

УДК 004

Д.Скрипник, магістр гр. КІ-22МЗ,
Центральноукраїнський національний технічний університет

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ СИСТЕМИ СТРУКТУРНОГО КОДУВАННЯ ДАНИХ У КОМП'ЮТЕРНИХ МЕРЕЖАХ АВТОМАТИЗОВАНИХ СИСТЕМ УПРАВЛІННЯ

У статті розроблено програмне забезпечення, яке призначено для системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Метою розробки є дослідження та програмна реалізація системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Об'єктом дослідження є процес структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Предметом дослідження є методи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Методи дослідження базуються на методах теорії кодування, методах математичної статистики, методах розробки програмного забезпечення. Результат роботи – програмна реалізація системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. В процесі роботи над програмною моделлю виконано аналіз існуючих апаратних та програмних засобів. В повній мірі описані всі компоненти розробленого програмного забезпечення.

Постановка проблеми. Стиснення даних має дуже широке застосування в різних аспектах життя, таких як медицина [1], мультимедіа та обробка зображень [2-4], бездротова сенсорна мережа (WSN), яка є дуже цікавою мережевою технологією [5], а також комп'ютерна мережа, математичні обчислювальні методи історично перевершували вдосконалення топології та з'єднань мережі. Кількість балів на обчислювальний вузол зросла, пропускна здатність мережі, доступна для обчислювального вузла, зросла, однак недостатньо швидко, щоб не відставати від зростаючої обчислювальної продуктивності вузла. Очікується, що такий розрив збільшиться при вході в обчислювальну сферу з суперкомп'ютерами, які містять сотні обчислювальних ядер у кожному обчислювальному вузлі. У цьому новому полі пропускна здатність з'єднання між вузлами та різними файловими системами перевищить преміум-клас. Стиснення даних перед надсиланням їх на вузли введення/виведення призводить до зменшення навантаження на мережу. Навіть у випадку, коли вузли вводу/виводу вимагають декомпресії даних перед надсиланням їх у зовнішню мережу, стиснення все одно може бути корисним через те, що деякі підходи до стиснення є асиметричними відповідно до пропускної здатності, декомпресія зазвичай відбувається на порядок швидше, ніж стиснення, гарантуючи той факт, що декомпресія даних з використанням обмеженої кількості вузлів введення/виведення не повертається у вузьке місце [6]. Було визначено два методи стиснення, які включають без втрат або з втратами. Стиснення без втрат призводить до зменшення кількості бітів через ідентифікацію та усунення статистичної надлишковості, і нічого не втрачено в інформації. Другий підхід – це стиснення з втратами, яке зменшує кількість бітів шляхом видалення менш важливої або непотрібної інформації [7-9]. Зазвичай будь-який пристрій, що використовує стиснення даних, відомий як кодер, який може використовувати багато різних методів, навіть нейронну мережу [10] на одному або кількох рівнях [11]. Компресія хороша тим, що вона зменшує ресурси, необхідні для зберігання та передачі даних. Деякі з обчислювальних джерел використовуються в процесах стиснення та розпакування, які вважаються одним із компромісів стиснення. Конструкція системи стиснення даних має певні компроміси між різними факторами, які включають ступінь стиснення, ступінь викривлення, який було

введено, особливо у випадку використання стиснення даних із втратою даних, і обчислювальні джерела, необхідні для стиснення та розпакування даних [12]. Стиснення та кодування даних вважаються хорошим інструментом для підвищення продуктивності мережі, оскільки це передбачає надсилання або зберігання меншої кількості бітів, що потребує менше часу для надсилання даних [13].

Таким чином, магістерська робота присвячена підвищенню ефективності функціонування АСУ на основі новітніх методів стиснення інформаційних потоків.

Аналіз останніх досліджень і публікацій. При аналізі останніх досліджень і публікацій [1-20] було виявлено певні прогалини у забезпеченні системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.

Мета й завдання дослідження. Метою роботи є дослідження та програмна реалізація системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.

Для досягнення поставленої мети визначена програма дослідження, що складається з наступних завдань:

- Огляд існуючих систем структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.
- Дослідження системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.
- Програмна реалізація системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.

Об'єктом дослідження є процес структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.

Предметом дослідження є методи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління.

Методи дослідження базуються на методах теорії кодування, методах математичної статистики, методах розробки програмного забезпечення.

Виклад основного матеріалу. У цьому розділі аналізується вплив стиснення даних на продуктивність мережі. На першому етапі файл даних стискається за допомогою кодування DCT і Хаффмана, а комп'ютерні програми пишуться за допомогою програми `mat lab`. Дві програми написані для виконання етапів стиснення кожного методу окремо та отримання результату стиснення. Також надайте інформацію та подробиці про основні ідеї методів стиснення з втратами та без втрат, їх подробиці та математичну формулу для Хаффмана як методу без втрат і DCT із втратами, а також ілюструйте основні показники, що використовуються для вимірювання продуктивності мережі.

Стиснення даних

Стиснення даних із втратами та без втрат стиснення широко використовувалося в бездротових мережах, таких як мережі 3-G, і для оптимізації веб-сайтів для скорочення наскрізного часу передачі [14-20]. У цих ситуаціях стиснення призвело до зменшення затримки передачі та покращення часу відповіді. Стиснення призводить до покращення наскрізного часу передачі в мережі, пропускну здатності мережі та, відповідно, пропускну здатності файлу введення/виведення. Було проведено широкий спектр досліджень ефектів стиснення для мінімізації споживання енергії. Існує 2 класи стиснення [11-15].

