 1$, $a_2 = 2 - \sqrt{3}$; * - with parameters $a_1 = \frac{2}{3}, a_2 = \frac{1}{3}$

The convergence of the template recovery process occurs with a larger rate for values of parameters $a_1 = \sqrt{3} - 1, a_2 = 2 - \sqrt{3}$.

Conclusion

This paper introduces an algorithm of definition of parameters' values in system of self-organization, which provide the maximal rate of recovering of the damaged template. The structure of a mathematical model is suggested for describing of systems of self-organization. For choosing of model's parameters is suggested a criteria: the maximum of rate of the convergence of recovering the template. The limit value of convergence rate is defined by scatter of system's multipliers. The algorithm is formulated for definition of system's parameters which optimize the rate of recovering the template with given value of scatter of multipliers. In particular for the logistic map we found the optimal coefficients and we show that the rate of recovering the template could be increased by 11.5%. The same result was obtained the Markov-Ferhulst model.

References

1. May, Robert M. Biological Populations Obeying Difference Equations: Stable Points, Stable Cycles, and Chaos / Robert M. May // J. theor. Biol. – 1975. – No. 51. – PP. 511-524.
2. Андриевский, А.Е. Управления хаосом. Методы и приложения. Часть 1. Методы. / А.Е. Андриевский, А.Д. Фрадков // АиТ. – 2003. – № 5. – С. 3-45.
3. Pyragas, K. Delayed feedback control of chaos / K. Pyragas // Phil. Trans. R. Soc. A. – 2006. – No. 364. – PP. 2309-2334.
4. Dmitrishin, D. On the generalized linear and non-linear DFC in non-linear dynamics [Electronic resource] / D. Dmitrishin, A. Khamitova, A. Stokolos // arXiv:1407.6488 [math.DS]. Access to resource: https://arxiv.org/abs/1407.6488
5. Дмитиришин, Д. Перемешивание как способ управления хаосом // Д. Дмитиришин, И. Скринник // Информатика и математические методы моделирования. – 2016. –Т.6. №1. – С. 11-18.
6. Ott, E. Controlling chaos / E. Ott, C. Grebodgi, J.A. Yorke. // Physical Review Letters. – 1990. – Vol. 64. – No. 11. – PP. 1196 - 1199.
7. Dmitrishin, D. Methods of harmonic analysis in nonlinear dynamics / D. Dmitrishin, A. Khamitova // Comptes Rendus Mathematique. – 2013. – Volume 351. – Issue 9-10. – PP. 367-370.
8. Dmitrishin, D. On the stability of cycles by delayed feedback control / D. Dmitrishin, P. Hagelstein, A. Khamitova, A. Stokolos // Linear and Multilinear Algebra. – 2016. – Vol. 64. – Iss. 8. - PP. 1538-1549.
9. Tian, L. Predictive control of sudden occurrence of chaos / L. Tian, G. Dong // Int. J. Nonlinear Science. – 2008. – No. 5(2). – Pp. 99-105.
10. Dmitrishin, D. Fejer polynomials and Chaos / D. Dmitrishin, A. Khamitova, A. Stokolos // Springer Proceedings in Mathematics and Statistics. – 2014. – No. 108. – PP. 49-75.

ОПТИМІЗАЦІЯ ПАРАМЕТРІВ В СИСТЕМАХ ОПТИМИЗАЦІЇ З САМООРГАНІЗАЦІЄЮ

О.Д. Франжева

Одеський національний політехнічний університет,
пр. Шевченка, 1, м. Одеса, 65044, Украина; e-mail: franjeva.lena@gmail.com

У статті розглядаються моделі систем з самоорганізацією, в яких використовуються стани системи в попередні моменти часу. Ці стани входять в систему у вигляді лінійних комбінацій з коефіцієнтами, що підлягають означенню. Ці лінійні комбінації можна розглядати як стабілізуючі управління за принципом зворотнього зв'язку. Так як ці управління можна обирати не єдиним способом, необхідно накладати допоміжні умови на властивості управлінь у вигляді критеріїв оптимізації. Запропоновано новий критерій параметричної оптимізації моделювання нелінійних систем, що саморганізуються. Вказано алгоритм конструювання таких параметричних управлінь в системах малих порядків. Приведені приклади.

Ключові слова: нелінійні динамічні системи, оптимальна стабілізація, самоорганізація, моделювання

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ В САМООРГАНИЗУЮЩИХСЯ СИСТЕМАХ

Е.Д. Франжева

Одесский национальный политехнический университет,
пр. Шевченко, 1, г. Одесса, 65044, Украина; e-mail: franjeva.lena@gmail.com

В статье рассматриваются модели систем с самоорганизацией, в которых используются состояния системы в предшествующие моменты времени. Эти состояния входят в систему в виде линейных комбинаций с коэффициентами, подлежащими определению. Эти линейные комбинации можно рассматривать как стабилизирующие управления по принципу обратной связи. Так как эти управления можно выбирать не единственным образом, необходимо накладывать дополнительные условия на свойства управлений в виде критериев оптимизации. Предложен новый критерий параметрической оптимизации моделирования самоорганизующихся нелинейных систем. Указан алгоритм конструирования таких параметрических управлений в системах малых порядков. Приведены примеры.

Ключевые слова: нелинейные динамические системы, оптимальная стабилизация, самоорганизация, моделирование

УДК 004.056:004.7

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 290-296

ВПЛИВ НЕЛІНІЙНОСТІ НЕЙРОНА НА ЦИКЛІЧНУ СИСТЕМУ УПРАВЛІННЯ

В.Г. Кононович, О.Ю. Козлова, О.Ю. Кунянский

Одеський національний політехнічний університет
просп. Шевченко, 1, Одеса, 65044, Україна; e-mail: vl_kononovich@ukr.net

Досліджується поведінка нейрона, який є основним елементом нейронних мереж. Розглядається логістична модель нейрона як найпростіший нелінійний елемент. Представлена математична модель нейрона в циклічному управлінні для випаду з вхідним потоком. Визначені умови і характер можливих нерегулярних коливань у нейроні. Отримані результати дозволяють підвищити ефективність роботи циклічних систем управління.

Ключові слова: нейрон, процеси управління, нелінійний елемент, нелінійна динаміка, комп'ютерне моделювання, динамічний хаос, біфуркації.

Вступ

Для поступового поліпшення різних видів діяльності застосовують процесно-орієнтовані підходи до управління. При автоматизації певних етапів застосовують алгоритми штучного інтелекту, нейронні мережі з нелінійними елементами. У той же час, нелінійна динаміка та синергетика дають приклади складної поведінки систем, складених з простих нелінійних елементів, здатних генерувати динамічний хаос. Виникає проблема вивчення умов впливу нелінійних явищ на системи, у які входять нелінійні елементи – в нашому випадку, нейронів. Проблемам прийняття рішень, управління та застосування нейронних мереж присвячено гігантський обсяг літератури. Вкажемо деякі з них. Моделі та теорія систем прийняття рішень представлені в [1, 2]. Приклади автоматизації управління безпекою та застосування циклічних процедур управління наведені в [3, 4]. Хорошим посібником по нейронним мережам та нейрокомп'ютерам є [5]. Одна з новітніх моделей нейрокомп'ютера Хопфілда описана в [6]. Все ширше застосовуються синергетичні методи. Але динамічні властивості моделі нейрона та можлива його складна поведінка досліджені ще недостатньо і дана тема є актуальною.

Метою даної роботи є дослідження процесів поведінки моделі нейрона як нелінійного елемента у складі нейромережі або нейрокомп'ютера та визначення його нормальних режимів.

Циклічні процеси в системах управління

Циклічні процеси повсюдно поширені у природі. Цикли важливі при вирішенні практичних задач. Циклічність і прогнозування за допомогою минулих циклів майбутніх станів складають суть сучасного підходу до процесно-орієнтованого управління. Під *процесно-орієнтованим підходом* розуміють ідентифікацію процесів та управління ними. Передусім, забезпечується взаємодія процесів. *Ідентифікація* основних процесів представляє собою їх перелік, визначення меж, стратегічної значимості кожного процесу та аналіз потреби їх оптимізації.

У стандарті ISO 9001:2008 сформульовані вимоги до системи менеджменту якості, щодо постійного поліпшення її результативності на основі процесно-орієнтованого підходу [7]. Процесно-орієнтований підхід дає можливість в умовах обмеженості засобів та ресурсів досягати потрібних результатів з мінімальними витратами. Всім процесам виробництва продукції чи послуг в організації притаманні певні відхилення від заданих значень внаслідок багатьох причин. Для зниження відхилень виробництва застосовують в управлінні, наприклад, концепцію PDCA (плануй (Plan), роби (Do), перевіряй (Check), дій (Act)). Результативність процесів забезпечується управляючим зворотним зв'язком по критерію, який необхідно поліпшити із заданою точністю.

