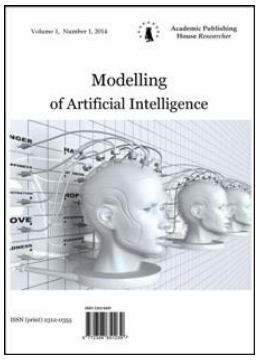

Published in the Russian Federation
Modeling of Artificial Intelligence
Has been issued since 2014.
ISSN: 2312-0355
E-ISSN: 2413-7200
2017, 4(1): 29-38

DOI: 10.13187/mai.2017.1.29
www.ejournal11.com

Search Fuzzy Image of the Attacker Based on the Use of Automatic Classification Methods

Simon Zh. Simavoryan [a, *], Arsen R. Simonyan [a], Elena I. Ulitina [a], Rafik A. Simonyan [b], Ellina A. Pilosyan [a], Nadezhda A. Kornienko [a]

[a] Sochi State University, Russian Federation
[b] Kuban State University, Russian Federation

Abstract

Work is devoted to development of methods of search of an indistinct image of the malefactor on the basis of use of methods of automatic classification. This type of search is used by initial search when the inquiry is set in the form of the description of signs of malicious action. Search is carried out on data from the specialized knowledge base on malicious actions (templates) and the knowledge base on regular situations and values of their admissible characteristics. Methods of automatic classification meet as well other names: objective classification, splitting, taxonomy, diagonalization of matrixes of communication, etc. When using these methods, given about malicious actions, join in the knowledge base as an image (n + 1) member malicious action. After that some algorithm of automatic classification which is carrying out splitting the knowledge base into groups of images of malicious actions "uniform" somewhat is put in action. As result of initial search of an image of the malefactor system the images of malefactors carried by an algorithm of automatic classification in the same group, as an image of the malefactor for whom the image of malicious action is looked for are given. For increase in a noise stability of the procedure of initial search of an indistinct image of the malefactor it is possible to carry out in parallel splitting on various algorithms of automatic classification then as an indistinct image of the malefactor to give association (for increase in completeness of initial search), crossing (for completeness of accuracy), or composition of subsets in which has been carried (n + 1) member an image from the knowledge base.

Keywords: antagonism of malefactors and service of information security, search of an indistinct image of the malefactor, methods of automatic classification.

1.Введение

Анализ немногочисленных публикаций по системам обнаружения злоумышленных действий в сети вычислительных систем показывает, что задача поиска образа злоумышленника на основе использования методов автоматической классификации является слабо структурированной и требует дальнейших исследований и разработок. Так, например, в докладе (Лебедев, 2012) приведена методика формирования

* Corresponding author
E-mail addresses: simsim58@mail.ru (S.Zh. Simavoryan), oppm@mail.ru (A.R. Simonyan), elenaulitina@mail.ru (E.I. Ulitina), raf55@list.ru (R.A. Simonyan), azalto@mail.ru (E.A. Pilosyan), kornienko_nadja@mail.ru (N.A. Kornienko)

автоматизированного поиска злоумышленников и формирования перечня уязвимостей оконечных узлов в вычислительных сетях. В работе (Ажмухамедов, 2012) поиск злоумышленных действий ведется на основе аномалий сетевого трафика. В работе (Беляев, 2009) приведено обоснование актуальности исследований в области обнаружения вторжений в вычислительные системы, область таких исследований названа «социальной инженерией».

Поиск образа злоумышленника является неотъемлемой частью интеллектуальной деятельности службы защиты информации, поэтому разработка методов автоматизированного поиска образа интеллектуального злоумышленника в вычислительных сетях является актуальной и насущной задачей. Весьма перспективным представляется использование при первоначальном поиске алгоритмов размытой автоматической классификации. В статье (Вятченин, 2004) рассматривается расширение методов поиска злоумышленника до уровня методов поиска нечеткого образа злоумышленника в вычислительной сети, основанная на идее первоначального поиска алгоритмов размытой автоматической классификации.

Основными источниками для написания данной статьи стали работы ведущих ученых в области следующих дисциплин: информационная безопасность в вычислительных сетях и методы автоматической классификации.

2. Обсуждение

Основная идея методов автоматической классификации применительно к поиску нечеткого образа злоумышленника заключается в следующем. Пусть имеются следующие специализированные базы данных: 1) специализированная база данных (знаний) о злоумышленных действиях, и 2) база данных (знаний) о штатных ситуациях и значений их допустимых характеристик. Построение баз знаний осуществляется по принципам построения объектно-характеристических таблиц (ОХТ) (Герасименко, 1996). Разработка структуры информационного кадастра может быть осуществлена в следующей последовательности: 1) на основе анализа целей злоумышленника составляется наиболее точный список всех возможных злоумышленных действий по каждому из каналов несанкционированного получения информации; 2) составленный список классифицируется на однородные группы для каждого класса из злоумышленников в отдельности, причем основным критерием однородности должны быть идентичность характеристик злоумышленных действий; 3) для каждой группы злоумышленников составляется перечень характеристик, по которым должны собираться сведения о них; 4) для всех групп устанавливаются их взаимосвязи, на основе чего и строится упорядоченная структура всех ОХТ.

В работах (Симаворян, 2013; Симаворян, 2015; Simavoryan, 2015) приводится система требований к защите информации - общетеоретические требования и прикладные требования. Поэтому крайне необходимо соблюдать эти требования при формировании специализированной базы знаний о злоумышленных действиях и базы знаний о штатных ситуациях и значений их допустимых характеристик. Выполнение общетеоретических требований обеспечивает полноту, т.е. достаточность данных для решения задач поиска нечеткого образа злоумышленника; непротиворечивость и логическую стройность баз знаний. Выполнение прикладных требований обеспечивает унифицированность баз знаний, т.е. способность обеспечивать потребности решения задач поиска; и адекватность не только для настоящих, но и будущих потребностей и условий решения этих задач. При таком представлении информационного кадастра обеспечиваются хорошие условия для автоматизированного поиска нечеткого образа злоумышленных действий. Эти таблицы составляют основу как специализированной базы знаний о злоумышленных действиях, так и базу знаний о штатных ситуациях и значениях их допустимых характеристик. При построении таблиц необходимо также учитывать следующие особенности: особенности интеллектуального противоборства злоумышленников и службы безопасности (Симаворян, 2014; Симаворян, 2015), уровни интеллектуальной защиты информации.

Таблицы строятся по принципу объект-признак. Обычно, в общем случае, подразумевается, что имеется k признаков, признаки принято классифицировать на следующие типы: количественные (численные), качественные (лингвистические) и

ранговые, значения которых $x_j^{(i)}$ ($i = 1,..., k,\ j = 1,..., n$). Так, например, в режиме оперативно-диспетчерского управления защитой информации (Симаворян, 2015) (значения параметров изменяются со временем $x_j^{(i)}(t)$, поэтому необходимо построить такую структуру ОХТ, чтобы для она сжатой, содержательно хорошо интерпретируемой, для описания исследуемых злоумышленных действий с целью идентификации основных характеристик, их выявления и прогнозирования интегральных показателей поведения злоумышленников во времени, поиска закономерностей их взаимодействия, проявления и т.д. При этом выявление структуры необходимо производить по следующим признакам: структура злоумышленных действий в различных подпространствах параметров, структура взаимосвязи параметров по данным имеющих баз знаний, структура динамических характеристик злоумышленных действий.

Особенность алгоритма автоматизированного поиска злоумышленника в вычислительной сети заключается в том, что этот тип поиска используется при первоначальном поиске, когда запрос задается в виде описания значений признаков некоторого объекта. Методы автоматической классификации встречаются также и с другими наименованиями: объективная классификация, разбиение, таксономия, диагонализация матриц связи и т.д. При использовании этих методов, представленных в запросе данные о некотором объекте, включаются в архив в качестве образа (n+1)-го объекта. После этого приводится в действие некоторый алгоритм автоматической классификации, осуществляющий разбиение архива на группы «однородных» в некотором смысле образов объектов. В качестве результата первоначального поиска образа злоумышленника, системой выдаются образы, отнесенные алгоритмом автоматической классификации в ту же группу, что и образ злоумышленника, для которого ищется образ злоумышленного действия. Для повышения помехоустойчивости процедуры первоначального поиска размытого образа злоумышленника можно параллельно осуществить разбиение по различным алгоритмам автоматической классификации, после чего в качестве размытого образа злоумышленника выдавать объединение (для повышения полноты первоначального поиска), пересечение (для полноты точности) или композицию подмножеств, в которые был отнесен (n+1)-ый образ из базы знаний.

Для первоначального поиска размытого образа злоумышленника успешно могут быть использованы также алгоритмы иерархической классификации (Вятченин, 2004). При этом в качестве размытого образа злоумышленника система может выдавать нечеткое подмножество, функция принадлежности которому $\mu(u_i)$ численно будет равна номеру того уровня иерархии, на котором i-ый и $(n+1)$-ый образы впервые попадают в один класс. Некоторые алгоритмы иерархической классификации: «объединение», «ближнего соседа», «средней связи», «дальнего соседа» и «групповых центров» приведены в работе (Вятченин, 2004).

Далее рассмотрим первоначальный поиск по неявным запросам. Этот тип первоначального поиска используется, когда вместо задания запроса в явном виде, то есть в виде описания значений признаков некоторого объекта, пользователь, просматривая часть базы знаний, указывает с его точки зрения, соответствующие запросу образы объектов. Иначе говоря, пользователь делит просмотренную часть базы знаний на две группы: «соответствующих» и «несоответствующих» запросу. Такой вид неявного запроса может оказаться очень полезным, поскольку обладает рядом достоинств: во-первых, разбиение, осуществляемое пользователем путем просмотра части базы знаний, фактически неявно выражает цель, для которой ищется размытый образ злоумышленника. Таким образом отпадает необходимость проведения первоначального разбиения базы знаний в зависимости от цели поиска образа злоумышленника, что требуется при весовом поиске. Во-вторых, для каждого типа злоумышленников, кроме данных о злоумышленнике, отраженных в базе данных, иногда дополнительно существует такая трудно формализуемая и не формализуемая информация (обычно хранимая в текстовом или визуальном виде), которая может быть использована сотрудником службы безопасности при разбиении просматриваемой части базы знаний. В-третьих, поскольку злоумышленные действия непредставимы в виде простой «суммы» элементарных действий, сотрудник службы

защиты информации, для исследования и описания этих объектов пользуется некоторыми целостными, интегральными понятиями, которые он не всегда может однозначно выразить посредством элементарных описаний. В рамках всех рассматриваемых ранее методов первоначального поиска нет возможности адекватного использования этих интегральных понятий. Рассматриваемый механизм поиска, включая в качестве обязательного звена специалиста службы защиты информации, в определенных пределах позволяет учесть интегральные понятия, которыми оперирует злоумышленник, и использовать их при дальнейшем поиске. Поиск по неявным запросам заключается в анализе слабых, малозаметных закономерностей, присущих группам «соответствующих» и «несоответствующих» запросу образов злоумышленников, и в нахождении размытого образа злоумышленника во всей базе знаний с учетом этих слабых и малозаметных закономерностей, с целью дальнейшего исследования. В такой постановке эта задача может встречаться под названием «поиск разделяющего правила» (Вятченин, 2004). Эффективность этого поиска зависит от корректности разделяющего правила, которая выражается степенью точности аппроксимации этим правилом разбиения, осуществленного специалистом по защите информации. Разделяющее правило будем искать во множестве правил вида: если значение некоторой линейной функции $\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i)$ больше некоторого порога a, то этот объект включается в группу «соответствующих», а если меньше - то в группу «несоответствующих». Для каждого разбиения, осуществляемого службой безопасности, будем искать значения λ_j^l: j=1,2,...,k; l=1,2,...,m_i, при которых разделяющее правило максимально точно аппроксимирует это разбиение.

Отметим, что некоторые рассматриваемые далее методы поиска разделяющего правила применимы также для разделяющих правил более сложного вида. Для упрощения изложения, не нарушая общности, будем предлагать, что специалист службы защиты информации, посмотрев первые P образов злоумышленных действий из базы знаний, P_1 образов объектов $U_1,U_2,...,U_{p1}$ отнес в группу «соответствующих», а $P_2=P-P_1$ образов объектов $U_{p1+1},U_{p1+2},...,U_p$ в группу «несоответствующих».

Правила, в точности аппроксимирующего разбиение, необходимо найти значение λ_j^l, при которых система p неравенств совместна.

$$\begin{cases} \sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) > a; & i = 1,2,\dots p_1 \\ \sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) < a; & i = p_1+1, p_1+2, \dots, p \end{cases} \quad (1)$$

Поскольку в общем случае система неравенств (1) можем быть несовместной, то мы предварительно будем рассматривать задачу нахождения максимальной совместной подсистемы системы неравенств (1) с целью дальнейшего поиска разделяющего правила, максимальной аппроксимирующего рассматриваемое разбиение.

Введём в рассмотрение p булевы переменные $x_1, x_2 \dots, x_p$. Тогда в каждой подсистеме системы неравенств (1) можно поставить во взаимно-однозначное соответствие некоторую комбинацию значений переменных $x_1, x_2 \dots, x_p$, такую, что x_i=1, если i-тое неравенство системы (1) включено в подсистему, и x_i=0 - в противном случае.

В качестве области определения функции алгебры логики от p переменных можно рассматривать множество вершин P -мерного единого куба. Упорядочим вершины этого куба следующим образом: будем говорить, что точка $\alpha=(\alpha_1, \alpha_2, \dots, \alpha_p)$ предшествует точке $\beta=(\beta_1, \beta_2..., \beta_p)$, если $\alpha_i \le \beta_i$ при $1 \le i \le p$. Тот факт, что α предшествует β, будем обозначать $\alpha \le \beta$, если

$\alpha \le \beta$ и $\alpha \ne \beta$, то будем писать $\alpha < \beta$.

Рассмотрим следующее определение (Вятченин, 2004), (Любомудров, 2015): функция алгебры логики $\varphi(x_1, x_2 \dots, x_p)$ называется монотонной, если из того, что $\alpha < \beta$ следует, что $\varphi(\alpha_1, \alpha_2 \dots \dots \alpha_p) \le \varphi(\beta_1, \beta_2 \dots \dots \beta_p)$.

Определим некоторую функцию алгебры логики от p булевых переменных

$$\varphi(x_{1,}x_{2,\dots}x_{p,}) = \begin{cases} 0, \text{если соответствующая значениям признаков } x_1, x_2, \dots x_p \\ \text{подсистема неравенств совместна,} \\ 1 - \text{в противном случае} \end{cases}$$

Очевидно, что рассматриваемая функция φ является монотонной, поскольку при добавлении неравенства к некоторой несовместной подсистеме неравенств эта подсистема остаётся несовместной, а при исключении из подсистемы совместных неравенств одного неравенств подсистема остается совместной. Тогда задача нахождения максимальной совместной подсистемы неравенств (1) сводится к задаче поиска максимального верхнего нуля монотонной функции алгебры логики (Любомудров, 2015). Алгоритм решения этой задачи путём ограниченного перебора является очень трудоёмким, так как верхняя оценка числа обращений к процедуре анализа совместимости некоторой подсистемы неравенств равно $C_p^{[\frac{p}{2}]} + 1$. Поэтому целесообразнее использовать огрубленные или эвристические методы, позволяющие находить квазимаксимальные, то есть близкие к максимальной, совместные подсистемы неравенств. Рассмотрим некоторые из этих методов.

Заметим, что систему (1) можно представить в виде

$$\begin{cases} \sum_{j=1}^{k}\sum_{l=1}^{m_j} \lambda_j^l * f_j^l(u_i) - 1 > 0; \qquad i = 1,2, \dots p_1; \\ \sum_{j=1}^{k}\sum_{l=1}^{m_j} \lambda_j^l * f_j^l(u_i) + 1 > 0; \qquad i = p_1 + 1, p_1 + 2, \dots, p. \end{cases} \quad (2)$$

Для упрощения изложения введем обозначения: левую часть i-того неравенства системы (2) обозначим C_i. Тогда задача на хождения максимальной совместимой подсистемы (2) эквивалентна задаче нахождения минимума функции

$$-\sum_{i=1}^{p} sign^*(C_i), \quad (3)$$

где $sign^*(x) = \begin{cases} 1, \text{если } x > 0 \\ 0, \text{ если } x < 0 \\ -1, \text{ если } x = 0 \end{cases}$

Минимизируемую функцию (3) с точностью до операции сдвига заменим ее гладким приближением

$\sum_{i=1}^{p}(1 + e^{\alpha \cdot C_i})^{-1} + 6\sum_{i=1}^{p}(1 + e^{\alpha \cdot C_i})^{-1} \cdot (1 + e^{-\alpha \cdot C_i})^{-1}$, (4)

где α-достаточно большое число.

Покажем, что эта функция аппроксимирует (3)

При x>0 $[1 + e^{\alpha x}]^{-1} \approx 0$, при x=0 $[1 + e^{\alpha x}]^{-1} \approx 1$, при x=0 $[1 + e^{\alpha x}]^{-1} = [1 + e^{-\alpha x}]^{-1} = \frac{1}{2}$. Следовательно, если левая часть одного из неравенств системы больше нуля, то значение минимизуруемого функционала не изменится, если меньше нуля – то значение функционала увеличивается на I, а если равно нулю – то увеличится на 2. Результат локальной минимизации функционала (4) во многом зависит от удачного выбора начальных значений λ_j^l, при которых достигает локального минимума функционал $\sum_{i=1}^{p}(1 + e^{\alpha C_i})^{-1}$.

Так как найденные при минимизации функционала (2) значения λ_j^l характеризируют лишь локальный минимум, что требует дополнительного проведения корректировки порога. Для этого полученные значения λ_j^l подставляются в неравенства системы (2), после чего система неравенств приводится к виду

$$\begin{cases} a < \gamma_i; i = 1,2, \dots p_1, \\ a > \gamma_i; i = p_1 + 1, \ p_1 + 2, \dots p, \end{cases}$$

где γ_i – некоторые значения.

После этого путем перебора (p-1) интервалов между каждым $\gamma_i < max\{\gamma_l\}$ и $\gamma_j = min\{\gamma_l \mid \gamma_l > \gamma_i\}$ находится интервал (γ', γ"), из которого, если выбрать порог α, будет выполнятся максимальное число неравенств. Разделяющее правило, построенное на основе полученных значений λ_j^l и скорректированного порога, уже на этом этапе может быть использовано для первоначального поиска размытого образа злоумышленника.

Однако поскольку для квазимаксимальной совместной подсистемы системы неравенств (2) имеется множество возможных значений λ_j^l , при которых эти неравенства удовлетворяются, то представляет интерес вопрос: какие же значения λ_j^l выбрать. В этом плане нам кажется целесообразным осуществлять выбор, исходя из решения одной из следующих оптимизационных задач. Первая задача состоит в минимизации суммарной степени нарушения неравенств, не вошедших в квазимаксимальную совместную подсистему. Допустим, в квазимаксимальную совместную подсистему после корректировки порога не вошли $q_1 < p_1$ неравенств вида

$$\sum_{j=1}^{k}\sum_{l=1}^{m}\lambda_j^l \cdot f_j^l(u_i) > a$$

и $q_2 < p_2$ неравенства вида

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) < a$$

Тогда, не нарушая общности рассуждения, будем предполагать, что в квазимаксимальную совместную подсистему не вошли неравенства системы (2) с номерами 1,2,…,$q_1, p_1+1, p_2+2, \dots, p_1+q_2$. К левой части неравенств с номерами i = 1, 2, …, q_1 добавим переменные z_i, а к левой части неравенств с номерами i=$p_1+1, p_2+2, \dots, p_1+q_2$ - переменные $(-z_i)$. Тогда рассматриваемую оптимизационную задачу можно записать в виде, легко сводимом к классической задаче линейного программирования: найти $\lambda_j^l, j = 1,2,\dots,k; l = 1,2,\dots,m_j$ минимизирующие

$$\sum_{i=1}^{q_1} z_i + \sum_{j=p_1+1}^{p_1+q_2} z_i$$

при ограничениях

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) + z_i - a = 0; \qquad i = 1,2,\dots,q_1,$$

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) - a > 0; \qquad i = q_1+1, q_2+2, \dots, p_1,$$

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) - z_i - a = 0; \qquad i = p_1+1, p_2+2, \dots, p_1+q_2,$$

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i) - z_i - a < 0; \qquad i = p_1+q_2+1, p_1+q_2+2, \dots, p,$$

$$z_i > 0; \qquad i = 1,2,\dots,q_1-1, q_1, p_1+1, p_1+2, \dots, p_1+q_2-1, p_1+q_2$$

Представляет также интерес несколько иная задача: найти удовлетворяющие квазимаксимальной совместной подсистеме значения λ_j^l, максимизирующие суммарное значение, которое показывает, насколько $\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i)$ больше порога для $i = q_1+1, q_2+2, \dots, p_1$ и насколько $\sum_{j=1}^{k}\sum_{l=1}^{m_j}\lambda_j^l * f_j^l(u_i)$ меньше порога для $i = p_1+q_2+1, p_1+q_2+2, \dots, p$. Иными словами, мы ищем такое разделяющее правило, при котором неравенства, входящие в квазимаксимальную совместную подсистему, будут выполняться по возможности более «строго». Эту задачу такие можно записать в виде, легко сводимом к задаче линейного программирования. Для этого к левой части неравенств системы (2) с номерами $i = q_1+1, q_2+2, \dots, p_1$ добавим

переменные $(-z_i)$, а к левой части неравенств с номерами $i = p_1 + q_2 + 1, p_1 + q_2 + 2, \dots, p$ добавим переменные z_i. Тогда рассматриваемая задача будет формулироваться следующим образом: найти значения λ_j^i максимизирующие

$$\sum_{i=q_1+1}^{p_1} z_i + \sum_{i=p_1+q_2+1}^{p} z_i$$

при ограничениях

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j} \lambda_j^l * f_j^l(u_i) - z_i - a = 0; \qquad i = q_1 + 1, q_2 + 2, \dots, p_1$$

$$\sum_{j=1}^{k}\sum_{l=1}^{m_j} \lambda_j^l * f_j^l(u_i) + z_i - a = 0; \qquad i = p_1 + q_2 + 1, p_1 + q_2 + 2, \dots, p$$

$$z_i > 0;$$

$$i = q_1 + 1, q_2 + 2, \dots, p_1 - 1, p_1, p_1 + q_2 + 1, p_1 + q_2 + 2, \dots p$$

Эту же задачу можно свести к задаче минимизации функционала

$$-\sum_{i=q_1+1}^{p_1} C_i * (1 + e^{-\alpha * C_i})^{-1} + \sum_{i=p_1+q_1+1}^{p} C_i * (1 + e^{\alpha * C_i})^{-1}$$

$$-\beta \cdot \sum_{i=q_1+1}^{p_1}(1 + e^{\alpha C_i})^{-1} + \beta \cdot \sum_{i=p_1+q_2+1}^{p_1} C_i \cdot (1 + e^{-\alpha \cdot C_i})^{-1}, \qquad (4)$$

где α и β – достаточно большие числа, а значение порога в выражении C_i выбирается произвольным образом из предварительно определенного интервала (γ', γ").

Представляет интерес также следующий эвристический метод построения разделяющего правила: для каждой из групп, на которые разбита просмотренная пользователем часть базы знаний, конструируются агрегированные показатели (фактор-проекции или фактор-проекторы) и в качестве λ_j^l берётся разность вкладов f_j^l в агрегированные показатели для групп «соответствующих» и «несоответствующих» образов злоумышленников. После этого следует осуществлять корректировку порога.

3. Результаты

Предложенные механизмы первоначального поиска злоумышленников представляют большой практический интерес. Их практическая реализация позволит последовательно и планомерно на регулярной основе обнаруживать потенциально возможные злоумышленные (как преднамеренные, так и не преднамеренные) действия. Предложенные механизмы позволяют учесть интегральные понятия, которыми оперирует злоумышленник, что облегчает задачу нахождения размытого образа злоумышленника с учетом проявления слабых и малозаметных его закономерностей.

4. Заключение

Разработанные практические особенности первоначального поиска образа злоумышленника позволяет перейти к разработке комплекса программ по построению специализированных баз данных, а именно: 1) специализированной базы данных (знаний) о злоумышленных действиях, и 2) базы данных (знаний) о штатных ситуациях и значений их допустимых характеристик, с целью применения методов автоматической классификации для поиска нечеткого образа злоумышленника (злоумышленных действий).

5. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-01-00527.

Литература

Ажмухамедов, Марьенков, 2012 – *Ажмухамедов И.М., Марьенков А.Н.* Поиск и оценка аномалий сетевого трафика на основе циклического анализа // Инженерный вестник Дона. 2012. Т. 20. № 2. С. 17-26.

Беляев и др., 2009 – *Беляев А.В., Петренко С.А., Обухов А.В.* Современное состояние проблемы обнаружения вторжений // Защита информации. Инсайд. 2009. № 4 (28). С. 21-31.

Вятченин, 2004 – *Вятченин Д.А.* Нечеткие методы автоматической классификации: Монография. Мн.: УП «Технопринт», 2004, 219 с.

Герасименко, 1996 – *Герасименко В.А.* Основы информационной грамоты. М.: Энергоатомиздат, 1996. 320 с.

Лебедев, 2012 – *Лебедев Р.В.* Методика формирования исходных данных для моделирования сетевых атак // Решетневские чтения. 2012. Т. 2. № 16. С. 663-665.

Любомудров, 2015 – *Любомудров А.А.* Подход к определению принадлежности функций алгебры логики к классу линейных функций // Электронная книга в сборнике «Прикладная информатика» № 2 (56) 2015: http://fictionbook.ru/author/a_a_lyubomudrov/podhod_k_opredeleniyu_prinadlejnosti_fun/ режим доступа 03.10.2016.