Стиснення без втрат

Це залежить від зменшення кількості бітів через ідентифікацію та усунення статичної надмірності. Кількість інформаційних бітів однакова до та після стиснення, кодування Лемпеля-Зіва, Хаффмана та кодування довжини серії є методами стиснення без втрат [26]. Метод кодування Хаффмана Код Хаффмана було зіставлено із символами фіксованої довжини на коди різної довжини. Алгоритм Хаффмана був заснований на списку всіх даних або символів, які були впорядковані в порядку спадання ймовірностей. На основі цих ймовірностей було призначено кодові слова. Довші кодові слова були призначені для менших ймовірностей, а коротші кодові слова для вищих. Після цього він створює бінарне

дерево методом «знизу вгору» з символом на кожному листі. У якому є кілька кроків, на кожному з яких вибрано 2 символи з мінімальними значеннями частоти, а потім додано до вершини часткового дерева. Вибрані символи мінімальної частоти були вилучені зі списку, замінені другорядним символом, який позначає 2 оригінальні символи. Ось чому список було зменшено до одного вторинного символу, що вказує на те, що дерево завершено. Нарешті, кожному листку або символу призначається кодове слово відповідно до шляху від кореневого вузла до символів у цьому списку. Нижче наведено кроки логарифма Хаффмана, які використовуються для стиснення файлу даних, який містить зображення:

1. Упорядкування можливостей частоти символів для форми завантаження та розгляд її як кінцевої серії у формі дерева.

2. Повторіть цей процес у випадку, якщо на діаграмі дерева є додаткова одна нота.

А. кожен 2 ноти з нижчою ймовірністю повторення частоти, що дорівнює кількості повторень ноти в результаті.

В. Кодування кожної пари гілок деревоподібної діаграми у подвійній формулі. Перевагою цього підходу є стабільність довжини коду, а також те, що між кодами немає інтервалів, що призводить до неоднозначності під час відкриття коду та повернення файлу до вихідного стану, на малюнку 1 показано кроки кодування Хаффмана [17, 18].

Стиснення з втратами

При стисненні з втратами числовий біт зменшується шляхом ідентифікації необхідної інформації та її видалення. Цей метод стиснення вимагає певної втрати даних, які неможливо відновити до початкових даних. Це вказує на те, що спотворення даних досить низьке. Менше відволікання призводить до кращих результатів, і серед численних логарифмів є дискретне косинусне перетворення (DCT), яке виражає кінцеву послідовність точок даних на основі підсумовування косинусних функцій, які коливаються на різних частотах. DCT, був визначений як один із широко використовуваних методів перетворення в стисненні даних і обробці сигналу. Він використовується в більшості цифрових засобів масової інформації, включаючи цифрові зображення (наприклад, HEIF і JPEG, у яких невеликі компоненти високої частоти можуть не враховуватися), цифрове аудіо (наприклад, MP3, Dolby Digital і AAC), цифрове відео (наприклад, H.26x і MPEG), цифрове радіо (наприклад, DAB+ і AAC+), цифрове телебачення (наприклад, HDTV, VOD і SDTV) і мова. кодування (наприклад, Siren, Opus і AACLD). DCT схожий на DFT: він виконує перетворення зображення або сигналу.

Параметри продуктивності мережі

Поняття продуктивності мережі пов'язане з послугами, які вона надає користувачам. Наприклад, мережа доставки контенту (CDN) надає послуги для доставки контенту користувачам замість постачальника контенту. Поняття продуктивності CDN залежить від типу вмісту, який він надає. Для CDN, яка надає користувачам сайт електронної комерції, зазвичай ефективність вимірюється на основі часу завантаження сторінки, а також доступності сторінки. Доступність є мірою відсотка часу, протягом якого користувач здатний завантажити сторінку веб-сайту без збоїв, а час завантаження є мірою швидкості, з якою сторінка була завантажена та відтворена браузером [11, 12]. Продуктивність можна виміряти багатьма матрицями, такими як затримка (затримка), пропускна здатність, тремтіння, втрата пакетів і пропускна здатність.

Пропускна здатність

Пропускна здатність (також звану швидкістю передачі даних) можна визначити як здатність мережі переміщувати обсяг даних по черзі протягом одиниці часу, вона представляє швидкість, з якою мережа передає трафік програми.

Затримка

Затримку можна визначити як час, необхідний для передачі пакету від пристрою-джерела до пристрою призначення. Її також можна визначити як часову затримку для проходження через мережу, яка включає в себе всі проміжні вузли від джерела до місця призначення. Для потоку пакетів, джиттер може бути визначений як міра затримки пакетів від перевантаження мережі в проміжних вузлах або з'єднаннях, які не мають необхідної

пропускної здатності, є фактичним показником швидкості передачі даних через мережу. Незважаючи на те, що спочатку пропускна здатність у бітах на секунду та пропускна здатність здаються ідентичними, однак це не так. Ширина смуги може бути визначена як потенційне вимірювання зв'язку, а пропускна здатність представляє фактичне вимірювання швидкості передачі даних [10, 11].

Встановлено, що суть кодування полягає у формуванні коду-номера $N(m, \Lambda, \mathcal{G})_j$ всієї двійкової послідовності із заданим значенням структурної ознаки. Визначено, що структурними ознаками є: вектор \mathcal{S} заборон появи на певній позиції одиничного елемента ($\overline{\mathcal{S}} = \{s_i\}, i = \overline{1, m}$; s_i – ознака заборони появи на i -й позиції одиничного елемента; якщо $s_i = 0$, те на i -й позиції заборонена поява одиниці й навпаки); кількість серій одиниць \mathcal{G} у двійковій послідовності. Кодування з урахуванням заборон на позиції одиниць відповідає структурному представленню двійкових даних. Двійкова послідовність чисел розглядається, як структурне число елементи якого набувають значення $\{0; 1\}$.

За результатами досліджень, отримано систему виразів одноозначового структурного кодування двійкових даних у структурному просторі, що забезпечує формування коду-номера відповідно двійковому структурному числу, елементи якого задовольняють заданій структурній ознаці (система виразів, що визначає номер заданої двійкової послідовності в безлічі одноозначових структурних чисел).

