Литература: ***1.*** *The CERT Coordination Center (CERT/CC), a center of Internet security expertise, http://www.cert.org/stats/cert_stats.html, 2005.* ***2.*** *Ю. Машевский, Перелом в развитии вредоносных программ, http://www.viruslist.com/ru/analysis.pubid=166969803, 29. 07. 2005.* ***3.*** *Lindqvist, Ulf & Porras, Phillip A. Detecting Computer and Network Misuse Through the Production-Based Expert System Toolset (P-BEST). Proceedings of the 1999 IEEE Symposium on Security and Privacy. Oakland, CA, May 9 – 12, 1999.* ***4.*** *Mukherjee, Biswanath; Heberlein, L. Todd; & Levitt, Karl N. (University of California, Davis). Network Intrusion Detection. IEEE Network 8, 3 (May/June 1994): 26 - 41.* ***5.*** *Vigna, Giovanni & Kemmerer, Richard A. (University of California, Santa Barbara). NetSTAT: A Network-Based Intrusion Detection Approach. Proceedings of the 14th Annual Computer Security Applications Conference. Scottsdale, AZ, Dec. 1998.* ***6.*** *Paxson, Vern. (Lawrence Berkeley National Laboratory). Bro: A System for Detecting Network Intruders in Real-Time, Proceedings of 7th USENIX Security Symposium. San Antonio, TX, January 1998.* ***7.*** *J. P. Anderson, Computer Security Threat Monitoring and Surveillance, James P. Anderson Co., Fort Washington, Pa. 1980.* ***8.*** *D. E. Denning, An Intrusion Detection Model, IEEE Trans. Software Eng., vol. 13, no. 2, Feb. 1987.* ***9.*** *H. S. Javitz, A. Valdes, D. E. Denning, and P. G. Neumann. Analytical Techniques Development for a Statistical Intrusion Detection System (SIDS) Based on Accounting Records. Technical report, SRI International, Menlo Park, California, July 1986.* ***10.*** *T. F. Lunt, J. van Horne, and L. Halme. Automated Analysis of Computer System Audit Trails. In Proceedings of the Ninth DOE Computer Security Group Conference, May 1986.* ***11.*** *Lunt, Teresa F., et al. (SRI International). A Real-Time Intrusion Detection Expert System (IDES).****12.*** *An Application of Pattern Matching in Intrusion Detection. Sandeep Kumar, Eugene H.Spafford, 1994.* ***13.*** *Temporal Sequence Learning and Data Reduction for Anomaly Detection Terran Lane, Carla E. Brodley, 1998.* ***14.*** *Machine Learning Techniques for the Domain of Anomaly Detection for Computer Security Terran Lane, 1998.* ***15.*** *The Base-Rate Fallacy and its Implications for the Difficulty of Intrusion Detection Stefan Axelson, 1999.* ***16.*** *S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff. A sense of self for UNIX processes. In Proceedings of the 1996 IEEE Symposium on Security and Privacy, pages 120 – 128, Los Alamitos, CA, 1996. IEEE Computer Society Press.* ***17.*** *S. A. Hofmeyr, S. Forrest, and A. Somayaji. Intrusion detection using sequences of system calls. Journal of Computer Security, 6:151 – 180, 1998.* ***18.*** *M. Damashek. Gauging similarity with n-grams: Languageindependent categorization of text. Science, 267:843–848, Feb. 1995.* ***19.*** *Anderson, Debra, et al. (SRI International). Detecting Unusual Program Behavior Using the Statistical Component of the NextGeneration Intrusion Detection Expert System (NIDES) (SRICSL-95 -06). Menlo Park, CA: Computer Science Laboratory, SRI International, May 1995.* ***20.*** *W. Lee and S. J. Stolfo. Data mining approaches for intrusion detection. In Proceedings of the 7th USENIX Security Symposium, 1998.* ***21.*** *R. C. Carrasco and J. Oncina. Learning stochastic regular grammars by means of a state merging method. In Proceedings of the Second International ICGI Colloquium on Grammatical Inteference and Applications, pages 139 – 152, Alicante, Spain, 1994.* ***22.*** *D. Ron, Y. Singer, and N. Tishby. The power of amnesia: Learning probabilistic automata with variable memory length. Machine Learning, 25, 1996.* ***23.*** *L. R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE, 77(2):257 – 286, 1989.* ***24.*** *А. Новиков, С. Кащенко, Моделирование поведения программного обеспечения с точки зрения безопасности, Правове, нормативне та метрологічне забезпечення систем захиту інформації в Україні, вип. 2, 2001 р.* ***25.*** *Политика безопасности. Краткий обзор защитных политик, http://www.securitylab.ru/analytics/216207.php, 06. 06. 2002.* ***26.*** *А. Астахов, Разработка эффективных политик информационной безопасности, http://www.osp.ru/cio/2004/01/070.htm, 30. 01. 2004.* ***27.*** *D. Wagner, D. Dean, Intrusion Detection via Static Analysis, Proceedings of the 2001 IEEE Symposium on Security and Privacy, May 14 – 16. 2001, p. 156 – 169.* ***28.*** *Wahbe, Lucco, Anderson, Efficient Software-Based Fault Isolation, Proceedings of the Symposium on Operating System Principles, 1993.* ***29.*** *William E. Perry. Effective Methods for Software Testing, Second Edition, Wiley, 2001.*

УДК 681.3

УЗАГАЛЬНЕНІ ЗАВАДОСТІЙКІ КОДИ В ЗАДАЧАХ ЗАБЕЗПЕЧЕННЯ ЦІЛІСНОСТІ ІНФОРМАЦІЙНИХ ОБ'ЄКТІВ В УМОВАХ ПРИРОДНИХ ВПЛИВІВ

Вячеслав Василенко
Національний авіаційний університет

Анотація: **Для використання в задачах забезпечення цілісності інформаційних об'єктів в умовах природних впливів пропонуються узагальнені завадостійкі коді.**

Summary: **For the use in the tasks of providing of integrity of information's on lding object in the conditions of influence of natural factors the generalized antigambling codes are offered.**
Ключові слова: **Виявлення викривлень, виправлення викривлень, контроль цілісності, завадостійкі корегуючі коди.**

І Задачі захисту цілісності інформаційних об'єктів телекомунікаційних мереж

Відповідно до термінології нормативних документів Департаменту спеціальних телекомунікаційних систем і захисту інформації Служби безпеки України [1] під цілісністю інформації розуміється її властивість, яка полягає у тому, що інформація не може бути модифікована неавторизованим користувачем або процесом. Іншими словами, під цілісністю інформації розуміється відсутність в ній будь-яких викривлень (модифікацій), які не були санкціоновані її власником, незалежно від причин або джерел виникнення таких викривлень. Під доступністю інформації розуміється властивість інформації, що полягає у тому, що вона знаходиться у вигляді, необхідному користувачу (процесу), в місці, необхідному користувачу (процесу), і в той час, коли вона йому необхідна. Звідси можна зробити висновок про те, що *порушення цілісності* інформації неминуче приводить і до *порушення її доступності*.

Викривлення інформації, тобто порушення її цілісності, можливі на будь-якому етапі її циркуляції у 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 і елементів електричних ланцюгів, порушення реєструючого шару магнітних або оптичних носіїв 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 в цифровому представленні інформації, незалежно від використаної системи числення або форми представлення інформації і, в цьому значенні, є загрозами функціональним властивостям захищеності інформаційних ресурсів – їх цілісності і доступності. Надалі розглядаються проблеми забезпечення цілісності інформаційних об'єктів в умовах природних впливів.

