

## МЕТОДИ ВИЯВЛЕННЯ СТЕГАНОГРАФІЧНОГО ПРИХОВАННЯ ІНФОРМАЦІЇ В ЗОБРАЖЕННЯХ

© Швідченко І.В., 2015

**This paper is devoted to a problem of detecting the fact of hidden information transfer. Various steganalysis methods for graphic files have been considered and analyzed.**

**Keywords - steganography, steganalysis, cover medium, message.**

**Стаття присвячена проблемі виявлення факту прихованої передачі інформації. Розглянуто і проаналізовано різні методи стеганоаналізу для графічних файлів.**

**Ключові слова - стеганографія, стеганоаналіз, контейнер, повідомлення.**

### Вступ

В сучасних комп'ютерних мережах передаються дуже великі потоки мультимедійних повідомлень, які потенційно можуть бути використані для приховання інформації. Враховуючи можливість використання стеганографії для реалізації злочинних намірів, досить актуальними є питання стеганоаналізу. Стеганоаналіз – наука про вивчення методів виявлення існування секретної інформації у відкритих повідомленнях. В стеганоаналізі розрізняють дві основних стратегії дій противника (стеганоаналітика): активну та пасивну. При активній стратегії противник намагається знищити секретну інформацію в відкритому повідомленні, при пасивній – виявити факт існування і саму секретну інформацію.

На сьогодні найпоширенішим методом стеганографічного приховання є метод заміни найменш значущих бітів. Ідея методу полягає в заміні від одного до чотирьох молодших бітів у байтах кольорного подання пікселів вихідного зображення бітами повідомлення, яке приховується. Можливість такої заміни обумовлена наявністю в зображеннях структурної надлишковості. Метод застосовується до растрових зображень, представлених у форматі без компресії. Одним з таких форматів є BMP. Перевагою BMP є висока якість зображення, а також простота формату, що робить його популярним для застосування як контейнер.

Ще один метод стеганографічного перетворення інформації, базується на використанні особливостей файлів, стиснутих із втратою даних. Популярним графічним форматом, що використовує алгоритм стискання даних зі втратами є JPEG. При прихованні в JPEG файли інформація приховується не в значення кольорних складових окремих пікселів, а в біти квантованих дискретних косинусних коефіцієнтів. З позиції стеганографії файли даного формату дозволяють приховувати порівняно великі обсяги інформації [1, 2].

В статті розглянуто методи аналізу, які використовуються для виявлення факту прихованої передачі інформації в графічних файлах BMP і JPEG формату.

### Методи стеганоаналізу

Найпростішими методами аналізу контейнерів-зображень є **візуальні методи**. Візуальні методи намагаються виявити існування стеганографічного вкраплення за допомогою візуального контролю (неозброєним оком) або за допомогою автоматизованих процесів. Візуальний контроль за допомогою неозброєного ока матиме успіх, коли стеганографічні дані вкраплені у однотонні фрагменти зображення. Автоматизовані комп'ютерні додатки дозволяють розкласти зображення на його індивідуальні бітові площини. Бітова площина складається з одного біту пам'яті для кожного пікселя в зображенні, і є типовим місцем зберігання інформації, прихованої за допомогою стеганопрограм. Будь-який незвичний зовнішній вигляд у відображенні площини молодшого

двійкового розряду буде, ймовірно, означати існування вкраплених стеганографічних даних.

Для методу візуального аналізу бітових площин велике значення має те, який метод використовувався при вкрапленні інформації. Якщо приховання інформації здійснювалося за допомогою методу послідовної заміни (рис. 1, а), або методу розподіленого вкраплення (на основі генератора псевдовипадкових чисел) (рис. 1, б), то факт приховання може бути встановлений з великою ймовірністю. Також, візуально можна визначити наявність вкрапленої інформації у випадку використання методу вкраплення повідомлення із заповненням. Оскільки ймовірнісні характеристики повідомлення не збігаються з ймовірнісними характеристиками молодших бітів порожнього контейнера, то при перегляді бітового зрізу з вкрапленими даними буде чітко видно границю між заповненою і не задіяною вкрапленням частиною (рис. 1, в).