Основними елементами циклу PDCA являються (рис. 1):

P – визначення цілей та прийняття рішення щодо необхідних змін (розроблення плану);

D – здійснення змін (втілення плану);

C – вимірювання та аналіз результатів (контроль виконання плану);

A – проведення необхідних дій, якщо результати не відповідають запланованим, або стандартизація дій у випадку успіху (виправлення плану).

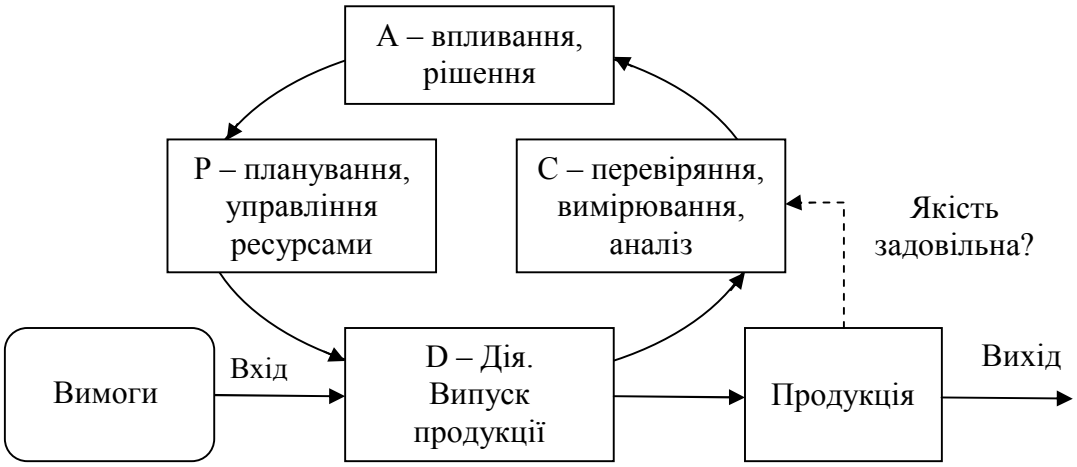


Рис. 1. Цикл постійного поліпшення системи менеджменту якості

У міжнародному стандарті ISO 9000 дано таке визначення: «*Процес* – це сукупність взаємозв'язаних або взаємодіючих видів діяльності, які перетворюють входи у заплановані виходи». *Процес управління* – це певна сукупність управлінських дій, які направлені на досягнення цілей шляхом перетворення ресурсів на «вході» (фінансові, матеріально-технічні та кадрові) у потрібний результат забезпечення життєдіяльності на «виході» системи. В середині процесу відбувається застосування та переробка матеріальних або фінансових потоків, інформації в інші потоки або послуги. Аналіз та поліпшення виконують по відношенню до процесу в цілому для досягнення цілі в умовах, що склалися.

Усе, що отримали на виході процесу, має бути перевірено. Для прийняття рішення власник процесу повинен отримати інформацію про хід процесу, про результати процесу та інформацію від споживача, щодо степені його задоволеності продуктом. Власник процесу контролює із встановленою періодичністю хід процесу та приймає рішення у випадках відхилення ходу процесу від нормального. Власник процесу в ході управління планує (Plan) розподіл ресурсів для досягнення цілей процесу з максимальною ефективністю. В ході виконання (Do) процесу власник перевіряє (Check) хід процесу на основі інформації, яка поступає в результаті вимірювання параметрів процесу. Власник процесу веде оперативне управління процесом, коригуючи (Act), змінюючи хід процесу. Діяльність власника процесу носить циклічний характер при

нормальному ході процесу або аперіодичний – у випадках виникнення проблемних ситуацій, які вимагають термінового втручання. Процес – це модель реальної діяльності. Окремі кроки контуру управління процесом можна автоматизувати із застосуванням нейронних мереж.

Застосування нейромережних технологій

Одна із численних структурних схем нейрона [8] показана на рис. 2, де: $x_1,\ldots,x_n$ – вхідні сигнали; $w_1,\ldots,w_n$ – вагові коефіцієнти; b – зміщування; S – суматор; F – функція активації нейрона; Y – вихідний сигнал нейрона. На рис. 2 вхідні сигнали $x_1,\ldots,x_n$, які поступають у нейрон, помножуються за допомогою помножувачів на вагові коефіцієнти $w_1,\ldots,w_n$, потім результат підсумовується в блоці S, і попадає на вхід функціонального перетворювача, який реалізує функцію активації нейрона F. Додавання навчального зміщування b нейрону дозволяє зсувати початок відліку логічної функції, що дає ефект, аналогічний налаштуванню порога нейрона, та приводить до прискорення процесу навчання. Зміщування b підсумовується з вагою w і приводить до зміщування аргументу функції активації на величину b

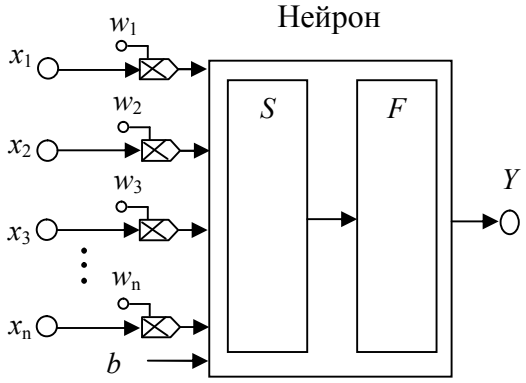


Рис. 2. Структурна схема нейрона

Головною властивістю нейронної мережі є її властивість до навчання. Процес навчання зводиться до змінювання вагових коефіцієнтів *w*. Однією із задач та головною особливістю нейронних мереж (НМ) являється здатність до навчання по деяким прикладам, які складають навчальну множину. Навчальний процес НМ представляється як налаштування архітектури та вагових коефіцієнтів синаптичних зв'язків у відповідності з даними навчаючої множини так, щоб поставлена задача була вирішена ефективно [9].

На виході суматора сигнал враховує значення вхідних сигналів і вагові коефіцієнти. Величина сигналу на виході суматора розраховується так:

$$S = x_1 w_1 + x_2 w_2 + x_3 w_3 + \ldots + x_n w_n. \qquad (1)$$

Отриманий від суматора сигнал передається до функції активації. Результат цих обчислень являється вихідним сигналом нейрона Y. Активаційні функції вводять у нейронну мережу нелінійність. Без нелінійності приховані шари у мережі не зможуть давати корисного ефекту, у порівнянні з просто лінійними персептронами (вони не мають прихованих шарів, тільки вхідний і вихідний). Вихідний сигнал нейрону

$$Y = F(S) = F(x_1 w_1 + x_2 w_2 + x_3 w_3 + \ldots + x_n w_n). \quad (2)$$

Враховуючи зміщування функція суматора буде виглядати так:

$$S_i = \sum_{i=1}^{n} w_i \cdot x_i + b_i, \quad (3)$$

де, i – номер шару нейрона.

У НМ використовуються гладкі нелінійні функції для збільшення потужності персептрону: гіперболічний тангенс або класичний сигмоїд. У випадку гіперболічного тангенсу [10]

$$\frac{dy_j}{dS_j} = 1 - S_j^2, \quad (4)$$

де S – вихід суматора нейрона.

Найчастіше у якості активаційної функції використовують сигмоїд, який має таких вигляд [11]:

$$y(S_j) = \frac{1}{1 + e^{-\alpha S_j}}. \quad (5)$$

Сигмоїд має просту похідну

$$\frac{dy_j}{dS_j} = \alpha S_j (1 - S_j). \quad (6)$$

Сигмоїдальні функції монотонно зростаючі, вони мають похідні, які відмінні від нуля по всій області визначення. При зменшенні параметра α сигмоїд стає більш пологим, вироджуючись у горизонтальну лінію на рівні 0.5 при $\alpha = 0$. При збільшенні α сигмоїд все більше наближається до функції одиничного скачка. Дані характеристики забезпечують правильне навчання та функціонування мережі [9].

