УДК 004.056 + 004.056.55 Вестник УрФО № 3(21) / 2016, с. 24–29

Стюгин М. А.

*УСЛОВИЯ СУЩЕСТВОВАНИЯ АБСОЛЮТНО ЗАЩИЩЕННЫХ ОТ ИССЛЕДОВАНИЯ СИСТЕМ**

В данной статье рассматривается метод построения защищенных информационных систем путем затруднения их исследования со стороны потенциального злоумышленника. Разработана формализованная модель исследователя по отношению к объекту, имеющему выход, некоторое количество входов и набор дискретных внутренних состояний. С использованием данной модели можно моделировать любые цифровые информационные системы в виде алгоритмов, интерфейсов и пр. Дано определение защищенной от исследования системы и абсолютно защищенной от исследования системы. Получена теорема условий существования абсолютно защищенной от исследования системы. Частным случаем данной теоремы является теорема Шеннона идеальной секретности шифров в криптографии.

Ключевые слова: *информационная безопасность, защита от исследования, шифрование, идеальная секретность, абсолютные шифры, неразличимые информационные системы.*

Styugin M. A.

CONDITIONS FOR CREATING PERFECTLY SECURE SYSTEMS

The present paper reviews a method for establishing secure information systems by complicating the possibility to research them for potential adversaries. A formalized model of a researcher and a definition of a research secure system are presented. A theorem for conditions required for creating a system perfectly secured from research. The Shannon's theorem of absolute security of perfect secrecy ciphers in cryptography is an instance of the theorem presented in the paper.

Keywords: *information security, protection from research, encryption, perfect secrecy, perfect ciphers, indistinguishable information systems.*

* Работа поддержана грантом РФФИ 16-29-09456\16

Введение

В последнее время мы можем наблюдать все возрастающую сложность информационных систем. Увеличивается сложность как аппаратного, так и программного обеспечения. Этот тренд влияет также и на развитие технологий информационной безопасности. Поскольку уязвимости в сложных системах не могут быть исчерпывающе найдены и устранены на этапе проектирования, возникает необходимость скрывать информацию о принципах и алгоритмах работы системы. Технологии такого «сокрытия» основываются на методах защиты информационных систем от исследования. Последние несколько лет появилось огромное множество работ в этом направлении и каждый год их число увеличивается.

Только за 2015 год было найдено более 150 статей, посвященных различным методам сокрытия данных/процессов в области информационных систем и злонамеренного программного обеспечения. Здесь можно выделить и некоторые достаточно специфические проблемы, которые достаточно часто затрагиваются в исследованиях, в том числе презентуемые на ведущих конференциях отрасли (top-tier conferences).

В качестве одного из трендов можно выделить технологии затрудняющие для противника сбор каких-либо данных с использованием методов Data Mining [1] и технологии сокрытия действий пользователей. Наиболее известные исследования с такой постановкой задачи были сделаны Steven Alpern [2] и Shmuel Gal [3].

Также стоит упомянуть технологию непрерывного изменения систем с целью защиты их от исследования (Moving Target Defence - MTD) [4] и технологию защиты программного кода и алгоритмов. В последнее время появилось более 150 различных техник MTD [5] затрагивающих такие направления как защита локальной сети [6], защита от инъекций программного кода [7], защита от XSS-атак [8], защита от DDoS-атак [9] и пр.

Но несмотря на лавинообразный рост публикации в этой области, до сих пор нет точных математических методов анализа и сравнения систем защищенных от исследования. Существующие математические модели рассматривают только отдельные характеристики систем в целях сравнения их эффективности, как например, работы [10, 11]. Однако не существует общей модели, которая могла бы определить идеальные объекты, такие как абсолютно защищенная от исследования система. Например, формулировка абсолютно неразличимого канала передачи информации в работе [12] фактически не дает формального определения его неразличимости.

Для разработки подобной теории, нам необходимо получить условия идеальной секретности, существующие в криптографии применительно к информационным системам в целом и сформулировать теорему секретности Шеннона [13] в общем виде, для которой теорема абсолютной секретности Шеннона была бы частным случаем.

1. Формализация задачи защиты от исследования

Ключевым аспектом данной работы является формализация процесса исследования и информационных ограничений исследователя. Такая формализация позволит нам дать определения защищенной от исследования системы и определить условия ее существования.

