ISSN 2519-2310

COMPUTER SCIENCE AND CYBERSECURITY

ISSUE 2(10) 2018

V. N. Karazin Kharkiv National University Publishing

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені В.Н.КАРАЗІНА
ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени В.Н. КАРАЗИНА
V.N. KARAZIN KHARKIV NATIONAL UNIVERSITY

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА
КОМПЬЮТЕРНЫЕ НАУКИ И КИБЕРБЕЗОПАСНОСТЬ
COMPUTER SCIENCE AND CYBERSECURITY
(CS&CS)

Issue 2(10) 2018

Заснований 2015 року

Міжнародний електронний науково-теоретичний журнал
Международный электронный научно-теоретический журнал
International electronic scientific journal

The journal publishes research articles on theoretical, scientific and technical problems of effective facilities development for computer information communication systems and on information security problems based on advanced mathematical methods, information technologies and technical means.

Journal is published quarterly.
Approved for placement on the Internet by Academic Council of the Karazin Kharkiv National University (Nov. 26, 2018, protocol No.12)

Editorial office:
Karazin Kharkiv National University
Svobody Sq., 6, office 315a, Kharkiv, 61022, Ukraine (*North building of University, 3th floor*)
Phone: +38 (057) 705-10-83
E-mail: *cscsjournal@karazin.ua*
Web-page: *http://periodicals.karazin.ua/cscs (Open Journal System)*

Published articles have been internally and externally peer reviewed

TABLE OF CONTENTS

Issue 2(10) 2018

УДК 681.3.06

СТРАТЕГІЯ ВИБОРУ S-БЛОКІВ ДЛЯ НЕЛІНІЙНОГО ПЕРЕТВОРЕННЯ ШИФРУ СТРУМОК

Костянтин Лисицький

Харківський національний університет імені В.Н. Каразіна, майдан Свободи 4, м. Харків, 61022, Україна
lisickiy@ukr.net

Рецензент: Олександр Кузнецов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна, майдан Свободи 6, м. Харків, 61022, Україна
kuznetsov@karazin.ua

Надійшло в квітні 2018

Анотація: *Обговорюються особливості побудування S-перетворення шифру Струмок. Зокрема увага зосереджується на виборі S-блоків для цього перетворення. Оцінюються показники його випадковості, зокрема визначається мінімальне число циклів повторного використання цього перетворення, після котрого результат зашифрування приходить до показників стаціонарного стану випадкової підстановки. Розглядається можливість використання в S перетворенні випадкових S-блоків. Робиться висновок, що випадкові S-блоки не погіршують показників його випадковості. Пропонується удосконалення цієї конструкції S-перетворення, яке дозволяє покращити показники його випадковості. Воно будується на використанні замість паралельного набору S-блоків ланцюжка з керованих підстановок з додатним змішуючим лінійним перетворенням сегментів блоків даних на вході шару керованих підстановок. Показується, що ця конструкція дозволяє вже при однократному проході активізувати мінімум чотири S-блоки, замість одного S-блоку в вихідній конструкції. Запропонована удосконалена конструкція S-перетворення представляється більш швидкодіючою і дозволяє за три цикли активізувати мінімум 20-ть S-блоків замість 16-ти в вихідній конструкції, що додатково покращує показники випадковості удосконалення.*

Ключові слова: *потоковий шифр; показники випадковості; випадкові S-блоки; активні S-блоки; керовані підстановки.*

1 Вступ

В даній роботі розглянемо модифікований потоковий SNOW-подібний шифр Струмок, який був розроблений з метою підвищення швидкодії вихідної конструкції [1]. Ця мета досягається тим, що в шифрі використовується лінійний рекурентний регістр над полем GF(264). Уся інша частина шифру повторює оригінальну конструкцію, за виключенням того, що S-перетворення тепер береться не 32-ох бітним а 64-ох бітним. У представленій роботі зосереджується увага саме на конструкції S-перетворення. Відповідно до пропозиції розробників, воно повторює 64-х бітну «цеглинку» шифру Калина.

Відомо, що розробники шифру Калина при виборі S-блоків однією з основних вимог взяли, вимогу, щоб показники нелінійності булевих функцій, які входять в S-блоки, були не менше ніж 104 [2]. В своїх дослідженнях [3] вони довели, що ймовірність знайти підстановку з таким показником нелінійності дорівнює 0,0000007. В цій же роботі підкреслено, що для породження оптимального S-блоку необхідно в середньому перебрати 1.100.000 підстановок.

Аналогічним шляхом пішли і при виборі S-блоків в шифрах “Кузнечик” та білоруському шифрі. Але, якщо це було важливо для великих шифрів (*для забезпечення мінімізації числа циклів їх приходу до стану випадкової підстановки* [4]), то в даному випадку для 64-х бітного блоку даних, відкривається можливість без втрати стійкості використати S-блоки без такого жорстокого їх відбору, про котрий було зазначене вище.

В першій частині статті буде показано, що показники випадковості цього перетворення зберігаються і при використанні в ньому випадкових S-блоків, а в другій частині пропонується удосконалена конструкція S-перетворення для шифру Струмок.

2 Опис S перетворення шифру Струмок

Нелінійне перетворення шифру Струмок будується на основі 64-х бітної конструкції шифру Калина, наведеної на рис. 1 [2].

Вхідне 64-бітове значення ділиться на 8-м байтів, кожен з яких замінюється відповідно до заданої таблиці підстановки. У перетворенні використовується 8-м різних таблиць, по одній на кожен байт.

Таблиці підстановок перетворення повторюють таблиці підстановок шифру Калина.

Операція лінійного розсіювання (перемішування в колонці) використовує поліноміальний уявлення байтів в поле GF (28), утвореному неприводимим поліномом

$$m(x)=x^8+x^4+x^3+x^2+1,$$

або $\{01\}\{1d\}$ в шістнадцятковому представленні.

Слід зазначити, що цей незвідний поліном в шифрі Калина не збігається з утворюючим поліномом шифру Rijndael/AES.

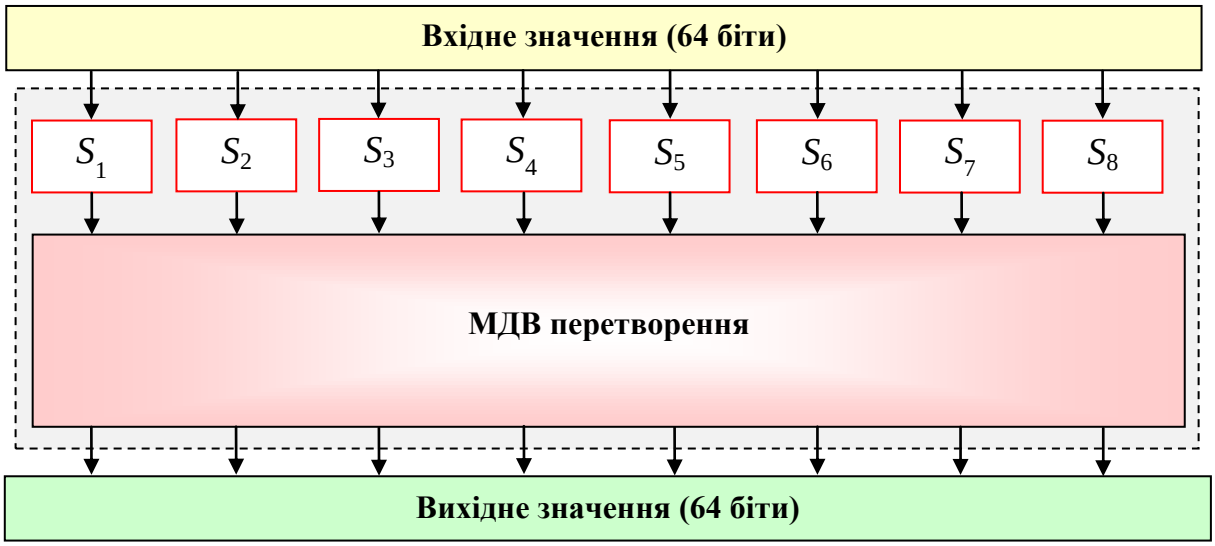

Рис. 1 – S перетворення шифру Струмок

Перемішування в колонках (MixColumns)

В ході перетворення MixColumns виконується послідовна обробка всіх колонок поточного стану. Кожна 8-байтна колонка розглядається як поліном над полем GF(28) с 8-ма термами, а в ході перетворення виконується множення цього поліному за модулем x^8+1 на фіксований поліном $c(x)$, де $c(x)=\{01\}x^7+\{05\}x^6+\{01\}x^5+\{08\}x^4+\{06\}x^3+\{07\}x^2+\{04\}x+\{01\}$.

Ця операція еквівалентна матричному множенню над GF(28) вихідного 8-байтного вектору на фіксовану матрицю, результат заноситься в 8-байтний вектор b (див. рис. 2).

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} = \begin{bmatrix} 01 & 01 & 05 & 01 & 08 & 06 & 07 & 04 \\ 04 & 01 & 01 & 05 & 01 & 08 & 06 & 07 \\ 07 & 04 & 01 & 01 & 05 & 01 & 08 & 06 \\ 06 & 07 & 04 & 01 & 01 & 05 & 01 & 08 \\ 08 & 06 & 07 & 04 & 01 & 01 & 05 & 01 \\ 01 & 08 & 06 & 07 & 04 & 01 & 01 & 05 \\ 05 & 01 & 08 & 06 & 07 & 04 & 01 & 01 \\ 01 & 05 & 01 & 08 & 06 & 07 & 04 & 01 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix}$$

Рис. 2 – Матричне множення над GF(28) вихідного 8-байтного вектору

Порядок обчислення елементів підсумкового вектору b пояснює рис. 3, де всі операції множення виконуються над полем GF(2^8).

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} = \begin{bmatrix} 01\cdot a_0 \oplus 01\cdot a_1 \oplus 05\cdot a_2 \oplus 01\cdot a_3 \oplus 08\cdot a_4 \oplus 06\cdot a_5 \oplus 07\cdot a_6 \oplus 04\cdot a_7 \\ 04\cdot a_0 \oplus 01\cdot a_1 \oplus 01\cdot a_2 \oplus 05\cdot a_3 \oplus 01\cdot a_4 \oplus 08\cdot a_5 \oplus 06\cdot a_6 \oplus 07\cdot a_7 \\ 07\cdot a_0 \oplus 04\cdot a_1 \oplus 01\cdot a_2 \oplus 01\cdot a_3 \oplus 05\cdot a_4 \oplus 01\cdot a_5 \oplus 08\cdot a_6 \oplus 06\cdot a_7 \\ 06\cdot a_0 \oplus 07\cdot a_1 \oplus 04\cdot a_2 \oplus 01\cdot a_3 \oplus 01\cdot a_4 \oplus 05\cdot a_5 \oplus 01\cdot a_6 \oplus 08\cdot a_7 \\ 08\cdot a_0 \oplus 06\cdot a_1 \oplus 07\cdot a_2 \oplus 04\cdot a_3 \oplus 01\cdot a_4 \oplus 01\cdot a_5 \oplus 05\cdot a_6 \oplus 01\cdot a_7 \\ 01\cdot a_0 \oplus 08\cdot a_1 \oplus 06\cdot a_2 \oplus 07\cdot a_3 \oplus 04\cdot a_4 \oplus 01\cdot a_5 \oplus 01\cdot a_6 \oplus 05\cdot a_7 \\ 05\cdot a_0 \oplus 01\cdot a_1 \oplus 08\cdot a_2 \oplus 06\cdot a_3 \oplus 07\cdot a_4 \oplus 04\cdot a_5 \oplus 01\cdot a_6 \oplus 01\cdot a_7 \\ 01\cdot a_0 \oplus 05\cdot a_1 \oplus 01\cdot a_2 \oplus 08\cdot a_3 \oplus 06\cdot a_4 \oplus 07\cdot a_5 \oplus 04\cdot a_6 \oplus 01\cdot a_7 \end{bmatrix}.$$

Рис. 3 – Порядок обчислення елементів підсумкового вектору b

3 Показники випадковості S перетворення

Наведемо результати оцінки очікуваних параметрів переходу модифікованого S перетворення шифру SNOW до стану випадкової підстановки. Будемо оцінювати показники його випадковості як циклової функції ітеративного шифру.

Відповідно до [5] необхідно виконати оцінку мінімального числа активних (*задіяних S-блоків*), після проходження яких S перетворення стає випадковою підстановкою. Це мінімальне число визначається диференціальними і лінійними показниками самих S-блоків, що застосовуються в S перетворенні, конструкціями і властивостями подальшого лінійного перетворення, а також значеннями показників доказової стійкості перетворення в цілому, які залежать від розміру його бітового входу. В роботі [5] цей зв'язок між зазначеними показниками визначений у вигляді двох співвідношень:

$$IPS_D=(DP^{\pi}_{\max})^k,\ IPS_L = 2^{k-1}\cdot\left(LP^{\pi}_{\max}\right)^k, \quad (1)$$

де: – $DP^{\pi}_{\max}$ і $LP^{\pi}_{\max}$ – максимальні значення диференціальної і лінійної ймовірностей підстановлювальних перетворень $\pi(x)$; – *IPSD* (*Differential Indicator of Provable Security*) – диференційний показник доказовою безпеки; – *IPSL* (*Linear Indicator of Provable Security*) – лінійний показник доказовою безпеки; – $k = k_{\min}$ – мінімальне число активних S-блоків, що беруть участь у формуванні переходу шифру до випадкової підстановці.

Користуючись розрахунковими співвідношеннями, встановленими в роботі [6], можна прийти до висновку, що очікуване значення максимуму диференціального переходу для шифру з 64-*x* бітовим входом (*S перетворення*) виявляється близьким до 68, а очікуване значення максимуму зміщення лінійного корпусу для шифру з 64-*x* бітовим входом виявляється близьким до 2^{65}. Відповідно отримаємо, що максимальні значення лінійної і диференціальної ймовірності для багатоциклового перетворення з 64-*x* бітовим входом виходять близькими один до одного і рівними приблизно 2^{-58}.

Виходячи з наведених вище співвідношень можна зробити висновок, що для перетворення з 64-*x* бітовим входом потрібно для приходу до стану випадкової підстановки за диференціальними показниками при використанні S-блоків з граничними показниками δ-рівномірності рівними $DP^{\pi}_{\max}= 2^{-6}$, (у відповідності з рівністю $2^{-58} = (2^{-6})^k$) $k_{\min}$=10 S-блоків.

Аналогічно, для приходу до стану випадкової підстановки за лінійними показниками при використанні S-блоків з граничними показниками нелінійності рівними $LP^{\pi}_{\max}=2^{-4,8}$ буде потрібно ($2^{-58} = 2^{k-1}\cdot\left(2^{-4,8}\right)^k$) $k_{\min}$=15 S-блоків.

У нашому випадку число активних S-блоків одноциклового перетворення дорівнює одному (на вході першого циклу перетворення активізується мінімум один S-блок і далі через те, що S перетворення містить 8-м S-блоків для двоциклового перетворення число S-блоків, що активізуються буде вже рівним 9-ти). Це означає, що для S-блоків з граничними диференціальними і лінійними показниками (для S-блоків з $DP^{\pi}_{\max}=LP^{\pi}_{\max}=2^{-6}$) S перетворення буде приходити до випадкової підстановці з запасом за три цикли.

4 Перспективи використання випадкових S-блоків

Методика виконання розрахунків представлена в роботі [7]. Нижче, у табл. 1 представлені результати розрахунків числа переходів різного типу в 17-ти рядках диференціальної таблиці випадкової байтової підстановки.

Таблиця 1 – Розрахунок числа переходів різного типу

Значення переходу таблиці	Число переходів диференціальної Таблиці	Число переходів в рядку	Число переходів в 17-ти рядках
12	1	0,003906	0,0664
10	10	0,039065	0,664
8	104	0,40625	6,906
6	830	3,24218	55,117

З представлених результатів випливає, що для 17-ти активних S-блоків при виборі в рядках максимально можливих переходів можна очікувати при випадкових входах в S-блоки:

- один перехід зі значенням 10;
- сім переходів зі значенням 8;
- дев'ять переходів зі значенням 6.

Всього 17-ть переходів (17-ть активних S-блоків). У припущенні, що вхід в перший S-блок (вхід в перший цикл) може бути обраний максимально можливим, обчислення в цьому випадку призводять до результату:

$$\left(\frac{12}{256}\right)\times\left(\frac{10}{256}\right)\times\left(\frac{8}{256}\right)^{7}\times\left(\frac{6}{256}\right)^{8}=2^{-87,4}.$$

Це означає, що випадкові S-блоки для цього шифру з великим запасом забезпечують за три цикли його прихід до стану випадкової підстановки.

Зауважимо, що для двох циклів маємо 9-ть активних S-блоків, і розрахунки для цього випадку представлені в табл. 2.

Таблиця 2 – Число переходів різного типу

Значення переходу таблиці	Число переходів диференціальної Таблиці	Число переходів в рядку	Число переходів в 9-ти рядках
12	1	0,003906	0,035154
10	10	0,039065	0,35158
8	104	0,40625	3,656
6	830	3,24218	29,17

З аналізу даних табл. 2 випливає, що для 9-ти активних S-блоків при виборі в рядках максимально можливих переходів можна очікувати при випадкових входах в S-блоки:

- нуль переходів зі значенням 10;
- чотири переходи зі значенням 8;

- п'ять переходів зі значенням 9.

У припущенні, що вхід в перший S-блок (вхід в перший цикл) може бути обраний максимально можливим, обчислення в цьому випадку призводять до результату:

$$\left(\frac{12}{256}\right)\times\left(\frac{8}{256}\right)^4\times\left(\frac{6}{256}\right)^4=2^{-46},$$

тобто двох циклів в цьому разі не вистачає для приходу до випадкової підстановки.

Будемо вважати, що процедура проходження S-блоків є випадковою і статистично незалежною. Методика розрахунків для цього випадку представлена в роботі [7].

У таблиці 3 представлені результати оцінки числа переходів і їх значень в 17 випадково взятих рядках таблиці ЛАТ. З результатів випливає, що для 17 активних S-блоків при використанні максимально можливих переходів можна очікувати при випадкових входах в випадкові S-блоки:

Таблиця 3 – Число переходів різного типу в 17-ти рядках лінійної таблиці

Значення переходу	Число переходів в таблиці ЛАТ	Число переходів в рядку таблиці ЛАТ	Число переходів в 17 випадково взятих рядках таблиці ЛАТ
±34	1,998	0,0078	0,1326
±32	4	0,0156	0,2652
±30	10	0,0392	0,6664
±28	28	0,1098	1,8666
±26	65	0,2588	4,3996
±24	146	0,572	9,724
±22	298	1,164	19,788

- один перехід зі значенням 30;
- два переходи зі значенням 28;
- чотири переходи зі значенням 26;
- дев'ять переходів зі значенням 24.

Найперший (один) S-блок береться з максимально можливим значенням переходу 34.

Вважаючи далі, що рядки в S-блок вибираються зі всієї безлічі 256-ти рядків, при цьому переходи по S-блокам йдуть в довільному порядку і здійснюються за найбільш ймовірного шляху, можемо виконати оцінку ймовірності приходу шифру до стану випадкової підстановки з випадковими S-блоками. Обчислення для значення $k = 17$ призводять до результату

$$2^{16}\times\left(\frac{34}{128}\right)^2\times\left(\frac{30}{128}\right)^2\left(\left(\frac{28}{128}\right)^2\right)^2\times\left(\left(\frac{26}{128}\right)^2\right)^4\times\left(\left(\frac{24}{128}\right)^2\right)^9=2^{16-74}=2^{-58}.$$

Таким чином, і в цьому випадку мінімальне число S-блоків, що активізуються задовольняє граничному значенню $k_{min}=17$. При цьому слід нагадати, що шифр Калина приходить до стану випадкової підстановки, як за диференціальним, так і за лінійними показниками на третьому-четвертому циклі.

5 Альтернативна конструкція S перетворення

Сама конструкція, що пропонується представлена на рис. 4. У своєї основі вона повторює конструкцію першого циклу шифру ШУП [7], тільки в цьому випадку замість SL перетворень тут виступають байтові S-блоки (*керовані підстановки*), а замість підсумування за модулем 2 сегментів на вході першого SL перетворення, використовується інша, більш ефективна схема змішування, яка заснована на багатошаровому підсумуванні за модулем 2 сегмен-

тів вхідного блоку даних. Крім того, в даному випадку підсумування виходу останнього S-блоку виконується тільки з виходом першого S-блоку.

Спочатку здійснюється розбивка вхідного блоку даних з 64 бітів на лівий і правий 32-*х* бітні підблоки і формується новий 64-*х* бітний блок даних. Він складається з нового лівого 32-*х* бітного підблоку (*який одержується за допомогою підсумування за модулем 2, лівого і правого 32-х бітних підблоків вихідного блоку даних*) та правого підблоку, шо повторює старий правий 32-*х* бітний підблок. Потім здійснюються аналогічні операції з новим лівим напівблоком і далі з новим лівим підблоком чергового напівблоку де він зводиться до байтового розміру.

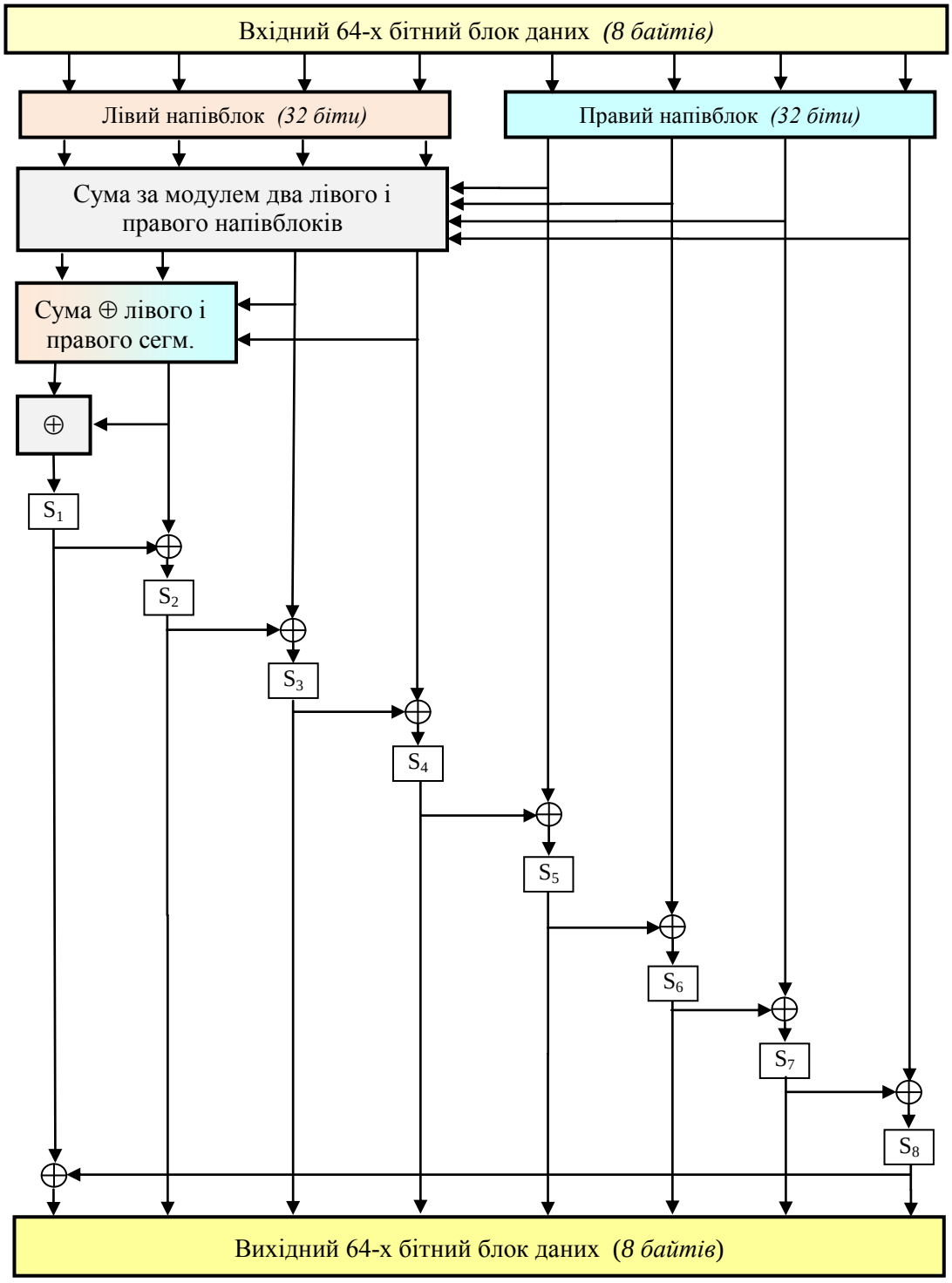

Рис. 4 – Альтернативна схема S перетворення на керованих підстановках

Розміщення рядків з сум байтів після додавання за модулем 2 ілюструється нижче.

$$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$$

Після першого XOR: $X_1 \oplus X_5$, $X_2 \oplus X_6$, $X_3 \oplus X_7$, $X_4 \oplus X_8$, X_5, X_6, X_7, X_8;

Після другого XOR: $X_1 \oplus X_5 \oplus X_3 \oplus X_7$, $X_2 \oplus X_6 \oplus X_4 \oplus X_8$, $X_3 \oplus X_7$, $X_4 \oplus X_8$, X_5, X_6, X_7, X_8;

Після третього XOR: $X_1 \oplus X_5 \oplus X_3 \oplus X_7 \oplus X_2 \oplus X_6 \oplus X_4 \oplus X_8$, $X_2 \oplus X_6 \oplus X_4 \oplus X_8$, $X_3 \oplus X_7$, $X_4 \oplus X_8$,

В результаті на вході першого S-блоку маємо $X_1 \oplus X_5 \oplus X_3 \oplus X_7 \oplus X_2 \oplus X_6 \oplus X_4 \oplus X_8$ – сума всіх байтів входу.

В гіршому випадку буде мати ненульову різницю байт X_5 ($X_1 \oplus X_5 = 0$, $X_1 = X_5 \neq 0$, $X_3 = X_7 = X_2 = X_6 = X_4 = X_8 = 0$, котрий буде активізувати на першому циклі 4-ри S-блоки.