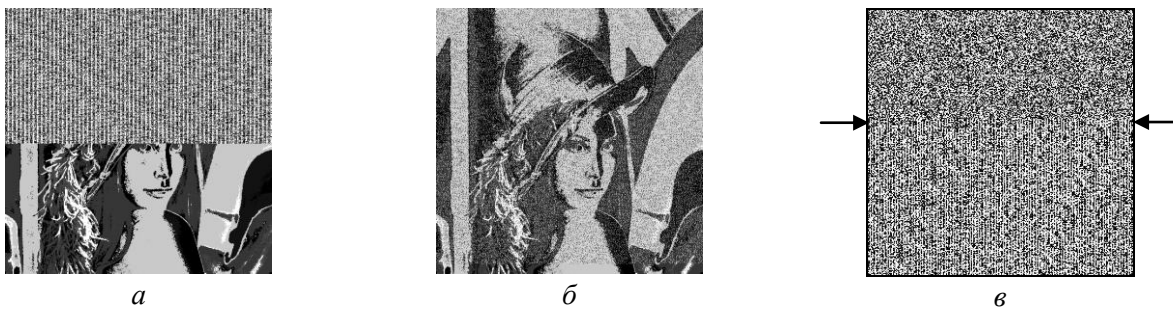


Рис. 1. Бітові площини стеганоконтейнерів: а – метод послідовної заміни; б – метод розподіленого вкраплення; в – метод вкраплення з заповненням

Існують стеганографічні програми, які використовують метод дописування даних у кінець файлу-контейнера завдяки використанню системи маркерів. Всі стандартні програми перегляду, доходючи до маркера “кінець зображення” припиняють роботу, і прихована інформація залишається непізнаною. Цим способом можна розмістити досить багато інформації. Однак такий метод приховання є уразливим до **методів структурного аналізу**. Виявити зміни у форматі файлу даних можливо за допомогою шістнадцятиричного редактора. Використовуючи інформацію заголовка, можна виявити “кінець зображення”, а отже визначити місце розташування позиції вкраплених даних.

Деякі стеганографічні програми залишають після себе сигнатури – певні послідовності байтів, які завжди з’являються у файлі після вкраплення інформації. **Методи сигнатурного аналізу** зображення дозволяють відшукати бітові послідовності, специфічні для певних програм стеганоприховання (рис. 2). Стеганоаналіз на основі сигнатур може вимагати дуже багато часу, тому що спочатку потрібно розпізнати сигнатуру для певної стеганографічної програми з великої вибірки файлів, які були вкраплені з її допомогою. Крім того, повинні бути використані автоматизовані процеси для пошуку всіх потенційних файлів-контейнерів для цієї певної сигнатури.

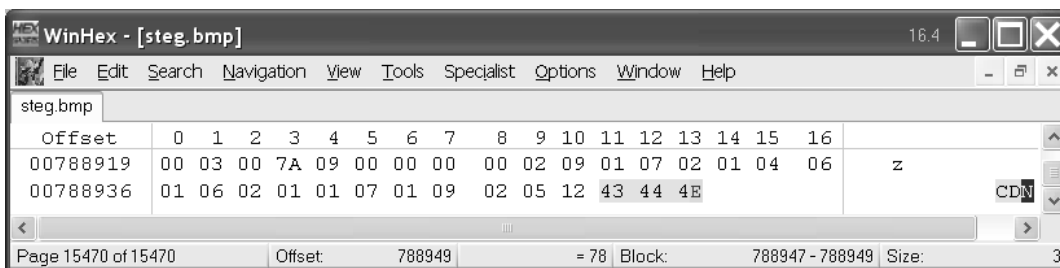


Рис. 2. Перегляд зображення в шістнадцятиричному редакторі на наявність сигнатур. CDN – сигнатура стеганографічної програми Hiderman

До переваг сигнатурних методів відноситься можливість отримання результату, який однозначно характеризує застосовану для приховання даних стеганосистему. Основним недоліком є невелике (менш 10 %) число стеганопрограм, що залишають у контейнерах свої сигнатури.

**Статистичні методи аналізу** намагаються виявити найменші зміни в статистичній поведінці файлу, викликані вкрапленою стеганографією. Суть статистичних методів полягає в оцінюванні ймовірності існування стеганографічного приховання з невідомою стеганосистемою на основі критерію оцінки наближення досліджуваного контейнера до “природного”.

