DOI: 10.18372/2410-7840.20.12449
УДК 004.056

СОВЕРШЕНСТВОВАНИЕ КИБЕРЗАЩИТЫ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМ ТРАНСПОРТА ЗА СЧЕТ МИНИМИЗАЦИИ ОБУЧАЮЩИХ ВЫБОРОК В СИСТЕМАХ ВЫЯВЛЕНИЯ ВТОРЖЕНИЙ

Берик Ахметов

Последние десятилетия ознаменовались бурным ростом темпов использования информационных технологий в разнообразных областях цивилизационных практик. Это лишь подтверждает курс на цифровизацию бизнес процессов, в частности в транспортной отрасли Республики Казахстан. В условиях роста количества дестабилизирующих воздействий на информационно-коммуникационные системы, в том числе транспортные, с задействованием самых разных цифровых технологий, необходимы дальнейшие научные изыскания, направленные на развитие теоретических и методологических основ синтеза интелектуализированных, способных к самообучению систем выявления вторжений. Показано, что процесс киберзащиты для информационно-коммуникационных систем, в частности транспортных, контролируется и анализируется по значениям нескольких параметров признаков аномалий, кибератак и угроз. Предложены дополнения к используемым в системах обнаружения вторжений методам выбора информативных признаков для обучающих выборок. Показано, что наиболее простыми, и в тоже время эффективными, с точки зрения аппаратно-программной реализации в подобных системах являются методы-фильтры. Это позволит более результативно выполнять ранжирование информативных признаков для аномалий, атак и угроз в интелектуализированных системах кибербезопасности.

Ключевые слова: *кибербезопасность, защита информации, распознавания угроз, аномалии, минимизация признаков, методы фильтры, информационно-коммуникационные системы.*

Введение. Задача классификации конкретных типов аномалий, атак и киберугроз, является одной из главных составляющих процесса распознавания деструктивного вмешательства или несанкционированного доступа (НСД) в работу информационно-коммуникационных систем (ИКС) [1] (далее ИКС транспорта – ИКСТ). При этом, чем сложнее архитектура ИКСТ тем сложнее обобщить и на этой основе разработать методологию процедур качественного распознавания объектов (аномалий, кибератак, угроз – далее объектов распознавания (ОР или ОРааy)). В бизнес-процессах транспортной отрасли используется достаточно большое количество самых разнообразных информационных технологий и систем [2, 3]. В настоящее время в системах распознавания (или выявления) вторжений (СВВ) активно применяются методы классификации ОР на основе теории обучения машин (ТОМ) [4]. Заметим, что на этапе синтеза обучающих выборок [5, 6] для СВВ, основополагающей стадией распознавания и классификации ОР, является использование минимизированного набора признаков (атрибутов [4, 5]) ОРааy. Минимизация набора информативных признаков для распознавания ОРааy, необходима для увеличения быстродействия СВВ [7]. Вполне очевидно, что используя алгоритм простого перебора признаков, мы сжижаем быстродействие СВВ. Далее полагаем, что под признаками ОРааy будут трактоваться результаты измерений априори известных характеристик объекта.

Таким образом, одним из перспективных и актуальных направлений исследований систем интеллектуального выявления вторжений (СИВВ) становиться решение задачи оптимизации набора признаков ОР. Это позволит не только сократить размерность пространства признаков ОРааy, но и исключить атрибуты, которые могут повлиять на точность распознавания.

Анализ литературных данных и постановка проблемы. Проблематике синтеза СИВВ, и в частности, подсистем обнаружения аномалий, кибератак, угроз (ОРааy) посвящено достаточно большое количество публикаций. В силу того, что данное научное направление активно развивается, остановимся лишь на некоторых публикациях последних. В [8, 9] выполнен анализ существующих методов распознавания (обнаружения) угроз, кибератак и аномалий в ИКС, в частности для ИКСТ [1, 10, 11]. В [12, 13] рассмотрены существующие подходы в задачах классификации методов обнаружения, и, в частности, базирующиеся на методах ТОМ. В [14, 15] рассмотрены различные технологии обучения СВВ. Практически всеми авторами отмечается, что типичным недостатком большинства СВВ являются ошибочные срабатывания [16, 17]. Это, не в последнюю очередь, связано с отсутствием в СВВ реализованных алгоритмов, предотвращающих использование в обучающих и тестовых выборках избыточности признаков в методах-фильтрах [18, 19].

