

Model decision rules to detect anomalies in Information Systems

B. Akhmetov, A. Korchenko, N. Zhumangalieva, J. Kultan

Abstract:

The disadvantage of modern intrusion detection systems, built on the principle of identifying the abnormal condition is that they are mainly focused on the use of mathematical models that require a lot of time to prepare statistics. Mathematical models based on expert approaches in this regard are more effective, but for the performance of its functions require the use of appropriate decision rules. For solving this problem, we propose a model of decision rules on fuzzy logic, which through the use of a plurality of pairs of "invasion: the value" and "Invasion: the set of conjugate pairs", as well as models of reference values allows you to display an abnormal condition, generates a certain type of cyber attack in computer network. Based on this model there have been developed examples of rules to detect such intrusions as scanning, spoofing and Dos-attacks that can practically be used to improve real systems intrusion detection mechanism is used to identify anomalies generated by the actions of attacking computer systems.

Key words:

Cyber attack, intrusion detection systems, network traffic anomaly, anomaly detection in computer systems, the set of conjugate pairs, decision rules, expert evaluation

ACM Computing Classification System:

Combinatorial algorithms, Algebraic algorithms, Nonalgebraic algorithms, Symbolic calculus algorithms, Exact arithmetic algorithms, Hybrid symbolic-numeric methods.

Введение

Стремительное развитие информационных технологий (ИТ) в свою очередь породило большое количество угроз ресурсам информационных систем (РИС). Одним из решений обеспечения безопасности РИС, являются системы обнаружения вторжений (СОВ) представляющие собой программные или аппаратные средства, ориентированные прежде всего на выявление фактов неавторизованного доступа. Следует отметить, что современные СОВ основываются на сигнатурном (шаблонном) и аномальном принципах.

Первый базируется на представлении каждого вторжения в виде определенного шаблона (модели, сценария, правила, сигнатуры) отражающего характеристики и сценарии возможных несанкционированных воздействий. Поэтому такие системы с достаточно высокой точностью выявляют тип кибератак и практически функционируют без ложных срабатываний. Анализ сетевого трафика с использованием сигнатурного принципа характерен тем, что распознавание возможно только при известных кибератаках, а для этого необходимо постоянно обновлять и расширять наборы шаблонов. Кроме неустойчивости к новейшим типам вторжений, такие системы сильно зависят от скорости разработки и обновления сигнатур. Также известно, что например, для таких вторжений как сложные распределенные атаки проверка известных шаблонов является достаточно сложной задачей.

Второй принцип основан на выявлении аномального состояния системы порожденного кибератакой и ориентирован на контроль активности в среде окружения, например, наблюдение за значениями величин сетевого трафика. Преимущества систем, реализирующих этот принцип, в первую очередь связано с тем, что они могут обнаруживать не только новые виды кибератак, но и те, которые характеризуются большой продолжительностью во времени.

1. Описание отдельных методов

Существующие СОВ аномального принципа в основном ориентированы на использование таких математических моделей, которые требуют много времени на подготовку статистических данных, что не требуют более эффективные в этом отношении экспертные подходы, преимущества которых показаны в (Корченко 2006). В связи с этим актуальной задачей при разработке СОВ является создание моделей обнаружения аномалий на основе экспертных оценок. В работе (Ахметов, Корченко, Жумангалеева 2016) предложена модель базовых величин (МБВ), которая за счет множества пар «вторжение : величины» и «вторжение : множество сопряженных пар» позволяют отображать аномальное состояние, порождаемое определенным типом вторжения в компьютерной сети. Также, известна модель эталонных величин (МЭВ) (Ахметов, Корченко, Жумангалеева 2015), которая за счет данных экспертных оценок и МБВ позволяет формировать множества эталонных величин характерных для определенного типа вторжения.

Применение этих моделей при построении СОВ, базирующихся на втором принципе, связано с необходимостью формирования правил, направленных на выявления аномального состояния порожденного атакующими действиями. В связи с этим, целью данной работы является разработка математической модели

используемой при формировании соответствующих решающих правил для идентификации аномального состояния в среде окружения. Под средой окружения будем подразумевать совокупность значений сформированных переменных (например, время обработки запроса, загруженность процессора, количество обращений к ресурсу, число подключений и др.), которые можно использовать для оценивания протекающих процессов в информационной системе (ИС) с целью выявления ее аномального состояния. Отображением среды окружения в данном случае могут быть величины входящие в множество ***V*** (Ахметов, Корченко, Жумангалеева 2016). Для решения поставленной задачи необходимо построить решающие правила, представляющие собой некоторые утверждения, основанные на результате обобщения определенных теоретических и экспериментальных знаний (данных) и отражающие интуитивное суждение лица, принимающего решение, для обеспечения поиска рационального смыслового решения слабо формализованных задач.

2. Математические определения

Построение решающих правил можно осуществить с помощью соответствующей модели, для создания которой введем множество нечетких идентификаторов (fuzzy identifiers)

$$\boldsymbol{FI}=\bigcup_{i=1}^{d} FI_i = \{FI_1, FI_2, FI_3, \ldots, FI_d\}, \ (i=\overline{1,d}), \tag{1}$$

где d – количество элементов множества, необходимое для отображения аномального состояния, а $FI_i \ (i=\overline{1,d})$ – элементы ***FI***, каждый из которых принимает одно из текстовых значений, характеризующих в нечеткой форме уровень аномального состояния системы, которое может быть порождено определенными вторжениями. Например, при $d=5$ выражение (1) можно определить как:

$$\boldsymbol{FI}=\bigcup_{i=1}^{5} FI_i = \{FI_1, FI_2, FI_3, FI_4, FI_5\}=\{L, LTH, HTTL, H, LIM\}, \tag{2}$$

где $FI_1=L$, $FI_2=LTH$, $FI_3=HTTL$, $FI_4=H$ и $FI_5=LIM$ соответственно отображаются текстовыми значениями

- «Low (L)» – «Низкий»,
- «Lower than high (*LTH*)» – «Больше низкий чем высокий»,
- «Higher than the lowest (*HTTL*)» – «Больше высокий чем низкий»,
- «High (*H*)» – «Высокий»,
- «Limits (*LIM*)» – «Предельный».

Далее на основе множеств нечетких идентификаторов ***FI*** и сопряженных пар ***MP*** [2] построим множество решающих правил (solution rule)

$$\boldsymbol{SR}=\{\bigcup_{i=1}^{n} \boldsymbol{SR_i}\}=\{\boldsymbol{SR_1}, \boldsymbol{SR_2}, \boldsymbol{SR_3}, \ldots, \boldsymbol{SR_n}\}, \ (i=\overline{1,n}), \tag{3}$$

где $\boldsymbol{SR_i}$ – подмножество возможных правил для выявления i-го аномального состояния, порожденного i-м вторжением, при этом

$$\bigcup_{i=1}^{n} \boldsymbol{SR_i} = \bigcup_{i=1}^{n} \{ \bigcup_{j=1}^{r_i} SR_{ij} \} = \{SR_{11}, SR_{12}, SR_{13}, \ldots, SR_{1r_1}\},$$
$$\{SR_{21}, SR_{22}, SR_{23}, \ldots, SR_{2r_2}\}, \{SR_{31}, SR_{32}, SR_{33}, \ldots, SR_{3r_3}\}, \ldots,$$
$$\{SR_{n1}, SR_{n2}, SR_{n3}, \ldots, SR_{nr_i}\}, (i = \overline{1,n},\ j = \overline{1,r_i}), \quad (4)$$

где SR_{ij} – j-е правило i-го подмножества возможных правил, а r_i $(i = \overline{1,n})$ – общее количество возможных правил, направленных на обнаружение i-й аномалии.

Отметим, что каждому SR_{ij} соответствует решающее выражение (правило) т. е.:

$$\begin{array}{llll}
\{SR_{11} = (MP_{11} \in FI_{11}), & SR_{12} = (MP_{12} \in FI_{12}), & SR_{13} = (MP_{13} \in FI_{13}), & SR_{1r_1} = (MP_{1r_1} \in FI_{1r_1})\}, \\
\{SR_{21} = (MP_{21} \in FI_{21}), & SR_{22} = (MP_{22} \in FI_{22}), & SR_{23} = (MP_{23} \in FI_{23}), & SR_{2r_2} = (MP_{2r_2} \in FI_{2r_2})\}, \\
\{SR_{31} = (MP_{31} \in FI_{31}), & SR_{32} = (MP_{32} \in FI_{32}), & SR_{33} = (MP_{33} \in FI_{33}), & SR_{3r_3} = (MP_{3r_3} \in FI_{3r_3})\}, \\
\ldots, & & & \\
\{SR_{n1} = (MP_{n1} \in FI_{n1}), & SR_{n2} = (MP_{n2} \in FI_{n2}), & SR_{n3} = (MP_{n3} \in FI_{n3}), & SR_{nr_n} = (MP_{nr_n} \in FI_{nr_n})\}.
\end{array} \quad (5)$$

Обобщая выражение (5) с учетом (3) и (4) получим

$$\boldsymbol{SR} = \bigcup_{i=1}^{n} \{ \bigcup_{j=1}^{r_i} SR_{ir_j} \} = \bigcup_{i=1}^{n} \{ \bigcup_{j=1}^{r_i} (MP_{ir_j} \in FI_{ir_j}) \} =$$
$$\{ \bigcup_{i=1}^{n} \{ \bigcup_{j=1}^{r_i} SR_{ir_j} = (MP_{ir_j} \in FI_{ir_j}) \} \}, (i = \overline{1,n},\ j = \overline{1,r_i}), \quad (6)$$

где SR_{ir_j} есть r_j-е правило выявления аномалии порожденной i-м вторжением, которое буквально интерпретируется как: «Если MP_{ir_j} истинно, то уровень аномального состояния, который может быть порожден i-м вторжением, будет FI_{ir_j} ».

Построение правил обычно осуществляется на основе экспертного подхода, особенно это важно в тех случаях, когда необходимо дать предпочтение одной из альтернатив, например, при каком MP_{ir_j} (6) исход, связанный с FI_{ir_j} будет наиболее объективно отображать состояние системы. Рассмотрим процесс формирования предпочтения для набора альтернатив на конкретном примере.

Пусть для построения подмножества правил $\boldsymbol{SR_1}$ используется r_1 сопряженных пар и d (1) нечетких идентификаторов, один из которых наиболее объективно может отразить состояние среды окружения относительно наличия

аномалии. Итак, общее количество возможных альтернативных решений – $d \times r_1$, т. е. на составление каждого правила SR_{1j} $(j = \overline{1,r_1})$ необходимо рассмотреть d альтернативных вариантов правил, для выбора одного из которых воспользуемся методами определения коэффициентов важности (КВ) [4]. Воспользуемся методом ранговых преобразований (РП), поскольку он позволяет воспользоваться услугами нескольких экспертов, в качестве входных данных применяются табличные формы, выходная функция линейная, а трудоемкость низкая (Ахметов, Корченко, Ахметова, Жумангалеева 2014).

Далее, в качестве примера, определим $d=r_1=5$, тогда

$$\boldsymbol{MP_1}=\{\bigcup_{j=1}^{r_1} MP_{1j}\}=\{MP_{11}, MP_{12}, MP_{13}, MP_{14}, MP_{15}\}=$$

$$\{((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^e),\ ((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge$$

$$\underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^e),$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^e),$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^e),$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^e)\}, \qquad (6a)$$

а в качестве значений FI_{1j} $(j = \overline{1,5})$ воспользуемся данными из формулы (2). Таким образом, для каждого MP_{1j} $(j = \overline{1,5})$ возможны $d=5$ исходов выявления аномалий, связанных с конкретными значениями нечетких идентификаторов в (2). Наиболее объективный из исходов определим с помощью метода средних рангов (СР) (Ахметов, Корченко, Ахметова, Жумангалеева 2014).

