УДК 681.325

А.А. Горбунов

Алгоритмы структурной идентификации математических моделей криптосистем на основе определения базовых параметров

Описывается подход получения характеристик криптосистем при помощи методов идентификации математических моделей источников экспериментальных данных. Рассматриваются способы практической реализации предложенных алгоритмов идентификации.

Ключевые слова: алгоритм, идентификация, криптосистема, параметр, модель.

На современном уровне решения задач по моделированию систем защиты информации всё более актуальное звучание получают вопросы, связанные с областью средств криптографической защиты информации. При этом разработка соответствующего математического аппарата позволяет на адекватном уровне оценивать параметры стойкости различных криптосистем (КС), осуществлять сравнительный анализ их характеристик.

Существуют различные методы идентификации математических моделей (ММ) дискретных динамических объектов как источников данных или всевозможных преобразователей текстовых последовательностей, являющихся основой разнообразных систем управления. Оценивание параметров КС через методы структурной идентификации их ММ базируется на алгоритмах определения характеристик тех последовательностей, которые экспериментально были измерены (сняты) в контрольных точках тех или иных блоков КС (например, вход и выход шифратора, дешифратора, канала передачи). Описываемые в настоящей статье алгоритмы структурной идентификации ММ КС основаны на представлении шифратора криптосистемы (и соответствующего дешифратора) в виде гипотетического источника экспериментальных данных. При данном подходе на входе и выходе шифратора КС имеем последовательности открытого $\boldsymbol{u}(t) \equiv \boldsymbol{u}_t = \{u_0, u_1, \ldots, u_{M-1}\}$ и зашифрованного $\boldsymbol{y}(t) \equiv \boldsymbol{y}_t = \{y_0, y_1, \ldots, y_{M-1}\}$ (синхронного с открытым сообщением) текстов ($t = 0, 1, \ldots, M-1$).

В работе [1] введена ММ нестационарных источников экспериментальных данных. Для каждого из участков стационарности текстов, описываемых указанной моделью, определяется набор *базовых параметров* (БП) ММ источника текста. БП изначально квантованного по времени с интервалом $\Delta t = 1$ текстового сигнала представляются парой $\text{БП} = \{q, n\}$, где q – число уровней квантования текста; n – «сложность» гипотетического источника стационарного участка данных (или память автомата, порождающего текст на данном участке стационарности). По реализации сигнала (в данном случае текстового) Источника могут быть найдены значения оптимальных БП ($\text{ОБП} = \{q = q_{\text{opt}}, n = n_{\text{opt}}\}$), исходя из критерия минимизации энтропийной функции $E\big(q, n(q)\big) = n(q) \cdot \log q$.

В рамках рассмотрения с подобной точки зрения моделей шифратора (и соответствующего дешифратора) КС как неавтономных источников текстовых данных (см., например, [2]) появляется возможность оценивать параметры этих ММ как неавтономных синхронных дискретных автоматов; причём могут быть получены оценки оптимальных БП не только скалярных, но также и векторных входных $\boldsymbol{u}_t = \left[u_t^{(1)} \quad \ldots \quad u_t^{(k)}\right]^T$ и выходных $\boldsymbol{y}_t = \left[y_t^{(1)} \quad \ldots \quad y_t^{(r)}\right]^T$ сигналов.

Проводя структурную идентификацию ММ блоков КС по измеряемым практически экспериментальным данным и определяя их ОБП, можно получать различные характеристики изучаемых источников текстовых сигналов (например, для случая шифратора КС): входного, выходного, а также взаимного $\boldsymbol{uy}_t = \left[u_t^{(1)} \quad \ldots \quad u_t^{(k)}\, y_t^{(1)} \quad \ldots \quad y_t^{(r)}\right]^T$ векторных текстов. В целях практического нахождения значений таких параметров КС, необходимо реализовать эффективный алгоритм поиска ОБП на измеренных текстовых экспери-

ментальных данных. Для нахождения минимального значения функции энтропии $E(q, n(q))$ источника текста границы области поиска $q_{\min} \le q \le q_{\max}$ базового параметра q задаются исходя из априорных сведений об идентифицируемой модели объекта.

При этом реализацией алгоритма вычисления базового параметра $n = n(q)$ должен обеспечиваться по возможности наиболее быстродействующий вариант проверки соответствия значения этого параметра понятию «сложности» источника экспериментальных данных. Указанное соответствие обеспечивается для такого минимального значения n, при котором идентифицируемый источник текста являлся бы непротиворечивым прогнозирующим оператором рассматриваемого порядка n его текстовой q-уровневой последовательности. Это означает, что, например, для входного текстового сигнала $\boldsymbol{u}_t$ любым n_u идущим подряд символам $\{\ldots, u_{t-n_u+1}, \ldots, u_{t-1}, u_t, \ldots\}$ всегда должен соответствовать только единственный вариант последующего символа u_{t+1}, прогнозируемого по данной n_u-после-довательности.

