В дальнейшем планируется реализация возможности сокрытия зашифрованных данных посредством алгоритмов стеганографии, а также внедрение в систему функции добавления в зашифрованный файл скрытых цифровых водяных знаков. В итоге, конечный круг реализованных функций позволит снизить вероятность несанкционированного доступа к данным, и как следствие, повысить уровень её защищенности.

Литература

1. Панасенко, С. Алгоритмы шифрования. Специальный справочник [Текст] / С. Панасенко. – Санкт-Петербург: БХВ-Петербург, 2009. – 578 с.

2. Глухих, В.И. Информационная безопасность и защиты данных [Текст] / В.И. Глухих. – ИГТУ, 2011. – 248 с.

3. Блог Лаборатории Касперского [Электронный ресурс]. – URL: https://blog.kaspersky.ru/encryption-reasons/879/

4. Криптография и защита данных [Электронный ресурс]. – URL:http://www.crypto.com/report.html

5. Мобильная безопасность: Защита мобильных устройств в корпоративной среде [Электронный ресурс]. –URL: https://xakep.ru/2011/10/13/57058/

6. Шнайер, Б. Прикладная криптография [Текст] / Брюс Шнайер. – Триумф, 2012. – 815 с.

А.А. Сытник, И.В. Гвоздюк

ОБ ОДНОМ ПОДХОДЕ К АВТОМАТНОМУ МОДЕЛИРОВАНИЮ ПОВЕДЕНИЯ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМ

(Саратовский государственный технический университет имени Гагарина Ю.А)

Конечные автоматы составляют один из важнейших классов математических моделей для дискретных систем с конечным множеством состояний. Практическое и теоретическое значение автоматных моделей в решении задач проектирования и эксплуатации информационно-коммуникационных систем, формальных языков и трансляторов стало причиной интенсивных исследований по теории автоматов. Конечные автоматы, по сути, являются, практически, единственной математической моделью, способной эффективно использоваться при ппопытках формализации сложных информационных систем и программного обеспечения. Разнообразие возникших задач, подходов к их решению, научных позиций исследователей привело к выделению классов автоматов (автоматы типов Мили и Мура, автоматы Медведева, автономные автоматы, автоматы с конечной глубиной памяти, (n, m, l) – автоматы и т.д.), а также к разработке различных математических способов их задания (табличное задание, графы автоматов, автоматные матрицы, логические уравнения, формулы языка регулярных выражений, задание автомата композицией автоматов).

В данной статье исследуется задача распознавания автомата в заданном конечном семействе КДА $\alpha = \{ A_i \}_{i \in I}$, где для каждого $i \in I$ $A_i = (S_i , X, Y, \delta_i , \lambda_i)$, средствами условного эксперимента со специальным типом функционирования - функционирования на основе приложения чередующихся периодических входных последовательностей

$$P_1^{m_1} \; P_2^{m_2} \; \dots \; P_\alpha^{m_\alpha} , \text{ где } \; P_j^{m_j} \in X^* , \; m_j \in R^+ , \; 1 \le j \le \alpha .$$

Для интерпретации в приложениях в условия задачи неявно включается предположение о том, что автоматы из семейства α имеют множества состояний S_i, $i \in I$, с большим числом элементов. Разрабатываемый метод распознавания ориентируется на обоснованное сокращение множества состояний, реакции на которые учитываются при реализации метода распознавания (а не при построении распознающего эксперимента).

В данной работе, в отличии от теоремы Э.Мура явно выделяются свойства, на которых базируется решение установочной задачи:

- совпадение p - приемников у различных состояний автомата;
- несовпадение наблюдаемых реакций на одно и то же входное слово для различных состояний.

Первое свойство в частных случаях позволяет решать установочную задачу без наблюдения выходных слов. Критерием такого решения установочной задачи является условие

$$(\forall s , s` ) ( \exists \; p \in X^* ) \;\; \delta ( s , p ) \ne \delta ( s` , p).$$

Второе свойство оказывается основным и существенным при сведении распознавания автомата в заданном семействе автоматов к установочной задаче для "расщепляемого" (по А.Гиллу) автомату для семейства. В теореме 13 представлены оба случая проявления свойств состояний автоматов.

