## Стеганография

STEGANO – поиск информации, передаваемой в неявном виде. Картинка в картинке, зашифрованные послания в последних битах и т. д.

## Imaging Processing

# Литература

1. **Конахович Г. Ф., Пузыренко А. Ю.** Компьютерная стеганография. Теория и практика.— К.: "МК-Пресс", 2006. — 288 с, ил.

2. **Аграновский А.В.** Стеганография, цифровые водяные знаки и стеганоанализ. – «Вузовская книга», 2009. – 220 с.

**3. Ховард М., Лебланк Д., Вьега Дж.** 24 смертных греха компьютерной безопасности. Библиотека программиста. — СПб.: Питер, 2010. — 400 с.: ил.

4. **Ященко В. В.** Введение в криптографию.

5. ***Digital Image Forensics***: lecture notes, exercises, and matlab code for a survey course in digital image and video forensics. // www.cs.dartmouth.edu/farid

# *Digital Image Forensics. Возможные направления*

1. Проверить, соответствует ли сигнатура файла сигнатуре графического файла, например, с помощью программы __.
2. Определив по сигнатурам файлов границы графического файла, проверить, нет ли там дополнительной информации.
3. Представить файл в шестнадцатеричной или текстовой форме, например, с использованием программ ___.
4. Разложить файл на цветовые составляющие.
5. Попробовать усилить изображение или эквализировать.

- **Общие вопросы стеганографии**
- **Общие теоретические положения**
  - **Представление сигналов во временной (пространственной) и частотной областях**
  - **ГСЧ**
- **Скрытие данных в контейнерах различной природы**
  - **В неподвижных изображениях**
  - **В аудиосигналах**
  - **В тексте**
- **Общие положения анализа изображений**
  - **Форматы графических файлов**
  - **Вопросы обработки изображений**
  - ***Digital Image Forensics***

***Криптографическая*** (с греческого *ypvmoq* — "тайный", *ypdqxv* "пишу") **защита информации** (система изменения последней с целью сделать ее непонятной для непосвященных, сокрытие содержания сообщений за счет их шифрования), выражающаяся в наличии шифрованного сообщения, сама по себе привлекает внимание.

Скрытие же самого факта существования секретных данных при их передаче, хранении или обработке является задачей ***стеганографии*** (от греческого *arcyavog* — "скрытый") — науки, которая изучает способы и методы скрытия конфиденциальных сведений. Задача извлечения информации при этом отступает на второй план и решается в большинстве случаев стандартными криптографическими методами.

# Литература

.Какой сигнал: аналоговый или цифровой?

.Какова цель?

1. Зашифровать информацию саму по себе.
2. Спрятать одну информацию в другой.

.Что является носителем?

.В какой области работаем с сигналами?

1. В пространственной.
2. В частотной.

.Каков алгоритм?

## Области применения стеганографии

**Защита от копирования**

Электронная коммерция; контроль за тиражированием (DVD); распространение мультимедийной информации (видео по запросу)

**Скрытая аннотация документов**

Медицинские снимки; картография; мультимедийные базы данных

**Аутентификация**

Системы видеонаблюдения, электронной коммерции, голосовой почты; электронное конфиденциальное делопроизводство

**Скрытая связь**

Применение в военных и разведывательных целях, а также в случаях, когда нельзя использовать криптографию


**Взаимосвязь между устойчивостью стеганосистемы и объемом скрываемого сообщения при неизменном размере файла-контейнера**


**Структурная схема стеганосистемы как системы связи**

**Типы стеганографических систем:**

- бесключевые стеганосистемы;
- системы с секретным ключом;
- системы с открытым ключом;
- смешанные стеганосистемы.

**Прин́цип Керкго́ффса** — правило разработки криптографических систем, согласно которому в **засекреченном** виде держится только определённый набор параметров алгоритма, называемый **ключом**, а сам **алгоритм** шифрования должен быть **открытым**.

Впервые данный принцип сформулировал в XIX веке голландский криптограф Огюст Керкгоффс. Шеннон сформулировал этот принцип (вероятно, независимо от Керкгоффса) следующим образом: «Враг знает систему».