Зауважимо, що для 256-бітного шифру с 4-и байтовими SL перетвореннями маємо активізацію на першому циклі в гіршому випадку 13-ти S-блоків.

Природно, що при трьох циклах у S перетворення буде активізуватися 20 S-блоків, і воно з великим запасом буде становитися випадковою підстановкою і при випадково згенерованих S-блоках.

Ця додатна операція потребує для свого виконання три XOR-и (*для 32-ох бітної платформи*). Всього в конструкції, що наведена на рис. 4 виконується 11-ть XOR-ів. Для виконання 64-*х* бітного перетворення шифру Калина потрібно буде використати (*при програмному виконанні*) 56 XOR-ів, не кажучи вже про операції множення байтів.

7 Висновки

Показано, що перетворення, яке складається з 8-ми паралельних S-блоків з наступним множенням виходів S-блоків на МДВ матрицю розміру 8×8 забезпечує граничні показники випадковості і при використанні випадкових S-блоків. Тобто в якості S-блоків в такому перетворенні можна використовувати практично підстановки з виходу генератора випадкових підстановок.

В статті запропонована та розглянута удосконалена конструкція S перетворення, в котрому для 64-х бітного вхідного блоку даних активізується (*вже на першому циклі*) мінімум чотири S-блоки, чого не дозволяють відомі конструкції циклових перетворень сучасних шифрів (*за виключенням шифру Лабіринт, в котрому на вході першого циклу використовується додаткове нелінійне перетворення*).

Посилання

[1] Opys potochnogo shyfru Strumok.

[2] Informacijni tehnologii'. Kryptografichnyj zahyst informacii'. Algorytm symetrychnogo blokovogo peretvorennja: DSTU 7624:2014. – K.: Derzhspozhyvstandart Ukrai'ny, - 2015. 238 p. – (Nacional'nyj standart Ukrai'ny).

[3] Rodinko, M.Yu. Improvement of the method for optimal S-boxes generation / M.Yu. Rodinko, R.V. Oliynykov, T.O. Hrinenko // Applied Radio Electronics. – 2015. – V.14. – No.4. pp. 315-320.

[4] Gorbenko, I. D. O dinamike prikhoda blochnykh simmetrichnykh shifrov k sluchainoi podstanovke / I. D. Gorbenko, K. E. Lisitskii // Radiotekhnika – Vseukr. mezhved. nauchn.-tekhn. sb. – 2014. – Vip. 176. pp. 27-39.

[5] Gorbenko, I. D. On Ciphers Coming to a Stationary State of Random Substitution / I.D. Gorbenko, K.E. Lisitskiy, D.S. Denisov // Universal Journal of Electrical and Electronic Engineering, 2, - 2014. 206-215. doi: 10.13189/ujeee.2014.020409.

[6] Lisitskii, K. E. Maksimal'nye znacheniya polnykh differentsialov i lineinykh korpusov blochnykh simmetrichnykh shifrov / K. E. Lisitskii // Tekhnologicheskii audit i rezervy proizvodstva. – 2014. – № 1/1 (15), pp. 47-52.

[7] Dolgov, V. I. The new concept of block symmetric ciphers design / V.I. Dolgov, I.V. Lisitska, K.Ye. Lisitskyi // doi: 10.1615 / TelecomRadEng. V. 76. Issue. 2. pp. 157-184.

Reviewer: Alexandr Kuznetsov, Doctor of Technical Sciences, Prof., Academician of the Academy of Sciences of Applied Radio Electronics, V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, 61022, Ukraine.
E-mail: kuznetsov@karazin.ua

Received: April 2018.

Authors:
Konstantin Lisitzky, postgraduate student of the Department of Information Systems and Technologies Security, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine.

E-mail: lisickiy@ukr.net

Strategy of selection of S-blocks for nonlinear transformation of cipher Strumok.

Abstract. The peculiarities of the construction of the S-transform of the Stream cipher "Strumok" are discussed. In particular, attention is focused on the choice of S-blocks for this transformation. The rates of its randomness are estimated, in particular, the minimum number of cycles of reuse of this transformation is determined, after which the result of encryption comes to the indicators of the stationary state of random substitution. Consideration of the possibility of using in the S transformation of random S-blocks. It is concluded that random S-blocks do not worsen the indicators of its randomness. It is proposed to improve this S-conversion design, which allows to improve its randomness. It is built on using instead of a parallel set of S-blocks of a managed substitution chain with a positive blended linear transformation of segments of data blocks at the input of a layer of controlled substitutions. It is shown that this design allows, at a single pass, to activate at least four S-blocks, instead of one S-block in the original design. The proposed enhanced S-conversion design appears to be more efficient and allows for at least 20 cycles of S-blocks to be activated in three cycles in the original design, which further improves the chance of improvement.

Keywords: stream cipher; random indices; random S-blocks; active S-blocks; controlled substitutions.

Рецензент: Александр Кузнецов, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы 6, г. Харьков, 61022, Украина.
E-mail: kuznetsov@karazin.ua

Поступила: Апрель 2018.

Авторы:
Константин Лисицкий, аспирант кафедры безопасности информационных систем и технологий, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы 4, г. Харьков, 61022, Украина.
E-mail: lisickiy@ukr.net

Стратегия выбора S-блоков для нелинейного преобразования шифра Струмок.

Аннотация. Обсуждаются особенности построения S-преобразования шифра "Струмок". В частности внимание сосредотачивается на выборе S-блоков для этого преобразования. Оцениваются показатели его случайности, в частности определяется минимальное число циклов повторного использования этого преобразования, после которого результат зашифрования приходит к показателям стационарного состояния случайной подстановки. Рассматривается возможность использования в S преобразовании случайных S-блоков. Делается вывод, что случайные S-блоки не ухудшают показателей его случайности. Предлагается усовершенствование этой конструкции S-преобразование, которое позволяет улучшить показатели его случайности. Оно строится на использовании вместо параллельного набора S-блоков цепочки из управляемых подстановок с положительным смешивающим линейным преобразованием сегментов блоков данных на входе слоя управляемых подстановок. Показывается, что эта конструкция позволяет уже при однократном проходе активизировать минимум четыре S-блоки, вместо одного S-блока в исходной конструкции. Предложенная усовершенствованная конструкция S-преобразования представляется более быстродействующей и позволяет за три цикла активизировать минимум 20-ть S-блоков вместо 16-ти в исходной конструкции, дополнительно улучшает показатели случайности совершенствования.

Ключевые слова: потоковый шифр; показатели случайности; случайные S-блоки; активные S-блоки; управляемые подстановки.

UDC 681.142.01

DATA SINGLE-ERROR CORRECTION METHOD OF A RESIDUE CLASS CODE

Viktor Krasnobayev[1], Sergey Koshman[1], Maksym Silenko[1], Sergey Moroz[2]

[1] V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkiv, 61022, Ukraine
v.a.krasnobaev@gmail.com, s_koshman@ukr.net, maximkasilenko@gmail.com

[2] Kharkiv Petro Vasylenko National Technical University of Agriculture, 19 Rizdviana St., Kharkiv, 61052, Ukraine
frost9i@ukr.net

Reviewer: Serhii Toliupa, Doctor of Sciences (Engineering), Full Prof., Taras Shevchenko National University of Kiev, 81 Lomonosova St., Kyiv, 03189, Ukraine.
tolupa@i.ua

Received on March 2018

Abstract: *This article describes the method of correcting of single errors in the residue class (RC). The study of this method makes it possible to create effective systems for monitoring data errors of computer systems in RC. The results of the analysis of the corrective capabilities of the arithmetic code in the RC showed a high efficiency of using non-positional code structures, due to the presence in the non-positional code structure of primary and secondary redundancy. The paper shows that the corrective capabilities of codes in the RC depend on the introduction of additional redundancy in the code structure. At the same time, if certain conditions are met, the data can be corrected by introducing only one test base. The article provides examples of monitoring and correcting of single errors were represented by codes in the RC.*

Keywords: *non-positional code structure; residue classes; positional numeral systems; minimum code distance; error-control coding; data diagnosing and correction.*

1 Introduction

In general, in order to verify, diagnose and correct errors a code structure requires a certain error-correcting capability. In this case, code is required to be introduced to data duplication, i.e. information redundancy should be implemented. All of the above fully refers to a non-positional code structure (NCS) in residue classes (RC) [1-3].

For each random RC the amount of redundancy $R = M_0 / M$ uniquely determines correction capability of a non-positional error-correcting code. Error correcting codes in RC can have any given values of minimum code distance (MCD) $d_{\min}^{(RC)}$, which depends on the value of redundancy R. The acquainted theorem [1] establishes a link between error-correcting code redundancy R, the value of MCD $d_{\min}^{(RC)}$, and the amount of RC check bases k.

Error-correcting code has MCD values $d_{\min}^{(RC)}$ in case when the degree of redundancy R is not less than the product $d_{\min}^{(RC)} - 1$ of RC bases. On the one hand we get $R \geq \prod_{i=1}^{d_{\min}^{(RC)}-1} m_{q_i}$, but on the other hand $R = M_0 / M = \prod_{i=1}^{n+k} m_i / \prod_{i=1}^{n} m_i = \prod_{i=1}^{k} m_{n+i}$. In this case, it's correct to state that $d_{\min}^{(RC)} - 1 = k$, or

$$d_{\min}^{(RC)} = k + 1. \quad (1)$$

There are two approaches to solve the problem of providing NCS with all required error-correcting properties in RC.

The first approach. If the requirements for error-correcting properties of NCS are known, for example, depending on amount of errors being detected $t_{\text{det.}}$ or corrected $t_{cor.}$ required information redundancy R should be introduced, using the amount of k or the value $\{m_{n+k}\}$ of check bases.

Redundancy R determines minimum code distance $d_{\min}^{(RC)}$ of NCS in RC.

Then, according to the error-control coding (ECC) theory for an ordered $(m_i < m_{i+1})$ RC we have that

$$t_{\text{det.}} \le d_{\min}^{(RC)} - 1, \tag{2}$$

$$t_{\text{det.}} \le k; \tag{3}$$

$$t_{cor.} \le \left[\frac{d_{\min}^{(RC)} - 1}{2} \right], \tag{4}$$

$$t_{cor.} \le \left[\frac{k}{2} \right]. \tag{5}$$

The second approach. For a given NCS $A_{RC} = (a_1 \| a_2 \| ... \| a_{i-1} \| a_i \| a_{i+1} \| ... \| a_n \| ... \| a_{n+k})$ (for a given value k) its error-correcting capabilities (determined by the $d_{\min}^{(RC)}$ value) of RC code are defined by the expressions (3) and (5).

Note that, if an ordered RC is extended by adding k check bases to n information modules, then MCD $d_{\min}^{(RC)}$ of the error-correcting code is increased by the value k (see expression (1)).

The values of $d_{\min}^{(RC)}$ can be also increased by decreasing the number n of information bases, i.e. by transitioning to less accurate calculations. It's clear that in RC between error-correcting R properties of error-control codes and calculation accuracy W inverse proportion exists. The same computer can perform arithmetical calculations or any other math operations both with high W accuracy but a low error-correcting R capability and with lower W accuracy, but with a higher capability R of error detection and correction in order to verify, diagnose and correct data faults, as well as to demonstrate higher data processing performance (the time to execute basic operations is inversely proportional to n information bases in RC) [2,4,5].

2 The main part

Now we'll analyze the process of single-error correcting data capability in RC given the minimal information redundancy by introduction of a single $(k = 1)$ check base. In this case, according to the error control coding theory in RC [1, 2], MCD is equal to the value $d_{\min}^{(RC)} = k + 1$. If $k = 1$, then MCD is $d_{\min}^{(RC)} = 2$, which, as according to the general error control coding theory, ensures any single-error detection (an error in one of the residues a_i $(i = \overline{1, n+1})$) in NCS.

In general, just as in the positional numeral system (PNS), the process of data error correction in RC consists of three stages. The first stage – data checking (correctness or incorrectness verification of the initial number A_{RC}). On the second stage diagnosing the false $\tilde{A}_{RC}$ number (detection of a single corrupted residue $\tilde{a}_i$ of the number $\tilde{A}_{RC}$ to the base m_i in RC). And, finally, on the third stage correcting the invalid residue $\tilde{a}_i$ to its true value a_i of the number, i.e. correcting false $\tilde{A}_{RC}$ number (getting the correct number $A_{RC} = \tilde{A}_{cor.}$).

The degree of information redundancy R (code error-correcting property) is estimated by the value of MCD $d_{\min}^{(PNS)}$. As previously noted, the value of MCD is defined by the ratio $d_{\min}^{(RC)} = k + 1$, where k is the amount of check bases in an ordered RC.

Let's start with the NCS $A_{RC} = (a_1 \| a_2 \| ... \| a_{i-1} \| a_i \| a_{i+1} \| ... \| a_n \| ... \| a_{n+k})$ in RC having a minimal $(k = 1)$ additional information redundancy. In this situation it's considered that $d_{\min}^{(RC)} = 2$.

According to the error control coding theory in PNS if the minimum code distance is granted to be $d_{\min}^{(PNS)} = 2$, a single error in a code structure is ensured to be detected. In PNS a single error is

understood as a corruption of a single information bit, for instance $0 \to 1$ or $1 \to 0$. In order to correct this single error it's required to ensure the condition, when $d_{\min}^{(PNS)} = 3$.

Contrary to PNS, a single error in RC is understood as a corruption of a single residue a_i modulo m_i. Inasmuch as the residue a_i of the number $A_{RC} = (a_1 \| a_2 \| ... \| a_{i-1} \| a_i \| a_{i+1} \| ... \| a_n \| a_{n+1})$ modulo m_i contains $z = \{[\log_2(m_i - 1)] + 1\}$ binary bits, then it's formally correct to be considered that if $d_{\min}^{(RC)} = 2$ $(k = 1)$ is within limits of a single residue a_i, an error cluster can be detected in RC, with its length not exceeding z binary bits. However in RC, as it is shown in literature [1, 2, 5], there are some cases when a single errors can be corrected while $d_{\min}^{(RC)} = 2$.

In the light of specific features and properties of NCS representation in RC an error-correcting capability given $d_{\min}^{(RC)} = 2$ can be explained in the following manner.

1. A single error in PNS and in RC are different concepts, as it was shown before. With that being said, MCD $d_{\min}^{(PNS)}$ for PNS and $d_{\min}^{(RC)}$ for RC has different meaning and measure.

2. Existing (implicitly) intrinsic (natural, primal) information redundancy in NCS, being stored in residues $\{a_i\}$ due to their forming procedure, has a positive effect (from the perspective of increasing data jam-resistance, transfer and processing reliability) that kicks in only with the presence of a subsidiary (artificial, secondary) information redundancy. An artificial information redundancy in NCS is being introduced by using (additionally to n information bases) k check bases in RC. A distinguishing feature of RC is its significant display of the intrinsic information redundancy only if the subsidiary one is also present, due to introduction of check bases.

3. As shown in [1,2,5], error control code in RC with mutually prime bases has the MCD value of $d_{\min}^{(RC)}$ only if the information redundancy level is not less than the product of any $d_{\min}^{(RC)} - 1$ bases of a given RC.

The availability and interaction of primary and secondary redundancies during the subsidiary tests (time redundancy usage) of error-correcting process, which may provide a single-error error-correcting capability in RC, while $d_{\min}^{(RC)} = 2$ (given $k = 1$).

Indeed, according to the expressions (3) and (5) for an ordered RC following conclusions can be made: with a single $(k = 1)$ check base m_{n+1} in RC, the NCS $A = (a_1 \| a_2 \| ... \| a_{i-1} \| a_i \| a_{i+1} \| ... \| a_n \| a_{n+1})$ can have several values of $d_{\min}^{(RC)}$. In this case, it depends on the value of check residue m_{n+1}. If, for every different RC modulus condition $m_i < m_{n+1}$ $(i = \overline{1, n})$ is met, then conclusion can be made that $d_{\min}^{(RC)} = 2$, as according to the expression (1), and $t_{\text{det.}} = 1$, according to the expression (2). If the condition $m_i \cdot m_j < m_{n+1}$ $(i, j = \overline{1, n};\ i \neq j)$ is met across the totality of $\{m_i\}$ information bases for a random modulus pair, then $d_{\min}^{(RC)} = 3$ and $t_{\text{det.}} = 2$.

Thus, for the NCS in RC given $k = 1$, the MCD $d_{\min}^{(RC)}$ can vary, depending on the value of RC check base m_{n+1}. Assume, RC is given information bases $m_1 = 3$, $m_2 = 4$, $m_3 = 5$, $m_4 = 7$ and moreover $m_k = m_{n+1} = m_5 = 11$. In this case error verification of any single corrupted NCS residue can be ensured.

Assume, for example, $m_k = m_{n+1} = 61$. Ad hoc, we'll draw up a Table 1, mapping information bases to check bases. As Table 1 shows, number representation specificity in RC in some cases allows not only to detect an error, but to find a place of its occurrence with the use of a single check base, which would be impossible to do in the PNS, utilizing existing methods of detecting and correcting errors.

Let's assume, that in the corrupted ($\tilde{A} \geq M$) number $\tilde{A} = (a_1 \| a_2 \| ... \| a_{i-1} \| \tilde{a}_i \| a_{i+1} \| ... \| a_n \| a_{n+1})$ the error $\tilde{a}_i = (a_i + \Delta a_i) \bmod m_i$ is verified to be present in the residue a_i modulo m_i.

Table 1 – Research results of error-correcting capabilities of error control codes in RC ($l=1$)

$m_k = m_{n+1} = m_5 = 61;\ d_{\min}^{(RC)} = k+1 = 2,\ \prod_{i=1}^{3} m_i \le m_5$.							Max. amount of detectable data errors in RC	Max. amount of correctable data errors in RC
RC information bases				$\prod_{r=1}^{k'} m_{i_r} \le m_{n+1}$	k'	$d_{\min}^{(RC)'} = k'+1$		
$m_1 = 3$	$m_2 = 4$	$m_3 = 5$	$m_4 = 7$					
+	–	–	–	$3 < 61$	1	2	1	0
–	+	–	–	$4 < 61$	1	2	1	0
–	–	+	–	$5 < 61$	1	2	1	0
–	–	–	+	$7 < 61$	1	2	1	0
+	+	–	–	$3 \cdot 4 = 12 < 61$	2	3	2	1
+	–	+	–	$3 \cdot 5 = 15 < 61$	2	3	2	1
+	–	–	+	$3 \cdot 7 = 21 < 61$	2	3	2	1
–	+	+	–	$4 \cdot 5 = 20 < 61$	2	3	2	1
–	+	–	+	$4 \cdot 7 = 28 < 61$	2	3	2	1
–	–	+	+	$5 \cdot 7 = 35 < 61$	2	3	2	1
+	+	+	–	$3 \cdot 4 \cdot 5 = 60 < 61$	3	4	3	2

We'll take a look at the ratio, which makes it possible to correct an error in a given residue $\tilde{a}_i$ [1].

It's clear that:

$$\tilde{A} = (A + \Delta A) \bmod M_0. \tag{6}$$

Basing on that the error magnitude can be equated to $\Delta A = (0 \| 0 \| ... \| 0 \| \Delta a_i \| 0 \| ... \| 0 \| 0)$, then the correct ($A < M$) number A can be expressed as follows:

$$A = (\tilde{A} - \Delta A) \bmod M_0 = \left[(a_1 \| a_2 \| ... \| a_{i-1} \| \tilde{a}_i \| a_{i+1} \| ... \| a_n \| a_{n+1}) - (0 \| 0 \| ... \| 0 \| \Delta a_i \| 0 \| ... \right.$$

$$\left. ... \| 0 \| 0) \right] \bmod M_0 = [a_1 \| a_2 \| ... \| a_{i-1} \| (\tilde{a}_i - \Delta a_i) \bmod m_i \| a_{i+1} \| ... \| a_n \| a_{n+1}] \bmod M_0.$$

We'll quantify the value of A. Inasmuch number A is correct, i.e. is contained in numerical interval $[0, M)$, then the following inequality will be fulfilled:

$$A = (\tilde{A} - \Delta A) \bmod M_0 < M\ . \tag{7}$$

Basing on the value of the error ΔA is equal to $\Delta A = \Delta a_i \cdot B_i$, then the inequality (7) will be expressed as:

$$\tilde{A} - \Delta a_i \cdot B_i - r \cdot M_0 < M \text{ or}$$

$$\tilde{A} - \Delta a_i \cdot B_i - r \cdot M_0 < M_0 / m_{n+1} (r = 1, 2, 3, ...),$$

$$\tilde{A} - (\tilde{a}_i - a_i) \cdot B_i - r \cdot M_0 < M_0 / m_{n+1},$$

$$\tilde{A} - (a_i - \tilde{a}_i) \cdot B_i - r \cdot M_0 < M_0 / m_{n+1},$$

$$(a_i - \tilde{a}_i) \cdot B_i < M_0 / m_{n+1} - \tilde{A} + r \cdot M_0,$$

$$a_i - \tilde{a}_i < (M_0 / m_{n+1}) / B_i - \tilde{A} / B_i + r \cdot M_0 / B_i,$$

$$a_i < \tilde{a}_i + (M_0 / m_{n+1}) / B_i - \tilde{A} / B_i + r \cdot M_0 / B_i. \tag{8}$$

Since the orthogonal base of RC module m_i takes the form of $B_i = \bar{m}_i \cdot M_0 / m_i$, then the expression (8) shows up as:

$$a_i < \tilde{a}_i + (m_i + r \cdot m_i \cdot m_{n+1}) / (\bar{m}_i \cdot m_{n+1}) - \tilde{A} / B_i \quad \text{or}$$
$$a_i < \tilde{a}_i + m_i (1 + r \cdot m_{n+1}) / (\bar{m}_i \cdot m_{n+1}) - \tilde{A} / B_i . \quad (9)$$

Inasmuch as the value of the residue a_i is a natural number, then the value of $m_i(1 + r \cdot m_{n+1}) / (\bar{m}_i \cdot m_{n+1}) - \tilde{A} / B_i$, as shown in the expression (9), should be an integer. Thus, taking an integral part of the last ratio, the formula for correcting error in the residue $\tilde{a}_i$ of the number $\tilde{A}$ will be:

$$a_i = (\tilde{a}_i + [m_i \cdot (1 + r \cdot m_{n+1}) / (\bar{m}_i \cdot m_{n+1}) - \tilde{A} / B_i) \bmod m_i]. \quad (10)$$

We'll have a look at the examples of error correction in RC.

Example №1. Perform data verification of the number $A_{RC} = (0 \| 0 \| 0 \| 0 \| 5)$ and correct it if required, when RC was given information $m_1 = 3$, $m_2 = 4$, $m_3 = 5$, $m_5 = 7$ and check $m_k = m_5 = 11$ bases. Thereby, $M = \prod_{i=1}^{n} m_i = \prod_{i=1}^{4} m_i = 420$ and $M_0 = M \cdot m_{n+1} = 420 \cdot 11 = 4620$. Orthogonal RC bases B_i $(i = \overline{1, n+1})$ are shown in Table 2.

I. Data verification of $A_{RC} = (0 \| 0 \| 0 \| 0 \| 5)$. According to the control procedure [1] the value will be defined as:

Table 2 – Orthogonal RC bases B_i $(l = 1)$

$B_1 = (1 \| 0 \| 0 \| 0 \| 0) = 1540, \quad \bar{m}_1 = 1$
$B_2 = (0 \| 1 \| 0 \| 0 \| 0) = 3465, \quad \bar{m}_2 = 3$
$B_3 = (0 \| 0 \| 1 \| 0 \| 0) = 3696, \quad \bar{m}_3 = 4$
$B_4 = (0 \| 0 \| 0 \| 1 \| 0) = 2640, \quad \bar{m}_4 = 4$
$B_5 = (0 \| 0 \| 0 \| 0 \| 1) = 2520, \quad \bar{m}_5 = 6$

$$A_{PNS} = \left(\sum_{i=1}^{n+1} a_i \cdot B_i \right) \bmod M_0 = \left(\sum_{i=1}^{5} a_i \cdot B_i \right) \bmod M_0 = (a_1 \cdot B_1 + a_2 \cdot B_2 + a_3 \cdot B_3 +$$
$$a_4 \cdot B_4 + a_5 \cdot B_5) \bmod M_0 = (0 \cdot 1540 + 0 \cdot 3465 + 0 \cdot 3696 + 0 \cdot 2640 + 5 \cdot 2520) \bmod 4620 =$$
$$= (5 \cdot 2520) \bmod 4620 = 12600 (\bmod 4620) = 3360 > 420 .$$

Thus, in the process of data verification it was evaluated, that $A_{RC} = 3360 > M = 420$. In this case, with the possibility of only single errors appearing, conclusion is made that the number in question $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$ is incorrect ($3360 > M = 420$).

In order to correct the number $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$ data is required to be verified first, i.e. corrupted residue $\tilde{a}_i$ has to be detected. Once done, the true value of the residue a_i modulo m_i needs to be defined, whereupon the corrupted residue $\tilde{a}_i$ should be corrected.

II. Data diagnosing of $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$. According to the mapping method [1, 2], possible projections $\tilde{A}_j$ of the number $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$ are:

$$\tilde{A}_1 = (0 \| 0 \| 0 \| 5), \ \tilde{A}_2 = (0 \| 0 \| 0 \| 5), \ \tilde{A}_3 = (0 \| 0 \| 0 \| 5),$$
$$\tilde{A}_4 = (0 \| 0 \| 0 \| 5) \text{ and } \tilde{A}_5 = (0 \| 0 \| 0 \| 0).$$

Computational formula for the values $\tilde{A}_{j\,PNS}$ of PNS number projections is written as [1]:

$$\tilde{A}_{j\,PNS} = \left(\sum_{\substack{i=1; \\ j=\overline{1,\,n+1}.}}^{n} a_i \cdot B_{ij} \right) \bmod M_j = (a_1 \cdot B_{1j} + a_2 \cdot B_{2j} + \ldots + a_n \cdot B_{nj}) \bmod M_j . \qquad (11)$$

According to the expression (11) we'll compute all the values of $\tilde{A}_{j\,PNS}$. Once done, we will make $(n+1)$ comparison of the $\tilde{A}_{j\,PNS}$ numbers to the number $M = M_0 / m_{n+1}$. If there are any numbers not being contained in the informational numeric interval $[0, M)$, which contains k correct numbers (i.e. $\tilde{A}_k \geq M$), among $\tilde{A}_i$ projections, then conclusion is made that these k residues of the number A are not corrupted. Only the residues among the rest $[(n+1)-k]$ number $\tilde{A}_{RC}$ residues can be false.