Исследованный метод впервые был указан в работе А.М. Богомолова и В.А. Твердохлебова [7]. В этой работе (теорема 23) показана общая структура процесса функционирования конечного детерминированного автомата под воздействиями периодических последовательностей. Здесь же приводится пример исключения из анализа вершин, не входящих в циклы, порождаемые приложением периодических последовательностей. Полного исследования, отвечающего на вопросы:

- как выбирать входные слова для формирования периодических входных последовательностей?
- для каких автоматов метод эффективен?
- как строить эксперимент по распознаванию автомата на основе приложения периодических входных последовательностей?
- какими алгоритмами возможна реализация условного эксперимента, построенного на основе этого метода?

- как оценить "неполноту" метода по отношению к методу (теоретически предполагаемых) установочного дерева?

На эти и другие вопросы в указанной работе ответов нет. В ряде дальнейших исследований, рассматриваются технические аспекты контроля и диагностирования с приложением периодических входных последовательностей. Однако окончательной теоретической проработки указанного подхода к диагностированию не содержится.

Закономерность связи ограничения воздействий на вход автомата (например, сведением к воздействию периодическими входными последовательностями) с ограничениями на функционирование и реакции автомата отмечалась специалистами по технической диагностике, но систематического исследования, закрывающего проблему, нет. В диссертации разработан метод периодических воздействий на вход автомата с целью выделения фрагментов его функционирования, позволяющих сокращать объем анализируемой информации с наименьшей (для периодических воздействий) потерей эффективности распознавания автомата в заданном семействе автоматов. Специфическим является сокращение анализируемой (и требующей построения) информацией как при разработке условного эксперимента, так и информации, используемой при фактическом проведении условного эксперимента.

При построении для конечного детерминированного автомата $A = (S, X, Y, \delta, \lambda)$ графа Gp зацикливания автомата по периодической входной последовательности p^m характерными путями в графе являются :

- пути с наибольшей длиной m_1 в цикле графа;
- пути длины $m = m_1 + m_2$ в графе Gp.

Информацию о заключительном состоянии автомата, получаемую как наблюдаемую реакцию на входную последовательность p^m, представляет слово $\lambda^0 (\delta (s, p^{m1}), p^{m2})$, где s - состояние, p^m преемником которого является заключенное состояние. Выходное слово по предположению определяется как проекция слова $\lambda (\delta (s, p^{m1}), p^{m2})$.

Ранее проекцию составляла первая буква, то есть $pr_1 \lambda (\delta (s, p^{m1}), p^{m2})$.

Это не обязательное ограничение. В общем случае будем рассматривать в качестве характеристики заключительного состояния слово

$$pr_{i, i2 .. ik} \lambda (\delta (s, p^{m1}), p^{m2}) ,$$

где $1 \le i_1 < i_2 < ... < i_k \le |p^{m2}|$ и $1 \le k \le |p^{m2}|$.

При решении задачи распознавания автомата в конечном семействе конечных детерминированных автоматов $\alpha = \{A_i\}_{i \in I}$, где для каждого $i \in I$ $A_j = (S_j , X, Y, \delta_j , \lambda_j)$ одним из фундаментальных методов является сведение задачи к установочной задаче для расщепляемого (по терминологии А.Гилла]) автомата

$$ą = (\bigcup_{i \in I} S_i , X, Y, \bigcup_{i \in I} \delta_i , \bigcup_{i \in I} \lambda_i)$$

В случае, когда множеств состояний $S = \bigcup_{i \in I} S_i$ расщепляемого автомата ą велико, практическое решение задачи распознавания (например, при техниче-

ском диагностировании на основе распознавания автоматов) оказывается невозможным. При сохранении основных идей, на которых построено сведение задачи, можно распознавать конечный детерминированный автомат приближенно с сокращением числа анализируемых состояний представленный конечным недетерминированным автоматом. На принципиальную возможность разработки такого метода указано в работе А.М. Богомолова, В.А. Твердохлебова [7].