*Сущность принципа* заключается в том, что чем меньше секретов содержит система, тем выше её безопасность. Так, если утрата любого из секретов приводит к разрушению системы, то система с **меньшим** числом **секретов** будет **надёжней**.

# Шесть требований Керкгоффса

1.Система должна быть физически, если не математически, невскрываемой;
2.Нужно, чтобы не требовалось сохранение системы в тайне; попадание системы в руки врага не должно причинять неудобств;
3.Хранение и передача ключа должны быть осуществимы без помощи бумажных записей; корреспонденты должны располагать возможностью менять ключ по своему усмотрению;
4.Система должна быть пригодной для сообщения через телеграф;
5.Система должна быть легко переносимой, работа с ней не должна требовать участия нескольких лиц одновременно;
6.Наконец, от системы требуется, учитывая возможные обстоятельства её применения, чтобы она была проста в использовании, не требовала значительного умственного напряжения или соблюдения большого количества правил.


**Модель анализа угрозы и оценки устойчивости стеганосистем**

## Этапы взлома стеганографической системы

### Пассивные атаки

1. Обнаружение факта присутствия скрытой информации.
2. Извлечение скрытого сообщения.

### Активные (или злонамеренные) атаки

3. Видоизменение (модификация) скрытой информации.
4. Запрет на выполнение любой пересылки информации, в том числе скрытой.

## Виды атак на стеганосистемы

- **Атака на основании известного заполненного контейнера.** В этом случае нарушитель имеет в своем распоряжении один или несколько заполненных контейнеров (в последнем случае предполагается, что встраивание скрытой информации выполнялось тем же самым способом). Задача нарушителя может заключаться в выявлении факта наличия стеганоканала (основное задание), а также в извлечении данных или определении ключа. Зная ключ, нарушитель имеет возможность анализа других стеганосообщений.

- **Атака на основании известного встроенного сообщения.** Этот тип атаки в большей мере характерен для систем защиты интеллектуальной собственности, когда в качестве ЦВЗ, например, используется известный логотип фирмы. Задачей анализа является получение ключа. Если соответствующий скрытому сообщению заполненный контейнер неизвестен, то задача является практически неразрешимой.
- **Атака на основании выбранного скрытого сообщения.** В этом случае нарушитель может предлагать для передачи свои сообщения и анализировать полученные при этом контейнеры-результаты.

- **Адаптивная атака на основании выбранного сообщения.** Эта атака является частным случаем предыдущей. При этом нарушитель имеет возможность выбирать сообщения для навязывания их передачи адаптивно, **в** зависимости от результатов анализа предшествующих контейнеров-результатов.
- **Атака на основании выбранного заполненного контейнера.** Этот тип атаки более характерен для систем ЦВЗ. У стеганоаналитика есть детектор заполненных контейнеров в виде "черного ящика" и несколько таких контейнеров. Анализируя продетектированные скрытые сообщения, нарушитель пытается раскрыть ключ.

**Атаки, не имеющих прямых аналогов в криптоанализе**:

- **Атака на основании известного пустого контейнера.** Если последний известен нарушителю, то путем сравнения его с подозреваемым на присутствие скрытых данных контейнером он всегда может установить факт наличия стеганоканала. Намного более интересным представляется сценарий, когда контейнер известен приблизительно, с некоторой погрешностью (например, если к нему добавлен шум). В этом случае существует возможность построения устойчивой стеганосистемы.

- **Атака на основании выбранного пустого контейнера.** В этом случае нарушитель способен заставить воспользоваться предложенным им контейнером. Последний, например, может иметь значительные однородные области (однотонные изображения), и тогда обеспечить секретность встраивания будет непросто.

- **Атака на основании известной математической модели контейнера или его части.** При этом атакующий пытается определить отличие подозреваемого сообщения от известной ему модели. Например, можно допустить, что биты в середине определенной части изображения являются коррелированными. Тогда отсутствие такой корреляции может служить сигналом о наличии скрытого сообщения. Задача того, кто встраивает сообщение, заключается в том, чтобы не нарушить статистики контейнера. Отправитель и тот, кто атакует, могут иметь в своем распоряжении разные модели сигналов, тогда в информационно-скрывающем противоборстве победит тот, кто владеет более эффективной (оптимальной) моделью.