The set of the active quotient residues for a given RC and the totality of the quotient B_{ij} orthogonal bases are shown in Table 3 and Table 4 respectively.

Table 3 – Set of the active quotient RC residues ($l = 1$)

j \ i	m_1	m_2	m_3	m_4	M_j
1	4	5	7	11	1540
2	3	5	7	11	1155
3	3	4	7	11	924
4	3	4	5	11	660
5	3	4	5	7	420

Table 4 – The totality of the quotient orthogonal RC bases B_{ij} ($l = 1$)

B_{ij} j \ i	1	2	3	4
1	385	616	1100	980
2	385	231	330	210
3	616	693	792	672
4	220	165	396	540
5	280	105	336	120

Now then (Table 4):

$$\tilde{A}_{1PNS} = \left(\sum_{i=1}^{4} a_i \cdot B_{i1} \right) \bmod M_1 = (a_1 \cdot B_{11} + a_2 \cdot B_{21} + a_3 \cdot B_{31} + a_4 \cdot B_{41}) \bmod M_1 =$$
$$= (0 \cdot 385 + 0 \cdot 616 + 0 \cdot 1100 + 5 \cdot 980) \bmod 1540 = 280 < 420 .$$

Arriving at conclusion, that the residue a_1 of the number $\tilde{A}_1$ is possibly a corrupted residue $\bar{a}_1$;

$$\tilde{A}_{2PNS} = \left(\sum_{i=1}^{4} a_i \cdot B_{i2} \right) \bmod M_2 = (a_1 \cdot B_{12} + a_2 \cdot B_{22} + a_3 \cdot B_{32} + a_4 \cdot B_{42}) \bmod M_2 =$$
$$= (0 \cdot 385 + 0 \cdot 231 + 0 \cdot 330 + 5 \cdot 210) \bmod 1155 = 1050 > 420 .$$

Hence, the residue a_2 is ensured being not corrupted;

$$\tilde{A}_{3PNS} = \left(\sum_{i=1}^{4} a_i \cdot B_{i3} \right) \bmod M_3 = (a_1 \cdot B_{13} + a_2 \cdot B_{23} + a_3 \cdot B_{33} + a_4 \cdot B_{43}) \bmod M_3 =$$
$$= (0 \cdot 616 + 0 \cdot 693 + 0 \cdot 792 + 5 \cdot 672) \bmod 924 = 588 > 420 .$$

Deduced, the residue a_3 is ensured being not corrupted;

$$\tilde{A}_{4PNS} = \left(\sum_{i=1}^{4} a_i \cdot B_{i4} \right) \bmod M_4 = (a_1 \cdot B_{14} + a_2 \cdot B_{24} + a_3 \cdot B_{34} + a_4 \cdot B_{44}) \bmod M_4 =$$

$$= (0 \cdot 220 + 0 \cdot 165 + 0 \cdot 369 + 5 \cdot 540) \bmod 660 = 60 < 420.$$

Conclusion: the residue a_4 modulo m_4 of the number $\tilde{A}_4$ is possibly a corrupted residue $\bar{a}_4$;

$$\tilde{A}_{5PNS} = \left(\sum_{i=1}^{4} a_i \cdot B_{i5} \right) \bmod M_5 .$$ Since $M_5 = M = 420$,

the residue $\bar{a}_5$ of the check module $m_k = m_5$ will be always among the totality of possibly corrupted residues $\bar{a}_i$ of RC number.

Overall conclusion. During data diagnosing of $\tilde{A} = (0 \| 0 \| 0 \| 0 \| 5)$ in NCS, the residues $a_2 = 0$ and $a_3 = 0$ were ensured not being corrupted. The residues to the bases m_1, m_4 and m_5 might be corrupted, i.e. the residues $\bar{a}_1 = 0$, $\bar{a}_4 = 0$ and $\bar{a}_5 = 5$. In this case it's required to correct the residues $\bar{a}_1$, $\bar{a}_4$ and $\bar{a}_5$.

III. Correcting data errors $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$. According to the acquainted [1] expression:

$$a_i = \left(\bar{a}_i + \left[\frac{m_i \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_i} - \frac{\tilde{A}}{B_i} \right] \right) \bmod m_i , \qquad (12)$$

we will correct possibly $\bar{a}_1$, $\bar{a}_4$ and $\bar{a}_5$ corrupted residues a_1, a_4 and a_5, where $r = 1, 2, 3, \ldots$.

It turns out that:

$$a_1 = \left(\bar{a}_1 + \left[\frac{m_1 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_1} - \frac{\tilde{A}}{B_1} \right] \right) \bmod m_1 = \left(0 + \left[\frac{3 \cdot (1 + r \cdot 11)}{11 \cdot 1} - \frac{3360}{1540} \right] \right) \bmod 3 =$$

$$= (0 + [3,27 - 2,18]) \bmod 3 = (0 + [1,09]) \bmod 3 = (0 + 1) \bmod 3 = 1;$$

$$a_4 = \left(\bar{a}_4 + \left[\frac{m_4 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_4} - \frac{\tilde{A}}{B_4} \right] \right) \bmod m_4 = \left(0 + \left[\frac{7 \cdot 12}{11 \cdot 4} - \frac{3360}{2640} \right] \right) \bmod 7 =$$

$$= (0 + [1,9 - 1,27]) \bmod 7 = (0 + [0,63]) \bmod 7 = (0 + 0) \bmod 7 = 0;$$

$$a_5 = \left(\bar{a}_5 + \left[\frac{m_{n+1} \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_{n+1}} - \frac{\tilde{A}}{B_5} \right] \right) \bmod m_{n+1} = \left(5 + \left[\frac{11 \cdot (1 + 11)}{11 \cdot 6} - \frac{3360}{2520} \right] \right) \bmod 11 =$$

$$= (5 + [2 - 1,3]) \bmod 11 = (5 + [0,7]) \bmod 11 = (5 + 0) \bmod 5 = 0.$$

With accordance to the computed residues $a_1 = 1$, $a_4 = 0$ and $a_5 = 0$ we are correcting (recovering) the corrupted number $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$, i.e. the corrected number becomes $\tilde{A}_{cor.} = (1 \| 0 \| 0 \| 0 \| 5)$.

To validate corrected data, as according to the acquainted [1] expression, we'll define the value of the number $\tilde{A}_{cor.} = (1 \| 0 \| 0 \| 0 \| 5)$ in the following way (see Table 2):

$$\tilde{A}_{cor.PNS} = \left(\sum_{i=1}^{5} a_i \cdot B_i \right) \bmod M_0 = (a_1 \cdot B_1 + a_2 \cdot B_2 + a_3 \cdot B_3 + a_4 \cdot B_4 + a_5 \cdot B_5) \bmod M_0 =$$

$$= (1 \cdot 1540 + 0 \cdot 3465 + 0 \cdot 3696 + 0 \cdot 2640 + 5 \cdot 2520) \bmod 4620 = 14140 (\bmod 4620) = 280.$$

Thus $280 < M = 420$, the number $\tilde{A}_{280} = (1 \| 0 \| 0 \| 0 \| 5)$ is correct.

In order to validate correctness of the number $\tilde{A}_{3360}$ we'll make a computation and comparison of the values to the correct residues $a_2 = 0$ and $a_3 = 0$. In this case they are $a_2 = \left(0 + \left[\frac{4 \cdot (1 + 11)}{11 \cdot 3} - \frac{3360}{3465} \right] \right) \bmod 4 = 0$ and $a_3 = \left(0 + \left[\frac{5 \cdot (1 + 11)}{11 \cdot 4} - \frac{3360}{3696} \right] \right) \bmod 5 = 0$. The result-

ed computations $a_2 = 0$ and $a_3 = 0$ of the residues modulo m_2 and m_3 in RC verified correctness of the corrupted number $\tilde{A}_{3360} = (0 \| 0 \| 0 \| 0 \| 5)$. Thus, the original number $\tilde{A}_{RC} = (0 \| 0 \| 0 \| 0 \| 5)$ is corrupted $\tilde{A}_{3360}$, wherein the single error $\Delta a_1 = 1$ occurred modulo m_1. This error made the correct number A_{280} being corrupted $\tilde{A}_{3360}$.

In order to verify if the correct number A_{280} is true, subsidiary tests on the process of corruption and correction of the number A_{280} modulo $m_1 = 3$ are required. The amount of possible N_{CC} incorrect (corrupted) $\tilde{A}_{RC}$ codewords (if only a single error occurred) for each correct A_{RC} number are $N_{CC} = \sum_{i=1}^{n+1} m_i - (n+1)$.

Test results have shown that corruption of the residue a_1 modulo $m_1 = 3$ of the correct number A_{280} can produce only two incorrect numbers: $\tilde{A}_{3360} = (\tilde{0} \| 0 \| 0 \| 0 \| 5)$ and $\tilde{A}_{1820} = (\tilde{2} \| 0 \| 0 \| 0 \| 5)$. This points to the fact that the corrected number $A_{cor.} = A_{280} = (1 \| 0 \| 0 \| 0 \| 5)$ is both correct (is contained in the interval [0, 420)) and true. The trueness of the resulted number $A_{280} = (\hat{1} \| 0 \| 0 \| 0 \| 5)$ is confirmed by the fact that the single error $\Delta a_1 = 2$ to the base $m_1 = 3$ converts $(\tilde{A} = (A + \Delta A) \bmod M_0 =$
$= (1 \| 0 \| 0 \| 0 \| 5) + (2 \| 0 \| 0 \| 0 \| 0) = \left[(1+2) \bmod 3 \| 0 \| 0 \| 0 \| 5\right] = (\tilde{0} \| 0 \| 0 \| 0 \| 5))$ this number to the unique incorrect number $\tilde{A}_{3360} = (\tilde{0} \| 0 \| 0 \| 0 \| 5)$.

Example №2. Assume, the correct number is $A_{280} = (1 \| 0 \| 0 \| 0 \| 5)$ and assume that $\Delta a_1 = 1$.

In this case $\tilde{A} = (A + \Delta A) \bmod M_0 = (1 \| 0 \| 0 \| 0 \| 5) + (1 \| 0 \| 0 \| 0 \| 0) =$
$= \left[(1+1) \bmod 3 \| 0 \| 0 \| 0 \| 5\right] = (\tilde{2} \| 0 \| 0 \| 0 \| 5)$. This RC number is relevant to the number 1820 in PNS, i.e. the number $\tilde{A}_{1820}$ is incorrect. We'll correct the number $\tilde{A}_{1820}$ now.

Data diagnosing should be made ahead of correcting the number $\tilde{A}_{1820}$. To do this we'll map projections A_j ($j = \overline{1, 5}$) of the number $\tilde{A}_{1820} = (2 \| 0 \| 0 \| 0 \| 5)$ first. Resulted RC code structures are: $\tilde{A}_1 = (0 \| 0 \| 0 \| 5)$, $\tilde{A}_2 = (2 \| 0 \| 0 \| 5)$, $\tilde{A}_3 = (2 \| 0 \| 0 \| 5)$, $\tilde{A}_4 = (2 \| 0 \| 0 \| 5)$ and $\tilde{A}_5 = (2 \| 0 \| 0 \| 0)$.

All the projections of $\tilde{A}_{j\,PNS}$ are:

$$\tilde{A}_{1\,PNS} = (5 \cdot 980) \bmod 1540 = 280 < 420 = M;$$
$$\tilde{A}_{2\,PNS} = (2 \cdot 385 + 5 \cdot 231) \bmod 1155 = 1925 \left(\bmod 1155\right) = 770 > 420 = M;$$
$$\tilde{A}_{3\,PNS} = (2 \cdot 616 + 5 \cdot 672) \bmod 924 = 4592 \left(\bmod 924\right) = 896 > 420 = M;$$
$$\tilde{A}_{4\,PNS} = (2 \cdot 220 + 5 \cdot 540) \bmod 660 = 3140 \left(\bmod 660\right) = 500 > 420 = M;$$
$$\tilde{A}_{5\,PNS} = 2 \cdot 280 (\bmod 420) = 560 \left(\bmod 420\right) = 140 < 420 = M.$$

Inasmuch as $\tilde{A}_{2\,PNS}$, $\tilde{A}_{3\,PNS}$ and $\tilde{A}_{4\,PNS} > 420$, the conclusion is made that the residues $a_2 = 0$, $a_3 = 0$ and $a_4 = 0$ of the number $\tilde{A}_5 = (2 \| 0 \| 0 \| 0 \| 5)$ are not corrupted. Only the residues a_1 and a_5 can be corrupted $\bar{a}_1 = 2$ and $\bar{a}_5 = 5$.

We obtain, that:

$$a_1 = \left(\bar{a}_1 + \left[\frac{m_1 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_1} - \frac{\tilde{A}}{B_1} \right] \right) \bmod m_1 = \left(2 + \left[\frac{3 \cdot (1+11)}{11 \cdot 1} - \frac{1820}{1540} \right] \right) \bmod 3 =$$
$$= \left(2 + [3,27 - 1,18]\right) \bmod 3 = \left(2 + [2,09]\right) \bmod 3 = (2+2) \bmod 3 = 4 (\bmod 3) = 1.$$

Hence, the corrected residue modulo m_1 is $a_1 = 1$. In a like manner the residue $a_5 = 5$.

Applying the results a_1 and a_5 the corrupted number $\tilde{A}_{1820} = (\tilde{2} \| 0 \| 0 \| 0 \| 5)$ is corrected. As a final result the corrected number is $A_{280} = (1 \| 0 \| 0 \| 0 \| 5)$.

Example №3. Performing verification of the number $A_{RC} = (0 \| 0 \| 0 \| 2 \| 1)$. In case corruption was detected, data diagnosing and correction should be made.

I. Data checking of $A_{RC} = (0 \| 0 \| 0 \| 2 \| 1)$. According to the acquainted control procedure A_{PNS} will be calculated using expression:

$$A_{PNS} = \left(\sum_{i=1}^{n+1} a_i \cdot B_i \right) \bmod M_0 = (0 \cdot 1540 + 0 \cdot 3465 + 0 \cdot 3696 + 2 \cdot 2640 + 1 \cdot 2520) \bmod 4620 =$$

$$= 7800 (\bmod 4620) = 3180 > 420.$$ This number is incorrect $\tilde{A}_{3180}$.

II. Data diagnosing of $\tilde{A}_{3180} = (0 \| 0 \| 0 \| 2 \| 1)$. All possible projections $\tilde{A}_j$ of the number $\tilde{A}_{3180}$ are: $\tilde{A}_1 = (0 \| 0 \| 2 \| 1)$, , $\tilde{A}_3 = (0 \| 0 \| 2 \| 1)$, $\tilde{A}_4 = (0 \| 0 \| 0 \| 1)$ and $\tilde{A}_5 = (0 \| 0 \| 0 \| 2)$.

Calculating the values of all of five projections $\tilde{A}_j$ in PNS:

$$\tilde{A}_{1RC} = (0 \| 0 \| 2 \| 1) = \tilde{A}_{1PNS} = (a_1 \cdot B_{11} + a_2 \cdot B_{21} + a_3 \cdot B_{31} + a_4 \cdot B_{41}) \bmod M_1 =$$
$$= (0 \cdot 385 + 0 \cdot 616 + 2 \cdot 1100 + 1 \cdot 980) \bmod 1540 = 100 < M = 420;$$
$$\tilde{A}_{2RC} = (0 \| 0 \| 2 \| 1) = \tilde{A}_{2PNS} = (a_1 \cdot B_{12} + a_2 \cdot B_{22} + a_3 \cdot B_{32} + a_4 \cdot B_{42}) \bmod M_2 =$$
$$= (0 \cdot 385 + 0 \cdot 231 + 2 \cdot 330 + 1 \cdot 210) \bmod 1155 = 870 > M = 420;$$
$$\tilde{A}_{3RC} = (0 \| 0 \| 2 \| 1) = \tilde{A}_{3PNS} = (a_1 \cdot B_{13} + a_2 \cdot B_{23} + a_3 \cdot B_{33} + a_4 \cdot B_{43}) \bmod M_3 =$$
$$= (0 \cdot 616 + 0 \cdot 693 + 2 \cdot 792 + 1 \cdot 672) \bmod 924 = 418 < M = 420;$$
$$\tilde{A}_{4RC} = (0 \| 0 \| 0 \| 1) = \tilde{A}_{4PNS} = (a_1 \cdot B_{14} + a_2 \cdot B_{24} + a_3 \cdot B_{34} + a_4 \cdot B_{44}) \bmod M_4 =$$
$$= (0 \cdot 220 + 0 \cdot 165 + 2 \cdot 396 + 1 \cdot 540) \bmod 660 = 540 > M = 420;$$
$$\tilde{A}_{5RC} = (0 \| 0 \| 0 \| 2) = \tilde{A}_{5PNS} = (a_1 \cdot B_{15} + a_2 \cdot B_{25} + a_3 \cdot B_{35} + a_4 \cdot B_{45}) \bmod M_5 =$$
$$= (0 \cdot 280 + 0 \cdot 105 + 2 \cdot 336 + 1 \cdot 120) \bmod 420 = 240 < M = 420.$$

The calculations of the $\tilde{A}_{j\,PNS}$ values and comparing them to the verification interval $[0, 420)$ range of correct RC numbers A_{RC} resulted in following. The totality of the residues $a_2 = 0$ and $a_4 = 0$ is correct (residues are not being corrupted), while the residues $\bar{a}_1 = 0$, $\bar{a}_3 = 0$ and $\bar{a}_5 = 1$ of the incorrect number $\tilde{A}_{3180} = (0 \| 0 \| 0 \| 2 \| 1)$ might be incorrect (could have been corrupted).

III. Correcting possibly corrupted residues $\bar{a}_1$, $\bar{a}_3$ and $\bar{a}_5$ of the number $\tilde{A}_{3180}$.

Possibly corrupted residues $\bar{a}_1 = 0$, $\bar{a}_3 = 0$ and $\bar{a}_5 = 1$ required to be corrected using expression $a_i = \left(\tilde{a}_i + \left[\frac{m_i \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_i} - \frac{\tilde{A}}{B_i} \right] \right) \bmod m_i$. Then:

$$a_1 = \left(\bar{a}_1 + \left[\frac{m_1 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_1} - \frac{\tilde{A}}{B_1} \right] \right) \bmod m_1 = \left(0 + \left[\frac{3 \cdot (1 + r \cdot 11)}{11 \cdot 1} - \frac{3180}{1540} \right] \right) \bmod 3 =$$
$$= (0 + [3,27 - 2,06]) \bmod 3 = (0 + [1,21]) \bmod 3 = (0 + 1) \bmod 3 = 1.$$

Hence, $a_1 = 1$.

For the value $\bar{a}_3$ it is:

$$a_3 = \left(\tilde{a}_3 + \left[\frac{m_3 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_3} - \frac{\tilde{A}}{B_3} \right] \right) \bmod m_3 = \left(0 + \left[\frac{5 \cdot (1 + r \cdot 11)}{11 \cdot 4} - \frac{3180}{3696} \right] \right) \bmod 5 =$$
$$= (0 + [1,36 - 0,86]) \bmod 5 = (0 + [0,5]) \bmod 5 = (0 + 0) \bmod 5 = 0.$$

In this case $a_3 = 0$.

For the residue $\bar{a}_5$ value is:

$$a_5 = \left(\tilde{a}_5 + \left[\frac{m_5 \cdot (1 + r \cdot m_{n+1})}{m_{n+1} \cdot \bar{m}_5} - \frac{\tilde{A}}{B_5} \right] \right) \bmod m_5 = \left(1 + \left[\frac{11 \cdot (1 + r \cdot 11)}{11 \cdot 6} - \frac{3180}{2520} \right] \right) \bmod 11 =$$
$$= (1 + [2 - 1,26]) \bmod 11 = (1 + [0,74]) \bmod 11 = (1 + 0) \bmod 11 = 1.$$

Obtaining that $a_5 = 1$.

Using the calculated values $a_1 = 1$, $a_3 = 0$ and $a_5 = 1$ of the recovered residues the corrupted number $\tilde{A}_{RC} = (0 \| 0 \| 0 \| 2 \| 1)$ can be corrected, becoming $A_{RC} = (1 \| 0 \| 0 \| 2 \| 1)$. Verified by $100 < 420$.

3 Conclusions of research

Contrary to PNS (positional numeral system), arithmetic RC (residue class) codes feature additional correcting properties. Thus, NCS (non-positional code structure) involves both intrinsic and subsidiary information redundancies, that in some cases results in allowing to correct single errors in RC, while MCD is $d_{\min}^{(RC)} = 2$. However, correcting single errors requires performing subsidiary tests of data checking, i.e. time redundancy usage, additionally to information redundancy. Examples of specific implementation of a single error correcting procedures were introduced, that prove reviewed method is possible to be implemented in order to correct data errors in RC.

References

[1] Akushskii, I. Ya. Mashinnaya arifmetika v ostatochnykh klassakh / I. Ya. Akushskii, D. I. Yuditskii – M.: Sov. Radio, 1968. – 440 p.
[2] Krasnobayev, V. A. A method for increasing the reliability of verification of data represented in a residue number system / V. A. Krasnobayev, S. A. Koshman, M. A. Mavrina // Cybernetics and Systems Analysis. 2014. – Vol. 50, Issue 6, pp. 969-976.
[3] Modeli i metody obrabotki dannykh v sisteme ostatochnykh klassov: [monografiya] / [Krasnobaev V. A., Koshman S., A. i dr.] – Khar'kov: OOO "V dele". 2017. – 197 p.
[4] Stasev, Yu.V., Kuznetsov, A.A., Nosik, A.M. Formation of pseudorandom sequences with improved autocorrelation properties // Cybernetics and Systems Analysis. 2007. Vol. 43, Issue 1, pp. 1-11.
[5] Karpenko O., Kuznetsov A., Sai V., Stasev Yu. Discrete Signals with Multi-Level Correlation Function // Telecommunications and Radio Engineering. 2012. – Vol. 71, Issue 1, pp. 91-98.

Рецензент: Сергій Толюпа, д.т.н., проф., Київський національний університет імені Т. Шевченка, м. Київ, Україна. E-mail: tolupa@i.ua

Надійшло: Березень 2018.

Автори:
Віктор Краснобаєв, д.т.н., проф., заслужений винахідник України, почесний радист СРСР, Харківський національний університет імені В. Н. Каразіна, м. Харків, Україна.
E-mail: v.a.krasnobaev@gmail.com
Сергій Кошман, к.т.н., доцент, Харківський національний університет імені В. Н. Каразіна, м. Харків, Україна.
E-mail: s_koshman@ukr.net
Максим Силенко, студент, Харківський національний університет імені В. Н. Каразіна, м. Харків, Україна.
E-mail: maximkasilenko@gmail.com
Сергій Мороз, к.т.н., Харківський національний технічний університет сільського господарства імені Петра Василенка, м Харків, Україна.
E-mail: frost9i@ukr.net

Метод виправлення однократних помилок даних, що представлені кодом класу лишків.

Анотація. У даній статті розглянуто метод виправлення одноразових помилок у класі лишків (КЛ). Дослідження даного методу дає можливість створювати ефективні системи контролю помилок даних комп'ютерних систем у КЛ. Результати аналізу коригувальних можливостей арифметичного коду у КЛ показали високу ефективність використання непозиційних кодових структур, за рахунок наявності у непозиційній кодовій структурі первинної та вторинної надмірності. У роботі показано, що коригувальні можливості кодів у КЛ залежать від введення додаткової надмірності у кодову структуру. При цьому при виконанні певних умов корекція даних може бути проведена введенням тільки однієї контрольної основи. У статті наведені приклади контролю та виправлення одноразових помилок даних, представлених кодами у КЛ.

Ключові слова: непозиційна кодова структура; клас лишків; позиційна система числення; мінімальна кодова відстань; завадостійке кодування; діагностика та корекція даних.

Рецензент: Сергей Толюпа, д.т.н., проф., Киевский национальный университет имени Т. Шевченко, г. Киев, Украина. E-mail: tolupa@i.ua

Поступила: Март 2018.

Авторы:

Виктор Краснобаев, д.т.н., проф., заслуженный изобретатель Украины, почетный радист СССР, Харьковский национальный университет имени В. Н. Каразина, г. Харьков, Украина.
E-mail: v.a.krasnobaev@gmail.com

Сергей Кошман, к.т.н., доцент, Харьковский национальный университет имени В. Н. Каразина, г. Харьков, Украина.
E-mail: s_koshman@ukr.net

Максим Силенко, студент, Харьковский национальный университет имени В. Н. Каразина, г Харьков, Украина.
E-mail: maximkasilenko@gmail.com

Сергей Мороз, к.т.н., Харьковский национальный технический университет сельского хозяйства имени Петра Василенка, г Харьков, Украина.
E-mail: frost9i@ukr.net

Метод исправления однократных ошибок данных, представленных кодом класса вычетов.

Аннотация. В данной статье рассмотрен метод исправления однократных ошибок в классе вычетов (КВ). Исследование данного метода даёт возможность создавать эффективные системы контроля ошибок данных компьютерных систем в КВ. Результаты анализа корректирующих возможностей арифметического кода в КВ показали высокую эффективность использования непозиционных кодовых структур, за счёт наличия в непозиционной кодовой структуре первичной и вторичной избыточности. В работе показано, что корректирующие возможности кодов в КВ зависят от введения дополнительной избыточности в кодовую структуру. При этом при выполнении определённых условий коррекция данных может быть проведена введением только одного контрольного основания. В статье приведены примеры контроля и исправления однократных ошибок данных, представленных кодами в КВ.

Ключевые слова: непозиционная кодовая структура; класс вычетов; позиционная система счисления; минимальное кодовое расстояние; помехоустойчивое кодирование; диагностика и коррекция данных.

