УДК 681.3

О. Я. Матов[1], В. С. Василенко[2]
[1]Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна
[2]Національний авіаційний університет
вул. Космонавта Комарова, 1, 03058 Київ, Україна

Узагальнені завадостійкі коди в задачах забезпечення цілісності інформаційних об'єктів. Код умовних лишків

Розглянуто задачі захисту цілісності інформаційних об'єктів телекомунікаційних мереж (ТКМ) і основні способи (механізми) забезпечення цілісності інформації в умовах природних дій (проблема завадостійкості) для каналів ТКМ. Для підвищення стійкості інформаційних об'єктів щодо порушення цілісності запропоновано збільшення двійкової довжини інформаційних символів та застосування узагальнених завадостійких корегуючих кодів, які були б спроможними забезпечити виявлення та виправлення пакетів викривлень значної тривалості.

***Ключові слова**: базове кодове слово, вірність інформації, викривлення, інформаційний обмін, телекомунікаційна система, цілісність інформаційних об'єктів.*

Вступ

Відповідно до термінології нормативних документів Департаменту спеціальних телекомунікаційних систем і захисту інформації Служби безпеки України [1] під цілісністю інформації розуміється її властивість, яка полягає в тому, що інформація не може бути модифікованою неавторизованим користувачем, або процесом. Іншими словами, під цілісністю інформації розуміється відсутність у ній будь-яких викривлень (модифікацій), які не були санкціоновані її власником, не залежно від причин або джерел виникнення таких викривлень.

Викривлення інформації, тобто порушення її цілісності, можливі на будь-якому етапі її циркуляції в обчислювальних мережах: при зберіганні, передачі або обробці. Причини таких викривлень можуть бути випадковими або навмисними. У свою чергу, випадкові викривлення можуть бути як природними, пов'язаними з дією природних чинників, так і штучними. До числа природних чинників відносяться атмосферні електромагнітні розряди, іскріння контактів в автомобілях, електротранспорті, недостатня надійність електронних елементів й елементів електричних ланцюгів, порушення реєструвального шару магнітних або оптичних но-

сіїв і багато що інше. Випадкові штучні викривлення пов'язані з діяльністю людей — з випадковими помилками персоналу. Навмисні викривлення завжди пов'язані з умисними діями порушників. І ті, і інші дії мають своїм наслідком викривлення того, або іншого числа символів у цифровому представленні інформації, незалежно від використовуваної системи числення, або форми представлення інформації, і, у цьому значенні, є загрозами функціональним властивостям захищеності інформаційних ресурсів — їхньої цілісності та доступності.

Задачі захисту цілісності інформаційних об'єктів телекомунікаційних мереж

Наслідком природних впливів у каналах телекомунікаційних мереж (ТКМ) є зменшення співвідношення енергетик сигнал/шум (сигнал/завада). Це відношення визначає вірність інформації, визначену, наприклад, через імовірність помилок двійкових символів (біт) $P_{пом}$, а також інтенсивність цих помилок. Слід підкреслити, що інтенсивність природних дій у каналах деяких ТКМ, яка визначається, в основному, цим співвідношенням, є настільки значною, що лише за їх рахунок, без урахування можливостей зловмисників по створенню загроз, наприклад, різного роду завад, середня вірогідність помилки двійкового символу (біта) $P_{пом}$ для телефонних кабельних каналів ТКМ складає від $1{,}29\cdot10^{-4}$ до $2\cdot10^{-3}$; для радіорелейних телефонних — від $2{,}66\cdot10^{-4}$ до $7{,}3\cdot10^{-4}$ відповідно. Відомо також, що із часом такі помилки групуються в пакети двох видів: «короткі» — тривалістю 2...10 мс і «довгі» — тривалістю 100…200 мс. «Короткі» пакети з'являються частіше, але більшість зафіксованих помилок зосереджено в «довгих» групах (75–90 %).

Використання викривленої інформації чревате наслідками (часто надзвичайно важкими) для власників або користувачів цієї інформації. Тому задача забезпечення цілісності та доступності інформаційних ресурсів є однією з найактуальніших при розробці й експлуатації АС і їхніх елементів. Ця необхідність підтверджується й вимогами щодо допустимої вірогідності P_n помилок у повідомленнях, яку слід трактувати як вірогідність порушення цілісності інформаційних об'єктів, які обробляються (якщо передача й обробка інформації здійснюється у вигляді повідомлень). Наприклад, вона може задаватися від 10^{-4} (у задачах оперативно-виробничого планування) до 10^{-6} (у задачах бухгалтерського обліку).

Для забезпечення контролю та поновлення цілісності інформаційних об'єктів, включаючи й відновлення зруйнованої інформації, до складу інформації, яка захищається, включають надмірну інформацію — ознаку цілісності або контрольну ознаку (залежно від прийнятої в задачах контролю цілісності або завадостійкого кодування термінології) — своєрідний образ відображення цієї інформації, процедура формування якого відома, і який із дуже високою вірогідністю відповідає інформації, що захищається.

При цьому між інформацією, що захищається, і ознаками цілісності, або контрольними ознаками встановлюється регулярний (функціональний) односторонній зв'язок (процедури розрахунку контрольної ознаки за початковою інформацією, що захищається, відомі, а процедури розрахунку початкової інформації за контрольними ознаками найчастіше не існує). Контроль цілісності (відсутність викривлень) зводиться при цьому до тих або інших процедур перевірки наявності

вказаного регулярного (функціонального) одностороннього зв'язку між ознаками цілісності та прийнятої з каналу зв'язку (або зчитаної із запам'ятовуючого пристрою (ЗП)) інформацією.

Механізми забезпечення цілісності істотно залежать від умов їхнього застосування, а саме від характеру впливу випадкових (природних) або штучних (зловмисних) викривлень.

Характерною особливістю випадкових викривлень є те, що вони, через відсутність навмисності, порушують регулярний (функціональний) односторонній зв'язок між прийнятою (або зчитаною з ЗП) інформацією й ознаками цілісності, сформованими перед передачею (перед записом у ЗП). Тому при виявленні порушення вказаного зв'язку встановлюється факт наявності таких викривлень, а за певних умов, і їхнього місця, і величини (характеру). За відсутності порушення цього зв'язку встановлюється факт відсутності викривлень.