Основная цель атаки на стеганографическую систему аналогична цели атак на криптосистему с той лишь разницей, что резко возрастает значимость активных (злонамеренных) атак. Любой контейнер может быть заменен с целью удаления или разрушения скрытого сообщения, независимо от того, существует оно в контейнере или нет. Обнаружение существования скрытых данных ограничивает время на этапе, их удаления, необходимое для обработки только тех контейнеров, которые содержат скрытую информацию.

$$\nabla^2 C_{x,y} = C_{x+1,y} + C_{x-1,y} + C_{x,y+1} + C_{x,y-1} - 4 \cdot C_{x,y}.$$

**Таблица 3.1. Наиболее распространенные показатели визуального искажения, основанные на анализе пиксельной структуры контейнера**

| Разностные показатели искажения | |
|---|---|
| Максимальная разность (Maximum Difference) | $MD = \max_{x,y} \left| C_{x,y} - S_{x,y} \right|$ (3.1) |
| Средняя абсолютная разность *(Average Absolute Difference)* | $AD = \frac{1}{XY} \cdot \sum_{x,y} \left| C_{x,y} - S_{x,y} \right|$ (3.2) |
| Нормированная средняя абсолютная разность *(Normalized Average Absolute Difference)* | $NAD = \sum_{x,y} \left| C_{x,y} - S_{x,y} \right| \Big/ \sum_{x,y} \left| C_{x,y} \right|.$ (3.3) |
| Среднеквадратическая ошибка *(Mean Square Error)* | $MSE = \frac{1}{XY} \cdot \sum_{x,y} \left( C_{x,y} - S_{x,y} \right)^2.$ (3.4) |
| Нормированная среднеквадратическая ошибка *(Normalized Mean Square Error)* | $NMSE = \sum_{x,y} \left( C_{x,y} - S_{x,y} \right)^2 \Big/ \sum_{x,y} \left( C_{x,y} \right)^2.$ (3.5) |
| $L^p$-норма *($L^p$-norm)* | $L^p = \left( \frac{1}{XY} \cdot \sum_{x,y} \left| C_{x,y} - S_{x,y} \right|^p \right)^{1/p}.$ (3.6) |
| Лапласова среднеквадратическая ошибка *(Laplacian Mean Square Error)* | $LMSE = \sum_{x,y} \left( \nabla^2 C_{x,y} - \nabla^2 S_{x,y} \right)^2 \Big/ \sum_{x,y} \left( \nabla^2 C_{x,y} \right)^2.$ (3.7) |

| | |
|---|---|
| Отношение "сигнал/шум" *(Signal to Noise Ratio)* | $$SNR = \sum_{x,y}(C_{x,y})^2 \Big/ \sum_{x,y}(C_{x,y} - S_{x,y})^2. \quad (3.8)$$ |
| Максимальное отношение "сигнал/шум" *(Peak Signal to Noise Ratio)* | $$PSNR = XY \cdot \max_{x,y}(C_{x,y})^2 \Big/ \sum_{x,y}(C_{x,y} - S_{x,y})^2. \quad (3.9)$$ |
| Качество изображения *(Image Fidelity)* | $$IF = 1 - \sum_{x,y}(C_{x,y} - S_{x,y})^2 \Big/ \sum_{x,y}(C_{x,y})^2. \quad (3.10)$$ |
| **Корреляционные показатели искажения** | |
| Нормированная взаимная корреляция *(Normalized Cross-Correlation)* | $$NC = \sum_{x,y} C_{x,y} \cdot S_{x,y} \Big/ \sum_{x,y}(C_{x,y})^2. \quad (3.11)$$ |
| Качество корреляции *(Correlation Quality)* | $$CQ = \sum_{x,y} C_{x,y} \cdot S_{x,y} \Big/ \sum_{x,y} C_{x,y}. \quad (3.12)$$ |