**Метод оцінки числа переходів значень молодших бітів у сусідніх елементах контейнера.** У методі використовується знання, що між молодшими бітами сусідніх елементів, і між ними і іншими бітами природних контейнерів є кореляційні зв’язки. При аналізі графічних файлів формату BMP в якості елементів послідовності, яка аналізується, вибираються найменш значущі біти (НЗБ) пікселів, що розташовуються поруч, кольорних складових зображення. При дослідженні файлів формату JPEG – молодші біти сусідніх коефіцієнтів ДКП, відмінних від 0 і 1.

Під “переходом” розуміють перехід значення  $i$ -го елемента послідовності в значення  $i + 1$  елемента послідовності  $x$ ,  $i = 1, 2, \dots, n - 1$ ,  $n$  – довжина послідовності. Оскільки послідовності є двійковими, то аналізуються чотири види переходів: з 0 в 0, з 0 в 1, з 1 в 0 і з 1 в 1. За отриманими результатами будується гістограма [3]. Для кожного розряду перший стовпець гістограми показує число переходів у потоці НЗБ із 0 в 0, другий стовпець – з 0 в 1, третій стовпець – з 1 в 0, четвертий стовпець – з 1 в 1.

Для порожнього контейнера і контейнера, що містить вкраплену інформацію, число переходів у потоці НЗБ буде різним. Розподіл НЗБ стеганоконтейнера має, як правило, випадковий характер. Відповідно число переходів у потоці НЗБ для всіх станів буде приблизно однаковим, що не властиво порожньому контейнеру (рис. 3, а, б).

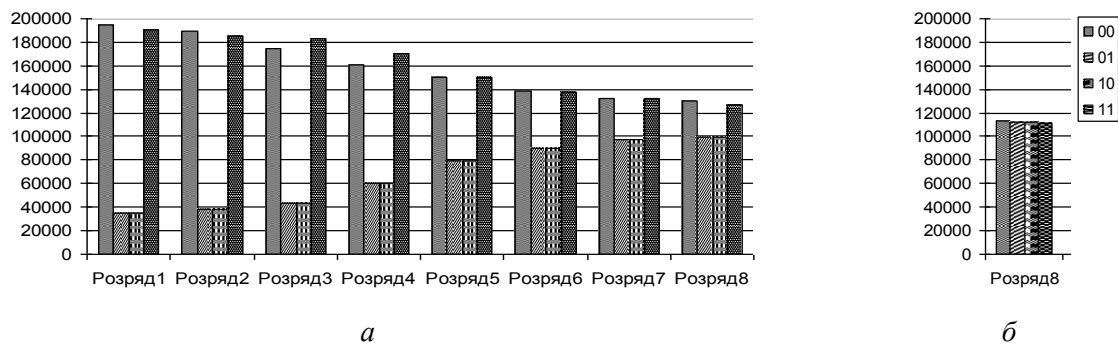


Рис. 3. Гістограма частот переходів бітових значень: а – порожнього контейнера, б – стеганоконтейнера (восьмий розряд контейнера, в який були внесені зміни)

Статистичний критерій для оцінки частот переходів бітових значень [3]: файл, що аналізується розбивається на  $K$  блоків однакової довжини і вибирається деяке порогове значення  $h_M$ . Обчислюються значення статистик

$$M_j = \left[ \frac{(\mu_{00} - \mu_{01})^2}{2} + \frac{(\mu_{11} - \mu_{10})^2}{2} \right], j = 1, 2, \dots, K$$

$\mu_{ij}$  – кількість переходів у потоці НЗБ з  $i$  в  $j$  ( $i, j = 0, 1$ ) У випадку якщо  $M_j < h_M$ , вважається, що в  $j$ -м блоці міститься прихована інформація.

Застосування цього критерію до JPEG файлів виявилось ефективним при значному заповненні зображень контейнерів. Діапазон зміни значень  $M_j$  для порожніх контейнерів становив від 40000 до 130000. Після приховання інформації значення  $M_j$  зменшувалися до 250-6000.

**Метод оцінки частот появи  $k$ -бітових серій у потоці НЗБ елементів контейнера.** Метод дозволяє оцінити рівномірність розподілу елементів в послідовності, яка досліджується, на основі аналізу частоти появи нулів і одиниць, і серій, що складаються з  $k$  бітів [4]. У бітовому поданні послідовності  $x$ , що досліджується, підраховується скільки разів зустрічаються нулі і одиниці ( $k = 1$ ), серії-двійки (00, 01, 10, 11:  $k = 2$ ), серії-трійки (000, 001, 010, 011, 100, 101, 110, 111:  $k = 3$ ) і т.д. На основі результатів будується гістограма.

