

Падалка І.О. *Льотна академія Національного авіаційного університету, Кропивницький*
Опенько П.В. *Національний університет оборони України імені Івана Черняхівського, Київ*
Руденко В.М. *Донбаська державна машинобудівна академія, Краматорськ*
Столяренко М.П. *Науково-дослідний центр ракетних військ і артилерії, Суми*

МЕТОД ВИЯВЛЕННЯ АНОМАЛЬНИХ ПОСЛІДОВНОСТЕЙ В ДІАГНОСТИЧНИХ ДАНИХ ТЕХНОЛОГІЧНОГО ОБЛАДНАННЯ ПОВІТРЯНОГО СУДНА ДЛЯ ПОПЕРЕДЖЕННЯ ОСОБЛИВИХ ВИПАДКІВ В ПОЛЬОТІ

Анотація: Прогнозування особливих випадків в польоті є основним завданням параметричного діагностування технологічного обладнання повітряного судна. Для вирішення цього завдання бортові засоби автоматизованого контролю, діагностики та управління бортовим обладнанням, розвантаження та інформаційної підтримки екіпажу використовують математичні моделі, засновані на тренд-аналізі деяких реєстрованих параметрів з напрацювання. Разом із тим, існуючі діагностичні моделі, що базуються на відповідних математичних моделях, не завжди дозволяють прогнозувати виникнення відмов технологічного обладнання. Це підтверджує, що завдання прогнозування особливих випадків в польоті є актуальним. У роботі запропоновано метод прогнозування особливих випадків в польоті на основі виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна. Для виявлення зазначених аномальних послідовностей пропонується використовувати гібридну стохастичну модель, засновану на об'єднанні марковських і продукційних моделей, що використовують темпоральні правила для уточнення перехідних ймовірностей між станами процесу. За рахунок включення в модель уточнюючих продукційних правил підвищується достовірність опису випадкових процесів, які не являються марковськими. Також з'являється можливість інтеграції в модель апріорних експертних знань, що є дуже важливим для прогнозування особливих випадків в польоті. Застосування запропонованого методу дозволить впровадити прогностичний принцип управління безпекою польотів, а також отримати економічний ефект від запобігання простою повітряного судна через раптові відмови обладнання.

Ключові слова: безпека польотів, особливі випадки в польоті, параметрична діагностика, прогнозування, аномальна послідовність, часовий ряд, темпоральний патерн.

Padalka I. *The Flight Academy of the National Aviation, Kropyvnytsky*
Open'ko P. *Ivan Cherniakhovskyi National Defense University of Ukraine, Kyiv*
Rudenko V. *Donbass State Engineering Academy, Kramatorsk*
Stoliarenko M. *Research Center for Missile Forces and Artillery, Sumy*

A METHOD FOR DETECTING ABNORMAL SEQUENCES IN THE DIAGNOSTIC DATA OF TECHNOLOGICAL EQUIPMENT OF AN AIRCRAFT TO PREVENT SPECIAL CASES IN FLIGHT

Abstract: Prediction of special cases in flight is the main task of parametric diagnostics of aircraft technological equipment. To solve this problem, on-board means of automated control, diagnostics and control of on-board equipment, unloading and information support of the crew use mathematical models based on the trend analysis of some registered operating time parameters. However, existing diagnostic models based on the corresponding mathematical models do not always allow predicting the occurrence of

process equipment failures. What makes the task of forecasting special cases in flight relevant. The paper proposes a method for predicting special cases in flight based on identifying abnormal sequences in the diagnostic data of aircraft technological equipment. To identify abnormal sequences, it is proposed to use a hybrid stochastic model based on a combination of Markov and production models using temporal rules to specify transition probabilities between process states. By including more precise production rules in the model, the probability of describing random processes that are not Markovian is increased, and it becomes possible to integrate a priori expert knowledge into the model, which is very important for predicting special cases in flight. The application of the proposed method will allow to implement the prognostic principle of flight safety management, as well as to obtain the economic effect of preventing aircraft downtime due to sudden equipment failure.

Keywords: *flight safety, special cases in flight, parametric diagnostics, forecasting, abnormal sequence, time series, temporal pattern.*

Падалка И.О. *Летная академия Национального авиационного университета, Кропивницкий*
Опенько П.В. *Национальный университет обороны Украины имени Ивана Черняховского, Киев*
Руденко В.Н. *Донбасская государственная машиностроительная академия, Краматорск*
Столяренко Н.П. *Научно-исследовательский центр ракетных войск и артиллерии, Сумы*

МЕТОД ВЫЯВЛЕНИЯ АНОМАЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ДИАГНОСТИЧЕСКИХ ДАННЫХ ТЕХНОЛОГИЧЕСКОГО ОБОРУДОВАНИЯ ВОЗДУШНОГО СУДНА ДЛЯ ПРЕДУПРЕЖДЕНИЯ ОСОБЫХ СЛУЧАЕВ В ПОЛЕТЕ

Аннотация: *Прогнозирование особых случаев в полете является основной задачей параметрической диагностики технологического оборудования воздушного судна. Для решения этой задачи бортовые средства автоматизированного контроля, диагностики и управления бортовым оборудованием, разгрузки и информационной поддержки экипажа используют математические модели, основанные на тренд-анализе некоторых регистрируемых параметров по наработке. Вместе с тем, существующие диагностические модели, основанные на соответствующих математических моделях, не всегда позволяют прогнозировать возникновение отказов технологического оборудования. Это подтверждает, что задача прогнозирования особых случаев в полете является актуальной. В работе предложен метод прогнозирования особых случаев в полете на основе выявления аномальных последовательностей в диагностических данных технологического оборудования воздушного судна. Для выявления аномальных последовательностей предлагается использовать гибридную стохастическую модель, основанную на объединении марковских и продукционных моделей, использующих темпоральные правила для уточнения переходных вероятностей между состояниями процесса. За счет включения в модель уточняющих продукционных правил повышается достоверность описания случайных процессов, которые не являются марковскими, а также появляется возможность интеграции в модель априорных экспертных знаний, что очень важно для прогнозирования особых случаев в полете. Применение предложенного метода позволит внедрить прогностический принцип управления безопасностью полетов, а также получить экономический эффект от предотвращения простоя воздушного судна вследствие внезапного отказа оборудования.*

Ключевые слова: *безопасность полетов, особые случаи в полете, параметрическая диагностика, прогнозирование, аномальная последовательность, временной ряд, темпоральный паттерн.*

1. Вступ

Ефективність застосування авіаційної техніки нерозривно пов'язана з проблемою безпеки польотів, успішне вирішення якої значною мірою визначає перспективи розвитку як цивільної,

так і державної авіації. Аналіз досвіду експлуатації авіаційної техніки свідчить про те, що однією з основних причин негативних наслідків особливих випадків в польоті є неправильні або несвоєчасно прийняті рішення, тобто "людський фактор". Для прийняття ефективних рішень, необхідне забезпечення повною, достовірною та якісною інформацією про стан елементів і всіх систем в цілому, а також прогнозування їх технічного стану, що є основним завданням параметричного діагностування технологічного обладнання повітряного судна. Для вирішення цього завдання бортові засоби автоматизованого контролю, діагностики та управління бортовим обладнанням, розвантаження та інформаційної підтримки екіпажу використовують математичні моделі, засновані на тренд-аналізі деяких реєстрованих параметрів з напрацювання. Однак існуючі діагностичні моделі, що базуються на відповідних математичних моделях, не завжди дозволяють прогнозувати виникнення відмов технологічного обладнання. Що робить актуальною задачу прогнозування особливих випадків в польоті на основі виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для підвищення оперативності прийняття рішення екіпажем, щодо виявлення, розпізнання та недопущення негативних наслідків особливих випадків в польоті.