| Другие показатели | |
|---|---|
| Структурное содержание *(Structural Content)* | $$SC = \sum_{x,y}(C_{x,y})^2 \Big/ \sum_{x,y}(S_{x,y})^2. \quad (3.13)$$ |
| Общее сигма-отношение "сигнал/шум" *(Global Sigma Signal to Noise Ratio)* | $$GSSNR = \sum_b \sigma_b^2 \Big/ \sum_b (\sigma_b - \tilde{\sigma}_b)^2, \quad (3.14)$$ где $\sigma_b = \sqrt{\frac{1}{n}\cdot \sum_{\text{блок } b}(C_{x,y})^2 - \left(\frac{1}{n}\cdot \sum_{\text{блок } b} C_{x,y}\right)^2}$. |
| Сигма-отношение "сигнал/шум" *(Sigma Signal to Noise Ratio)* | $$SSNR' = \frac{1}{n}\cdot \sum_b SSNR_b, \quad (3.15)$$ где $SSNR_b = 10\cdot \lg\left[\frac{\sigma_b^2}{(\sigma_b - \tilde{\sigma}_b)^2}\right]$. |
| Нормированное отношение "сигма/ ошибка" *(Normalized Sigma to Error Ratio)* | $$NSER = \frac{1}{\max(SER)}\cdot \sum_b SER_b, \quad (3.16)$$ где $SER_b = \sigma_b^2 \Big/ \left[\frac{1}{n}\cdot \sum_{\text{блок } b}(C_{x,y} - S_{x,y})^2\right]$. |
| Подобие гистограмм *(Histogram Similarity)* | $$HS = \sum_{c=0}^{255} \lvert f_C(c) - f_S(c)\rvert, \quad (3.17)$$ где $f_C(c)$ — относительная частота градации цвета $c$ в изображении с 256 уровнями цветов. |

Степень воспринимаемого человеком качества оперирует чувствительностью к контрасту и явлением маскировки системой восприятия человека и базируется на многоканальной модели человеческого пространственного зрения.

Шаги вычисления данного показателя:

- проведение крупношаговой сегментации изображения;
- разложение ошибки кодирования и первичного изображения на перцепционные (относящиеся к восприятию органами чувств) компоненты, используя гребенчатые фильтры;
- вычисление порога обнаружения для каждого пикселя, используя первичное изображение как маску;
- распределение фильтрованной ошибки с помощью порога принятия решения, объединение по всем цветовым каналам.

Единица измерения показателя определяется как единица *превышения порога,* что подразумевает под собой только значимое (заметное) отличие (***Just Noticeable Difference JND***). Общий показатель, скрытое максимальное соотношение "сигнал/шум" (***Masked Peak Signal to Noise Ratio — MPSNR***):

$$MPSNR = 10 \cdot \lg\left(\frac{255^2}{\varepsilon^2}\right),$$

где e — вычисленное искажение. Поскольку показатель качества не соответствует смыслу, заложенному в понятие "децибел", его называют *визуальным* или *зрительным децибелом* (ВдБ).

**Таблица 3.2. ITU-R Rec. 500. Оценка качества по шкале от 1 до 5**

| Оценка | Искажение | Качество |
|---|---|---|
| 5 | Незаметное | Отличное |
| 4 | Заметное, не раздражающее | Хорошее |
| 3 | Несущественно раздражающее | Удовлетворительное |
| 2 | Раздражающее | Неудовлетворительное |
| 1 | Чрезвычайно раздражающее | Крайне неудовлетворительное |

$$Q = \frac{5}{1 + N \cdot \varepsilon},$$

где $\varepsilon$ — вычисленное искажение; $N$ — нормирующий коэффициент, который зачастую выбирается таким, чтобы характеристика искажения отображала соответствующую качественную оценку.