За основу возьмем модель приведенную в работе [14] существенно ее модифицировав. В работе [14] рассмотрена модель исследователя по отношению к объекту, имеющему вход, выход и дискретное внутреннее состояние. Данная формализация удобна тем, что под нее подходит любая цифровая информационная система. На вход объекта подается множество параметров $par \in \{par\}$. Множество значений выхода $f \in \{f\}$. Функция связывающая вход и выход объекта - $f(par)$. Параметры входа (то есть то какие именно переменные являются входом для объекта) определяются переменной $<par> \in \{<par>\}$. Исследователь может ошибиться и выбрать неправильное множество входных параметров - $<par>' \neq <par>$. Кроме того, мы можем разделить наблюдаемое и ненаблюдаемое множество входных параметров и значений выхода:

$$\{par\} = \{par\}^{\mathbf{v}} \cup \{par\}^{\mathbf{vn}}, \{par\}^{\mathbf{v}} \cap \{par\}^{\mathbf{vn}} = \varnothing$$

$$\{f\} = \{f\}^{\mathbf{v}} \cup \{f\}^{\mathbf{vn}}, \ \{f\}^{\mathbf{v}} \cap \{f\}^{\mathbf{vn}} = \varnothing$$

В результате модель исследователя, как она представлена в [14] может быть выражена посредством кортежа из трех значений:

$$\{<par>' = <par>, par \in \{par\}^{\mathbf{v}}, f \in \{f\}^{\mathbf{v}}\}$$

Теперь сделаем существенные изменения модели обратив внимание на следующий

факт, всего у нас три характеристики объекта в отношении исследователя (вход, выход и функция). Каждая из характеристик может быть наблюдаема или ненаблюдаема и значение каждой из характеристик можно выбрать правильно или неправильно. Таким образом, расширим множество переменных кортежа:

$\{<par>' = <par>,\ par \in \{par\}^v,\ <f\ (par')>' = f(par), f(par) \in \{f(par)\}^v,\ <f>' = <f>, f \in \{f\}^v\}$.

В этой записи переменные $\{<par>' = <par>,\ <f\ (par')>' = f(par),\ <f>' = <f>\}$ являются предусловием истинности переменных $\{par \in \{par\}^v, f(par) \in \{f(par)\}^v, f \in \{f\}^v\}$. То есть если исследователь неправильно выбрал входные каналы для объекта, множество входных значений для него однозначно будет ненаблюдаемым, то есть

$$<par>' \neq <par> \Rightarrow par \notin \{par\}^v$$

$$<f(par')>' \neq f(par) \Rightarrow f(par) \notin \{f(par)\}^v$$

$$<f>' \neq <f> \Rightarrow f \notin \{f\}^v$$

В дальнейшем будем рассматривать только случаи когда условия $\{<par>' = <par>,\ <f\ (par')>' = f(par),\ <f>' = <f>\}$ выполняются. Для этого оставим в кортеже только три переменных принадлежности к наблюдаемым множествам:

$$\{par \in \{par\}^v, f(par) \in \{f(par)\}^v,\ \ f \in \{f\}^v\}$$

Для дальнейшего построения модели используем существующие в криптографии методы классификации исследователя (различителя – distinguisher [13]).

Каждая переменная может быть наблюдаема однократно, многократно или полностью. Это свойство динамической характеристики канала. Необходимо обратить внимание, что многократное или полное наблюдение переменных невозможно связать с определением «абсолютно защищенной от исследования системы», поскольку объект является детерминированным. Детерминированность значений выхода от значений входа позволяет, например, при повторном получении одного и того же значения, определить повторения входных параметров, что дает исследователю некую информацию. Поэтому, будем рассматривать только однократно наблюдаемые системы.

В качестве упрощения записи мы можем также упростить запись кортежа. Если переменная является наблюдаемой, то ставим 1, иначе 0. В результате идеальная для исследователя система записывается кортежем (1, 1, 1), что характеризует систему с выполненными условиями $\{par \in \{par\}^v,\ f(par) \in \{f(par)\}^v,\ f \in \{f\}^v\}$. Для начала определим, что означает выражение «решить задачу исследования». Решить задачу исследования – это значит привести систему к состоянию (1, 1, 1). Следовательно, защищенная от исследования система это такая, которая не может быть приведена исследователем в состояние (1, 1, 1). Заметим также, что в реальной системе видимость двух любых параметров всегда означает видимость третьего. Например, мы знаем переменную на входе и знаем функцию системы, таким образом, мы можем вычислить значение на выходе.

