УДК 681.142.33:681.14

С.А. Агеев

ПРИМЕНЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ МЕТОДОВ ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ ДЛЯ УПРАВЛЕНИЯ РИСКАМИ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В ЗАЩИЩЁННЫХ МУЛЬТИСЕРВИСНЫХ СЕТЯХ СПЕЦИАЛЬНОГО НАЗНАЧЕНИЯ

Агеев С.А. **Применение интеллектуальных методов представления информации для управления рисками информационной безопасности в защищённых мультисервисных сетях специального назначения.**
Аннотация. Рассматриваются основные способы применения интеллектуальных методов и алгоритмов, синтезированных на их основе, представления данных сетевого мониторинга для управления рисками информационной безопасности защищенных мультисервисных сетей (ЗМС). Разработана и исследована математическая модель интеллектуального представления данных для анализа и оценки риска информационной безопасности ЗМС.
Ключевые слова: защищенная мультисервисная сеть, кластеризация, классификация, автоматизация управления, математическая модель, интеллектуальное управление, нечеткий логический вывод, база знаний, функция принадлежности, нечеткие данные, лингвистическая переменная.

Ageev S. A. **Application of Intelligent Methods of Information Representation to Manage Information Security Risks in Secure Multiservice Special-Purpose Networks.**
Abstract. The article considers main methods to use intelligent techniques and algorithms, synthesized on their basis, as well as examinesdata presentation of network monitoring for IT security risk management of secure multiservice networks. The mathematical model of intelligent data presentation is developed and examined for IT security risk investigation and assessment.
Keywords: secure multiservice networks, clusterization, classification, control automation, mathematical model, intelligentcontrol, fuzzy inference, knowledgebase, membership function, fuzzy data, linguistic variable.

1. Введение. Важнейшей проблемой при создании и эксплуатации защищенных мультисервисных сетей специального назначения (ЗМС СН) является проблема обеспечения их безопасного функционирования и безопасности циркулирующей в ней информации [1, 2]. Для решения данной проблемы необходимо применение методов, которые позволяют оперативно оценивать риски информационных угроз с заданной степенью достоверности, а также на основании анализа этих оценок вырабатывать и реализовывать оперативное управление, направленное на их снижение [3]. Под оперативностью управления понимают [4, 5]:

$$P_{оу} = P\left\{ t_у \le T_{зу} \right\}, \qquad (1)$$

где P_{oy} - вероятность события заключающегося в том, что время цикла управления не превысит заданное время $T_{зу}$. В свою очередь, время цикла управления t_y складывается из времени сбора информации о состоянии сетевых элементов $t_{сб}$, поступающей от подсистемы сетевого мониторинга, времени анализа информации t_a, времени выработки решений t_p, времени доведения управляющей информации до соответствующих сетевых элементов $t_д$, времени реализации сетевыми элементами управленческих решений t_u и времени подтверждения сетевыми элементами выполнения управленческих решений t_n. Таким образом, можно записать соотношение:

$$t_y = t_{сб} + t_a + t_p + t_д + t_u + t_n . \quad (2)$$

Откуда следует, что повышение оперативности цикла управления заключается в снижении значений слагаемых в выражении (2).

Учитывая разноплановость, многокритериальность, большую размерность решаемых задач по управлению ЗМС СН, часть процедур управления предлагается реализовывать на основе технологии интеллектуальных мультиагентов (ИМА), основой которых является технология «агент-менеджер» [3, 6].

Данная работа посвящена решению проблемы оперативной обработки и представления исходных данных для подсистемы поддержки принятия решений при управлении информационной безопасностью (ИБ) в ЗМС СН.

2. Постановка задачи исследования. Основными способами повышения оперативности управления в данной работе предлагаются способы уменьшения времени анализа исходных данных и способы уменьшения времени выработки и принятия управленческих решений. При этом качество управления, а именно значения целевых функций, которые подлежат оптимизации, должны оставаться в области Парето-оптимальных значений. Для решения данной проблемы в работе предлагается применение методов нечётких кластеризации, классификации и ранжирования [3].