**Классификация методов компьютерной стеганографии**

# Скрытие данных в неподвижных изображениях

- Основные свойства ЗСЧ, которые необходимо учитывать при построении стеганоалгоритмов
- Скрытие данных в пространственной области изображения
- Скрытие данных в частотной области изображения
- Методы расширения спектра
- Другие методы скрытия данных в неподвижных изображениях
  - Статистические методы
  - Структурные методы

**Скрытие данных в аудиосигналах**

**Скрытие данных в тексте**

# Скрытие данных в пространственной области изображения

## [1] – найти аналоги в [2]

.Метод замены наименее значащего бита

.Метод псевдослучайного интервала

.Метод псевдослучайной перестановки

.Метод блочного скрытия

.Методы замены палитры

.Метод квантования изображения

.Метод Куттера-Джордана-Боссена

.Метод Дармстедтера-Делейгла-Квисквотера-Макка

# Скрытие данных в изображения

# [2]

.Метод сокрытия с использованием младших битов элементов палитры

.Метод сокрытия с использованием младших битов элементов палитры

.Метод сокрытия, основанный на наличии одинаковых элементов палитры

.Метод сокрытия путем перестановки элементов палитры

# Скрытие данных в частотной области изображения

**[1]**

.Метод относительной замены величин коэффициентов ДКИ (метод Коха и Жао)

.Метод Бенгама-Мемопа-Эо-Юнг

.Метод Хсу и Ву

.Метод Фридрих

# Скрытие данных в аудиосигналах

## [1] – найти аналоги в [2]

.Кодирование наименее значащих бит (временная область)

.Метод фазового кодирования (частотная область)

.Метод расширения спектра (временная область)

.Скрытие данных с использованием эхо-сигнала

# Скрытие данных в тексте

# [1]

.Методы произвольного интервала

1. Метод изменения интервала между предложениями
2. Метод изменения количества пробелов в конце текстовых строк
3. Метод изменения количества пробелов между словами выровненного по ширине текста

.Синтаксические и семантические методы

# Digital Image Forensics

## 1. Format-Based Forensics

1.1: Fourier

1.2: JPEG

1.3: JPEG Header

1.4: Double JPEG

1.5: JPEG Ghost

# Digital Image Forensics

## 3. Pixel-Based Forensics

3.1: Resampling

3.2: Cloning

3.3: Thumbnails

## 4. Statistical-Based Forensics

4.1: Principal Component Analysis

4.2: Linear Discriminant Analysis

4.3: Computer Generated or Photographic?

4.4: Computer Generated or Photographic: Perception

# Digital Image Forensics

5. Geometric-Based Forensics

5.1: Camera Model

5.2: Calibration

5.3: Lens Distortion

5.4: Rectification

5.5: Composite

5.6: Reflection

5.7: Shadow

5.8: Reflection Perception

5.9: Shadow Perception

# Digital Image Forensics

## 8. Printer Forensics

8.1: Clustering

8.2: Banding

8.3: Profiling

# Digital Image Forensics

## 9. MatLab Code

9.1 JPEG Ghost

9.2 Color Filter Array (1-D)

9.3 Chromatic Aberration

9.4 Sensor Noise

9.5 Linear Discriminant Analysis

9.6 Lens Distortion

9.7 Rectification

9.8 Enhancement

9.9 Clustering

# Глоссарий

**Common Weakness Enumeration** (**CWE,** общий перечень слабостей) is a universal online dictionary of weaknesses that have been found in computer software.

**Secure Sockets Layer** (**SSL,** уровень защищённых сокетов) — криптографический протокол, который подразумевает более безопасную связь.

TSL

RC4 - RC4 обычно используется со 128-разрядными ключами, — эффективная разрядность этих ключей составляет всего 30 бит.

# Некоторые уязвимости из CWE, связанные с некачественным генерированием случайных чисел

CWE-330: Использование недостаточно случайных значений.

CWE-331: Недостаточная **энтропия**.

CWE-334: Сокращенное пространство случайных значений.

CWE-335: Ошибка **инициализации** ГПСЧ.

CWE-338: Использование криптографически слабого ГПСЧ.

CWE-340: Проблемы **прогнозируемости**.

CWE-341: **Прогнозируемость** по наблюдаемому состоянию.