UDC 004.056.55

CODE-BASED SCHEMES FOR DIGITAL SIGNATURES

Alexandr Kuznetsov[1,2], Anastasia Kiian[1], Ivan Belozertsev[1], Mykola Pastukhov[3], Dmytro Prokopovych-Tkachenko[4]

[1] V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, 61022, Ukraine
[2] JSC "Institute of Information Technologies", 12 Bakulin St., Kharkiv, 61166, Ukraine
kuznetsov@karazin.ua, nastyak931@gmail.com, ivanbelozersevv.jw@gmail.com

[3] University of Customs and Finance, 2/4 Volodymyr Vernadsky St., Dnipro, 49000, Ukraine
denart66@gmail.com

[4] University of Customs and Finance, 2/4 Volodymyr Vernadsky St., Dnipro, 49000, Ukraine
me_dnepr@ua.fm

Reviewer: Ivan Gorbenko, Doctor of Sciences (Engineering), Full Professor, Academician of the Academy of Applied Radioelectronics Sciences, V. N. Karazin Kharkiv National University, 4 Svobody Sq., Kharkov, 61022, Ukraine
gorbenkoi@iit.kharkov.ua

Received on May 2018

Abstract. *This article is devoted to the features of construction and use of electronic digital signature schemes based on the use of error-correcting codes, namely the most common scheme, which is based on this approach, CFS and the new proposed scheme. A functioning of these schemes directly depends on used code cryptosystem: the first basically contains principles of Niederreiter code cryptosystem, the second involves use of McEliece cryptosystem, which until recently was considered impossible. Algorithms for generating and verifying signatures according to both schemes, described step by step, are considered in detail. The article studies the efficiency of algorithms in terms of volume of required keys and the length of generated signature, the results of which are presented using analytical ratios and in graphical form for specific examples. The resistance of the considered schemes to classical and quantum cryptanalysis was also analyzed, the latter of which is a actual topic in the era of the rapid development of the sphere of post-quantum cryptography. Both schemes have provable resistance to both types of cryptanalysis, but when using quantum computers it is necessary to significantly increase the key lengths, which is a great shortcoming. It has been revealed that the proposed scheme has an indisputable advantage over the used CFS scheme - protection from specific attacks such as a simultaneous replacement of two signature elements and rapid falsification, by adding an additional element to the generated signature. During the study, the advantages, disadvantages and prospects of using the proposed scheme and the CFS scheme in terms of use of quantum computers are highlighted.*

Keywords: *post-quantum cryptography; digital signature; code-based cryptography; quantum security.*

1 Introduction

In today's increasingly tumultuous world, information is gaining in value. Today, it is perhaps the most expensive resource of mankind. That is why the issue of information security plays an important role and raises serious discussions around it. To date, there is a certain set of proprietary security algorithms that are used during manipulations with information on conventional computers. Despite this, the situation can change radically in the near future, as active work is under way in the quest for the development of a quantum computer [1-6].

A quantum computer is a computing device that works on the basis of the phenomena of quantum confusion and quantum superposition, and allows you to override options and perform complex calculations much faster. For this reason, existing algorithms and ciphers whose security is based on such mathematical problems as factorization of large numbers, discrete logarithms, and others, will lose their security [5].

From this perspective, algorithms designed for the formation and verification of digital signatures also become vulnerable to various types of attacks [3]. This fact will lead to the fact that an digital signature will not guarantee the integrity of the document and reliably confirm the identity of its author. From the above it follows the relevance of the comprehensive study of alternative schemes of digital signature and assessment of their capabilities. One of the most promising areas of research, from the standpoint of post-quantum efficiency, is cryptography, based on error-correcting codes [7-19]. In this work, we will consider two code schemes of a digital signature, we will carry

out their comparative analysis and estimation of possibilities of their application in the post-quantum period.

2 A classic example of a code-based digital signature scheme

A classic example of a digital signature scheme based on error-correcting codes is the CFS scheme, named after the initials of its inventors – Courtois, Finiasz, Sendrier [13].

CFS involves the use of algebraic (n, k, d) code from the class ($n=2^m$, $k=n-mt$, $2t+1$) of non-idle Goppa codes. The formation of the public and private keys of the scheme is in accordance with the principle used in the Niderraiter cryptosystem, which is discussed in detail in the papers [3, 7-11]. Hence, the private keys are matrices X, of size $(n-k)\times n$, and P, of size $n\times n$, which are similar to the Niderraiter scheme defined as a random inverse matrix and a random matrix of permutations respectively, as well as a fast algorithm for decoding an algebraic code. A private key is a matrix $H_X = X \cdot H \cdot P$, where H – the verification $(n-k)\times n$ matrix of the algebraic code and the correcting ability of the code t. The input data for using the CFS is a hash function h, a fast algorithm for decoding an algebraic code and a message (plain text). The hash function is intended to convert a message of arbitrary length. The output of the hash function is a hash value $h(x)$ of bit length $n-k$. A quick algorithm for decoding the algebraic code, that is, having a polynomial complexity, is applied to the syndromic sequence. $s=(s_0, s_1, ..., s_{n-k-1})$. In this case, one of the situations is possible:

- If decoding is successful the vector of errors $e=(e_0, e_1, ..., e_{n-1})$ corresponding to the syndrome will be displayed.
- If decoding is unsuccessful an error message will be displayed.

The signature formation algorithm consists of gradual execution of several steps. Initially, the hashing of the plain text M and the assignment to the counter i of value i=1. The hash value $h(M)$ and counter i are represented as bit sequences, from concatenation of which, the new hash value $h(h(M)\|i)$ is calculated. The latter should be interpreted as a syndromic sequence $s_X=(s_0, s_1, ..., s_{n-k-1})$ calculated for some arbitrary code word and error vector $e=(e_0, e_1, ..., e_{n-1})$. Since it was suggested that $h(h(M)\|i)$ is a syndrome, which is calculated according to the check matrix H of algebraic code, we need to build a vector:

$$s_X^{*T} = X^{-1} \cdot s_X{}^T = X^{-1} \cdot H_X \cdot e^T = H \cdot P \cdot e^T = H \cdot \overline{e}^T.$$

Then we can apply a fast decoding algorithm to find the vector $\overline{e}^T = P \cdot e^T$. If decoding fails, then you need to increment the value of the counter i and perform all actions, starting with the concatenation of the counter and the hash value of the message, until the derived vector $\overline{e}^T = P \cdot e^T$ is deduced, that corresponds to vector s_X^*. When such a vector is found, you need to go to the next step and calculate the value:

$$e^T = P^{-1} \cdot \overline{e}^T = P^{-1} \cdot P \cdot e^T.$$

The final signature for a message consists of two parts: the value of the counter and the vector e, $Y=(e,i)$. Formally, you we write the generated signature as

$$Y=(e,i): H_X \cdot e^T = \left(h(h(M)\|i)\right)^T.$$

In order to verify the authenticity of the signature, it must be ensured that the result of the hashing $h(h(M)\|i)$ is a syndromic sequence that was calculated according to the vector $e=(e_0, e_1, ..., e_{n-1})$, the latter is interpreted as a vector of errors.

A user who wants to verify the authenticity of the signature has an input of an a public key con-

sisting of t a matrix H_X, hash function *h*, the signature itself $Y=(e,i)$ and a message M. In order to verify a signature, you need to calculate the values of two vectors: $(s_X^{'})^T = H_X \cdot e^T$, $(s_X^{''})^T = h(h(M)\|i)$.

A digital signature is considered correct only if these two vectors are the same [3,13].

Therefore, the essence of the CFS scheme is the repeated hashing of a message that is encapsulated with a randomized counter value in order to identify the correct syndromic sequence. The stability of this scheme is based on the complexity of solving the problem of syndromic decoding.

3 An alternative scheme for formation and verification of digital signature

The CFS scheme is the most commonly code-based digital signature scheme. However, this scheme has certain disadvantages. In 2017, an alternative to this scheme was proposed. An alternative to the CFS, unlike the original scheme, is based on the use of the one-sided function of McEliece, which is considered in the works [16-19]. As a result, the private keys of this scheme are the matrixes X and P (in the case of non-binary codes, the matrix D is added), which are an invertible matrix and a permutation matrix, respectively, as well as a fast algorithm for decoding the algebraic code. The public key is a number *t*, which characterizes the corrective ability of algebraic (*n*, *k*, *d*) code from the class of irreducible Goppa codes. For a binary case, the code parameters are related to this relationship: $n = 2^m$, $k = n-mt$, $2t+1$. Also, the component of the public key is the matrix G_x, which is formed according to the rule $G_X = X \cdot H \cdot P \cdot D$, where G is a generating matrix of algebraic code.

When forming a signature, the same hash function *h* as in the CFS scheme is used, which was described in detail earlier, so we will not focus on it. The decoding algorithm is the ability to find the vector of errors $e=(e_0,e_1,...,e_{n-1})$ and the vector $I=(I_0,I_1,...,I_{k-1})$ according to the original code word with errors $c_X^* = (c*_0, c*_1, ..., c*_{n-1})$, taking into account the equation $c_X^* = I \cdot G_X + e$.

When signing a message, the user first has to find the hash code from its content and determine the value of the counter i equal to 1. Then, as in the CFS scheme, the concatenation of the hash value of the message and the counter occurs followed by the hashing of the generated sequence, which results in $h(h(M)\|i)$. $h(h(M)\|i)$ is interpreted as a codeword with errors $c_X^* = (c*_0, c*_1, ..., c*_{n-1})$, which is calculated for some values of the vectors $I=(I_0,I_1,...,I_{k-1})$ and $e=(e_0,e_1,...,e_{n-1})$, provided that $c = IG_X$ and $c_X^* = c+e$ are equal. The next step is to calculate the value of the vector $\bar{c}^* = c_X^* \cdot D^{-1} \cdot P^{-1}$. It is assumed that the value of this vector represents the distorted codeword of the algebraic code no more than in t digits, that is, the distortion does not exceed the correction ability. A similar code word can be decoded using a polynomial complexity algorithm. Therefore, it is assumed that:

$$\begin{aligned}\bar{c}^* &= c_X^* \cdot D^{-1} \cdot P^{-1} = (I \cdot G_X + e) \cdot D^{-1} \cdot P^{-1} = \\ &= (I \cdot X \cdot H \cdot P \cdot D + e) \cdot D^{-1} \cdot P^{-1} = \\ &= I \cdot X \cdot H + e \cdot D^{-1} \cdot P^{-1}.\end{aligned}$$

By applying a polynomial complexity algorithm and decoding the code word. $\bar{c}^* = I' \cdot G + e'$, $e' = e \cdot D^{-1} \cdot P^{-1}$, we can find vector $I' = I \cdot X$. If decoding was successful, then the corresponding values I' and e' will be displayed. If the decoding failed, you need to increment the value of the counter and repeat all steps in the signature formation, starting with the concatenation initially, until the values I' and e' are successfully decoded. After finding such values, the vectors $I = I' X^{-1}$ and $e = e' \cdot D \cdot P$ are calculated.

The signature of a message in this case can be formally defined as

$Y=(I,e,i): IG_X+e=\left(h(h(M)\|i)\right)$ that is, it consists of a counter value i, for which $h(h(M)\|i)$ will be interpreted as a code word with errors a vector of errors e and information vector I. The complexity of calculating vectors I and e using known hash value $h(h(M)\|i)$ for an illegitimate user is a NP-complete task.

To verify the authenticity of the signature, you need to make sure that the result of the hashing $h(h(M)\|i)$ is a codeword with errors, which is calculated using the values of the vectors I and e. For the purpose of verification, it is necessary to calculate the values of two vectors $c_X^*{}' = IG_X + e$ and $c_X^*{}'' = h(h(M)\|i)$.

If the values of these vectors coincide $c_X^*{}' = c_X^*{}''$ and the Hamming weight of the vector e does not exceed the correcting ability of the code $w(e) \le t$, then the signature can be considered true. If at least one of the declared requirements is not fulfilled, the signature is rejected. Therefore, the essence of the CFS schema-alternative is to interpret the hex value of the sequence that came out by combining the value of the counter and the hex value of the message as a codeword with errors. At the same time, the verification procedure is radically different from the original scheme by adding another condition that provides an alternative scheme with advantages over CFS, which will be discussed in the next section.

4 The comparative analysis of schemes for the formation and verification of digital signatures

The lengths comparison of key scheme parameters

The CFS scheme and its alternative are based on two different approaches: the first is to use the function of the Niderraiter scheme, the second one of the McEliece scheme, on which the volumes of key signature data schemes depend directly.

Public Key:

– The CFS public key length is determined by the number of cells in the matrix $H_X = X \cdot H \cdot P$.

$$l_{PB.K.} = (n-k)\cdot n = n^2 - kn = m\cdot t\cdot 2^m.$$

– The length of the public key of the alternative scheme is determined by the number of cells in the matrix $G_X = X \cdot G \cdot P$

$$l_{PB.K.} = k\cdot n = (2^m - m\cdot t)\cdot 2^m.$$

Private Key:

– The length of the private key CFS is determined by the sum of the number of binary cells of the matrix (the size of $(n-k)\times(n-k)$) and the length of n integers in range $0,1,..,n-1$ for determining the matrix P and is calculated:

$$l_{PR.K.} = (n-k)^2 + n\cdot\lceil \log_2 n\rceil = (m\cdot t)^2 + 2^m\cdot m.$$

– The length of the private key of the alternative scheme is determined by the sum of the number of binary cells of the matrix X (the size of $(k\times k)$) and the length of n integers in range $0,1,..,n-1$ for determining the matrix P. This length can be calculated according to [17]:

$$l_{PR.K.} = k^2 + n\cdot\lceil \log_2 n\rceil = (2^m - m\cdot t)^2 + 2^m\cdot m.$$

In order to demonstrate the differences between the alternative scheme and the CFS scheme more clearly, we present a graphic representation (Fig. 1-2). After analyzing of the data, we can conclude that the graph of the private and public keys of the alternative scheme is declining, and the graph of the CFS scheme is increasing. Up to a certain point, the length of the private and public keys of the alternative scheme will exceed the values for the CFS scheme.

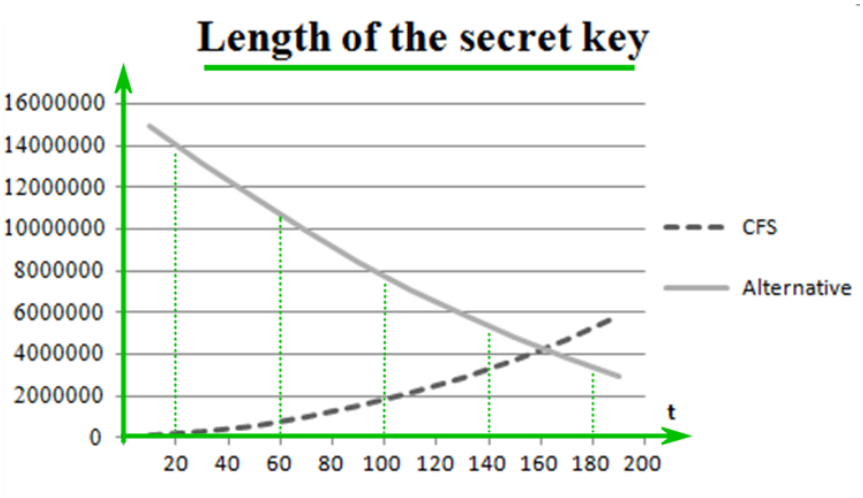

Fig. 1 – Comparison of the length of private keys

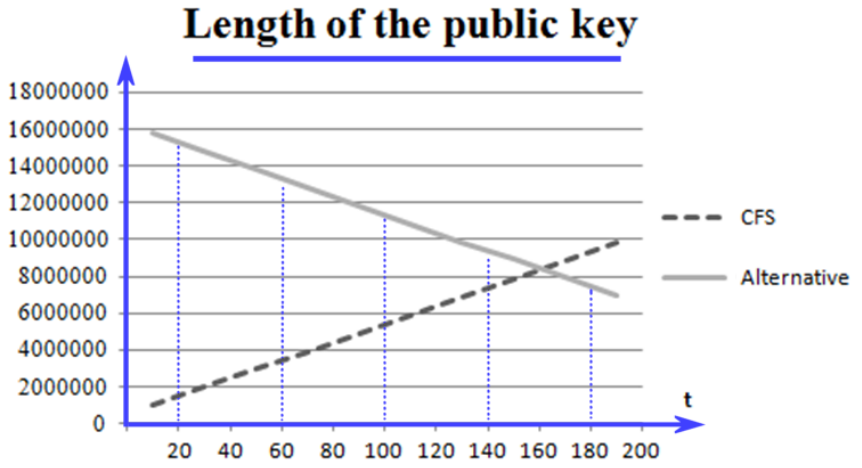

Fig. 2 – Comparison of the length of public keys

The comparison of lengths of signatures, which are formed according to both schemes

According to the CFS scheme, the signature $Y=(e,i)$ contains two components: the binary vector e, which has the length of n, and an integer i. The latter may acquire the values in the range $0,1,..,2^{n-k}-1$. From here, we have that the bit length of the signature is determined according to the expression: $l_{DS}=2\cdot n-k=2^m+m\cdot t$.

Vector e can acquire a limited number of values. The limitation is imposed according to the correcting ability of the code used. The number of possible vectors e is defined as:

$$N_{w(e)\le t}=\sum_{i=0}^{t}C_n^i .$$

Because the vector e corresponds to the condition above, it can be transformed into a break-even sequence $e*$ with bit length of $\lceil \log_2(N_{w(e)\le t})\rceil$. Then we have:

$$l_{DS}*=\left\lceil \log_2\left(\sum_{i=0}^{t}C_n^i\right)\right\rceil+n-k=\left\lceil \log_2\left(\sum_{i=0}^{t}C_{2^m}^i\right)\right\rceil+m\cdot t .$$

Using the expression for the upper bound of Hamming, the expression can be transformed:

$$l_{DS}*\le\left\lceil \log_2\left(2^{n-k}\right)\right\rceil+n-k=2\cdot m\cdot t .$$

In the case of an alternative scheme, the components of the signature become larger $Y=(I,e,i)$: the vector I (the bit length of k), the vector e and an integer i, whose length is determined in the same way as in the CFS scheme. From here, we have the length of the signature $Y=(I,e,i)$ is determined by the expression: $l_{DS}=2\cdot n=2^{m+1}$. If we make a break-even transform of vector e, then this estimate can be rewritten as:

$$l_{DS}*=\left\lceil \log_2\left(\sum_{i=0}^{t}C_n^i\right)\right\rceil+n=\left\lceil \log_2\left(\sum_{i=0}^{t}C_{2^m}^i\right)\right\rceil+2^m.$$

Similarly, to the consideration of CFS, using Hamming's upper boundary we have [20]:

$$l_{DS}*\le\left\lceil \log_2\left(2^{n-k}\right)\right\rceil+n=m\cdot t+2^m.$$

Let us demonstrate the resulting estimates through a graphical representation of an example code with a parameter $n = 12$ (Fig. 3).

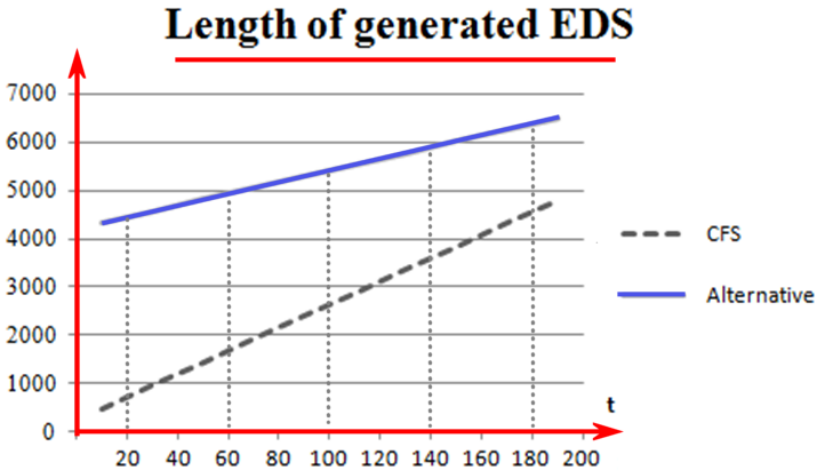

Fig. 3 – Comparison of lengths of signatures

By analyzing the obtained data, we can conclude that the length of the signature formed by the alternative scheme significantly exceeds the length of the signature CFS scheme. In particular, an increase in the length of the signature takes place by adding the vector I.

Cryptographic robustness of the signature schemes

As noted earlier, the CFS scheme is based on the use of a one-way function from the Niderraiter cryptosystem. The robustness of this function can be defined as the number of roofing sets in which it is possible to fix all combinations of t errors without knowledge of the private key [20]:

$$N\ge\frac{C_n^t}{C_{n-k}^t}=\frac{\frac{n!}{t!(n-t)!}}{\frac{(n-k)!}{t!(n-k-t)!}}=\frac{n!(n-k-t)!}{(n-t)!(n-k)!}.$$

In order to form a signature $Y'=(e',i')$ for an altered message M' the attacker needs to implement the decoding of a random code on average *t!* times. Taking into account this fact, the evaluation of digital signature robustness under the CFS scheme can be defined as:

$$N_c\ge t!\frac{C_n^t}{C_{n-k}^t}=t!\frac{n!(n-k-t)!}{(n-t)!(n-k)!}=t!\frac{2^m!(mt-t)!}{(2^m-t)!(mt)!}.$$

In a number of studies that have been carried out, the equivalence of McEliece and Niderraiter cryptosystems has been demonstrated. Hence, we can assume that the robustness of the CFS

schemes and their alternatives is also equivalent.

In the case of the use of quantum cryptanalysis, the estimation of the robustness of both schemes acquires a different character. Using one of the most popular quantum algorithms, Grover's algorithm, it is possible to determine the number of iterations to decode random code, which must be performed $t!$ times: $C^{\frac{2}{2\log n}}$, $C = 1/(1-R)^{1-R}$.

Assume that a quantum algorithm can be used to find the value of a counter i, by checking values (*brute force attack*), which requires an average $\frac{\pi}{4}\sqrt{t!}$ attempts. Therefore, the robustness of digital signature schemes in terms of quantum computers can be defined:

$$N_{PR.K.} \geq \frac{\pi}{4}\sqrt{t!}\left(\frac{1}{(1-R)^{1-R}}\right)^{\frac{n}{2\log n}} = \frac{\pi}{4}\sqrt{t!}\left(\left(1-\frac{k}{n}\right)^{\frac{k}{n}-1}\right)^{\frac{n}{2\log n}} =$$

$$= \frac{\pi}{4}\sqrt{t!}\left(1-\frac{k}{n}\right)^{\frac{k-n}{2\log n}} = \frac{\pi}{4}\sqrt{t!}\left(\frac{m\cdot t}{2^m}\right)^{t/2}.$$

When analyzing an alternative scheme, it is worth noting that it has a significant advantage over CFS, since it is able to provide security against fast signature falsification on the basis of the addition of an arbitrary codeword. An attack of this type with respect to CFS can be organized through the following actions:

- Select an arbitrary codeword from the code (n,k,d), that has a check matrix H_x. In this case the equation $H_x \cdot c^T = 0$ is true. We get a formed signature $Y = (e,i)$.
- Perform a codeword addition:

$$Y = (e+c,i) : H_x \cdot (e+c)^T =$$
$$= H_x \cdot e^T + H_x \cdot c^T = H_x \cdot e^T = (h(h(M)\,||\,\mathrm{i}))^T.$$

Changing the last expression with respect to the alternative scheme, we obtain: $Y = (I, e+c, i) : IG_X + e + c \neq \left(h(h(M)\|i)\right)$. That is, a quick falsification of the signature in this case is impossible. This property is also enhanced by additional testing of Hamming's error vector during the signature verification process. It also protects against other hypothetical attacks, such as simultaneous falsification of two signature elements, etc.

5 Conclusions

In the modern world, digital signature plays an important role and serves as the confirmation of the author's personality, and the integrity of the document. Considering two code-based digital signature schemes, namely CFS and its alternative, one can conclude that both schemes are comparable in length of key parameters, the latter depending on the parameters of the chosen code. The length of the formed signature according to the alternative scheme is slightly larger, but this increase is not critical. It is also worth noting that the robustness of the schemes against classical and quantum cryptanalysis is equivalent, which follows from the estimates equivalence of the robustness of McEliece and Niderraiter schemes on which the work of the considered digital signature algorithms is based. However, the alternative scheme has a significant advantage over the CFS common scheme: it is able to provide protection against specific attacks of fast signature falsification and simultaneous falsification of two signature components.

As the disadvantages of both cryptosystems, it is worth noting the large volumes of key data, which, according to researchers, will need to increase more than three times in the post-quantum period. The ability to reduce the key length while maintaining the robustness of signatures remains a promising area of research.