Множник $\frac{\partial s_j}{\partial w_{ij}} = y_j^{(n-1)}$, $y_j^{(n-1)}$ – вихід нейрону попереднього шару. Проведемо заміну

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j} \quad (7)$$

і отримаємо рекурсивну формулу для розрахунків величин $\delta_j^{(n)}$ шару n із величин $\delta_k^{(n+1)}$ більш старшого шару $n+1$.

$$\delta_j^{(n)} = \left[\sum_k \delta_k^{(n+1)} \cdot w_{jk}^{(n+1)} \right] \cdot \frac{dy_j}{ds_j}. \quad (8)$$

Для вихідного шару

$$\delta_j^{(n)} = (y_i^{(n)} - d_i) \cdot \frac{dy_i}{ds_i}. \quad (9)$$

Проаналізуємо більш детально поведінку активаційної функції (6).

Дослідження поведінки функції активації

Існує немало реалізацій моделі нейрону, як апаратних так і програмних. Нас цікавлять програмні реалізації, в яких функції реалізуються цифровими методами [12]. Нехай НМ має сигмоїдальну функцію активізації. Будемо розглядати перехідний процес від моменту часу, коли на початку циклу управління в момент t_0 на вхід НМ подавався вектор вхідних сигналів $\mathbf{X}_0 = \{x_{01}, x_{02}, ...\}$, обчислена сума S_0, на виході НМ маємо вихідний сигнал $Y_0(S_0)$. Припустимо, що сума S та вихідний сигнал $Y_0(S_0)$ запам'ятовується у системі. Саме така вимога необхідна, щоб інтелектуальна динамічна система не «забувала» свою, принаймні найближчу, історію. Далі виконуються дії відповідних елементів циклу PDCA. Функціонал перетворень у цьому циклі нам невідомий. Щоб вивчити вклад функції активації в процес управління, будемо вважати, що, у найпростішому випадку, всі інші перетворення у циклі PDCA однозначні, мають адитивний характер та лінійні з точністю до константи. Тоді, в кінці циклу управління PDCA в момент часу t_1, після закінчення перехідних процесів, маємо вектор вхідних сигналів $\mathbf{X}_1 = \{x_{11}, x_{12}, ...\}$, встановлюється сума S_1, на виході НМ маємо вихідний сигнал $Y_1(S_1)$. Якщо на етапі випуску продукції в циклі PDCA (див. рис. 1) нема відхилень, то $S_1 = S_0$ і $Y_1(S_1) = Y_0(S_0)$. У противному разі $S_1 = S_0 + s_{in}$, де s_{in} – зміна суми S під впливом потоку управління. Потік управління може бути постійним, коли управління спрямоване на постійне підвищення обсягу продукції або якості продукції; періодичним зі зміною знака; хаотичним або шумовим, тощо.

Правило переходу від $Y_0(S_0)$ до $Y_1(S_1)$ задається в диференціальній формі виразом (7). Це дає можливість зміни суми S в одному циклі PDCA записати так:

$$S_1 = \frac{dy}{dS} = \alpha S_0 (1 - S_0) + s_{in}. \quad (10)$$

Провівши ітерації, у загальному вигляді отримуємо формулу

$$S_{n+1} = \alpha S_n (1 - S_n) + s_{in}, \quad (11)$$

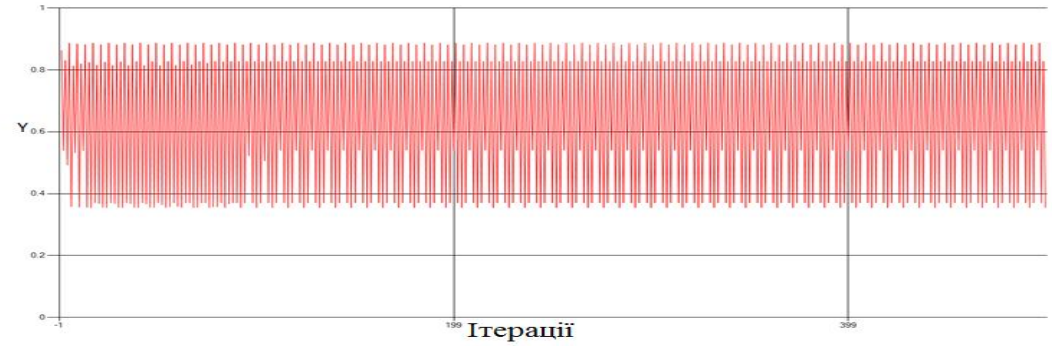
де n – номер ітерації, яке можна вважати модельним часом.

Отримане одномірне відображення із вхідним потоком. Якщо $s_{in} = 0$, то маємо одномірне відображення, яке називають *логістичним відображенням* і яке ретельно проаналізоване у численних публікаціях [13, гл. 7]. Зокрема відомо, що при невеликих значеннях α $(0 < \alpha < 1)$ $S_n \to 0$ при $n \to \infty$. Незалежно від вибору S_0. Існують нерухомі точки, для яких справедлива рівність

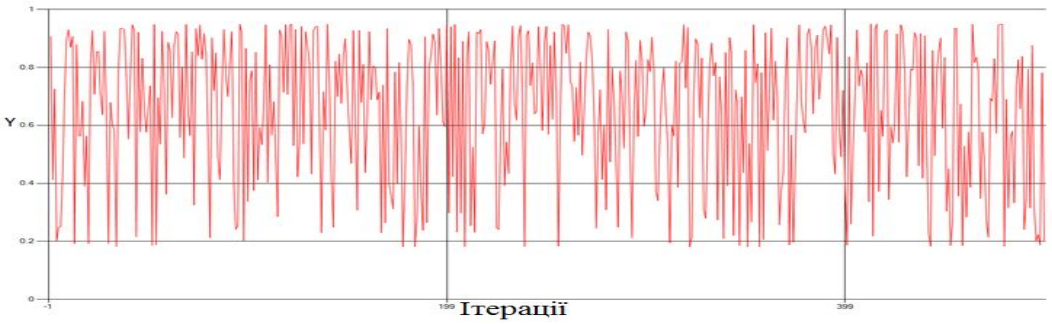
$$S^* = \alpha S^* (1 - S^*). \quad (12)$$

При $\alpha < 1$ квадратне рівняння (6) має один не від'ємний корінь $S^* = 0$. Нерухома точка стійка. При $\alpha > 1$ не від'ємних коренів два: $S^* = 0$ і $S^* = (\alpha - 1)/\alpha$. При $\alpha = 1$

(саме це значення вказується для формули (6) у більшості посібників) нерухома точка $S^* = 0$ втрачає стійкість, а нова нерухома точка стає стійкою. Аналіз одномірного логістичного відображення із вхідним потоком зустрічаються не часто. Дослідимо поведінку нейрона у залежності від коефіцієнту росту α за виразом (12). Ці рівняння показані на рис. 3, відповідно: 4а – при $\alpha = 3.551$; 4б – при $\alpha = 3.8$.



а



б

Рис. 3. Сигнал на виході НМ у залежності від при значення параметру α

При великих значеннях коефіцієнту росту виникають спочатку регулярні коливання, а при збільшенні параметру α – детермінований хаос. При малих значеннях коефіцієнту росту неколивальний характер перехідного процесу. При подальшому збільшенні α здійснюються виникають наступні біфуркації подвоєння періоду за сценарієм Фейгенбаума та хаос.

Висновки

В результаті дослідження процесів поведінки нейрона як нелінійного елемента у складі нейромережі або нейрокомп'ютера, знайдена математична модель нейрона в циклічному управлінні для випадку з вхідним потоком, визначена область його нормальних режимів, знайдені умови біфуркацій. При певних умовах нейрон може перейти в режим регулярних та нерегулярних коливань і хаосу. Отримані результати дозволяють підвищити ефективність роботи циклічних систем управління та формалізувати напрямки подальших досліджень щодо розробки ефективних систем управління з використанням методів нелінійної динаміки. У розглянутому випадку простої моделі нейрона зрив коливань робиться просто – розриваються ланцюги зворотного зв'язку для суматора. Для цього досить починати розраховувати підсумок у кожному циклі заново. Але розглянута нелінійність може бути присутня у будь-якому циклі PDCA. Тому у подальшому необхідно продовжити дану тематику досліджень.