Встановлено, що безлічі двійкових послідовностей у структурному просторі за кількістю серій одиниць $V(m, \Lambda, \mathcal{G})$ відповідає пронумерована множина двійкових послідовностей (1, 2), що містять задану кількість серій одиниць та не проходять через позиції із заборонами одиниць.

$$V(m, \Lambda, \mathcal{G}) = \sum_{k=1}^K V(\Theta^{(k)}) = \sum_{k=1}^K \prod_{z=1}^Z V(\mathcal{G}_z^{(k)}, \Theta^{(k)}) \quad (1)$$

$$V(\mathcal{G}_z^{(k)}, \Theta^{(k)}) = \binom{m_z + 1}{2\mathcal{G}_z^{(k)}} = \frac{(m_z + 1)!}{(2\mathcal{G}_z^{(k)})! (m_z + 1 - 2\mathcal{G}_z^{(k)})!}, \quad (2)$$

де

$\mathcal{G}_z^{(k)}$ – значення числа серій одиниць для \mathbf{Z} -ї припустимої зони двійкової послідовності A ;

$\Theta^{(k)}$ – вектор, елементами якого є k -та комбінація кількостей серій одиниць $\mathcal{G}_z^{(k)}$ у припустимих зонах $\Theta^{(k)} = \{\mathcal{G}_1^{(k)}, \dots, \mathcal{G}_z^{(k)}, \dots, \mathcal{G}_Z^{(k)}\}$, $k = \overline{1, K}$;

Z – кількість припустимих зон у двійковій послідовності;

m_z – кількість двійкових елементів в z -й припустимій зоні;

$V(\mathcal{G}_z^{(k)}, \Theta^{(k)})$ – кількість припустимих двійкових послідовностей, отриманих для z -ї припустимої зони за кількістю серій одиниць, рівній $\mathcal{G}_z^{(k)}$ для вектора $\Theta^{(k)}$.

Показано, що величина $V(\mathcal{G}_z^{(k)}, \Theta^{(k)})$ дорівнює кількості двійкових підпослідовностей, що утворюють відповідну z -у зону, у яких число серій одиниць дорівнює $\mathcal{G}_z^{(k)}$. У зв'язку із цим z -м елементом безлічі $\Psi(\Theta^{(k)})$ є величина коду-номера $N(\mathcal{G}_z^{(k)}, \Theta^{(k)})$, для якого виконується нерівність

$$N(\mathcal{G}_z^{(k)}, \Theta^{(k)}) \leq V(\mathcal{G}_z^{(k)}, \Theta^{(k)}) - 1. \quad (3)$$

Визначено, що кодування двійкових структурних чисел за кількістю серій одиниць є система виразів, що дозволяє визначити код-номер $N(m, \Lambda, \mathcal{G})_j$ оброблюваної двійкової

послідовності $A^{(j)} = \{a_{ij}\}_{i=1, \overline{m}}$ з виявленими обмеженнями на позиції одиниць $\Lambda = \{\lambda_i\}_{i=1, \overline{m}}$ та на кількість серій одиниць \mathcal{G} у безлічі двійкових структурних чисел, що містять задану кількість серій одиниць.

Встановлено, що jednoznakove подання двійкових даних у структурному просторі є взаємооднозначним, а саме: за заданим кодом номером $N(m, \Lambda, \mathcal{G})_j$ можна відновити тільки одну вихідну двійкову послідовність $A^{(j)} = \{a_{ij}\}_{i=1, \overline{m}}$ відповідну даній кодовій конструкції без внесення похибки, тобто

$$N(m, \Lambda, \mathcal{G})_j = \varphi_k(A^{(j)}) ; A^{(j)*} = \varphi_d(N(m, \Lambda, \mathcal{G})_j) \quad (4)$$

де

$A^{(j)} = \{a_{izj}\}_{i=1, \overline{m}}$, $A^{(j)*} = \{a_{izj}^*\}_{i=1, \overline{m}}$, $a_{izj}^* = a_{izj}$, $i = \overline{1, m}$, $z = \overline{1, Z}$, φ_k , φ_d – відповідно оператори кодування й декодування двійкових структурних чисел з урахуванням обмеженої кількості серій одиниць;

$A^{(j)}$, $A^{(j)*}$ – відповідно вихідна й відновлена двійкові послідовності;

a_{izj} , a_{izj}^* – i, z, j -е елементи відповідно до вихідної і відновленої двійкових послідовностей, що належать Z -й припустимій зоні.

Розроблено систему правил та алгоритмів відновлення двійкових структурних чисел з обмеженою кількістю серій одиниць.

Встановлено, що вихідну послідовність $A^{(j)} = \{a_{izj}\}_{i=1, \overline{m}}$, яка задовольняє обмеженням:

$$0 \leq a_{ij} \leq s_i, s_i = \max_{1 \leq j \leq n} \{a_{ij}\}, a_{ij}; s_i \in \{0; 1\}; \mathcal{G} = \sum_{z=1}^Z \mathcal{G}_z^{(k)} ; 0 \leq \mathcal{G}_z^{(k)} \leq \min\{\mathcal{G}; [\frac{m_z + 1}{2}]\}$$

можна відновити без внесення похибки на основі значень коду-номера $N(m, \Lambda, \mathcal{G})_j$,

обмежень на позиції одиниць $\Lambda = \{\lambda_i\}_{i=1, \overline{m}}$ і на кількість серій одиниць \mathcal{G} , тобто виконується умова $A_\alpha^{(u)} = A_\alpha^{(j)}$, де $A_\alpha^{(u)}$ і $A_\alpha^{(j)}$ вихідна й відновлена двійкові послідовності, $\alpha = \overline{1, \zeta}$.

Встановлено, що мінімальне значення коефіцієнта стиснення k_{\min} у результаті jednoznakove кодування двійкових даних у структурному просторі дорівнює:

$$k_{\min} = \frac{\log_2 m}{\log_2 V(m, \Lambda, \mathcal{G})} \quad (5)$$

де $V(m, \Lambda, \mathcal{G})$ – кількість двійкових чисел довжиною m елементів з кількістю серій одиниць \mathcal{G} у структурному просторі із заданим вектором обмежень Λ .

Встановлено, що величина k_{\min} є верхньою межею для значення коефіцієнта стиснення k_{cm} :

$$k_{cm} = \frac{\log_2 m}{\log_2 N(m, \Lambda, \mathcal{G})} < k_{\min} \quad (6)$$

За рахунок розробленого кодування забезпечується додаткове підвищення ступеня стиснення щодо структурного кодування й структурного кодування за кількістю серій при цьому, виконуються відповідні нерівності:

$$\log_2 V(m, \Lambda, \mathcal{G}) \leq \log_2 V(m, \Lambda) \quad (7)$$

$$\log_2 V(m, \Lambda, \mathcal{G}) \leq \log_2 V(m, \mathcal{G}) \quad (8)$$

де $\log_2 V(m, \Lambda)$ й $\log_2 V(m, \mathcal{G})$ – кількість розрядів, що відведено для подання відповідно структурного й структурного числа з обмеженою кількістю серій одиниць та складені із m двійкових елементів кожне.