Пусть с выходов сенсоров систем обнаружения и предупреждения вторжений (СОВ, СПВ), расположенных в составе встроенных в сетевые элементы (СЭ) интеллектуальных

мультиагентов [3, 6], за время ΔT_i поступают данные в виде совокупности обнаруженных признаков текущего состояния ИБ $X_i = \{x_{i1}, x_{i2},\ldots, x_{ij}, x_{in}\}$, где $i = \overline{1,M}$. Элементы множества $\{X_i\}$ могут иметь различную физическую природу, различную размерность, а также могут иметь различные системы измерений и шкалирования. Подготовка данных для функционирования подсистемы принятия решений об уровне риска угроз ИБ заключается в выполнении следующих процедур:

– производится сортировка на классы предполагаемых угроз одним из известных методов [3, 7, 8], например, определяется атакуется ли СУБД конкретного СЭ, операционная система СЭ, сетевой интерфейс СЭ и т. д.;

– производится кластеризация входных данных для каждого класса угроз ИБ СЭ ЗМС СН;

– производится классификация полученных кластеров на соответствие известным классам и видам деструктивных воздействий на сетевые элементы ЗМС СН;

– производится ранжирование по степени риска ИБ полученных данных.

Выполнение вышеперечисленных процедур должно быть реализовано за время меньшее, чем интервал ΔT_i. Данный подход иллюстрирует рисунок 1. На этом рисунке X - признаки состояния ИБ СЭ, Y-кластеры известных угроз ИБ СЭ, Z- классы неизвестных угроз ИБ СЭ. Очевидно, что выполняется следующее отношение:

$$Y \subseteq X, Z \subseteq X, \tag{3}$$

$W = Y \cup Z$ – множество угроз ИБ СЭ ЗМС СН.

Таким образом, необходимо построить ранжированный ряд угроз ИБ $r_1 \prec r_2 \prec \ldots \prec r_n$, где r_k - ранжированное значение угрозы, k– класс угрозы, N_k - количество угроз данного класса, w_k - весовой коэффициент угрозы, ΔT_i - период времени, за который появились угрозы данного класса, то есть:

$$r_k = \frac{N_k \times w_k}{\Delta T_i}. \tag{4}$$

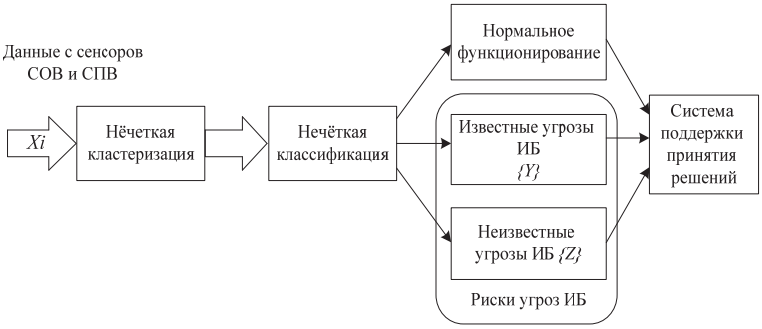

Рис. 1. Обобщённая структура взаимосвязи процедур кластеризации, классификации и ранжирования

Полученные данные поступают на вход системы поддержки принятия решения, которая автоматически генерирует управляющее воздействие для данного СЭ, направленное на уменьшение риска ИБ, а также информирует об этом вышестоящий интеллектуальный менеджер управления [3, 6].

3. Применение методов нечётких кластеризации и классификации для обработки первичной информации СОВ и СПВ. Для решения сформулированной выше задачи предлагается использовать методы нечёткой кластеризации и нечёткой классификации [3, 6, 7], которые позволяют улучшить показатели оперативности цикла управления рисками ИБ ЗМС СН.

Процедура кластеризации заключается в разбиении всего множества признаков состояния ИБ на группы по некоторым признакам. Полученные группы называют кластерами. Таким образом, формальная постановка задачи кластеризации имеет следующий вид.