В результате, защищенная от исследования система может существовать только в состояниях с одной видимой переменной. Такие состояния имею следующую структуру:

$(0, 0, 1)$ - $\{par \notin \{par\}^v, f(par) \notin \{f(par)\}^v, f \in \{f\}^v\}$

$(0, 1, 0)$ - $\{par \notin \{par\}^v, f(par) \in \{f(par)\}^v, f \notin \{f\}^v\}$

$(1, 0, 0)$ - $\{par \in \{par\}^v, f(par) \notin \{f(par)\}^v, f \notin \{f\}^v\}$

Определение 1. *Защищенная от исследования система* – это такая система, в которой знание (видимость) одной из характеристик (значение входа, значение выхода, функция зависимости выхода от входа) не позволяют вычислить две другие характеристики системы.

2. Абсолютно защищенная от исследования система

По каждому из условий мы можем выразить абсолютно защищенной от исследования системы по аналогии с существующим в шифровании принципом Shannon Secrecy [13].

Определение 2.1. *Абсолютно защищенная от исследования система в состоянии (0, 0, 1)* это система, для которой соблюдаются условия (1):

$$\Pr(f = \{f\}) = \Pr(f = \{f\} \mid par = \{par\})$$

$$\Pr(f = \{f\}) = \Pr(f = \{f\} \mid f(par) = \{f(par)\})$$

Определение 2.2. *Абсолютно защищенная от исследования система в состоянии (0, 1, 0)* это система, для которой соблюдаются условия (2):

$$\Pr(f(par) = \{f(par)\}) = \Pr(f(par) = \{f(par)\} \mid par = \{par\})$$

$$\Pr(f(par) = \{f(par)\}) = \Pr(f(par) = \{f(par)\} \mid f = \{f\})$$

Определение 2.3. *Абсолютно защищенная от исследования система в состоянии (1, 0, 0)* это система, для которой соблюдаются условия (3):

$$\Pr(par = \{par\}) = \Pr(par = \{par\} \mid f = \{f\}).$$

$$\Pr(par = \{par\}) = \Pr(par = \{par\} \mid f(par) = \{f(par)\})$$

Далее определим при каких условиях система может находиться в состояниях (1), (2) и (3). Обозначим как $S(1) = 1$ выполнение условий (1) и как $S(1) = 0$ – невыполнение. Аналогично $S(2)$ и $S(3)$.

Теорема 1.1.

$S(1) = 1 \Rightarrow |\{f(par)\}| \geq \max(|\{f\}|, |\{par\}|)$

То есть система может находиться в состоянии (1) тогда, когда множество различных функций между входом и выходом системы не меньше, чем максимум между размерами множеств значений выхода или входа.

Для доказательства рассмотрим два условия.

1. Рассмотрим случай когда $|\{par\}| < |\{f\}|$ и рассмотрим невыполнение первой части условия:

$$|\{f(par)\}| < |\{f\}|$$

Определим множество $\{f_{\underline{par}}\}$ таким образом, чтобы в него попали все значения f от конкретного параметра $\underline{par}$ путем перебора всех возможных функций из $\{f(par)\}$. В результате получаем

$\{f_{\underline{par}}\} = \{f \mid f = f(\underline{par})$ для некоторого $f(par) \in \{f(par)\}\}$

Очевидно, что $|\{f_{\underline{par}}\}| \leq |\{f(par)\}|$. Отсюда следует что существует как минимум один $f' \in \{f\}$, такой что $f' \notin \{f_{\underline{par}}\}$. Получается, что

$\Pr(\{f\} = f' \mid \{par\} = \underline{par}) = 0 \neq \Pr(\{f\} = f')$

Полученное выражение противоречит условию S(1), в котором определено

$$\Pr(f = \{f\}) = \Pr(f = \{f\} \mid par = \{par\})$$

2. Рассмотрим случай когда $|\{par\}| > |\{f\}|$ и рассмотрим невыполнение первой части условия:

$$|\{f(par)\}| < |\{par\}|$$

Определим множество $\{par_{\underline{f}}\}$ таким образом, чтобы в него попали все значения par из которых можно получить результат $\underline{f}$ путем перебора всех возможных функций из $\{f(par)\}$. В результате получаем

$\{par_{\underline{f}}\} = \{\underline{par} \mid f(par) = f$ для некоторого $f(par) \in \{f(par)\}\}$.

Очевидно, что $|\{par_{\underline{f}}\}| \leq |\{f(par)\}|$. Отсюда следует что существует как минимум один $par' \in \{par\}$, такой что $par' \notin \{par_{\underline{f}}\}$. Получается, что

$\Pr(\{par\} = par' \mid \{f\} = \underline{f}) = 0 \neq \Pr(\{par\} = par')$

Полученное выражение противоречит условию $S(1)$, в котором определено

$$\Pr(f = \{f\}) = \Pr(f = \{f\} \mid par = \{par\})$$

что равносильно

$$\Pr(par = \{par\}) = \Pr(par = \{par\} \mid f = \{f\})$$

■

Теорема 1.2

$S(2) = 1 \Rightarrow |\{par\}| \geq |\{f\}|$

Теорема 1.3

$S(3) = 1 \Rightarrow |\{f(par)\}| \geq \max(|\{f\}|, |\{par\}|)$

3. Условия, накладываемые на абсолютно защищенную от исследования систему

Все приведенные условия являются необходимыми, но не достаточными для построения абсолютно защищенной от исследования системы.

В соответствии с принятой в литературе терминологией будем обозначать как «uniform» - абсолютно случайную функцию или последовательность бит в которой каждый последующий бит может быть предсказан не более чем с вероятностью ½ вне зависимости от асимптотической сложности вычислений.

Определим далее понятие uniform-функции.

$$f(par) - \text{uniform} \Leftrightarrow \forall par: \Pr(f) = 1/|\{f\}|$$

То есть функция является uniform если она выбрана таким образом, что для исследователя любое значение функции на выходе является равновероятным. То есть сама функция детерминированная, а не вероятностная. Но при этом она выбрана случайным способом, не дающим какой-либо информации исследователю.

Теорема 2

1. Если $f(par)$ - uniform и $|\{f\}| \geq |\{par\}|$ то $S(1) = 1$ тогда и только тогда когда для любого $par \in \{par\}$ существуют $f_1, \ldots, f_n$, где для каждой f_i, $i = 1 \ldots n$ существуют $f_1^i(par), \ldots, f_m^i(par)$, где $m = |\{f(par)\}|/|\{f\}|$ такие что $f_i = f_k^i(par)$, $i = 1 \ldots n$, $k = 1 \ldots m$.

2. Если $f(par)$ - uniform и $|\{f\}| \leq |\{par\}|$ то $S(1) = 1$ тогда и только тогда когда для любого $f \in \{f\}$ существуют $par_1, \ldots, par_m$, где для каждой par_i, $i = 1 \ldots n$ существуют $f_1^i(par), \ldots, f_m^i(par)$, где $m = |\{f(par)\}|/|\{par\}|$ такие что $f_i = f_k^i(par)$, $i = 1 \ldots n$, $k = 1 \ldots m$.

Первая часть данной теоремы говорит, что для каждого par существует n результатов

функций и для каждого параметра и результата функции существует m функций которые с ними можно сопоставить. Вторая часть данной теоремы говорит что для каждого значения на выходе f существует n различных параметров из которых ее можно получить m различными функциями.

Приведем общую логику построения доказательства. Рассмотрим первую часть теоремы когда $|\{f\}| \geq |\{par\}|$. В этом случае мы должны иметь такое количество функций по преобразованию par в f чтобы иметь возможность получить каждый из f из каждого par. В противном случае мы приходи к факту что условие $S(1)$ не соблюдается. Если же количество функций более чем достаточно чтобы получить из каждого par каждый f, то они также должны быть равномерно распределены по всем парам для того чтобы не сделать вероятность появления одной из пар больше чем других, что также приводит к нарушению условия $S(1)$. Поэтому количество функций $f(par)$ должно быть кратно количеству результатов на выходе – f.

Сделаем доказательство от противного. Допустим, что условие 1 не соблюдается. Следовательно существуют такие par_1, $par_2 \in \{par\}$, для которых существуют $f_1^1 \ldots f_n^1$ и $f_1^2 \ldots f_l^2$, такие что $n \neq l$. Отсюда мы приходим к тому, что существует такой f_u^i, который принадлежит одному множеству, но не принадлежит другому, поэтому $\Pr(f_u^i = \{f\}) \neq \Pr(f_u^i = \{f\} \mid par = \{par\})$. Пришли к невыполнению условия $S(1)$.

Допустим теперь, что выполняется требование «для любого $par \in \{par\}$ существуют $f_1, \ldots, f_n$», но не выполняется требование для каждой f_i, $i = 1 \ldots n$ существуют $f_1^i(par), \ldots, f_m^i(par)$, где $m = |\{f(par)\}|/|\{f\}|$ такие что $f_i = f_k^i(par)$, $i = 1 \ldots n$, $k = 1 \ldots m$. Здесь мы приходим к невыполнению второго условия в $S(2)$, где $\Pr(f = \{f\}) = \Pr(f = \{f\} \mid f(par) = \{f(par)\})$.