Характерною ж особливістю навмисних викривлень є те, що зловмисник прагне забезпечити, зімітувати наявність регулярного (функціонального) зв'язку між модифікованою ним початковою інформацією, прийнятою (або зчитаною з ЗП), і ознаками цілісності. Із цією метою порушник, використовуючи знання процедур формування контрольних ознак, після необхідної для його цілей модифікації початкової інформації перед передачею одержувачу (перед записом у ЗП) забезпечує формування відповідних ознак. При успішному формуванні вказаних ознак, розкрити наявність модифікації неможливо. Для боротьби із цим власнику (або авторизованому користувачу) необхідно використовувати або секретні (невідомі потенційним порушникам) процедури формування контрольних ознак (що дуже складно забезпечити), або вводити в загальновідомі процедури формування контрольних ознак секретні параметри (ключі перетворення). Не знаючи цих секретних параметрів, порушник не зуміє забезпечити, зімітувати наявність регулярного (функціонального) зв'язку між модифікованою ним початковою інформацією, прийнятою (або зчитаною з ЗП), і ознаками цілісності.

Виділяють дві основні причини виникнення природних викривлень у процесі циркуляції інформації в мережах:

— збої в якійсь частині устаткування мережі, або виникнення несприятливих об'єктивних подій у мережі (наприклад, колізій при використанні методу випадкового доступу до мережі). Як правило, система передачі даних готова до такого роду проявів і усуває їх за допомогою планово передбачених засобів;

— завади, викликані зовнішніми джерелами й атмосферними явищами.

Труднощі боротьби із завадами полягають у безладності, нерегулярності та в структурній схожості завад з інформаційними сигналами. Тому захист інформації від викривлень внаслідок шкідливого впливу завад має велике практичне значення, і є однією із серйозних проблем сучасної теорії та техніки інформаційного обміну в каналах ТКМ.

Серед основних способів (механізмів) забезпечення цілісності (і в певному значенні — доступності) інформації в умовах природних дій (проблема завадостійкості) для каналів ТКМ (взагалі для мереж передачі даних) слід виділяти наступні.

1. Збільшення вже згаданого співвідношення сигнал/завада за рахунок підвищення енергетики сигналу (велика початкова потужність, регенерація на пунктах підсилення та ретрансляції, що вимагає значних енергетичних або матеріальних витрат.

2. Збільшення співвідношення сигнал/завада за рахунок зниження рівня завад (шумів) шляхом використання спеціальних ліній зв'язку, кабельних ліній зв'язку з низьким рівнем власних шумів, наприклад, оптоволоконних, що також вимагає значних матеріальних витрат, і може бути реалізованим лише в кабельних лініях зв'язку.

3. Забезпечення хоча б задовільної узгодженості смуги пропускання $П$ каналу зі спектром сигналу, який визначається параметрами сигналу, у першу чергу, його тривалістю $\tau \approx 1/B$, де τ — тривалість сигналу, а B — технічна швидкість передачі інформації (швидкість по символьній передачі) у даному каналі. Задовільною найчастіше вважають таку узгодженість, коли $П \geq 2B$.

4. Застосування групових (мажоритарних) методів захисту, які ґрунтуються на використанні декількох каналів зв'язку (3...5), що є фізично (найчастіше, навіть, географічно) рознесеними, якими передається одна й та ж інформація, або на багатократній передачі (3...5 раз) однієї і тієї ж інформації одним каналом зв'язку. У першому випадку необхідні істотні матеріальні витрати, а в другому значно зменшується пропускна спроможність каналу зв'язку (у 3...5 раз), а час затримки передавання інформаційних об'єктів може стати неприпустимо великим. Із цих причин у системах передачі даних використання цих методів є не завжди доцільним.

5. Застосування різного роду завадостійких кодів із виявленням помилок у прийнятій (зчитаній) інформації, які дозволяють реалізувати програмні, апаратурні або програмно-апаратурні засоби виявлення викривлень. Це, у свою чергу, дає можливість застосування способів передачі повідомлень із різного роду зворотним зв'язком (інформаційним — деякий аналог мажоритарного методу з багатократною передачею інформації й зворотним прийомом й ухваленням рішення щодо правильності передачі на стороні передавача, або з вирішальним зворотним зв'язком — багатократній, за необхідності, передачі з ухваленням рішення щодо правильності передачі на стороні приймача). Недоліки таких способів забезпечення цілісності зводяться до необхідності організації другого (зворотного) каналу зв'язку, тобто до істотних матеріальних витрат, а також до збільшення часу затримки передавання інформаційних об'єктів, який може бути неприпустимо великим.

6. Застосування різного роду завадостійких корегуючих кодів, які дозволяють реалізувати програмні, апаратурні або програмно-апаратурні засоби виявлення й усунення викривлень.

Останній зі способів (механізмів) забезпечення цілісності інформаційних об'єктів — із застосуванням завадостійких корегуючих кодів наразі знайшов широке застосування в стандартах радіозв'язку, у тому числі мобільного, стільникового зв'язку. Він не потребує зворотного каналу й забезпечує, як правило, прийнятне значення часу затримки передавання інформаційних об'єктів. Тому, чи не єдиною проблемою в цих та інших ТКМ із використанням телефонних кабельних

та радіоканалів є проблема забезпечення цілісності інформаційних об'єктів в умовах впливу навіть природних (не говорячи вже про штучні, навмисні завади) пакетних викривлень, як «коротких» (тривалістю 2...10 мс) так і особливо «довгих» (тривалістю 100…200 мс). Це є особливо актуальним і для вже згаданих систем стільникового зв'язку. Наприклад, у стандартах CDMA базовий цифровий потік розбивається на пакети тривалістю по 20 мс і подається на згорточний кодер с половинною швидкістю [2]. При цьому тривалість пакета викривлень може бути порівняною чи, навіть, значно перевищувати тривалість інформаційного пакета, що може суттєво вплинути на результативність процедур інформаційного обміну.

Як вихід із таких ситуацій може розглядатися можливість збільшення кількості двійкових символів в інформаційних символах (збільшення їхньої довжини в двійкових одиницях) та застосування завадостійких корегуючих кодів, які були б спроможними забезпечити виявлення та виправлення викривлень у таких інформаційних символах, що є еквівалентним забезпеченню корегування пакетів викривлень значної тривалості. Як такий у статті пропонується узагальнений завадостійкий корегуючий код умовних лишків.

Узагальнений завадостійкий код умовних лишків

Під узагальненими розумітимемо коди, призначені для виявлення (виявлення й виправлення) пакетних викривлень, у яких використовуються алгоритми кодування й декодування, аналогічні відповідним алгоритмам двійкових кодів, але по відношенню до b-розрядних, узагальнених символів (УС).