CWE-342: **Прогнозирование** точного значения по предыдущим значениям.

CWE-343: **Прогнозирование** диапазона значений по предыдущим значениям.

## Разновидности генераторов случайных чисел

- Не-криптографические генераторы псевдослучайных чисел (не-криптографические ГПСЧ).
- Криптографические генераторы псевдослучайных чисел (КГСЧ).
- «Чистые» генераторы случайных чисел (ЧГСЧ), также называемые *энтропийными генераторами.*

# Генераторы случайных чисел

- Одной **статистической случайности** недостаточно - необходимо, чтобы атакующий не смог угадать генерируемые числа, даже если он видит **некоторую часть** серии.
- Конечная цель заключается в том, чтобы при **неизвестном начальном** числе атакующий не мог угадать генерируемые числа даже в том случае, если он располагает **большим количеством** уже сгенерированных чисел.
- Чтобы КГСЧ был безопасным, **начальное** число должно быть **трудно прогнозируемым**.

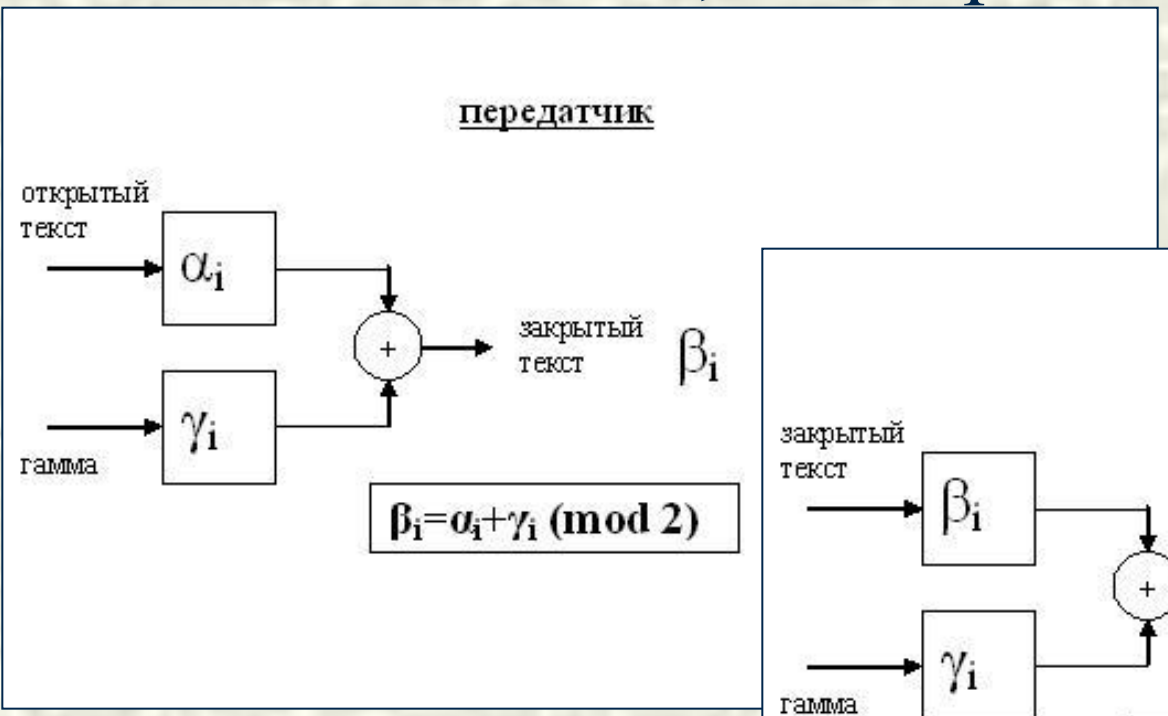
На практике это означает, что **безопасность КГСЧ** никогда не бывает **лучше безопасности начального числа**. Если атакующий с вероятностью 1/224 угадывает начальное число, то он сможет с вероятностью 1/224 угадать поток генерируемых чисел. В этом случае безопасность системы находится на 24-разрядном уровне, даже если нижележащий криптографический алгоритм обеспечивает 128-разрядную безопасность.