Пусть $A = (S, X, Y, \delta, \lambda)$ - конечный детерминированный автомат. Как показано в работе [4], для любого входного слова $p \in X^*$ связи состояний из множества состояний S, представленные состояниями и их p - преемниками, определяются графом со свойствами:

- граф имеет конечное число связных подграфов;
- каждый подграф обязательно имеет один цикл (или петлю), каждая вершина которого может быть корнем дерева:
- граф ориентированный с направлением дуг от висячих вершин деревьев к их корнями с обходом контуров.

Основная идея сокращения числа состояний автомата, требующих анализа при распознавании, состоит в исключений из анализа всех вершин деревьев, кроме корней деревьев. Корни деревьев образуют циклы или петли. Следовательно, в предлагаемом методе распознавания анализируется наблюдаемое поведение автомата, представленное циклами и переходимы из одного цикла в другой.

Рассмотрим основные положения метода.

Функционирование автомата на основе его зацикливания.

Пусть $A = (S, X, Y, \delta, \lambda)$ - конечный детерминированный автомат. Во множестве вариантов его функционирования выделим такие, которые определяются следующими ограничениями:

1. К автомату A прикладываются только периодические входные последовательности вида p^m , где $p \in X^*$ и $m \in N^+$.
2. При эксперименте с автоматом A наблюдаются только некоторые выходные сигналы, выдаваемые автоматами при изменении его состояний в цикле.

Таким образом, функционированию автомата сопоставляется его проекция, в которой представлены

- связи состояний s и $\delta (s, p)$.
- выходные слова, образованные выходными сигналами $pr_{|p|}$ $\lambda (s, p)$ для состояний, входящих в цикл.

Сокращение действий по наблюдению и анализу выходных сигналов, выдаваемых автоматом до достижения циклов и наблюдение только последнего выходного сигнала для каждого приложения слова p, связано с потерей информации об объекте эксперимента.

В связи с этим возникают новые критерии существования решений установочных задач и задач распознавания автоматов, новые методы построения

экспериментов и оценок их длины. Выделяются критерии эффективности методов.

Основную идею разработанного метода составляет построение по заданному конечному детерминированному автомату $A = (S, X, Y, \delta, \lambda)$ такого автомата Ω , у которого:

1. Входными сигналами являются слова $p_i \in X^*$ используемые как периоды периодических входных последовательностей.
2. Выходными сигналами являются слова $q_j \in Y^*$, построенные выделением некоторых выходных сигналов, выдаваемых автоматом при изменении его состояний в цикле.
3. Множеством состояний автомата полагаются только те состояния, которые образуют циклы.
4. Задачи для автомата A решаются как задачи для автомата

$$\Omega = ([S], \{p_i , i \in I\} , \{q_j , j \in J\} , \tilde{\delta} , \tilde{\lambda})$$

где [S] – множество состояний, входящих в циклы, а $\tilde{\delta}$ и $\tilde{\lambda}$ - новые функции переходов и выходов.

Для автомата $A = (S, X, Y, \delta, \lambda)$ используются два способа расширения функции выходов λ. В первом случае функция λ расширяется до функции вида $\lambda : S \times X^* \to Y^*$ по правилу:

$(\forall s \in S) (\forall x \in X) (\forall p \in X^*) \lambda (s, x p) = \lambda (s, x) \lambda (\delta (s, x), p)$

С целью уменьшения объема наблюдаемых выходных сигналов автомата, используется также расширение функции λ до функции $\lambda`$ вида $\lambda` : S \times X^* \to Y$ при котором наблюдается только последний выходной сигнал:

$(\forall s \in S) (\forall x \in X) (\forall p \in X^*) \lambda` (s, x p) = \lambda (\delta (s, p), x)$

где функция δ расширена до функции вида $\delta : S \times X^* \to Y$ по правилу:

