определяют возможность ее применения в технологиях формирования электронных документов для представления структур документов-шаблонов, форматом которых является произвольный описательный язык разметки.

1. Разработка модели модифицированной технологии формирования электронных документов на основании шаблонов в WEB-ориентированных информационных системах / С.Ф. Чалый, Д.Л. Кравченко, Е.А. Моспан // Сборник научных трудов Харьковского университета воздушных сил. – 2008. – Вип. 3 (18). – С. 135–138. 2. Левыкин В. М., Моспан Е. А. Разработка модели формирования электронных документов в WEB-ориентированных информационных системах // АСУ и приборы автоматики. – 2008. – Вып. 144. – С. 54–58. 3. Markup Languages: Theory and Practice: Journal. – Cambridge: MIT Press, 1999. – 120 p. 4. Молли Э Хольцшлаг Использование HTML и XHTML: Using HTML and XHTML : Спец. изд.: Пер. с англ. – Издательский дом Вильямс, 2004, 728 с. 5. Charles F. Goldfarb, Yuri Rubinsky The SGML Handbook. – Oxford University Press, 1990, 663 p.

УДК 519.15:621

О. Різник, Б. Балич, І. Вербенко
Національний університет “Львівська політехніка”,
кафедра автоматизованих систем управління

ВИКОРИСТАННЯ ШУМОПОДІБНИХ КОДІВ ДЛЯ ЗАДАЧ СТЕГАНОГРАФІЇ

Розглянуто можливість використання шумоподібних кодів для задач стеганографії. Розроблено методику побудови кодових комбінацій чисел на основі теорії числових в’язанок, що дає можливість представлення кодових комбінацій чисел у вигляді шумоподібного коду для приховання інформації в найменш значущих бітах графічного формату BMP. Для цих цілей використовується технологія на основі моделі числової в'язанки, яка зводиться до заміни певних пікселів у зображенні, що дає змогу створювати ефективні алгоритми із завадостійким кодуванням та декодуванням при перетворенні графічних форматів з BMP в JPEG та інші і навпаки.

In the article the use of noise codes is examined for the tasks of steganography. The developed method of construction of code combinations of numbers is on the basis of theory of numerical bundles, which enables presentation of code combinations of numbers as a noise code for the information hiding in the the least meaningful bats of graphic format of BMP. For these aims technology is used on the basis of model of numerical bundle, which is taken to replacement of certain pels in an image, that allows to create effective algorithms with a antijammingness code and decoding at transformation of graphic formats from BMP in JPEG et al and vice versa.

Вступ

Запропоновано нетрадиційний підхід до використання шумоподібних кодів для кодування переданих даних.

Запропонований підхід відрізняється від звичайних методів кодування тим, що кодова послідовність використовується як достатньо складна кодувальна функція.

Ключовою проблемою технічних засобів захисту інформації є генерація насправді випадкової послідовності бітів. Річ у тім, що генератори випадкових послідовностей, які використовуються для загальних цілей, є псевдовипадковими генераторами, оскільки в принципі існує кінцева, а не безконечна безліч станів ЕОМ, і, як би складно не формувалося в алгоритмі число, воно все одно має відносно невелику кількість бітів інформаційної насиченості. Якіснішими генераторами випадкових чисел є генератори, засновані на фізичних процесах. Ідея використання шумоподібніх кодів як псевдовипадкових послідовностей виникає з припущення можливості опису поведінки фізичних і природних систем за їх допомогою [2]. Шумоподібні коди належать до множини з украй нерегулярною розгалуженою структурою. Основні поняття теорії поки що знаходяться в процесі становлення та розвитку, але поле їх застосування безперервно розширюється. Великий інтерес до цих кодів пов'язаний з тим, що їх аналоги, такі як квазікоди Баркера, лінійки Голомба, числові в'язанки використовуються в реальних завданнях, причому в типових, а не в екзотичних ситуаціях.

Розглянуто новий підхід приховування інформацій у графічних зображеннях на основі шумоподібних кодів.

Постановка проблеми

Швидкий розвиток алгоритмів стиснення зображень привів до зміни уявлень про саму техніку впровадження секретної інформації. Пропонується включати інформацію до найменш значущих бітів для зменшення помітності для стороннього спостерігача, а саму приховану інформацію кодувати за допомогою шумоподібних кодів.