У цих кодах початкова двійкова кодова послідовність — базове кодове слово $I_1, I_2, \ldots, I_n$ розбивається на $n = k/b$ узагальнених b-розрядних символів — груп двійкових розрядів з розрядністю b, в яких передбачається виявлення та виправлення викривлень:

$$\underbrace{I_1 \ldots\ldots I_b}_{\text{1-й УС}},\ \underbrace{I_{b+1} \ldots\ldots\ldots\ldots I_{2b}}_{\text{2-й УС}}, \ldots, \underbrace{I_{m-b+1} \ldots\ldots \mathrm{I}_n}_{n\text{-й УС}}.$$

Двійкові символи, що входять до однієї b-розрядної групи, розглядаються як b-значний УС, який може приймати будь-яке із s значень від 0 до $s - 1$, де $s = 2^b$.

Код умовних лишків (лишків умовних код — ЛУ-код) є одним із прикладів узагальнених кодів [3]. Теоретичною основою ЛУ-коду є теорія лишкових класів. Із цієї теорії [4] відомо, що будь-яке число можна представити у вигляді набору лишків від розподілу цього числа на набір взаємно простих чисел, які мають назву основ системи числення, — p_i, де $i = 1, 2, \ldots, n$, n — кількість таких основ. Вибір величини n здійснюється з умови, яка викладена нижче. Тоді:

$$A = \alpha_1, \alpha_2, \ldots, \alpha_n, \tag{1}$$

де $\alpha = A - [A/p_i] \cdot p_i$, а позначка $[A/p_i]$ означає операцію розрахунку цілої частини від дробового числа A/p_i. При цьому між числом A та його уявленням (1) існує взаємна однозначна відповідність якщо:

$$A \le P = \prod_{i=1}^{n} p_i \cdot$$

У цьому виразі величина P — діапазон представлення або робочий діапазон чисел. Звернемо увагу на те, що величина α_i представляє собою групу двійкових розрядів, кількість яких не перевищує розрядності відповідної основи p_i.

Чудовою властивістю системи лишкових класів (СЛК) є те, що до неї легко вводяться властивості виявлення та виправлення викривлень. Відомо, що якщо ввести ще одну, контрольну, основу p_k, то уявлення A в розширеному діапазоні $R = P \cdot p_k$, у вигляді

$$A = \alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_k, \qquad (2)$$

де α_k — лишок по основі p_k, має чудову для побудови корегуючих кодів властивість: при $p_k > p_n$ будь-яке викривлення в одному з лишків α_i може бути знайденим, а при $p_k > 2 \cdot p_n \cdot p_{n-1}$, де p_n, p_{n-1} — найбільші з основ, може бути й виправленим. Це означає, що при представленні чисел у вигляді (2) створюється завадостійкий код із можливостями або виявлення викривлень, або їхньої корекції.

Такий код має принаймні 2 недоліки. Перший із них пов'язаний із необхідністю роботи із числами в системі числення в залишкових класах, а другий — з тим, що можливі викривлення знаходяться й виправляються (викривлений символ поновлюється) тільки в тому випадку, якщо викривлений лише один із символів α_i, тобто викривлення повинні бути фіксованими в межах однієї з груп розрядів. Цей недолік достатньо просто усувається в коді умовних лишків, який уводиться наступним чином.

Нехай існує код деякого числа A, представленого в будь-якій системі числення, зокрема, позиційній, наприклад, двійковій. Для визначеності, нехай це число A представлено послідовністю з нулів і одиниць. Розіб'ємо цю послідовність певним (у загальному випадку довільним) чином на n груп, як і для решти узагальнених кодів.

Як показано вище, код кожного i-го УС слід розглядати як s-значний УС α_i, який може приймати будь-яке з s значень від 0 до $s - 1$, де $s = 2^b$. Будемо умовно вважати цей код лишком деякого умовного числа A по основі p_i. Оскільки величина α_i, як елемент початкового числа

$$0 \le \alpha_i \le s - 1,$$

а як лишок від ділення A на p_i

$$0 \le \alpha_i \le p_i,$$

то для представлення коду будь-якої групи у вигляді лишку по основі p_i необхідно, щоб виконувалася умова:

$$p_i > s - 1,$$

інакше в групу з b розрядів може бути записаним код $\alpha_i \geq p_i$, що в лишкових класах не припустимо.

При такому підході будь-які комбінації початкового коду числа A «вписуються» в систему числення з основами p_i (i = 1, 2,...). Якщо розширити систему основ на контрольну p_k і для одержаного набору умовних лишків α_i (i = 1, 2,...) розрахувати умовний лишок α_k, то на одержане умовне число

$$A = \alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_k \qquad (3)$$

розповсюджуються всі можливості СЛК по виявленню й виправленню викривлень, тобто одержаний код (3) має всі властивості коду (2), але останній код може бути отриманим для будь-якої двійкової послідовності, а не тільки по відношенню до чисел у лишкових класах. Відзначимо, що таким чином усунений перший недолік коду (2).

Оскільки для отримання контрольної ознаки, тобто для кодування будь-якої послідовності двійкових цифр завадостійким кодом, умовно, не реально, не фізично, групи розрядів початкового числа розглядаються як деякі лишки, то такий код одержав найменування коду умовних лишків.

Слід звернути увагу на те, що при кодуванні ЛУ-кодом початкова послідовність не змінюється, до неї тільки приформовуються додаткові, обчислені за окремими правилами, контрольні символи.

Такий код дозволяє знаходити й виправляти b-розрядні пакети викривлень, згруповані в межах будь-якого з n УС, і потребує при цьому надмірності біля $r \approx 2b + 1$ двійкових розрядів (оскільки $p_k \approx 2p_n p_{n-1}$, $r = [\log_2 p_k] + 1$. У конкретних випадках ця надмірність може відхилятися в ту або іншу сторону, що залежить також від алгоритмів кодування–декодування.

Виявлення порушення цілісності із застосуванням коду умовних лишків

Нагадаємо, що ЛУ-код відноситься до узагальнених кодів, у яких усі операції з кодування–декодування здійснюються не над окремими двійковими розрядами, а над їхніми групами — узагальненими символами. В основі ЛУ-коду лежать властивості системи лишкових класів, тому в ньому принципово можуть бути використані відомі [4] алгоритми кодування–декодування. В основі цих алгоритмів лежить той факт, що будь-яке викривлення в одній із груп розрядів α_i переводить початкове число з робочого діапазону [0, $P = \prod_{i=1}^{n} p_i$) до діапазону [P, $R = q \cdot P$), тобто приводить до збільшення початкового числа $A < P$ на деяку величину $l_i \cdot R_i$. Тут q — контрольна, надлишкова основа така, що її значення перевищує значення будь-якої з основ, що утворюють робочий діапазон P, тобто $q > p_i$, i = 1, 2, …, n [3]; l_i і R_i — цілі числа ($R_i = R/p_i$); R_i — основні константи такої системи числення. Дійсно, якщо вихідне число

$$A = \alpha_1, \alpha_2, \ldots, \alpha_i, \ldots, \alpha_n, \alpha_k$$

є викривленим по основі p_i і має вигляд:

$$A' = \alpha_1, \alpha_2, \ldots, \tilde{\alpha}_i, \ldots, \alpha_n, \alpha_k,$$