Список літератури

1. Ситник, В.Ф. Системи підтримки прийняття рішень: Навч. посіб. / В.Ф. Ситник. – К.: КНЕУ, 2009. – 614 с.
2. Орлов, А.И. Теория принятия решений. Учебное пособие / А.И. Орлов. – М.: Издательство Экзамен, 2005. – 656 с.
3. Прохоров, С.А. Автоматизация комплексного управления безопасностью предприятия / С.А. Прохоров, А.А. Федосеев, А.В. Иващенко – Самара: СНЦ РАН, 2008 – 55 с.
4. Кононович, В.Г. Технічна експлуатація систем захисту інформації телекомунікаційних мереж загального користування. Частина 4 – Інформаційна безпека комунікаційних мереж та послуг. Реагування на атаки: навч. посібник / В.Г. Кононович, С.В. Гладиш; За ред. чл.-кор. МАЗ В.Г. Кононовича. – Одеса: ОНАЗ ім. О.С. Попова, 2009. – 208 с.
5. Круг, П.Г. Нейронные сети и нейрокомпьютеры: Учебное пособие по курсу «Микропроцессоры» / П.Г. Круг. – М.: Издательство МЭИ, 2002. – 176 с.
6. Чернавский, Д.С. Синергетика и информация (динамическая теория информации) / Д.С. Чернавский // Изд. 2. – М.: Едиториал УРСС, 2004. – 288 с.
7. ДСТУ ISO 9001:2009 Системи менеджменту якості. Вимоги [Аналог ISO 9001:2008]. – 32 с.
8. Фролов, А.В. Синтез и распознавание речи. Современные решения. Введение в нейронные сети [Электронный ресурс] / А.В. Фролов, Г.В. Фролов. / Библиотека братьев Фроловых. – Режим доступа: http://www.frolov-lib.ru/books/hi/ch04.html.
9. Расширенная модель искусственного нейрона [Электронный ресурс]. // Информац.-познават. журнала «Виктория».– Режим доступа: http://www.victoria.lviv.ua/html/oio/html/theme6_rus.htm (Дата посещения: 05.02.2015)
10. Короткий, С. Нейронные сети: алгоритм обратного распространения [Электронный ресурс] / С. Короткий. // «Лаборатория искусственного интеллекта». – Режим доступа: http://www.shestopaloff.ca/kyriako/Russian/Artificial_Intelligence/Some_publications/Korotky_Neuron_network_Lectures.pdf. (Дата посещения: 05.02.2015).
11. Нейронные сети – математический аппарат [Электронный ресурс]. // «BaseGroup: технологии анализа данных». – Режим доступа: http://www.basegroup.ru/library/analysis/neural/math (Дата посещения: 05.02.2015).
12. Шампайн, Л.Ф. Решение обыкновенных дифференциальных уравнений с использованием МАТЛАБ: Учебное пособие / Л.Ф. Шампайн, И. Гладвел, С. Томпсон. Пер. с англ. – СПб.: Издательство «Лань», 2009. – 304 с.
13. Малинецкий Г.Г. Математические основы синергетики. Хаос, структуры, вычислительный эксперимент. / Г.Г. Малинецкий. – М.: КомКнига, 2005. – 312 с.

ВЛИЯНИЕ НЕЛИНЕЙНОСТИ НЕЙРОНА НА ЦИКЛИЧЕСКУЮ СИСТЕМУ УПРАВЛЕНИЯ

В.Г. Кононович, О.Ю. Козлова, О.Ю. Кунянский

Одесский национальный политехнический университет,
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: vl_kononovich@ukr.net

Исследуется поведение нейрона, который является основным элементом нейронных сетей. Рассматривается логистическая модель нейрона как наипростейший нелинейный элемент. Представлена математическая модель нейрона в циклическом управлении для случая со входным потоком. Определены условия и характер возможных нерегулярных колебаний в нейроне. Полученные результаты позволяют повысить эффективность работы циклических систем управления.

Ключевые слова: нейрон, процессы управления, нелинейный элемент, нелинейная динамика, компьютерное моделирование, динамический хаос, бифуркации

INFLUENCE OF NON-LINEARITY OF NEURON ON CYCLIC SYSTEM MANAGEMENTS

V.G. Kononovich, O.U. Kozlova, O.U. Kunjanskij

Odessa National Polytechnic University,
Shevchenko Ave, 1, Odessa, 65044, Ukraine; E-mail: vl_kononovich@ukr.net

Behavior of neuron that is the basic element of neural networks is investigated. The logistic model of neuron as most simple nonlinear element is examined. The mathematical model of neuron is presented in a cyclic management for a case with an input stream. Terms and character of possible irregular vibrations are certain in a neuron. They got results allow increasing efficiency of work of cyclic control system.

Keywords: neuron, control processes, non-linear element, non-linear dynamics, computer simulation, dynamic chaos, bifurcation

UDC 519.688:004.032.26

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 297-301

DEVELOPMENT OF EFFECTIVE VOCABULARY STRUCTURES FOR THE SPEECH RECOGNITION TASKS

D.V. Zahanich, I.E. Mazurok

Odessa I.I. Mechnikov National University,
Dvoryanskaya st., 2, Odessa, 65026, Ukraine; e-mail: igor@mazurok.com

In this paper we describe a speech recognition method, which is optimized for mobile devices with limited computing power. This article is focused on reducing the size of a necessary dictionary and development of method of finding dictionary strings that match the input speech data. The proposed approach allows to solve the problem of speech recognition on mobile devices offline.

Keywords: speech recognition system, dictionary structure, fuzzy search.

Introduction

In recent decades mankind faced the problem of creating effective tools for data input. Keyboards and touch screens don't solve the problem efficient enough, because these tools are not natural for humans. Today personal computers are used not only by specialists, who are capable of fast typing, but also by regular users with poor computer skills. To solve this problem, various techniques and special devices for rapid data entry were developed. For example, sign writing received a new life and many devices allow you to enter words with line drawing on the touch screen. However, the core problem remains the same, because such methods also require special skills.

There are cases in which a user can't perform data input with regular tools or it is very difficult to do so. For example, it is quite difficult to type a message while driving a car. Moreover, the number of such scenarios increases with the spread of portable devices such as smartphones, automobile navigation devices, "smart home" systems, etc.

Conventional data input methods require performing some actions by user's hands and often prohibits user from performing others tasks at the moment of data input. However, speech recognition methods don't have these disadvantages. It took almost half a century to improve speech recognition systems for widespread usage [1]. Researchers and developers of such systems had to solve a lot of problems [2]:

- the problem of the extraction of the desired information from an audio signal;
- the problem of the classification of the extracted information from an audio signal;
- necessity of extensive training of speech recognition systems because of a large number of words pronunciation variants;
- low overall performance of speech recognition systems.

In recent years, speech recognition systems have become much more popular than before. This trend started due to progress made in speech technology area and accumulation of large amounts of data in the Internet. However, modern speech models require a considerable memory capacity and computation power. In this regard, the modern speech recognition systems perform most of calculations on the server side with multi-core processors and CPU/GPU clusters with great computational power [3]. Thus, the problem of low performance of speech recognition systems has not been solved yet. This becomes the major problem in case of mobile devices that can't be constantly connected to the Internet.

The primary *goal* of this article is to develop speech recognition method, which is capable of running on the mobile device with limited computational power. We focus on reducing the size of the dictionary, which is the key component of every speech recognition system.

Article

A proposed method of speech recognition is based on fuzzy string search. Thus, the speech recognition is achieved by searching the recognized phonetic transcription of the spoken word in the phonetic dictionary.

Speech sounds are infinitely varied. Accurate physical analysis may reveal that a person never produces exactly the same sounds. For example, the “k” sound in the word “kit” and “k” sound in word “skill” are not identical. However, these sounds are still represent the same phoneme /k/.

There are many sets of phonemes that are used in speech recognition tasks. In this article we use the following set of phonemes of the English language, based on Arpabet [7]. Arpabet is phonetic transcription code developed by Advanced Research Projects Agency (ARPA) as part of their Speech Understanding Project (1971-1976). This code represents each phoneme in the American dialect of English language as a distinct sequence of ASCII characters. The current phoneme set contains 39 phonemes (or more accurately, "phones"). Basically, "phones" are more or less similar classes of sounds.

Table 1.

Phones of American dialect of English and their corresponding examples of pronunciation

Phoneme	Example	Phoneme	Example
AA	odd	L	lee
AE	at	M	me
AH	hut	N	knee
AO	ought	NG	ping
AW	cow	OW	oat
AY	hide	OY	toy
B	be	P	pee
CH	cheese	R	read
D	dee	S	sea
DH	thee	SH	she
EH	Ed	T	tea
ER	hurt	TH	theta
EY	ate	UH	hood
F	fee	UW	two
G	green	V	vee
HH	he	W	we
IH	it	Y	yield
IY	eat	Z	zee
JH	gee	ZH	seizure
K	key		

Example 1. Phonetic transcription of word “University” is represented in the form of code "Y UW N AH V ER S AH T IY".

To recognize a spoken word we acquire set of spoken phones and compare it with the phonetic transcriptions of the words, which are presented in the dictionary.

