УДК 004.056:004.942:519.852 https://doi.org/10.15407/pp2026.03.045

А.І. Ісмагілов, А.В. Шевченко, Р.М. Федоренко, П.П. Ігнатенко

МОДЕЛЬ ЗАПАСУ ФУНКЦІОНАЛЬНОЇ СТІЙКОСТІ ІНФОРМАЦІЙНОЇ СИСТЕМИ В УМОВАХ ІНЦИДЕНТІВ НУЛЬОВОГО ДНЯ

Робота присвячена створенню моделі запасу функціональної стійкості інформаційної системи в умовах інцидентів нульового дня. Актуальність досліджень визначена зростанням кількості інцидентів функціональної стійкості інформаційних систем, зокрема зростанню інцидентів нульового дня. Мета роботи – розвиток моделі запасу функціональної стійкості в умовах інцидентів нульового дня. В роботі запропонована модель системи забезпечення функціональної стійкості інформаційної системи на основі гібридної адитивно-мінімізуючої згортки з параметром керування рівнем взаємозамінності підсистем від повної взаємозамінності до повної невзаємозамінності. Як критерії якості використані показники функціональної стійкості, запасу функціональної стійкості, корисного ефекту та ефективності системи. Виявлення інцидентів нульового дня виконується шляхом визначення аномалій поведінки системи за допомогою множини методів кластеризації. Для кластеризації станів системи використаний набір агломеративних методів кластеризації в зваженій та незваженій формах (найближчого сусіда, найдальшого сусіда, попарного середнього, центроїдний метод, метод Варда та його модифікації), а також різні метрики відстані (Чебишева, Хемінга, Евкліда та квадрат метрики Евкліда). Всі методи і метрики використовуються для виконання однієї задачі багато разів різними способами. Якщо хоч одна перевірка виявить кластер, який буде визначений як аномальний, то такий кластер обов'язково буде досліджений додатково і більш детально. Підхід був апробований на модельних, реальних навчальних та на контрольних даних. Для визначення рівня небезпечності кластера використана оцінка запасу функціональної стійкості на основі відстаней поточного кластера до раніше ідентифікованих кластерів. Запропонований алгоритм визначення запасу функціональної стійкості інформаційної системи.

Ключові слова: модель, прогноз, інформаційна система, функціональна стійкість, кластеризація, аномалії, великі дані.

A.I. Ismagilov, A.V. Shevchenko, R.M. Fedorenko, P.P. Ignatenko

MODEL OF FUNCTIONAL RESILIENCE RESERVE OF AN INFORMATION SYSTEM IN THE CONDITIONS OF ZERO-DAY INCIDENTS

The work is devoted to the creation of a model of the functional stability reserve of an information system in the conditions of zero-day incidents. The relevance of the research is determined by the increase in the number of functional stability incidents of information systems, in particular the increase in zero-day incidents. The purpose of the work is to develop a model of the functional stability reserve in the conditions of zero-day incidents. The work proposes a model of a system for ensuring the functional stability of an information system based on a hybrid additive-minimizing convolution with a parameter controlling the level of interchangeability of subsystems from full interchangeability to full non-interchangeability. The indicators of functional stability, functional stability reserve, useful effect and system efficiency are used as quality criteria. Detection of zero-day incidents is performed by identifying anomalies in the system behavior using a set of clustering methods. For clustering the system states, a set of agglomerative clustering methods in weighted and unweighted forms (nearest neighbor, farthest neighbor, pairwise average, centroid, Ward's method and its modifications) were used, as well as various distance metrics (Chebyshev, Hamming, Euclidean and square of the Euclidean metric). All methods and metrics are used to perform the same task many times in different ways. If at least one check reveals a cluster that would be determined as anomalous, then such a cluster would necessarily be investigated additionally and in more detail. The approach was tested on model, real training and control data. To determine the level of cluster danger, an estimate of the functional stability reserve was used based on the distances of the current cluster to previously identified clusters. An algorithm for determining the functional stability reserve of an information system is proposed.

Keywords: short sequences, binary sequences, structural-probabilistic assessment, sequential aggregation, multicomponent model, system state change detection, structural atypicality, IoT monitoring, streaming data, forecasting.

Вступ.

Основною характеристикою цифрової трансформації є зміна статусу цифрових технологій. Тепер це не просто цифрова автоматизація діяльності, а побудова всіх інших процесів навколо цифрових технологій. Процеси стають цифрозалежними. Це неминуча плата за переваги, що їх надають цифрові технології. Цим користаються зловмисники тому, що тепер для виведення з ладу багатьох суто фізичних інфраструктур, достатньо організувати атаку на інформаційну інфраструктуру відповідного фізичного об'єкту. Виникає потреба у збереженні функціональності інформаційних систем, навіть в умовах атак. Це є питанням забезпечення функціональної стійкості (ЗФС) інформаційних систем. Тобто - забезпечення запасу функціональної стійкості (ЗФС), який забезпечить роботу інформаційної системи (ІС), якщо частка інформаційних ресурсів (ІР) буде знищена, пошкоджена або заблокована внаслідок інциденту. Тому інциденти інформаційної безпеки будемо розглядати як інциденти функціональної стійкості (ІФС) інформаційних систем.