Дано конечное множество объектов $X_i = \{x_{i1}, x_{i2},\ldots, x_{ij}, , x_{in}\}$, где $i = \overline{1,M}$. Каждый из объектов характеризуется m-компонентным признаковым описанием $(p_1, p_2,\ldots, p_k,.., p_m)$, $p_k \in P_k$, где P_k – допустимое множество значений признака. Требуется построить множество кластеров (построить разбиение множества X_i) $\{C_i\}$, где $i = \overline{1,c}$ и отображение f: $X \to C$, со следующими свойствами:

$$\bigcup_{i=\overline{1,c}} C_i = X, C_i \cap C_j = \emptyset,\ i,j = \overline{1,c},\ i \neq j,$$
$$\emptyset \subset C_i \subset X,\ i = \overline{1,c}. \qquad (5)$$

При этом для процедур субтрактивной кластеризации количество кластеров c подлежит определению. Для оценки качества разбиения

применяется критерий разброса, показывающий сумму расстояний в выбранной метрике от координат признаков до координат центра своего кластера. Например, для евклидовой метрики получаем следующий критерий разбиения:

$$\sum_{i=\overline{1,c}} \sum_{X_k \in C_i} \left\| V_i - X_i \right\|^2 \to min\,, \tag{6}$$

где: $C_i = \{x_p : \varphi_{k_i} = 1, k = \overline{1,M}\}$ – i–ый кластер; $V_i = \frac{1}{|C_i|} \sum_{x_k \in C_i} x_k$ - центр i-го кластера; $|C_i|$ - мощность множества признаков кластера C_i.

Формальная постановка задачи нечёткой кластеризации имеет следующие особенности.

Нечёткие кластеры определяются матрицей нечёткого разбиения:

$$F(X) = [\mu_{ki}],\ \mu_{ki} \in [0, 1]: k = \overline{1,M},\ i = \overline{1,c}\,, \tag{7}$$

в которой k-ая строка содержит степень принадлежности объекта X_k кластерам $\{C_i\}$. При этом должны выполняться условия:

$$\sum_{i=1,c} \mu_{ki} = 1,\ k = \overline{1,M}, \quad 0 < \sum_{k=1,M} \mu_{ki} < M,\ i = \overline{1,c}\,. \tag{8}$$

Критерий качества нечёткой кластеризации в простейшем случае имеет вид:

$$\sum_{i=\overline{1,c}} \sum_{k=1,M} (\mu_{ki})^m \left\| V_i - X_k \right\|^2 \to min, \tag{9}$$

где: $V_i = \frac{\sum_{k=1,M} (\mu_{ki})^m X_k}{\sum_{k=1,M} (\mu_{ki})^m}$ - центр нечёткого кластера; $m \in (1, \infty)$-экспоненциальный вес. Схема алгоритма субтрактивной (горной) кластеризации представлена на рисунке 2. Данный алгоритм является чётким алгоритмом, но, при соответствующем выборе шкал признаков кластеров, он может применяться и для обработки нечётких исходных данных.

Рис. 2. Схема алгоритма горной кластеризации

Алгоритм нечёткой классификации предлагается реализовывать на основе метода нечёткого логического вывода Мамдани [7, 8], имеющего вид:

$$\bigcup_{i=1}^{c} \left(\bigcap_{j=1}^{m} x_i = a_{i,j} \times p_j \right) \rightarrow y_j = d_j,\ j = \overline{1,m}, \tag{10}$$

где x_i - набор входных признаков;

y_j - выходная переменная j- го правила;

$a_{i,j}$ - нечёткий терм, которым оценивается переменная x_i в правиле j базы знаний;

p_j - весовой коэффициент правила j;

d_j - набор значений выходной переменной y_j.

Степени принадлежности $\mu_{dj}(X^*)$признака классифицируемой угрозе ИБ вычисляется по методам, изложенным в [3, 7, 8]. Признаки ИБ тоже являются нечёткими величинами. Каждый признак характеризуется степенью наблюдаемости [7, 8], которым присваиваются лингвистические термы. Ранжирование классифицируемых рисков угроз ИБ СЭ ЗМС СН осуществляется в соответствие с (2). При реализации процедур нечёткой классификации возможно предварительное обучение нечёткого классификатора [7, 8].

В процессе функционирования корректировка баз знаний возможна в соответствие с методами, изложенными в [8].

Исходные данные, поступающие на вход алгоритма кластеризации должны удовлетворять следующим свойствам [7]:

– функция распределения вероятностей разброса признаков относительно центра кластера должна аппроксимироваться гауссовой функцией распределения;

– признаки состояния системы должны быть некоррелированными;

– значения признаков должны быть либо безразмерными, либо иметь одни и те же единицы измерения;

– распределение значений признаков должны быть устойчивы к влиянию случайных факторов;

– совокупности распределений значений признаков должны быть однородны и не содержать неких выбросов.