Таким образом, возникает задача выявления совпадающих между собой n-последовательностей символов (длиной n_u, n_y и n_{uy} символьных отсчетов для открытого $\boldsymbol{u}_t$, шифрованного $\boldsymbol{y}_t$ и взаимного $\boldsymbol{uy}_t$ текстов соответственно) с целью дальнейшей проверки непротиворечивости последующих символов, прогнозируемых по ним. Можно выделить, по крайней мере, три подхода построения алгоритмов для решения этой задачи в зависимости от способа организации доступа к набору всех n-последовательностей, получаемых по анализируемой текстовой последовательности данных.

1. Алгоритмы, основанные на непосредственном переборе всех n-последовательностей в тексте. Здесь при последовательном посимвольном просмотре изучаемой текстовой последовательности сравнивается набор из n символов, относящийся к текущему просматриваемому символу со всеми предыдущими n-символьными наборами. Так как время работы подобных алгоритмов можно оценить как $O(M^2)$, то область их практической применимости ограничивается сравнительно короткими текстовыми последовательностями. При этом дополнительные затраты памяти на организацию подобного перебора n-последовательностей в составе имеющегося в распоряжении текста, как правило, минимальны.

2. Бинарный поиск среди n-последовательностей. При реализации в рамках данного подхода алгоритмов поиска совпадающих n-последовательностей последние индексируются в соответствии с некоторым правилом упорядочивания. В результате появляется возможность для каждой очередной n-последовательности осуществлять поиск другой совпадающей с ней n-последовательности среди ранее просмотренных за логарифмическое время. Общее же время работы таких алгоритмов можно оценить уже как $O(M \cdot \log M)$, что позволяет говорить о возможности производить при помощи их обработку значительно более длинных последовательностей в сравнении с алгоритмами непосредственного перебора. Однако при данном подходе построения алгоритмов возникают дополнительные затраты памяти, связанные с поддержкой упорядоченности обработанных n-последовательностей (например, для размещения в памяти ссылок на n-последовательности символов текста, организованных в виде сбалансированного бинарного дерева).

3. Построение суффиксного дерева по обрабатываемой текстовой последовательности (например, при помощи алгоритма Укконена [3]) позволяет за линейное относительно её длины время определить минимальное значение величины n, при котором у совпадающих n-последовательностей отсутствует противоречие в последующих (прогнозируемых) символах. Таким образом, и всё время работы подобных алгоритмов можно оценить как $O(M)$, что, с теоретической точки зрения, является неким нижним пределом для всех алгоритмов, обращающихся к каждому символу обрабатываемого текста. Тем не менее, практическая реализация построения суффиксного дерева требует существенных затрат памяти как при достаточно большой длине M обрабатываемой текстовой последовательности, так и при большой размерности q её алфавита. Указанное обстоятельство до определенной степени ограничивает универсальность алгоритмов, реализуемых в рамках данного подхода.

Литература

1. Кирьянов К.Г. Выбор оптимальных базовых параметров источников экспериментальных данных при их идентификации // Труды III Междунар. конф. «Идентификация систем и задачи управления SICPRO'04». – М.: ИПУ РАН, 2004. – С. 187–208.
2. Горбунов А.А. Связь функции ненадёжности и расстояния единственности криптосистем с базовыми параметрами шифратора в форме математической модели синхронного автомата Хаффмана / А.А. Горбунов, К.Г. Кирьянов // Вестник Нижегор. ун-та им. Н.И. Лобачевского. Сер. Радиофизика. – 2005. – Вып. 1(3). – С. 185–193.
3. Гасфилд Д. Строки, деревья и последовательности в алгоритмах. – СПб.: Невский диалект; БХВ-Петербург, 2003. – 654 с.

Горбунов Александр Александрович
ГОУ ВПО «Нижегородский государственный университет им. Н.И. Лобачевского»,
ассистент Центра БИСК радиофизического факультета
Эл. адрес: aagor@rf.unn.ru

A.A. Gorbunov

Structural identification algorithms of cryptosystems' mathematical models by way of base parametrs finding

The paper deals with approach to finding cryptosystems' characteristics by way of identification for mathematical models of the experimental data Sources. The practical realization ways of offered identification algorithms is considered.

Keywords: algorithm, identification, cryptosystem, parameter, model.