На жаль, вивчення даних провідних агенцій, які узагальнюють статистичні дані щодо інцидентів, свідчить про постійне зростання кількості та важкості ІФС. Наприклад, експерти Verizon Communication inc. тільки за перші кілька місяців 2026 року проаналізували понад 31 тис. інцидентів, які призвели до більше 22 тис. фактів компрометації даних [1, 2]. Серед основних видів інцидентів 2026 року вони виділяють такі: вторгнення в систему (61%), соціальна інженерія (17%), атаки на веб (10%), помилки (8%), зловживання привілеями (3%). Динаміка зростання частки вторгнень в систему за роками теж не втішає: 2024 – 36%, 2025 – 53%, 2026 – 61%.

Найбільш небезпечними серед ІФС є інциденти нульового дня (ІНД), тобто інциденти, які з'явилися вперше. Тому дослідження шляхів забезпечення запасу функціональної стійкості інформаційних систем в умовах інцидентів нульового дня є актуальною задачею.

Аналіз існуючих досліджень і публікацій.

Для виявлення інцидентів нульового дня немає ані сигнатур, ані еталонів поведінки, ані інших еталонів ідентифікації. Доводиться виявляти ІНД за ознаками аномалій в роботі ІС та зовнішнього оточення ІС. Алгоритми виявлення аномалій зазвичай реалізує система виявлення вторгнень (IDS, Intrusion Detection System).

Алгоритми можуть базуватися на різних методичних підходах. Огляд багатьох з них представлений в [3] (Chandola V., Banerjee A., Kumar V.), [4] (Pang G., Shen C., Cao L., van den Hengel A.).

Багато підходів базується на кластерізації простору станів, серед яких потрібно виявити аномальні кластери. Ізоляцію аномалій випадковим розбиттям простору ознак запропонували (Liu F.T., Ting K.M., Zhou Z.-H.) [5, 6]. Критерій аномальності – мінімум кількості розбиттів, потрібних для ізоляції кластера. Інкрементна кластеризація розглянута в [7] (Burbeck K., Nadjm-Tehrani S.). Кластеризація та опорно-векторні машини використані в [8] (Song J., Takakura H., Okabe Y., Kwon Y.) та в [9] (Tang C., Xiang Y., Wang Y., Qian J., Qiang B.). Кластеризація C-Mean з перемаркуванням нормальних об'єктів використана в [10] (Shin G., Kim D., Kim S., Han M.). Поєднання C-Mean та Spatial Location Constraint Prototype Loss виконане в [11] (Nguyen T.-L., Kao H., Nguyen T.-T., Horng M.-F., Shieh C.-S.). В [12] (Cevallos M. J.F., Rizzardi A., Sicari S., Coen-Porisini A.) аномалії спочатку кластеризуються, а потім оцінюються щодо рівня небезпеки.

Кластеризація – ефективний інструмент розділення нормальних та аномальних об'єктів. Але у всіх розглянутих роботах відсутній аналіз метрики запасу функціональної стійкості або функціональної стійкості.

Мета роботи – розвиток моделей запасу функціональної стійкості в умовах інцидентів нульового дня.

Виклад основного матеріалу.

Моделі функціональної стійкості та корисних ефектів.

Моделі **підсистем** побудовані на основі логістичних залежностей SL_{i} корисних ефектів Ef_{i} від вхідних ресурсів R підсистем

$$Ef_{\text{i}}(R,P) = SL_{\text{i}}(R,P),$$

де P – параметри логістичного рівняння. Модель системи створена на основі моделей підсистем Ef_{i} за допомогою адитивної та мінімізуючої згорток

$$Ef_s = k_{Add} \sum_{i=1}^{n} \beta_{\text{i}}\, Ef_{\text{i}} + (1 - k_{Add}) \min_{i=\overline{1,n}} \left(\frac{Ef_{\text{i}}}{w_{\text{i}}}\right);$$

де n – кількість підсистем, β_{i} – вагові коефіцієнти адитивної згортки, що забезпечує повну взаємозамінність підсистем, w_{i} – вагові коефіцієнти мінімізуючої згортки, що забезпечує повну невзаємозамінність підсистем, k_{Add} – коефіцієнт взаємозамінності, який надає системі властивості повної взаємозамінності підсистем при $k_{Add} = 1$ або властивості повної невзаємозамінності підсистем при $k_{Add} = 0$, або властивості часткової взаємозамінності підсистем при $0 < k_{Add} < 1$. Вагові коефіцієнти мають бути нормовані

$$\sum_{i=1}^{n} \beta_{\text{i}} = 1.$$

Функціональна стійкість системи дорівнює корисному ефекту системи

$$FS = Ef_s.$$

Запас функціональної стійкості знаходимо як різницю між поточним значенням функціональної стійкості FS та межею функціональної стійкості mFS (мінімально припустимий рівень ФС, за якого ще зберігається працездатність системи)

$$\text{z}FS = FS - mFS.$$

За критерій якості, залежно від постановки задачі, використовується рівень корисного ефекту (функціональна стійкість) або запас функціональної стійкості

$$FS = Ef_s \to max.$$
$$zFS \to max.$$

Такі критерії найбільш доречні в задачах, коли потрібно якісно виконувати конкретні набори задач у певний зазвичай не дуже тривалий період. Натомість у постановках задач, які передбачають тривалу безперебійну роботу системи, важливішим є ефективність, тобто відношення досягнутого корисного ефекту (або запасу ФС) до ресурсів, які були на це витрачені

$$Efc = \frac{Ef_s}{R} \to max.$$

Як система, для якої ми створюємо модель ФС, розглядається система забезпечення функціональної стійкості інформаційної системи. Після створення моделі функціональної стійкості системи захисту переходимо до створення моделі запасу функціональної стійкості інформаційної системи в умовах дії ІНД.