Human speech is a complex phenomenon. Variants of pronunciation of even the same word may be quite different from each other depending on a speaker, a speaker accent, level of external noise, and many other factors. In this regard, even completely correct set of recognized spoken phones will rarely be the same as conventional phonetic transcription of actually spoken word. To solve this problem, every word in the dictionary is represented as a set of possible pronunciations variants in form of phonetic transcriptions. Phonetic transcriptions of the words are encoded as ASCII characters, so we can use methods of fuzzy string search.

However, we must take into account the phenomenon of coarticulation. It leads us to the necessity of storing a large number of possible pronunciations of every word, which differ in similar-sounding phonemes.

To solve this problem we provide phonetic encryption algorithm based on Soundex.

In this modification of Soundex, similar-sounding phonemes are replaced to ASCII characters, and then every sequence of the similar characters is reduced to one character. The resulting line is called phonetic hash string.

Table 2.

Similar phonemes and their encoding characters

Character	Phonemes
A	AA, AO, HH, OW, AY, OY
B	B, P
F	F, TH
K	K
S	S, SH, Z, ZH
G	G, JH
V	V, W
T	D, DH, T, CH
L	L
N	M, N, NG
R	ER, R
E	EH, EY, AH, AE

Example 2. Pronunciations "Y UW N AH V ER S AH T IY" and "UW N EH V R S EH T IY" of word “University” are encoded into the same phonetic hash string "INEVRSETI".

The proposed phonetic encryption algorithm allows us to present all known pronunciations of the word in the dictionary as theirs phonetic hash strings, greatly reducing the size of dictionary.

The final stage of the proposed speech recognition method consists in searching an encoded phonetic string of spoken word in the dictionary. We use fuzzy string search algorithm, which is known as the FB-Trie algorithm (Eng. Forward-backward trie). It was described in the "Fast approximate search in large dictionaries" [4].

The task of fuzzy string search in the dictionary is to choose for a given search query W subset P of all the words from the dictionary D, which has distance p to the search query and p does not exceed a certain threshold N:

$$P = \{P_i \mid P_i \in D \bigcap p(P_i, W) \le N\}.$$

We use the Damerau–Levenshtein distance because it is capable of detection up to 80% of all human misspellings.

The Damerau–Levenshtein distance is the distance between two strings, i.e., finite sequence of symbols, given by counting the minimum number of operations needed to

transform one string into the other, where an operation is defined as an insertion, deletion, or substitution of a single character, or a transposition of two adjacent characters.

FB-Trie algorithm is based on the concept of universal Levenshtein automaton and it applies a dictionary structure in the form of direct and reverse prefix tree in conjunction with splitting of keyword W into two approximately equal parts W1 and W2. Reverse trie contains inversions of all the vocabulary words.

The choice of this algorithm is based on the comparative analysis in “Indexing methods for approximate dictionary searching” [5]. According to this research, the FB-Trie algorithm provides a linear search time depending on the length of the search query and dictionary structure consumes only 300% of the memory, which is occupied by the raw dictionary. Pronunciation dictionary of 20,000 words, which corresponds to the number of the most frequently used words in the English language, requires amount of memory that does not exceed 5-6 megabytes.

For comparison purposes, we take the statistics of RAM limit for applications on the Android mobile platform. According to the document "Android Application Memory Limit" [6], to support all possible device configurations speech recognition system should not consume more than 16 megabytes of RAM. Thus, the proposed method of speech recognition can be implemented even on the weakest of modern mobile devices.

Conclusion

This paper presents a method for speech recognition, which is based on fuzzy string search in pronunciation dictionary. The main advantage of this method is relatively small amount of memory, which is consumed by dictionary component, so it could be implemented on the devices with small computing power. The second advantage is simplified training process, which can be done even without recorded audio data. Efficiency and a small amount of memory required to build the software allow the implementation of the proposed algorithm for mobile communication devices and low-power computers. However, the proposed method doesn’t take into account context of speech and language grammar model, so it may show high error rate it continuous speech.

The proposed solution to speech recognition task can be used effectively in cases where the speech recognition system has limited computational resources and is not supposed to recognize complex grammatical structures. Such cases may include remote control systems with a small vocabulary of control commands.

References

1. Xuedong Huang. Spoken Language Processing: A Guide to Theory, Algorithm and System Development / Xuedong Huang, Alex Acero, Hsiao-Wuen Hon. – Prentice Hall PTR, 2001. – 480 p.
2. Zakaria Kurdi. Automatic Speech Processing and Natural Languages / Zakaria Kurdi // ISTE Wiley. Volume 1. – 2016. – 720 p.
3. Dong Yu. Automatic Speech Recognition: A Deep Learning Approach / Dong Yu, Li Deng. – Springer, 2014. – 328 p.
4. Mihov, S. Fast approximate search in large dictionaries / Stoyan Mihov, Klaus U. Schulz // Computational Linguistics. – 2004. – V.30. – No.4. - pp. 18-23.
5. Boytsov L. Indexing methods for approximate dictionary searching: Comparative analysis. / Leonid Boytsov // ACM J.Exp. Algor. 16. – 2011. – Vol. 1 – C. 47-65.
6. Android Application Memory Limit [Electronic resourse]. Available from: https://drive.google.com/file/d/0B7Vx1OvzrLa3Y0R0X1BZbUpicGc/view
7. Yukio Tono. Developmental and Crosslinguistic Perspectives in Learner Corpus Research / Yukio Tono, Yuji Kawaguchi, Makoto Minegishi. – John Benjamins Publishing, 2012. – 263 p.

РОЗРОБКА ЕФЕКТИВНИХ СТРУКТУР СЛОВНИКА ДЛЯ ЗАДАЧ РОЗПІЗНАВАННЯ МОВЛЕННЯ

Д.В. Заганич, І.Є. Мазурок

Одеський національний університет ім. І.І. Мечникова,
вул. Дворянська, 2, Одеса, 65026; e- mail: igor@mazurok.com

У статті пропонуються модифікації алгоритмів розпізнавання мови, оптимізованих для реалізації на мобільних пристроях з обмеженою обчислювальною потужністю. Головна увага приділяється скороченню розміру необхідного словника і алгоритму пошуку введених мовних даних. Запропонований підхід дозволяє вирішувати задачі розпізнавання мови на мобільних пристроях у автономному режимі.
Ключові слова: система розпізнавання мови, структура словника, нечіткий пошук

РАЗРАБОТКА ЭФФЕКТИВНОЙ СТРУКТУРЫ СЛОВАРЯ В ЗАДАЧЕ РАСПОЗНАВАНИЯ РЕЧИ

Д.В. Заганич, И.Е. Мазурок

Одесский национальный университет им. И.И. Мечникова,
ул. Дворянская, 2, Одесса, 65026; e-mail: igor@mazurok.com

В статье предлагаются модификации алгоритмов распознавания речи, оптимизированные для реализации на мобильных устройствах с ограниченной вычислительной мощностью. Основное внимание уделяется сокращению размера необходимого словаря и алгоритму поиска вводимых речевых данных. Предложенный подход позволяет решать задачи распознавания речи на мобильных устройствах в автономном режиме.
Ключевые слова: система распознавания речи, структура словаря, нечеткий поиск

УДК 004

Informatics and Mathematical Methods in Simulation
Vol. 6 (2016), No. 3, pp. 302-309

СРАВНЕНИЕ СВОЙСТВ НОМИНАЛЬНОГО ТИПА ОБЪЕКТОВ РАЗЛИЧНЫХ ПРЕДМЕТНЫХ ПОДОБЛАСТЕЙ В РЕЛЯЦИОННЫХ БАЗАХ ДАННЫХ

М.Г. Глава

Одесский национальный политехнический университет,
пр. Шевченко, 1, Одесса, 65044, Украина; e-mail: glavamaria@mail.ru

Рассматривается проблема объединения моделей предметной области (ПрО). Предлагается сопоставлять объекты ПрО на основе значений свойств экземпляров этих объектов. Методы сопоставления свойств различаются в зависимости от типа шкал, в которых измеряются их значения. Предлагается построить онтологическую модель для свойств номинального типа и обработать ее концепты с помощью построения таблицы сопряженности признаков и метода анализа соответствий.

Ключевые слова: модель предметной области, предметная подобласть, свойства номинального типа, таблица сопряженности признаков, анализ соответствий.

Введение

На сегодняшний день работа любой организации любой сферы деятельности не возможна без использования информационных технологий. Огромный поток данных хранится и обрабатывается с помощью баз и хранилищ данных, что существенно упрощает управление и контроль деятельностью. Учитывая экономическое положение страны и анализируя рынок, подверженный реорганизации предприятий, можно сделать вывод о существовании проблемы объединения информационных систем.