Використання шумоподібних систем, заснованих на шумоподібних кодах, дає потенційну перевагу над традиційними системами псевдовипадкових послідовностей. Вихідна концепція виражається в розвитку ідеї “чому вдень не видно зірок на небі”, тобто в штучному додаванні до початкових даних шумоподібного сигналу. Аби “побачити зірки”, потрібно “вимкнути” сонце, що “закриває” їх.

Найнаочніше принцип роботи цього методу можна проілюструвати однобарвними растровими зображеннями. Хай вихідне растрове зображення представлене відліками яскравості, розташованими в матриці прямокутного вигляду. Підбираємо за початковими умовами параметри шумоподібного коду на основі числової в'язанки (ЧВ). Потім, застосовуючи деяку функцію (наприклад, порозрядну суму за модулем 2, до пар значень точок вихідного зображення і шумоподібного коду, отримаємо нове зображення, яке і передається по каналу зв'язку. Для розшифрування повідомлення потрібно, знаючи параметри шумоподібного коду, відновити сам шумоподібний код і, застосовуючи операцію, зворотну відносно операції передавальної сторони (у нашому прикладі це також сума за модулем 2), відновити вихідне зображення.

Для ускладнення параметрів шумоподібних кодів можна використовувати вибір різних варіантів параметрів, напряму і початку відліку числової в'язанки. У методі використовується приховування даних із спотворенням контейнера, що засноване на особливостях людського зору.

Приховування повідомлення виконується так:

• беремо повідомлення, заздалегідь готуємо його: шифруємо і архівуємо. Цим досягаємо відразу дві мети – зменшення розміру і збільшення стійкості системи.

• далі беремо контейнер і впроваджуємо оброблене в першому пункті повідомлення в його байтовий контекст.

Так розкладаємо упаковане повідомлення в бітову послідовність; замінюємо надлишкові біти контейнера бітами повідомлення, представленими шумоподібними кодами.

Існують спеціальні програми, які аналізують зображення на наявність прихованої інформації. Як показали дослідження, використання шумоподібних кодів на основі числових в’язанок дає змогу обійти виявлення спеціальними програмами стегоаналізу.

Розв’язання задачі

Задача шумоподібного кодування дає змогу переформулювати задачу стеганографії в термінах складності за А.Н. Колмогоровим [2] як побудову мінімального алгоритму, що породжує псевдовипадкову послідовність при відомих параметрах ключа і неможливість побудови короткого

алгоритму відновлення послідовності при невідомому ключі. При цьому складність даних, за Колмогоровим, визначається як мінімально можлива довжина алгоритму для машини Т'юринга, яку може згенерувати даний набір даних.

Дослідження показали відсутність різних наборів начальних значень, що приводять до побудови шумоподібного коду, який використовується для кодування, що дає змогу стверджувати значущість всіх бітів ключа, а отже, і довжину (за Колмогоровим) алгоритму дешифрування.

Шумоподібне кодування на основі числових в'язанок також дає можливість передачі додаткової інформації без збільшення обсягу переданих даних за рахунок завадостійких властивостей числової в'язанки (ЧВ).

Завадостійкість ЧВ можна оцінити за співвідношенням кількості помилкових кодових комбінацій, які піддаються виявленню t_1 або виправленню t_2 до загальної кількості усіх робочих кодових комбінацій заданої розрядності шумоподібного коду N. Кожна з $S_N(S_N\text{-}1)/2$ різних пар кодових комбінацій містить точно R із N одиничних символів в однойменних розрядах, що випливає з властивостей ЧВ. Решта $N\text{-}R$ символів однієї і стільки ж іншої кодової комбінації відрізняються від символів, що містяться в однойменних розрядах. Тому мінімальна кодова відстань для цього коду визначається як [3]:

$$d_{\min} = 2(N\text{-}R).$$

Кількість помилок, які можна виявити t_1, і кількість помилок, що можна виправити t_2 за допомогою коригувального коду, визначається мінімальною кодовою відстанню залежностями:

$$t_1 \le d_{\min}\text{-}1,\ t_2 \le (t_1\text{-}1)/2.$$

Співвідношення між параметрами N і R, коли код набуває здатності виявляти та виправляти максимально можливу кількість помилок:

$$R = \begin{cases} n/2, & n - парне \\ (n-1)/2, & n - непарне \end{cases}.$$

Властивістю пропонованого методу кодування є прояв хаотичних властивостей шумоподібного коду (аналог великої довжини криптографічного ключа) лише при використанні початкових значень, відповідних числовій в'язанці, що вимагає попереднього складання каталога в'язанок, що задовольняють заданим властивостям для ефективного вибору ключа для кодування.