Аналогично доказывается и вторая часть теоремы. ■

Теорема 2 представляет из себя обобщенную теорему Шеннона [13]. Если в условиях теоремы мы приравняем мощность множества значений выхода (шифртекстов) и множество значений входа (исходных текстов), то получим что на каждое входное значение и каждое значение выхода у нас есть только одна функция преобразования.

Теперь мы можем удостовериться, например, что определение абсолютно неразличимого канала передачи данных в [12] действительно удовлетворяет условиям абсолютно защищенных от исследования систем.

Выводы

В данной работе нам удалось формализовать исследователя информационной системы с учетом существующих у него ограничений. На основании информационных ограничений исследователя мы можем определить защищенную от исследования систему и абсолютно защищенную от исследования систему. Нам удалось получить обобщенную теорему, формулирующую необходимые и достаточные условия к абсолютно защищенной от исследования системе. Частным случаем данной теоремы является теорема абсолютной секретности Шеннона в криптографии.

Полученная модель подтверждает результат, приведенный в работе [12]. Полученные теоретические результаты могут быть использованы для проектирования неразличимых информационных систем если задачей является затруднение их исследования со стороны внешнего наблюдателя.

Примечания

1. Chengy, P., Lee, I, Pan, J.-S., Lin, C.-W., Roddick, J.F. Hide association rules with fewer side effects // IEICE Transactions on Information and Systems. Volume E98D, Issue 10, 1 October 2015, Pages 1788-1798

2. S. Alpern and S. Gal. Searching for an agent who may or may not want to be found. Operations Research, 50(2):311–323, 2002.

3. S. Gal. On the optimality of a simple strategy for searching graphs. International Journal of Game Theory, 29(4):533–542, 2001.

4. Jajodia, S. Moving Target Defense. Creating Asymmetric Uncertainty for Cyber Threats. Series: Advances in Information Security, 2011 [Текст] / S. Jajodia, A.K.Ghosh, V. Swarup, C. Wang, X.S. Wang. - 184 p.

5. Okhravi H., Hobson T., Bigelow D., Streilein W., Finding Focus in the Blur of Moving-Target Techniques // IEEE Security and Privacy. 2014. vol. 12. pp. 16-26.

6. Carvalho M., Ford R., Moving-target defenses for computer networks // IEEE Security and Privacy. 2014. Vol. 12(2). pp. 73-76.

7. Larsen P., Brunthaler S., Franz M. Automatic Software Diversity // IEEE Security and Privacy. 2015. Vol. 13 (2). pp. 30-37.

8. Portner J., Kerr J., Chu B. Moving target defense against cross-site scripting attacks // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2015. Vol. 8930. pp. 85-91.

9. Ma D., Xu Z., Lin D. Defending blind DDoS attack on SDN based on moving target defense // Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST. 2015. Vol. 152. pp. 463-480.

10. Jin B. Hong, Dong Seong Kim. Scalable Security Models for Assessing Effectiveness of Moving Target Defenses // Dependable Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on, 23-26 June 2014. pp. 515 - 526

11. Jin Hong, Dong-Seong Kim. HARMs: Hierarchical Attack Representation Models for Network Security Analysis // Originally published in the Proceedings of the 10th Australian Information Security Management Conference, Novotel Langley Hotel, Perth, Western Australia, 3rd-5th December, 2012

12. Mikhail Styugin. Absolutely Indiscernible Data Transfer Channel // Proceedings of The 14th European Conference on Cyber Warfare and Security (ECCWS-2015). pp. 286-293

13. Katz J., Lindell Y. Introduction to Modern Cryptography, Second Edition. N.Y.: Chapman and Hall/CRC, 2014. 603 p.

14. Styugin M. Protection against System Research // Cybernetics and Systems: An International Journal. 2014. Vol. 45 (4). pp. 362-372.

Стюгин Михаил Андреевич, старший научный сотрудник Научно-исследовательского управления Сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева, кандидат технических наук. 660037, г. Красноярск, проспект им. газеты «Красноярский рабочий», д. 31, а/я 1075. E-mail: styugin@gmail.com

Styugin Mikhail Andreevich, research fellow of Research Department of Siberian State Aerospace University, PhD in computer science. 660014 Krasnoyarsk, Russia Office A-406, 31, Krasnoyarsky Rabochy Av. E-mail: styugin@gmail.com.