Другим аспектом, подтверждающим актуальность данной проблемы, можно назвать реформы, проводимые в Украине, направленные на информатизацию учета имущественных прав, налоговой реформы и т.д. влекущие за собой создание единых информационных хранилищ страны, объединяя базы данных регионов.

Поскольку построение любой информационной системы начинается с описания предметной области (ПрО) и построения ее модели, решение поставленной проблемы сводится к объединению моделей ПрО рассматриваемых информационных систем. Результатом объединения моделей ПрО будет модель ПрО более высокого $(j+1)$-го порядка, а модели более низкого порядка j будем называть моделями предметных подобластей (ПрПО). Для получения модели $(j+1)$-го порядка, адекватной исследуемой ПрО [1], необходимо определить объекты, находящиеся на пересечении объединяемых ПрПО, т.е. определить подобные объекты, описываемые подобными свойствами. Предложенный подход позволит избежать избыточности и несогласованности данных ПрО $(j+1)$-го порядка, а также потери данных, наработанных за период существования объединяемых информационных систем, за счет выделения/объединения подобных объектов/свойств объектов и дополнения модели ПрО отличающимися объектами ПрПО.

В работе [2] предложена технология поиска проекций одних и тех же ПрПО (объектов ПрПО), в которой предлагается сопоставлять объекты на основе значений свойств экземпляров этих объектов. Алгоритмы сопоставления различаются в

зависимости от типа данных конкретного свойства. В данной работе предлагается алгоритм сопоставления свойств номинального типа.

Согласно предложенной технологии объекты потенциально подобных ПрПО необходимо подготовить к сопоставлению: выделить существенные свойства, основываясь на количестве информации каждого свойства и оценке экспертов; проранжировать по значимости объекты каждой сравниваемой ПрПО, основываясь на количестве и степени важности связей определенного объекта с другими в этой же ПрПО, и количестве значимых свойств, измеряемых определенной шкалой (порядковой, номинальной, числовой); отсортировать кортежи по значениям порядковых и номинальных свойств, соблюдая полученный ранее ранг свойств.

Алгоритм сопоставления свойств номинального типа

Алгоритм сопоставления свойств номинального типа выполняется после анализа порядковых свойств. Можно предположить, что предыдущие шаги сблизили ранг потенциально подобных объектов и их свойств. А в силу работы алгоритма сопоставления свойств порядкового типа, приблизилось соответствие кортежей сравниваемых объектов, соответственно, значения свойств номинального типа сопоставляются для выровненных мощностей множеств кортежей. Под рангом понимается место объекта/свойства в последовательности рассматриваемых объектов/свойств, определяемое при помощи порядковой шкалы.

Поскольку информацию об одном и том же объекте или действии можно представить разными номинальными значениями, посимвольное сравнение отклоняется. Сравнение семантических значений экземпляров свойств для решения представленной проблемы является нетривиальной задачей и практически не выполнимой в силу своей сложности и объема работы. А также семантический анализ исключает универсальность технологии сопоставления ПрПО, поскольку потребует огромных усилий и затрат времени экспертов в конкретной ПрО/ПрПО.

Для сопоставления свойств номинального типа предлагается построить модель онтологии, которая характеризует любые номинальные свойства.

На сегодняшний день термин «онтология» не имеет устоявшегося определения, но в информационных технологиях чаще всего используется определение, сформулированное Т. Р. Грубером: «Онтология – это спецификация концептуализации» [4]. Под концептуализацией понимается представление понятий, которые классифицируют объекты ПрО и связи между ними.

Выделяют следующие типы онтологий [5–7]: метаонтология (онтология верхнего уровня) – оперирует общими концептами и отношениями, которые не зависят от конкретной ПрО; предметная онтология – содержит понятия, описывающие конкретную ПрО и отношения между ними; онтология задач – содержит функции, с помощью которых производится преобразование входных данных в выходные.

Для решения поставленной задачи будем использовать метаонтологию, поскольку необходимо создать технологию сопоставления ПрПО, независящую от самих ПрПО.

Процесс создания метаописаний также называют аннотированием, которое может происходить как с помощью человека, так и с помощью специальных алгоритмов, реализованных программно [8].

В метаописаниях выделяют три типа: системные (служебные) метаданные – предназначены для функционирования информационных систем и систем управления знаниями; структурные метаданные – содержат справочную информацию об объектах, т.е. описания, использующиеся при идентификации и категоризации объектов в тех или иных целях; семантические метаописания – включают концептуальное (аннотированное) изложение содержания и смысла информации об объекте.

Остановимся на структурных метаданных поскольку данный подтип позволит создать модель онтологии номинальных свойств, не зависящую от конкретной ПрО/ПрПО, т.е. позволит измерить близость (подобие) свойств объектов разных ПрПО.

Для решения поставленной задачи к структурным метаданным номинальных свойств отнесем, например, количества пробелов в значениях свойств, прописных букв, знака пунктуации «.», знака пунктуации «,», знака пунктуации «-»; наличие аббревиатур; наличие кавычек; часть речи и т.д.

На следующем шаге необходимо обработать каждое номинальное свойство каждого объекта потенциально подобных ПрПО, заполнив индивиды (экземпляры) концептов (классов) онтологии. Мощность множества концептов онтологии будет равна количеству номинальных свойств, участвующих в сопоставлении. Мощность множества индивидов в каждом концепте онтологии будет равна мощности множества кортежей соответствующего свойства.

Далее предлагается сопоставлять попарно концепты онтологии двух сравниваемых ПрПО d_i и d_j в порядке, определенном на этапе ранжирования объектов и их свойств методами, представленными ниже. При выявлении низкой степени соответствия, сравнить текущий концепт ПрПО d_i, со следующим по рангу концептом ПрПО d_j для исключения ошибки в выборе ранга свойства. Если же и на этом этапе мера подобия низкая, то исключить данные свойства из рассмотрения.

После отбора свойств с высокой и средней мерой подобия, анализировать их с привлечением экспертов, поскольку программный метод не дает гарантии исключения ошибок при анализе номинальных свойств, но значительно сократит время работы экспертов и снизит вероятность ошибок 1-го рода.

Методы обработки концептов онтологии

Для определения меры подобия свойств необходимо обработать индивиды концептов онтологии и составить таблицу сопряженности (или кросстаблицу) по следующим признакам:

Количество пробелов = 0; 0 < Количество пробелов < 3; Количество пробелов ≥ 3; Количество прописных букв = 0; 0 < Количество прописных букв < 4; Количество прописных букв ≥ 4; Количество знака пунктуации «.» = 0; 0 < Количество знака пунктуации «.» < 3; Количество знака пунктуации «.» ≥ 3; Количество знака пунктуации «,» = 0; Количество знака пунктуации «,» = 1; Количество знака пунктуации «,» > 1; Количество знака пунктуации «-» = 0; Количество знака пунктуации «-» ≥ 1; Наличие аббревиатур = «есть»; Наличие аббревиатур = «нет»; Наличие кавычек = «есть»; Наличие кавычек = «нет»; Часть речи = «имя существительное»; Часть речи = «имя прилагательное»; Часть речи = «глагол». Таблица сопряженности или кросстабуляции – это таблица совместного распределения частот двух и более номинативных признаков, измеренных на одной группе объектов [9].

Мера сходства между концептами онтологии, а, соответственно, и между номинальными свойствами по комплексу признаков, анализировалась следующими методами: таксономический анализ Е. С. Смирнова [10], коэффициент взаимной сопряженности Пирсона [11], критерий согласия χ^2 Пирсона [12].

Таксономический анализ Е. С. Смирнова. В таксономическом анализе Е.С. Смирнова предполагается, что вес модальностей признаков различен в зависимости от частот их встречаемости. Чем реже встречается модальность в выборке, тем её вес больше и наоборот. При этом различают веса по присутствию и отсутствию одной и той же модальности. Следовательно, учитываются совпадения не только по присутствию тех или иных модальностей признаков, но и по их отсутствию. Всякому

несовпадению двух объектов по модальностям приписывается один и тот же вес «–1». Коэффициент сходства T_{ij} между *i*-ми концептами онтологий ПрПО d_i и d_j равен $T_{ij} = \frac{1}{M}\sum_{k=1}^{M} w_k$, где M – общее количество модальностей по всем признакам; w_k – вес k-й модальности либо по присутствию ее, либо по отсутствию, либо по несовпадению их.

Вес по присутствию k-ой модальности w_k^+, определяют по формуле: $w_k^+ = n_k / N_k$, а вес по отсутствию w_k^-: $w_k^- = N_k / n_k$, где N_k – число концептов, у которых данная модальность присутствует; n_k – число концептов, у которых модальность отсутствует.