Вихідна інформація про поведінку динамічного процесу зазвичай представляється у вигляді часових рядів (ЧР) даних, одержуваних від датчиків первинної інформації. Аномальні стани і їх передвісники представляються у вигляді відрізків часових рядів – аномальних темпоральних патернів (ТП), які потрібно виявляти в потоці даних. Прогнозування появи аномальних ТП в даних, що описують поведінку складних технологічних процесів, ускладнюється слабкою структурованістю первинної інформації, яка обумовлена тим, що вимірювання ведуться на різних режимах роботи та в різних умовах, при наявності шумів, перешкод, а також нестаціонарністю, нелінійністю і невідповідністю типовим законам розподілу процесів, що породжують первинну інформацію. Виділяють три основні проблеми в виявленні ТП особливого типу. По-перше, складність визначення типового профілю, що описує всі варіанти нормального протікання процесу. Друга проблема полягає у відсутності маркованих даних для побудови множин навчальних прикладів. По-третє, наявність невизначеностей, викликаних шумами і спотвореннями, які часто виявляються схожими на аномалії і відповідно важко піддаються виявленню [1].

2. Аналіз літературних даних і постановка проблеми

Існуючі підходи до виявлення аномалій можуть бути згруповані у дві категорії – "з учителем", або на основі класифікації, і "без вчителя". Підходи "з учителем" навчаються на маркованих прикладах, що належать до класів нормальних або аномальних ТП, і потім в процесі роботи здійснюється класифікація тестових зразків [2-5]. Методи детектування аномалій з категорії "без вчителя" будують темпоральний профіль у вигляді моделі, яка описує нормальний розвиток процесу, на основі заданої множини прикладів і потім перевіряють ймовірність того, що вхідний тестовий зразок був породжений саме цією моделлю. Дані методи працюють при допущенні, що більшість зразків в даних належать до нормального класу, в іншому випадку вони викликають велику кількість хибнопозитивних спрацьовувань [6, 7]. Так як методи "без вчителя" не вимагають наявності маркованих даних, в практичному плані вони є більш затребуваними, ніж підходи, засновані на навчанні "з учителем" [8]. Детектування аномалій з використанням технологій м'яких обчислень базується на методах нечіткої логіки [9], нейромережевих моделях [10] і штучних імунних системах [11, 12]. Нейромережеві моделі використовуються, як правило, при розробці алгоритмів детектування аномалій за схемою навчання "з учителем", нечіткі системи і імунологічні методи

використовуються в обох підходах. Також існують методи навчання "частково з учителем", що було викликано бажанням полегшити процес маркування даних [13]. На відміну від методів "з учителем" і "без вчителя" методи "частково з учителем" можуть використовувати частково марковані дані для підвищення ефективності роботи методів "без вчителя". Особливу групу методів в області детектування аномалій представляють стохастичні моделі, які використовуються в гібридному підході до виявлення аномалій в темпоральних даних. Серед стохастичних моделей слід виділити стохастичні автомати [8], марковські моделі (ММ) [13].

Постановка проблеми. В даній роботі пропонується підхід до розробки методу виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для попередження особливих випадків в польоті.

3. Мета дослідження

Метою дослідження є розробка методу прогнозування особливих випадків в польоті на основі виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна; з метою підвищення безпеки польотів.

4. Виклад основного матеріалу.

Розглянемо випадок, коли поведінка повітряного судна оцінюється на основі спостережень за значеннями одного конкретного інтегрального параметру. Послідовність значень такого параметру, отримана на основі показань датчиків за певний часовий інтервал, являє собою ЧР, аналіз якого дозволить судити про стан і зміну стану повітряного судна. Для представлення часового процесу функціонування технологічного обладнання повітряного судна використаємо модель на основі аналізу даних параметричної діагностики.

Гібридна стохастична модель (ГСМ) здатна адекватно представляти будь-які стохастичні процеси, в тому числі і процеси, які не є марковськими. ГСМ заснована на об'єднанні стохастичної ММ і продукційної моделі, представленою у вигляді системи продукційних темпоральних правил (ПТП), що уточнюють значення перехідних ймовірностей, які залежать від передісторії процесу. Система уточнюючих ПТП формується таким чином.

У загальному випадку ЧР TS – це впорядкована послідовність значень єдиного параметру $TS = \langle ts_1, ts_2, \dots, ts_i, \dots, ts_q \rangle$, що описує перебіг будь-якого тривалого процесу, де індекс i відповідає позначці часу, кінцевої довжини N . $p = (a, b, \dots, r)$ – ТП тривалості k . Підтримкою ТП p в X називається величина $Sup(p)$, що дорівнює числу входжень ТП p в X .

Таким чином умова марковості для кінцевого ЧР X запишеться у вигляді

$$\frac{Sup(b, c)}{Sup(b)} = \frac{Sup(a, \dots, b, c)}{Sup(a, \dots, b)} \quad (\forall p = (a, \dots, b, c) \in X | Sup(p) > \omega). \quad (1)$$

де ω – значення порога підтримки.

Для всіх ТП, що не задовольняють умові (1), формуються уточнюючі ПТП, що встановлюють для передкінцевих станів b ТП (a, \dots, b, c) нові значення перехідних ймовірностей $P(x / b) = P(x / a, \dots, b)$, обчислені на основі формули

$$P(c | a, \dots, b) = \frac{Sup(a, \dots, b, c)}{Sup(a, \dots, b)}. \quad (2)$$

Нехай $p = (a, \dots, b, c)$ – ТП в X довжиною k і $p^* = (\dots, b, c)$ – його частина (підпаттерн) довжиною $l < k$. Тоді ПТП $R(p)$, відповідне ТП p , називається домінуючим над ПТП $R(p^*)$, відповідним підпаттерну p^* .

ПТП, відповідні ТП у ЧР X , що мають достатній рівень підтримки, і не задовольняють умові (1), утворюють систему уточнюючих продукційних правил для ГСМ. Продукційні правила коректують значення перехідних ймовірностей вихідної ММ, що обчислені на основі формули

$$P(a, b) = \frac{Sup(a, b)}{Sup(a)}. \quad (3)$$

ГСМ представляється четвіркою: $ГСМ = \langle A, \|P\|, Q, P \rangle$, де A – множина станів; $\|P\|$ – матриця перехідних ймовірностей; $Q = [q_1, q_2, \dots, q_s]$ – вектор початкового розподілу ймовірностей; P – система ПТП [14].

Функціонування ГСМ в процесі надходження на її вхід потоку даних, представленого у вигляді ЧР X , здійснюється на основі наступного алгоритму.

1. Для поточного стану процесу $b = x(i)$ в потоці даних X формується множина ТП P , що включає δ попередніх даному стану ТП довжиною $1, 2, \dots, \delta$.

2. З множини уточнюючих продукційних правил P , що входять в ГСМ, формується підмножина правил P^* , відповідних ТП з P .

3. Якщо P^* виявляється порожнім, то перехідні ймовірності для поточного стану b визначаються на основі матриці перехідних ймовірностей ММ, інакше перехід до п. 4.

4. Якщо P^* не порожнє, то з усіх, що входять в нього ПТП вибирається правило, домінуюче над усіма ПТП множини P^* . На основі даного правила встановлюються перехідні ймовірності для стану b .

Ймовірність підтримки ТП ГСМ обчислюється за формулою

$$P\{p \in X\} = P(a, b) \cdot P(b, c) \cdot \dots \cdot P(d, e). \quad (4)$$

Детектування виробляється на основі ковзаючого вікна, для кожного символу враховувалась його участь у ТП на всій довжині вікна.

На рис. 1 представлена схема методу виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для попередження особливих випадків в польоті.

На рис. 2 представлені результати оцінки тестової послідовності довжиною 41 символ методом виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для попередження особливих випадків в польоті. Для побудови ГСМ використовувалась тренувальна послідовність з 700 символів. Рис. 2 наочно ілюструє, що установка ефективного порогу відповідності та використання продукційних правил дозволяє ефективно провести виявлення аномальних ТП. Початок тестової послідовності вважається аномальним, а починаючи з 11 символу вважається таким, що підтримується ГСМ. Поява аномального ТП свідчить про можливість виникнення особливих випадків в польоті, через не нормальну роботу технологічного обладнання повітряного судна.