Таким образом, необходимо на стадии обучения СВВ минимизировать количество информативных атрибутов ОРааy по различным критериям.

Постановка цели и задач исследования. Цель исследования – разработка модели, которая выявляет оптимальные наборы атрибутов для разнотипных ОР в системах выявления вторжений в ИКС.

Оптимизационная модель для сокращения наборов признаков разнотипных ОР в интелектуализированных системах выявления вторжений.

Приняты следующие обозначения параметров в модели: MI – общее число ОРааy в СИВВ (ОР – аномалии, атаки, угрозы); PA – число возможных целей злоумышленников (внешних и внутренних) в процессе нападения на ИКС; B_{p_a} – множество номеров киберугроз, реализуемых злоумышленником при достижении $p_a - \ddot{u}$ цели (в ходе кибератаки).

В общем случае, проблематика распознавания ОРааy для ИКСТ, заключается в следующем [20, 21]. Исследуем множество PA, которое характеризуются совокупностью атрибутов $\{\mathrm{p}_{\mathrm{ax1}},...,\mathrm{p}_{\mathrm{axn}}\}$. Дополнительно полагаем, с учетом [5, 6], что множество PA представлено в виде объединения подмножеств ОРааy для ИКСТ – $(B_{p_{a1}},...,B_{p_{al}})$. Подмножества можно трактовать как классы ОРааy $(KL_1,...,KL_l)$. Сделано допущение, что классы не пересекаются. В СИВВ имеется конечная группа объектов $\{\mathrm{sp}_{\mathrm{a1}},...,\mathrm{sp}_{\mathrm{am}}\}$ из PA. Про объекты известно, к каким классам ОРааy их можно отнести. Таким образом, это т.н. прецеденты. Прецедент – это объект для обучения СИВВ (ОИ) [1, 5, 6]. СИВВ необходимо по имеющемуся множеству атрибутов, т. е. используя только описания объекта sp_{an} из PA, типизировать ОРааy. А затем выбрать метод или средство противодействия. Т.е. по итогам классификации конфигурируют работу систем киберзащиты. СИВВ первоначально неизвестно к какому классу, подклассу или типу относится ОРааy.

Результаты исследования, изложенного в статье, являются продолжением работ [5, 6, 22] выполненных автором ранее, в том числе в соавторстве с коллегами из Украины и Казахстана.

Проанализируем случай, когда ОРааy из исследуемого множества PA описываются показателями, каждый из которых принимает значения из множества $\{0,1,..,k_{p_a}-1\}$. Это сделано для того чтобы представить информацию в СИВВ в двоичной форме, что удобно для схемотехнической или программной реализации СИВВ.

Введем следующие обозначения:

– $\mathrm{NP}_{\mathrm{p_a}}$ – конечная подборка из r_{p_a} разных атрибутов вида $\mathrm{NP}_{\mathrm{p_a}} = \{p_{ax_{j1}},..., p_{ax_{jr}}\}$, $\sigma_{\mathrm{DOP}} = (\sigma_{\mathrm{DOP_1}}, ..., \sigma_{\mathrm{DOP_r}})$;

– $\sigma_{\mathrm{DOP_i}}$ – допустимое значение показателя ОР $p_{ax_i}, i = 1,2,...,r_{p_a}$.

Заметим, что набор $\sigma_{\mathrm{DOP_i}}$ это т.н. элементарный классификатор (ЭК) [1]. ЭК может быть получен по показателям из $\mathrm{NP}_{\mathrm{p_a}}$;

$\mathrm{PA}_{\mathrm{mn}}^{\mathrm{k_{Pa}}}$, $k_{p_a} \geq 2$ – множество всех матриц, характеризующих объекты анализа в СВВ (уязвимости, угрозы, шаблоны кибератак [1, 3]) с элементами $m \times n$ и $\{0,1,.., k_{p_a}-1\}$;