де

$$\tilde{\alpha}_i = \{\alpha_i + \Delta\alpha_i\} \pmod{p_i},$$

то це в системі числення в лишкових класах є еквівалентним наступному перетворенню:

$$A' = (\alpha_1, \alpha_2, \ldots, \{\alpha_i + \Delta\alpha_i\} \pmod{p_i}, \ldots, \alpha_{n1}, \alpha_k) =$$
$$= (\alpha_1, \alpha_2, \ldots, \alpha_i, \ldots, \alpha_n, \alpha_k) + (0, 0, \ldots, \Delta\alpha_i, \ldots, 0, 0).$$

При цьому величина ΔA викривлення перевищує величину робочого діапазону P:

$$\Delta A = (0, 0, \ldots, \Delta\alpha_i, \ldots, 0, 0) > P.$$

Це пов'язано з тим, що тільки число виду $\Delta A = l_i \cdot R_i = l_i \cdot R/p_i$ має всі лишки, окрім лишка по основі p_i такими, що дорівнюють нулю. Але величина $R/p_i > R/q$ із тієї причини, що $q > p_i$, і тоді, навіть при $l_i = 1$, величина викривлення $\Delta A = l_i \cdot R_i > P = R/q$.

Відтак, сума $A' = A + \Delta A > P$, тобто викривлене число (рис. 1) виходить за межі робочого діапазону P і попадає до діапазону $[P, R)$, що може бути певним чином виявленим.

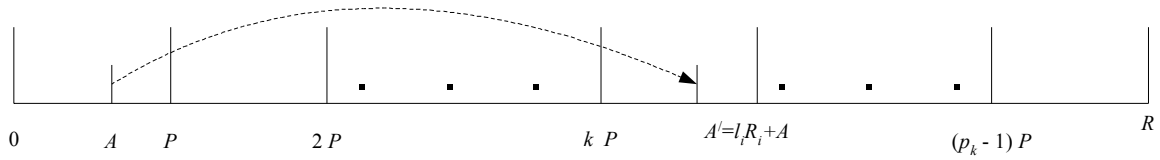

Рис. 1. До виходу викривленого числа за межі робочого діапазону

Отже, для виявлення наявності порушень цілісності досить установити факт виходу прийнятого (зчитаного) числа за межі робочого діапазону. Запропоновані нижче алгоритми контролю цілісності (кодування–декодування) також використовують цей факт.

Алгоритм контролю цілісності з використанням процедури нулевізації

Алгоритм із використанням процедури нулевізації є одним із механізмів контролю цілісності (наявності викривлень в інформаційних об'єктах внаслідок природних чи штучних впливів), що використовують властивості завадостійкого ко-

дування, які має система лишкових класів. Тому ці алгоритми передбачають наявність процедури кодування–декодування, яка складається із двох етапів.

На першому етапі алгоритму (при формуванні ознаки цілісності, контрольної ознаки чи кодуванні) виконується послідовність операцій (процедура нулевізації), яка зводиться до того, що по першим n лишкам α_i ($i = 1, 2, \ldots, n$) числа

$$A' = (\alpha_1, \alpha_2, \ldots, \{\alpha_i + \Delta\alpha_i\} \pmod{p_i}, \ldots, \alpha_n, \alpha_k)$$

послідовно формуються, так звані мінімальні числа вигляду:

$$t_1 = (\alpha_1, \alpha_2^{(1)}, \alpha_3^{(1)}, \ldots, \alpha_n^{(1)}, \alpha_k^{(1)}),$$
$$t_2 = (0, (\alpha_2 - \alpha_2^{(1)}) \pmod{p_2}, \alpha_3^{(2)}, \ldots, \alpha_n^{(2)}, \alpha_k^{(2)}),$$
$$t_3 = (0, 0, (\alpha_3 - \alpha_3^1 - \alpha_3^2) \pmod{p_3}, \alpha_4^{(3)}, \ldots, \alpha_n^{(3)}, \alpha_k^{(3)}),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$t_n = (0, 0, 0, \ldots, (\alpha_n - \sum_{j=1}^{n-1} \alpha_n^{(j)}) \pmod{p_n}, \alpha_k^{(n)}).$$

Звернемо увагу на те, що кожне з таких мінімальних чисел може бути представленим у вигляді:

$$t_i = v_i \cdot \prod_{j=1}^{i-1} p_i .$$

З урахуванням того, що в системі лишкових класів

$$t_i \pmod{p_i} = \alpha_i^{i-1} = \{\alpha_i - \sum_{j=1}^{i-1} \alpha_i^{(j)}\} \pmod{p_i} = v_i \cdot \prod_{j=1}^{i-1} p_i \pmod{p_i}$$

величину v_i можна визначити як

$$v_i = \{\alpha_i^{i-1} / \prod_{j=1}^{i-1} p_i\} \pmod{p_i} = \{(\alpha_i - \sum_{j=1}^{i-1} \alpha_i^{(j)}) / \prod_{j=1}^{i-1} p_i\} \pmod{p_i},$$

для всіх лишків α_i з номерами $i > 1$, а для першого з лишків α_1 значення $v_1 = 1$.

Сума цих чисел $T = \sum_{i=1}^{n} t_i$ має наступні дві властивості [4]. По-перше, лишки цієї суми по всім основам, окрім p_k, завжди дорівнюють лишкам вихідного числа A. По-друге, величина цієї суми завжди є меншою ніж величина робочого діапазону: $T < P$, тобто величина T лежить у межах робочого діапазону й для не викривлених чисел $T = A$.

Таким чином, процес отримання величини $T = A$ є *процесом кодування* вихідного числа ЛУ-кодом, при чому значення A залежить лише від цього вихідного числа, і не залежить від невідомої при кодуванні величини лишку по контрольній

основі p_k. Цей лишок α_k (контрольна ознака, ознака цілісності, що розшукується) дорівнює при цьому сумі за модулем p_k проміжних величин $\alpha_k^{(i)}$ (i = 1, 2, ..., n) тобто:

$$\alpha_k = T \,(\mathrm{mod}\, p_k) = (\sum_{i=1}^{n} \alpha_k^{(i)}) \,(\mathrm{mod}\, p_k).$$

На другому етапі алгоритму (контроль цілісності чи декодування) виконується віднімання із числа A' величини T, що призводить до того, що отримана різниця

$$\Gamma = A' - T = k \cdot P$$

має по всім основам, окрім контрольної, лишки, що дорівнюють нулю, а по контрольній — лишок, величина якого:

$$\gamma = (\alpha_k - (T \bmod p_k)) \,(\mathrm{mod}\, p_k) = (k \cdot P) \,(\mathrm{mod}\, p_k).$$

Величина Γ при запису в лишкових класах має вигляд:

$$\gamma = (A' - T) \,(\mathrm{mod}\, p_k)) = (0, 0, \ldots, 0, \ldots 0, k \cdot P \,(\mathrm{mod}\, p_k)),$$

де k = 0, 1, 2, …, p_k – 1.