Модель запасу функціональної стійкості інформаційної системи в умовах інцидентів нульового дня.

Модель ЗФС ІС в умовах ІНД потребує інформації про рівень небезпеки стану, в якому опинилась ІС внаслідок дії ІНД. Потрібно оцінити рівень небезпеки кожного нового стану, який описується за допомогою певних показників роботи ІС. В нашому випадку для цього використані такі показники мережевої ІС: кількість пакетів/сек, кількість байт/сек, кількість з'єднань за хвилину.

Означені показники дозволяють визначати стани системи та зовнішнього оточення. Основна задача – виявлення небезпечних станів. Друга задача (але не менш важлива) – визначення рівня небезпеки, що його створює для системи ситуація, якій відповідає певний набір показників.

Але перед тим, як оцінювати стан на основі набору показників, потрібно цей стан відокремити від інших станів (інших наборів показників). Для цього потрібно виконати кластеризацію відповідних наборів даних.

У даному дослідженні для кластеризації станів системи використаний набір

агломеративних методів кластеризації в зваженій та незваженій формах (найближчого сусіда, найдальшого сусіда, попарного середнього, центроїдний метод, метод Варда та його модифікації), а також різні метрики відстані (Чебишева, Хемінга, Евкліда та квадрат метрики Евкліда). Перелік методів та метрик може розширюватися або звужуватися залежно від результатів їх застосування на певних наборах даних. Усі методи і метрики використовуються для виконання однієї задачі багато разів різними способами. Якщо хоч одна перевірка виявить кластер, який буде визначений, як аномальний, то такий кластер обов'язково буде досліджений додатково і більш детально.

В агломеративних методах, що були використані, кожна точка первинно була розміщена в окремому кластері. Далі об'єднувалася та пара кластерів, що мала найменшу величину критерію, який був унікальним для кожного методу та представляв певну агрегацію відстаней між окремими точками та/або іншими характеристиками кластерів.

Підхід був апробований на модельних, реальних навчальних та на контрольних даних.

Навчальна вибірка реальних даних містила дані, що відповідали таким характерним станам ІС.

Нормальні (безпечні) стани.

1. Нормальна робота (Normal Activity, n) системи з типовими стабільними показниками без надмірних коливань.

2. Нормальне пікове навантаження (Normal Peak Load, nP) на систему, яке пов'язане з виконанням ресурсоємних завдань, добовими коливаннями, резервуванням даних.

Аномальні (небезпечні) стани.

3. Сканування портів (Scan Port, aS) на підготовчих етапах атак. Невеликий трафік. Велика кількість з'єднань.

4. Підбір паролів прямим перебором (Brute Force Password, aB). Невеликий трафік. Велика кількість з'єднань. Висока активність автентифікації.

5. Періодичний сигнальний обмін (Beaconing Periodical, aP) невеликими пакетами між скомпрометованим вузлом та зовнішнім сервером. Трафік невеликий. Стабільний періодичний обмін даними.

6. Атака низької інтенсивності (Attack Low Slow, aL) утримує низький рівень параметрів, щоб не перевищити поріг виявлення.

7. Розподілена атака відмови в обслуговуванні (DDoS, aD). Інтенсивне зростання показників.

8. Витік інформації (Data Exfiltration, aE) з високим трафіком та малою кількістю з'єднань.

Для кожного набору даних було виконано 48=4x12 варіантів обчислювальних експериментів (рис.1). Найбільша чутливість щодо виявлення аномалій була досягнута при комбінації декількох методів.

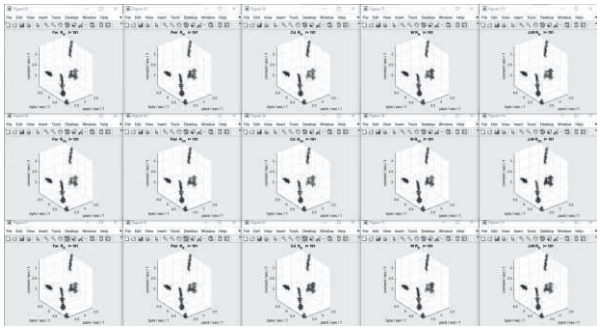

Рис. 1. Скріншот роботи програми щодо кластеризації станів ІС різними методами на реальних даних.

На основі бази знань ідентифікованих станів корпоративної інформаційно-комунікаційної системи ТОВ «САