УДК 004.056.53(045)

Ахметов Б.С.[1*], Корченко А.А.[2**], Жумангалиева Н.К.[1***]

[1]Казахский национальный исследовательский технический университет имени К.И.Сатпаева, Республика Казахстан, г. Алматы

[2]Национальный авиационный университет, кафедра безопасность информационных технологий, Украина, Киев

E-mail: *bakhytzhan.akhmetov.54@mail.ru, **annakor@ukr.net, ***nazym_k.81@mail.ru

Технология выявления аномального состояния для систем обнаружения вторжений

Одним из решений обеспечения безопасности, являются системы обнаружения вторжений, построенные по аномальному принципу. Такие системы обычно основываются на математических методах, требующих много времени на подготовку статистических данных. Поэтому необходимы более эффективные методы, основанные на экспертных подходах. Для решения этой задачи предлагается технология, базирующаяся на математических моделях и методах нечеткой логики, и содержащая восемь базовых этапов (выбор метода обработки нечетких данных, выбор метода определения коэффициента важности, формирование множеств вторжений и величин, формирование эталонов величин, фаззификация величин, формирование множества решающих правил, определение матриц инициализации, формирование результата), раскрывающие процесс выявления аномального состояния, порождаемого определенным типом кибератак в информационных системах. Эту технологию можно использовать для создания или усовершенствования существующих систем выявления кибератак в компьютерных сетях.

Ключевые слова: кибератака, системы обнаружения вторжений, обнаружение аномалий в компьютерных системах, решающие правила, модель базовых величин, модель эталонных величин, модель решающих правил, построение решающих правил, технология выявления аномалий, технология обнаружения вторжений.

Akhmetov B.S., Korchenko A.A., Zhumangaliyeva N.K.

Technology of abnormal states for detection of intrusion systems

One of the security solutions are detection of intrusion systems based on the anomalous principle. Such systems are usually based on mathematical methods that require a lot of time for preparing statistics. That's why, a need for more effective methods based on expert approaches. In order to solve this problem, technology is proposed, based on mathematical models and methods of fuzzy logic, and contains eight basic steps (selection of fuzzy data processing method, the choice of method for determining the importance of the factor, the formation of sets of invasions and values, the formation of standards of size, fuzzification values, forming a plurality of critical rules, the definition of initialization matrix formation results), revealing the process of identifying an abnormal condition, generated by a specific type of cyber attacks in the information systems. This technology can be used to create or enhance existing detect systems of cyber attacks on computer networks.

Key words: cyber attack, detection of intrusion system, the detection of anomalies in computer systems, decision rules, the model of base units, a model of reference values, model of decision rules, construction of decision rules, anomaly detection technology, detection of intrusion technology.

Ахметов Б.С., Корченко А.А., Жумангалиева Н.К.

Шабуылдарды анықтау жүйесінің ауытқымалығының жағдайын анықтау технологиясы

Қауіпсіздікті қамтамасыз етудің бір шешімі ретінде ауытқымалығының принципі бойынша құрылған шабуылдарды анықтауға арналған жүйені атауға болады. Бұл жүйелер әдетте статистикалық мәліметтерді дайындауға көп уақыт қажет ететін математикалық әдістерге негізделеді. Сондықтан сарапшылық ұстанымдарға негізделген нәтижесі көбірек әдістер қажет. Бұл міндетті шешу үшін математикалық моделдер мен айқын емес қисын әдістеріне негізделген және ақпараттық жүйедегі кибершабуылдардың белгілі бір түрі туғызатын ауытқымалық жағдайды анықтау үрдісінің сегіз негізгі кезеңдерден тұратын (айқын емес мәліметтерді өңдеу әдісі, маңыздылық коэффициентін анықтау әдісі, шабуылдар жиыны мен шамаларды қалыптастыру,шама эталондарын қалыптастыру, шамалардың айқындалмауы, шешуші ережелер жиынтығын қалыптастыру, инициализация қалыптамаларын айқындау,нәтижені қалыптастыру) технология ұсынылады. Бұл технологияны компьютерлік жүйелердегі кибершабуылдарды анықтауға арналған жүйені құру немесе жетілдіру үшін пайдалануға болады.
Түйін сөздер: кибершабуыл, шабуылдарды анықтау жүйесі, компьютерлік жүйедегі аномалияларды анықтау, шешуші ережелер, негізгі шамалардың моделі, шама эталондарының моделі, шешуші ережелерді құру, аномалияларды анықтау технологиясы, шабуылдарды анықтау технологиясы.

1 Введение

Интенсивное развитие информационных технологий оказало положительное влияние на все сферы человеческой деятельности. Вместе с этим наблюдаются и побочные эффекты, в первую очередь в связи с тем, что ресурсы информационных систем (РИС) все больше подвергаются воздействиям кибератак, под которыми понимаются меры, предпринимаемые для подрыва безопасности информационной системы (ИС) или реализация угроз характеристикам безопасности РИС посредством использования их уязвимостей. Современный спектр вторжений на РИС достаточно широкий и только основываясь на базовые признаки их можно классифицировать по: автоматизации; взаимодействию с политикой безопасности; дистанционности; действию, порожденному несанкционированным доступом; внешнему проявлению; инициализационному условию; инструментальным средствам; наличию обратной связи; нарушению базовых характеристик безопасности; природе взаимодействия; реляционным признакам; специфике реализации; направленности результата; степени сложности; типу базового ресурса; семиуровневой эталонной модели [1]. В стремительно развивающейся информационной среде появляются новые виды угроз, порождающие новые виды кибератак на ее ресурсы. В этой связи существует потребность в системах безопасности построенных на основе технологий, позволяющих анализировать, контролировать, прогнозировать и блокировать такие вторжения. Одним из решений защиты РИС от указанных кибератак, являются системы обнаружения вторжений (СОВ), построенные по аномальному принципу. Такие системы обычно основываются на математических методах, требующих много времени на подготовку статистических данных. Поэтому необходимы более эффективные технологии основанные на экспертных подходах.

Отметим, что несанкционированные воздействия на РИС оказывают влияние на среду их окружения и порождают в ней определенные аномалии. Такая среда обычно слабоформализованная, нечетко определенная и для выявления вторжений, породивших ано-

малии в этой среде необходимы соответствующие технологии. В работах [1-3] показана эффективность применения математического аппарата нечетких множеств для решения такого рода задач, а его использование для формализации подхода к выявлению вторжений, позволит повысить эффективность разрабатываемых СОВ. В этой связи, целью данной работы является разработка технологии выявления аномалий, использование которой позволит синтезировать эффективно функционирующие системы, осуществляющие обнаружение вторжений по аномальному состоянию величин (например, сетевого трафика), характеризующих среду окружения. Под такой средой 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 ИС с целью выявления ее аномального состояния. В работах [4-6] предложена модель базовых величин (МБВ), модель эталонных величин (МЭВ) и модель решающих правил (МРП), которые возьмем за основу разработки соответствующей технологии. Реализация технологии осуществляется за восемь базовых этапов: 1) выбор метода обработки нечетких данных, 2) выбор метода определения коэффициента важности (КВ), 3) формирование множеств вторжений и величин, 4) формирование эталонов величин, 5) фаззификация величин, 6) формирование множества решающих правил (РП), 7) определение матриц инициализации, 8) формирование результата, которые представлены на рис. 1. Опишем каждый из них.

Этап 1 – выбор метода обработки нечетких данных. На этом этапе осуществляется выбор методов обработки нечетких данных относительно заданных критериев. В работе [7] рассмотрены три базовые группы соответствующих методов – формирования функций принадлежности (ФП) (четырнадцать методов – МФФП_1, МФФП_2, МФФП_3, $\ldots, \text{МФФП}_{14}$, например, метод корректировки параметров (КП), метод интервальных оценок (МИО), метод лингвистических термов с использованием статистических данных (МЛТС) и др.), сравнения функций принадлежности (восемь методов – МСФП_1, МСФП_2, МСФП_3, $\ldots, \text{МСФП}_8$, например, α – уровневое расстояние (АУР), функция упорядочения нечетких подмножеств (ФУ), метод поиска "центра тяжести"(ЦТ) и др.) и нечеткой арифметики (четырнадцать методов – МНА_1, МНА_2, МНА_3, $\ldots, \text{МНА}_{14}$, например, максминная композиция (ММК), α – уровневый принцип обобщения (АУПО), метод линейной аппроксимации по локальным максимумам (ЛАЛМ) и др.), из которых посредством процедур выбора МФФП, МСФП и МНА отбирается один из представителей. Процесс выбора осуществляется на основе заданных критериев. Так для всех групп методов базовыми критериями являются – используемый класс ФП и экспертная информация, для МФФП – использование ранговых оценок и число привлекаемых экспертов, а для МСФП – применение α – уровневого подхода. Если несколько методов будут отвечать установленным критериям, то окончательное решение о выборе будет основываться на предпочтении эксперта. Например, согласно принятых критериев для каждой группы возможных методов МФФП_i ($i = \overline{1,14}$), МСФП_j ($j = \overline{1,8}$) и МНА_k ($k = \overline{1,14}$), после реализации процедуры выбора определяется соответственно метод ЛАЛМ, АУР и МЛТС, которые совместно будут использоваться для обработки нечетких данных при решении задачи выявления аномального состояния в компьютерных системах.

Этап 2 – выбор метода определения коэффициента важности (МОКВ).

Этап ориентирован на выбор (согласно установленным критериям) метода формирования КВ из заданного множества. В работе [8] рассмотрено двадцать пять МОКВ ($МОКВ_1$, $МОКВ_2$, $МОКВ_3$, ... $МОКВ_{25}$, например, метод средних рангов (СР), мультипликативная свертка Кини (МСК), метод случайных векторов (СЛВ) и др.), среди которых в процессе реализации процедуры выбора определяется рабочий метод. Если несколько методов будут отвечать установленным критериям, то в данном случае окончательным решением о выборе будет принимать эксперт. Приоритет метода определяется посредством процедуры выбора МОКВ согласно таких критериев как: форма выражения входных (ВхД) и выходных (ВыхД) данных; трудоемкости и рекомендуемой шкалы [8]. Например, согласно установленных критериев и приоритетов эксперта из множества $МОКВ_i$ $(i = \overline{1, 25})$ выбирается метод СР.

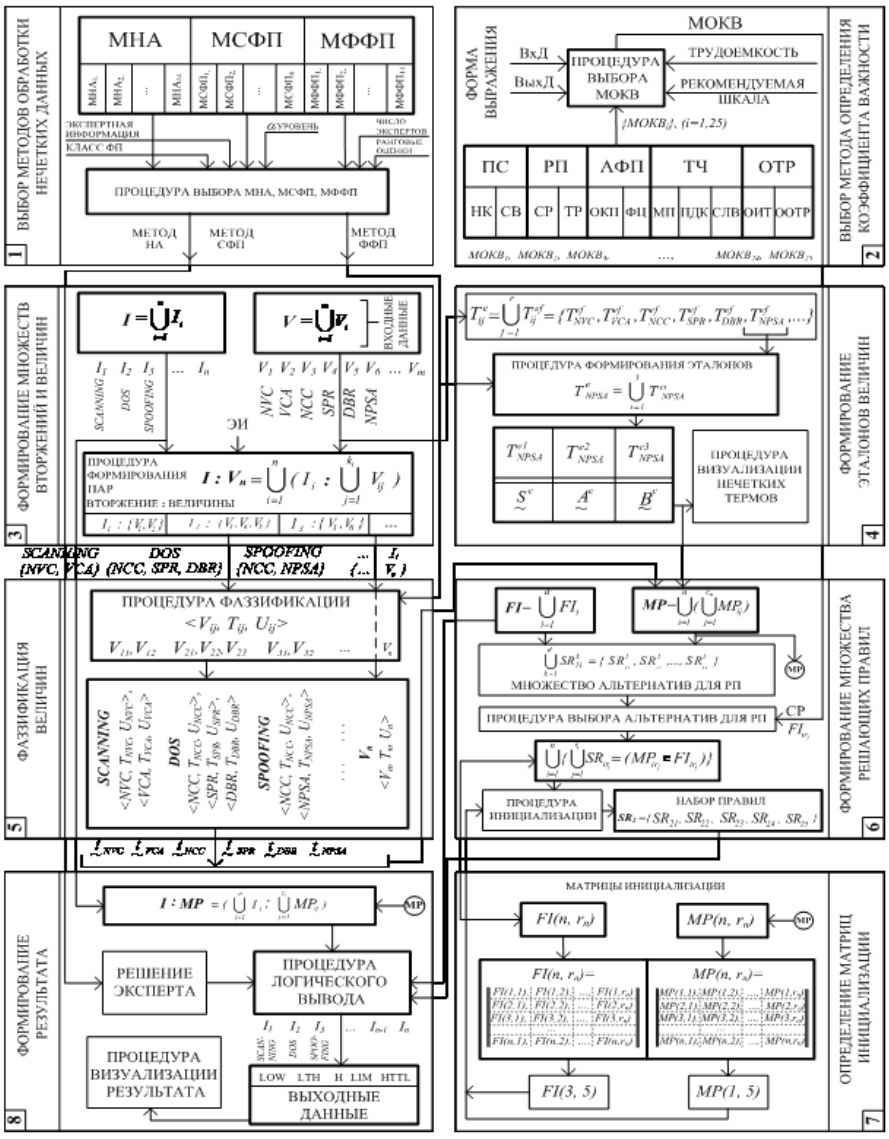

Рисунок 1 – Схема отображения технологии идентификации аномалий

Этап 3 – формирование множеств вторжений и величин. Этап предназначен для формирования множества вторжений и соответствующего им множества величин

для выявления аномального состояния. На основании входных величин среды окружения с использованием МБВ [4] формируются множество возможных вторжений $\mathbf{I} = \bigcup_{i=1}^{n} I_i$ и соответствующее им множество возможных величин $\mathbf{V} = \bigcup_{i=1}^{m} V_i$, согласно значений которых (например, $V_1 = NVC$, $V_2 = VCA$, $V_3 = NCC$, $V_4 = SPR$, $V_5 = DR$, $V_6 = NPS$, $\ldots$, V_m) с учетом решений экспертов можно выявить аномальное состояние, порождаемое определенным элементом из множества $\mathbf{I}$, например, ($I_1 = SCANNING$, $I_2 = DOS$, $I_3 = SPOOFING$, $\ldots, I_n$) [4]. Для выявления аномального состояния каждому типу вторжения $\mathbf{I}$ ставится в соответствие подмножество набора величин $\mathbf{V_n}$ из множества $\mathbf{V}$, по которым можно обнаружить подозрительную активность в системе. Таким образом, формируется множество пар – "вторжение : величины" $\mathbf{I} : \mathbf{V_n} = \bigcup_{i=1}^{n}(I_i : \bigcup_{j=1}^{k_i} V_{ij})$, в котором каждому вторжению будет соответствовать набор величин $(I_1 : \{V_1, V_2\})$, $(I_2 : \{V_3, V_4, V_5\})$, $(I_3\{V_3, V_6\})$, $\ldots, (I_n\{\ldots, V_n\})$, например, $(SCANNING\{NVC, VCA\})$, $(DS : \{NCC, SPR, DR\})$ и $(SPOOFING\{NCC, NPS\})$.