На цьому ж етапі здійснюється аналіз величини γ, отриманої внаслідок нулевізації інформаційного об'єкта. Звернемо увагу на те, що для не викривлених чисел величина k = 0, а отже, і γ = 0, для викривлених $k \neq 0$ і $\gamma \neq 0$.

Таким чином, аналіз величини γ, отриманої внаслідок нулевізації інформаційного об'єкта, дозволяє встановити факт наявності чи відсутності порушень цілісності.

Для ілюстрації можливостей даного алгоритму розглянемо два приклади.

Приклад 1. Нехай необхідно сформувати ознаку цілісності (контрольну ознаку) — закодувати з використанням алгоритму нулевізації вихідний код 110110, вважаючи, що можлива довжина пакета викривлень b = 2. Тоді можливе розбиття вихідного коду на три (n = 3) дворозрядні групи α_1 = 11, α_2 = 01, α_3 = 10, s = 4, а в якості умовних основ можна вибрати p_1 = 4, p_2 = 5, p_3 = 7. При цьому значення контрольної основи ($p_k > 2 \cdot p_n \cdot p_{n-1} = 2 \cdot 5 \cdot 7 = 70$) можна вибрати величиною p_k = 71, що потребує для свого відображення семи розрядів. Визначимо також основні константи цієї системи числення для обраної сукупності основ: R_1 = 2485; R_2 = = 2088; R_3 = 1420; $R_4 = P$ = 140; m_1 = 1; m_2 = 2; m_3 = 6; m_4 = 35.

Тоді для формування ознаки цілісності (контрольної ознаки) слід сформувати код наступного вигляду:

$$A = 11.01.10.0000000.$$

При цьому, перше мінімальне число t_1 повинно мати лишок по першій основі, що дорівнює $11_{(2)} = 3_{(10)}$. Таким числом є $t_1 = 3$, або при представленні в СЛК із вибраними основами: $t_1 = 11.11.11.0000011$.

Друге мінімальне число t_2 повинно мати лишок по першій основі, який дорівнює нулю, а по другій:

$$((\alpha_2 - \alpha_2^{(1)}) \pmod{p_2} = (1 - 3) \pmod 5 = 11_{(2)}.$$

Мінімальним числом, яке має такі лишки по першій і другій основам, є $t_2 = 8$, тобто: $t_2 = 00.11.01.0001000$.

Трете мінімальне число t_3 повинно мати нульові лишки по першим двом основам, а по третій:

$$((\alpha_3 - \alpha_3^{(1)} - \alpha_3^{(2)}) \pmod{p_3} = (2 - 3 - 1) \pmod 7 = 5 = 101_{(2)}.$$

Мінімальним числом, що має такі лишки, є $t_3 = 40$, тобто: $t_3 = 00.00.101.0101000$.

Тоді сума цих чисел $T = \sum_{i=1}^{3} t_i$ дорівнює 51, тобто:

$$T = 11.01.10.0110011.$$

Код T і є результатом кодування. У цьому коді лишок по контрольній основі $\gamma = 0110011$ і є шуканим значенням ознаки цілісності (контрольної ознаки).

Приклад 2. Нехай треба здійснити контроль цілісності (декодувати) з використанням алгоритму нулевізації для умов наведеного вище прикладу 1 код $A' = 11.01.01.0110011$, у якому викривлена третя пара розрядів. Як і раніше

$$t_1 = 011.011.011.0000011,\ t_2 = 000.011.001.0001000.$$

Для третього мінімального числа t_3:

$$((\alpha_3 - \alpha_3^{(1)} - \alpha_3^{(2)}) \pmod{p_3} = (1 - 3 - 1) \pmod 7 = 4 = 100_{(2)}.$$

Мінімальним числом, що має такі лишки, є $t_3 = 60$, тобто: $t_3 = 000.000.100.0111100$.

При цьому:

$$T = \sum_{i=1}^{3} t_i = 71,$$

тобто оскільки $T \pmod{71} = 0$, тоді $T = 11.01.01.0000000$, і

$$\gamma = \Gamma \pmod{p_k} = (\alpha_k - (T \bmod p_k)) \pmod{p_k} = (0110011 - 0000000) \pmod{71} = 51.$$

Оскільки ознака цілісності (контрольна ознака) є відмінною від нуля — $\gamma \neq 0$, то робимо висновок про виявлення порушення цілісності, чи про наявність викривлення в об'єкті, що декодується.

Алгоритм контролю та поновлення цілісності з використанням процедури нулевізації

Відомо, що [4] при $p_k > 2 \cdot p_n \cdot p_{n-1}$ є взаємно однозначна відповідність між величиною викривлення $\Delta\alpha_i$ і величиною γ, що дає змогу, отримавши γ, визначити місце та величину викривлення, тобто здійснити її виправлення.

Для виявлення можливостей побудови алгоритму контролю та поновлення цілісності з використанням процедури нулевізації нагадаємо, що на числовій осі величина викривлення $l_i \cdot R_i$ відображається точкою в деякому піддіапазоні «контрольного» діапазону $[(P + 1), R)$.

Відповідно, процес викривлення початкового числа A відобразиться переміщенням точки A з робочого діапазону $[0, P)$ до деякого іншого піддіапазону. Звернемо увагу на те, що в залежності від величини початкового числа (рис. 2), викривлене число (A_1 чи A_2) може попасти до одного із суміжних діапазонів із номерами k або $(k - 1)$.