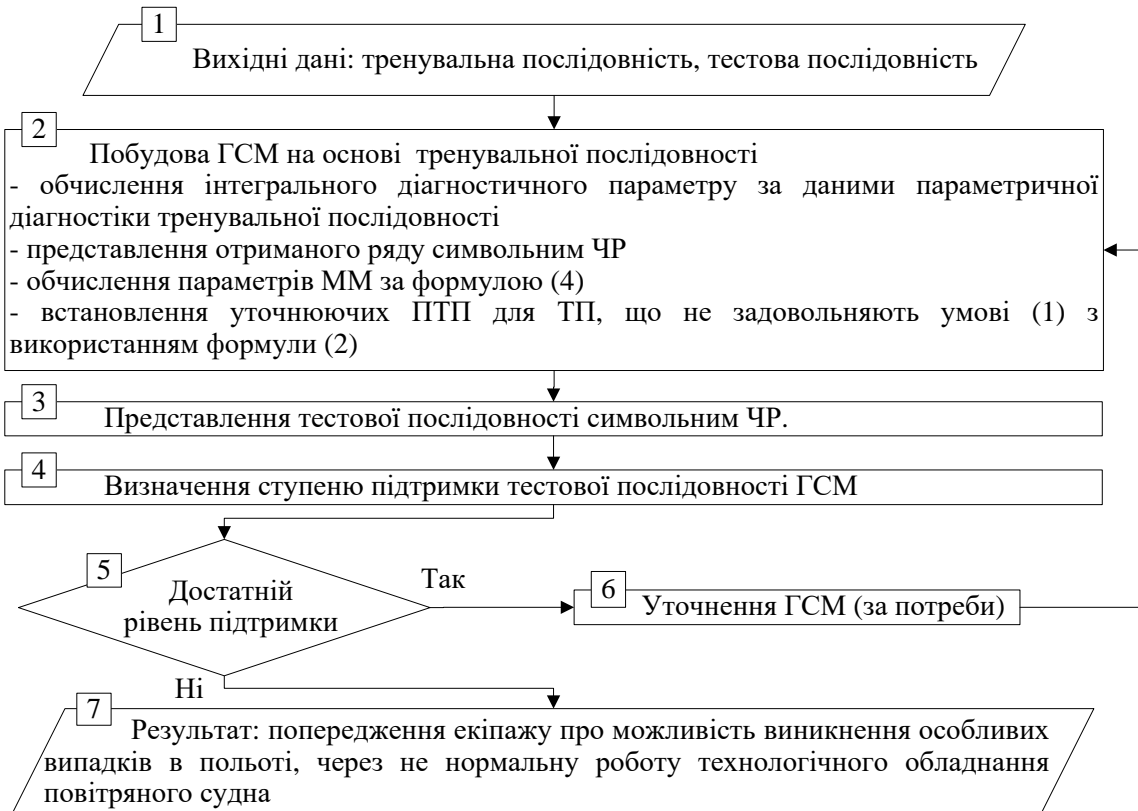


Рис. 1. Схема методу виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для попередження особливих випадків в польоті

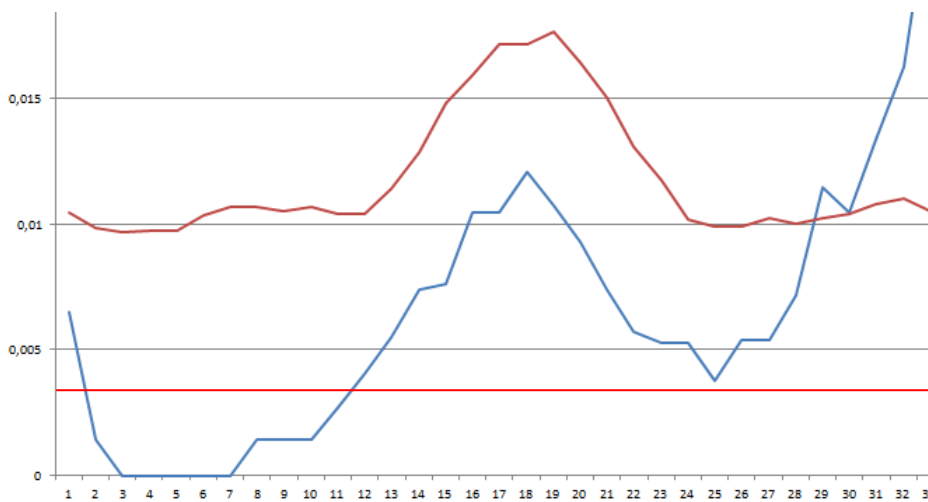


Рис. 2. Результати оцінки тестової послідовності довжиною 41 символ (синій колір), в порівнянні з випадковою частиною тренувальної послідовності тієї ж довжини (темно-червоний колір), пороговий рівень підтримки (червоний колір)