Этап 4 – формирование эталонов величин. Этот этап направлен на получение эталонов, которые необходимы для измерения текущих значений величин характеризующих среду окружения. На основании входных данных (см. этап 3) $\mathbf{V} = \bigcup_{i=1}^{m} V_i$, выбранного на первом этапе МФФП и с помощью процедуры формирования эталонных величин получаем соответствующие значения эталонов лингвистических переменных (ЛП) для всех $T_{ij}^{e} = \bigcup_{f=1}^{r} T_{ij}^{ef}$, например, $\{T_{NVC}^{ef}, T_{VCA}^{ef}, T_{NCC}^{ef}, T_{SPR}^{ef}, T_{DBR}^{ef}, T_{NPSA}^{ef}, \ldots\}$. Так, например, для NPSA [4] с использованием $\text{МФФП}_6 = \text{МЛТС}$ [1] можем получить эталонные значения $T_{NPSA}^{e} = \bigcup_{i=1}^{3} T_{NPSA}^{ei}$ и осуществить визуализацию лингвистических термов для **NPSA** – $\{T_{NPSA}^{e1}, T_{NPSA}^{e2}, T_{NPSA}^{e3}\} = \{\underset{\sim}{S}^{e}, \underset{\sim}{A}^{e}, \underset{\sim}{B}^{e}\}$. Далее с помощью процедуры визуализации формируется графическое представление эталонов лингвистических термов $\{\underset{\sim}{S}^{e}, \underset{\sim}{A}^{e}, \underset{\sim}{B}^{e}\}$.

Этап 5 – фаззификация величин. На этом этапе осуществляется преобразование набора подмножеств величин, характеризующих текущее состояние системы, в соответствующие им текущие значения нечетких переменных. На основании МБВ [4], выбранного (на первом этапе) метода получения ФП и с помощью процедуры фаззификации, реализующей один из МФФП формируется набор ЛП, каждая из которых представляется кортежем $< V_{ij}, T_{ij}, U_{ij} >$. Далее на основе процедуры, связывающей с каждым вторжением из множества $\mathbf{I}$ конкретный набор величин из множества $\mathbf{V}$, получаем множества пар [4] $\mathbf{I} : \mathbf{V_n} = \bigcup_{i=1}^{n}(I_i : \bigcup_{j=1}^{k_i} V_{ij})$. Так, например, с использованием множества пар "вторжение : величины", $\text{МФФП}_6 = \text{МЛТС}$ (см. этап 1) и набора кортежей, отображающих соответствующие значения ЛП для вторжения SCANNING (при V_{11}, V_{12} – $< NVC, T_{NVC}, U_{NVC} >$, $< VCA, T_{VCA}, U_{VCA} >$), DOS (при V_{21}, V_{22}, V_{23} – $< NCC, T_{NCC}, U_{NCC} >$, $< SPR, T_{SPR}, U_{SPR} >$, $< DR, T_{DBR}, U_{DBR} >$) и SPOOFING (при V_{31}, V_{32} – $< NCC, T_{NCC}, U_{NCC} >$, $< NPS, T_{NPSA}, U_{NPSA} >$) формируются текущие зна-

чения нечетких переменных среды окружения $t_{\sim NVC}$, $t_{\sim VCA}$, $t_{\sim NCC}$, $t_{\sim SPR}$, $t_{\sim DBR}$, и $t_{\sim NPSA}$, которые соответственно отражают величины NVC, VCA, NCC, SPR, DBR и NPSA.

Этап 6 – формирование множества решающих правил (РП). Этап ориентирован на формирование РП необходимых для измерения текущего состояния системы относительно эталонных величин. На основании множеств нечетких идентификаторов $\mathbf{FI} = \bigcup_{i=1}^{d} FI_i$ [4] и сопряженных пар $\mathbf{MP} = \bigcup_{i=1}^{n}(\bigcup_{j=1}^{c_n} MP_{ij})$ [4] (использующих конкретные значения лингвистических термов, определенных на четвертом этапе) формируется множество альтернатив SR_{ij}^{k} ($i = \overline{1,n}$, $k = \overline{1,d}$, $j = \overline{1,r_n}$, где n – количество вторжений, r_n – количество правил для выявления i – го вторжения, а d – количество альтернативных вариантов для формирования одного правила). Например, для первого вторжения и первого правила это будет $\bigcup_{k=1}^{d} SR_{11}^{k} = \{SR_{11}^{1}, SR_{11}^{2}, SR_{11}^{3}, SR_{11}^{4}, SR_{11}^{5}\}$. Для построения РП, отображаемых выражением $\bigcup_{i=1}^{n}\{\bigcup_{j=1}^{r_i} SR_{ir_j} = (MP_{ir_j} \in FI_{ir_j})\}$ [6]. Формирование правил осуществляется на основе множества альтернатив с помощью процедуры их выбора, которая базируются на одном из методов формирования КВ (см. этап 2). Далее, отобранные FI_{ir_j} на этапе 7 используются в качестве данных для матриц инициализации, которые посредством процедуры инициализации передают конкретные значения в MP_{ir_j} и FI_{ir_j}, формируя таким образом непосредственные наборы РП, например,

