Лабораторная работа №1

Тестирование псевдослучайных последовательностей

Цель работы: Научиться тестировать последовательность бит на равномерность и случайность

Теоретическая часть

Существует достаточно много различных наборов тестов, оценивающих «случайность» псевдослучайной последовательности. Наиболее известны тесты, описанные Кнутом, статистические тесты американского национального института стандартов и технологий NIST, батарея тестов Diehard и др. В этой лабораторной работе будут рассматриваться некоторые тесты из набора NIST.

Пусть дана последовательность, состоящая из n бит. Почти все эти тесты начинаются с преобразования входной последовательности 0 и 1 (будем обозначать ее ε) в последовательность -1 и 1 (будем обозначать ее X) соответственно:

$$X_i = 2 * \varepsilon_i - 1$$

Частотный тест

Этот тест оценивает пропорцию нулей и единиц в проверяемой последовательности. Тест определяет, является ли количество нулей и единиц в последовательности приблизительно таким же, как должно быть в истинно случайной последовательности.

Шаги алгоритма:

1. Входная последовательность, состоящая из 0 и 1 (будем обозначать ее ε), преобразовывается в последовательность -1 и 1 (будем обозначать ее X) соответственно:
$$X_i = 2 * \varepsilon_i - 1$$
2. Вычисляется сумма $S_n = X_1 + X_2 + X_3 + \cdots + X_n$, где n – количество элементов проверяемой последовательности.

3. Вычисляется статистика $S = \frac{|S_n|}{\sqrt{n}}$

4. Если $S \leq 1.82138636$, то тест считается успешно пройденным, иначе делается вывод о том, что последовательность является неслучайной.

Примечание: Если проверяемая последовательность не прошла данный тест, то проходить остальные тесты, проверяющие случайность, нет необходимости, т.к. уже ясно, что последовательность не является равномерно распределенной.

Тест на последовательность одинаковых бит

Этот тест анализирует количество цепочек в проверяемой последовательности, где цепочка – это непрерывная последовательность одинаковых бит. Под **цепочкой длиной k** понимается цепочка, состоящая из ровно k бит и ограниченная до и после битами с противоположным значением. Тест определяет, является ли количество цепочек из нулей и единиц различной длины в последовательности приблизительно таким же, как должно быть в истинно случайной последовательности

Шаги алгоритма:

1. Вычисляется частота, с которой в проверяемой последовательности встречаются единицы:
$$\pi = \frac{1}{n} * \sum_{j=1}^{n} \varepsilon_j$$
2. Вычисляется значение $V_n = 1 + \sum_{k=1}^{n-1} r(k)$, где r(k) = 0, если $\varepsilon_k = \varepsilon_{k+1}$ и r(k)=1 иначе
3. Вычисляется статистика $S = \frac{|V_n - 2*n*\pi*(1-\pi)|}{2*\sqrt{2*n}*\pi*(1-\pi)}$
4. Если $S \leq 1.82138636$, то тест считается успешно пройденным, иначе делается вывод о том, что последовательность является неслучайной.

Расширенный тест на произвольные отклонения

Этот тест оценивает общее число посещений определенного состояния при произвольном обходе кумулятивной суммы. Цель этого теста – определить отклонения от ожидаемого числа посещений различных состояний при произвольном обходе. Фактически данный тест состоит из 18 тестов, по одному для каждого состояния: $-9, -8, \dots -1, 1, 2, \dots, 9$.

Шаги алгоритма:

1. Входная последовательность, состоящая из 0 и 1 (будем обозначать ее ε), преобразовывается в последовательность -1 и 1 (будем обозначать ее X) соответственно:
$$X_i = 2 * \varepsilon_i - 1$$
2. Вычисляются суммы S_i последовательно удлиняющихся подпоследовательностей, начинающихся с X_1
$$S_1 = X_1$$
$$S_2 = X_1 + X_2$$
$$S_3 = X_1 + X_2 + X_3$$
$$\dots$$
$$S_n = X_1 + X_2 + X_3 + \dots + X_n$$
3. Формируется новая последовательность $S' = 0, S_1, S_2, \dots, S_n, 0$
4. Вычисляется $L = k - 1$, где k – количество нулей в полученной последовательности S'.
5. Для каждого из 18 состояний вычисляется ξ_j, которое показывает, сколько раз состояние j встречалось в последовательности S'. Здесь j = $-9, -8, \dots -1, 1, 2, \dots, 9$.
6. Вычисляются 18 статистик $Y_j = \frac{|\xi_j - L|}{\sqrt{2*L*(4*|j|-2)}}$ для каждого состояния j = $-9, -8, \dots -1, 1, 2, \dots, 9$.
7. Если все статистики $Y_j \leq 1.82138636$, то тест считается успешно пройденным, если же хотя бы для одной статистики Y_j это условие не выполнилось, то делается вывод о том, что последовательность является неслучайной.

Задание

Реализовать приложение, позволяющее выполнять следующие действия:

1. Задавать длину генерируемой последовательность в битах (при тестировании рекомендуется задавать длину последовательности не менее 10 000 бит)
2. Генерировать псевдослучайную последовательность 0 и 1 с помощью стандартного алгоритма генерации случайных чисел
3. Загружать последовательность из текстового файла
4. Сохранять полученную последовательность в файл и выводить ее на экран приложения
5. Проверять полученную последовательность с помощью реализованных тестов. Результат проверки должен отображаться в приложении

Дополнительные требования к приложению

- Программа должна быть оформлена в виде удобной утилиты с интерактивным интерфейсом пользователя
- Текст программы оформляется прилично (удобочитаемо, с описанием ВСЕХ функций, переменных и критических мест).
- В процессе работы программа ОБЯЗАТЕЛЬНО выдает информацию о состоянии процесса генерации / тестирования (если процесс занимает длительное время)
- Интерфейс программы может быть произвольным, но удобным и понятным (разрешается использование библиотек GUI)
- Среда разработки и язык программирования могут быть произвольными

Примечания:

1. Задание является дифференцированным.
   - На оценку «Удовлетворительно» достаточно реализовать только «Частотный тест»
   - На оценку «Хорошо» необходимо реализовать «Частотный тест» и «Тест на последовательность одинаковых бит»
   - На оценку «Отлично» необходимо реализовать все три теста: «Частотный тест», «Тест на последовательность одинаковых бит» и «Расширенный тест на произвольные отклонения»
2. В первой лабораторной работе разбивка по вариантам не предусматривается.

Требования для сдачи лабораторной работы:

- Демонстрация работы, реализованной вами системы.
- АВТОРСТВО
- Теория (ориентирование по алгоритмам и теоретическим аспектам методов тестирования)
- Оформление и представление письменного отчета по лабораторной работе, который содержит:
  - Титульный лист
  - Задание на лабораторную работу
  - Описание используемых алгоритмов шифрования
  - Листинг программы