References

[1] A.J. Menezes, P.C. van Oorschot, S.A. Vanstone. Handbook of Applied Cryptography. CRC Press, 1997, 794 p.
[2] N. Ferguson and B. Schneier. Practical Cryptography. John Wiley & Sons, 2003, 432 p.
[3] D. Bernstein, J. Buchmann and E.Dahmen. Post-Quantum Cryptography. Springer-Verlag, Berlin-Heidleberg, 2009, 245 p.
[4] J. Proos and Ch. Zalka. "Shor's discrete logarithm quantum algorithm for elliptic curves". [On-line]: https://arxiv.org/abs/quant-ph/0301141.
[5] D. Deutsch and R. Jozsa. "Rapid solutions of problems by quantum computation". Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences, Vol. 439, No. 1907. – 1992. pp. 553-558.
[6] P.W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". Foundations of Computer Science: Conference Publications, 1997, pp. 1484-1509.
[7] Niederreiter, H. "Knapsack-type cryptosystems and algebraic coding theory". Problem Control and Inform Theory, 1986, V. 15. pp. 19-34.
[8] Yu.V. Stasev, A.A. Kuznetsov. "Asymmetric code-theoretical schemes constructed with the use of algebraic geometric codes". Kibernetika i Sistemnyi Analiz, No. 3, pp. 47-57, May-June 2005.
[9] A. Kuznetsov, I. Svatovskij, N. Kiyan and A. Pushkar'ov. "Code-based public-key cryptosystems for the post-quantum period, "2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 125-130.
[10] A. Kuznetsov, R. Serhiienko and D. Prokopovych-Tkachenko. "Construction of cascade codes in the frequency domain, "2017 4-th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, 2017, pp. 131-136.
[11] Yu.V. Stasev, A.A. Kuznetsov. "Asymmetric Code-Theoretical Schemes Constructed with the Use of Algebraic Geometric Codes." Cybernetics and Systems Analysis, Vol. 41, Issue 3, pp. 354-363, May 2005.
[12] M. Finiasz and N. Sendrier. "Security bounds for the design of codebased cryptosystems". In M. Matsui, ed., Advances in Cryptology, ASIACRYPT 2009, Vol. 5912 of Lecture Notes in Computer Science. -Springer Berlin Heidelberg, 2009, pp. 88-105.
[13] N.Courtois, M.Finiasz and N.Sendrier. "How to achieve a McEliece-based digital signature scheme". In Advances in Cryptology - ASIACRYPT 2001, Vol. 2248, pp. 157–174.
[14] M. Finiasz "Parallel-CFS: Strengthening the CFS McEliece-based signature scheme". In Biryukov A., Gong G., Stinson D., eds.: Selected Areas in Cryptography. Vol. 6544 of LNCS., Springer (2010) pp.159-170.
[15] J. Stern "A new identification scheme based on syndrome decoding". In Advances in Cryptology - CRYPTO'93, Vol. 773 of LNCS. Springer Verlag (1994).
[16] R. J. McEliece "A public-key cryptosystem based on algebraic coding theory". DSN Progress Report 42-44, Jet Propulsion Lab., Pasadena, CA, January-February, 1978, pp. 114-116.
[17] V. M. Sidel'nikov "Kriptografiya i teoriya kodirovaniya". Materialy konferentsii «Moskovskii universitet i razvitie kriptografii v Rossii», MGU. – 2002. – 22 p. (in Russian).
[18] S. Sander "Study of McElice cryptosystem". [On-line]: https://courses.cs.ut.ee/MTAT.07.022/2015_spring/uploads/Main/sander-report-s15.pdf.
[19] Li, Y. X., Deng, R.H., Wang, X.M. "On the equivalence of McEliece's and Niederreiter's public-key cryptosystems". [On-line]: https://ieeexplore.ieee.org/document/272496/.
[20] Clark, G.C., Cain, J.B. Error-Correction Coding for Digital Communications. Springer, 1981, 432 p.

Рецензент: Іван Горбенко, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, 61022, Україна.
Email: gorbenkoi@iit.kharkov.ua.

Надійшло: Травень 2018.

Автори:

Олександр Кузнецов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет ім. В. Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна; АТ «Інститут інформаційних технологій», вул. Бакуліна, 12, м. Харків, Україна. E-mail: kuznetsov@karazin.ua

Анастасія Киян, студентка, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна. E-mail: nastyak931@gmail.com

Іван Білозерцев, студент, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна. E-mail: ivanbelozersevv.jw@gmail.com

Микола Пастухов, к.т.н., доцент, Університет митної справи та фінансів, вул. Володимира Вернадського 2/4, м. Дніпро, 49000, Україна. E-mail: denart66@gmail.com

Дмітро Прокоповіч-Ткаченко, к.т.н., завідувач кафедрою кібербезпеки, Університет митної справи та фінансів, вул. Володимира Вернадського 2/4, м. Дніпро, 49000, Україна. E-mail: me_dnepr@ua.fm

Схеми на основі кодів для цифрових підписів.

Анотація. Стаття присвячена розгляду особливостей побудови і використання схем електронного цифрового підпису, заснованих на використанні кодів, виправляючих помилки, а саме найпоширенішої схеми, яка базується на цьому підході, CFS і нової запропонованої схеми. Функціонування цих схем безпосередньо залежить від використовуваної кодової криптосистеми: перша в своїй основі містить принципи кодової криптосистеме Нідеррайтера, друга передбачає використання криптосистеми Мак-Еліса, що до недавнього моменту вважалося неможливим. Детально розглянуті алгоритми формування та перевірки підпису згідно обох схем, описані покроково. У роботі проведені дослідження ефективності алгоритмів з точки зору обсягу необхідних ключових даних і довжини сформованого підпису, результати якого представлені за допомогою

аналітичних співвідношень і на конкретних прикладах в графічному вигляді. Також було проаналізовано стійкість розглянутих схем до класичного і квантового криптоаналізу, останній з яких є актуальною тематикою в еру стрімкого розвитку сфери пост-квантової криптографії. Розглянуты схеми мають доказову стійкість до обох видів криптоаналізу, однак при використанні квантових комп'ютерів необхідно значно збільшувати довжини ключів, що є вагомим недостакі. Виявлено факт, що запропонована схема має незаперечну перевагу перед використовуваної схемою CFS - захист від специфічних атак таких, як одночасна заміна двох елементів підпису і швидка фальсифікація, за рахунок додавання додаткового елемента в сформований підпис. Протягом дослідження виділені переваги, недоліки і перспективи використання запропонованої схеми і схеми CFS в умовах застосування квантових комп'ютерів.

Ключові слова: постквантова криптографія; цифровий підпис; криптографія на основі кодів; квантова безпека.

Рецензент: Иван Горбенко, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы, 4, г. Харьков, 61022, Украина.
Email: gorbenkoi@iit.kharkov.ua.

Поступила: Май 2018.

Авторы:

Александр Кузнецов, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина; АО «Институт информационных технологий», ул. Бакулина, 12, Харьков, Украина. E-mail: kuznetsov@karazin.ua

Анастасия Киян, студентка, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина. E-mail: nastyak931@gmail.com

Иван Белозерцев, студент, Харьковский национальный университет имени В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина. E-mail: ivanbelozersevv.jw@gmail.com

Николай Пастухов, к.т.н., доцент, Университет таможенного дела и финансов, ул. Владимира Вернадского, 2/4, г. Днепр, 49000, Украина. E-mail: denart66@gmail.com

Дмитрий Прокопови-Ткаченко, к.т.н., заведующий кафедрой кибербезопасности, Университет таможенного дела и финансов, ул. Владимира Вернадского, 2/4, г. Днепр, 49000, Украина. E-mail: me_dnepr@ua.fm

Схемы на основе кодов для цифровых подписей.

Аннотация. Статья посвящена рассмотрению особенностей построения и использования схем электронной цифровой подписи, основанных на использовании кодов, исправлющих ошибки, а именно наиболее распространенной схемы, которая базируется на этом подходе, CFS и новой предложенной схемы. Функционирование данных схем напрямую зависит от используемой кодовой криптосистемы: первая в своей основе содержит принципы кодовой криптосистеме Нидеррайтера, вторая предусматривает использование криптосистемы Мак-Элиса, что до недавнего момента считалось невозможным. Подробно рассмотрены алгоритмы формирования и проверки подписи согласно обоим схемам, описанные пошагово. В работе произведены исследования эффективности алгоритмов с точки зрения объема требуемых ключевых данных и длины формируемой подписи, результаты которого представлены с помощью аналитических соотношений и на конкретных примерах в графическом виде. Также была проанализирована стойкость рассмотренных схем к классическому и квантовому криптоанализу, последний из которых является актуальной тематикой в эру стремительного развития сферы пост-квантовой криптографии. Рассмотренные схемы имеют доказуемую стойкость к обоим видам криптоанализу, однако при использовании квантовых компьютеров необходимо значительно увеличивать длины ключей, что является весомым недостаком. Выявлено факт, что предложенная схема имеет неоспоримое преимущество перед используемой схемой CFS - защита от специфических атак таких, как одновременная подмена двух элементов подписи и быстрая фальсификация, за счет добавления дополнительного элемента в сформированую подпись. В течении исследования выделены достоинства, недостатки и перспективы использования предложенной схемы и схемы CFS в условиях применения квантовых компьютеров.

Ключевые слова: постквантовая криптография; цифровая подпись; криптография на основе кодов; квантовая безопасность.

UDC 004.056.55

PROBABILISTIC MINUTIA DISTRIBUTION IN BIOMETRIC FINGERPRINT IMAGES

Sergey Rassomakhin[1], Alexandr Kuznetsov[1], Vladimir Shlokin[1], Ivan Belozertsev[1], Roman Serhiienko[2]

[1] V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, 61022, Ukraine
rassomakhin@karazin.ua, kuznetsov@karazin.ua, vshlokin@ukr.net, ivanbelozersevv.jw@gmail.com

[2] National Army Academy named after Hetman Petro Sahaidachnyi, 32 Heroes of Maidan St., Lviv, 79012, Ukraine
romanserg69@gmail.com

Reviewer: Vyacheslav Kharchenko, Doctor of Technical Sciences, Professor, Academician of the Academy of Sciences of Applied Radio Electronics, National Aerospace University named after. M. E. Zhukovsky, Kharkiv, Ukraine.
E-mail: v_s_kharchenko@ukr.net

Received on June 2018

Abstract. *The analysis of the fingerprint scanning results shows an extremely small degree of similarity among the obtained images. The cause of the problem mentioned above is not only the complexity of the procedure itself, but also the imperfection of the used recognition algorithms. The research involves development of a mathematical model for the probabilistic minutia distribution in biometric fingerprint images. The suggested model is based on heuristic analysis of the fingerprint scanning results with account for the nature of the potential errors. She allows to model a typical minutiae behavior in the biometric fingerprint images.*

Keywords: *biometric authentication; fingerprint images; minutia.*

1 Introduction

Dactyloscopy occupies a special place among the known methods of biometric authentication [1-8]. Biometric fingerprint image minutiae processing which underlines the above method, enables robust and efficient identification of individuals.

At the same time, minutiae distribution of certain implementations may be described with rather complex dependencies. This is explained by the significant differences in the number of minutiae and their placement [9-12]. Error types and their distribution functions are also rather ambiguous due to the multiple nature of possible causes.

The choice between simplicity and adequacy of the models, describing minutiae placement and errors, is a compromise option. However, the closed nature of existing fingerprint recognition algorithms makes it impossible to collect an amount of statistics enough for solving the problem in a straightforward way. Therefore, the research and development of mathematical models for the probabilistic minutiae distribution in biometric fingerprint images is an important and relevant scientific problem. The models represented in this paper were obtained through heuristic analysis of the fingerprint scanning results with account for the nature of the potential errors.

2 The analysis of biometric fingerprint images

Let us use database DB1_1 [13] for the analysis of characteristic and error distributions which may occur during fingerprint image processing. This database contains 8 images (files 101_1.tif – 101_8.tif) of the same fingerprint. The goal of the analysis is to make a preliminary conclusion about the nature of the errors typical for minutia recognition.

The original fingerprint images are shown on fig. 1. Fig. 2 shows the processing results of the given samples using SourceAFIS.FingerprintAnalysis [14-15].

The results represent detected minutiae which correspond to the endings and bifurcations of the ridges. The arrows represent the angles of given minutiae.

Variation of fingerprint orientations and their displacement, as well as the changes in contrast and brightness, cause the significant differences of processing results which translates to the varia-

tion of the number of minutiae and their positioning. The following figure shows the circles that correspond to the same area of the fingerprint, but displaced and rotated during the scanning process.

Fig. 1 – Biometric images of a single finger

Fig. 2 – Fingerprint minutiae extraction results

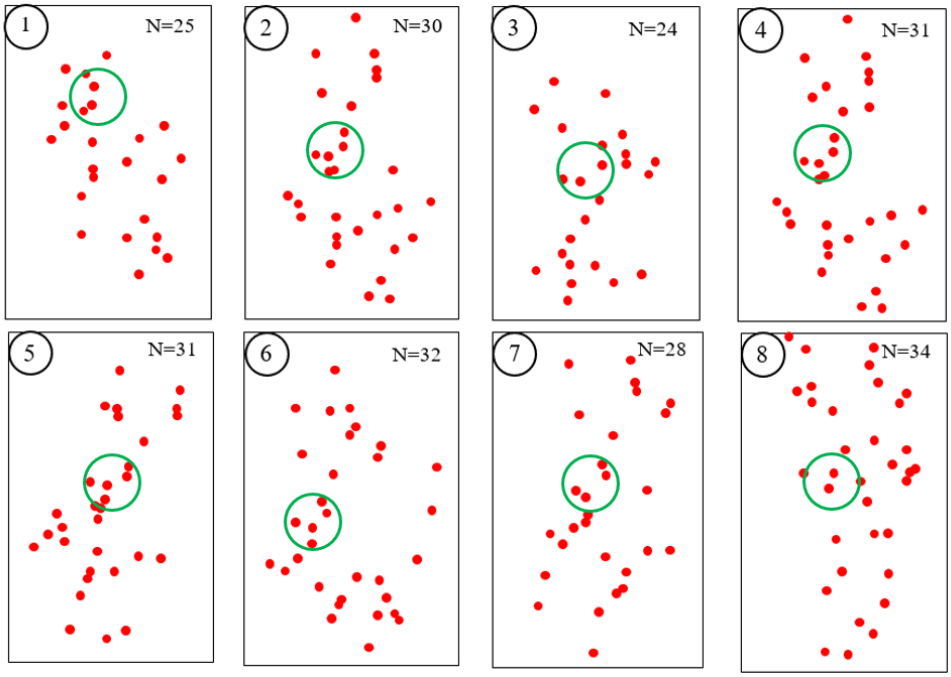

Fig. 3 – Plain portraits of the extracted minutiae distribution

Figure 3 represents plane portraits of minutiae placement for the given fingerprint images. As we can see, the degree of similarity of the given portraits is rather low. Visual similarity takes place only in case of similar scanning conditions (*on Fig. 3 it is the pairs 2 and 4 or 3 and 7*). Apparently, the cause of the problem mentioned above is not only the complexity of the procedure itself, but also the imperfection of the used recognition algorithm implemented in SourceAFIS. Fingerprint Analysis [14-15].

3 Mathematical model for the probabilistic minutiae distribution

The analysis of the portraits mentioned above indicates the following *features* that we can base our empiric choice of the type of minutiae distribution on:

- the density of distribution of points along the horizontal (X) and vertical (Y) axes is roughly uniform in the central part of the frame and slightly decreases to its edges;
- linear displacements of the center of the fingerprint horizontally and vertically do not imply the appearance of zones free of minutiae at the edges of the frame (new points may enter the scanning area);
- the distribution of minutiae angles is approximately uniform in the range of $[0,\ 2\pi]$.

Let us use the following assumptions to construct a model of minutiae distribution according to the features mentioned above:

- the portrait coordinates of the fingerprint X, Y, as well as the minutiae angles values are normalized in the range $[-0.5; +0.5]$, while the geometric center of the image has zero coordinates on the plane $[0; 0]$, and the portrait itself is placed in a unit square area covering all 4 quadrants of the image plane;
- for the primary generation of random numbers necessary to obtain the distribution of minutiae coordinates on the fingerprint image portraits, a uniformly distributed (continuous) random number generator in range $[0;1]$: $f(x_i, y_i)$:*unif* $[0,1]$, $i \in 1 \ldots N$, where N – the number of minutiae in the portrait, a random value that does not go out of range $[15; 60]$ with a mathematical expectation $m_N = 25 \div 35$ and unimodal distribution.

The analysis of features which were discussed before, as well as taking into account the assumptions made above, allow us to use the dependency shown on Fig. 4 to describe the probability density function (*PDF*) $f(x)$ and $f(y)$ Cartesian coordinates of the minutiae on the plain portraits This type of PDF provides a uniform points distribution in the central part of the unit square and the decreasing probability of point appearance at the edges of the square area of the portrait. The area of non-zero PDF values $[-0.75; +0.75]$ is 0.25 in both directions beyond the unit square, which provides a non-zero probability of the point appearance in the border areas of the portrait. The errors appear in the form of possible geometric center drifting. The choice of this PDF is, of course, not the only one possible, however, in our opinion, is acceptable, considering the tradeoff between the simplicity and features mentioned above.

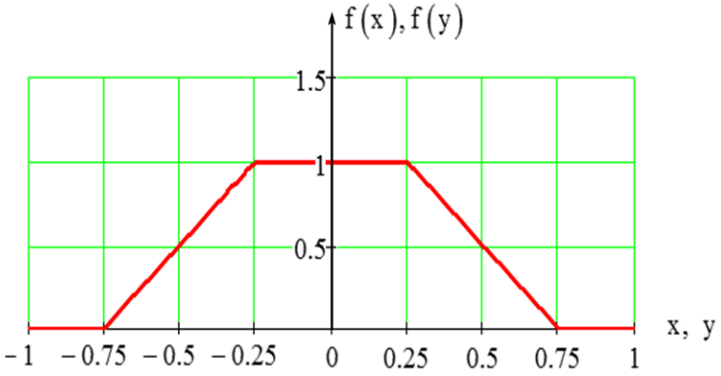

Fig. 4 – Probability density function of the minutiae coordinates

To obtain test samples of minutia placement portraits the generator of random numbers, distributed according to PDF $f(x)$, $f(y)$ (Fig. 4), is required. Considering the identity of the distributions along the coordinates of the plane when using the normalized unit square of the portrait, in the following we shall consider only the function $f(x)$:

$$f(x)=\begin{cases} 2x+1.5 & if \quad -0.75 \le x < -0.25; \\ 1 & if \quad -0.25 \le x \le 0.25; \\ -2x+1.5 & if \quad 0.25 < x \le 0.75; \\ 0 & if \quad |x| > 0.75. \end{cases} \tag{1}$$

To generate a random variable subject to distribution (1), one can use the functional result transformation of the standard for most programming systems of a random number generator located continuously uniformly in the range $[0,1]$. We use the inverse function method: if z:*unif* $[0,1]$ then the random variable x obtained by a functional transformation z in the form of

$$x=\begin{cases} \sqrt{z}-0.75 & if \quad 0 \le z < 0.25; \\ z-0.5 & if \quad 0.25 \le z \le 0.75; \\ -\sqrt{1-z}+0.75 & if \quad 0.75 < z \le 1; \end{cases} \tag{2}$$

will have a PDF (1).

Figure 5 shows the histogram of the statistical tests of the functional transformation (2) from *unif* $[0,1]$ the number of trials equal to 30,000 and dividing the interval $[-0.75, +0.75]$ into 100 equal subintervals. The dashed line in Fig. 5 shows the envelope (1).

The resulting algorithm for random number generation will be used later to obtain the coordinates of the characteristic points of the normalized square fingerprint portraits.

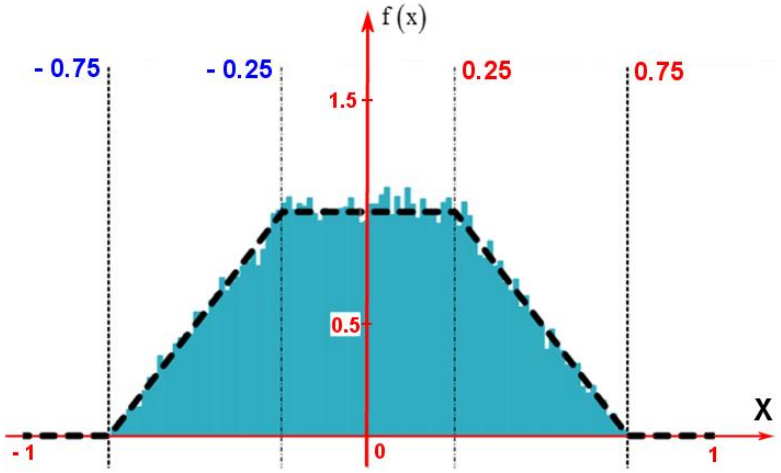

Fig. 5 – Result of statistical tests of the random coordinate sensor

Let us choose the use of a discrete (integer) random value N from the range of integers $[15, 45]$ with a discrete normal truncated distribution and the following numerical characteristics to generate a random variable – number of minutiae in a fingerprint portrait sample:

- mathematical expectation $m_N \approx 30$;
- standard deviation $\sigma \approx 2 \div 5$.

Let us again use the functional transformation data of the generator $unif\,[0,1]$ to obtain a random number of minutiae on a fingerprint portrait sample . We simulate the samples of a random variable N based on the central limit theorem. We proceed to the discrete form of uniformly distributed numbers using the operation of integer rounding and centering:

$$z' = round\,(z) - 0.5\ , \text{where}\ z{:}unif\,[0,1]. \tag{3}$$

Then, limiting the number of terms to $m_N = 30$, the random number of minutiae in the portrait can be determined as the sum

$$N = \sum_{i=1}^{30} z' + 30. \tag{4}$$

A discrete random variable N can take integer values from a range $[15, 45]$. The truncated normal function of the PDF of this random variable is approximated by weighted binomial coefficients:

$$Q(N_i) = \binom{i}{30} \cdot \left(\frac{1}{2}\right)^{30},\ i \in [0,30],\ N_i \in [15,45], \tag{5}$$

where $Q(N_i)$ – is the probability that the number of minutiae on a portrait (taking into account points masked outside the unit square) will be a value N_i.

The form and numerical characteristics of the distribution (5) are shown in Fig. 6.

To simulate the random values of the minutiae angles normalized in the unit square, it is expedient to use a random variable $z{:}unif\,[0,1]$ uniformly distributed over a unit interval:

$$\varphi = z. \tag{6}$$

The true minutia angle is determined on the basis of the normalized value (6): $\Phi = 2\pi \cdot \varphi$.

Table 1 presents the results of modeling a normalized portrait based on the distributions (1), (4), and (6). The highlighted rows in Table 1 correspond to points that did not fall into a unit square. Therefore, in spite of the fact that during the experiment we obtained $N = 28$, only 21 points were found in the unit square (Fig. 7). "Masked" points can appear in case the shifts along the axes X and Y or the image rotation occur.

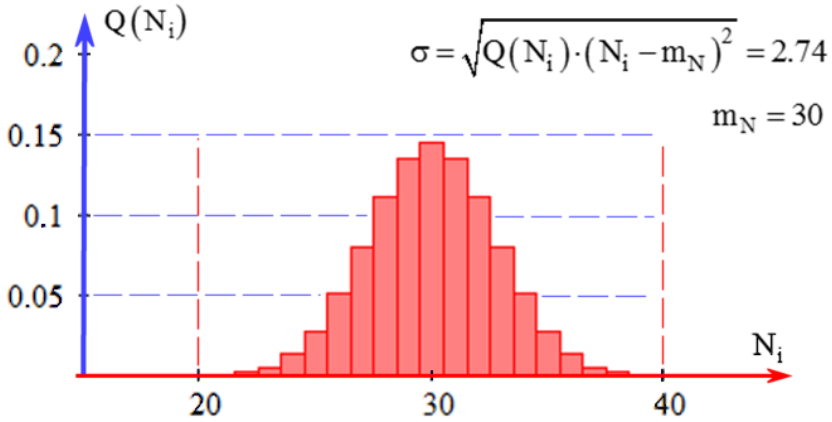

Fig. 6 – PDF of the number of minutiae on a normalized fingerprint portrait

Table 1 –Portrait matrix

№	X	Y	φ	№	X	Y	φ
1	0.21	-0.07	0.6	15	0.01	0.39	0.15
2	0.28	0.18	0.58	16	-0.05	-0.55	0.08
3	0.12	-0	0.49	17	0.51	0.5	0.64
4	0.19	0.31	0.74	18	-0.25	-0.34	0.55
5	0.07	-0.68	0.62	19	0.23	-0.41	0.41
6	0.05	-0.12	0.8	20	0.13	-0.41	0.47
7	-0.25	0.33	0.58	21	0.14	0.08	0.15
8	-0.01	0.42	0.91	22	0.06	0.23	0.74
9	0.17	0.15	0.73	23	-0.05	0.17	0.83
10	0.12	0.23	0.67	24	0.69	0.31	0.87
11	0.3	0.54	0.32	25	0.1	0.7	0.3
12	0.64	-0.14	0.31	26	-0.33	-0.3	0.13
13	0.13	0.44	0.11	27	-0.17	-0.11	0.78
14	0.09	-0.05	0.85	28	0.15	0.08	0.61

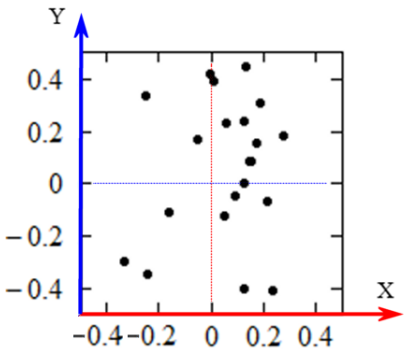

Fig. 7 – Random points distribution sample

In case of necessity, it is possible to consider a three-dimensional space for point placing by adding a third coordinate for the normalized angle φ in the corresponding processing algorithm.

4 Conclusions

The analysis of the fingerprint scanning results shows an extremely small degree of similarity among the obtained images. Considering the nature of the possible errors, visual similarity takes

place only in case of similar scanning conditions. The cause of the problem mentioned above is not only the complexity of the procedure itself, but also the imperfection of the used recognition algorithms. The research involves development of a mathematical model for the probabilistic minutiae distribution in biometric fingerprint images. The suggested model is based on heuristic analysis of the fingerprint scanning results with account for the nature of the potential errors. She allows to model a typical minutiae behavior in the biometric fingerprint images.

References

[1] Xudong Jiang and Wei-Yun Yau "Fingerprint minutiae matching based on the local and global structures, "Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, 2000, pp. 1038-1041 Vol. 2.
[2] D. Maltoni, D.Maio, A.K.Jain, S.Prabhakar. Handbook of Fingerprint Recognition, Springer, New York, 2003.
[3] "Privacy Enhancing Technologies for Biometric Data", 2015. [On-line]. Internet: http://www.cs.haifa.ac.il/~orrd/PrivDay/2015/
[4] "Privacy Enhancing Technologies for Biometric Data", 2016. [On-line]. Internet: http://www.cs.haifa.ac.il/~orrd/PrivDay/
[5] N. K. Ratha, K. Karu, Shaoyun Chen and A. K. Jain, "A real-time matching system for large fingerprint databases," in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 8, pp. 799-813, Aug 1996.
[6] ISO/IEC 19794-2. Information technology – Biometric data interchange formats – Part 2: Finger minutiae data.
[7] ISO/IEC 19794-3. Information technology – Biometric data interchange formats – Part 3: Finger pattern spectral data.
[8] ISO/IEC 19794-4. Information technology – Biometric data interchange formats – Part 4: Finger image data.
[9] ISO/IEC 19794-5. Information technology – Biometric data interchange formats – Part 5: Face image data.
[10] Craw, I., Costen, N.P., Kato, T., Akamatsu, S., "How should we represent faces for automatic recognition?", IEEE Trans. Pat. Anal. Mach. Intel. 21725–736, 1999.
[11] Draper, B.A., Baek, K., Bartlett, M.S., Beveridge, J.R., "Recognizing faces with PCA and ICA", Computer Vision and Image Understanding, 91:115-137, 2003.
[12] C. Xiang, X.A. Fan, T.H. Lee. "Face recognition using recursive Fisher linear discriminant." Communications, Circuits and Systems. – 2004. – Vol. 2. – pp. 27-29.
[13] "FVC2004. Fingerprint Verification Competetion. Databases". [On-line]. Internet: http://bias.csr.unibo.it/fvc2004/databases.asp.
[14] "SourceAFIS for Java and .NET". [On-line]. Internet: https://sourceafis.machinezoo.com/
[15] "SourceAFIS Fingerprint recognition library for .NET and experimentally for Java". [On-line]. Internet: https:// source-forge.net/projects/sourceafis/

Рецензент: Вячеслав Харченко, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Національний аерокосмічний університет ім. М. Є. Жуковського, м. Харків, Україна.
E-mail: v_s_kharchenko@ukr.net

Надійшло: Червень 2018.