$E_k^{r_{pa}}, k_{p_a} \geq 2, r_{p_a} \leq n,$ множество всех $\mathrm{k}_{\mathrm{p_a}}$-ых наборов, имеющих длину r_{pa};

$\mathrm{Q_p}(\sigma_{DOP})$, $\sigma_{DOP} \in E_k^{r_{pa}}$, $\sigma_{DOP} = (\sigma_{\mathrm{DOP_1}},..., \sigma_{\mathrm{DOP_r}})$, $p \in \{1,2,...,r_{p_a}\}$, – множество всех наборов вида $\beta_1,...,\beta_{\mathrm{r}}$ в $E_k^{r_{pa}}$ для которых выполняются условия $\beta_p \neq \sigma_{DOP_p}$ и $\beta_j = \sigma_{DOP_j}$ при $j \in \{1,2,...,r_{pa}\}$;

$CU(\mathrm{LU}, \sigma_{\mathrm{DOP}})$ – множество всех пар вида (HU, σ_{DOP});

HU-σ_{DOP} – покрытие матрицы $LU \in PA_{mn}^k$;

$\mathrm{BU}(\mathrm{LU}, \sigma_{\mathrm{DOP}})$ – множество всех пар вида (HU, σ_{DOP});

$SU(\mathrm{LU}, \sigma_{\mathrm{DOP}})$ – совокупность всех σ_{DOP} – подматриц матрицы LU.

Также принимаем, что:

$$CU(LU) = \bigcup_{r_{pa}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} CU(LU, \sigma_{DOP}), \quad (1)$$

$$BU(LU) = \bigcup_{r_{pa}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} BU(LU, \sigma_{DOP}), \quad (2)$$

$$SU(LU) = \bigcup_{r_{pa}=1}^{n=MI} \bigcup_{\sigma_{DOP} \in E_k^r} CU(LU, \sigma_{DOP}). \quad (3)$$

Как было показано в [1, 2, 5, 6, 22], существует устойчивый тренд на увеличение количества типов ОРааy. С точки увеличения результативности распознавания ОРааy, наибольший интерес в задачах синтеза алгоритмов распознавания для СИВВ представляет исследование асимптотики значений параметров |CU(LU)|и |SU(LU)|. В частности, необходимо изучить ситуацию, связанную с утверждением, что почти всех матриц LU из $\mathrm{PA}^{\mathrm{k}}_{\mathrm{un}}$ при $n = MI \to \infty$ выполняется свойство β. Следовательно, доля тех матриц из $\mathrm{PA}^{\mathrm{k}}_{\mathrm{mn}}$ для которых с ε-приемлемой точностью выполняется свойство β, стремиться к единице (1) и, одновременно, $\varepsilon \to 0$ при $n = MI \to \infty$.

Положим, что $\mathrm{PA}^{\mathrm{k}}_{\mathrm{mn}} = \{\mathrm{LU}\}$ пространство элементарных событий в СИВВ. При этом каждое событие LU происходит с вероятностью $1/|\mathrm{PA}^{\mathrm{k}}_{\mathrm{mn}}|$. Математическое ожидание случайной величины $\hbar(\mathrm{LU})$ обозначать как $M\hbar(LU)$, а дисперсию – $D\hbar(LU)$.

Проанализируем, сколькими способами можно построить матрицу из $\mathrm{PA} = \mathrm{PA}_{(\mathrm{n}\nu_1, m\omega_1, \sigma_{\mathrm{DOP1}})} \cap \mathrm{PA}_{(\mathrm{n}\nu_2, m\omega_2, \sigma_{\mathrm{DOP_2}})}$.

Этап 1. Выбираем те элементы, которые располагаются на пересечении строк с номерами из $\mathrm{n}\nu_1$ и, столбцов с номерами из $m\omega_1$. Это задание можно выполнить $(\mathrm{k_{p_a}}\text{-}1)^{r_{p_a}}$ способами.

Этап 2. Выбираем только те элементы матрицы, которые располагаются на пересечении строк с номерами из $n\nu_2$, и столбцов с номерами из $m\omega_2$. Необходимо учесть, что *ab* из них расположены одновременно на пересечении строк с номерами из $n\nu_1$ и столбцов с номерами из $m\omega_2$, ($(\mathrm{k_{p_a}}\text{-}1)^{l-a}$ – способов).

Этап 3. Произвольно доопределив строки матрицы с номерами из $\mathrm{n}\nu_1 \cup n\nu_2$ ($\mathrm{k_{p_a}}^{(r_{p_a}+l-a)MI+ab-r^2_{p_a}-l^2}$ – способов), выбираем оставшиеся строки ($\mathrm{k_{p_a}}^{\mathrm{PA\cdot MI}-(r_{p_a}+l-a)MI}$ – способов).

Этап 4. Полагаем, что для всех матриц, характеризующих объекты анализа в СИВВ (атрибуты ОР) $\mathrm{LU} \in \mathrm{PA}^{\mathrm{k}}_{\mathrm{mn}}$ при $n = MI \to \infty$ справедливо отношение:

$$|SU_1(LU)| = 0.$$

Этап 5. На основании лемм, рассмотренных в работах [1, 7, 10], выполнен анализ следующих ситуаций.

5.1. Если $\mathrm{n}^{\alpha} \le m \le k^{n^{\beta}}$, $\alpha > 1, \beta < 1$, то имеет место

$$PA\eta_1(LU) \approx PA\eta_2(LU) \approx$$
$$\approx \sum_{r_{p_a} \in \hat{\lambda}_1} CU_m^{r_{p_a}} \cdot CU_n^{r_{p_a}} \cdot r_{p_a}(k_{p_a}-1)^{r_{p_a}} \cdot k_{p_a}^{r_{p_a}-r^2_{p_a}}, \quad (4)$$
$$\text{при } n = MI \to \infty;$$

5.2. Если $\mathrm{n}^{\alpha} \le m \le k^{n^{\beta}}$, $\alpha > 1, \beta < 1$, то имеет место

$$\frac{\mathrm{D}\hbar\eta_2(LU)}{PA(\eta_2(LU))^2} \to 0 \text{ при } n = MI \to \infty, \quad (5)$$

где $\eta_{(\mathrm{nv},\mathrm{m}\omega)}(LU, \sigma_{DOP})$ – случайная величина равная 1, если $LU \in PA_{(nv,m\omega,\sigma_{DOP})}$ и 0 в противном случае, получим

$$PA\eta_3(LU) = \sum_{r_{p_a} \ge r_{pa_1}} \sum_{\substack{nv \in V_r^m \\ m\omega \in W_r^n}} \sum_{\sigma_{DOP} \in E_k^r} P(\eta_{(nv,m\omega)}(LU, \sigma_{DOP}) = 1), \quad (6)$$

где $\mathrm{nv} \in \mathrm{V}^{\mathrm{m}}_{\mathrm{r_{pa}}}$, $mw \in W^n_{\mathrm{r_{pa}}}$, $\sigma_{DOP} \in E_k^{r_{p_a}}$; $P(\eta_{(nv,m\omega)}(LU, \sigma_{DOP}) = 1)$ – вероятность того, что $\eta_{(\mathrm{nv},m\omega)}(LU, \sigma_{DOP}) = 1$.

Следовательно, в силу, $|PA_{mv,m\omega,\sigma_{DOP}}| = (k_{p_a} - 1)^{r_{p_a}} \cdot k_{p_a}^{mn-r^2_{p_a}}$, получим

$$PA\eta_3(LU) = \sum_{r_{p_a} \ge r_{p_a1}} CU_{n=MI}^{r_{p_a}} \cdot CU_m^{r_{p_a}} \cdot r_{p_a}!(k_{p_a}-1)^{r_{p_a}} \cdot k_{p_a}^{r_{p_a}-r^2_{p_a}}. \quad (7)$$

Так как при $r_{p_a} \ge r_{p_a}1$:

$$\mathrm{CU}^{\mathrm{r_{p_a}}}_{\mathrm{n=MI}} \cdot \mathrm{CU}^{\mathrm{r_{p_a}}}_{\mathrm{m}} \cdot r_{p_p}!(k_{p_a}-1)^{r_{p_a}} \cdot k_{p_a}^{r_{p_a}-r^2_{p_a}} \le \frac{(mn)^{r_{p_a}}}{r_{p_p}!} \cdot r^{2\cdot r_{p_a}-r^2_{p_a}} \le (\frac{k^2{}_{p_a} \cdot e}{r_{p_a}})^{r_{p_a}}, \quad (8)$$