Наслідком природних впливів в каналах телекомунікаційних мереж (ТКМ) є зменшення співвідношення сигнал/шум (сигнал/завада). Як відомо, це відношення визначає вірність інформації, яка визначиться, наприклад, через ймовірність помилок двійкових символів (біт) $P_{пом}$, а також інтенсивність цих помилок. Слід підкреслити, що інтенсивність природних дій в каналах деяких ТКМ, яка визначається, в основному, цим співвідношенням, є настільки значною, що лише за їх рахунок, без урахування можливостей зловмисників зі створення загроз, наприклад різного роду завад, середня вірогідність помилки двійкового символу (біта) $P_{пом}$ для телефонних кабельних каналів ТКМ складає від $1{,}29\times10^{-4}$ до 2×10^{-3}; для радіорелейних телефонних – від $2{,}66\times10^{-4}$ до $7{,}3\times10^{-4}$ відповідно. Відомо також, що з 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 тягне за собою наслідки (часто надзвичайно важкі) для власників або користувачів цієї інформації. Тому задача забезпечення цілісності і доступності інформаційних ресурсів є однією з найактуальніших при розробці і експлуатації АС і їх елементів. Ця необхідність підтверджується і вимогами щодо допустимої вірогідності $P_п$ помилок в повідомленнях, яку слід трактувати як вірогідність порушення цілісності інформаційних об'єктів, які обробляються (якщо передача і 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).

Використовуючи зв'язок між допустимою вірністю інформації $P_п$ і середньою вірогідністю помилки символу $P_{пом}$ ($P_{пом} = P_п /n$, де n – розрядність повідомлення), одержимо, що середня допустима вірогідність викривлення символу, у разі відсутності засобів захисту від помилок, складе значення від 10^{-8} до 10^{-6} – при обробці повідомлень з довжиною 100 двійкових символів (біт), і від 10^{-9} до 10^{-7} – при обробці повідомлень з довжиною 1000 символів. Слід підкреслити, що ці вимоги не є найвищими. Наприклад, в системах з криптографічним захистом наслідки викривлень рівноцінні втраті повідомлення, не говорячи вже про втрати, які пов'язані з навмисним викривленням фінансової інформації, навмисним викривленням команд, розпоряджень і т. п. Останнє іноді приводить до постановки задачі передачі інформації з абсолютною цілісністю. Аналіз цих даних дозволяє зробити висновок про необхідність використовування механізмів, засобів або алгоритмів, які дозволяють забезпечити значне підвищення цілісності інформації, яка приймається.

Для забезпечення контролю та поновлення цілісності інформаційних об'єктів, включаючи і 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 з 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 по контрольних ознаках найчастіше не існує). Контроль цілісності (на відсутність викривлень) зводиться при цьому до тих або інших процедур перевірки наявності вказаного регулярного (функціонального) одностороннього зв'язку між ознаками цілісності і прийнятої з каналу зв'язку (або зчитаної із запам'ятовуючого (ЗП) пристрою) інформацією.

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

Характерною особливістю випадкових викривлень є те, що вони, через їх хаотичність, відсутність навмисності, порушують регулярний (функціональний) односторонній зв'язок між прийнятою (або зчитаною із ЗП) інформацією і ознаками цілісності, сформованими перед передачею (перед записом в ЗП). Тому при виявленні порушення вказаного зв'язку встановлюється факт наявності таких викривлень, а за певних умов, і їх місця та величини (характер). 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 одержувачем (або зчитаною з ЗП), і ознаками цілісності. З цією метою порушник, використовуючи знання процедур формування контрольних ознак після необхідної для його цілей модифікації початкової інформації перед передачею одержувачу (перед записом в ЗП), забезпечує формування відповідних ознак. При успішному формуванні вказаних ознак розкрити наявність модифікації неможливо. Для боротьби з 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 (ключів перетворення), порушник не зуміє забезпечити, зімітувати наявність регулярного (функціонального) зв'язку між модифікованою ним початковою інформацією, прийнятою (або зчитаною із ЗП), і ознаками цілісності.

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

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

• завади, викликані зовнішніми джерелами і атмосферними явищами.

Завади − це електричні збурення, що виникають в самій апаратурі або потрапляють в неї ззовні. Найпоширенішими є флуктуаційні (випадкові) завади. Вони є послідовністю імпульсів, що мають випадкову амплітуду і виникають один за одним через різні проміжки часу. Прикладами таких завад можуть бути атмосферні і індустріальні завади. У приймачі завади можуть настільки ослабити інформаційний сигнал, що він або взагалі не буде виявленим, або буде викривленим так, що "одиниця" може перейти в "нуль" і навпаки.

Труднощі боротьби з завадами полягають в безладності, нерегулярності і в структурній схожості завад з інформаційними сигналами. Тому захист інформації від викривлень і шкідливого впливу завад має велике практичне значення і є однією з серйозних проблем сучасної теорії і техніки зв'язку.

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:

- 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);
- збільшення співвідношення сигнал/завада за рахунок зниження рівня завад (шумів) шляхом використання спеціальних, кабельних ліній зв'язку з низьким рівнем власних шумів, наприклад,

оптоволоконних, що також вимагає значних матеріальних витрат, і може бути реалізованим лише в таких лініях зв'язку;

- забезпечення хоча б задовільної узгодженості смуги пропускання П каналу із спектром сигналу, який визначається параметрами сигналу, в першу чергу його тривалістю $\tau \approx 1/B$, де τ − тривалість сигналу, а В − технічна швидкість передачі інформації (швидкість посимвольної передачі) в даному каналі; задовільною найчастіше вважають таку узгодженість, коли $П \geq 2B$;
- 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 і 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 можливість каналу зв'язку (у 3...5 раз), а час затримки передавання інформаційних об'єктів може стати неприпустимо великим. З цих причин в системах передачі даних (СПД) використання цих методів не завжди доцільне;
- застосування різного роду завадостійких кодів з виявленням помилок в прийнятій (зчитаній) інформації, які дозволяють реалізувати програмні, апаратурні або програмно-апаратурні засоби виявлення викривлень; це, в свою чергу дає можливість застосування передачі повідомлень з різного роду зворотним зв'язком (інформаційним – деякий аналог мажоритарного методу з багатократною передачею інформації і зворотним прийомом і ухваленням рішення щодо правильності передачі на стороні передавача, або з вирішальним зворотним зв'язком (ВЗЗ) – багатократній, при 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;
- застосування різного роду завадостійких корегуючих кодів (ЗКК), які дозволяють реалізувати програмні, апаратурні або програмно-апаратурні засоби виявлення і усунення викривлень.

Останній із способів (механізмів) забезпечення цілісності інформаційних об'єктів − із застосуванням завадостійких корегуючих кодів – наразі знайшов широке застосування в стандартах радіо зв'язку, наприклад, стільникового. Він не потребує зворотного каналу і забезпечує, як правило, прийнятне значення часу затримки передавання інформаційних об'єктів. Тому, чи не єдиною проблемою в цих та інших ТКМ з 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, в стандартах CDMA базовий цифровий потік розбивається на пакети тривалістю по 20 мс и подається на згорточний кодер с половинною швидкістю [2]. При цьому тривалість пакету викривлень може бути порівняною чи, навіть, значно більшою. тривалості інформаційного пакету, що може суттєво вплинути на результативність процедур інформаційного обміну.

Як вихід із таких ситуацій може розглядатися можливість збільшення тривалості інформаційних пактів із одночасним застосуванням перемежування потрібної глибини та завадостійких корегуючих кодів, які були б спроможними забезпечити виявлення та виправлення пакетів викривлень значної тривалості. Як такі коди в статті пропонуються узагальнені завадостійкі корегуючі коди.