Пропонується застосувати для розподілу змінених бітів комбінаторну модель шумоподібних кодів у вигляді числової в'язанки. Числова в'язанка з параметрами (S_N, N, R) – це алгебраїчна структура, утворена на послідовності N цілих додатних чисел, значення яких, як і значення сум поруч розміщених між собою чисел, вичерпують числа натурального ряду не більше одного разу $(R = 1)$. Елементи ЧВ розташовані один біля одного у вигляді замкнутого ланцюжка, і їх сума дорівнює S_N [2, 3, 4].

Пропонується застосування для розподілу найменш значущих бітів (НЗБ) комбінаторної моделі шумоподібних кодів. Одне з перших формулювань деяких засадничих правил для статистичних властивостей шумоподібних кодів (періодичних псевдовипадкових послідовностей великого періоду) було представлено Соломоном Голомбом.

Три основні правила здобули популярність як постулати Голомба.

1. Кількість "1" в кожному періоді повинна відрізнятися від кількості "0" не більш ніж на одиницю.

2. У кожному періоді половина серій (з однакових символів) повинна мати довжину один, одна чверть повинна мати довжину два, одна восьма повинна мати довжину три і так далі. Більш того, для кожної з цих довжин має бути однакова кількість серій з "1" і "0".

3. Передбачимо, у нас є дві копії однієї і тієї ж послідовності періоду p, зрушені відносно один одного на деяке значення d. Тоді для кожного d, $0 \le d \le p\text{-}1$, ми можемо підрахувати кількість узгоджень між цими двома послідовностями A_d, і кількість неузгодженостей D_d.

Коефіцієнт автокореляції для кожного d визначається співвідношенням $(A_d - D_d)/p$, і ця функція автокореляції набуває різних значень у міру того, як d проходить всі допустимі значення.

Послідовність, що задовольняє правила 1–3, часто іменується шумоподібним кодом, або "ПШ-послідовністю", де ПШ означає "псевдошумова".

Оскільки ми кодуємо таблицю кодів ASCII, яка містить 256 кодів, а шумоподібний код завдовжки 256 сьогодні не знайдено, то треба використовувати шумоподібний код з більшою довжиною.

Оскільки ми кодуємо таблицю кодів ASCII, яка містить 256 варіантів, а в'язанки з сумою, що дорівнює 256, не існує, то необхідно використовувати ЧВ з більшою сумою, наприклад, ЧВ (283, 20, 1) [4].

За вхідними даними складаємо таблицю кодів. Вона складається з масиву, кожен елемент якого має вигляд (0, 0, .., 0, 0) за таким принципом:

• беремо порядковий номер символу заданого алфавіту, якщо цей номер наявний серед елементів в'язанки, то в елемент масиву, в позицію потрібного елементу в'язанки, пишеться “1”;

• якщо заданий порядковий номер в таблиці не знайдений, то “1” пишуться в позиції елементів в'язанки, які в сумі дають потрібний номер (слід пам'ятати, що підсумовують за правилами, визначеними для ЧВ).

Після цього проводиться посимвольне зчитування вихідної інформації з файла, потім визначається номер зчитаного символу в заданому алфавіті, визначається код, який відповідає цьому номеру і записується в проміжний файл. Код з проміжного файла додається послідовно в наймолодші біти RGB файла BMP.

Щоб декодувати таке зображення, необхідно з послідовності символів молодших бітів RGB файла BMP вирахувати оригінальні значення молодших бітів RGB файла BMP, після чого відповідно до порядку N ЧВ зчитати послідовності за N символів. Для зчитаних N символів шукаємо в таблиці кодів ЧВ, який символ ASCII алфавіту був закодований, і отримуємо файл результату.

Наприклад, розглянемо побудову шумоподібних кодів за допомогою ЧВ (1, 1, 1, 2, 2, 5, 1, 3, 3) порядка $N = 9$ кратності $R = 4$, виділимо рядок із $S_N = 19$ пронумерованих у зростаючому порядку клітинок одновимірного масиву і заповнимо інформаційними "одиницями" клітинки, номери яких збігаються з числами, визначеними з ЧВ.