то при достаточно большом количестве объектов (киберугроз, уязвимостей, кибератак) описанных с помощью своих «матриц показателей» и $n = MI \to \infty$ будем иметь

$$\sum_{\mathrm{r_{p_a}} \ge \mathrm{r_{p_a}1}}^{n=MI} \mathrm{CU}^{\mathrm{r_{p_a}}}_{\mathrm{n=MI}} \cdot \mathrm{CU}^{\mathrm{r_{p_a}}}_{\mathrm{m}} \cdot r_{p_p}!(k_{p_a}-1)^{r_{p_a}} \cdot k_{p_a}^{r_{p_a}-r^2_{p_a}} \le n(\frac{k^2{}_{p_a} \cdot e}{\log_{k_{p_a}} nm})^{\log_{k_{p_a}} nm} \to 0. \quad (9)$$

Таким образом, $|BU(LU)| \le CU_{n=MI}^{[\log_{k_{Pa}} nm]} mn \le \frac{n^{\log_{k_{Pa}} mn} \cdot mn}{[\log_{k_{Pa}} mn]!}$.

Также имеем

$$|\mathrm{SU(LU)}| \ge \sum_{r_{Pa} \in \lambda_1} \frac{(nm)^{r_{Pa}-1}}{r_{p_a}!}(1-\frac{r_{p_a}}{n})^{r_{Pa}} \cdot (1-\frac{r_{p_a}}{m})^{r_{Pa}} \ge \sum_{r_{Pa} \in \lambda_1} \frac{(nm)^{r_{Pa}-1}}{[\log_k mn]!}, \quad (10)$$

где $r_{p_{a1}} = \frac{1}{2} \cdot \log_{k_{Pa}} mn - \frac{1}{2} \cdot \log_{k_{Pa}} \log_{k_{Pa}} mn - \log_{k_{Pa}} \log_{k_{Pa}} \log_{k_{Pa}} n$; λ_1 – интервал.

Отсюда имеем

$$\frac{|\mathrm{SU(LU)}|}{|\mathrm{BU(LU)}|} \ge \frac{(mn)^{r_{Pa1}-2}}{n^{\log_{k_{Pa}} mn}} \ge n^{(\alpha+1)(r_{Pq1}-2)-\log_{k_{Pa}} mn} \to \infty, \text{ при } n = MI \to \infty. \quad (11)$$

Соответственно, когда количество строк матрицы, содержащей типовые показатели объектов (уязвимости, аномалии, кибератаки и т.п.), по порядку имеется большее количество столбцов, величина $|SU(LU)|$ по порядку больше величины $|BU(LU)|$.

Таким образом, изложенные выкладки подтверждают предположение о том, что дополнения к методам фильтров в задачах минимизации обучающих выборок в СИВВ, позволит более результативно выполнять ранжирование информативных атрибутов OPaay, см. рис. Аналитически подтверждено, что методы-фильтры, позволяют результативно реализовывать оценивание информативности для подмножеств. В частности, для сокращения малоинформативных признаков, анализ которых затрудняет работу СИВВ.

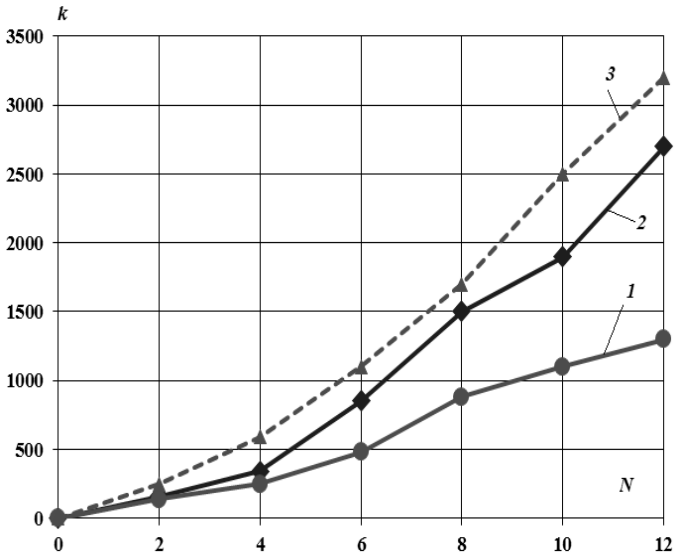
Перспективы дальнейших исследований в области синтеза СИВВ для ИКСТ состоят в том, чтобы усовершенствовать базу знаний системы, а также провести исследования на большем количестве классов кибератак на ИКСТ.