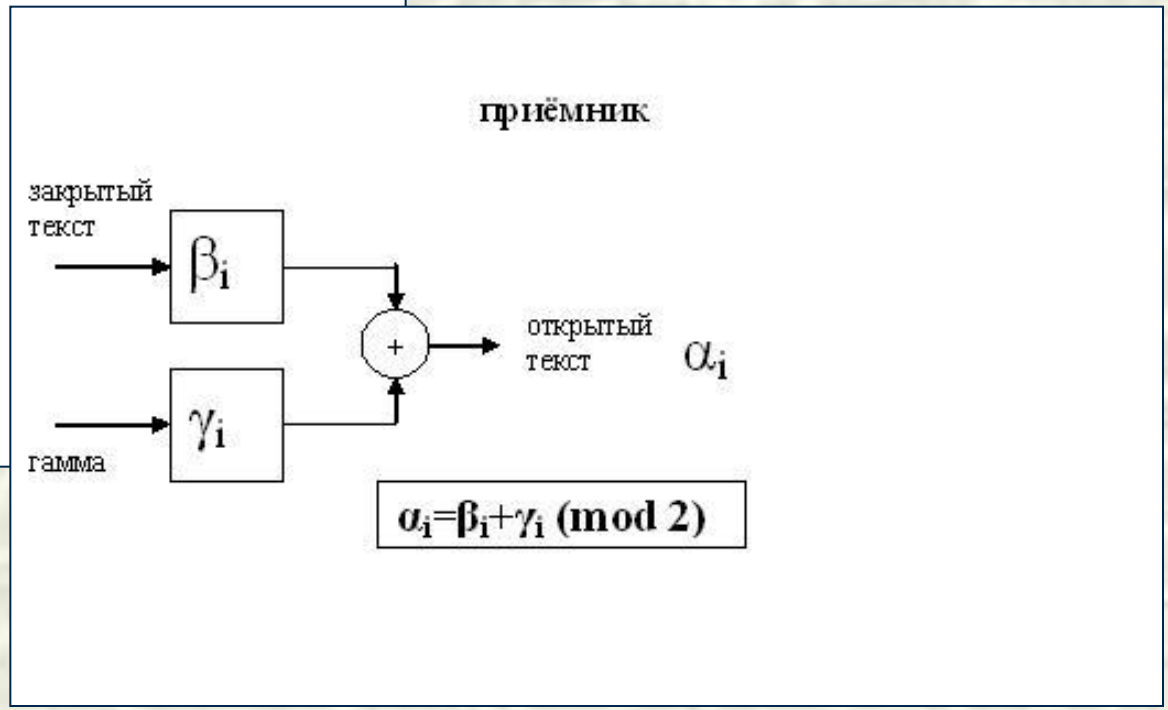
# Генераторы случайных чисел

| Некриптографические ГПСЧ | Криптографические ГПСЧ | |
|---|---|---|
| Строят длинную серию чисел по заданному начальному числу. Если несколько раз задать генератору одно начальное число, он сгенерирует одинаковые серии чисел. | | |
| •В большинстве не-криптографических генераторов существуют **232** возможных состояния. Каждый раз, когда программа сообщает пользователю **один** бит информации о случайном числе (обычно четное оно или нет), атакующий в общем случае может исключить **половину** состояний. Таким образом, при наличии даже минимальной информации полное состояние раскрывается после относительно небольшого | Единственное реальное различие заключается в том, что если атакующий не знает начальное число, вы можете сообщить ему первые 4000 сгенерированных чисел, и он не сможет угадать 4001-е число с вероятностью, сколько-нибудь превышающей | |

**Криптографические ГПСЧ** включают в себя инфраструктуру изменения начального числа; непосредственно КГСЧ - аналог поточных шифров;

**Гамми́рование** —метод симметричного шифрования, заключающийся в «наложении» последовательности, состоящей из случайных чисел, на открытый текст.



# Поточное шифрование. Сверточные коды.