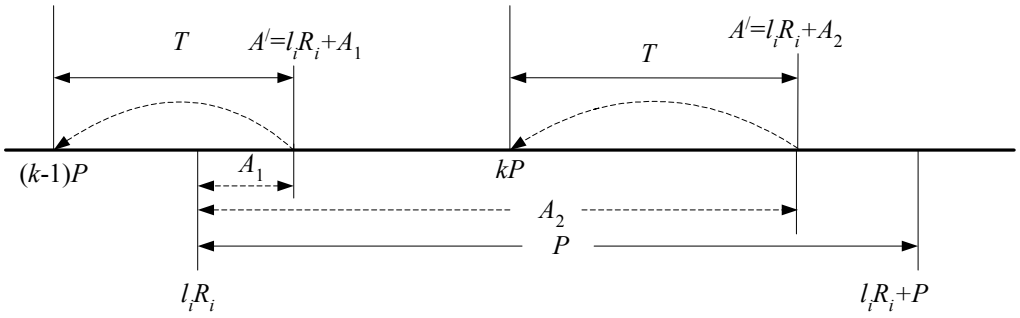

Рис. 2. Ілюстрація процесу нулевізації

Зокрема, при

$$A = A_1 \leq k \cdot P - l_i \cdot R_i,$$

це буде (в уже прийнятих позначеннях) діапазон $[(k - 1) \cdot P, k \cdot P)$, тобто діапазон із номером $(k - 1)$, а при

$$A = A_2 > k \cdot P - l_i \cdot R_i$$

це буде діапазон із номером k.

Унаслідок операції нулевізації із числа A', яке контролюється, віднімаються відповідно числа $T = A' - (k - 1) \cdot P < P$, чи $T = A' - k \cdot P < P$. При цьому по контрольній основі q отримується результат γ такий, що відповідає лівій межі (див. рис.

2) піддіапазону $[(k-1)\cdot P, k\cdot P)$, тобто величині $(k-1)\cdot P$, або ж такий, що відповідає лівій межі піддіапазону $[k\cdot P, (k+1)\cdot P)$, а саме величині $k\cdot P$.

Тобто маємо:

$$\gamma = \{k\cdot P\}_q, \text{ або } \gamma = \{(k-1)\cdot P\}_q,$$

звідки, за правилами СЛК, отримаємо:

$$k = \{\gamma/\{P\}_q\}_q, \tag{4а}$$

чи

$$(k-1) = \{\gamma/\{P\}_q\}_q. \tag{4б}$$

Таким чином, використовуючи вирази (4а), (4б) завжди можна визначити номер того діапазону — число k, у який потрапили викривлене число та результат нулевізації.

Звернемо увагу на те, що величина викривлення $l_i\cdot R_i$ і результат нулевізації $k\cdot P$ є близькими, тобто їхня різниця є меншою за величину робочого діапазону P. Це й надає принципову можливість визначити місце викривлення.

Оскільки подальші міркування певним чином залежать від співвідношення величин $k\cdot P$ та $l_i\cdot R_i$, розглянемо два наступних випадки.

У першому випадку, коли $k\cdot P > l_i\cdot R_i$, значення R_i, яке характеризує величину й місце викривлення, можна визначити з очевидної нерівності:

$$k\cdot P - [k\cdot P/R_i]\cdot R_i < P. \tag{5}$$

Підставимо в (5) замість R_i його значення у вигляді:

$$R_i = P\cdot q/p_i.$$

Тоді:

$$k\cdot P - [k\cdot P/ R_i]\cdot R_i = k\cdot P - [k\cdot P\cdot p_i / P\cdot q]\cdot P\cdot q/ p_i = k\cdot P - [k\cdot p_i /q]\cdot q\cdot P / p_i < P.$$

Розділивши обидві частини правої частини останньої нерівності на величину P, отримаємо:

$$k - [k\cdot p_i /q]\cdot q/p_i < 1.$$

Після множення обох частин останньої нерівності на величину p_i, одержимо:

$$k\cdot p_i - [k\cdot p_i /q]\cdot q < p_i, \tag{6}$$

або

$$\{k\cdot p_i\}_q < p_i. \tag{7}$$

Звернемо увагу на те, що в (5), (6) вирази у квадратних дужках є не що інше як величина l_i, оскільки:

$$[k \cdot P / R_i] = [k \cdot p_i / q] = l_i. \tag{8}$$

Вирази (5) та еквівалентні їм вирази (6), (7) утворюють системи нерівностей по n нерівностей у кожній (величина i може приймати значення $i = 1, 2, \ldots, n$), у яких справедливим є лише одна нерівність для того номера i та того значення основи p_i, по яким має місце викривлення.

Таким чином, унаслідок розв'язання систем нерівностей (6), (7) щодо змінної p_i, місце викривлення стає виявленим. Для визначення ж його величини проаналізуємо величини $T = A' - k \cdot P$, чи $T = A' - (k - 1) \cdot P$, які формуються по всім лишкам окрім лишку по контрольній основі в ході операції нулевізації числа, яке контролюється.

Як видно з рис. 2, величина сформованого в ході нулевізації числа T є меншою вихідного числа A_2 на величину $(k \cdot P - l_i \cdot R_i)$, тобто:

$$T = A' - k \cdot P = A_2 - (k \cdot P - l_i \cdot R_i) < A_2, \tag{9}$$

та

$$\Delta \widetilde{A} = (k \cdot P - l_i \cdot R_i),$$

а величина скорегованого числа повинна визначатися як:

$$A_2 = T + (k \cdot P - l_i \cdot R_i).$$

Тобто величина скорегованого значення лишку:

$$\alpha_i = \{\widetilde{\alpha}_i + \Delta\alpha_i) = \{ T + (k \cdot P - l_i \cdot R_i)\} \bmod p_i = \{\widetilde{\alpha}_i - \{l_i \cdot R_i\} \bmod p_i\} \bmod p_i,$$

або з урахуванням (8):

$$\alpha_i = \{\widetilde{\alpha}_i - \{[k \cdot p_i / q] \cdot R_i\} \bmod p_i\} \bmod p_i. \tag{10}$$

Приклад 3. Нехай у СЛК із основами 2, 3, 5, 17 вихідне число $18_{10} = 0, 0, 3, 1$ внаслідок викривлення перетворилося на $0, 0, 0, 1 = 120_{10}$.

Результат нулевізації має значення:

$$\Gamma = 0, 0, 0, 1, \gamma = 1,$$

звідки

$$k = \{\gamma / \{P\}_q\}_q = \{1/13\}_{17} = (1 + 3 \cdot 17)/13 = 52/13 = 4.$$

Пошук місця викривлення дає:

$$\{k \cdot p_i\}_q < p_i,$$

для $k = 4$:

$$\{4 \cdot 5\}_{17} < 5,$$

тобто виявлене викривлення по основі p_3.

Розрахунок скорегованого лишку по основі p_3:

$$\alpha_3 = \{0 - \{[20/17] \cdot 42\}_5\}_5 = 5 - 2 = 3.$$

Видно, що корекція викривлення здійснена правильно.