Таблица 1. Ранги SR_{1j}^k и КВ

SR_{1j}^k	j	k	Эксперты				x_{1j}^k	λ_{1j}^k
			1	2	3	4		
SR_{11}^1	1	1	1	3	1	2	1,75	0,18
SR_{11}^2		2	2	1	3	2	2	0,2
SR_{11}^3		3	3	2	2	2	2,25	0,23
SR_{11}^4		4	2	4	3	3	3	0,3
SR_{11}^5		5	4	4	3	4	3,75	0,38
SR_{12}^1	2	1	2	3	1	2	2	0,2
SR_{12}^2		2	1	2	1	2	1,5	0,15

SR_{12}^{3}		3	3	1	2	3	2,25	0,23
SR_{12}^{4}		4	3	4	2	2	2,75	0,28
SR_{12}^{5}		5	3	2	3	4	3	0,3
SR_{13}^{1}	3	1	2	3	2	4	2,75	0,28
SR_{13}^{2}		2	3	2	2	1	2	0,2
SR_{13}^{3}		3	2	3	1	1	1,75	0,18
SR_{13}^{4}		4	3	4	3	4	3,5	0,35
SR_{13}^{5}		5	4	3	2	4	3,25	0,33
SR_{14}^{1}	4	1	4	2	2	4	3	0,3
SR_{14}^{2}		2	2	4	3	2	2,75	0,28
SR_{14}^{3}		3	3	1	2	2	2	0,2
SR_{14}^{4}		4	1	2	3	1	1,75	0,18
SR_{14}^{5}		5	2	4	4	3	3,25	0,33
SR_{15}^{1}	5	1	4	4	3	3	3,5	0,35
SR_{15}^{2}		2	2	4	4	3	3,25	0,33
SR_{15}^{3}		3	2	4	3	3	3	0,3
SR_{15}^{4}		4	4	3	2	3	3	0,3
SR_{15}^{5}		5	2	2	4	3	2,75	0,28

Согласно этого метода, в качестве примера, воспользуемся суждениями 4-х экспертов относительно *d=5* возможных исходов SR_{1j}^{k} $(k=\overline{1,d},\ j=\overline{1,r_1})$ по каждому *j*-му правилу. Например, для первого правила подмножество альтернативных решений будет

$$\bigcup_{k=1}^{d} SR_{11}^{k} = \{ SR_{11}^{1},\ SR_{11}^{2},\ SR_{11}^{3},\ SR_{11}^{4},\ SR_{11}^{5} \} =$$

$$\{((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^{e}) \in L,$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^{e}) \in LTH,$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^{e}) \in HTTL, \qquad (6\ b)$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^{e}) \in H,$$

$$((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^{e}) \in LIM\ \}.$$

Далее на основе РП определим КВ, которые отражаются величиной λ. Его минимальное значение свидетельствует о большей предпочтительности альтернативы, т. е. ее КВ более высокий. Для правила SR_{11} произведем расчеты значений x_{1j}^{k} и λ_{1j}^{k} по каждому из возможных исходов SR_{11}^{k} $(k = \overline{1,5})$: x_{11}^{1}=(1+3+1+2)/4=1,75; x_{11}^{2}=(2+1+3+2)/4=2; x_{11}^{3}= (3+2+2+2)/4=2,25; x_{11}^{4}=(2+4+3+3)/4=3; x_{11}^{5}=(4+4+ 3+4)/4=3,75. Значение КВ определяется как $\lambda_{1j}^{k} = x_{1j}^{k} / N$, где N – сумма всех рангов ($N=10$). По результатам, занесенным в табл. 1 видно, что лучший исход имеет, SR_{11}^{1} поскольку $\bigwedge_{k=1}^{5} \lambda_{11}^{k} = \lambda_{11}^{1} = 0,18$.