Симаворян и др., 2013 – *Симаворян С.Ж., Симонян А.Р., Улитина Е.И., Симонян Р.А.* Системный подход к проектированию интеллектуальных систем защиты информации // *Известия Сочинского государственного университета*, 2013, № 4-2(28), с. 128-132.

Симаворян и др., 2014 – *Симаворян С.Ж., Симонян А.Р., Улитина Е.И., Симонян А.Р.* Исследование интеллектуального противоборства злоумышленников и службы защиты информации в АСОД // *Известия Сочинского государственного университета*, 2014, № 4-1 (32). С. 15-23.

Симаворян и др., 2015 – *Симаворян С.Ж., Симонян А.Р., Улитина Е.И., Симонян А.Р.* К вопросу о разработке методологии проектирования интеллектуальных систем защиты информации // Материалы Международной научно-практической конференции «Актуальные задачи математического моделирования и информационных технологий», г. Сочи, 15-24 мая 2015 г./ Соч. гос. ун-т; Сочи, 2015. С. 123-125.

Simavoryan et al., 2014 – *Simavoryan S.Sy., Simonyan A.R., Ulitina E.I., Simonyan R.A.* About one Approach to a Question of Classification of Intellectual System of Information Security // *Modeling of Artificial Intelligence*, 2014, Vol (1), Is 1, p. 29-44.

Simavoryan et al., 2015a – *Simon Zh. Simavoryan, Arsen R. Simonyan, Elena I. Ulitina, Rafik A. Simonyan*. Research of the Intellectual Antagonism of Malefactors and Service of Information Security in the ADPS // *Modeling of Artificial Intelligence*, 2015, Vol (5), Is. 1, p. 33-41.

Simavoryan et al., 2015b – *Simon Zh. Simavoryan, Arsen R. Simonyan, Elena I. Ulitina, Rafik A. Simonyan*. Projecting Intelligent Systems to Protect Information in Automated Data Processing Systems (Functional Approach) // *Modeling of Artificial Intelligence*, 2015, Vol (7), Is.3, pp. 212-220.

References

Azhmukhamedov, Mar'enkov, 2012 – *Azhmukhamedov I.M., Mar'enkov A.N.* (2012). Poisk i otsenka anomalii setevogo trafika na osnove tsiklicheskogo analiza [Search and estimation of anomalies of network traffic on the basis of cyclic analysis]. Inzhenernyi vestnik Dona. T. 20. № 2. S. 17-26.

Belyaev i dr., 2009 – *Belyaev A.V., Petrenko S.A., Obukhov A.V.* (2009). Sovremennoe sostoyanie problemy obnaruzheniya vtorzhenii [Current state of the problem of intrusion detection]. *Zashchita informatsii. Insaid*. № 4 (28). S. 21-31.

Gerasimenko, 1996 – *Gerasimenko V.A.* (1996). Osnovy informatsionnoi gramoty [The fundamentals of Information literacy]. M.: Energoatomizdat, 320 s.

Lebedev, 2012 – *Lebedev R.V.* (2012). Metodika formirovaniya iskhodnykh dannykh dlya modelirovaniya setevykh atak. Reshetnevskie chteniya [The technique of the formation of initial data for simulation of network attacks] T. 2. № 16. S. 663-665.

Lyubomudrov, 2015 – *Lyubomudrov A.A.* (2015). Podkhod k opredeleniyu prinadlezhnosti funktsii algebry logiki k klassu lineinykh funktsii [The approach to the definition of the membership of the functions of the algebra of logic to the class of linear functions]. Elektronnaya kniga v sbornike «Prikladnaya informatika» № 2 (56): http://fictionbook.ru/author/a_a_lyubomudrov/podhod_k_ opredeleniyu_prinadlejnosti_fun/ rezhim dostupa 03.10.2016.

Simavoryan i dr., 2013 – *Simavoryan S.Zh., Simonyan A.R., Ulitina E.I., Simonyan R.A.* (2013). Sistemnyi podkhod k proektirovaniyu intellektual'nykh sistem zashchity informatsii [The systematic approach to the design of intelligent systems of information protection]. *Izvestiya Sochinskogo gosudarstvennogo universiteta*, № 4-2(28), s. 128-132.

Simavoryan i dr., 2014 – *Simavoryan S.Zh., Simonyan A.R., Ulitina E.I., Simonyan A.R.* (2014). Issledovanie intellektual'nogo protivoborstva zloumyshlennikov i sluzhby zashchity informatsii v ASOD [The investigation of intellectual confrontation of malefactors and information security services in ASAD]. *Izvestiya Sochinskogo gosudarstvennogo universiteta*, № 4-1 (32). C. 15-23.