Для JPEG зображень гістограма будується за значеннями частот появи бітових серій у потоці НЗБ коефіцієнтів ДКП відмінних від  $-1, 0, 1$ .

Для незаповнених BMP і JPEG зображень не є характерним, щоб значення частот всіх компонентів знаходились досить близько (рис. 4, а). При вкрапленні інформації, значення частот зближуються (рис. 4, б). Цей факт використовується при аналізі.

Результати роботи методу залежать від стеганографічного перетворення і від обсягу даних, що приховуються. Як правило, виявлення факту приховання здійснимо при заповненні контейнера на 60% і вище.

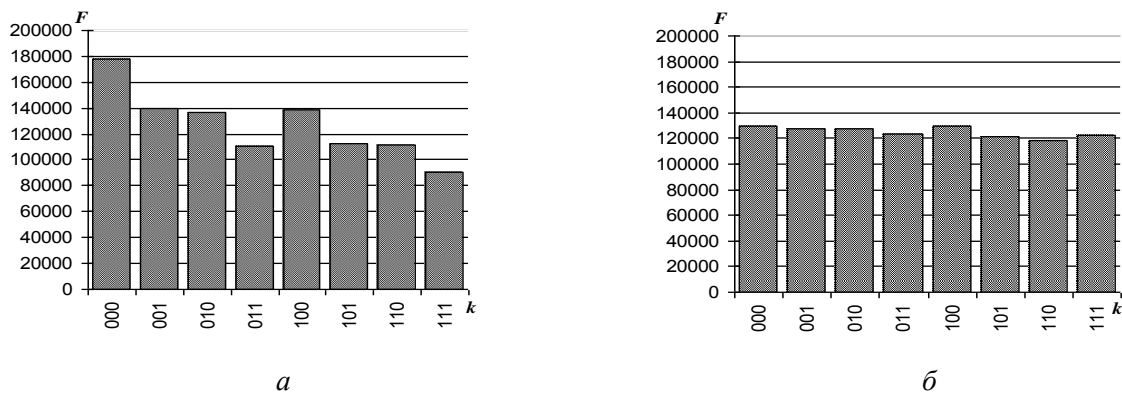


Рис. 4. Гістограма частот серії-трійки ( $k = 3$ ) у потоці НЗБ:  
а – пустого контейнера, б – стеганоконтейнера

**Метод аналізу розподілу пар значень на основі критерію  $\chi^2$**  (хі-квадрат). У методі використовується аналіз гістограми, отриманої за елементами зображення і оцінка розподілу пар значень цієї гістограми [1]. Для BMP файлів пари значень формуються значеннями пікселів зображення, для JPEG – квантованими коефіцієнтами дискретного косинусного перетворення, які відрізняються за молодшим бітом. Молодші біти зображень не є випадковими. Частоти двох сусідніх елементів контейнера мають перебувати досить далеко від значення частоти середнього арифметичного цих елементів. В “пустому” зображенні ситуація, коли частоти елементів зі значеннями  $2N$  і  $2N+1$  близькі за значенням, зустрічається досить рідко. При вкрапленні інформації дані частоти зближуються або стають рівними. Ідея атаки  $\chi^2$  полягає в пошуку цих близьких значень і підрахунку ймовірності вкраплення на основі того, як близько розташовуються значення частот парних і непарних елементів аналізованого контейнера. Особливістю алгоритму є послідовний аналіз всього зображення і, відповідно, накопичення частот елементів.

Метод  $\chi^2$  є універсальним, оскільки підходить для аналізу зображень, в які інформація вкраплювалася за допомогою різних стеганографічних програм. Однак результати роботи методу за критерієм  $\chi^2$  значною мірою залежать від методу приховання даних. При послідовній заміні НЗБ елементів контейнера і вкрапленні повідомлення з заповненням метод виявляє наявність прихованих даних (рис. 5, а, б), а при псевдовипадковому виборі молодших бітів (розподіленому вкрапленні) метод не спрацьовує.