На підставі математичного моделювання й проведеного аналізу побудовані графіки залежностей величини мінімального коефіцієнта стиснення k_{cm} від m_z й $\mathcal{G}_z^{(k)}$ (рисунок 1).

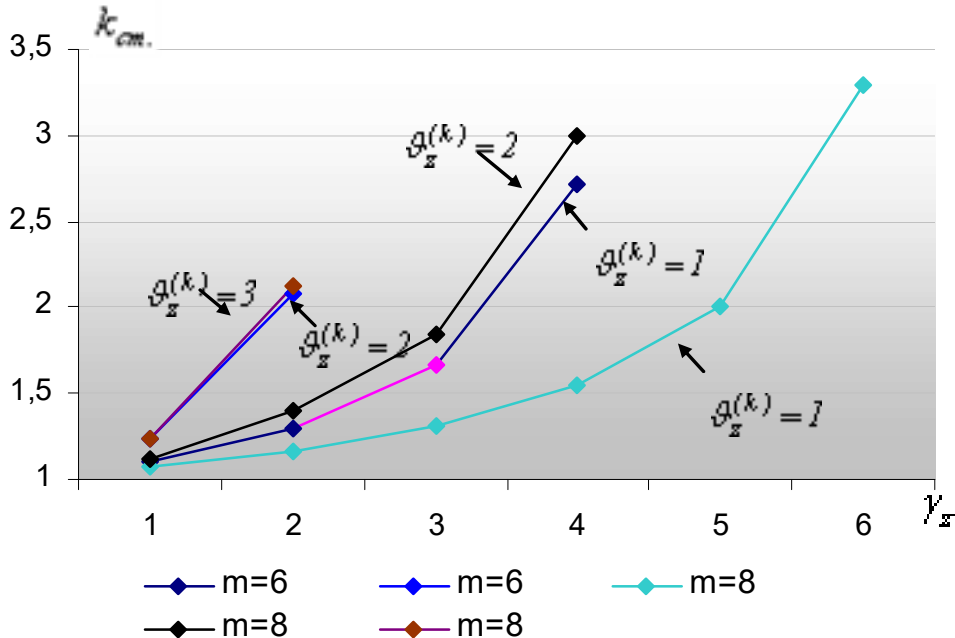


Рисунок 1 – Графіки залежності коефіцієнта стиснення k_{cm} від довжини двійкової послідовності m_z , числа серій одиниць $\mathcal{G}_z^{(k)}$ і γ_z - кількість заборонених позицій одиниць в z -й зоні

З аналізу отриманих результатів впливає таке.

По-перше, максимальне значення додаткового коефіцієнта стиснення k_{cm} забезпечується зі збільшенням кількості позицій із забороненою появою одиниць.

Для $m_z=8$ й $\mathcal{G}_z^{(k)}=1$ при переході від $\gamma_z=1$ до $\gamma_z=6$ величина k_{cm} збільшується в 1 рази, для $m_z=8$ й $\mathcal{G}_z^{(k)}=2$ при переході від $\gamma_z=1$ до $\gamma_z=4$ величина k_{cm} збільшується в 2,67 рази, а для $m_z=6$ й $\mathcal{G}_z^{(k)}=1$ при переході від $\gamma_z=1$ до $\gamma_z=4$ величина k_{cm} збільшується в 2,47 рази.

По-друге, для фіксованих значень кількості заборонених позицій одиниць в z -й зоні найбільші значення величини k_{cm} досягаються із збільшенням кількості серій одиниць.

Для $m_z=8$ й $\gamma_z=2$ зі збільшенням кількості серій одиниць від $\mathcal{G}_z^{(k)}=1$ до $\mathcal{G}_z^{(k)}=3$ величина k_{cm} збільшується в 1,98 рази. Аналіз отриманих результатів показує, що величина k_{cm} для $\mathcal{G}_z^{(k)}=1$, $\gamma_z=1$ збільшується в 1,15 разів при зменшенні довжини припустимої зони від $m_z=8$ до $m_z=4$ (рисунок 2).

Отже, за результатами досліджень сформовано метод однознакового структурного кодування двійкових даних по кількості серій одиниць.

Встановлено, що розроблена система правил є взаємодозначною для процедур кодування і декодування інформаційних потоків в АСУ.

Встановлено, що додатковий облік обмежень на кількість серій одиничних елементів у двійкових послідовностях чисел забезпечує збільшення ступеня стиснення даних в АСУ без втрат інформації.

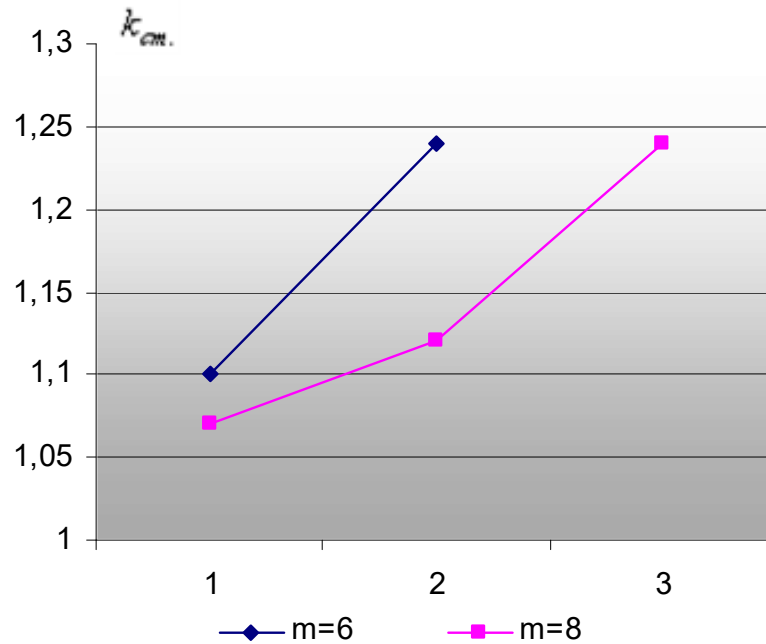


Рисунок 2. – Графіки залежності коефіцієнта стиснення k_{cm} від m_z і $g_z^{(k)} = 1$ для $\gamma_z = 1$

Розробка структурної схеми

У даному розділі наведено розробку методів двоозначового структурного кодування двійкових даних з метою додаткового збільшення ступеня стиснення інформації в АСУ.