Simavoryan i dr., 2015 – *Simavoryan S.Zh., Simonyan A.R., Ulitina E.I., Simonyan A.R.* (2015). K voprosu o razrabotke metodologii proektirovaniya intellektual'nykh sistem zashchity informatsii [To the question of developing a methodology for the design of intelligent information security systems]. Materialy Mezhdunarodnoi nauchno-prakticheskoi konferentsii «Aktual'nye zadachi matematicheskogo modelirovaniya i informatsionnykh tekhnologii», g. Sochi, 15-24 maya 2015 g./ Soch. gos. un-t; Sochi, S. 123-125.

Vyatchenin, 2004 – *Vyatchenin D.A.* (2004). Nechetkie metody avtomaticheskoi klassifikatsii [Fuzzy methods of automatic classification]: Monografiya. Mn.: UP «Tekhnoprint», 219 s.

Simavoryan et al., 2014 – *Simavoryan S.Sy., Simonyan A.R., Ulitina E.I., Simonyan R.A.* (2014). About one Approach to a Question of Classification of Intellectual System of Information Security. *Modeling of Artificial Intelligence*, Vol (1), Is 1, p. 29-44.

Simavoryan et al., 2015a – *Simon Zh. Simavoryan, Arsen R. Simonyan, Elena I. Ulitina, Rafik A. Simonyan* (2015). Research of the Intellectual Antagonism of Malefactors and Service of Information Security in the ADPS. *Modeling of Artificial Intelligence*, Vol (5), Is. 1, p. 33-41.

Simavoryan et al., 2015b – *Simon Zh. Simavoryan, Arsen R. Simonyan, Elena I. Ulitina, Rafik A. Simonyan* (2015). Projecting Intelligent Systems to Protect Information in Automated Data Processing Systems (Functional Approach). *Modeling of Artificial Intelligence*, Vol (7), Is.3, pp. 212-220.

Поиск нечеткого образа злоумышленника на основе использования методов автоматической классификации

Симон Жоржевич Симаворян [a, *], Арсен Рафикович Симонян [a], Елена Ивановна Улитина [a], Рафик Арсенович Симонян [b], Эллина Анатольевна Пилосян [a], Надежда Андреевна Корниенко [a]

[a] Сочинский государственный университет, Российская Федерация,
[b] Кубанский государственный университет, Российская Федерация

Аннотация. Работа посвящена разработке методов поиска нечеткого образа злоумышленника на основе использования методов автоматической классификации. Этот тип поиска используется при первоначальном поиске, когда запрос задается в виде описания признаков злоумышленного действия. Поиск осуществляется на данных из специализированной базы знаний о злоумышленных действиях (шаблонах) и базы знаний о штатных ситуациях и значений их допустимых характеристик. Методы автоматической классификации встречаются также и с другими наименованиями: объективная классификация, разбиение, таксономия, диагонализация матриц связи и т.д. При использовании этих методов, данные о злоумышленных действиях, включаются в базу знаний в качестве образа (n+1) –го злоумышленного действия. После этого приводится в действие некоторый алгоритм автоматической классификации, осуществляющий разбиение

[*] Корреспондирующий автор
Адреса электронной почты: simsim58@mail.ru (С.Ж. Симаворян), oppm@mail.ru (А.Р. Симонян), elenaulitina@mail.ru (Е.И. Улитина), raf55@list.ru (Р.А.Симонян), azalto@mail.ru (Е.А.Пилосян), kornienko_nadja@mail.ru (Н.А. Корниенко)

базы знаний на группы «однородных» в некотором смысле образов злоумышленных действий. В качестве результата первоначального поиска образа злоумышленника системой выдаются образы злоумышленников, отнесенные алгоритмом автоматической классификации в ту же группу, что и образ злоумышленника, для которого ищется образ злоумышленного действия. Для повышения помехоустойчивости процедуры первоначального поиска размытого образа злоумышленника можно параллельно осуществить разбиение по различным алгоритмам автоматической классификации, после чего в качестве размытого образа злоумышленника выдавать объединение (для повышения полноты первоначального поиска), пересечение (для полноты точности) или композицию подмножеств, в которые был отнесен (n+1) –ый образ из базы знаний.

Ключевые слова: противоборство злоумышленников и службы защиты информации, поиск нечеткого образа злоумышленника, методы автоматической классификации.