4. Анализ результатов математического моделирования. На рисунке 3 приведены некоторые результаты программной реализации предложенных методов кластеризации и классификации для двумерного случая.

Признаки каждого кластера генерировались с помощью двумерного гауссова распределения, имеющего вид:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} exp\{-(\frac{(x_i - x_u)^2}{2\sigma_x^2} + \frac{(y_i - y_u)^2}{2\sigma_y^2})\}, \quad (11)$$

где σ_x - среднеквадратическое отклонение признака относительно центра по оси x; σ_y - среднеквадратическое отклонение признака

относительно центра по оси y; $(x_ц, y_ц)$ – координаты центра кластера; (x_i, y_i) - координаты центра i–го признака данного кластера.

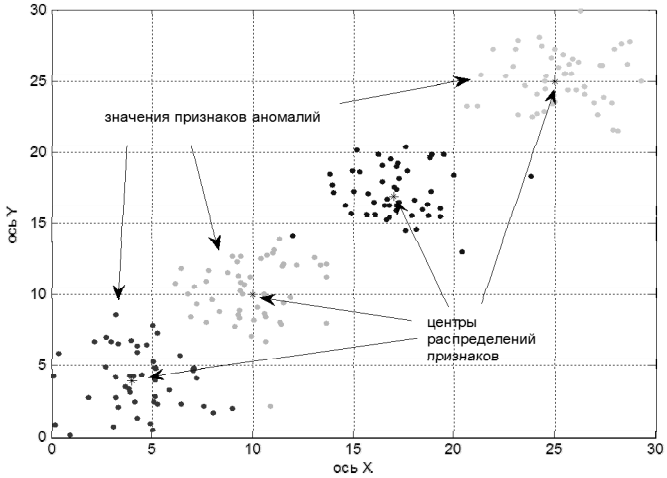

Рис. 3. Распределение признаков, центры кластеров

В таблице 1 приведены характеристики исследуемой модели признаков. Результаты кластеризации приведены на рисунках 4 и 5. На этих рисунках приведён результат выполнения алгоритма субтрактивной кластеризации для двумерного случая.

Таблица 1. Характеристики модели признаков угроз ИБ

Параметры	Первый кластер	Второй кластер	Третий кластер	Четвёртый кластер
Координаты центров распределения признаков в кластерах	(4, 4)	(10, 10)	(17, 17)	(25, 25)
Среднеквадратические отклонения по осям x и y	(2, 2)	(2, 2)	(2, 2)	(2, 2)
Координаты центров кластеров	(1.8; 2.8)	(8.4; 8.9)	(17.2; 18.3)	(25.8; 25.6)
Значение потенциала кластера	0.4	1	0.82	0.64
Количество признаков в кластере	50	50	50	50
Расстояние от центра распределения до центра кластера	2.51	1.94	1.32	1.0