Выводы. В результате выполненных исследований:

– показано, что процесс киберзащиты ИКС, в частности транспортных, контролируется и анализируется по значениям нескольких параметров признаков аномалий, кибератак и угроз;

– рассмотрены дополнения к используемым в системах обнаружения вторжений методам выбора информативных атрибутов для обучающих выборок. Показано, что наиболее простым путем реализации СИВВ, и в тоже время эффективным, с точки зрения аппаратно-программной реализации в подобных системах является применение методов-фильтров;

– предложены дополнения к методам фильтров в задаче минимизации обучающих выборок в системах обнаружения аномалий, атак и угроз. Это позволит более результативно выполнять ранжирование информативных признаков для аномалий, атак и угроз в интелектуализированных системах кибербезопасности. Показано, что методы-фильтры, позволяют достаточно результативно выполнять оценку информативности для подмножеств, в частности для сокращения малоинформативных признаков, анализ которых затрудняет работу систем обнаружения и классификации аномалий, кибератак и угроз.



(***N*** – количество признаков в наборе для обучения СИВВ; ***k*** – количество шагов обучения СИВВ с допускаемой точностью выявления атаки не менее 80 %)

1 – предлагаемая модель; 2 – методы прогнозирования состояний; 3 – последовательный перебор элементарных классификаторов и соответствующих признаков

Рис. Оптимизация наборов признаков в процессе распознавания кибератак класса «Отказ в обслуживании»

ЛИТЕРАТУРА

[1]. A. Petrov, V. Lakhno, A. Korchenko, "Models, methods and information technologies of protection of corporate systems of transport based on intellectual identification of threats", *Decision Making in Manufacturing and Services*, vol. 9, no. 2, pp. 117-135, 2016.

[2]. В. Лахно, "Інформаційна безпека інтелектуальних транспортних систем", *Захист інформації*, Т. 17, № 4, С. 298-305, 2015.

[3]. M. M. Al Hadidi, Y. K. Ibrahim, V. Lakhno, A. Korchenko, A. Tereshchuk, A. Pereverzev, "Intelligent Systems for Monitoring and Recognition of Cyber Attacks on Information and Communication Systems of Transport", *International Review on Computers and Software (IRECOS),* vol. 11, no. 12, pp. 1167-1177, 2016.

[4]. R. Abidar, K. Moummadi, F. Moutaouakkil, H. Medromi, "Intelligent and Pervasive Supervising Platform for Information System Security Based on Multi-Agent Systems", *International review on computers and software*, Vol. 10, Issue 1, pp. 44-51, 2015.

[5]. Г. Бекетова, Б. Ахметов, О. Корченко, В. Лахно, "Розробка моделі інтелектуального розпізнавання аномалій і кібератак з використанням логічних процедур, які базуються на покриттях матриць ознак", *Безпека інформації,* Т. 22, №. 3, С. 242-254, 2016.

[6]. G. Beketova, G. B. Akhmetov, A. Korchenko, V. Lakhno, A. Tereshuk., "Cyber intelligence systems based on adaptive regression splines and logical procedures of attack recognition", *Computer modelling and new technologies*, vol. 21, no. 2, pp. 7-16, 2017.

[7]. A. A. El Hassani, A. A. El Kalam, A. Bouhoula, R. Abassi, A. A. Ouahman, "Integrity-OrBAC: a new model to preserve Critical Infrastructures integrity", *International Journal of Information Security*, vol. 14, issue 4, pp. 367-385, 2015.

[8]. Р. Грищук, "Атаки на інформацію в інформаційно-комунікаційних системах", *Сучасна спеціальна техніка*, no. 1 (24), С. 61-66, 2011.

[9]. Petrov, B. Borowik, M. Karpinskyy, *Immune and defensive corporate systems with intellectual identification of threats*, Pszczyna: Śląska Oficyna Drukarska, 2016, 222 p, ISBN: 978-83-62674-68-8.

[10]. V. Lahno, "Ensuring of information processes' reliability and security in critical application data processing systems", *MEST Journal*, vol. 2, issue 1, pp. 71-79, 2014.

[11]. N. Manap, S. Basir, S. Hussein, P. Tehrani, A. Rouhani, "A. Legal Issues of Data Protection in Cloud Computing", *International Journal of Soft Computing,* vol. 8, issue 5, pp. 371-376, 2013.

[12]. J.A. George, M. Hemalatha, "Improving Authentication and Authorization for Identity Based Cloud Environment Using OAUTH with Fuzzy Based Blowfish Algorithm", *International review on computers and software*, vol. 10, issue 7, pp. 783-788, 2015.

[13]. H.-H. Li, C.-L. Wu, "Study of Network Access Control System Featuring Collaboratively Interacting Network Security Components", *International review on computers and software*, vol. 8, issue 2, pp. 527-532, 2013.

[14]. R. Geetha, E. Kannan, Secure "Communication Against Framing Attack in Wireless Sensor Network", *International review on computers and software,* vol. 10, issue 4, pp. 393-398, 2015.

[15]. S. Shamshirband, N. B. Anuar, M. L. Kiah, A. Patel, "An appraisal and design of a multiagent system based cooperative wireless intrusion detection computational intelligence technique", *Engineering Applications of Artificial Intelligence*, vol. 26, issue 9, pp. 2105-2127, 2013.

[16]. L. Keunsoo, J. Kim, K. Hoon Kwon, Y. Han, S. Kim, "DDoS attack detection method using cluster analysis", *Expert Systems with Applications,* vol. 4, issue 3, pp. 1659-1665, 2008.

[17]. S. Dilek, H. Çakır, M. Aydın, "Applications of artificial intelligence techniques to combating cybercrimes: A review", *International Journal of Artificial Intelligence & Applications*, vol. 6, issue 1, pp. 21-39, 2015.

[18]. A. Patel, M. Taghavi, K. Bakhtiyari, J. Celestino Junior, "An intrusion detection and prevention system in cloud computing: A systematic review", *Journal of Network and Computer Applications,* vol. 36, issue 1, pp. 25-41, 2013.

[19]. D. K. Barman, G. Khataniar, "Design of Intrusion Detection System Based On Artificial Neural Network and Application of Rough Set", *International Journal of Computer Science and Communication Networks,* vol. 2, issue 4, pp. 548-552, 2012.

[20]. J. Raiyn, "A survey of Cyber Attack Detection Strategies", *International Journal of Security and Its Applications,* vol. 8, issue 1, pp. 247-256, 2014.

[21]. S. Mukkamala, A.H. Sung, A. Abraham, V. Ramos, "Intrusion detection systems using adaptive regression splines", *Sixth International Conference on Enterprise Information Systems,* part 3, pp. 211-218, 2006.

[22]. B. Akhmetov, "Designing a decision support system for the weakly formalized problems in the provision of cybersecurity", *Eastern-European Journal of Enterprise Technologies,* vol. 1, no. 2 (85), pp. 4-15, 2017.

ВДОСКОНАЛЕННЯ КІБЕРЗАХИСТУ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМ ТРАНСПОРТУ ШЛЯХОМ МІНІМІЗАЦІЇ НАВЧАЛЬНИХ ВИБІРОК В СИСТЕМАХ ВИЯВЛЕННЯ ВТОРГНЕНЬ

Останні десятиліття ознаменувалися бурхливим зростанням темпів використання інформаційних технологій у різноманітних областях цивілізаційних практик. Це лише підтверджує курс на цифровізацію бізнес процесів, зокрема у транспортній галузі Республіки Казахстан. Тенденція на цифровізацію економіки та інтеграція в бізнес процеси різних інформаційно-комунікаційних систем, зобов'язує враховувати попутні ризики. В першу чергу, ту їх частину, яка пов'язана із захистом інформації та кібербезпекою цифрових си-