Структурними ознаками для двійкових послідовностей виступають: кількість серій одиниць у кожній припустимій зоні й межі припустимих зон.

Встановлено, що додаткове збільшення ступеня стиснення двійкових даних без внесення похибки в АСУ досягається в результаті обліку припустимих комбінацій тільки для одного вектора обмежень на число серій одиниць у припустимих зонах.

Виведемо аналітичні вирази, що дозволяють формувати код-номер для двійкових послідовностей, що задовольняють одночасно двом структурним ознакам.

Встановимо, що двоозначове подання двійкових структурних чисел є взаємодозначним, тобто відновлення вихідних даних здійснюється без внесення похибки в інформаційний потік АСУ.

Розробимо метод відновлення двійкових даних на основі двоозначового структурного декодування. Покажемо, що двійкові послідовності, що задовольняють системі структурних обмежень на число серій одиниць у припустимих зонах є двоозначовими двійковими структурними числами.

Множина, складена із двійкових послідовностей, що одночасно задовольняє структурним обмеженням:

- на позиції із припустимою появою одиничних елементів, що задається вектором обмежень на діапазон значень оброблюваних елементів Λ ;

- сумарне число серій одиниць \mathcal{G} у всій оброблюваній послідовності (перша структурна ознака);

– число серій одиниць $g_{z,j}$ у кожній припустимій зоні (друга структурна ознака – задається вектором $\Theta^{(k)}$ значень величин $g_z^{(k)}$) – визначена, як множина двоознакових двійкових структурних чисел.

Встановлено, що двоознаковою структурною нумерацією даних у двійковому структурному просторі є процес обчислення порядкового номера відповідно до двоознакового двійкового структурного числа в припустимій безлічі, що складається з двійкових послідовностей задовольняючій системі обмежень:

$$\begin{cases} 0 \leq a_{ij} \leq s_i, i = \overline{1, m}; \\ g = \sum_{z=1}^Z g_z^{(k)}; \\ g_z^{(k)} = g_{z,j}, z = \overline{1, Z}, \end{cases} \quad (9)$$

де

s_i – обмеження на позиції, що допускають появу одиничних елементів;

$g_z^{(k)}$ – припустиме значення обмеження на число серій одиниць в Z -й зоні;

$g_{z,j}$ – значення числа серій одиниць в Z -й зоні, обчислене для конкретної оброблюваної J -ї двійкової послідовності.

На основі аналізу встановлено, що значення коду-номера $N(m, \Lambda, \Theta^{(k)})_j$ двоознакового двійкового структурного числа буде меншим у порівнянні зі значенням коду-номера $N(m, \Lambda, g)_j$ одноознакового структурного числа $N(m, \Lambda, \Theta^{(k)})_j \leq N(m, \Lambda, g)_j$.

Встановимо взаємооднозначність двоознакового структурного подання двійкових даних у структурному просторі. Для обраного лексикографічного правила нумерації, заданих обмежень на число серій одиниць у кожній припустимій зоні й обмежень на розташування одиничних елементів у двійковій послідовності $A(m, \Theta^{(k)})_j$, можна сформувати тільки один код-номер $N(m, \Lambda, g)_j$ відповідно до даної послідовності.

Визначено, що для двійкової послідовності $A(m, \Theta^{(k)})$ розглянутої, як двоознакове структурне число із заданими структурними параметрами й обмеженнями, що накладені на них (довжина послідовності m , обмеження Λ на позиції із припустимою появою одиничних елементів та обмеження $\Theta^{(k)}$ на число серій одиниць у припустимих зонах) – можна сформувати тільки один код-номер $N(m, \Lambda, \Theta^{(k)})_j$ відповідно до цієї послідовності.

Розроблено систему правил відновлення двійкових двоознакових структурних чисел у вигляді структурної схеми системи (рисунок 3). Двійкову послідовність $A(m, \Theta^{(k)})_j = \{a_{izj}\}_{i=\overline{1, m}}$, що задовольняє системі обмежень можна відновити без внесення

похибки на основі значень коду-номера $N(m, \Lambda, \Theta^{(k)})_j$, з урахуванням відомих значень величин: довжини послідовності m , вектора обмежень на позиції одиниць $\Lambda = \{\lambda_i\}_{i=\overline{1, m}}$ і

вектора $\Theta^{(k)}$ обмежень на число серій одиниць у припустимих зонах. На основі визначених переваг двоознакового відновлення відносно одноознакового поелементного відновлення встановлено, що є можливість здійснювати відновлення елементів Z -ї припустимої зони незалежно від процесу відновлення попередніх $(z-1)$ зон.

Таким чином, для додаткового підвищення ступеня компактного подання й зниження часу на кодування необхідно на однознакові структурні двійкові послідовності накладати додаткове обмеження на комбінацію числа серій одиниць у припустимих зонах.



Рисунок 3 – Структурна схема системи

Встановлено, що: верхньою межею обсягу безлічі двоозначових чисел є обсяг безлічі однознакових структурних чисел; кількість розрядів, що відведено на подання коду-номера двоозначового двійкового структурного числа не буде перевищувати кількість розрядів, що витрачаються на подання коду-номера однознакового двійкового числа в структурному просторі.

Розробимо методіку оцінки ступеня компактного подання двоозначових структурних чисел. Здійснимо порівняльну оцінку характеристик розробленого компактного подання даних з відомими методами. На основі досліджень та математичного моделювання побудуємо графіки залежності величин мінімальних ступенів стиснення.

На основі проведених досліджень встановлено наступне.