$(\forall s \in S) (\forall x \in X) (\forall p \in X^*) \delta (s, x p) = \delta (\delta (s, x), p)$

Предположим, что $X = \{ x_1, x_2 , \dots , x_m \}$.Входная последовательность x^k_i при достаточно большом k (k > n) для любого начального состояния сформирует последовательность состояний

$$s, \delta (s, x_i), \delta (s, x^2_i), \dots , \delta (s, x^k_i), \qquad (1)$$

в которую обязательно входит цикл. Рассматривая последовательность (1), как путь в графе и совмещая все пути с одним и тем же циклом, получаем граф G_{x_i} , состоящий из связных подграфов:

$$G^1_{x_i}, G^2_{x_i}, \dots, G^{w_i}_{x_i}$$

соответственно с циклами $C^1_{x_i}, C^2_{x_i}, \dots, C^{w_i}_{x_i}$.

В предлагаемом методе распознавания поведения автомата A может быть приближено (с потерей некоторых вариантов функционирования) представлено поведением на циклах $C^1_{x_i}, C^2_{x_i}, \dots, C^{w_i}_{x_i}$ и исключением из анализа вершин, не входящих в циклы.

Литература

1. Салин В.С., Папшев С.В., Сытник А.А. Практическое применение метода BorderFlow в задаче автоматизированной семантической кластеризации веб-сайта. // Научно-методический журнал «Информатизация образования и науки» № 3(27)/2015. ФГАУ ГНИИ ИТТ «Информика». С. 65-73.
2. Федеральное агентство по техническому регулированию и метрологии ГОСТ-28806-90: Качество программных средств. Термины и определения // Информационный портал по стандартизации. – Стандартинформ, 2017 – Режим доступа: http://standard.gost.ru/ (дата обращения: 10.01.2017).
3. Alexander A. Sytnik, Sergey V. Papshev. Semantic Segmentation of Hypertext on the Basis of Automata Model. International Journal of Computing Anticipatory Systems, v. 28, 2014, D.M. Dubois (Ed.), CHAOS, Liège, Belgium, ISSN 1373-5411, ISBN 2-930396-17-2. P.109-115.
4. Сытник А.А., Шульга Т.Э. Математические модели адаптивных дискретных систем. Монография //Саратов: Сарат. гос. техн. ун-т, 2015. 272с. ISBN 978-5-433-2947-2.
5. Сытник А.А. Перечислимость при восстановлении поведения автоматов //Доклады РАН. 1993. Т.238. N1. С.25-26
6. Богомолов А.М., Твердохлебов В.А. Диагностика сложных систем. Киев. Наукова Думка. 1974. 128 с.
7. Богомолов А.М., Твердохлебов В.А. Целенаправленное поведение автоматов. Киев. Наукова Думка. 1975. 123 с.
8. Сытник А.А. Методы и модели восстановления поведения автоматов. //Автоматика и телемеханика. 1992. N 11.

А.А. Сытник, С.В. Папшев, Т.Э. Шульга

ОБ ОДНОМ ПОХОДЕ К СЕМАНТИЧЕСКОЙ КЛАСТЕРИЗАЦИИ

(Саратовский государственный технический университет имени Гагарина Ю.А)

Аннотация – В статье предлагается решение актуальной проблемы семантической кластеризации нетекстовых веб-документов за счет использования статистики посещения и гиперссылок. Результаты дополняют известные методы семантической кластеризации текстовых документов и предоставить возможность классифицировать текстовые и нетекстовые объекты в рамках единой системы на предварительном этапе интеллектуальной обработки данных на основе автоматных моделей.

Ключевые слова – Семантическая кластеризация, семантический веб, гипертекст, нетекстовый документ, автомат, система, модель; алгоритм.

Abstract: The article proposes the solution of the actual problem of semantic clustering of non-text web documents by using statistics of visits and hyperlinks. The results complement the known methods of semantic clustering of text documents and provide an opportunity to classify text and non-text objects within a single system at