II Узагальнені завадостійкі коди. Лишкові-Хеммінгові та лишкові-матричні коди

Під узагальненими розумітимемо коди, призначені для виявлення (виявлення і виправлення) пакетних викривлень з кратністю b, в яких використовуються алгоритми кодування і декодування щодо узагальнених b – розрядних символів.

В цих кодах початкова двійкова кодова послідовність − базове кодове слово (БКС) $I_1\ I_2\ \ldots\ I_k$ розбивається на n = k/b груп двійкових розрядів з розрядністю b, в яких передбачається виявлення та виправлення викривлень:

$$\underbrace{I_1 \ldots\ldots\ldots I_b}_{1\text{ – а група}},\ \underbrace{I_{b+1} \ldots\ldots\ldots I_{2b}}_{2\text{ – а група}} \ldots\ldots\ldots\ldots \underbrace{I_{k-b+1} \ldots\ldots I_k}_{n\text{ – а група}}.$$

Двійкові символи, що входять в одну b − розрядну групу, розглядаються як b – значний узагальнений символ, який може приймати будь-яке із s значень від 0 до (s – 1), де

$$s = 2^b.$$

При кодуванні та декодуванні операції над узагальненими символами пропонується виконувати за деяким модулем, тобто розшукувати лишок від ділення результату операції на деякий модуль. Це дало автору можливість (в разі застосування алгоритмів, які можуть бути аналогічними відповідним алгоритмам двійкових кодів, але по відношенню до узагальнених символів) для відмінності відповідних узагальнених кодів від двійкових ввести в їх назву слово "лишок", тобто говорити про лишкові-Хеммінгові (ЛХ), лишкові-матричні (ЛМ), лишково-згорточні (ЛЗ) чи лишкові-ланцюгові (ЛЛ) та інші коди.

Принципи побудови та застосування таких кодів розглянемо на прикладі лише деяких із таких кодів. При потребі читач самостійно може застосувати викладені підходи і до інших кодів цього класу.

У лишково-Хеммінгових кодах двійкові базові кодові слова, розбиті на b-розрядні узагальнені символи, записуються у вигляді $\alpha_1, \alpha_2, \ldots \alpha_n$, де $\alpha_i \le 2^b - 1$, а $N = b \cdot n$. Так само, як і в двійковому коді Хеммінга (класична форма запису коду) узагальнені символи α_i з номерами $i = 2^j$ ($j = 0, 1, \ldots$) є перевірочними, решта символів – інформаційні. Причому для отримання перевірочних символів при кодування використовується алгоритм, аналогічний алгоритму для двійкового коду Хеммінга, але відносно узагальнених символів. При цьому усі необхідні для кодування і декодування операції здійснюються за деяким модулем. Тобто, в ЛХ-коді для отримання першого перевірочного символу необхідно скласти за деяким модулем (одержати лишки від суми) всі узагальнені символи базового кодового слова, що мають в коді свого номера одиницю в першому (молодшому) розряді; для отримання другого перевірочного символу – скласти за модулем усі символи, що мають в коді свого номера одиницю в другому розряді і т. д.

Як модуль для отримання контрольних символів досить зручно використовувати величину $s = 2^b$, тобто

$$\alpha_1 = \{\alpha_3 + \alpha_5 + \alpha_7 + \ldots\ldots\ldots\ldots\}_s,$$
$$\alpha_2 = \{\alpha_3 + \alpha_5 + \alpha_6 + \alpha_7 + \ldots\ldots\ldots\}_s,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

При такому значенні модуля потрібна розрядність перевірочних символів не відрізняється від розрядності узагальнених символів b.

При декодуванні зберігається той же алгоритм розрахунку перевірочних αi символів, що і при кодуванні, але при додаванні за модулем використовуються і контрольні символи. Знов одержані перевірочні символи порівнюються з відповідними перевірочними символами, обчисленими при кодуванні. При їх відповідності робиться висновок про відсутність викривлення, в решті випадків – про наявність викривлення.

Якщо приписати результатам відповідності значення 0, а результатам не відповідності – значення 1, то одержана сукупність нулів і одиниць утворює код, який також, як і в двійковому коді Хеммінга, є номером викривленого символу.

Приклад. Хай необхідно закодувати ЛХ-кодом восьмирозрядну (N = 8) послідовність 10001101. Якщо код орієнтований на виправлення двократних викривлень, то b = 2, кількість узагальнених символів n = N/b = 4. Як модуль для отримання контрольних символів використаємо величину s = 4. Відомо, що в коді Хеммінга при n = 4 потрібно три перевірочні символи $\alpha_1, \alpha_2, \alpha_4$, а інформаційними символами є $\alpha_3 = 10$, $\alpha_5 = 00$, $\alpha_6 = 11$, $\alpha_7 = 01$ Для отримання першого перевірочного символу складемо за модулем чотири $\alpha_3, \alpha_5, \alpha_7$.

$$\alpha_1 = \{\alpha_3 + \alpha_5 + \alpha_7\}_4 = 11$$

Аналогічно цьому

$$\alpha_2 = \{\alpha_3 + \alpha_6 + \alpha_7\}_4 = 10,$$
$$\alpha_4 = \{\alpha_5 + \alpha_6 + \alpha_7\}_4 = 00.$$

Після кодування одержано код

$$11.\ 10.\ 10.\ 00.\ 00.\ 11\ 01 = \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7,$$

який повинен бути записаним у запам'ятовуючий пристрій (ЗП), переданим в канал зв'язку і т. д.

Хай зчитаний або прийнятий з каналу зв'язку код має викривлення у п'ятій групі:

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \acute{\alpha}_5, \alpha_6, \alpha_7 = 11.10.10.00.01.11.01.$$

Після обчислення нових контрольних символів, одержимо

$$\alpha_1 = \{\alpha_3 + \acute{\alpha}_5 + \alpha_7\}_4 = 00$$
$$\alpha_2 = \{\alpha_3 + \alpha_6 + \alpha_7\}_4 = 10,$$
$$\alpha_4 = \{\acute{\alpha}_5 + \alpha_6 + \alpha_7\}_4 = 01.$$

Результати порівняння дадуть код 101, оскільки перший а третій перевірочні символи не співпадають. Це свідчить про виявлення помилки в п'ятому символі, що і було насправді.

Неважко визначити і величину викривлення. Дійсно, будь-який з перевірочних символів, наприклад α_i, при викривленні деякого інформаційного, наприклад α_j, що приймає участь у формуванні символу α_i має величину

$$\acute{\alpha}_i = \{\alpha_c + \alpha_d + \ldots + \{\alpha_j + \Delta\alpha_j\} + \ldots\}_s = \{\alpha_i + \Delta\alpha_j\}_s . \quad (1)$$

Звідки

$$\Delta\alpha_j = \{\acute{\alpha}_i - \alpha_i\}_s. \quad (2)$$

Для вищерозглянутого прикладу

$$\Delta\alpha_5 = \{\acute{\alpha}_1 - \alpha_1\}_4 = \{00 - 11\}_4 = 01,$$

або

$$\Delta\alpha_5 = \{\acute{\alpha}_4 - \alpha_4\}_4 = \{01 - 00\}_4 = 01.$$

Знаючи величину ($\acute{\alpha}_i$) і місце викривлення (i), легко здійснити корекцію, оскільки із (2) витікає

$$A_i = \{\acute{\alpha}_i - \Delta\alpha_j\}_s.$$

У нашому прикладі

$$\alpha_5 = \{\acute{\alpha}_5 - \Delta\alpha_5\}_4 = \{01 - 01\}_4 = 00,$$

що і є насправді.

Алгоритм декодування ЛХ-коду може бути спрощеним, якщо при кодуванні замість перевірочних символів αi в записану або передану послідовність записувати величину

$$\Delta\alpha_i = \{s - \Delta\alpha_j\}s.$$

Тоді для вже розглянутого прикладу (α_1 =11, α_2 =10, α_4 = 00) $\Delta\alpha_1$ = 01, $\Delta\alpha_2$ = 10, $\Delta\alpha_4$ = 00 і записувати (передавати) необхідно код:

$$\Delta\alpha_1, \Delta\alpha_2, \alpha_3, \Delta\alpha_4, \alpha_5, \alpha_6, \alpha_7 = 01.\ 10.\ 10.\ 00.\ 00.\ 11.\ 01.$$

Якщо зчитано або прийнято слово з тим же викривленням, що і раніше, тобто

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \acute{\alpha}_5, \alpha_6, \alpha_7, = 01.10.10.00.01.11.01,$$

то після декодування отримаємо

$$\Delta\alpha_1 = \{\alpha_1 + \alpha_3 + \acute{\alpha}_5 + \alpha_7\}_4 = 01,$$
$$\Delta\alpha_2 = \{\alpha_2 + \alpha_3 + \alpha_6 + \alpha_7\}_4 = 00,$$
$$\Delta\alpha_4 = \{\alpha_4 + \acute{\alpha}_5 + \alpha_6 + \alpha_7\}_4 = 01.$$

При цьому, якщо відмінним від нуля перевірочним символам приписати значення 1, а іншим – значення 0, то одержимо код i = 101, що визначає місце викривлення, величина якого дорівнює значенню будь-якого ненульового перевірочного символу. Для розглянутого прикладу величина викривлення $\Delta\alpha_i$ = 01, корекція якого нескладна.

У лишкові-матричних кодах БКС розбивається на узагальнені символи, які зводяться в прямокутну матрицю розмірності $m \times n$ (m стовпців і n рядків) вигляду (див. рис. 1)

$$\begin{Vmatrix} \alpha_1 & \alpha_2 & . & . & . & \alpha_m \\ \alpha_{m+1} & \alpha_{m+2} & . & . & . & \alpha_{2m} \\ . & . & . & . & . & . \\ \alpha_{m(n-1)+1} & \alpha_{m(n-1)+2} & . & . & . & \alpha_{mn} \end{Vmatrix}.$$

Рисунок 1 – Представлення БКС у вигляді матриці

Ця матриця при кодуванні розширюється на один рядок і один стовпець за рахунок перевірочних символів, кожний з яких є доповненням до s суми по модулю s елементів відповідного рядка або відповідного стовпця, при цьому одержують нову розширену матрицю (див. рис. 2), яка записується в ЗП (передається в канал зв'язку).

При декодуванні викривлених БКС ті перевірочні елементи з додаткових рядка і стовпця, які відповідають рядку або стовпцю, що містить викривлені символи, відрізнятимуться від нуля, що дає можливість визначати місце викривлення. Якщо викривленим є тільки один елемент в рядку і стовпці, то ненульове значення відповідних перевірочних символів визначить величину цієї помилки. У цьому значенні можливості ЛМ-коду до відношенню до узагальнених символів повністю співпадатимуть з можливостями з виявлення і виправлення викривлень двійкового матричного коду по відношенню до двійкових символів.

$$\left\|\begin{array}{cccc} \alpha_1 & \alpha_2 & \ldots & \alpha_m & s-\sum_{i=1}^{m}\alpha_i(\mathrm{mod}\,s) \\ \alpha_{m+1} & \alpha_{m+2} & \ldots & \alpha_{2m} & s-\sum_{i=1}^{m}\alpha_{m+i}(\mathrm{mod}\,s) \\ . & . & \ldots & . & . \\ \alpha_{m(n-1)+1} & \alpha_{m(n-1)+2} & \ldots & \alpha_{nm} & s-\sum_{i=1}^{m}\alpha_{m(n-1)+i}(\mathrm{mod}\,s) \\ s-\sum_{k=1}^{n}\alpha_{m(k-1)+1}(\mathrm{mod}\,s) & s-\sum_{k=1}^{n}\alpha_{m(k-1)+2}(\mathrm{mod}\,s) & \ldots & s-\sum_{k=1}^{n}\alpha_{nm}(\mathrm{mod}\,s) & \end{array}\right\|$$

Рисунок 2 – Представлення БКС у вигляді розширеної матриці

Приклад. Хай необхідно закодувати ЛМ-кодом восьми розрядне БКС 10.00.11.10 Для b = 2 і s = 4 одержимо матрицю

$$\left\|\begin{matrix}\alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4\end{matrix}\right\| = \left\|\begin{matrix}10 & 00 \\ 11 & 01\end{matrix}\right\|.$$

Після кодування розширена матриця буде мати вигляд:

$$\left\|\begin{array}{ccc} \alpha_1 & \alpha_2 & s-(\alpha_1+\alpha_2)(\mathrm{mod}\,s) \\ \alpha_3 & \alpha_4 & s-(\alpha_3+\alpha_4)(\mathrm{mod}\,s) \\ s-(\alpha_1+\alpha_3)(\mathrm{mod}\,s) & s-(\alpha_2+\alpha_4)(\mathrm{mod}\,s) & \end{array}\right\| = \left\|\begin{matrix}10 & 00 & 10 \\ 11 & 01 & 00 \\ 11 & 11 & \end{matrix}\right\|.$$

Хай зчитана (прийнята з каналу зв'язку) послідовність, що відповідає наступній матриці (викривлений елемент першого рядка другого стовпчика):

$$\left\|\begin{matrix}10 & 01 & 10 \\ 11 & 01 & 00 \\ 11 & 11 & \end{matrix}\right\|.$$

В результаті декодування шляхом додавання за модулем s усіх елементів (включно із додатковими) відповідних рядків та стовпців, одержуємо матрицю:

$$\left\|\begin{matrix}10 & 01 & 01 \\ 11 & 01 & 00 \\ 00 & 01 & \end{matrix}\right\|,$$

з якої виходить, що викривленим є елемент першого рядка і другого стовпця, а величина викривлення дорівнює 01. Після чого корекція b − розрядного викривлення стає тривіальною.

II Узагальнений завадостійкий код умовних лишків

Ще одним із прикладів узагальнених кодів є код умовних лишків (лишків умовних код, ЛУ-код). Теоретичною основою ЛУ-коду є теорія лишкових класів [3]. З цієї теорії 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, …, n, n – кількість таких основ. Вибір величини n здійснюється з умови, яка викладена нижче. Тоді

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

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 А і його уявленням (3) існує взаємна однозначна відповідність, якщо

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

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 в неї легко вводяться властивості виявлення і виправлення викривлень. Відомо, що якщо ввести ще одну, надлишкову, основу $p_к$, то уявлення A в розширеному діапазоні $R = P\cdot p_к$, у вигляді

$$A = \alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_k, \quad (4)$$

де $\alpha_к$ – лишок по основі $p_к$, має чудову для побудови корегуючих кодів властивість: при $p_k > p_n$ будь-яке викривлення в одному з лишків α_i може бути знайденою, а при $p_к > 2 \cdot p_n \cdot p_{n-1}$, де p_n, p_{n-1} – найбільші із основ, може бути і виправленою. Це означає, що при представленні чисел у вигляді (4) створюється завадостійкий код з можливостями або виявлення викривлень, або і їх корекції.