По-перше, при фіксованих параметрах $m=8$, $Z=2$, $\mathcal{G}=2$ значення мінімального ступеня стиснення для двоозначового подання збільшується в середньому на 30% при збільшенні кількості заборонених двійкових елементів в інформаційних потоках АСУ (рисунок 4).

Коефіцієнти $k(m, \Lambda)_{\min}$, $k(m, \mathcal{G})_{\min}$, $k(m, \mathcal{G}, \Lambda)_{\min}$, $k(m, \Lambda, \Theta^{(k)})_{\min}$ показані на рисунку 4, характеризують стиснення інформаційного потоку на основі структурного кодування, кодування з урахуванням обмеженого числа серій одиниць, однознакового й двоозначового кодування у двійковому структурному просторі відповідно. Дані коефіцієнти є функції від параметрів: кодування довжини двійкової послідовності m , числа серій одиниць \mathcal{G} , кількості припустимих зон Z , кількості заборонених двійкових елементів γ і вектора обмежень $\Theta^{(k)}$ на число серій одиниць у припустимих зонах.

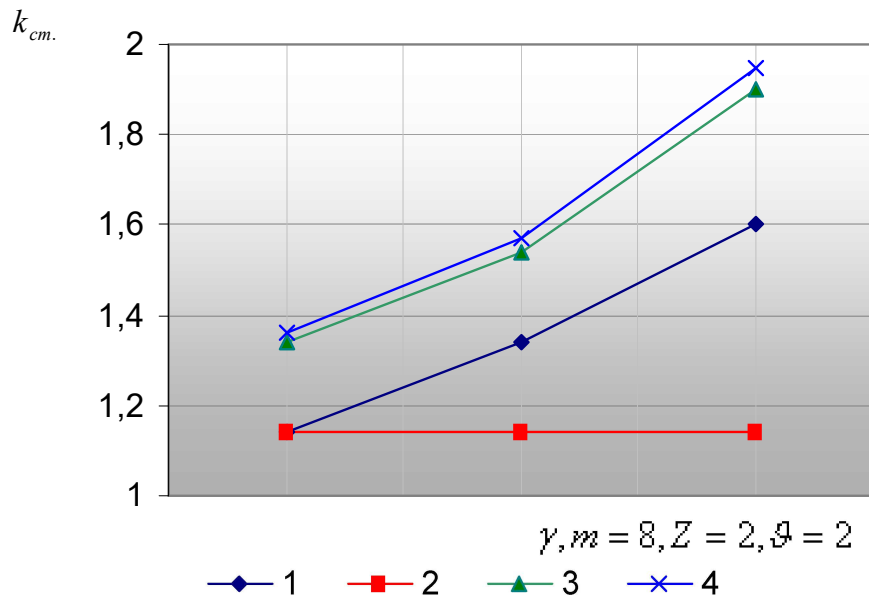


Рисунок 4 – Графіки залежності величин $k(m, \Lambda)_{\min}$, $k(m, \vartheta)_{\min}$, $k(m, \vartheta, \Lambda)_{\min}$ і $k(m, \Lambda, \Theta^{(k)})_{\min}$ від γ й $m=8$, $Z=2$, $\vartheta=2$ для 1 – структурного кодування; 2 – кодування з обмеженням на число серій одиниць; 3 – однознакового структурного кодування; 4 – двознакового структурного кодування

По-друге, при фіксованих параметрах $m=16$, $\vartheta=2$ і γ мінімальне значення ступеня стиснення для двознакового структурного кодування збільшується на 20% із збільшенням кількості припустимих зон Z (рисунок 5).

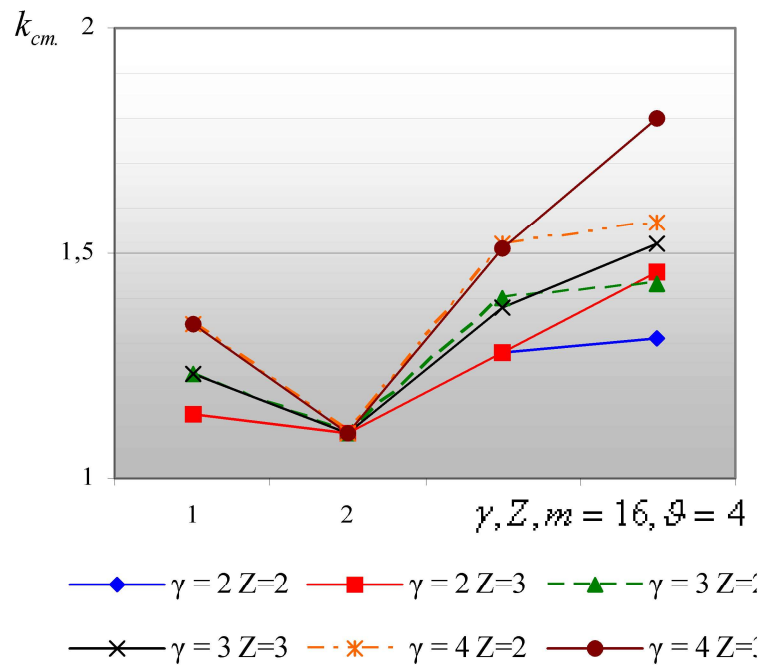


Рисунок 5 – Графіки залежності величин $k(m, \Lambda)_{\min}$, $k(m, \vartheta)_{\min}$, $k(m, \vartheta, \Lambda)_{\min}$ і $k(m, \Lambda, \Theta^{(k)})_{\min}$ від γ й Z , $m=16$, $\vartheta=4$ для: 1-структурного кодування; 2- кодування з обмеженням на число серій одиниць; 3-однознакового структурного кодування; 4- двознакового структурного кодування

По-третє, у випадку обробки двійкового подання компонент ортогональних перетворень, двоознакове кодування досягає найбільшого значення мінімального ступеня стиснення – 6,25 разів.

При цьому перевага за мінімальним значенням ступеня стиснення для двоознакового кодування щодо структурного й кодування за кількістю серій досягає відповідно 70% і 90% для інформаційних мереж АСУ.

Існує тенденція збільшення мінімального ступеня стиснення для двоознакового структурного подання при збільшенні довжини оброблюваної послідовності.