У другому випадку, коли результат нулевізації — число $(k - 1) \cdot P$ (див. рис. 2) є меншим за величину викривлення $l_i \cdot R_i$, величина різниці за виразом (6)

$$k \cdot p_i - [k \cdot p_i / q] \cdot q < p_i$$

є від'ємною, отже, обрахування місця й величини викривлення за виразами (6), (7) призведе до невірних результатів.

Тоді, з урахуванням властивостей операцій у лишкових класах, для визначення місця та величини викривлення замість системи нерівностей (7)

$$\{k \cdot p_i\}_q < p_i$$

слід скористатися системою нерівностей:

$$\{q - \{k \cdot p_i\}_q\}_q < p_i. \tag{11}$$

У разі вірності однієї із цих нерівностей, правомочним є висновок про те, що

$$\gamma = \{(k - 1) \cdot P\}_q,$$

а отже

$$k = \{\gamma / \{P\}_q\}_q + 1. \tag{12}$$

Як видно з рис. 2, у цьому разі величина викривлення $l_i \cdot R_i > (k - 1) \cdot P$. Тоді величина сформованого в ході нулевізації числа T є більшою вихідного числа A_2 на величину $[l_i \cdot R_i - (k - 1) \cdot P]$, тобто:

$$T = A' - (k - 1) \cdot P = A_1 + [l_i \cdot R_i - (k - 1) \cdot P] < A_1. \tag{13}$$

Останній вираз може бути представленим у вигляді:

$$T = A' - k \cdot P = A_1 - [(k - 1) \cdot P - l_i \cdot R_i].$$

Неважко помітити, що вирази (9) та (13) є тотожними, якщо вважати, що номер діапазону в обох випадках має значення k. І, хоча значення викривлення при цьому

$$\Delta A' = -(k \cdot P - l_i \cdot R_i),$$

величина скорегованого числа повинна визначатись, як і раніше, з виразу:

$$A_1 = T + ((k-1) \cdot P - l_i \cdot R_i).$$

Тобто скореговане значення символу α_i обчислюється як:

$$\alpha_i = \{\tilde{\alpha}_i + \Delta\alpha_i\} = \{T + ((k-1) \cdot P - l_i \cdot R_i)\} \bmod p_i = \{\tilde{\alpha}_i - \{l_i \cdot R_i\} \bmod p_i\} \bmod p_i,$$

або з урахуванням (8) та згідно з виразом (10) отримаємо:

$$\alpha_i = \{\tilde{\alpha}_i - \{[k \cdot p_i / q] \cdot R_i\} \bmod p_i\} \bmod p_i.$$

***Приклад* 4.** Нехай у СЛК із основами 2, 3, 5, 17 вихідне число $0_{10} = 0, 0, 0, 0$ внаслідок викривлення перетворилося на $0, 0, 2, 0 = 102_{10}$.

Результат нулевізації дає:

$$\Gamma = 0, 0, 0, 5, \gamma = 5,$$

звідки

$$k = \{\gamma / \{P\}_q\}_q = \{5/13\}_{17} = (5 + 2 \cdot 17)/13 = 39/13 = 3.$$

Пошук місця викривлення дає:

$$\{k \cdot p_i\}_q < p_i,$$

для $k = 3$:

$$\{3 \cdot 5\}_{17} = 15 > 5,$$

а величина $\{q - \{k \cdot p_i\}_q\}_q$ при $(k-1) = 3$ має значення, яке є меншим величини основи p_3: $17 - 5 = 2 < 5$, тобто, як і в попередньому прикладі, виявленим є викривлення по основі p_3, але з урахуванням (12) при $(k-1) = 3$, чи при $k = 4$.

Розрахунок скорегованого лишку по основі p_3:

$$\alpha_3 = \{2 - \{[20/17] \cdot 42\} \bmod 5\} \bmod 5 = 2 - 2 = 0.$$

Видно, що корекція викривлення здійснена правильно.

Отже алгоритм контролю та поновлення цілісності з використанням процедури нулевізації зводиться до послідовного виконання наступних операцій:

1) здійснення нулевізації прийнятого (зчитаного із запам'ятовуючого пристрою) інформаційного об'єкта з визначенням величин γ та k;

2) перевірки для кожної з основ p_i справедливості хоча б однієї з нерівностей:

$$\{k \cdot p_i\}_q < p_i,$$

чи

$$\{q - \{k \cdot p_i\}_q\}_q < p_i,$$

а отже, визначення таким чином того значення p_i, а отже й того номера i, де є викривлення символу α_i, та уточнення значення величини k. У разі справедливості першої із цих нерівностей попередньо обраховане значення k є правильним, а при справедливості другої із цих нерівностей попередньо обраховане значення k слід збільшити на одиницю;

3) здійснення корекції визначеного викривлення:

$$\alpha_i = \{\widetilde{\alpha}_i - \{[k \cdot p_i / q] \cdot R_i\} \bmod p_i\} \bmod p_i.$$

Приклад 5. Нехай для умов прикладів 1, 2 потрібно декодувати з використанням останнього алгоритму код числа A = 51 = 11.01.10.0110011, у якому викривлена третя пара розрядів так, що $\widetilde{A}$ = 11.01.01.0110011. Врахуємо, що перша операція вже виконана (приклад 2), і отримані наступні результати: γ = 51 та k = 10, а також те, що R_1 = 2485; R_2 = 2088; R_3 = 1420; $R_4 = P$ = 140, m_1 = 1; m_2 = 2; m_3 = 6; m_4 = 35 (рис. 3).

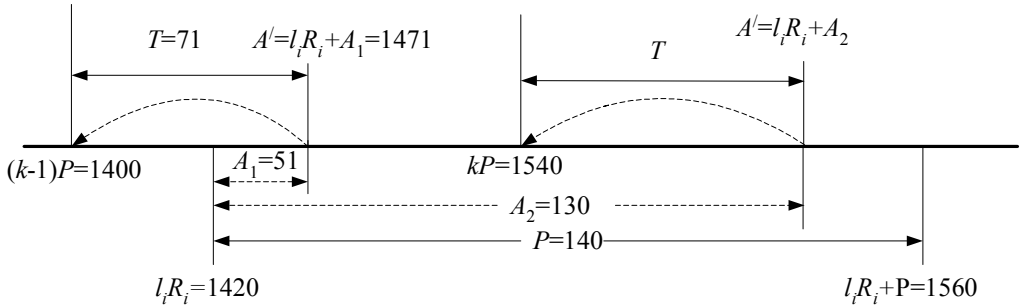

Рис. 3. Ілюстрація процесів нулевізації для прикладів 5, 6

Тоді друга операція щодо перевірки для кожної з основ p_i справедливості хоча б однієї з нерівностей:

$$\{k \cdot p_i\}_q < p_i,$$

чи

$$\{q - \{k \cdot p_i\}_q\}_q < p_i,$$

зведеться до наступного.