Такий код має принаймні 2 недоліки. Перший з них пов'язаний з тим, що можливі викривлення знаходяться і виправляються (викривлений символ поновлюється) тільки в тому випадку, якщо викривлений лише один з символів α_i, тобто викривлення повинні бути фіксованими в межах однієї із груп розрядів. Такий недолік є притаманним і будь-якому іншому коду і тому усувається відомими способами − застосуванням перемежування з глибиною не меншою ніж два. Другий недолік пов'язаний з необхідністю роботи з числами в системі числення в залишкових класах. Цей недолік достатньо просто усувається в коді умовних лишків, який вводиться таким чином.

Хай є код деякого числа А (БКС), представленого в будь-якій системі числення, зокрема позиційної, наприклад двійкової. Для визначеності, хай це число А представлено послідовністю з нулів і одиниць. Розіб'ємо цю послідовності певним (у загальному випадку довільним) чином на n узагальнених символів, як і для решти узагальнених кодів.

Як і раніше код кожного i–го узагальненого символу розглядатимемо як s–значний розряд α_i, який може приймати будь-яке з s значень від 0 до (s − 1), де $s = 2^b$. Вважатимемо цей код лишком деякого *умовного числа* А за основою p_i. Оскільки величина α_i, як елемент початкового числа

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

а як лишок від ділення А на p_i

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

то для представлення коду будь-якої групи у вигляді лишку за основою p_i необхідно, щоб виконувалася умова

$$p_i > s - 1,$$

інакше в групу із b розрядів може бути записаним код $\alpha_i \ge p_i$, що в лишкових класах не допустимо.

Приклад. Хай b = 3, s = 7, тоді α_i може приймати значення 000, 001, 010,..., 110. При $p_i = 5$ максимальне значення α_i обмежується кодом 100, тобто коди 101, 110, 111 є "неправильними". Якщо ж взяти $p_i > 7$, наприклад $p_i = 9$, то максимальне значення α_i обмежується не величиною p_i, а розрядністю групи b, тобто $\alpha_{max} = 111$.

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

$$A' = \alpha_1, \alpha_2, \ldots, \alpha_{n1}, \alpha_к \quad (5)$$

розповсюджуються усі можливості СЛК з виявлення і виправлення викривлень, тобто одержаний код (5) має всі властивості коду (4), але останній код може бути отриманим для будь-якої двійкової послідовності, а не тільки щодо чисел в лишкових класах. Відзначимо, що таким чином усунено другий недолік коду (4).

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.

Таким чином, ЛУ-код дозволяє знаходити і виправляти b − розрядні пакети викривлень, згруповані в межах будь-якого з n узагальнених символів і потребує при цьому надмірність біля

$$r \approx 2b + 1$$

двійкових розрядів (оскільки $p_k \approx 2p_n p_{n+1}$, $r = [\log 2p_k] + 1$). В конкретних випадках ця надмірність може відхилятися в ту або іншу сторону, що залежить також від алгоритмів кодування-декодування.

III Алгоритми кодування-декодування ЛУ-коду

Оскільки в основі ЛУ-коду лежать властивості СЛК, то в цьому коді принципово можуть бути використані відомі алгоритми кодування-декодування. До таких алгоритмів відносяться алгоритм нулізації, усі відомі версії якого дозволяють лише виявляти факт наявності викривлень [3], і, так званий Z-алгоритм [4].

В основі цих алгоритмів лежить той факт, що будь-яке викривлення в одній з груп розрядів α_i

переводить початкове число з робочого діапазону [0, P = $\prod_{i=1}^{k} p_i$) в діапазон [P, R = p_k·P), тобто призводить (рис. 3) до збільшення початкового числа $A' < P$ на деяку величину $l_i \cdot R_i$. Тут l_i і $R_i = R/ p_i$ – цілі числа. 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 вид

$$\widetilde{A} = \alpha_1, \alpha_2, \ldots, \widetilde{\alpha}_i, \ldots, \alpha_n, \alpha_к$$

де

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

то це є еквівалентним наступному перетворенню

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

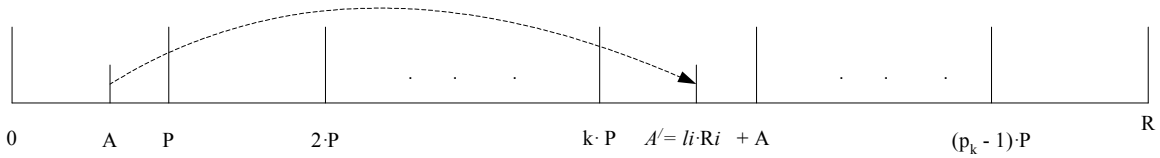

Рисунок 3 – До виходу викривленого числа за межі робочого діапазону

При цьому величина викривлення перевищує величину робочого діапазону P

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

оскільки тільки число виду

$$\Delta A = l_i \cdot R_i = l_i \cdot R/p_i$$

має всі лишки, окрім лишка за основою p_i, такими, що дорівнюють нулю. Але $\Delta A = l_i \cdot R_i > P = R/p_k$ тобто, навіть при $l_i = 1$ величина $R/p_i > R/p_k$ з тієї причини, що $p_k > p_i$.

Відтак, сума $\widetilde{A} = A' + \Delta A > P$, тобто викривлене число вийшло за межі робочого діапазону P і попало в діапазон [P, R).

Згадані алгоритми кодування-декодування як раз і використовують цей факт.

Використання для кодування-декодування алгоритму нулізації

Суть алгоритму нулізація зводиться до того, що як при кодуванні, так і при декодуванні числа

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

за лишками усіх *n* основ, що утворюють робочий діапазон α_i (i = 1, 2, …, n), послідовно формуються так звані мінімальні числа виду

$$t_1 = (\alpha_1, \alpha_2^{/}, \alpha_3^{/}, \ldots, \alpha_n^{/}, \alpha_k^{/}),$$

$$t_2 = (0, (\alpha_2 - \alpha_2^{/}) \pmod{p_2}, \alpha_3^{(2)}, \ldots, \alpha_n^{(2)}, \alpha_k^{(2)}),$$

$$t_3 = (0, 0, (\alpha_3 - \alpha_3^{/} - \alpha_3^{(2)}) \pmod{p_3}, \alpha_4^{(3)}, \ldots, \alpha_n^{(3)}, \alpha_k^{(3)}),$$

……..

$$t_n = (0, 0, 0, \ldots, (\alpha_n - \sum_{j=1}^{n-1} \alpha_n^{(j)}) \pmod{p_n}, \alpha_k^{(n)}),$$

Кожне із 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:

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

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

Підсумок цих чисел $T = \sum_{i=1}^{n} t_i$ має наступні дві властивості [3]. По-перше, лишки цієї суми за всіма основам, окрім p_k, завжди дорівнюють лишкам вихідного числа Ă. По-друге, величина цієї суми завжди є меншою ніж величина робочого діапазону T < P, тобто величина T лежить в межах робочого діапазону і для не викривлених чисел T = *A'*.

Неважко помітити, що *процес отримання величини* T= *A' є процесом кодування вихідного* числа ЛУ-кодом, тобто значення *A'* залежить лише від цього вихідного числа і не залежить від невідомої при кодуванні величини лишку за контрольною основою $p_к$. Цей лишок (контрольна ознака, що розшукується) α_k при цьому дорівнює сумі за модулем $p_к$ усіх проміжних величин $\alpha_k^{(i)}$ (i = 1, 2, ..., n) тобто

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

При декодуванні ж віднімання з числа Ă величини T приводить до того, що отримана різниця

$$\widetilde{A} - T = k \cdot P$$

має за всіма основами, окрім контрольної, лишки, що дорівнюють нулю, а за контрольною

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

тобто при запису в лишкових класах має вид

$$\widetilde{A} - T = (0, 0, \ldots, 0, \ldots 0, (k \cdot P) \ (\mathrm{mod}\, p_k)),$$

де k = 0, 1, 2, …, p_k.