5. Висновки

У роботі запропоновано метод виявлення аномальних послідовностей в діагностичних даних технологічного обладнання повітряного судна для попередження особливих випадків в польоті. Для виявлення аномальних послідовностей пропонується використовувати гібридну

стохастичну модель, засновану на об'єднанні марковських і продукційних моделей, що використовують темпоральні правила для уточнення перехідних ймовірностей між станами процесу. За рахунок включення в модель уточнюючих продукційних правил підвищується вірогідність опису випадкових процесів, які не є марковськими, а також з'являється можливість інтеграції в модель апріорних експертних знань, що є дуже важливим для прогнозування особливих випадків в польоті. Застосування запропонованого методу дозволить впровадити прогностичний принцип управління безпекою польотів, а також отримати економічний ефект від запобігання простою повітряного судна через раптової відмови обладнання.

Список використаної літератури

1. Chandola, V. Anomaly detection: a survey / V. Chandola, Arindam Banerjee, Vipin Kumar // *ACM Computing Surveys*. – 2009. – P. 1–72.
2. Joshi, M.V. Predicting rare classes: can boosting make any weak learner strong? / M.V. Joshi, R.C. Agarwal, V. Kumar // *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM*. – New York, NY, USA, 2002. – P. 297–306.
3. Chawla, N.V. Editorial: special issue on learning from imbalanced data sets / N.V. Chawla, N. Japkowicz, A. Kotcz // *SIGKDD Explorations* 6 (1). – 2004. – P. 1–6.
4. Steinwart, I. A classification framework for anomaly detection / I. Steinwart, D. Hush, C. Scovel // *Journal of Machine Learning Research* 6. – 2005. – P. 211–232.
5. Tajbakhsh, A. Intrusion detection using fuzzy association rules / A. Tajbakhsh, Mohammad Rahmati, Abdolreza Mirzaei // *Applied Soft Computing* 9 (2). – 2009. – P. 462–469.
6. Xu, X. Adaptive network intrusion detection method based on PCA and support vector machines / X. Xu, X.N. Wang // *ADMA 2005, Lecture Notes in Artificial Intelligence, LNAI 3584*. – 2005. – P. 696–703.
7. Jha, S. Markov chains, classifiers, and intrusion detection / S. Jha, K. Tan, R. Maxion // *Proceedings of the Computer Security Foundations Workshop (CSFW)*. – 2001, June.
8. Specification based anomaly detection: a new approach for detecting network intrusions / R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, S. Zhou // *Proceedings of the 9th ACM Conference on Computer and Communications Security, ACM Press*. – 2002. – P. 265–274.
9. Shah, H. Fuzzy clustering for intrusion detection / H. Shah, J. Undercoffer, A. Joshi // *Proceedings of the 12th IEEE International Conference on Fuzzy Systems*. – 2003. – P. 1274–1278.
10. Mahoney, M. Learning nonstationary models of normal network traffic for detecting novel attacks / M. Mahoney, P. Chan // *Proceedings of 8th International Conference on Knowledge Discovery and Data Mining*. – 2002. – P. 376–385.
11. Forrest, F. Using genetic algorithms to explore pattern recognition in the immune system / F. Forrest, B. Javornik, A.S. Perelson // *Evol.* – 1993. – Vol. 1. – № 3. – P. 191–211.
12. Self-nonsel discrimination in a computer / Ф. Forrest 5, A.S. Perelson, L. Allen, R. Cherukuri // *Proceedings of IEEE symposium on research in security and privacy*. – Oakland, CA. – 16–18 May 1994. – P. 202–212.
13. Yeung, D.Y. Host-based intrusion detection using dynamic and static behavioral models / D.Y. Yeung, Y.X. Ding // *Pattern Recognition* 36. – 2003. – P. 229–243.
14. Гибридная стохастическая модель обнаружения особых типов паттернов в темпоральных данных / С.М. Ковалев, А.Н. Гуда, М.А. Бутакова // *Вестник Ростовского государственного университета путей сообщения*. – 2013. – №3(51). – С. 36-42.