По-четверте, при обробці на двійковому рівні кодових комбінацій архіваторів і графічних форматів найбільше значення мінімального ступеня стиснення для двоознакового кодування досягає 5 разів.

При цьому вираш за мінімальним значенням ступеня стиснення для двоознакового кодування відносно структурного й кодування за кількістю серій досягає відповідно 2,5 й 4,5 разів.

Існує тенденція зниження значення мінімального ступеня стиснення для двоознакового структурного подання при збільшенні кількості припустимих зон у середньому на 10% на одну зону.

За рахунок наявності структурних закономірностей в оброблюваних даних, мінімальне значення ступеня стиснення збільшується відповідно для ортогональних перетворень у 3,5 рази й для кодових комбінацій архіваторів у 2,5 рази (рисунок 6.). Таким чином, розроблено методи двоознакового структурного кодування, що забезпечують додаткове підвищення ступеня стиснення даних в комп'ютерних мережах АСУ.

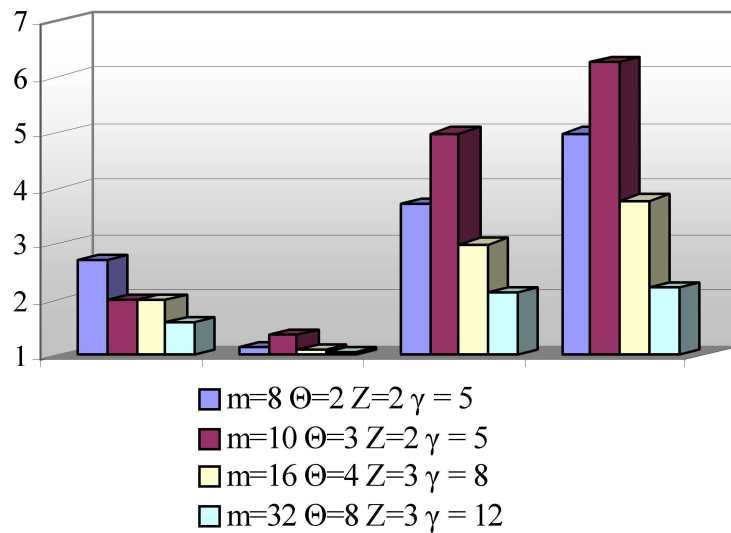


Рисунок 6 – Графіки залежності величин $k(m, \Lambda)_{\min}$, $k(m, \mathcal{G})_{\min}$, $k(m, \mathcal{G}, \Lambda)_{\min}$ і $k(m, \Lambda, \Theta^{(x)})_{\min}$ від γ , Z , m , \mathcal{G} для компактного подання компонентів ортогональних перетворень на основі: 1 – структурного кодування; 2 – кодування з обмеженням на число серій одиниць; 3 – одноознакового структурного кодування; 4 – двоознакового структурного кодування

Встановлені залежності для мінімального значення ступеня стиснення двоознакового структурного кодування в двійковому просторі від наявності структурних особливостей в оброблюваних даних залежно від структурних параметрів: довжина двійкової послідовності m , число серій одиниць \mathcal{G} , кількість припустимих зон Z , кількість заборонених двійкових елементів γ і компоненти вектора обмежень $\Theta^{(k)}$ щодо числа серій одиниць у припустимих зонах.

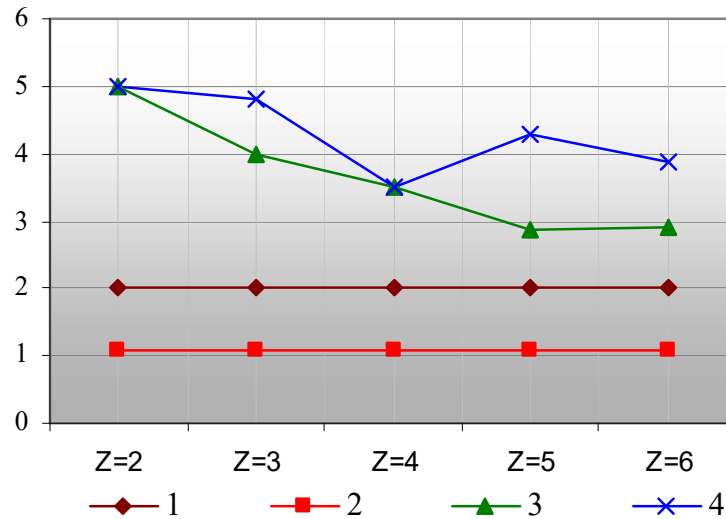


Рисунок 7 – Графіки залежності величин $k(m, \Lambda)_{\min}$, $k(m, \mathcal{G})_{\min}$, $k(m, \mathcal{G}, \Lambda)_{\min}$ і $k(m, \Lambda, \Theta^{(k)})_{\min}$ від Z , $m = 32$, $\mathcal{G} = 8$, $\gamma = 16$ для компактного подання кодових комбінацій архіваторів на основі: 1 – структурного кодування; 2 – кодування з обмеженням на число серій одиниць; 3 – однознакового структурного кодування; 4 – двознакового структурного кодування.

Висновки. У статті наведені теоретичне узагальнення й рішення наукового завдання дослідження методів структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Рішення даного завдання полягало у вирішенні наступних задач: Був проведений огляд існуючих систем структурного кодування даних у комп'ютерних мережах автоматизованих систем управління; Досліджена система структурного кодування даних у комп'ютерних мережах автоматизованих систем управління; На основі отриманих результатів досліджень створена програмна реалізація системи структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Розроблені під час виконання випускної кваліфікаційної роботи за другим (магістерським) рівнем вищої освіти алгоритми дозволяють успішно вирішувати завдання структурного кодування даних у комп'ютерних мережах автоматизованих систем управління. Проведено аналіз предметної галузі в ході якого були виявлені об'єкти, взаємодія яких носить істотний характер для функціональної діяльності предметної галузі, і їхні основні характеристики; побудована алгоритм і вибраний середовище розробки.