Аналогично произведем расчеты для

SR_{1j}^{k} $(j = \overline{2,5})$: SR_{12}^{k} – x_{12}^{1} =(2+3+1+2)/4=2; x_{12}^{2} = (1+2+1+2)/4=1,5;

x_{12}^{3} =(3+1+2+3)/4=2,25; x_{12}^{4} =(3+ 4+2+2)/4=2,75; x_{12}^{5} =(3+2+3+4)/4=3; SR_{13}^{k} –

x_{13}^{1} = (2+3+2+4)/4=2,75; x_{13}^{2} =(3+2+2+1)/4=2; x_{13}^{3} =(2+3 +1+1)/4=1,75;

x_{13}^{4} =(3+4+3+4)/4=3,5; x_{13}^{5} =(4+3+2 +4)/4=3,25; SR_{14}^{k} – x_{14}^{1} =(4+2+2+4)/4=3;

x_{14}^{2} =(2+4 +3+2)/4=2,75; x_{14}^{3} =(3+1+2+2)/4=2; x_{14}^{4} =(1+2+3+ 1)/4=1,75;

x_{14}^{5} =(2+4+4+3)/4=3,25; SR_{15}^{k} – x_{15}^{1} =(4+ 4+3+3)/4=3,5; x_{15}^{2} =(2+4+4+3)/4=3,25;

x_{15}^{3} =(2+4+3 +3)/4=3; x_{15}^{4} =(4+3+2+3)/4=3; x_{15}^{5} =(2+2+4+3)/4= 2,75. (6 с)

По результатам вычислений (см. табл. 1) видно, что лучший исход для правил SR_{12}, SR_{13}, SR_{14}, SR_{15} имеют соответственно альтернативные варианты SR_{12}^{2}, SR_{13}^{3}, SR_{14}^{4}, SR_{15}^{5}.

Полученные данные можно использовать в качестве конкретных значений при построении реальных правил в практических СОВ. С этой целью, с учетом (6), осуществим структурирование необходимых данных путем ввода матриц инициализации (MI) для множеств ***FI*** и ***MP***, которые обозначим соответственно $FI(n, r_n)$ и $MP(n, r_n)$, т. е.

$$FI(n, r_n) = \begin{Vmatrix} FI(1,1), & FI(1,2), & FI(1,3), & \ldots, & FI(1,r_n) \\ FI(2,1), & FI(2,2), & FI(2,3), & \ldots, & FI(2,r_n) \\ FI(3,1), & FI(3,2), & FI(3,3), & \ldots, & FI(3,r_n) \\ & & \ldots & & \\ FI(n,1), & FI(n,2), & FI(n,3), & \ldots, & FI(n,r_n) \end{Vmatrix} u$$

$$MP(n, r_n) = \begin{Vmatrix} MP(1,1), & MP(1,2), & MP(1,3), & \ldots, & MP(1,r_n) \\ MP(2,1), & MP(2,2), & MP(2,3), & \ldots, & MP(2,r_n) \\ MP(3,1), & MP(3,2), & MP(3,3), & \ldots, & MP(3,r_n) \\ & & \ldots & & \\ MP(n,1), & MP(n,2), & MP(n,3), & \ldots, & MP(n,r_n) \end{Vmatrix}. \quad (7)$$

Например, при n=3 и r_n=5 на основе экспертных оценок [4] были определены следующие MI *FI(3, 5)* и *MP(3, 5)*, т. е.

$$FI(3,5)=\begin{matrix} LOW & LTH & HTTL & H & LIM \\ LOW & LOW & HTTL & H & LIM \\ LOW & LTH & HTTL & H & H \end{matrix} \; И$$

$$(\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee (\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee (\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee (\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee (\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e} \vee$$

$$\underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^{e}) \wedge$$

$$\underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^{e}$$

$$MP(3,5)= \underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge \underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge \underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge \underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge \underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^{e} \wedge$$

$$\underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^{e} \quad \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^{e}$$

$$\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge$$

$$\underset{\sim}{t}_{NVC} \cong \underset{\sim}{VS}^{e} \quad \underset{\sim}{t}_{NVC} \cong \underset{\sim}{S}^{e} \quad \underset{\sim}{t}_{NVC} \cong \underset{\sim}{A}^{e} \quad \underset{\sim}{t}_{NVC} \cong \underset{\sim}{B}^{e} \quad \underset{\sim}{t}_{NVC} \cong \underset{\sim}{VB}^{e})$$

(8)