У клітинки, що залишилися незаповненими, занесемо "нулі". Утворена послідовність одиниць і нулів є S_N-розрядною шумоподібною кодовою послідовністю, циклічним зсувом якої можна одержати й решту дозволених $S_N - 1$ комбінацій (таблиця).

Шумоподібна кодова послідовність на основі ЧВ з $N = 9$ та $R = 4$

1	1	1	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0
0	1	1	1	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0
0	0	1	1	1	1	0	1	0	1	0	0	0	0	1	1	0	0	1
1	0	0	1	1	1	1	0	1	0	1	0	0	0	0	1	1	0	0
0	1	0	0	1	1	1	1	0	1	0	1	0	0	0	0	1	1	0
0	0	1	0	0	1	1	1	1	0	1	0	1	0	0	0	0	1	1
1	0	0	1	0	0	1	1	1	1	0	1	0	1	0	0	0	0	1
1	1	0	0	1	0	0	1	1	1	1	0	1	0	1	0	0	0	0
0	1	1	0	0	1	0	0	1	1	1	1	0	1	0	1	0	0	0
0	0	1	1	0	0	1	0	0	1	1	1	1	0	1	0	1	0	0
0	0	0	1	1	0	0	1	0	0	1	1	1	1	0	1	0	1	0
0	0	0	0	1	1	0	0	1	0	0	1	1	1	1	0	1	0	1
1	0	0	0	0	1	1	0	0	1	0	0	1	1	1	1	0	1	0
0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	1	1	0	1
1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	1	1	0
0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	1	1
1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	1
1	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1
1	1	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1

Будь-яка з $S_N(S_N\text{-}1)/2$ різних пар кодових комбінацій містить точно R із N одиничних символів в однойменних розрядах, що випливає з властивостей ЧВ. Решта $N-R$ символів однієї і стільки ж іншої шумоподібної кодової послідовності відрізняються від символів, що містяться в однойменних розрядах.

Висновки

Аналіз результатів дослідження шумоподібних кодів на основі ЧВ підтверджує актуальність і перспективність створення на їхній основі нових інформаційних технологій з використанням спеціальних систем кодування інформації з високою захищеністю від завад й стороннього декодування, що дає змогу розширити сферу застосувань комбінаторних методів оптимізації в інформатиці і комп'ютерній техніці [1].

Стегостійкість запропонованого методу, а також оптимальні види в'язанок, що дають мінімальний розмір довжини шумоподібного коду при максимальній стійкості шифросистеми, є предметом подальших досліджень.

Отже, використання шумоподібних кодів у стеганографії може дати потужний і ефективний механізм приховання інформації, який сильно залежить від параметрів використаного коду. Наведений метод шумоподібного кодування розкриває інший концептуальний підхід до організації процедури як приховання, так і можливого додаткового впровадження інформаційного вмісту.

Переваги:

- Метод використовує тільки арифметичні операції, що дає максимальну швидкодію.
- Використання шумоподібних кодів на основі числових в’язанок значно ускладнює стегоперевірку зображень на наявність прихованої інформації.

Недоліки:

- Малюнок-контейнер має бути не менш ніж 24-бітним.
- Безпосередньо реалізований метод працює з BMP файлами, але при їх перетворенні на інший графічний формат і зворотно прихована інформація відновлюється за рахунок завадостійких властивостей числових в’язанок.

1. Дурняк Б.В., Різник О.Я., Різник В.В., Кісь Я.П., Парубчак В.О. Захист даних методом комбінаторної оптимізації // Праці третьої міжнародної наукової конференції ISDMIT'2007, м.Євпаторія. – Т. 2. – С. 152, 153. 2. Колмогоров А. Н. Три подхода к количественному определению информации // Проблемы передачи информации. Т. 1. 1965. – Вып. 1. 3. Різник В.В. Синтез оптимальних комбінаторних систем. – Львів, 1989. 4. Різник О.Я., Парубчак В.О., Скрибайло-Леськів Д.Ю. Використання числових в’язанок для кодування інформації // Праці міжнародної конференції "Сучасні комп'ютерні системи та мережі: розробка та використання" (ACSN'2007). – С. 112–114.