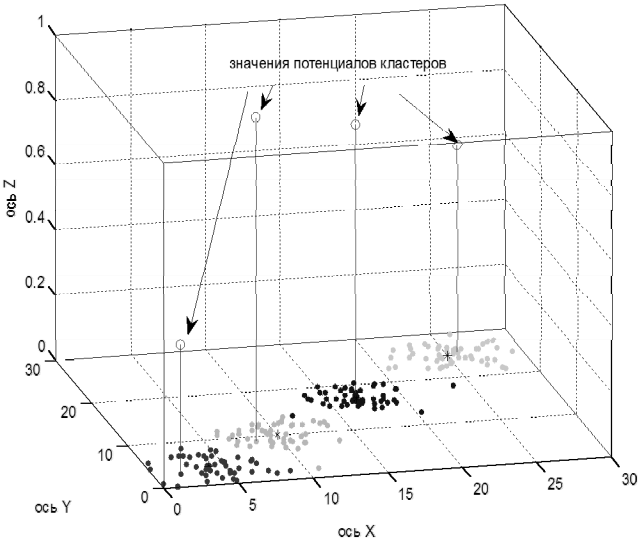

Рис. 4. Признаки кластеров, их центры, значение потенциалов кластеров

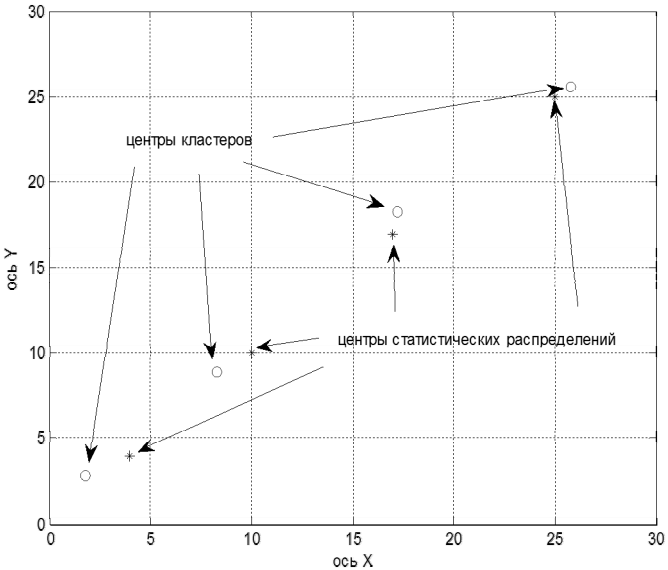

Рис. 5. Результаты кластеризации

Проведённые исследования алгоритма субтрактивной кластеризации показали его высокую устойчивость к различным

изменениям параметров распределения признаков. Из рисунка 3 видно, что области проекций значений признаков на координатные оси перекрываются для различных кластеров. Тем не менее, производится верное определение числа кластеров. Дальнейшим направлением исследования применения алгоритмов кластеризации данного класса (класс алгоритмов классификации без учителя) могут быть исследования, направленные на автоматическую адаптацию его параметров в зависимости от характеристик совокупности анализируемых признаков.

Правила базы данных процедуры классификации (10) предлагается реализовывать в следующем виде:

$$ЕСЛИ < (C_j \in R_j) \; И \; (R_j \in (L)) > ТО < УГРОЗА \; J - ИЗВЕСТНА >, \qquad (12)$$

где L – номер области пространства признаков. То есть, если центр кластера находится в достаточной близости от центра кластера известной угрозы, то считается, что риск идентифицирован с определённым значением коэффициента ранжирования. Если центр кластера находится на большом расстоянии от центров известных угроз, то независимо от того в какой области он находится, риск считается максимальным, так как характеристики угрозы не определены. На рисунке 6 приведён пример разбиения области пространства признаков по степени значимости угроз. Область 1 – риск допустим, область 2 – средняя степень риска, область 3 – высокая степень риска.

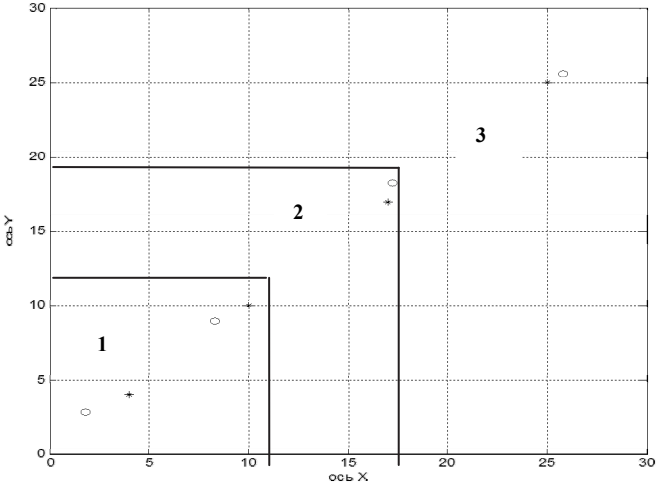

Рис. 6. Возможное разбиение пространства признаков и кластеров для оценки степени риска ИБ ЗМС СН

Далее, полученные данные поступают на вход системы поддержки принятия решений [3].

Признаки состояния для каждого класса угроз в данном численном эксперименте представлялись лингвистическими термами с соответствующими функциями принадлежности.

Анализ алгоритмов классификации и ранжирования показал их высокую устойчивость и эффективность для различных групп кластеров, для различного количества кластеров и различных их пространственных положений.