Для не викривлених чисел, тобто при k = 0, величина $\gamma = 0$, для викривлених $\gamma \neq 0$. *Таким чином, установлюється факт наявності чи відсутності викривлень*.

Для ілюстрації можливостей алгоритму розглянемо два приклади.

Приклад 1. Хай необхідно закодувати з використанням алгоритму нулізації вихідний код 110110, вважаючи, що можлива довжина пакета викривлень b = 2. Тоді можливе розбиття вихідного коду на три (n = 3) двох розрядні групи α_1 =11, α_2 =01, α_3 =10, s = 4, а як умовні основи можна вибрати $p_1 = 4$, $p_2 = 5$, $p_3 = 7$. При цьому значення контрольної основи ($p_к > 2 \cdot p_n \cdot p_{n-1} = 2 \cdot 5 \cdot 7 = 70$) можна вибрати $p_к = 71$, що потребує для свого відображення семи розрядів. Внаслідок цього для кодування формується код

A' = 011.001.010.0000000.

Перше мінімальне число t_1 повинно мати лишок за першою основою, що дорівнює $11_{(2)} = 3_{(10)}$. Таким числом є $t_1 = 3$ або при представленні в СЛК з вибраними основами

t_1 = 011.011.011.0000011.

Друге мінімальне число t_2 повинно мати лишок за першою основами, який дорівнює нулю, а за другою

$$(\alpha_2 - \alpha'_2) \ (\mathrm{mod}\, p_2) = (1 - 3) \ (\mathrm{mod}\, 5) = 11_{(2)}.$$

Мінімальним числом, яке має такі лишки за першою і другою основою, є $t_2 = 8$, тобто

t_2 = 000.011.001.0001000.

Трете мінімальне число t_3 повинно мати нульові лишки за першими двома основами, а за третьою

$$(\alpha_3 - \alpha'_3 - \alpha^{(2)}_3) \ (\mathrm{mod}\, p_3) = (2 - 3 - 1) \ (\mathrm{mod}\, 7) = 5 = 101_{(2)}.$$

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

t_3 = 000.000.101.0101000.

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

T = 11.01.10.0110011.

Код T є результатом кодування.

Звернемо увагу на те, що величини t_1, t_2, t_3 формуються послідовно, а також на їх досить велику розрядність.

Приклад 2. Декодувати з використанням алгоритму нулезації для умов наведеного вище прикладу код $\widetilde{A}$ = 11.01.01.0110011, в 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

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

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) (mod 71) = 0, то

$$T = 110110.0000000$$

і

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

Оскільки $\gamma \neq 0$, то робіться висновок про наявність в числі, що декодується, викривлення. Оскільки, при цьому

$$\gamma = (k \cdot P) \pmod{p_k} = (k \cdot 140) \pmod{71},$$

то k = 10, тому що 1400 (mod 71) = 51.

Нескладно упевнитися в тому, що число A = 1471, що декодується, може бути отриманим тільки в наслідок викривлення вихідного числа *A'* = 51 на величину ΔA = 1420 = 000.000.110.0000000, тобто при викривленні третьої групи розрядів на величину $\Delta\alpha_3 = 110$, після чого корекція результату здійснюється просто:

$$\alpha_3 = (\widetilde{\alpha}_3 - \Delta\alpha_3) \pmod 7 = (1 - 6) \pmod 7 = 2 = 10_{(2)}.$$

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

Зрозуміло, однак, що корегування викривлень тим шляхом, який розглянуто вище, тобто шляхом підбирання величини викривлення, є цілком неефективним.

Для виявлення можливостей алгоритму щодо корегування викривлень нагадаємо:

1. відомий факт, що при $p_k > p_n \cdot p_{n-1}$ між величиною викривлення $\Delta\alpha_i$ і величиною γ є взаємно однозначна відповідність, що дає змогу сподіватися в тому, що отримавши γ можна якимось чином визначити місце і величину викривлення, тобто здійснити її виправлення;

2. 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 із робочого діапазону [0, P) в деякий інший піддіапазон. Звернемо увагу на те, що залежно від величини початкового числа (див. рис. 4), викривлене число (A_1 чи A_2) може попасти в один із суміжних діапазонів із номерами k або (k – 1). Зокрема, при

$$A = A1 \leq k \cdot P - li \cdot Ri,$$

це буде (в уже прийнятих позначеннях) діапазон ((k – 1)·P, k·P), тобто діапазон із номером (k – 1), а при

$$A = A2 > k \cdot P - li \cdot Ri$$

це буде діапазон (k·P, (k + 1)·P), тобто діапазон із номером k.

Внаслідок операції нулізації 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 $q = p_k$ одержується результат γ такий, що відповідає лівій межі (див. рис. 4) піддіапазону [(k – 1)·P, k·P), тобто величині (k – 1)·P, або ж такий, що відповідає лівій межі піддіапазону [k·P, (k + 1)·P) , тобто величині k·P.

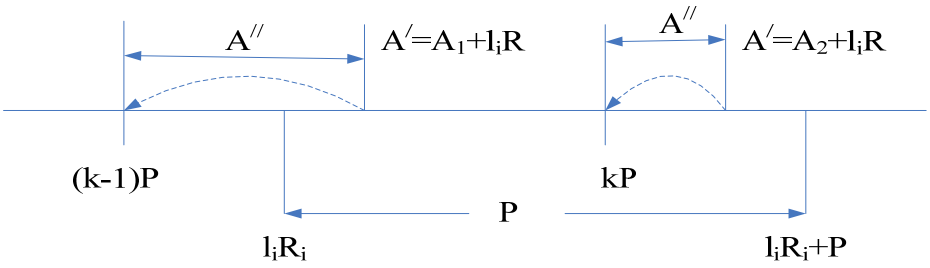

Рисунок 4 – Ілюстрація процесу нулізації

Тобто маємо

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

Звідси, за правилами СЛК, отримаємо

$$k = \{\gamma/\{P\}_q\}_q, \text{ чи } (k-1) = \{\gamma/\{P\}_q\}_q . \quad (6)$$

Тобто, використовуючи вирази (6) завжди можна визначити номер того діапазону, в який потрапило викривлене число та результат нулізації – число k·P. Оскільки величина викривлення $l_i\cdot R_i$ і результат нулізації k·P є близькими, тобто їх різниця є меншою за величину робочого діапазону P, то це надає принципову можливість визначити місце викривлення.

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

В першому випадку, при $k\cdot P > l_i\cdot R_i$ значення R_i, яке характеризує величину і місце викривлення, можна визначити із очевидної нерівності

$$k\cdot P - [k\cdot P/ R_i]\cdot R_i < P . \quad (7)$$

Підставимо в (7) замість R_i його значення у вигляді

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

Тоді

$$k\cdot P - [k\cdot P/ Ri]\cdot Ri = k\cdot P - [k\cdot P\cdot p_i / (P\cdot q)]\cdot P\cdot q/ p_i = k\cdot P - [k\cdot p_i /q]\cdot P\cdot q/ p_i < P.$$

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

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

Помножимо обидві частини останньої нерівності на величину p_i, одержимо:

$$k\cdot p_i - [k\cdot p_i /q]\cdot q < p_i \quad (8)$$

або

$$\{k\cdot p_i\}_q < p_i. \quad (9)$$

Звернемо увагу на те, що в (8, 9) вирази в квадратних дужках є не що інше як величина l_i:

$$[k\cdot P/ R_i] = [k\cdot p_i /q] = l_i \quad (10)$$

Вирази (7) та еквівалентні їм вирази (8, 9) утворюють системи нерівнянь по n (i = 1, 2, …, n) нерівнянь в кожній, в яких справедливим є лише одне е рівнянь для того номера i та значення основи p_i, за якою має місце викривлення.