15. Ma, J. Time-series novelty detection using one-class support vector machines / J. Ma, S. Perkins // Proceedings of the International Joint Conference on Neural Networks. – July 2003. – Vol. 3. – pp. 1741–1745.

References

1. Chandola, V. Anomaly detection: a survey / V. Chandola, Arindam Banerjee, Vipin Kumar // ACM Computing Surveys. – 2009. – P. 1–72.
2. Joshi, M.V. Predicting rare classes: can boosting make any weak learner strong? / M.V. Joshi, R.C. Agarwal, V. Kumar // Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM. – New York, NY, USA, 2002. – P. 297–306.
3. Chawla, N.V. Editorial: special issue on learning from imbalanced data sets / N.V. Chawla, N. Japkowicz, A. Kotcz // SIGKDD Explorations 6 (1). – 2004. – P. 1–6.
4. Steinwart, I. A classification framework for anomaly detection / I. Steinwart, D. Hush, C. Scovel // Journal of Machine Learning Research 6. – 2005. – P. 211–232.
5. Tajbakhsh, A. Intrusion detection using fuzzy association rules / A. Tajbakhsh, Mohammad Rahmati, Abdolreza Mirzaei // Applied Soft Computing 9 (2). – 2009. – P. 462–469.
6. Xu, X. Adaptive network intrusion detection method based on PCA and support vector machines / X. Xu, X.N. Wang // ADMA 2005, Lecture Notes in Artificial Intelligence, LNAI 3584. – 2005. – P. 696–703.
7. Jha, S. Markov chains, classifiers, and intrusion detection / S. Jha, K. Tan, R. Maxion // Proceedings of the Computer Security Foundations Workshop (CSFW). – 2001, June.
8. Specification based anomaly detection: a new approach for detecting network intrusions / R. Sekar, A. Gupta, J. Frullo, T. Shanbhag, A. Tiwari, H. Yang, S. Zhou // Proceedings of the 9th ACM Conference on Computer and Communications Security, ACM Press. – 2002. – P. 265–274.
9. Shah, H. Fuzzy clustering for intrusion detection / H. Shah, J. Undercoffer, A. Joshi // Proceedings of the 12th IEEE International Conference on Fuzzy Systems. – 2003. – P. 1274–1278.
10. Mahoney, M. Learning nonstationary models of normal network traffic for detecting novel attacks / M. Mahoney, P. Chan // Proceedings of 8th International Conference on Knowledge Discovery and Data Mining. – 2002. – P. 376–385.
11. Forrest, F. Using genetic algorithms to explore pattern recognition in the immune system / F. Forrest, B. Javornik, A.S. Perelson // Evol. – 1993. – Vol. 1. – № 3. – P. 191–211.
12. Self-nonsel discrimination in a computer / Ф. Forrest 5, A.S. Perelson, L. Allen, R. Cherukuri // Proceedings of IEEE symposium on research in security and privacy. – Oakland, CA. – 16–18 May 1994. – P. 202–212.
13. Yeung, D.Y. Host-based intrusion detection using dynamic and static behavioral models / D.Y. Yeung, Y.X. Ding // Pattern Recognition 36. – 2003. – P. 229–243.
14. Hybrid stochastic model for detecting special types of patterns in temporal data / S.M. Kovalev, A.N. Guda, M.A. Butakova // Vestnik Rostovskogo gosudarstvennogo universiteta putey soobshcheniya. – 2013. – №3(51). – S. 36-42.
15. Ma, J. Time-series novelty detection using one-class support vector machines / J. Ma, S. Perkins // Proceedings of the International Joint Conference on Neural Networks. – July 2003. – Vol. 3. – pp. 1741–1745.