Коэффициенты взаимной сопряженности Пирсона и Чупрова. Рассмотрим метод расчета коэффициента взаимной сопряженности Пирсона $K_{\text{П}}$. Коэффициент взаимной сопряженности Пирсона применяется, если необходимо оценить тесноту связи между неальтернативными признаками (т.е. имеющими более 2-х групп в таблице сопряженности), которые могут принимать любое число вариантов значений; $K_{П} = \sqrt{\varphi^2 / (1 + \varphi^2)}$, где φ^2 — показатель взаимной сопряженности, φ определяется как сумма отношений квадратов частот каждой клетки таблицы к произведению итоговых частот соответствующего столбца и строки. Вычитая из этой суммы 1, получим величину φ^2. Чем ближе коэффициент взаимной сопряженности Пирсона к 1, тем связь показателей больше: $\varphi^2 = \sum n_{xy}^2 / (n_x n_y) - 1$.

Наряду с коэффициентом взаимной сопряженности Пирсона для определения тесноты связи номинальных признаков используют аналогичный коэффициент Чупрова, который рассчитывается по следующей формуле: $K_{Ч} = \sqrt{\varphi^2 (K_1 - 1)^{-1} (K_2 - 1)^{-1}}$, где K_1 и K_2 – число групп первого и второго признака соответственно. Интерпретация результата расчета коэффициента сопряженности Чупрова аналогична коэффициенту Пирсона, чем ближе значение коэффициента к 1, тем больше связь показателей.

Критерий согласия χ^2-квадрат Пирсона. Критерий согласия хи-квадрат Пирсона – это непараметрический метод, который позволяет оценить значимость различий между фактическим (выявленным в результате исследования) количеством исходов или качественных характеристик выборки, попадающих в каждую категорию, и теоретическим количеством, которое можно ожидать в изучаемых группах при справедливости нулевой гипотезы. То есть, метод позволяет оценить статистическую значимость различий двух или нескольких относительных показателей (частот, долей).

Для расчета критерия χ^2 необходимо рассчитать ожидаемое количество наблюдений для каждой из ячеек таблицы сопряженности путем перемножения сумм рядов и столбцов с последующим делением полученного произведения на общее число наблюдений. Значение критерия χ^2 находится по следующей формуле: $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} (O_{ij} - E_{ij})^2 / E_{ij}$, где i – номер строки (от 1 до r), j – номер столбца (от 1 до c), O_{ij} – фактическое количество наблюдений в ячейке ij, E_{ij} – ожидаемое число наблюдений в ячейке ij. Для принятия решения о зависимости признаков необходимо сравнить значение расчетного критерия χ^2 с критическим значением при числе степеней свободы f. Число степеней свободы рассчитывается по формуле: $f = (r-1) \times (c-1)$. Критическое значение критерия χ^2-квадрат Пирсона определяется по таблице. В том случае, если полученное значение критерия χ^2 больше критического, гипотеза H_0 отклоняется и принимается альтернативная H_1. В качестве

гипотезы H_0 как правило принимается наиболее вероятный исход. В поставленной задаче предполагается, что свойства подобны, соответственно гипотезы имеют следующие значения: H_0 – свойства подобны; H_1 – свойства различны.

Апробация предложенных методов

Проверим работоспособность предложенных методов на примерах.

Предположим, дано 4 номинальных свойства: $ФИО_1$, $ФИО_2$, Город, Организация. Не будем учитывать ПрПО, объекты и ранги свойств для того, чтобы проверить качество работы предложенного подхода, сравнив заведомо подобные и разные свойства. Заполним индивиды (экземпляры) концептов (классов) онтологии по представленным свойствам. Построим таблицу сопряженности (кросстаблицу), подсчитав количество индивидов концепта онтологии, подходящих под определенный признак таблицы сопряженности.

Таксономический анализ Е. С. Смирнова. Определим меру сходства между концептами с помощью таксономического анализа Е. С. Смирнова. Выполним полный перебор концептов для определения максимальной и минимальной меры подобия. Признаки, по которым значения совпадают, отбрасываются. Отметим, что в данном методе расчет модальностей для двух концептов не целесообразен в силу своей специфики. Коэффициенты сходства T_{ij} представлены в табл. 1.

Таблица 1.

Коэффициенты сходства концептов

$ФИО_1$ и $ФИО_2$	0.0000
$ФИО_1$ и Город	-0.3333
$ФИО_1$ и Организация	-0.5333
$ФИО_2$ и Город	0.0000
$ФИО_2$ и Организация	-0.6000
Город и Организация	-0.5333

Проанализировав полученные результаты, можно сделать вывод, что наибольшее сходство между собой имеют свойства $ФИО_1$ и $ФИО_2$; $ФИО_2$ и Город. Последнее, как видно, ложно. Наименьшее сходство имеют свойства $ФИО_2$ и Организация; $ФИО_1$ и Организация; Город и Организация. Отметим, что мера подобия свойств $ФИО_1$ и Организация; $ФИО_2$ и Организация близки, что можно отнести к плюсам метода.

Коэффициенты взаимной сопряженности Пирсона и Чупрова. Рассчитаем коэффициенты взаимной сопряженности Пирсона $K_{\text{П}}$ и Чупрова $K_{\text{ч}}$. Строки со значениями признаков равными нулю, отбрасываем. Для удобства расчетов разместим данные в таблице, просуммировав количество значений признаков по строкам и столбцам. Затем рассчитаем частное от деления квадрата значения каждого признака на произведение сумм в соответствующие строке и столбце.

Для нахождения φ^2 необходимо просуммировать полученные итоги и вычесть 1:

$$\varphi^2 = 0.5019 + 0.5019 - 1 = 0.0038,\ K_{П} = \sqrt{0.0038/(1+0.0038)} = 0.0614,$$
$$K_{Ч} = \sqrt{0.0038/((11-1)(2-1))} = 0.0195.$$

Рассчитаем коэффициенты взаимной сопряженности Пирсона и Чупрова для других пар концептов. Для анализа результатов сведем результаты расчетов в табл. 2.

Таблица 2.

Коэффициенты взаимной сопряженности Пирсона и Чупрова

Свойства	$K_{\text{П}}$	$K_{\text{ч}}$
ФИО_1 и ФИО_2	0.0614	0.0195
ФИО_1 и Город	0.5360	0.1914
ФИО_1 и Организация	0.4635	0.1351
ФИО_2 и Город	0.5319	0.2093
ФИО_2 и Организация	0.4652	0.1357
Город и Организация	0.4912	0.1456

Коэффициенты у пар свойств ФИО_1 и Город, ФИО_2 и Город приблизительно равны, что подтверждает не случайность полученного результата. Эти же пары свойств имеют наиболее высокие коэффициенты. Наименьшие коэффициенты у пары свойств ФИО_1 и ФИО_2, причем значительно отличающиеся от коэффициентов других пар. Согласно интерпретации результатов, чем ближе коэффициенты к 1, тем выше зависимость признаков. Для поставленной задачи коэффициенты взаимной сопряженности Пирсона и Чупрова дают ложный результат.

Критерий согласия хи-квадрат Пирсона. Определим меру сходства между концептами с помощью критерия согласия χ^2 Пирсона.

Рассчитаем ожидаемые значения признаков, предварительно отбросив строки с нулевыми значениями признаков.

Определим значение χ^2 согласно формуле:

$$\chi^2 = \frac{50-50}{50} + \frac{50-50}{50} + \frac{49-49.5}{49.5} + \ldots + \frac{50-50}{50} + \frac{50-50}{50} + \frac{50-50}{50} = 3.0303.$$

Число степеней свободы будет $(11-1)(2-1)=10$. Критическое значение χ^2 распределения при уровне значимости $\alpha = 0.05$ составляет 18.307. Сравнив полученное значение χ^2 распределения и критическое, можем сделать вывод о том, что гипотеза H_0 о подобии концептов ФИО_1 и ФИО_2 принимается с вероятностью статистической ошибки первого рода не более 5%.

Выполним расчеты для других пар концептов и результаты сведем в табл. 3.

Таблица 3.

Расчетные и критические значения χ^2 распределения

Концепты	$\chi^2_{расч}$	$\chi^2_{крит}$	Число степеней свободы
ФИО_1 и ФИО_2	3.0303	18.307	10
ФИО_1 и Город	282.1724	19.675	11
ФИО_1 и Организация	218.9259	24.996	15
ФИО_2 и Город	276.1068	16.919	9
ФИО_2 и Организация	220.9956	24.996	15
Город и Организация	222.5498	24.996	15

При сравнении других пар свойств полученное значение χ^2 распределения значительно превышает критическое значение, следовательно, гипотеза H_0

отклоняется, исходя из чего можем сделать вывод о том, что это абсолютно разные свойства, поскольку наблюдаемые и ожидаемые частоты значительно отличаются.