$$\mathbf{SR_2} = \{SR_{21} = (t_{\sim NPSA} \cong \underset{\sim}{B}^{e} \wedge t_{\sim NCC} \cong \underset{\sim}{VS}^{e}) \in L,$$

$$SR_{22} = (t_{\sim NPSA} \cong \underset{\sim}{B}^{e} \wedge t_{\sim NCC} \cong \underset{\sim}{S}^{e}) \in LTH,$$

$$SR_{23} = (t_{\sim NPSA} \cong \underset{\sim}{B}^{e} \wedge t_{\sim NCC} \cong \underset{\sim}{A}^{e}) \in HTTL,$$

$$SR_{24} = (t_{\sim NPSA} \cong \underset{\sim}{B}^{e} \wedge t_{\sim NCC} \cong \underset{\sim}{B}^{e}) \in H,$$

$$SR_{25} = (t_{\sim NPSA} \cong \underset{\sim}{B}^{e} \wedge t_{\sim NCC} \cong \underset{\sim}{VB}^{e}) \in LIM\} \quad [6].$$

Этап 7 – определение матриц инициализации. Этап предназначен для формирования исходных данных (в виде набора матриц) для процедуры инициализации РП. На основе полученных конкретных значений всех FI_{ir_j} с помощью процедуры выбора альтернатив для РП и данных по конкретным парам FI_{ij} (см. этап 6) соответственно определяем матрицы инициализации для нечетких идентификаторов $FI(n, r_n)$ и сопряженных пар $MP(n, r_n)$, где n – количество вторжений, а r_n – количество правил для выявления i – го вторжения. Например, такие матрицы для использования на этапе 6 при построении РП имеют вид – $FI(3, 5)$ и $MP(3, 5)$, а их конкретные элементы отображены в [6].

Этап 8 – формирование результата. Этот этап направлен на получение выходных данных, характеризующих аномальное состояние. На основе сформированных множеств возможных вторжений (см. этап 3) и сопряженных пар (см. этап 6), формируется множество пар – "вторжение : множество сопряженных пар" $\mathbf{I} : \mathbf{MP} = (\bigcup_{i=1}^{n} I_i : \bigcup_{j=1}^{c_i} MP_{ij})$ [4]. Посредством этого множества, сформированных РП и множества **FI** (см. этап 6), с

помощью процедуры логического вывода (функционирующей на основе выбранных по решению эксперта МНА и МСФП) определяются конкретные значения нечетких идентификаторов, характеризующих уровень аномального состояния, который может быть порожден конкретной кибератакой. Другими словами каждому I_i присваивается один из FI_i. Так, например, вторжениям $I_1 = SCANNING$, $I_2 = DOS$ и $I_3 = SPOOFING$ соответственно будет определен уровень LOW, LTH и H [6].

Выводы: После определения этих результатов осуществляется их визуализация в виде эталонных лингвистических термов, на фоне которых идентифицируется значение переменной, характеризующей текущее состояние системы относительно аномалий.

Предложенная в работе технология базируется на математических моделях и методах нечеткой логики, и содержит восемь базовых этапов, раскрывающих процесс выявления аномального состояния, порождаемого определенным типом кибератак в ИС. На основе этой технологии можно создавать или усовершенствовать реальные системы обнаружения вторжений применяющих механизмы выявления аномалий, порожденных атакующими действиями в компьютерных сетях.