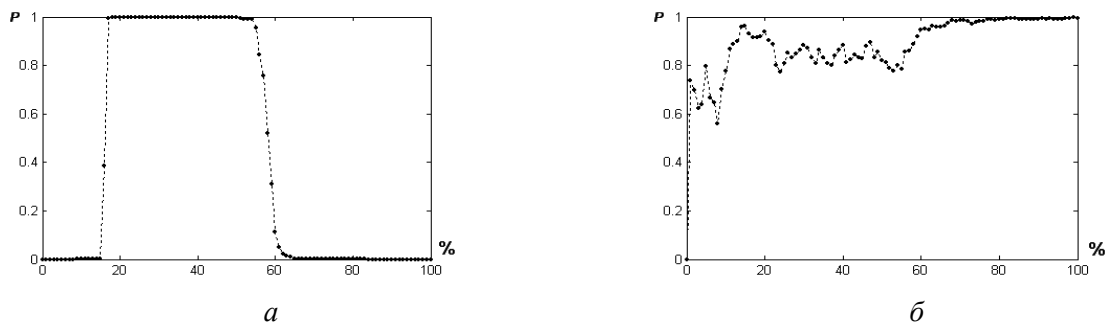


Рис. 5. Ймовірність вкраплення за критерієм  $\chi^2$ : а – послідовне вкраплення, б – вкраплення з заповненням

**Метод аналізу гістограм, побудованих за частотами елементів зображення.** Метод дозволяє оцінити рівномірність розподілу елементів зображення, що аналізується, а також визначити частоту появи конкретного елемента.

Якщо частоти двох сусідніх елементів ВМР зображення близькі за значенням і/або розташовані з різницею в одиницю (наслідок використання класичного методу НЗБ), то контейнер містить приховані дані (рис. 6, б). В іншому випадку контейнер вважається пустим (рис. 6, а) [4].

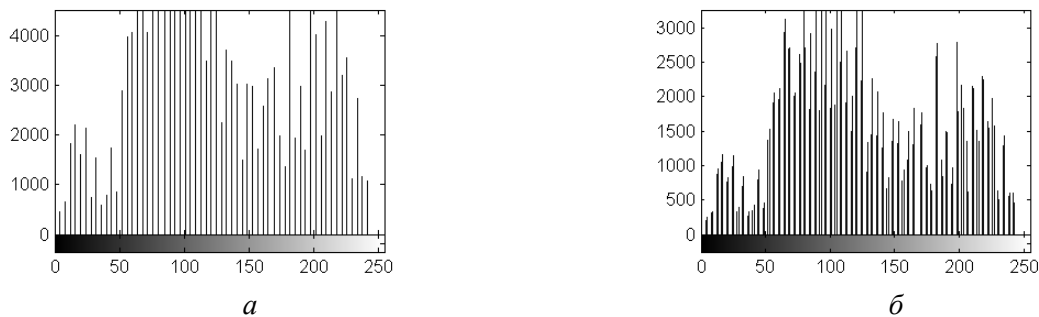


Рис. 6. Гістограма частот пікселів: а – вихідного ВМР зображення, б – ВМР зображення, що містить приховану інформацію

Для зображень в JPEG форматі будується гістограма частот квантованих коефіцієнтів ДКП. Експериментально виявлено, що огинаюча гістограми пустого зображення має більш гладкий характер (рис. 7, а), у порівнянні з гістограмами зображень, що містять приховані дані (рис. 7, б).

