РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА СТАТИСТИЧЕСКОГО ТЕСТИРОВАНИЯ КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ НА ОСНОВЕ МАРКОВСКИХ МОДЕЛЕЙ

М. Ю. Деркач

Белорусский государственный университет, г. Минск;
fpm.derkach@bsu.by
науч. рук. – Ю. С. Харин, д-р физ.-мат. наук, чл.-корр. НАН Беларуси

Проблема защиты информации затрагивает практически все сферы деятельности человека. Среди способов защиты информации важнейшим считается криптографический [1]. Надежность любой системы криптографической защиты информации (СКЗИ) в значительной степени определяется качеством используемых генераторов случайных и псевдослучайных последовательностей.

Генератор, используемый в СКЗИ, должен порождать выходную последовательность, неотличимую от равномерно распределенной случайной последовательности (РРСП) [1]. Для обнаружения отклонения от модели РРСП используются статистические тесты. Статистические свойства последовательностей определяются числовыми характеристиками. На основе оценочных критериев делаются заключения о степени близости свойств анализируемой случайной последовательностей и РРСП.

Для выявления зависимостей высокого порядка и выявления скрытых зависимостей требуются дополнительные исследования. Основными математическими моделями, используемыми в таких исследованиях, являются марковские модели. В НИИ ППМИ БГУ были разработаны методы и алгоритмы статистического тестирования выходных последовательностей, основанные на цепи Маркова порядка s с r частичными связями (ЦМ(s,r)) и цепи Маркова условного порядка (ЦМУП).

Данный доклад посвящен разработке программного комплекса, реализующего эффективные алгоритмы статистического анализа выходных последовательностей, основанные на оценивании таких марковских моделей, как однородная цепь Маркова, однородная цепь Маркова порядка s, скрытая марковская модель, двойная марковская модель.

Ключевые слова: Марковские модели; криптографические генераторы; алгоритмы статистического тестирования; алгоритмы статистического оценивания параметров марковских моделей.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Использовались следующие математические модели входных последовательностей $x_t \in A$:

M_1: Однородная цепь Маркова.

M_2: Однородная цепь Маркова S-ого порядка.

M_3: Скрытая марковская модель.

M_4: Двойная марковская модель.

МЕТОДЫ СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ ПАРАМЕТРОВ И ПРОВЕРКИ ГИПОТЕЗ

Для статистического оценивания параметров этих моделей использовались следующие алгоритмы:

1. Метод оценки максимального правдоподобия для моделей M_1, M_2.
1. Метод статистического бутстрэпа для моделей M_1, M_2.
2. Метод сглаживания оценки максимального правдоподобия для моделей M_1, M_2.
3. Обобщенный EM-алгоритм (алгоритм Баума-Велша) для моделей M3, M4.
4. Метод оценивания оптимальной последовательности скрытых состояний для M3, M4.

Введем гипотезы $H_0 =$ {гипотеза о том, что выходная последовательность – «чисто случайная} = {гипотеза о том, что выходная последовательность – РРСП} и $H_1 = \overline{H_0}$.

В параметрическом виде гипотеза H_0 представима следующим образом:

$$P^0 = (p^0{}_{ij}),\ p^0{}_{ij} \equiv \frac{1}{N}, i, j \in A,$$

$$C^0 = (c^0{}_{ij}),\ c^0{}_{ij} \equiv \frac{1}{M}, i \in A,\ j \in B.$$

H_0:

Рис.1. Структура программного комплекса

РЕЗУЛЬТАТЫ НА МОДЕЛЬНЫХ ДАННЫХ

Используя модуль генерации модельных данных были сгенерированы следующие выходные последовательности:

$$S_1 = \{\pi_1, P_1\} : \pi_1 = (0.5, 0.5)^T, P_1 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, T_1 = 20.$$

$$S_2 = \{\pi_2, P_2\} : \pi_2 = (0.5, 0.5)^T, P_2 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, T_2 = 100.$$

$$S_3 = \{\pi_3, P_3\} : \pi_3 = (0.5, 0.5)^T, P_3 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, T_3 = 100000.$$

$$S_4 = \{\pi_4, P_4\} : \pi_4 = (0.45, 0.55)^T, P_4 = \begin{pmatrix} 0.43 & 0.57 \\ 0.58 & 0.42 \end{pmatrix}, T_4 = 100.$$

$$S_5 = \{\pi_5, P_5\} : \pi_5 = (0.45, 0.55)^T, P_5 = \begin{pmatrix} 0.43 & 0.57 \\ 0.58 & 0.42 \end{pmatrix}, T_5 = 1000.$$

$$S_6 = \{\pi_6, P_6\} : \pi_6 = (0.53, 0.47)^T, P_6 = \begin{pmatrix} 0.73 & 0.27 \\ 0.20 & 0.80 \end{pmatrix}, T_6 = 20.$$

$$S_7 = \{\pi_7, P_7\} : \pi_7 = (0.53, 0.47)^T, P_7 = \begin{pmatrix} 0.73 & 0.27 \\ 0.20 & 0.80 \end{pmatrix}, T_7 = 100.$$

	Данные	Длина последовательности	Результаты тестирования			
			M_1	M_2	M_3	M_4
	S_1	20	H_0	H_1	H_1	H_1
	S_2	100	H_0	H_0	H_0	H_0
	S_3	100000	H_0	H_0	H_0	H_0
	S_4	100	H_0	H_0	H_0	H_0
	S_5	1000	H_1	H_1	H_1	H_1
	S_6	20	H_0	H_0	H_0	H_0
	S_7	100	H_1	H_1	H_1	H_1

РЕЗУЛЬТАТЫ НА РЕАЛЬНЫХ ДАННЫХ

Для экспериментов использовались следующие выходные последовательности физического генератора, полученные с сайта Humboldt University of Berlin (http://qrng.physik.hu-berlin.de/download):

$$S_1, N = 2, T_1 = 8388608(1MB).$$

$$S_2, N = 2, T_2 = 125829120(15MB).$$

$$S_3, N = 2, T_3 = 125829120(100MB).$$

№	Данные	Длина последовательности	Результаты тестирования			
			M_1	M_2	M_3	M_4
1	S_1	20	H_0	H_0	H_0	H_0
2	S_2	100	H_0	H_0	H_0	H_0
3	S_3	100000	H_0	H_0	H_0	H_0

Библиографические ссылки

1. *Харин Ю. С. и др.* Криптология. 2013. Минск.
2. *Berchtold A.* The double chain Markov model. 1999. Department of Statistics University of Washington Seattle.