Метод «Анализ соответствий». Наиболее информативным из предложенных методов оказался критерий согласия хи-квадрат Пирсона. На основе этого критерия разработан метод «Анализ соответствий».

Анализ соответствий – это разведочный метод анализа, позволяющий визуально и численно исследовать структуру таблиц сопряженности большой размерности [13].

Строки и столбцы исходной таблицы представляются точками пространства, между которыми вычисляется расстояние хи-квадрат. Далее требуется найти пространство небольшой размерности, в котором вычисленные расстояния минимально искажаются, и в этом смысле максимально точно воспроизвести структуру исходной таблицы с сохранением связей между признаками.

Сравним концепты онтологии используя функцию программного обеспечения STATISTICA компании StatSoft [14].

Подготовим данные аналогично предыдущим методам. Исключим из таблицы сопряженности строки со значениями признаков равными «0» по обоим концептам. Создадим таблицу данных в программе или импортируем подготовленную заранее, например, из Excel. На вкладке «Анализ» выберем «Многомерный анализ -> Анализ соответствий». В появившемся окне в качестве входных данных выберем пункт «Частоты без группирующих переменных», т.к. исходные данные представлены таблицей сопряженности, и нажмем кнопку «ОК». На экране появится окно с кратким результатом.

Согласно методу анализа соответствий свойств $ФИО_1$ и $ФИО_2$ значение p-value значительно больше 0.05, следовательно, гипотезу H_0 о подобии свойств следует принять. Все другие пары имеют значение p-value равное 0, что говорит о том, что гипотезу H_0 следует отклонить, т.е. среди свойств подобие не найдено.

Вывод

Исходя из результатов тестирования предложенных методов для сопоставления номинальных свойств объектов различных предметных подобластей предлагается использовать модель онтологии и метод «Анализ соответствий» в комплексе с подходами, предложенными в технологии, описанной в [2]. Но для уточнения полученных результатов необходимо провести дополнительные исследования на других номинальных свойствах. Возможно, для повышения достоверности данных, необходимо для сопоставления свойств применять несколько предложенных методов.

Применение предложенного подхода автоматизирует поставленную задачу, что значительно сократит время работы для лица, принимающего решение, и снизит вероятность ошибок 1-го рода в отличие от автоматического решения задачи.

Список литературы

1. Малахов, Е.В. Оценка степени адекватности баз данных как информационных моделей предметных областей / Е.В. Малахов // Тр. Одес. политехн. ун-та.– 2004.– Вып. 1(21).– С. 82–86.
2. Glava, M. Searching Similar Entities in Models of Various Subject Domains Based on the Analysis of Their Tuples / M. Glava, E. Malakhov // 2016 International Conference on Electronics and Information Technology (EIT´16), May 23—27, 2016, Odesa, Ukraine, 2016. - pp. 97–100. ISBN 978-1-5090-2224-3 (DOI: 10.1109/ICEAIT.2016.7501001; EID: 2-s2.0-84979503116).
3. Колесникова, С.И. Методы анализа информативности разнотипных признаков / С.И. Колесникова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2009. – №1(6). – С. 69–80.

4. Никоненко, А.А. Обзор баз знаний онтологического типа / А.А. Никоненко // Штучний інтелект. — 2009. — № 4. — С. 208-219.
5. Бухановский, А.В. Метаонтология исследовательского проектирования морских динамических объектов / А.В. Бухановский, Ю.И. Нечаев // Онтология проектирования. – 2012. – № 1. – С. 53–63.
6. Полетаева, Е.В. Принципы построения онтологии предметной области машиностроения [Электронный ресурс] / Е.В. Полетаева // Электронный научный журнал: Программные продукты, системы и алгоритмы.– 2015.– № 1. – С. 1–3 Режим доступа: http://swsys-web.ru/ontology-building-mechanical-engineering.html.
7. Палагин, А.В. Онтологические методы и средства обработки предметных знаний: монография / А.В. Палагин, С.Л. Крывый, Н.Г. Петренко. – Луганск: ВНУ им. В. Даля. – 2012. – 324 с.
8. Тузовский, А.Ф. Системы управления знаниями (методы и технологии) / А.Ф. Тузовский, С.В. Чириков, В.З. Ямпольский.– Под общ. ред. В.З. Ямпольского. – Томск: Изд-во НТЛ, 2005.– 260 с.
9. Наследов, А.Д. Математические методы психологического исследования. Анализ и интерпретация данных / А.Д. Наследов.– Учебное пособие.– СПб.: Речь, 2007.– 36 с.
10. Марков, А.А. Методика сравнительного анализа информационных объектов на базе экспертных методов многокритериальной оценки: автореф. дис. … к-та тех. наук : 05.13.01 / Марков, А. А. – М., 2003.– 166 с.
11. Громыко, Г.Л. Теория статистики: Учебник / Г.Л. Громыко.– Т11, 2-е изд., перераб. и доп.– М.: ИНФРА-М. – 2005.– 476 с.
12. Лапач, С.Н. Статистика в науке и бизнесе / С.Н. Лапач, А.В. Чубенко, П.Н. Бабич. – Киев: Морион, 2002.– 640 с.
13. Вуколов, Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие / Э.А. Вуколов. – 2-е изд.– М.: Форум, 2008.– 464 с.
14. Официальный сайт компании StatSoft [Электронный ресурс].– Режим доступа: http://www.statsoft.com.

ПОРІВНЯННЯ ВЛАСТИВОСТЕЙ НОМІНАЛЬНОГО ТИПУ ОБ'ЄКТІВ РІЗНИХ ПРЕДМЕТНИХ ПІДОБЛАСТЕЙ В РЕЛЯЦІЙНИХ БАЗАХ ДАНИХ

М.Г. Глава

Одеський національний політехнічний університет,
пр. Шевченко, 1, Одеса, 65044, Україна; e-mail: glavamaria@mail.ru

Розглядається проблема об'єднання моделей предметної області (ПрО). Пропонується зіставляти об'єкти ПрО на основі значень властивостей екземплярів цих об'єктів. Методи зіставлення властивостей розрізняються залежно від типу шкал, в яких вимірюються їх значення. Пропонується побудувати онтологічну модель для властивостей номінального типу і обробити її концепти за допомогою побудови таблиці спряженості ознак і методу аналізу відповідностей.
Ключові слова: модель предметної області, предметна підобласть, властивості номінального типу, таблиця спряженості ознак, аналіз відповідностей.

COMPARISON OF THE NOMINAL TYPE PROPERTIES OF OBJECTS OF DIFFERENT SUBJECT SUBDOMAINS IN RELATIONAL DATABASES

M. G. Glava

Odessa National Polytechnic University,
1, Shevchenko Ave., Odessa, 65044, Ukraine; e-mail: glavamaria@mail.ru

This article discusses the problem of the subject domain (SD) models merge. It is proposed to compare SD objects on the basis of the properties values of the tuples of these objects. The methods for comparing the properties of objects differ depending on the type of scales in which their values are measured. It is proposed to construct the ontological model for the nominal type properties and process its concepts by constructing contingency tables and correspondence analysis method.
Keywords: subject domain, model of subject domain, nominal type properties, contingency tables, correspondence analysis.

ІНФОРМАТИКА ТА МАТЕМАТИЧНІ МЕТОДИ В МОДЕЛЮВАННІ

Том 6, номер 3, 2016. Одеса – 110 с., іл.

ИНФОРМАТИКА И МАТЕМАТИЧЕСКИЕ МЕТОДЫ В МОДЕЛИРОВАНИИ

Том 6, номер 3, 2016. Одесса – 110 с., ил.

INFORMATICS AND MATHEMATICAL METHODS IN SIMULATION

Volume 6, No. 3, 2016. Odesa – 110 p.

Засновник: Одеський національний політехнічний університет

Зареєстровано Міністерством юстиції України 04.04.2011р.

Свідоцтво: серія КВ № 17610 - 6460Р

Друкується за рішенням Вченої ради Одеського національного політехнічного університету (протокол №1 від 06.09.2016)

Адреса редакції: Одеський національний політехнічний університет,

проспект Шевченка, 1, Одеса, 65044 Україна

Web: http://www.immm.opu.ua

E-mail: immm.ukraine@gmail.com