Таким чином, внаслідок розв'язання будь-якої із систем е рівнянь (7) – (9) щодо змінної p_i *місце викривлення стає виявленим*.

Для визначення ж його величини проаналізуємо величини T = A/ – k·P, чи T = A/ – (k – 1)·P, які формуються за всіма лишками окрім лишку за контрольною основою в ході операції нулізації числа, яке контролюється.

Як видно з рис. 4, вирази (7) – (9) є справедливими в разі, коли величина викривлення $l_i\cdot R_i < k\cdot P$. В цьому випадку величина сформованого в ході нулізації числа T є меншою вихідного числа A_2 на величину $(k\cdot P - l_i\cdot R_i)$, тобто

$$T = A/ - k\cdot P = A2 - (k\cdot P - l_i\cdot R_i) < A2 \quad (11)$$

та

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

а величина скорегованого числа має визначатися як:

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

Тобто величина скорегованого значення лишку:

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

або із урахуванням (10):

$$\alpha_i = \{\tilde{\alpha}_i - \{[k\cdot p_i / q]\cdot R_i\} \bmod p_i\} \bmod p_i. \qquad (12)$$

Приклад. Нехай в 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 нулізації дає

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

Звідки

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

Пошук місця викривлення із

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

для k = 4 дає

$$\{4\cdot 2\}_{17} < 2 \text{ – не вірно},$$
$$\{4\cdot 3\}_{17} < 3 \text{ – не вірно},$$
$$\{4\cdot 5\}_{17} < 5 \text{ – вірно},$$

тобто виявлене викривлення за основою p_3 =5.

Розрахунок скорегованого лишку за основою p_3:

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

Видно, що корекція викривлення здійснена правильно.

В другому випадку, коли результат нулізації – число (k – 1)·P (див. рис. 4) є меншим за величину викривлення $l_i\cdot R_i$, обрахування місця і величини викривлення за виразами (11) – (12) призведе до невірних результатів. Тоді, з урахуванням властивостей операцій в лишкових класах, для визначення місця та величини викривлення слід скористатися виразом:

$$Q - \{(k - 1)\cdot p_i\}q < p_i. \qquad (13)$$

В разі вірності цієї нерівності за однією із основ p_i, правомочним є висновок про те, що

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

а отже

$$K = \{\gamma/\{P\}_q\}_q + 1. \qquad (14)$$

Як видно з рис. 4, в цьому разі величина викривлення $l_i\cdot R_i > (k - 1)\cdot P$. Тоді величина сформованого в ході нулізації числа T є більшою вихідного числа A2 на величину $[l_i\cdot R_i - (k - 1)\cdot P]$, тобто

$$T = A^{/} - (k - 1)\cdot P = A1 + [l_i\cdot R_i - (k - 1)\cdot P] < A1. \qquad (15)$$

Останній вираз може бути представленим у вигляді

$$T = A^{/} - k\cdot P = A1 - [(k - 1)\cdot P - l_i\cdot R_i].$$

Неважко помітити, що вирази (11) та (15) є тотожними, якщо вважати, що номер діапазону в обох випадках має значення – k. І, хоча значення викривлення при цьому

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

величина скорегованого числа має визначатися, як і раніше, з виразу:

$$A1 = T + ((k - 1)\cdot P - l_i\cdot R_i).$$

Тобто величина скорегованого значення лишку:

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

або із урахуванням (12) отримаємо, як і раніше:

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

Приклад. Нехай в 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 нулізації дає

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

За правилами виконання операцій в СЛК,

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

Пошук місця викривлення з

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

для k = 3 дає:

$$\{3\cdot 5\}_{17} = 15 < 5 \text{ – не вірно,}$$

а

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

для (k –1) = 3

$$17 - 5 = 2 < 5 \text{ – вірно,}$$

тобто, як і в попередньому прикладі, виявлене викривлення по основі p_3, але з урахуванням (14) при (k –1) = 3, тобто при k = 4.

Розрахунок скорегованого лишку за основою p_3:

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

Видно, що корекція викривлення здійснена правильно.

IV Використання для кодування - декодування z-алгоритму

Для виявлення викривлень в z-алгоритмі використовується відмічений вище факт, що викривлене число виходить за межі робочого діапазону, тобто

$$\widetilde{A} \geq P. \tag{16}$$

Скористаємось відомим співвідношенням для переводу чисел із СЛК в позиційну систему числення

$$\widetilde{A} = \sum_{i=1}^{i=n+1} \alpha_i B_i - [(1/R) \sum_{i=1}^{i=n+1} \alpha_i B_i] \times R, \tag{17}$$

де: B_i – константа системи числення, її ортогональний базис, причому

$$B_i = R\cdot m_i/p_i, \ (i = 1, 2, \ldots, n + 1), \tag{18}$$

$(n + 1)$ – число умовних основ, включаючи контрольну;
m_i– ціле позитивне число, "вага" ортогонального базису B_i, таке, при якому

$$m_i B_i \ (\text{mod } p_i) = 1.$$

Підставивши вираз (17) в (16) з врахуванням (18), отримаємо

$$\sum_{i=1}^{i=n=1} \alpha_i R\cdot m_i/p_i - [(1/R) \sum_{i=1}^{i=n+1} \alpha_i R\cdot m_i/p_i] \times R > R/p_k. \tag{19}$$

Скоротивши обидві частини (19) на R отримаємо, що в разі наявності викривлень,

$$z > 1/p_к. \tag{20}$$

де

$$Z = \sum_{i=1}^{n+1} \alpha_i m_i / p_i - [\sum_{i=1}^{n+1} \alpha_i m_i / p_i]. \tag{21}$$

Вирази (20 – 21) визначають z – алгоритм декодування для ЛУ-коду, який лише визначає наявність викривлень. Цей алгоритм включає $(n + 1)$ незалежних (при необхідності одночасних) операцій множення коду i-ої групи ($i = 1, \ldots, n+1$) на відповідну константу і потім додавання $(n+1)$ отриманих добутків.

Для побудови алгоритму, здатного не лише визначати наявність, але й виправляти викривлення, скористаємось наступними міркуваннями.
Оскільки викривлення за i – тою основою, як показано вище, має величину $\Delta A = l_i\cdot R_i = l_i\cdot R/p_i$ то очевидним є нерівність

$$\widetilde{A} - l_i\cdot R_i < P, \tag{22}$$

причому величина $l_i\cdot$ визначається з виразу

$$[\widetilde{A} / R_i] = [(A + l_i\cdot R_i) / R_i] = l_i. \tag{23}$$

Тоді з врахуванням (17) – (19, (22) вираз (23) набуде вигляду

$$z\cdot p_i - [z\cdot p_i] < p_i/p_к. \tag{24}$$

Ясно, що вираз (19) і еквівалентний йому вираз (24) справедливі лише для тієї основи p_i, в лишку якої є викривлення. Відтак, вираз (24) дозволяє визначити місце (номер групи), де виникло викривлення. Неважко упевнитися, що величина цього викривлення

$$\Delta\alpha_i = \{[\widetilde{A}/R_i]R_i\}_{p_i} = \{[zp_i]R_i\}_{p_i}.$$

Власне виправлення зводиться до операції

$$\alpha_i = \{\widetilde{\alpha}_i - \Delta\alpha_i\}_{p_i}. \qquad (25)$$

Таким чином, вирази (21), (24), (25) визначають z-алгоритм декодування для корегуючого ЛУ-коду.