Литература

[1] *Корченко О.Г.* Построение систем защиты информации на нечетких множествах // Теория и практические решения / О.Г.Корченко. – К.: МК-Пресс, 2006. – 320 с.

[2] *Волянська В.В.* Система виявлення аномалій на основі нечітких моделей [Текст] / В. В. Волянська, А. О. Корченко, Є. В. Паціра // Зб. наук. пр. Інституту проблем моделювання в енергетиці НАН України ім. Г. Є Пухова. – Львів : ПП "Системи, технології, інформаційні послуги 2007. – [Спец. випуск]. – Т.2. – С. 56-60.

[3] *Корченко О.Г.* Системи захисту інформації [Текст] : Монографія / О. Г. Корченко. – К.: НАУ, 2004. – 264 с.

[4] *Ахметов Б.С., Корченко А.А., Жумангалиева Н.К.* Модель базовых величин для контроля аномальности состояния среды окружения // Вестник НАН РК. – 2016. – № 1(305) – 26 с.

[5] *Ахметов Б.С., Корченко А.А., Жумангалиева Н.К.* Базовые модели эталонных величин для систем обнаружения вторжений / Вестник МКТУ Х.А.Ясави. – 2015. – № 4.

[6] *Ахметов Б.С.* Использование методов нечетких множеств в системах обнаружения вторжений / Б.С. Ахметов, А.А. Корченко, Н.К. Жумангалиева // Інформаційна безпека. – 2014. – № 1 (13); № 2 (14). – С. 42-55.

[7] Использование методов экспертного оценивания в системах обнаружения вторжений / Б.С. Ахметов, А.А. Корченко, С.Т. Ахметова, Н.К. Жумангалиева // Інформаційна безпека. – 2014. – № 3 (15); № 4 (16). – С. 34-43.

References

[1] *Korchenko O.G.* Postroenie sistem zaschityi informatsii na nechetkih mnozhestvah // Teoriya i prakticheskie resheniya / O.G.Korchenko. – K.: MK-Press, 2006. – 320 s.

[2] *Volyanska V.V.* Sistema viyavlennya anomaliy na osnovi nechitkih modeley [Tekst] / V. V. Volyanska, A. O. Korchenko, E. V. PatsIra // Zb. nauk. pr. Institutu problem modelyuvannya v energetitsI NAN Ukrayini Im. G. E Puhova. – LvIv : PP "Sistemi, tehnologii, Informatsiyni poslugi 2007. – [Spets. vipusk]. – T.2. – S. 56-60.

[3] *Korchenko O.G.* Sistemi zahistu Informatsiyi [Tekst] : Monografiya / O. G. Korchenko. – K.: NAU, 2004. – 264 s.

[4] *Ahmetov B.S., Korchenko A.A., Zhumangalieva N.K.* Model bazovyih velichin dlya kontrolya anomalnosti sostoyaniya sredyi okruzheniya // Vestnik NAN RK. – 2016. – No 1(305) – 26 s.

[5] *Ahmetov B.S., Korchenko A.A., Zhumangalieva N.K.* Bazovyie modeli etalonnyih velichin dlya sistem obnaruzheniya vtorzheniy / Vestnik MKTU H.A.Yasavi. – 2015. – No 4.

[6] *Ahmetov B.S.* Ispolzovanie metodov nechetkih mnozhestv v sistemah obnaruzheniya vtorzheniy / B.S. Ahmetov, A.A. Korchenko, N.K. Zhumangalieva // InformatsIyna bezpeka. – 2014. – No 1 (13); No 2 (14). – S. 42-55.

[7] Ispolzovanie metodov ekspertnogo otsenivaniya v sistemah obnaruzheniya vtorzheniy / B.S. Ahmetov, A.A. Korchenko, S.T. Ahmetova, N.K. Zhumangalieva // Informatsiyna bezpeka. – 2014. – No 3 (15); No 4 (16). – S. 34-43.