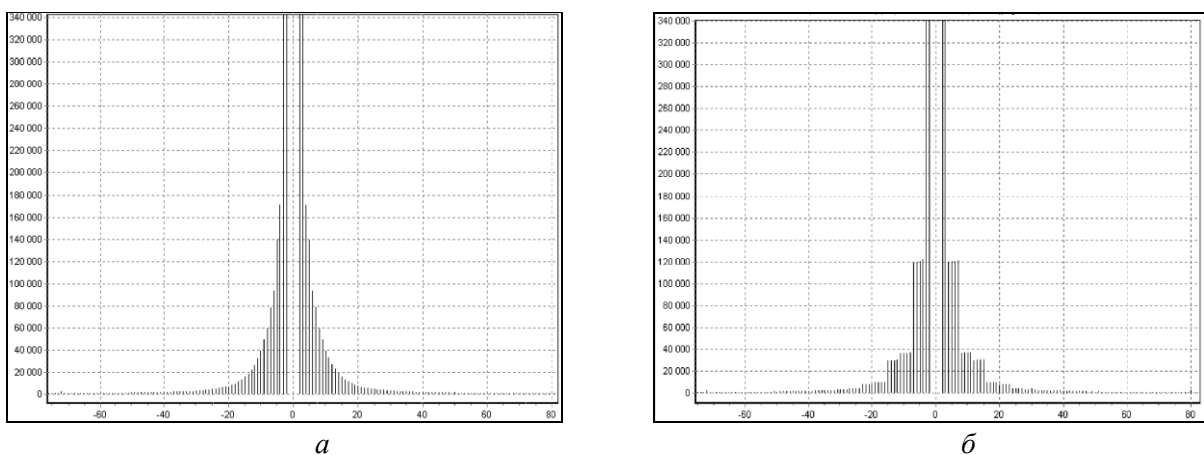


Рис. 7. Гістограма частот коефіцієнтів ДКП: а – вихідного JPEG зображення, б – JPEG зображення, що містить приховану інформацію

Звичайно, залежно від характеру і рівня стиснення зображення, гістограми можуть змінюватися – у них можуть з’являтися стрибки і провали, але важливо те, що приховання інформації змінює загальний вид гістограм. Більшість стеганографічних програм, що працюють із JPEG, приховують дані в молодших бітах коефіцієнтів ДКП відмінних від 0 і 1. Як наслідок частоти 0-х і 1-х ДКП не змінюються, в той час як всі інші частоти або зменшуються, або збільшуються залежно від алгоритму вкраплення. При значних обсягах приховуваної інформації гістограми часто мають ступеневий характер, що нетипово для “природних” JPEG зображень.

**Метод аналізу розподілу елементів зображення на площині.** Метод призначений для визначення залежностей між елементами послідовності, що досліджується.

На площину (поле) розміром  $(2^R - 1) \times (2^R - 1)$ , де  $R$  – розрядність елемента послідовності, наносяться точки з координатами  $(x_i, x_{i+1})$ ,  $x_i$  – елементи послідовності  $x$ , що досліджується,  $i = 1, 2, \dots, n - 1$ ,  $n$  – довжина послідовності [4]. За отриманим результатом проводиться аналіз.



Рис. 8. Розподіл на площині елементів: а – вихідного зображення, б – зображення, що містить приховану інформацію

Якщо точки по всьому полю розташовані хаотично, то між елементами послідовності відсутні залежності, що характерно для контейнерів з вкрапленими даними (рис. 8, б). У випадку пустого контейнера точки на полі будуть розташовані нерівномірно або утворювати “візерунки” (рис. 8, а).

Крім наведених вище, існує багато статистичних методів, заснованих на різних математичних моделях процесу стеганографії. Статистичні методи не є засобом, який дозволяє з 100% надійністю визначати наявність прихованої інформації. Вони дають можливість стеганоаналітику з певною ймовірністю судити про те, використовувалось стеганографічне перетворення чи ні.

### Висновки

Виявлення прихованої передачі даних, прихованих одним із багатьох існуючих методів стеганографії в різні формати контейнерів є досить складним процесом. Для вирішення цієї задачі необхідний комплексний підхід. Такий підхід спрямований на аналіз всіх можливих методів приховання, починаючи від найпростіших і закінчуючи найбільш складними. Таким чином, на сьогодні актуальною задачею є удосконалення існуючих і створення нових методів стеганоаналізу, а також розробка на їх основі програмного комплексу, за допомогою якого з деякою ймовірністю можна визначити наявність прихованої інформації в контейнері або її відсутність.

1. Аграновский А.В., Балакин А.В., Грибунин В.Г., Сапожников С.А. *Стеганография, цифровые водяные знаки и стеганоанализ*. – М.: Вузовская книга, 2009. – 220 с. 2. Конахович Г.Ф., Пузыренко А.Ю. *Компьютерная стеганография. Теория и практика*. – Киев: МК-Пресс, 2006. – 288 с. 3. Барсуков В.С., Романцов А.П. *Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации // Специальная Техника*. – 2000. – № 1. 4. Иванов М.А., Чугунков И.В. *Теория, применение и оценка качества генераторов псевдослучайных последовательностей*. – М.: КУДИЦ-ОБРАЗ, 2003.