где $\underset{\sim}{t}_{NCC}$, $\underset{\sim}{t}_{SPR}$, $\underset{\sim}{t}_{DBR}$, $\underset{\sim}{t}_{NPSA}$, $\underset{\sim}{t}_{VCA}$, $\underset{\sim}{t}_{NVC}$ – текущие значения величин «Number of concurrent connections to the server (*NCC*)» – «Количество одновременных подключений к серверу», «Speed of processing requests from the clients (*SPR*)» – «Скорость обработки запросов от клиентов», «The delay between requests from the single user (*DBR*)» – «Задержка между запросами от одного пользователя», «Number of packages with the same sender and receiver address (*NPSA*)» – «Количество пакетов с одинаковым адресом отправителя и получателя», «Virtual Channel Age (*VCA*)» – «Возраст виртуального канала», «Numbers of Virtual channels (*NVC*)» – «Количество виртуальных каналов» и являются идентификаторами величин [2] в среде окружения. Используемый в (8) знак «≅» – интерпретируется как «Нечеткое равно» и указывает на то, что текущее значение величины (например, $\underset{\sim}{t}_{SPR}$) находящегося слева от «≅» наиболее близко к одному из элементов (например, $\underset{\sim}{L}^{e}$) из заданного множества (например, T^{e}_{SPR}={ $\underset{\sim}{L}^{e}$, $\underset{\sim}{A}^{e}$, $\underset{\sim}{H}^{e}$ }), который указывается справа от «≅», т. е. запись $\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^{e}$ можно интерпретировать как: « $\underset{\sim}{t}_{SPR}$ наиболее близко расположен к $\underset{\sim}{L}^{e}$ входящего в T^{e}_{SPR}».

Далее с учетом MI (при $i=1,\ j=\overline{1,5}$) для $FI(n, r_n)$ и $MP(n, r_n)$ на основе (7) и (8) построим подмножество правил $\boldsymbol{SR_1}$ для выявления аномального состояния, которое может быть порождено таким вторжением, как Dos (DDos) атака.

$$\boldsymbol{SR_1} = \left\{ \begin{array}{l} SR_{11}=((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^e) \in L, \\ SR_{12}=((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^e) \in LTH, \\ SR_{13}=((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^e) \in HTTL, \\ SR_{14}=((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^e) \in H, \\ SR_{15}=((\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e \vee \underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e) \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^e) \in LIM. \end{array} \right\} \quad (9)$$

Отметим, что правило SR_{15} в (9) буквально можно интерпретировать как: «Если $\underset{\sim}{t}_{SPR} \cong \underset{\sim}{L}^e$ или $\underset{\sim}{t}_{DBR} \cong \underset{\sim}{S}^e$ и при этом $\underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^e$, то уровень аномального состояния, который может быть порожден Dos-атакой, будет LIM (Предельный)».

Из подмножества правил (9) видно, что для каждой сопряженной пары из $SR_{1j}\ (j=\overline{1,5})$ определены конкретные значения из $\boldsymbol{FI}$ согласно расчетов КВ с помощью метода РП. Используя эти данные по аналогии можно составить правила для выявления аномалий порожденных спуфингом и сканированием [2, 3]. Так с учетом (7) и (8) при $i=\overline{2,3}$ и $j=\overline{1,5}$ наборы правил $\boldsymbol{SR_2}$ (10) и $\boldsymbol{SR_3}$ (11) будут иметь следующий вид:

$$\begin{array}{l} \boldsymbol{SR_2}=\{ SR_{21}=(\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^e \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VS}^e) \in L, \\ SR_{22}=(\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^e \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{S}^e) \in LTH, \\ SR_{23}=(\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^e \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{A}^e) \in HTTL, \\ SR_{24}=(\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^e \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{B}^e) \in H, \\ SR_{25}=(\underset{\sim}{t}_{NPSA} \cong \underset{\sim}{B}^e \wedge \underset{\sim}{t}_{NCC} \cong \underset{\sim}{VB}^e) \in LIM \} \text{ и} \end{array} \quad (10)$$