5. Заключение. Проведённые в работе исследования предложенных методов показали их высокую эффективность, их высокие точностные характеристики, простоту программной реализации, принципиальную возможность их функционирования в режиме, близком к режиму реального времени. Оперативность выработки управленческих решений при применении данных методов может быть улучшена по сравнению со статистическими методами приблизительно на 8%-13%, так как уменьшается время получения оценок значений признаков.

Литература

1. *Агеев С.А., Бушуев А.С., Егоров Ю.П., Саенко И.Б.* Концепция автоматизации управления информационной безопасностью в защищенных мультисервисных сетях специального назначения // Автоматизация процессов управления. 2011. №1. С.50–57.
2. *Саенко И.Б., Агеев С.А., Шерстюк Ю.М., Полубелова О.В.* Концептуальные основы автоматизации управления защищенными мультисервисными сетями // Проблемы информационной безопасности. Компьютерные системы. 2011. №3. С. 30–39.
3. *Агеев С.А., Саенко И.Б., Егоров Ю.П., Гладких А.А., Богданов А. В.* Интеллектуальное иерархическое управление рисками информационной безопасности в защищенных мультисервисных сетях специального назначения // Автоматизация процессов управления. 2014. № 3 (37). С.78–88.
4. Автоматизация управления и связь в ВМФ. Изд. 2-е / под общей редакцией Ю.М. Кононова. СПб.: «Элмор». 2001. 512 с.
5. *Азаров Г.И.* Теоретические основы анализа оперативности передачи информации в системах управления и связи // М.: Академия ГПС МЧС России. 2012. 62 с.
6. *Котенко И.В.* Интеллектуальные механизмы управления кибербезопасностью // Труды ИСА РАН. 2009. Т.41. С. 74–103.
7. *Белов В.В., Смирнов А.Е., Чистякова В.И.* Распознавание нечетко определяемых состояний технических систем // М.: Горячая линия–Телеком. 2012. 138 с.
8. *Штовба С. Д.* Проектирование нечётких систем средствами MATLAB // М.: Горячая линия – Телеком. 2007. 288 с.
9. *Платонов В.В.* Программно-аппаратные средства защиты информации // М.: Издательский центр «Академия». 2013. 336 с.

References

1. Ageev S.A., Bushuev A.S., Egorov Ju.P., Saenko I.B. [Concept of Automation of Information-Security Control in Protected Special-Purpose Multi-Service Networks] *Avtomatizacija processov upravlenija – Control Systems Automation*. 2011. vol. 1. pp. 50–57. (In Russ.).
2. Saenko I.B., Ageev S.A., Sherstjuk Ju.M., Polubelova O.V. [Conceptual basics of automation control of protected multi-service networks]. *Problemy informacionnoj bezopasnosti. Komp'juternye sistemy – Information security. Computer systems.* 2011. vol. 3. pp. 30–39. (In Russ.).
3. Ageev S.A., Saenko I.B., Egorov Ju.P., Gladkih A.A., Bogdanov A. V. [Intelligent hierarchical management of the information security risks in protected special purpose multi-service networks.] *Avtomatizacija processov upravlenija – Control Systems Automation*. 2014. vol. 3(37). pp. 78–88. (In Russ.).
4. *Avtomatizacija upravlenija i svjaz' v VMF. Izd. 2-e. Pod obshhej redakciej Ju.M. Kononova* [Control automation and connection at Navy. 2nd ed. Editeb by Ju.M. Kononov]. SPb.: Jelmor Publ. 2001. 512 p. (In Russ.).
5. Azarov G.I. *Teoreticheskie osnovy analiza operativnosti peredachi informacii v sistemah upravlenija i svjazi* [Theoretical bases analysis of efficiency transferring information in systems of control and communications]. Moscow: Academy GPS EMERCOM of Russia Publ. 2012. 62 p. (In Russ.).
6. Kotenko I.V. [Intelligent control mechanisms of cyber security]. *Trudy ISA RAN – ISA RAS Proceedings.* 2009. vol. 41. pp. 74–103. (In Russ.).
7. Belov V.V., Smirnov A.E., Chistjakova V.I. *Raspoznavanie nechetko opredeljaemyh sostojanij tehnicheskih sistem.* [Identification in outline specify the state of technological ensembles]. Moscow: Hotline Telecom Publ. 2012. 138 p. (In Russ.).
8. Shtovba S. D. *Proektirovanie nechjotkih sistem sredstvami MATLAB*. [Design fuzzy systems by MATLAB]. Moscow: Hotline Telecom. 2007. 288 p. (In Russ.).
9. Platonov V.V. *Programmno-apparatnye sredstva zashhity informacii.* [Hardware and software of information security]. Moscow: Academy Publ. 2013. 336 p. (In Russ.).