стем. У завданнях кіберзахисту все частіше застосовують когнітивні технології виявлення й розпізнавання вторгнень. В умовах зростання кількості впливів що дестабілізують інформаційно-комунікаційні системи, зокрема транспортні, із застосуванням різних цифрових технологій, необхідні подальші наукові дослідження, спрямовані на розвиток теоретичних і методологічних основ синтезу інтелектуалізованих, здатних до самонавчання систем виявлення вторгнень. Показано, що процес кіберзахисту для інформаційно-комунікаційних систем, зокрема транспортних, контролюється та аналізується за значеннями багатьох параметрів ознак аномалій, кібератак та загроз. В роботі розглянуті доповнення до поширених у системах виявлення вторгнень методам вибору інформативних ознак для навчальних вибірок. Показано, що найбільш простими, й водночас ефективними, з точки зору апаратно-програмної реалізації у подібних системах є методи-фільтри. Запропоновано доповнення до методів фільтрів в задачах мінімізації навчальних вибірок в системах виявлення аномалій, атак і загроз. Це дозволить більш результативно виконувати ранжування інформативних ознак для аномалій, атак й загроз в інтелектуальних системах кібербезпеки. Показано, що методи-фільтри, дозволяють досить результативно виконувати оцінку інформативності для підмножин ознак, зокрема для скорочення малоінформативних ознак, аналіз яких ускладнює роботу систем виявлення та класифікації аномалій, кібератак і загроз.
Ключові слова: інформаційно-комунікаційна система, кібербезпека, захист інформації, розпізнавання загроз, аномалії, мінімізація ознак, методи фільтри.

IMPROVEMENT OF CYBERSPACE OF INFORMATION AND COMMUNICATION SYSTEMS OF TRANSPORT FOR THE ACCOUNT OF MINIMIZATION OF TRAINING SELECTIONS IN SYSTEMS OF INVESTIGATION OF INVASIONS

The last decades were marked by a rapid growth in using of information technologies in various areas of civilizational practices. This confirms the course on digitalization of business processes, particular in the transport industry of the Republic of Kazakhstan. The tendency to digitalization of the economy and integration into business processes of various information and communication systems, obliges to take into account emerging incidental risks. First of all, the part, which is connected with the protection of information and the cybersecurity of digital systems. In the tasks of cyber defense, cognitive technologies for detecting and recognizing intrusions are increasingly being used. In the conditions of growing number of destabilizing influences on information and communication systems, including transport, involving a variety of digital technologies, further scientific research is needed to develop the theoretical and methodological foundations for the synthesis of intelligent, self-taught intrusion detection systems. It is shown that the process of cyber defense for information and communication systems, in particular transport, is controlled and analyzed according to the values of several parameters of the signs of anomalies, cyber-attacks and threats. There are considered additions to methods of selecting informative features for training samples used in intrusion detection systems in this work. It is shown that the most simple, and at the same time effective, from the point of view of hardware and software implementation in such systems are filter methods. Additions are proposed to filter methods in the tasks of minimizing training samples in systems for detecting anomalies, attacks and threats. It is shown that the most simple, and at the same time effective, from the point of view of hardware and software implementation in such systems are filter methods. It is shown that the filtering methods allow to perform the estimation of informativeness for the subset of characteristics sufficiently, in particular to reduce the low-information characteristics, the analysis of which makes the detection and classification of anomalies, cyber-attacks and threats difficult.
Keywords: information and communication system, cybersecurity, information protection, threat recognition, anomalies, minimization of signs, filter methods.

Ахметов Берик Бахытжанович, кандидат технических наук, доцент, ректор Каспийского государственного университета технологий и инжиниринга имени Ш. Есенова.
E-mail: 007berik@mail.ru.
Ахметов Берік Бахитжанович, к.т.н., доцент, ректор Каспійського державного університету технологій та інжинірингу ім. Ш. Есенова.
Akhmetov Berik, PhD, Assistant professor, Rector of the Caspian State University of Technologies and Engineering named after Sh. Yessenov.