$$\begin{array}{l} \boldsymbol{SR_3}=\{ SR_{31}=(\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^e \wedge \underset{\sim}{t}_{NVC} \cong \underset{\sim}{VS}^e) \in L, \\ SR_{32}=(\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^e \wedge \underset{\sim}{t}_{NVC} \cong \underset{\sim}{S}^e) \in LTH, \end{array} \quad (11)$$

$$SR_{33}=(\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{NVC} \cong \underset{\sim}{A}^{e}) \in HTTL,$$

$$SR_{34}=(\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{NVC} \cong \underset{\sim}{B}^{e}) \in H,$$

$$SR_{35}=(\underset{\sim}{t}_{VCA} \cong \underset{\sim}{S}^{e} \wedge \underset{\sim}{t}_{NVC} \cong \underset{\sim}{VB}^{e}) \in LIM\ \}.$$

Заключение

Предложенная в работе модель решающих правил на нечеткой логике, позволяет за счет использования множества пар «вторжение : величины», «вторжение : множество сопряженных пар» и МЭВ отображать аномальное состояние, порождаемое определенным типом кибератак в компьютерной сети. На основе этой модели были разработаны примеры правил для обнаружения таких вторжений как сканирование, спуфинг и Dos-атака, которые могут практически быть использованы для усовершенствования реальных систем обнаружения вторжений применяющих механизмы выявления аномалий, порожденных атакующими действиями в компьютерных системах.

Литература

[1] Корченко, Г. 2006. Построение систем защиты информации на нечетких множествах [Текст]: Теория и практические решения: МК-Пресс, 2006. 320 p.

[2] Ахметов Б.С., Корченко А.А., Жумангалиева Н.К. 2016 Модель базовых величин для контроля аномальности состояния среды окружения In: Вестник НАН РК. 2016. n. 1.

[3] Ахметов Б.С., Корченко А.А., Жумангалиева Н.К. 2015 - Базовые модели эталонных величин для систем обнаружения вторжений / Вестник МКТУ Х.А.Ясави . 2015. n. 4.

[4] Ахметов, Корченко, Ахметова, Жумангалеева. 2014. Использование методов экспертного оценивания в системах обнаружения вторжений In: Інформаційна безпека. 2014. n. 3 (15); n. 4 (16). p. 34-43.

Модель решающих правил для обнаружения аномалий в информационных системах

Аннотация:

Недостатком современных систем обнаружения вторжений, построенных на принципе идентификации аномального состояния является то, что они в основном ориентированы на использование таких математических моделей, которые требуют много времени на подготовку статистических данных. Математические модели, основанные на экспертных подходах в этом отношении являются более эффективными, но для выполнения своих функций необходимо использование соответствующих решающих правил. Для решения этой задачи в работе предложена модель решающих правил на нечеткой

логике, которая за счет использования множества пар «вторжение : величины» и «вторжение : множество сопряженных пар», а также модели эталонных величин позволяет отображать аномальное состояние, порождаемое определенным типом кибератак в компьютерной сети. На основе этой модели были разработаны примеры правил для обнаружения таких вторжений как сканирование, спуфинг и Dos-атака, которые могут практически использоваться для усовершенствования реальных систем выявления вторжений применяющих механизмы выявления аномалий, порожденных атакующими действиями в компьютерных системах.

Ключевые слова:

Кибератака, системы обнаружения вторжений, аномалия в сетевом трафике, обнаружение аномалий в компьютерных системах, сопряженная пара, решающие правила, экспертная оценка.

.....

Ахметов Бахытжан Сражатдинович, д. т. н., проф.
КазНИТУ им. К. И. Сатпаева bakhytzhan.akhmetov.54@mail.ru

Корченко Анна Александровна к. т. н., доц.
Национальный авиационный университет Украина г. Киев annakor@ukr.net

Жумангалиева Назым Кенжегалиевна, магистр
КазНИТУ им. К. И. Сатпаева
nazym_k.81@mail.ru

Dr. Ing. Jaroslav Kultan, PhD.
Ekonomická univerzita v Bratislave
Fakulta hospodárskej informatiky
jkultan@gmail.com