Список літератури

1. Kuznetsov, O., Kryvinska, N., Pchenko, O., Smirnova, T., Ulianovska, Y. «Comparative Analysis of Cryptocurrency Trading Platforms Using the Analytic Hierarchy Process». CEUR Workshop Proceedings, 2023, 3628, pp. 106-115.
2. Al-Mudhafar Aqeel, A.M., Smirnova, T., Buravchenko, K., Smirnov, O. «The method of assessing and improving the user experience of subscribers in software-configured networks based on the use of machine learning». Advanced Information Systems, 2023, 7(2), pp. 49-56.
3. Smirnov, O., Sydorenko, V., Aleksander, M., Zhyharevych, O., Yenchov, S. «Simulation of the cloud IoT-based monitoring system for critical infrastructures». CEUR Workshop Proceedings, Volume 3530, 2023, pp. 256-265.
4. Smirnov, O., Odarchenko, R., Smirnova, T., Bondar, S., Volosheniuk, D. «Optimal Structure Construction of Private 5G Network for the Needs of Enterprises». Lecture Notes on Data Engineering and Communications Technologies, 2023, 178, pp. 208–223.
5. Smirnova, T., Gnatyuk, S., Yudin, O., Sydorenko, V., Polozhentsev, A., «The Model for Calculating the Quantitative Criteria for Assessing the Security Level of Information and Telecommunication Systems». CEUR Workshop Proceedings Volume 3156, 2022, Pages 390-399.
6. Smirnova T., Gnatyuk S., Berdibayev R., Avkurova Zh., Iavich M. «Cloud-Based Cyber Incidents Response

- System and Software Tools». *Communications in Computer and Information Science*, 2021, vol 1486. Springer, Cham. pp 169-184.
7. Smirnov O., Kuznetsov A., Kiian A., Kuznetsova T. «Non-binary constant weight coding technique». *CEUR Workshop Proceedings*. Volume 2740, 2020, Pages 102-114.
 8. Smirnov O., Alimseitova Zh., Adranova A., Akhmetov B., Lakhno V., Zhilkishbayeva G. «Models and algorithms for ensuring functional stability and cybersecurity of virtual cloud resources». *Journal of theoretical and applied information technology* Vol.98. No 21, 2020, P. 3334-3346.
 9. Smirnov O., Kuznetsov A., Kiian A., Cherep A., Kanabekova M., Chepurko I. «Testing of code-based pseudorandom number generators for post-quantum application». 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), Ukraine, Kyiv, May 14-18. 2020. P. 172-177.
 10. Smirnov O., Kuznetsov A., Pushkar'ov A., Serhiienko R., Babenko V., Kuznetsova T., «Representation of Cascade Codes in the Frequency Domain». In: Radivilova T., Ageyev D., Kryvinska N. (eds) *Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies*, vol 48. Springer, Cham. 2021. pp 557-587.
 11. Smirnov, O., Markovets, O. Vovk, N., Turchyn, Y., «Model of informational support for social network administrators' content creation». *CEUR Workshop Proceedings Volume 2616*, 2020, Pages 125-136.
 12. Smirnov, O., Drieieva, H., Drieiev, O., Polishchuk, Y., Brzhanov, R., Aleksander, M. «Method of fractal traffic generation by a model of generator on the graph». *CEUR Workshop Proceedings Volume 2616*, 2020, Pages 366-379.
 13. Smirnov, O., Drieieva, H., Drieiev, O., Simakhin, V., Bondar, S., Odarchenko, R. «Managing multifractal properties of the binary sequence generated with the Markov chains», *CEUR Workshop Proceedings Volume 2608*, 2020, Pages 633-645.
 14. Smirnov O. Kuznetsov A., Zaichenko Yu., Pastukhov M., Oleshko O., Kuznetsova K., «Formation of Discrete Signals with Special Correlation Properties». *International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2019*; Odessa; Ukraine; 9-13 September 2019. P.22-28.
 15. Smirnov, O., Kuznetsov, A., Kolovanova, I., Kuznetsova, T., «Noise immunity of the algebraic geometric codes». *International Journal of Computing*; 2019, Volume 18, Issue 4 – Research Institute for Intelligent Computer Systems – 2019. – P. 393-407.
 16. Smirnov, O., Kuznetsov, A., Reshetniak, O., Ivko, N., Katkova, T., Kuznetsova, T., «Generators of Pseudorandom Sequence with Multilevel Function of Correlation». 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 8 – 11 October 2019 . P.517-522.
 17. Smirnov, O., Odarchenko, R., Abakumova, A., Usik, P., Kundydz, M., «QoE optimization technique for media delivery in 5G networks». 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 8 – 11 October 2019. P.597-601.
 18. Smirnov, O., Krasnobayev, V., Yanko, A., Kuznetsova, T. «Methods of nulling numbers in the system of residual classes». *CEUR Workshop Proceedings*, Vol 2588, P. 90-106, 2019.
 19. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Averchev, A., Pastukhov, M., Kuznetsova, K., «Formation of Pseudorandom Sequences with Special Correlation Properties», 2019 3rd International Conference on Advanced Information and Communications Technologies, AICT -2019/ Lviv, Ukraine, 2-6 July, 2019, P. 395-399.
 20. Smirnov, O., Kuznetsov, A., Kiian, A., Zamula, A., Rudenko, S., Hryhorenko, V., «Variance Analysis of Networks Traffic for Intrusion Detection in Smart Grids», 2019 IEEE 6th International Conference On Energy Smart Systems (2019 IEEE ESS), Kyiv, Ukraine April 17-19, 2019 P. 353-358.
 21. Smirnov, O., Kuznetsov, A., Kavun, S., Babenko, B., Nakisko, O., Kuznetsova, K., «Malware Correlation Monitoring in Computer Networks of Promising Smart Grids», 2019 IEEE 6th International Conference On Energy Smart Systems (2019 IEEE ESS), Kyiv, Ukraine April 17-19, 2019 P. 347-352.
 22. Smirnov, O., Kuznetsov, A., Kovalchuk, D., Pastukhov, M., Kuznetsova, K., Prokopovych-Tkachenko, D., «Discrete Signals with Special Correlation Properties», *CEUR Workshop Proceedings Volume 2353*, CEUR Workshop Proceedings 2019, Pages 618-629.