Автори:
Сергій Рассомахін, д.т.н., зав. кафедри Безпеки інформаційних систем і технологій, Харківський національний університет ім. В. Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна.
E-mail: rassomakhin@karazin.ua

Олександр Кузнєцов, д.т.н., проф., академік Академії наук прикладної радіоелектроніки, Харківський національний університет ім. В. Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна.
E-mail: kuznetsov@karazin.ua

Володимир Шлокін, директор Інноваційного центру, Харківський національний університет ім. В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна.
E-mail: vshlokin@ukr.net

Іван Білозерцев, студент, Харківський національний університет ім. В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022, Україна. E-mail: ivanbelozersevv.jw@gmail.com

Роман Сергієнко, Національна академія СВ ім. гетьмана П. Сагайдачного, вул. Героїв Майдану, 32, м. Львів, 79012, Україна.
E-mail: romanserg69@gmail.com

Імовірнісний розподіл мінуцій в біометричних зображеннях відбитків пальців.

Анотація. Аналіз результатів сканування відбитків пальців показує надзвичайно малу ступінь подібності отриманих зображень. Причиною зазначеної проблеми є не тільки складність самої процедури, а й недосконаліть алгоритмів розпізнавання, що використовуються. Дослідження включає в себе розробку математичної моделі для імовірнісного розподілу мінуцій в біометричних зображеннях відбитків пальців. Запропонована модель заснована на евристичному аналізі результатів сканування відбитків пальців з урахуванням характеру потенційних помилок. Вона дозволяє моделювати типову поведінку мінуцій в біометричних зображеннях відбитків пальців.

Ключові слова: біометрична автентифікація; зображення відбитків пальців; мінуція.

Рецензент: Вячеслав Харченко, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Национальный аэрокосмический университет им. М. Є. Жуковского, г. Харьков, Украина.
E-mail: v_s_kharchenko@ukr.net

Поступила: Июнь 2018.

Авторы:

Сергей Рассомахин, д.т.н., зав. кафедры Безопасности информационных систем и технологий, Харьковский национальный университет им. В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина.
E-mail: rassomakhin@karazin.ua

Александр Кузнецов, д.т.н., проф., академик Академии наук прикладной радиоэлектроники, Харьковский национальный университет им. В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина.
E-mail: kuznetsov@karazin.ua

Владимир Шлокин, директор Инновационного центра, Харьковский национальный университет им. В. Н. Каразина, пл. Свободы, 6, г. Харьков, 61022, Украина.
E-mail: vshlokin@ukr.net

Иван Белозерцев, студент, Харьковский национальный университет им. В. Н. Каразина, пл. Свободы, 6, 61022, г. Харьков, Украина.
E-mail: ivanbelozersevv.jw@gmail.com

Роман Сергиенко, Национальная академия СВ им. гетьмана П. Сагайдачного, ул. Героев Майдана, 32, г. Львов, 79012, Украина.
E-mail: romanserg69@gmail.com

Вероятностное распределение минуций в биометрических изображениях отпечатков пальцев.

Аннотация. Анализ результатов сканирования отпечатков пальцев показывает чрезвычайно малую степень сходства полученных изображений. Причиной упомянутой проблемы является не только сложность самой процедуры, но и несовершенство используемых алгоритмов распознавания. Исследование включает в себя разработку математической модели для вероятностного распределения минуций в биометрических изображениях отпечатков пальцев. Предложенная модель основана на эвристическом анализе результатов сканирования отпечатков пальцев с учетом характера потенциальных ошибок. Она позволяет моделировать типичное поведение минуций в биометрических изображениях отпечатков пальцев.

Ключевые слова: биометрическая аутентификация; изображения отпечатков пальцев; минуция.

UDC 004.9: 621.391.7

PRINCIPLES OF FORMATION, PROCESSING AND PROPERTIES OF OFDM SIGNALS

Alexandr Zamula[1], Vladislav Morozov[1], Vadim Serbin[2]

[1] V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, 61022, Ukraine
zamylaaa@gmail.com, ilissar@hotmail.com

[2] Yuzhnoye State Design Office, 3 Krivorozhskaya St., Dnipro, 49008, Ukraine
buba75@i.ua

Reviewer: Alexandr Potii, Doctor of Sciences (Engineering), Full Professor, Department of Information Systems and Technologies Security, V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, 61022, Ukraine
potav@ua.fm

Received on May 2018

Abstract. *The article discusses the technology of forming signals used in mobile, information and telecommunication systems, and also provides an analysis of promising technologies that can be used in wireless communication systems of broadband access. It is shown that the widely used modulation scheme with orthogonal frequency division (OFDM) has a number of drawbacks, which can lead to a decrease in system performance. Alternative technologies for generating signals are presented, in particular, a technology based on windowed signal processing (W-OFDM), a technology based on time division (w-OFDM); UFMC technology and others to eliminate the disadvantages of OFDM technology. New points of view are proposed on the use of multi-carrier transmission technology in the form of multiplexing with orthogonal frequency division (in order to increase the security of modern wireless broadband access communication systems from external and internal threats), a class of non-linear discrete cryptographic sequences to form a physical data carrier - signal. It is shown that the use of such signals will improve the security of these systems from inserting (imposing) false messages into the system, falsifying messages, as well as ensuring the integrity and confidentiality of data, receiving noise immunity and secrecy of the system.*

Keywords*: noise immunity; information security; broadband access; signal; integrity; noise immunity; cellular communication; frequency division; interference; peak factor.*

1 Introduction

Modern wireless systems (for example, satellite systems, mobile telephony systems) belong to multi-user systems. When designing such systems, the main problem is to choose method of multiple access, i.e. the possibility of simultaneous use by many subscribers of a communication channel with minimal mutual influence [1,2]. Broadband signals are widely used in modern high-speed cellular communication systems of the WiMax, Mobile WiMax, MBWA standards, wireless discrete communication systems, such as LTE and Wi-Fi, in the transmission of information from digital television (DVB-T) and radio (DRM, DAB), in radiolocation, etc. The use of signals with orthogonal frequency division multiplexing (OFDM), including in the specified information transmission systems, allows to increase not only the information capacity of the system in case of multipath propagation with limited bandwidth, but also the data transmission speed, bringing it closer to the channel capacity, increase the secrecy of transmission and noise immunity of the system. Currently, there is a rapid development, research and standardization of technologies for the fifth generation of cellular networks (5G). The most priority tasks in this direction are: to achieve the maximum data transfer rate (up to 20 Gbit / s); ensuring the density of user devices (up to 10^6 devices / km2); providing users with highly reliable low latency communication services (URLLC) (data transmission delay not more than 1 ms) [3-4].

In order to achieve the above objectives for 5G networks, the following are considered: use of the spectrum in the millimeter range [5]; new types of signal modulation and coding methods; multiple access methods; improved technologies for building antennas and networks architecture [5-6]. In addition, it is worth noting studies devoted to: orthogonal frequency division multiplexing with filtering (F-OFDM) [7-9]; spatial diversity technologies (MIMO) [10]; radiocommunication cloud networks (C-RAN) [11], orthogonal frequency division technology with coding (C-OFDM) [12] and many others.

2 Implementation principles for OFDM technology

The main idea of OFDM is to achieve a high transmission rate, in the frequency domain, by dividing full signal frequency range into a number of non-overlapping frequency subchannels with lower speeds. In addition, each subchannel (subcarrier) is modulated by a separate symbol, then these channels are multiplexed in frequency domain and data are transmitted in parallel in orthogonal subchannels. Compared to single carrier transmission, this approach provides enhanced resistance to narrowband interference and channel distortion. Specified, in particular, allows for a high level of system flexibility, since modulation parameters, such as constellation size, coding rate, can be independently selected for each subchannel.

The structure of the OFDM modem contains a transmitter and receiver. In the transmitter, the original serial stream of information bits (Fig. 1) is encoded with an error-correcting code (according to the recommendation of LTE 3GPP TS 36.211, a convolutional turbo code with a base rate of 1/3 is used), interleaved (I) and demultiplexed into N parallel substreams.

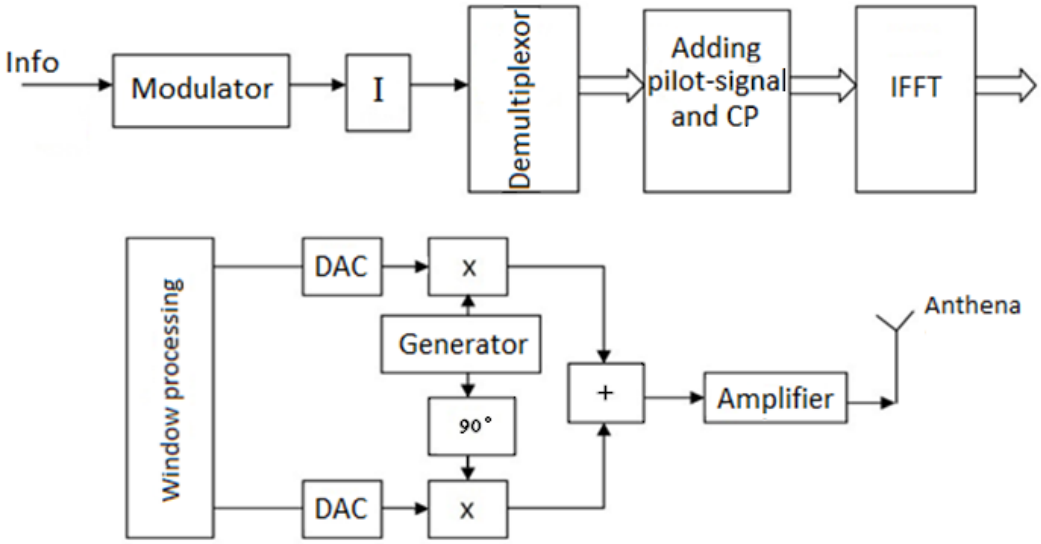

Fig. 1 – The structural circuits

Next, each of the streams is mapped to a stream of symbols using a phase modulation (BPSK, QPSK, 8-PSK) or amplitude-phase quadrature modulation (QAM). In case of BPSK modulation, a stream of binary numbers (1 and -1) is obtained, and for QPSK, 8-PSK, QAM – a stream of complex numbers. In addition to the information subcarriers, there are service subcarriers. These include guard intervals, pilot signals and additional overhead information for synchronization of the receiver and transmitter, and their modes of operation. Pilot signals may have a fixed position on the subcarriers, or variable, varying from symbol to symbol in OFDM frames. In this case, due to the insertion between the adjacent subchannels of a sufficiently long guard interval, the spectral overlap is excluded. In this case, the inter-channel interference (ICI) is reduced, the probability of bit error decreases, which means that the capacity of the wireless access system increases.

The operation of complex exponential multiplying with the corresponding subchannel frequency and then summing all the subchannels to form an OFDM signal is very similar to the operation of the inverse fast Fourier transform (IFFT). By using IFFT to form the required OFDM symbol, it greatly simplifies the implementation of modulators.

Maintaining orthogonality is necessary so that the receiver can correctly recognize the information on the subcarriers. To do this, you must fulfill the following conditions:

- the receiver and transmitter must be accurately synchronized;
- the analog components of the transmitter and receiver must be of very high quality;
- the channel should not be multipath.

Unfortunately, multipath distortion is almost inevitable in radio communication systems, which leads to a distortion of the received signal. To eliminate this type of interference, it is necessary to choose a guard interval, the duration of which is longer than the maximum propagation delay in the channel. Thus, it is possible to eliminate most types of inter channel interference (ICI) and between adjacent transmission units (i.e., inter-symbol interference (ISI)). To reduce the out-of-band emission of signals, window processing of the time signal is used, using a "raised cosine" window. Further, digital-to-analog converters (DAC) convert the real and imaginary components separately into analog form. After passing through the low-pass filter, the signal enters the quadrature mixer, which transfers the useful spectrum of the OFDM signal to the carrier frequency. These signals are then summed, amplified and the OFDM signal itself is formed.

The widespread use of the OFDM digital modulation scheme is due to a number of remarkable features of this technology:

- resistance to multipath effects;
- high noise immunity to narrowband interference;
- resistance to intersymbol interference due to the fact that the duration of the symbol in the auxiliary subcarrier is significantly longer compared to the propagation delay than in traditional modulation schemes;
- high spectral efficiency in comparison with traditional systems with frequency division of channels due to the large number of subcarriers;
- the ability to use different modulation schemes for different subcarriers, which allows you to adapt to the conditions of signal propagation and to different requirements for the quality of received signals;
- simple implementation using digital processing methods, etc.

3 Promising technologies for generating signals in modern mobile telecommunications systems

The effectiveness of the modern generation of mobile communications is largely based on the use of OFDM modulation. However, for further progress and transition to more advanced fifth-generation communication technologies, it is necessary to revise the OFDM technologies used, as well as explore other technologies. The following main differences between the 5G technology and previous generation mobile communication technologies [4-5] can be distinguished.

1. Mixed numerology. One of the goals of 5G is to ensure the use of various services, in particular eMBB, mMTC and URLLC. It is assumed that 5G technology should support more flexible use of the available frequency band to increase throughput. For this, it is necessary to develop and implement various options for using the available frequency and time resources for various services.
2. Increased bandwidth - a threefold increase in the efficiency of using the signal spectrum in 5G in comparison with eMBB services [3]. To increase the capacity in 5G networks, it is planned to reduce guard intervals [4].
3. With asynchronous data transmission in 4G networks, the base station is constantly synchronized with user equipment to reduce mutual interference between carriers [4]. Losses caused by such interference adversely affect the services, in particular, mMTC, which are associated with the mass connection of network subscribers. Thus, 5G support for asynchronous transmission is necessary in order to solve ICI-related problems and ensure operation over multiple connections [4].

As noted above, orthogonal frequency multiplexing is an access scheme that is used in modern 4G networks. Two separate signals are used to gain access to the network – an orthogonal frequency division multiplexing access (OFDMA) signal in the downlink and a single carrier frequency multiplex access (SC-FDMA) signal in the uplink. The advantages of this scheme are associated with the possibility of transmitting signals on multiple carriers. However, this OFDM scheme has a number of disadvantages, in particular: high sensitivity to frequency and clock frequency shifts; high ratio of peak signal power level to average – peak factor (PAPR); the use of guard intervals reduces spectral efficiency; sensitivity to the Doppler effect, which imposes some restrictions on its use in mobile networks; overlapping subcarrier bands leads to inter-bit interference; the OFDM sig-

nal is vulnerable to spectral conversion products caused by non-linear amplifiers, a constant component offset when using FFT. In addition, sensitivity to frequency and clock shifts necessitates the periodic addition of synchronization signals to the total amount of signals used and requires device and network synchronization before commencing communication (data exchange). The lack of continuity (phase transition) between two symbols during the generation of OFDM symbols triggers spectral spikes in the frequency domain, which leads to intense out-of-band emissions and else.

The limited capabilities of the signals based on the OFDM modulation scheme have become a prerequisite for research to select candidate signals for future generations of mobile communications, in particular 5G. In this regard, one of the tasks to be solved is to fulfill the requirement of a significant reduction in the delay in the introduction of new services and applications. Along with this, it becomes necessary to form a cyclic prefix and reduce the length of characters. These considerations led to the creation of a variety of signal conditioning technologies: with generalized frequency multiplexing (GFDM); filter bank multi-carrier (FBMC); time division OFDM (w-OFDM); universal multi-carrier filtered signal (UFMC); orthogonal frequency division multiplexing with F-OFDM filtering and others. Research is also being conducted on new multiple access schemes, including: sparse code multiple access (SCMA), non-orthogonal multiple access (NOMA) and resource-spread multiple access (RSMA).

FBMC is one of the most well-known spread spectrum modulation formats in wireless communications [14]. This modulation provides a significant advantage in the formation of each subcarrier and facilitates the flexible use of the spectral resource, allows you to meet various system requirements such as low latency, multiple access and others, which leads to improved system noise immunity under the signal scattering conditions in the time and frequency domains [15]. For example, rectangular filters are preferable for channels distributed in time, while a filter with a raised cosine characteristic is more resistant to frequency dispersion. Despite all the benefits of using FBMC, a considerable length of filters leads to a long symbol duration, which is a problem not only for applications with low latency requirements or a large number of users in communications, but also leads to an increase in computational complexity for MIMO detection technology, which, ultimately, will lead to problems in the operation of all major 5G applications.

The UFMC technology [13] is largely recommended for overcoming the ICI problem with multiple user access in asynchronous transmission mode and is based on frequency division and multiplexing by applying a subcarrier group filtering operation. UFMC is a generic version of the filtering technique for multiple sidebands. The sidebands are processed by the filter at the same time, instead of processing each power supply unit separately. Thus, mutual interference with power supply units is reduced in comparison with traditional OFDM. Also, the use of sidebands filtering operations is aimed at increasing the efficiency of a number of communications applications, such as systems with ultra-low packet latency. This type of modulation is more preferable for such applications in relation to the FBMC modulation scheme.

GFDM technology is a block modulation circuit with frequency channel multiplexing, designed to work with a variety of 5G applications, providing a variable waveform [16]. To improve reliability and communication latency without error correction, GFDM signals can be used along with the Walsh-Hadamard transform. When combining GFDM with quadrature amplitude modulation, in systems with multiple access, the problem of intra-system interference is solved subject to the use of non-orthogonal filters. From another point of view, GFDM can be considered as a scheme with flexible tuning of individual blocks, and not just one carrier as a whole. When manipulating the relevant GFDM signal parameters, it is possible to obtain various waveforms, such as OFDM, single-carrier frequency alignment (SC-FDE), etc. Despite the very promising possibilities that are opened by using signals with GFDM, this type of modulation is computationally complex [16].

F-OFDM technology is used in 4G downlink channels. For F-OFDM, the configured filter is applied to the OFDM symbol in the time domain to reduce the out-of-band emission level of the subband signal, while maintaining the orthogonality of the complex domains of the OFDM symbols. Since the filter bandwidth corresponds to the signal bandwidth, only a few subcarriers close to the edge are affected. The main consideration is that the filter length may exceed the cyclic prefix

length for F-OFDM [6]. This reduces the level of inter symbol interference due to the selected filter design using window processing (*with soft truncation*). The generation of the F-OFDM signal is based on the formation of a block of M nearby sidebands in a number of consecutive OFDM symbols [17]. In particular, during the processing of each symbol, the following parameters are formed in the transmitter: an inverse fast Fourier transform (IFFT) dimension value equal to *N*, a duration of *M* information symbols together with a cyclic prefix, where N > M. Information symbols can be constellation points as in OFDM. Analytically indicated can be represented as follows:

$$s(n) = \sum_{(l=0)}^{(L-1)} s_l(n - l(N + Ng)) \quad (1)$$

and

$$S_l(n) = \sum_{(m=m`)}^{(m'+M-1)} d_{l,m} e^{(j2\pi mn/N)}, -N_g \le n < N, \quad (2)$$

where *Ng* is the cyclic prefix length (CP), *d* – information symbol of the subcarrier m of the OFDM system, *L* denotes the number of OFDM symbols, and $\{m_0, m_{0+1}, \cdots, m_{0+M-1}\}$ is the selected set of subcarriers. The F-OFDM signal is generated by processing the *s*(*n*) signal using the appropriate filter, i.e.

$$\tilde{s}(n) = s(n) \cdot f(n). \quad (3)$$

The bandwidth of the filter is equal to the sum of the bandwidth capabilities of the selected sidebands and the time spent is the duration of the OFDM symbol. At the receiver, the received signal first passes through the filter $f(-n)$, which is identical to the filter of the transmitter. The received signal is processed using standard OFDM transforms, and then the filtered signal is divided into a sequence of individual OFDM symbols with the removal of the cyclic prefix. In this case, an FFT of dimension *N* is applied to each symbol and then informational symbols from the corresponding subcarriers are extracted.

The filter for F-OFDM must satisfy the following criteria: have a flat bandwidth over the subcarriers in the subband; have a sharp transitional strip to minimize guard bands. These criteria correspond to the filter with a rectangular frequency response. To meet these requirements, a low-pass filter is implemented using a "window", which effectively cuts off the impulse response and provides smooth transitions to zero at both ends [17]. Thus, the F-OFDM implementation adds to the existing CP-OFDM processing procedure, the filtering step on both the transmit side and the receive side.

To reduce the out-of-band emission of signals, window processing of the time signal is used, using a raised cosine window. It is known that the spectrum of the OFDM signal has many side lobes that slowly fade out in the frequency domain, which leads to an increase in out-of-band emission. To reduce out-of-band emission, OFDM symbols use guard subcarriers that are added along the edges of the OFDM signal. Window signal processing is used for the same purpose. Such signal processing allows a smooth transition between the end of the previous and the beginning of the next character. Such a transition is carried out by overlapping in time the prefix of the current character and the suffix of the previous character by summing them up.

In this case, the use of window processing for the formation of OFDM symbols can significantly reduce out-of-band emission. The out-of-band emission level is also influenced by the choice of guard interval between subcarriers. Studies have shown that the longer the guard interval, the lower the out-of-band emission level [18].

4 Security indicators evaluation of modern wireless broadband access communication systems based on OFDM technology features

State level of informatization is determined primarily by the development of information communications, as a set of network resources intended for the production and provision of telecommunications, information and other services. With the advent of new information and communication

technologies (ICT), the use of various transmission media (*optical fiber, radio frequency resources*), mobile communication systems, it has become possible to significantly increase the productivity, efficiency and quality of service of telecommunications networks, as well as expand the range of services that they provide. A variety of modern ICTs operate under conditions of external and internal influences caused, on the one hand, by natural interference, interference from other radio systems operating at close frequencies or in a common part of the frequency range, on the other hand, deliberate interference caused by counter stations for the purpose of electronic suppression of existing systems. Possible strategies of the counter station are: determining the content of messages when legal subscribers use cryptographic data protection algorithms; falsification of messages; violation of data integrity; statement of various types of interference, etc. Therefore, ICT, especially for critical purposes, have increasingly stringent requirements to ensure the effectiveness of their operation (*information transfer speed, reliability of information transfer, survivability, noise immunity, information security*). Increased requirements for quick decision-making and communicating information to executors (users) in the context of internal and external influences are largely not taken into account by existing information technologies. There is a contradiction between strict requirements to ensure secrecy, confidentiality, integrity, reliability of data stored and transmitted over wired and wireless communication lines, on one hand, and existing models, methods and technologies for managing telecommunications networks, information security, services and quality of service, on the other hand. The main ways to solve this contradiction is to increase the noise immunity (*in particular, noise immunity, energy, structural and information secrecy*) and information security of the ICS by improving the methodological foundations of ICT building by developing information exchange methods, synthesizing new classes of signals with the necessary ensemble, correlation and structural properties.

The development of wireless communications technologies has been constantly shaped based on studies of waveforms. As an example, we can use the technology of multiplexing signals with orthogonal frequency division multiplexing in modern wireless broadband access systems (WiMAX, Wi-Fi, LTE, etc.). The use of this technology allows to increase the information capacity of the system with a limited bandwidth, data reception and transmission speed, bringing it closer to the channel capacity, increasing the secrecy of transmission and noise immunity of signal reception, and as a result, to meet the ever-increasing needs of network users in high-speed connections services.

Analytically, the OFDM signal can be represented as [19]:

$$S(t) = \sum_{k=0}^{N-1} S_k(t) = \sum_{k=0}^{N-1} A_k e^{(j2\pi kf/T)}, 0 \le t \le T, \quad (4)$$

where k is the subcarrier index, $S_k(t)$ is the signal on the k-subcarrier, A_k is the amplitude component of the sequence of information symbols, N is the number of subcarriers, T is the duration of the information symbol.

The block diagram of an OFDM modulator is presented in Fig. 2 [20,21]. In the transmitter, the serial stream of binary symbols $s[n]$ is encoded with an error-correcting code, interleaved further, using inverse multiplexing (*demultiplexing*), turns into N parallel streams, each of which is matched (*complexly*) with the output stream s[n] using a certain constellation modulations (*quadrature modulation QAM, quadrature phase modulation QPSK, etc.*). The number of outputs of the demultiplexer is determined by the number of subcarrier frequencies. Next, the modulated $X_0,\ldots,X_{N-1}$ symbol streams undergo a fast inverse Fourier transform, which translates them into digital responses $X_0,\ldots,X_{N-1}$ (in general, complex numbers) in the time domain. The real ($\mathrm{Re}\{x_i\}$) and imaginary ($\mathrm{Im}\{x_i\}$) components of the response x_i ($i = 0, ..., N - 1$) are subjected to digital-to-analog conversion. The received analog signals are used for modulation in accordance with the sine wave and cosine wave (*obtained by shifting the sine wave by 90*) of the carrier frequency. After modulation, the signals are summed to form a signal s (t), which enters the communication channel.

Subcarriers orthogonality makes it possible to select each of them from the common signal at the reception even in the case of partial overlapping of their spectra. Since the subcarriers are located close to each other and even partially overlap, the spectral efficiency of the modulated OFDM sig-

nal is high. The parameters of the subcarrier signals are selected in such a way that they are orthogonal to each other, that is, the condition is met for them:

$$\int_0^T sin2\pi f_1(t) sin2\pi f_k(t) dt = 0, \quad (5)$$

where: t is the duration of the information symbol, f_1 and f_k are the frequencies of the l-th and k-th subcarriers, respectively.