Агеев Сергей Александрович — к-т техн. наук, доцент, начальник научно-технического центра, ОАО "НИИ "Нептун". Область научных интересов: искусственный интеллект, управление телекоммуникационными системами, адаптивные и интеллектуальные методы управления большими системами, применение математических методов в задачах управления системами связи. Число научных публикаций — 94. serg123_61@mail.ru; 7-я линия В.О., д. 80, к.1А, Санкт-Петербург, 199178; р.т.: +7(812) 327-09-72.

Ageev Sergey Aleksandrovich — Ph.D., associate professor, head of technical researcher centre, Scientific Research Institute NII Neptun. Research interests: adaptive and intelligent technique of control large systems, use mathematical methods in control problem of processing systems. The number of publications — 94. serg123_61@mail.ru; 80, 7th line of Vasilievsky Island, St.Petersburg, 199178 Russia; office phone: +7(812) 327-09-72.

РЕФЕРАТ

Агеев С.А. **Применение интеллектуальных методов представления информации для управления рисками информационной безопасности в защищённых мультисервисных сетях специального назначения.**

В статье предложены методы интеллектуального анализа данных, используемые для управления рисками информационных угроз элементам мультисервисных сетей специального назначения (ЗМС СН). Обосновывается, что скоротечность процессов в ЗМС СН, их многообразие, неточность и неполнота, а также большая размерность априорных данных о состояниях сетевых элементах и происходящих в них процессах приводят к необходимости применения интеллектуальных методов обработки данных и управления. Проведён анализ факторов, влияющих на время цикла управления информационной безопасностью (ИБ) ЗМС СН, предложена единая метрика оценки угроз информационной безопасности элементам ЗМС СН.

В представленных результатах исследований разработана единая математическая модель процедур кластеризации, классификации и ранжирования угроз ИБ, которые могут иметь ранее неизвестную структуру. Определены критерии качества функционирования предложенных процедур. Представлены результаты исследования разработанной математической модели. Сформулированы правила оценки степени угроз ЗМС СН. Приведён анализ полученных результатов математического моделирования и приведены рекомендации их практического применения.

В предложенной работе показывается, что применение методов и алгоритмов, основанных на интеллектуальных способах обработки данных совместно с технологией интеллектуальных агентов, позволяет существенно повысить оперативность управления ИБ ЗМС СН. Определены направления дальнейших исследований в данной предметной области.

SUMMARY

Ageev S.A. **Application of Intelligent Methods of Information Representation to Manage Information Security Risks in Secure Multiservice Special-Purpose Networks.**

The paper offers intellectual data analysis methods used for risk management of IT threats to the elements of secure multiservice special-purpose networks (SMSPN). It is proved that process speed in multiservice special-purpose networks, their manifold, inaccuracy and insufficiency, as well as high dimension of a priori data about network elements states and occurring in them processes – all create the need to use intelligent methods of data processing and control. The analysis of the factors influencing the cycle time of information security management of secure multiservice special-purpose networks is conducted. The unified metric for evaluating threats to information security of elements of SMSPN is offered.

The presented results of research show a developed unified mathematical model of procedures for clustering, classification and ranking of threats to information security, which may possess a previously unknown structure. Quality criteria of functioning of the proposed procedures are determined. The results of the study of the developed mathematical model are shown. Rules for assessing the degree of threats to secure multiservice special-purpose networks are formulated. The analysis of the obtained results of mathematical modeling and recommendations for their practical application are given.

In the proposed paper we show that the use of methods and algorithms based on intelligent data processing methods, together with the technology of intelligent agents, can significantly increase the efficiency of IT security management of secure multiservice special-purpose networks. Directions for further research in this subject area are determined.