Причому, оскільки лишки за будь-якими основам є рівноправними, то все сказане вище відноситься і до контрольної основи. Прийнявши на етапі кодування $\alpha_k = 0$, отримаємо

$$\alpha_k = (p_k - P\cdot[z\cdot p_k]) \pmod{p_k} \qquad (26)$$

і тоді вирази (21), (26) визначають z–алгоритм кодування.

Розглянемо приклади використання z–алгоритму стосовно $p_1 = 4$, $p_2 = 5$, $p_3 = 7$, $p_k = 71$, розрахувавши попередньо константи, які є необхідними для визначення змінних z. Для обраних умов отримаємо: P= 4·5·7 = 140; R = P·p_k = 9940.

При цьому $R_1 = 2485$; $R_2 = 1988$; $R_3 = 1420$; $R_4 = P = 140$, $m_1 = 1$; $m_2 = 2$; $m_3 = 6$; $m_4 = 3$. Позначивши значення m_i/p_i, як g_i отримаємо:

$$g_1 = 0{,}25;\ g_2 = 0{,}4;\ g_3 = 0{,}85714;\ g_4 = 0{,}493257.$$

Приклад. Закодувати повідомлення 11.01.10 з використанням z–алгоритму ЛУ-коду. Приймемо на етапі кодування $\alpha_4 = 0$. Із виразу (21) отримаємо

$$Z =]\alpha1\cdot g1 + \alpha2\cdot g2 + \alpha3\cdot g3 + \alpha4\cdot g4[=]3\cdot 0.25 + 1\cdot 0{,}4 + 2\cdot 0{,}857142 + 0\cdot 0{,}493257[=]2{,}86428[= 0{,}86428,$$

де позначка]x[означає обрахування дробової частини від величини x.

Тоді згідно з (26)

$$\alpha_4 = (p_4 - P\cdot[z\cdot p_4]) \pmod{p_4} = (71 - 140\cdot[0{,}86428\cdot 71]) \pmod{71} = 51_{(10)} = 110011_{(2)}.$$

Приклад. Знайти і виправити викривлення в повідомленні, що використане вище, де

$$\breve{A} = 11.01.01.\ 110011.$$

Тоді

$$Z =]3\cdot 0.25 + 1\cdot 0{,}4 + 1\cdot 0{,}857142 + 51\cdot 0{,}493257[=]27{,}147949[= 0{,}147949.$$

Оскільки, згідно з виразом (20),

$$z = 0{,}147949 > 1/p_к,$$

то робимо висновок про наявність викривлення в наданій кодовій комбінації.

Для виявлення місця викривлення оцінюємо справедливість нерівностей (24).

$$z\cdot p_1 - [z\cdot p_1] = 0{,}91796 < p_1/p_к = 0{,}09859,$$

нерівність не є справедливою.

$$z\cdot p_2 - [z\cdot p_2] = 0{,}739745 < p_2/p_к = 0{,}070422,$$

нерівність не є справедливою.

$$z\cdot p_3 - [z\cdot p_3] = 0{,}035643 < p_3/p_к = 0{,}09859,$$

нерівність є справедливою.

Звідси витікає висновок про викривлення в третій групі розрядів величиною

$$\Delta\alpha_3 = \{[zp_3]\cdot R_3\}_{p_3} = \{[1{,}03561.22]\cdot 1420\}_7 = \{1420\}_7 = 6,$$

тому

$$\alpha_3 = \{\alpha_3 - \Delta\alpha_3\}_{p_3} = \{1 - 6\}_7 = 2 = 10_{(2)}.$$

Порівнюючи отримане значення α_3 з вихідним (приклад 3) упевнюємося в правильній корекції знайденого викривлення.

Таким чином, застосування запропонованих узагальнених кодів дозволяє забезпечити виявлення та виправлення викривлень в b-розрядних узагальнених символах в кожному із базових кодових слів. З урахуванням перемежування глибиною λ довжина пакетів викривлень в узагальнених кодових словах, які можуть бути виправленими, може дорівнювати λ·b двійкових символів. Застосування таких кодів, на погляд автора, дозволить розв'язати сформульовану проблему щодо надійного забезпечення цілісності інформаційних об'єктів в умовах впливу пакетів викривлень значної тривалості.

*Література: **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.** Акушский И. Я., Юдицкий Д. И. Машинная арифметика в остаточных классах. // М.: Сов. радио, 1966. – 421 с. **4.** Василенко В. С., Будько М. М., Короленко М. П. Контроль и відновлення цілісності інформації в автоматизованих системах. К. НТУ "КПІ" // Правове, нормативне та метрологічне забезпечення Системи захисту інформації в Україні. Випуск 4// 2002, с. 119 – 128.*

УДК 621.391:519.2

НИЖНЯЯ ГРАНИЦА ВЕРОЯТНОСТИ РАЗЛИЧЕНИЯ ВНУТРЕННИХ СОСТОЯНИЙ КОМБИНИРУЮЩЕГО ГЕНЕРАТОРА ГАММЫ С НЕРАВНОМЕРНЫМ ДВИЖЕНИЕМ

Антон Алексейчук, Роман Проскуровский
СФ СБ Украины ВИТИ НТУУ "КПИ"

Аннотация: **Исследуется вероятностная модель функционирования комбинирующего генератора гаммы с неравномерным движением в режиме реинициализации начального состояния. Получена аналитическая нижняя граница вероятности правильного различения двух простых гипотез о распределении внутреннего состояния генератора.**
Summary: **A probabilistic model of combine generator is investigated with irregular clocking in procedure of reinitialization of initial state. The analytic lower bound is obtained for probability of the right distinguishing of two simple hypothesis about distribution of the inner state of generator.**
Ключевые слова: **Корреляционный криптоанализ, генератор гаммы с неравномерным движением, проверка гипотез.**

I Введение

Перспективный класс генераторов псевдослучайных последовательностей, используемых при построении поточных криптосистем, образуют генераторы гаммы (ГГ) с неравномерным движением, построенные на основе двоичных линейных регистов сдвига (ЛРС) [1, 2]. Известно, что неравномерность движения ЛРС генератора при определенных условиях улучшает его криптографические свойства (приводит к увеличению периодов и эквивалентных линейных сложностей его выходных последовательностей, повышает стойкость генератора относительно традиционных корреляционных атак) [3 – 6].

В настоящее время наиболее полно исследованны генераторы гаммы с неравномерным движением, состоящие из единственного ЛРС. Обзор методов их криптоанализа приведен в [5, 6]. Следует отметить, что большинство указанных методов основывается на упрощенном вероятностном описании функционирования ЛРС с неравномерным движением, что, отчасти, связано с аналитическими трудностями исследования свойств таких ЛРС в рамках более точных математичеких моделей.

Известным примером ГГ с неравномерным движением, состоящим из нескольких ЛРС, является генератор гаммы шифра A5/1. В [7] предложен статистический метод криптоанализа этого генератора, функционирующего в режиме реинициализации начального состояния (см. [1, 8]). Основным параметром, характеризующим практическую эффективность метода [7], является наименьшее число векторов инициализации (кадров или запусков ГГ), достаточное для восстановления начального состояния ГГ с заданной надежностью.

В настоящей статье описана вероятностная модель функционирования произвольного комбинирующего ГГ с неравномерным движением и равновероятной функцией усложнения, на основе которой получена нижняя граница вероятности различения внутренних состояний генератора по наблюдаемым знакам его выходных последовательностей в серии "независимых запусков". Получена также оценка надежности корреляционной атаки, аналогичной [7], на комбинирующий генератор гаммы с неравномерным движением и линейной функцией усложнения, функционирующий в режиме реинициализации начального состояния. В качестве следствия установлена верхняя граница наименьшего числа запусков ГГ, достаточного для успешной реализации указанной корреляционной атаки с заданной надежностью.