The orthogonality of the carrier signals guarantees the frequency independence of the channels from each other and, therefore, the absence of inter-channel interference. For the fast implementation of this procedure, the inverse fast Fourier transform algorithm is used, that is, the signal values at the input of the IFFT block belong to the frequency domain. At the output of the block IFFT receive the signal value in the time domain. Combining all the values, a complex OFDM signal is obtained. Taking into account the fact that IFFT works effectively with arrays of dimension 2^k, the number of subcarriers is chosen with the same multiplicity. For example, in WiMAX wireless communication systems, the number of subcarriers is chosen from 128 to 2048 and can occupy frequency bands from 1.25 MHz to 20 MHz. For each of the subcarriers, a different modulation type is used depending on the requirements and the type of interference in the channel. At the receiving end, the inverse operations are performed, in this case, instead of a digital-to-analog converter, an analog-to-digital converter (ADC) is used, instead of a reverse FFT, direct FFT.

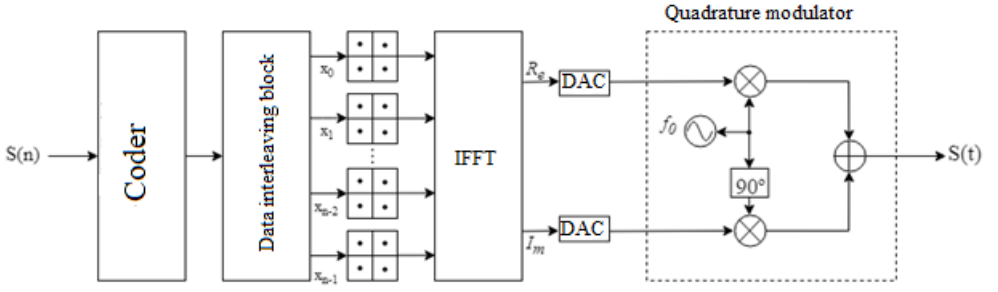

Fig. 2 – OFDM modulator circuit

The structure of the OFDM signal can be quite complex because it consists of many components:

- the structure of the time-frequency distribution, given by: initial frequency, frequency grid pitch, number of subcarriers;
- time slots specified by: the duration of the symbol, the duration of the guard interval;
- type of manipulation: phase (BPSK, QPSK, 8-PSK) or amplitude-phase quadrature modulation (QAM);
- discrete sequences that determine the law (rule) of manipulating the phase of the high-frequency carrier, and given the dimension of the signal space;
- type of symbol sync;
- the presence and type of noise-resistant coding (Reed-Solomon code, Bowes-Choudhury-Hokvingem code, turbo codes, etc.);
- the presence and type of data interleaving and so on.

The above features of the OFDM signal structure can be used in the construction of ICT, for which the ensuring requirements of the specified security indicators against the introduction (imposition) of spurious messages, falsification of messages; data integrity, confidentiality, noise immunity of reception, operation secrecy are decisive.

One of the components of information security (*along with information secrecy*) is imitation protection (*ensuring integrity*) of information. The mathematical apparatus of the imitation protection system includes a cryptographic algorithm for simulating the encryption of information (this may be an encryption algorithm, an authentication code, or another transformation) and an algorithm for deciding the truth of the information received, as well as a key system. In essence, imitational secu-

rity is a complex service that is provided by the provision of such services as integrity, authenticity, (truth), as well as the use of various cryptographic protocols with certain properties [22,23]. As studies [24,25] have shown, it is possible to provide the imitability necessary in ICS at the source level of complex signals by increasing the dimension of the signal space, the degree of correlation between them, the complexity of the laws of their construction. In accordance with the above definitions, the theory of authentication by J. Simons [26] can be used to quantify the simulated security. It was Symons who showed that for the quantitative assessment of authenticity one can use the probability of deception:

$$P_{обм} \geq 2^{-\Delta I(C,K)}, \quad (6)$$

where $\Delta I(C, K)$ is the amount of information on the authentication key K entered into the C cryptogram.

Let us analyze the expression (3).

1. Systems in which equality (3) is achieved are referred to as systems that are absolutely resistant to deception.

2. To reduce the probability of fraud, it is necessary to increase $\Delta I(C, K)$.

Taking into account the peculiarities of the OFDM signal structure, imitability (I_c) depends on: the dimension of the signal space (I), the number of attempts (C) of imposing (simulating), the space (Z) of the component of the OFDM signal structure (*in particular: initial frequency, frequency grid spacing, ensemble of discrete sequences (signals), number of subcarriers, etc.*), imposition strategies (X):

$$I_c = F(I, Z, C, X). \quad (7)$$

At the complex signals source level (physical level), the probability of cheating or imposing a false signal is defined as

$$P_{обм} \geq 2^{-l_i}, \quad (8)$$

where l_i is the length authentication code, the dimension of the signal space.

Let us show the possibility of increasing the imitability of wireless communication systems based on the use of various classes of discrete manipulating sequences (*hereinafter referred to as DS*). In [1,2,27-29] presents the results of research, which are devoted to the issues of synthesis, formation and study of the properties of a new class of DS, nonlinear discrete cryptographic sequences (hereinafter – CS). The synthesis of DS and signals obtained, for example, by manipulating the phase of a high-frequency carrier according to the CS law, is based on the use of random (pseudo-random) processes, including key data of cryptographic algorithms, and at the same time, the signals must have: absolute structural secrecy regarding their laws formations; improved ensemble properties (*exist for almost any period value, have a significant amount of signal system*); improved correlation properties, which will provide the necessary (*for a particular application of the ICS*) values of the indicators of noise immunity, information security and secrecy of the system. A special property of cryptographic signal systems is the possibility of their recovery in space and time using keys and a number of other parameters that are used in the process of synthesizing such signals.

We will evaluate the imitation resistance of the ICS radio channel for solving the problem of distinguishing signals when applying a dynamic mode of compliance change: the message bit is a complex signal, and various signal systems. Table 1 shows the results of the evaluation of the ensemble properties of various complex signal systems (M-sequences, sequences with a three-level value of the cross-correlation function (PCCF), cryptographic signals (CS)), the maximum achievable values of the side lobes of the cross-correlation function (*the so-called "dense border" packaging*) for the corresponding periods of the DS, as well as the values of the probabilities of imposing, obtained in accordance with the expression (4), when used in the ICS as a physical transfer data of the indicated classes of signals. Analysis of the data in Table 1 shows that the proposed method for the synthesis of complex non-linear discrete cryptographic signals allows the formation of large ensembles of discrete sequences.

Table 1 – Signal properties

Signal class	Sequence period	Dense packing value	Signal ensemble volume	Imposing probability value
M-sequences	31	9	3	$3\ 10^{-1}$
PCCF	31	9	495	$2\ 10^{-3}$
CS	31	9	1465137	$7\ 10^{-7}$
M-sequences	63	17	20	$5\ 10^{-2}$
PCCF	63	17	975	$1\ 10^{-3}$
CS	63	17	12 214 869	$8\ 10^{-7}$
M-sequences	127	27	36	$2\ 10^{-2}$
PCCF	127	17	11610	$8\ 10^{-5}$
CS	127	27	9006648	$1\ 10^{-7}$
M-sequences	255	36	28	$3\ 10^{-2}$
CS	255	36	17599	$5\ 10^{-5}$
M-sequences	511	63	276	$3\ 10^{-3}$
PCCF	511	33	147500	$6\ 10^{-6}$
CS	511	63	2666671	$3{,}7\ 10^{-7}$
M-sequences	1023	100	435	$2\ 10^{-3}$
PCCF	1023	65	338000	$3\ 10^{-6}$
CS	1023	100	5293538	$2\ 10^{-7}$

So for the period of the sequence $N = 63$ the number of pairs of CS satisfying the maximum value of the maximum PCCF side lobes – 17 is 12214869. For a representative of the class of linear sequences – sequences with a three-level cross-correlation function (Gold set), which are optimal from the point of view of cross-correlation functions [30], the number of pairs of signals corresponding to a given boundary is 975. The excess of the volume of a CS over an ensemble composed of M-sequences is more than 10^7 times. For the period of a sequence of 1023 elements, the number of pairs of gearboxes satisfying the limiting value for the side lobes of the cross-correlation function (CCF) 100 is 5293538, whereas for a representative of the class of linear sequences of M-sequences, the number of pairs that meet this boundary is 435, then there is an excess of the volume of the signal system is more than 10^5 times. With a slight decrease in the requirements for the limiting value of the maximum lateral peak of PCCF, according to which the selection of signals is carried out (in fact, a reduction in the noise immunity of reception), the performance of the ICS system can be significantly improved. So, for the period of the sequence $N=127$, increasing the limit value by 1.2 dB will increase the ensemble volume from $M = 11610$ (at the border of 17) to 9006648 signals, with a limit value of 27, that is, 776 times. As follows from the data Table. 1, the probability values of imposing in the case of the application of the CS is much less. So, with a period of the sequence $L = 1023$, it is four orders of magnitude less than using M-sequences and an order of magnitude less than when using sequences with a 3-level PCCF. An improvement in the imitation resistance index of the ICS is achieved due to the fact that the CS have improved ensemble properties in comparison with linear classes of signals, in particular, M-sequences.

In Table 2 shows the results of calculating the statistical characteristics of various correlation functions for discrete signals widely used in communication systems, including the characteristics of cryptographic DS. Calculations were carried out for different values of the DS period. The statistical characteristics of the correlation functions were selected: the value of maximum lateral emissions R_{max}, the value of the expectation of the emission module $m_{|R|}$, the value of the standard deviation of the emission module $D_{|R|}^{\frac{1}{2}}$ and emission values $D_{|R|}^{\frac{1}{2}}$.

Analysis of the data given in Table 2, suggests that maximum lateral emissions values of the CS, as well as the statistical characteristics of this class of signals are not inferior to the corresponding characteristics of the signals constructed using M-sequences and characteristic discrete signals [31]. This, in turn, indicates that the use of a CS provides noise immunity for receiving signals no worse than when applying the decree of signals based on linear formation laws.

Table 2 –Statistical characteristics of correlation functions

Signal type	Characteristics	$\frac{R_{макс}}{\sqrt{N}}$	$\frac{m_{\lvert R\rvert}}{\sqrt{N}}$	$\frac{D^{1/2}_{\lvert R\rvert}}{\sqrt{N}}$	$\frac{D^{1/2}_{(R)}}{\sqrt{N}}$
Characteristic discrete sequences	AACF	1,0 – 1,8	0,5	0,4	0,5
	PACF	0,1 – 1,9	0,2	0,1	0,2
	MIACF	1,4 – 2,6	0,6	0,5	0,8
	ACCF	1,9 – 3,2	1,0	0,8	1,0
	PCCF	2,5 – 3,6	1,0	0,8	1,2
	CCCF	2.1 – 5,0	0,9	0,7	1,1
M-sequence	AACF	0,7...1,25	0,32	0,26	0,41
	PACF	$1/\sqrt{N}$	$1/\sqrt{N}$	0	0
	MIACF	1,3...2,3	0,66	0,49	0,82
	ACCF	1,4...5,0	0,54	0,48	0,73
	PCCF	1,9...6,0	0,8	0,62	1,0
	CCCF	2,0...5,1	0,83	0,62	1
Cryptographic sequences	AACF	1,2 – 1,9	0,5	1	1,1
	PACF	0,2 - 1,9	0,6	0,4	0,7
	ACCF	1,4 – 3,4	0,5	0,4	0,6
	PCCF	1,9 – 5,2	0,7	0,5	0,8

From the data of Tables 1-2 it also follows that by varying the limiting values of the side-lobe level of the correlation function, depending on the requirements for the ICS, the tasks of achieving the required values of the noise immunity indicators for signal reception, imitation resistance and stealth of the ICS can be solved.

Let us make an ICS protection assessment from imposing false messages, for the case when the system uses a dynamic shift mode (according to the law of the control sequence) correspondence: the message bit is a complex signal. In this case, the value of the probability of imposing a false message ($P_{imp_{\overline{mes}}}$) (with equiprobable choice of characters of the control sequence) can be defined as:

$$P_{imp_{\overline{mes}}} = ((2)^{-k})^{n}, \tag{9}$$

where: the number of possible states of the source of the control sequence, which is determined by an ensemble of discrete signals of information carriers; n – message length provided in bits.

Table 3 shows the values of the probability of imposing $P_{imp_{\overline{mes}}}$. On the message for discrete signals obtained on the basis of carrier manipulation according to the law of M-sequences, PCCF and nonlinear cryptographic sequence. The message size is n = 32. In the calculations of $P_{imp_{\overline{mes}}}$, For the case of application in the system of nonlinear CS, sequences were selected whose correlation characteristics are close to the optimal limit values from the point of view of PCCF ($R_{max} \le 1,5\sqrt{N}$).

Data analysis from Table 3 shows that in ICS, which apply signal multiplexing technologies in orthogonal frequency division of channels, the value of $P_{imp_{\overline{mes}}}$ for nonlinear CS is much less than in the case of using linear classes of signals.

5 Models and methods for constructing a secure ICS based on the use of OFDM technology

The rapid development of communication systems and multimedia technologies has led to a significant increase in the amount of information transmitted and, consequently, the need to create reliable high-speed data transmission and information security technologies. Information security technologies in communication systems are widely used in the military sphere, copyright protection, personal data protection and in many other areas. Critical information should be inaccessible to attackers, which is one of the important requirements for modern communication systems.

Table 3 – Imposing message probability

Signal period	$P_{imp_{mes}}$ Value for signal systems:		
	M-sequences	PCCF	nonlinear CS
31	2^{-96}	2^{-288}	2^{-672}
63	2^{-96}	2^{-320}	2^{-768}
127	2^{-160}	2^{-448}	2^{-640}
1023	2^{-192}	2^{-608}	2^{-736}

When transmitting important information, it is necessary to use a secretive data transmission system, which involves ensuring protection both from unauthorized access to information and hiding the fact of transmission itself [32].

Orthogonal Frequency Division Multiplexing is a multichannel modulation method used in a variety of modern mobile communication standards. The basic principle of this method is to divide a message into multiple messages transmitted in parallel with a lower speed, which allows for an increase in the transmission speed. Through the use of frequency guard intervals, OFDM successfully counters the inter-channel and intersymbol interference that occurs in other modulation schemes with multiple carriers. Due to the advantages of this method, OFDM is widely used as a modulation and multiplexing method, in many telecommunications applications, including digital television (DAB, DVB), wireless networks (Wi-Fi) and mobile communications (LTE) [33].

6 Secured communications

To meet the growing demand for fast, reliable, and secure wireless data transmission, further research is needed in scope to hide the fact of data transmission. As a rule, the disadvantages of the OFDM method are compensated by the implementation of cryptographic methods at the application level [34]. To transmit sensitive information over public networks, a secure communications system is required. However, in wireless systems, an information signal can be received by anyone, including an attacker. The mere fact of the possibility that data transfer process can be detected, can be of great importance for communication systems in which information that is critical for its owners circulates. Using the means of steganography, it is possible to hide confidential data in other data, in particular - multimedia. Modern communication systems use channel noise to hide data transmissions [35]. The review, the main limitations and possibilities of the developed methods of covert data transmission are discussed in [36] and the channel capacity with additive white Gaussian noise in [37]. The widespread use of computer systems and open source software increases the ability of attackers to gain access to critical data. A possible solution to this problem is to modify the properties of the physical layer. A new solution is proposed to protect critical information transmitted over communication networks. The essence of this solution is to encrypt data at the physical level using OFDM technology.

7 OFDM wavelet transform

The main advantage of the traditional OFDM method with fast Fourier transform for wireless systems is its speed. However, FFT-OFDM requires a cyclic prefix to eliminate intersymbol interference, which reduces throughput. As an alternative to the fast Fourier transform, one can use the wavelet transform (WT) [38]. Unlike the fast Fourier transform, the wavelet transform uses the frequency and time characteristics of the signals. Based on the OFDM wavelet transform (WPT-OFDM) using batch conversion, it is more efficient to resist interference due to the high spectral qualities of the wavelet filters compared to FFT-OFDM [39,40]. WPT-OFDM does not use a cycle prefix, which will increase system throughput compared to FFT-OFDM [41,42]. In addition, the WPT-OFDM technology can significantly increase the noise immunity of signal reception [43,44].

8 Hidden communication using FFT-OFDM

The implementation of the hidden communication model in the FFT-OFDM system was modeled using MATLAB, and the overall efficiency is estimated by comparing the error probability at the corresponding signal-to-noise values. The block diagram of the developed communication model is shown in Fig. 3.

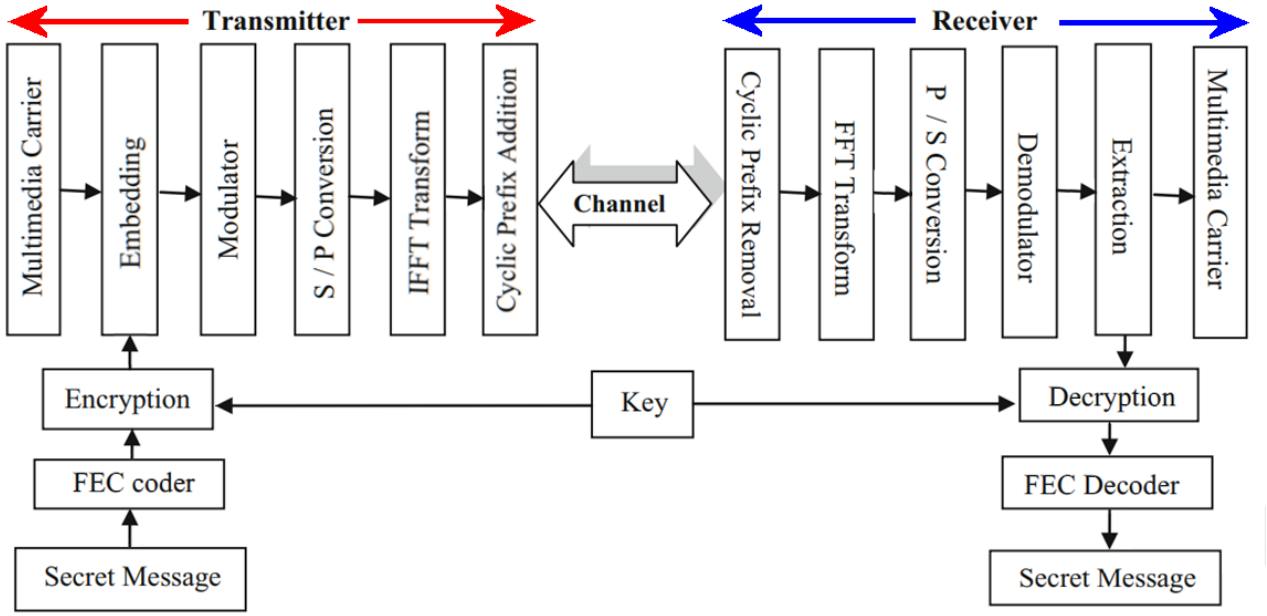

Fig. 3 – Hidden communication system model at the physical level FFT-OFDM

The message is encoded using Forward Error Correction (FEC). Using FEC, various encoding methods are supported, such as RS codes, convolutional codes (CC), etc. In this paper, we used convolutional codes to model the processing of information, which introduce redundancy in the transmitted message. In order to protect against unauthorized access to information, information bits are fed to the input of a data encryption device using a secret key. To decode the received message on the receiving side, the Viterbi algorithm is used. Encrypted data is embedded in a multimedia medium, in which image was used. The least-significant bit (LSB) replacement method was used to embed the message in the medium. The serial data stream is transmitted to the modulator and processed using M-ary PSK/QAM modulation. The pilot signal is used in the developed system to assess the quality of the channel and establish synchronism in the system. Modulated data is converted into multiple parallel streams with a low data rate, each of which is modulated by orthogonal carriers using an inverse fast Fourier transform (IFFT). The result of the IFFT transform is the summation of discrete signals in the time domain as follows [45]:

$$y_k = \frac{1}{N}\sum_{m=0}^{N-1} Y_m e^{\frac{j2\pi km}{N}} \qquad (10)$$

where $\{y_k|0\leq k \leq N–1\}$ is a sequence in a discrete time domain, $\{Y_m|0 \leq m \leq N–1\}$ – complex numbers in the discrete frequency domain.

The cyclic prefix (CP) replicates the "L" number of samples from the end of the "N" sample of the FFT frame and uses them at the beginning of each "N+1" FFT frame. The received OFDM symbol is then transmitted over the channel. On the receiving side, the CP is deleted, and the data is transmitted to the block performing the fast Fourier transform. The result of the FFT is the summation of the received signal in the frequency domain as follows

$$Y_M = \sum_{m=0}^{N-1} y_k e^{\frac{-j2\pi km}{N}} . \qquad (11)$$

At the receiving side, in order to recover information, the process is reversed to that performed in the transmitter. Further, the received data is decrypted using the secret key.

9 Hidden connection using WPT-OFDM

The structure of the transmitter and receiver of the communication system using WPT-OFDM is shown in Fig. 4.

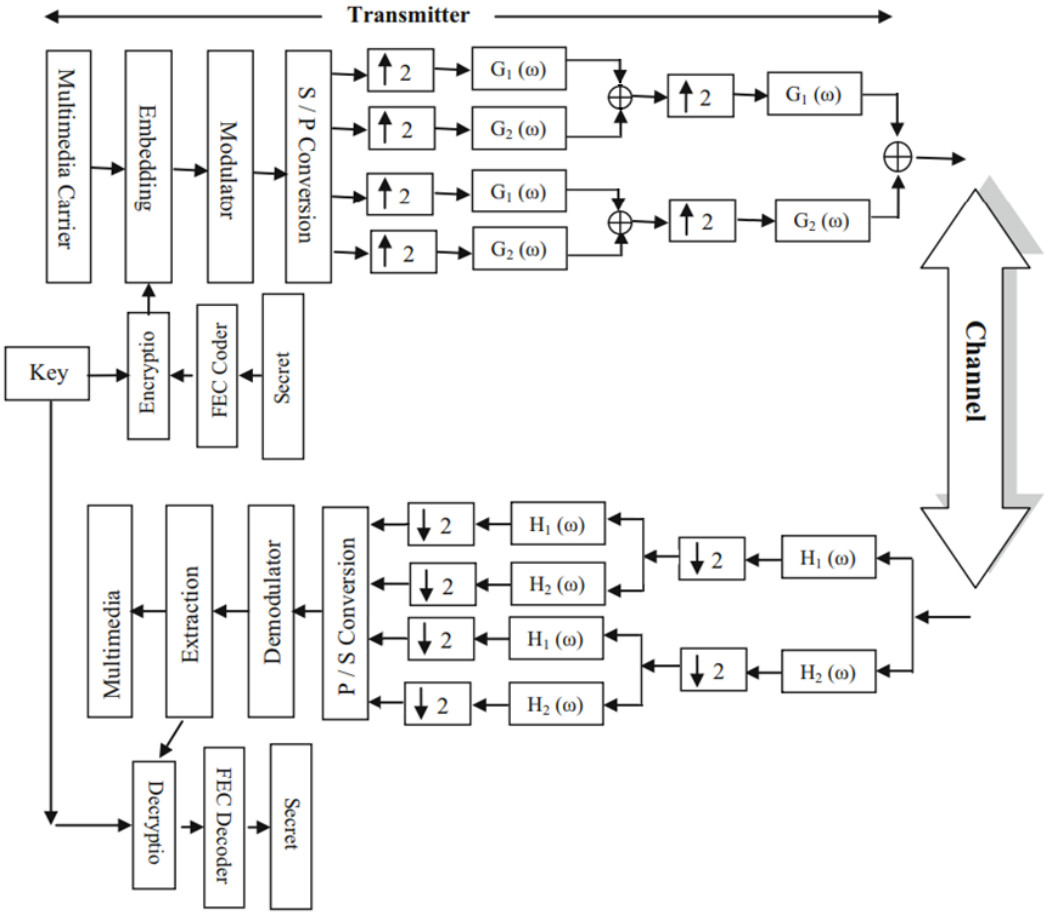

Fig. 4 – Block diagram of the WPT-OFDM communication model

The FFT technology is beneficial in terms of low computational complexity and cost, but at the same time, it has less bandwidth due to the addition of CP [46]. As shown in Fig. 4, the informational message is encoded, encrypted, and inserted into the image. In the model structure using the wavelet transform, the FFT and IFFT are replaced with WT and IWT, respectively. Wavelet-batch conversion decomposes the spectra of signals unevenly, limited to the low-frequency component of the signal. Wavelet transform (WT) is implemented using low-pass (LPF) and high-pass (HPF) filters, after which decimation is performed to increase the conversion efficiency. The WT signal 'z' is realized by passing it through a series of filters with a pulse response 'g' [45]:

$$y[n] = (z \otimes g)[n] = \sum_{k=-\infty}^{\infty} z[k]g[n-k] \quad . \tag{12}$$

The high-pass filter decomposes this signal and provides detailed coefficients. The output sampling rate of the output of the filter is reduced by 2 times. LPF and HPF results are as follows:

$$y_{LPF}[n] = (z \otimes g)[n] = \sum_{k=-\infty}^{\infty} z[k]g[2n-k] \tag{13}$$

$$y_{HPF}[n] = (z \otimes h)[n] = \sum_{k=-\infty}^{\infty} z[k]h[2n-k] \tag{14}$$

In the transmitter, encrypted and embedded data streams increase the sampling rate and are transmitted through the IWPT filter group. At the receiver, the data is transmitted through a filter, and then the encrypted data is extracted.

A communication system using FFT-OFDM has a lower throughput due to the use of the cyclic prefix. Moreover, the orthogonality of the carriers in FFT-OFDM systems is subject to channel attenuation. Resistance to such interference is higher in WPT-OFDM systems. In addition, the proposed WPT-OFDM model is more reliable for narrowband interference and multipath propagation.

10 Analysis of the effectiveness of the proposed covert communication systems

For comparison, BER estimates are presented with different SNR values for the carrier and messages in the developed OFDM system. Quality is estimated using the peak signal-to-noise ratio (PSNR), root mean square error (MSE) and average difference (AD) - the differences between the original image and the modified average difference between the transmitted and received data on the Rayleigh channel with fading:

$$MSE = \frac{\sum_{M,N}(T(r,c)-T'(r,c))^2}{M*N}, \quad (15)$$

$$PSNR = 10*\log_{10}[\frac{R^2}{MSE}], \quad (16)$$

$$AD = \frac{\sum_{M,N}(T(r,c)-T'(r,c))}{M*N}, \quad (17)$$

where: $T(r, c)$ is the original image; $T'(r, c)$ – modified image; r and c are the number of rows and columns in the input images, respectively; R – the maximum value of the image intensity.

An image of 512×512 in size was used as the transmitted container with critical information. The parameters of the communication system model are presented in Table 4.

Table 4 – Communication system model parameters

Parameter	Value
FFT size	256
CP size	1/4
Carrier quantity	12
Convolutional codes rate	1/2
Modulation	M-PSK
Modulation levels	2,4.8
Simulation channel	Rayleigh
Container	Image (512×512)
Cryptography method	XOR
Wavelet type	Haara

An increase in the signal-to-noise ratio results in a reduction in the amount of distortion when a message is extracted from the transmitted image. When used as a data channel WPT-OFDM, provides better communication quality and fewer errors. Fig. 5 shows a comparison of the BER parameter for hidden communication in FFT-OFDM and WPT-OFDM over the Rayleigh channel with fading for modulation of the form: BPSK, QPSK and 8-PSK. The assessment was carried out both for the transmitted container and for the critical message with different SNR values. Obviously, higher SNR values result in fewer errors in receiving the container, and therefore greater accuracy in retrieving the message. The results obtained demonstrate that hidden communication in WPT-OFDM provides better performance than hidden communication in a traditional OFDM system with different modulations, since the OFDM system is influenced by Doppler frequency changes. Fig. 5 (a) shows that to achieve a BER of 10^{-4} for data with modulated BPSK carrier, OFDM requires a SNR value of 9 dB, while for the proposed WPT-OFDM system, only 6 dB. To transmit a

secret message with a code rate of CC 1/2, for normal OFDM, a SNR of 6 dB is required, against 3 dB in WPT-OFDM. When using QPSK modulation, the container requires 12 dB for FFT-OFDM and 7 dB for WPT-OFDM. For the transmission of the most critical message with a CC 1/2 rate, the ratio is 10 to 7 dB.

It should be noted that using M-PSK modulation under the same conditions, the BER value of the system begins to deteriorate, but it is possible to compensate for this degradation with a high SNR value. Thus, covert communication using WPT-OFDM provides an improvement in the signal-to-noise ratio of 3-6 dB compared to FFT-OFDM when transmitting data over the Rayleigh channel. A comparison of the performance of hidden communication systems is given in Table 5. Fig. 6 and fig. 7 show a comparison of the PSNR values and the average difference between empty and full containers.

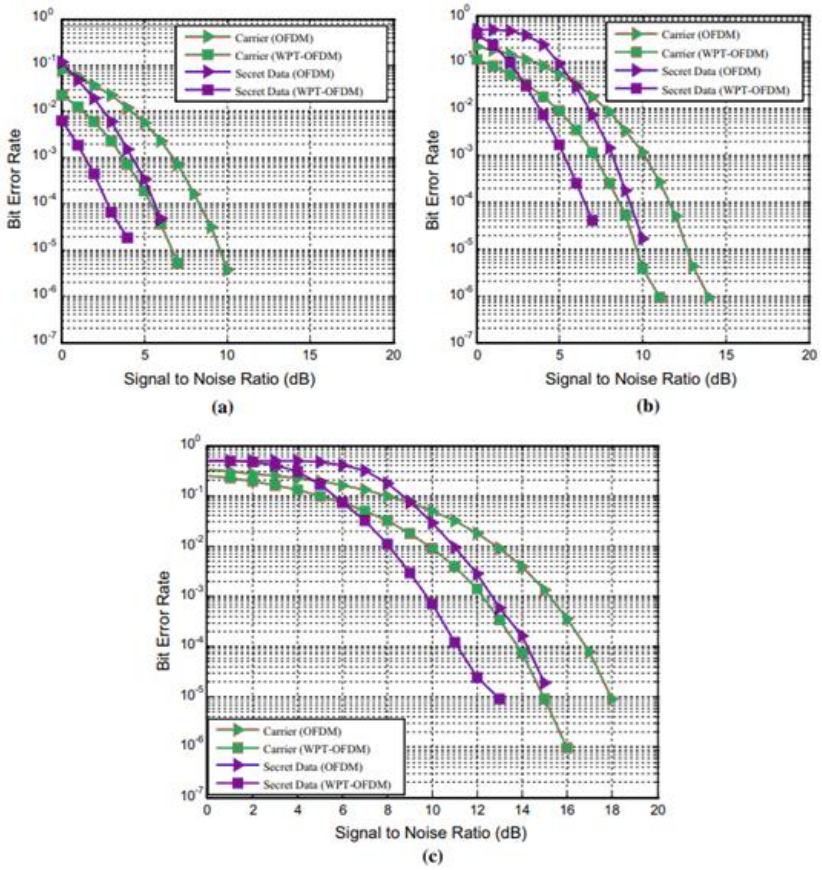

Fig. – 5 BER comparison for: BPSK *(a);* QPSK *(b);* 8-PSK *(c)*

Table 5 – Hidden communication systems performance

SNR	Modulation	Container BER	Message BER	PSNR	MSE	AD
1 dB	FFT	0.0775	0.1151	16.035	1.62×10^{3}	1.3103
	WPT	0.0228	0.0058	21.217	491.324	0.4497
2 dB	FFT	0.0551	0.0507	17.466	1.166×10^{3}	0.9215
	WPT	0.0124	0.0015	23.832	269.066	0.2489
4 dB	FFT	0.0221	0.0065	21.372	474.14	0.2763
	WPT	0.0023	4.691×10^{-5}	30.961	52.116	0.0490
6 dB	FFT	0.0057	4.7613×10^{-4}	27.189	124.24	0.0619
	WPT	1.864×10^{-4}	0	41.671	4.426	0.0027
8 dB	FFT	7.2861×10^{-4}	0	36.411	14.857	0.0073
	WPT	5.2452×10^{-6}	0	54.829	0.214	3.166×10^{-4}
10 dB	FFT	2.766×10^{-5}	0	51.291	0.4830	1.526×10^{-4}
	WPT	0	0	Inf	0	0

We used a message with a length of 243696 bits, which is transmitted in an image of 512×512 in size. Increasing the SNR value improves the values of the BER parameters for the received information sequence and the extracted message. When sending a message in a container, BERs are significantly better than direct transmissions in the same medium. Summarizing, we can say that hidden communication using WPT-OFDM provides better quality than conventional OFDM for all the cases considered.

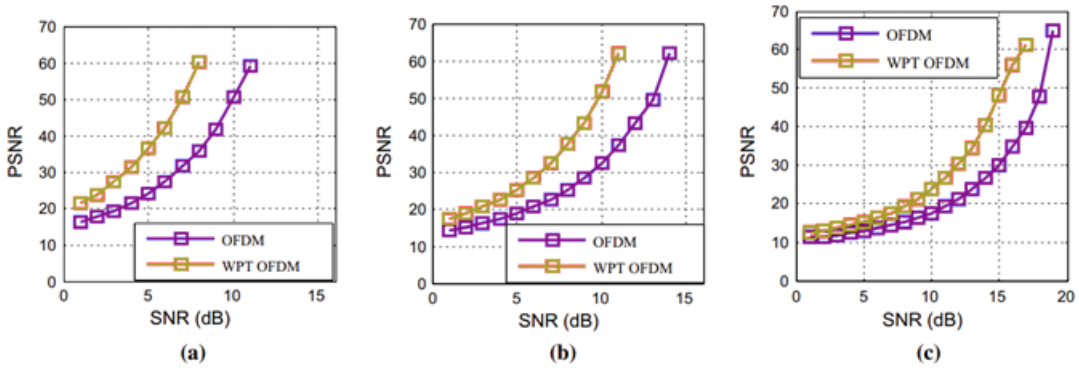

Fig. 6 – Comparison of PSNR for: BPSK *(a);* QPSK *(b);* 8-PSK *(c)*

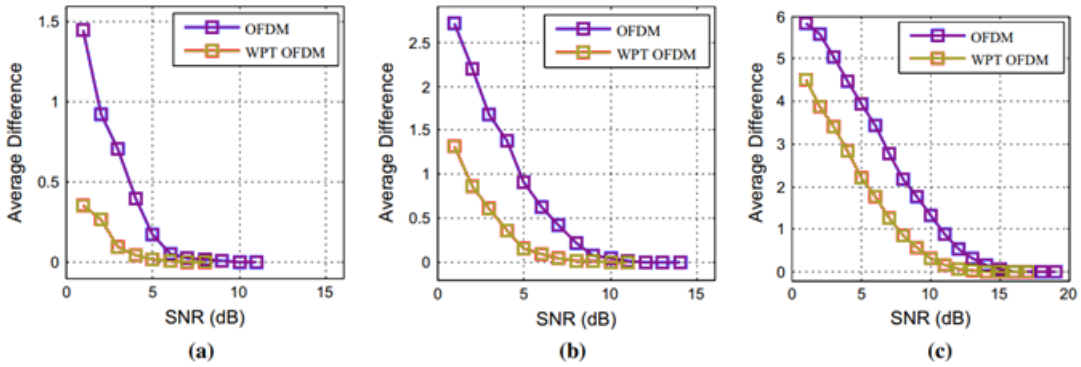

Fig. 7 – Comparison of AD for: BPSK *(a);* QPSK *(b);* 8-PSK *(c)*

11 Conclusions

This paper presents the technologies for generating signals that are already used in communication and telecommunication systems, and also provides an analysis of promising technologies that may be used in various new systems, including wireless broadband access communication systems. It is shown that the widely used OFDM modulation scheme has a number of shortcomings that can lead to a decrease in the performance indicators of the systems in which they are used, in particular: a reduction in the noise immunity of signal reception, due to distortions caused by multipathing when the electromagnetic field propagates between the base and mobile stations, as well as the effects of intersymbol and inter-channel interference; irrational, as compared with sequential waveforms, the use of transmitter power, which is associated with the use of a guard interval for protection against intersymbol interference and a high peak factor of a signal, etc. Alternative signal generation technologies are presented, in particular, signal generation technology based on window processing signals (W-OFDM) and providing a low level of out-of-band emission.

One of the main trends in the development of modern wireless broadband access communication systems is the rapid spread of technologies such as OFDM and MIMO. These technologies allow to achieve an increase in information efficiency in the conditions of multipath propagation and, as a result, to provide the ever-growing needs of wireless communication network users in high-speed connections and specific multimedia services. For some applications, ICS is a determining factor in their design and operation is the state of security of data processing and storage systems, which en-

sures the confidentiality, integrity and availability of information, as well as other properties of information and services: authenticity, observability, irrefutable and reliability. At the same time, the development of secure wireless communication systems that could reliably support multimedia applications faces a number of technological challenges that require serious research efforts. One of these calls is due to the choice of classes of discrete sequences, the properties of which largely determine the properties of physical data carriers in the ICS. In the work, on the basis of the analysis of the structure of the OFDM signal, the estimates of the security of the ICS from imposing false signals and messages by using non-linear discrete cryptographic signals as a physical data carrier are presented.

In this paper, a hidden communications system was proposed using WPT-OFDM. Hidden communications based on traditional OFDM are simpler in structure compared to WPT-OFDM, but the use of a protection bandwidth reduces bandwidth. The results showed that under equal conditions, the probability of an error in the transmitted data, the accuracy of extraction from the container and such quality parameters as PSNR, MSE and AD of the proposed system with wavelet transform are superior to traditional OFDM systems, and also allow increasing the system bandwidth.

References

[1] Gorbenko, I.D., Zamula, A.A., Semenko, Ye.A. Ensemble and correlation properties of cryptographic signals for telecommunication system and network applications // Telecommunications and Radio Engineering. – Vol. 75, 2016 Issue 2, pp. 169-178.
[2] I.D. Gorbenko, A.A. Zamula Cryptographic signals: requirements, methods of synthesis, properties, application in telecommunication systems Telecommunications and Radio Engineering. - Vol. 76, 2017 Issue 12, pp. 1079-1100.
[3] ITU-R, Recommendation M.2083-0, “IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond”, ITU recommendation, Sept. 2015.
[4] Pen Guan et al. “5G Field Trials: OFDM-Based Waveforms and Mixed Numerologies” IEEE Journal on Selected Areas in Communications, Vol. 35, No. 6, pp. 1234-1243, March 2017.
[5] T.S. Rappaport et al. “Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!”, IEEE Access, Vol. 1, pp. 335-349, 2013.
[6] J.G. Andrews et al. “What will 5G be?”, IEEE Journal on Selected Areas in Communications, Vol. 32, no. 6, pp. 1065-1082, June 2014.
[7] J. Abdoli, et al. “Filtered OFDM: A new waveform for future wireless systems”, Proc. IEEE SPAWC, pp. 66-70, Jun. 2015.
[8] X. Zhang et al. “Filtered-OFDM – Enabler for Flexible Waveformin The 5th Generation Cellular Networks”, Proc. IEEE GLOBECOM, pp. 1-6, Dec. 2015.
[9] Li. Jialing et al. “A resource block based filtered OFDM scheme and performance comparison”, Proc. IEEE ICT, pp. 1-5, May 2013.
[10] T. L. Marzetta “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas”, IEEE Transactions on Wireless Communications, Vol. 9, No. 11, pp. 3590-3600, Nov. 2010.
[11] China Mobile Research Institute, “C-RAN: The Road Towards Green RAN”, white paper, 2011. [On-line]. Available: http://labs.chinamobile.com/cran/
[12] H. Nikopour et al. “Sparse code multiple access”, Proc. IEEE PIMRC, pp. 332-336, Sept. 2013.
[13] 5G Forum. (2016, Mar.). 5G white paper: 5G vision, requirements, and enabling technologies [On-line]. Available: http://kani.or.kr/5g/whitepaper/5G%20Vision,%20Requirements,%20and%20Enabling%20Technologies.pdf.
[14] B. Farhang Boroujeny “Filter bank multicarrier modulation: a waveform candidate for 5G and beyond,”Advances in Electrical Engineering, vol. 2014, Dec. 2014. doi:10.1155/2014/482805.
[15] Zekeriyya Esat Ankaralı et al. “Enhanced OFDM for 5G RAN”, June 2017, doi: 10.3969/j. issn. 1673-5188. 2017. S1. 002.
[16] A. Şahin, I. Güvenç and H. Arslan “A survey on multicarrier communications: prototype filters, lattice structures, and implementation aspects, ”IEEE Communications Surveys & Tutorials, Vol. 16, No. 3, pp. 1312-1338, Aug. 2014. doi:10.1109/SURV.2013.121213.00263.
[17] Huawei and HiSilicon, “f-OFDM scheme and filter design,”3GPP Standard Contribution (R1-165425), Nanjing, China, May 2016.
[18] V.P. Fedosov, D.G. Kovtun, A.A. Legin, A.V. Lomakina Issledovanie modeli OFDM signala s malym urovnem vnepolosnogo izlucheniya / Izvestiya YuFU. Tekhnicheskie nauki. 2016, pp. 6-16.
[19] Harish Kumar Pal, Anand Kumar Singh PAPR Reduction technique using advanced peak windowing method of OFDM System. International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Vol.-3, Issue 2, May 2013.
[20] Bakulin, M. G., Kreindelin, V. B., Shloma, A. M., Shumov, A. P. Tekhnologiya OFDM. Uchebnoe posobie dlya vuzov. – M.: Goryachaya liniya – Telekom, 2015. - 360 p.
[21] Zamula O.A. Tekhnologii formirovaniya OFDM signalov v sovremennykh informatsionno-kommunikatsionnykh sistemakh Radiotekhnika: Vseukrainskii mezhvedomstvennyi nauchno – tekhnicheskii sbornik – 2018. – Vip. 193. pp. 152-159.
[22] Gorbenko, I.D., Gorbenko, Ju.I. Prykladna kryptologija. Teorija. Praktyka. Zastosuvannja: monografija / I.D. Gorbenko, Ju.I. Gorbenko. – Harkiv: Fort, 2012. – 880 p.
[23] Gorbenko, Ju.I. Metody pobuduvannja ta analizu, standartyzacija ta zastosuvannja kryptografichnyh system / Ju.I. Gorbenko. – Harkiv: Fort, 2016. – 959 p.
[24] I.D. Gorbenko, A.A. Zamula, V.L. Morozov Information security and noise immunity of telecommunication systems under conditions of various internal and external impacts // Telecommunications and Radio Engineering Vol. 76, 2017 Issue 19, pp 1705-1717.

[25] Zamula A.A., Morozov V.L. Informatsionnye tekhnologii peredachi dannykh v sovremennykh telekommunikatsionnykh sistemakh // Radiotekhnika: Vseukrainskii mezhvedomstvennyi nauchno – tekhnicheskii sbornik – 2016. – Vyp. 186. pp. 24-32.
[26] Simmons G. J. «Authentication theory coding theory» 1985.
[27] I. D. Gorbenko, A.A. Zamula, A. E. Semenko, V.L. Morozov Method for synthesis of performed signals systems based on cryptographic discrete sequences of symbols // Telecommunications and Radio Engineering Vol. 76, 2017 Issue 17, pp. 1523-1533.
[28] I.D. Gorbenko, A.A. Zamula, A.E. Semenko, V. L. Morozov Method for complex improvement of characteristics of orthogonal ensembles based on multiplicative combining of signals of different classes// Telecommunications and Radio Engineering Vol. 76, 2017 Issue 18, pp. 1581-1594.
[29] Gorbenko, I.D., Zamula, O.A. Modeli ta metody syntezu kryptografichnyh sygnaliv ta i'h optymizacija za kryterijem chasovoi' skladnosti // Matematychne ta komp'juterne modeljuvannja. Serija: Fizyko-matematychni nauky: zb. Nauk. prac' / Instytut kibernetyky imeni V.M. Glushkova Nacional'noi' akademii' nauk Ukrai'ny, 2017. Vyp. 15. 272 p.
[30] Sarvate D.V. Crosleration Properties of Pseudorandom and Related Sequences / D.V. Sarvate, M.V. Pursley // IEEE Trans. Commun, 1980. – Vol. Com 68 – pp. 59–90.
[31] Sverdlik, M. B. Optimal'nye diskretnye signaly. / Sverdlik M. B. M: Radio i Svyaz', 1975. – 200 p.
[32] Tew, Y. & Wong, K. (2014). An overview of data hiding in H.264/AVC compressed video. IEEE Transactions on Circuits and Systems for Video Technology, 24(2), 305–319.
[33] Sharma, V. & Singh, G. (2014). On BER assessment of conventional- and wavelet-OFDM over AWGN channel. Jounal of Optik, 125, 6071–6073.
[34] Gupta, M. K. & Tiwari, S. (2013). Performance evaluation of conventional and wavelet based OFDM system. International Journal of Electronics and Communications, 67(4), 348–354.
[35] Dutta, A. et al. (2013). Secret agent radio: Covert communication through dirty constellations. In M. Kirchner & D. Ghosal (Eds.), Information Hiding. IH 2012. Lecture notes in computer science (Vol. 7692, pp. 160-175). Berlin, Heidelberg: Springer.
[36] Bash, B. A., Goeckel, D., Towsley, D. & Guha, S. (2015). Hiding information in noise: Fundamental limits of covert wireless communication. IEEE Communications Magazine, 53(12), 26–31.
[37] Bloch, M. (2016). Covert communication over noisy memoryless channels: A resolvability perspective. IEEE Transactions on Information Theory, 62(5), 2334–2354.
[38] Bouhlel A., Sakly A. & Mansouri N. (2015). Performance comparison of DWT based MIMO OFDM and FFT based MIMO OFDM. Procedia Computer Science, 73, 266–273.
[39] Khan, A. R. & Gulhane, S. M. (2017). A highly sustainable multi-band orthogonal wavelet code division multiplexing UWB communication system for underground mine channel. Digital Communications and Networks, 3, 1–13.
[40] Kaur, H., Kumar, M., Sharma, A. K. & Singh, H. P. (2016). Performance analysis of DWT based OFDM over fading environments for mobile WiMax. Jounal of Optik, 127, 544–547.
[41] Mushtaq, A. S., Ihsan, A. A. & Qasim, N. (2015). 2D-DWT vs. FFT OFDM systems in fading AWGN channels. Radioelectronics and Communication Systems, 58(5), 228–233.
[42] Mahapatra, C., Leung, V. C. M. & Stouraitis, T. (2017). An orthogonal wavelet division multiple-access processor architecture for LTE-advanced wireless/radio-over-fiber systems over heterogeneous networks. EURASIP Journal on Advances in Signal Processing, 1, 1–16.
[43] Villalobos, S., Aldana, F., Ladino, I., Diaz, I. (2017). A waveletbased OFDM system implementation on GNURadio platform vs. an FFT-based. In Springer workshop on engineering applications (WEA-2017), Vol. 742, 201–211.
[44] Kumar, S. & Sharma, S. (2013). Performance evaluation physical layer of IEEE 802.11 standards under Fourier transform and wavelet transform over Rician fading channel. In IEEE, 2nd international conference on information management in the knowledge economy, pp. 69–74.
[45] Kumar, S., Singh, A. & Kumar, M. Wireless Netw (2018). https://doi.org/10.1007/s11276-018-1775-3
[46] Kansal, L., Sharma, V. & Singh, J. (2017). BER assessment of FFT-OFDM against WHT-OFDM over different fading channel. Wireless Networks, 23(7), 2189–2196.

Рецензент: Олександр Потій, д.т.н., проф., Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна.
E-mail: potav@ua.fm

Надійшло: Травень 2018.

Автори:
Олександр Замула, д.т.н., доцент, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна.
E-mail: zamulaaa@gmail.com

Владислав Морозов, аспірант, Харківський національний університет імені В.Н. Каразіна, м. Харків, Україна.
E-mail: ilissar@hotmail.com

Сербін Вадим, провідний фахівець, ДП «Конструкторське бюро «Південне», м. Дніпро, Україна.
E-mail: buba75@i.ua

Принципи формування, обробки та свойства OFDM сингалів.

Анотація. У статті розглянуті технології формування сигналів, використовуваних в системах мобільного зв'язку та інформаційно-телекомунікаційних системах, а також наводиться аналіз перспективних технологій, які можуть знайти застосування в бездротових системах зв'язку широкосмугового доступу. Показано, що широко використовувана схема модуляції з ортогональним частотним розділенням (OFDM) має низку недоліків, які можуть призвести до зниження показників ефективності систем. Представлені альтернативні технології формування сигналів, зокрема, технологія, заснована на віконної обробці сигналів (W-OFDM), технологія, заснована на тимчасовому поділі (w-OFDM); технологія UFMC і інші, що дозволяють усунути недоліки технології OFDM. Пропонуються нові погляди на використання в технології передачі з багатьма несучими в формі мультиплексування з ортогональним частотним розділенням (з метою підвищення захищеності сучасних бездротових систем зв'язку широкосмугового доступу від впливу зовнішніх і внутрішніх загроз), класу нелінійних дискретних крип-

тографічних послідовностей для освіти фізичного переносника даних - сигналів. Показано, що застосування таких сигналів дозволить поліпшити показники захищеності зазначених систем від введення (нав'язування) в систему помилкових повідомлень, фальсифікації повідомлень, а також показники забезпечення цілісності та конфіденційності даних, завадостійкості прийому і скритності функціонування системи.

Ключові слова: перешкодозахищеність; інформаційна безпека; широкосмуговий доступ; сигнал, цілісність; стійкість; частотний поділ; стільниковий зв'язок; інтерференція; пік-фактор.

Рецензент: Александр Потий, д.т.н., проф., Харьковский национальный университет имени В.Н. Каразина, г. Харьков, Украина.
E-mail: potav@ua.fm

Поступила: Май 2018.

Авторы:
Александр Замула, д.т.н., доцент, Харьковский национальный университет имени В.Н. Каразина, г. Харьков, Украина.
E-mail: zamulaaa@gmail.com

Владислав Морозов, аспирант, Харьковский национальный университет имени В.Н. Каразина, г. Харьков, Украина.
E-mail: ilissar@hotmail.com

Вадим Сербин, ведущий специалист, ГП «КБ «Южное», г.Днипро, Украина.
E-mail: buba75@i.ua

Принципы формирования, обработки и свойства OFDM сигналов.

Аннотация. В статье рассмотрены технологии формирования сигналов, используемых в системах мобильной связи и информационно-телекоммуникационных системах, а также приводится анализ перспективных технологий, которые могут найти применение в беспроводных системах связи широкополосного доступа. Показано, что широко используемая схема модуляции с ортогональным частотным разделением (OFDM) обладает рядом недостатков, которые могут привести к снижению показателей эффективности систем. Представлены альтернативные технологии формирования сигналов, в частности, технология, основанная на оконной обработке сигналов (W-OFDM), технология, основанная на временном разделении (w-OFDM); технология UFMC и другие, позволяющие устранить недостатки технологии OFDM. Предлагаются новые взгляды на использование в технологии передачи со многими несущими в форме мультиплексирования с ортогональным частотным разделением (в целях повышения защищенности современных беспроводных систем связи широкополосного доступа от воздействия внешних и внутренних угроз), класса нелинейных дискретных криптографических последовательностей для образования физического переносчика данных – сигналов. Показано, что применение таких сигналов позволит улучшить показатели защищенности указанных систем от ввода (навязывания) в систему ложных сообщений, фальсификации сообщений, а также показатели обеспечения целостности и конфиденциальности данных, помехоустойчивости приема и скрытности функционирования системы.

Ключевые слова: помехозащищенность; информационная безопасность; широкополосный доступ; сигнал; целостность; помехоустойчивость; частотное разделение; сотовая связь; интерференция; пик-фактор.

Статті пройшли внутрішнє та зовнішнє рецензування.

Наукове видання

КОМП'ЮТЕРНІ НАУКИ ТА КІБЕРБЕЗПЕКА

Випуск 2(10) 2018

Міжнародний електронний науково-теоретичний журнал

Англійською, українською, російською мовами

Комп'ютерне верстання – Федоренко В.В., Єсіна М.В.

61022, Харків, майдан Свободи, 6
Харківський національний університет імені В.Н. Каразіна

V. N. Karazin Kharkiv National University Publishing

CS&CS