Для першої з основ $p_1 = 4$:

$\{k \cdot p_i\}_q = \{10 \cdot 4\}_{71} = 40 < p_i = 4$ — нерівність не є справедливою;

$\{q - \{k \cdot p_i\}_q\}_q = \{71 - 40\}_{71} = 31 < 4$ — нерівність не є справедливою.

Для другої з основ $p_2 = 5$:

$\{k \cdot p_i\}_q = \{10 \cdot 5\}_{71} = 50 < 5$ — нерівність не є справедливою;

$\{q - \{k \cdot p_i\}_q\}_q = \{71 - 50\}_{71} = 21 < 5$, — нерівність не є справедливою.

Для третьої з основ $p_3 = 7$:

$\{k \cdot p_i\}_q = \{10 \cdot 7\}_{71} = \{70\}_{71} < 7$ — нерівність не є справедливою;

$\{q - \{k \cdot p_i\}_q\}_q = \{71 - 70\}_{71} = 1 < 7$ — **нерівність є справедливою**.

Надалі потрібно уточнення значення величини k. У разі справедливості першої із цих нерівностей попередньо обраховане значення k є правильним, а при справедливості другого — слід уточнити значення величини k. Отже, попередньо обраховане значення k слід збільшити на одиницю, отже, номер інтервалу набуває значення $k = 11$.

Таким чином, виявлено наявність викривлення в лишку по третій основі. Тому третьою операцією є корегування викривлення у відповідності з виразом:

$$\alpha_i = \{\widetilde{\alpha}_i - \{[k \cdot p_i / q] \cdot R_i\} \bmod p_i\} \bmod p_i = \{1 - \{[11 \cdot 7 / 71] \cdot 1420\}_7\}_7 =$$
$$= \{1 - \{[1{,}084] \cdot 1420\}_7\}_7 = \{1 - 6\}_7 = 2_{10} = 10_2.$$

Порівнявши отримане значення з вихідним, не викривленим (умови прикладу 1), упевнюємося в тому, що корекція здійснена вірно.

Приклад 6. Нехай для умов прикладу 5 потрібно декодувати з використанням останнього алгоритму код числа $A = 130 = 10.00.100.0111011$, у якому викривлена третя пара розрядів так, що $\widetilde{A} = 10.00.11.0110001$. Неважко впевнитися в тому, що результатом першої операції є наступні результати: $\gamma = 14$ та $k = 11$.

Тоді друга операція щодо перевірки для кожної з основ p_i справедливості хоча б однієї з нерівностей:

$$\{k \cdot p_i\}_q < p_i,$$

чи

$$\{q - \{k \cdot p_i\}_q\}_q < p_i,$$

зведеться до наступного.

Для першої з основ $p_1 = 4$:

$\{k \cdot p_i\}_q = \{11 \cdot 4\}_{71} = 44 < p_i = 4$ — нерівність не є справедливою;

$\{q - \{k \cdot p_i\}_q\}_q = \{71 - 44\}_{71} = 27 < 4$ — нерівність не є справедливою.

Для другої з основ $p_2 = 5$:

$\{k \cdot p_i\}_{\mathrm{q}} = \{11 \cdot 5\}_{71} = 55 < 5$ — нерівність не є справедливою;

$\{q - \{k \cdot p_i\}_q\}_q = \{71 - 55\}_{71} = 16 < 5$ — нерівність не є справедливою.

Для третьої з основ $p_3 = 7$:

$\{k \cdot p_i\}_q = \{11 \cdot 7\}_{71} = 6 < 7_i$ — **нерівність є справедливою**.

Оскільки справедливим є перша з нерівностей, попередньо обраховане значення $k = 11$ є правильним.

Таким чином, виявлено наявність викривлення в лишку по третій основі. Тому третьою операцією є корегування викривлення у відповідності з виразом:

$$\alpha_i = \{\widetilde{\alpha}_i - \{[k \cdot p_i / q] \cdot R_i\} \bmod p_i\} \bmod p_i = \{3 - \{[11 \cdot 7 / 71] \cdot 1420\}_7\}_7 =$$
$$= \{3 - \{[1{,}084] \cdot 1420\}_7\}_7 = \{3 - 6\}_7 = 4_{10} = 100_2.$$

Порівнявши отримане значення з вихідним, не викривленим (див. умови прикладу), упевнюємося в тому, що корекція здійснена вірно.

Перевагами алгоритму нулевізації є:

1) усі операції алгоритму здійснюються над числами з кінцевою, наперед відомою розрядністю;

2) результат операцій залежить від чисел із заданою розрядністю, що при кодуванні–декодуванні завжди приводить до правильних наслідків.

Недоліками цього алгоритму є:

1) обчислення величини T здійснюється послідовно, оскільки значення наступної величини t_i може бути визначеним лише при відомих попередніх t_i ($i = 1, 2, \ldots, i - 1$), що виключає можливість розпаралелювання процесу. За рахунок цього слід очікувати меншої швидкодії цього алгоритму порівняно з попереднім;

2) достатньо велика розрядність мінімальних чисел t_i. Якщо вважати, що кодуванню–декодуванню підлягають числа (блоки), які складаються з n лишків по b розрядів кожен, то перше мінімальне число потребує $N = n \cdot b$ розрядів, друге — $(N - b) = b \cdot (n - 1)$,…, i-е — $(N - (i - 1) \cdot b) = b \cdot (n - i + 1)$, … розрядів, що потребує наявності запам'ятовуючих пристроїв з відповідною ємністю.

Таким чином, у статті запропоновано використання завадостійкого корегуючого коду та відповідні алгоритми кодування–декодування для застосування в задачах контролю, чи контролю та поновлення цілісності інформаційних об'єктів в умовах пакетних викривлень. Здійснено їхній аналіз, показані переваги та недоліки.

1. НД ТЗІ 2.5-005-99 «Класифікація автоматизованих систем і стандартні функціональні профілі захищеності оброблюваної інформації від несанкціонованого доступу».

2. *Дубровский В.В.* CDMA — взгляд глазами профессионала. — mailto:v_dubrovskii@mail.ru.

3. *Василенко В.С., Будько М.М., Короленко М.П.* Механізми контролю цілісності інформації та її поновлення // Правове, нормативне та метрологічне забезпечення Системи захисту інформації в Україні. — К.: НТУУ «КПІ». — 2000. — С. 130–139.

4. *Акушский И.Я., Юдицкий Д.И.* Машинная арифметика в остаточных классах. — М.: Сов. радио, 1966. — 421 с.

Надійшла до редакції 21.08.2006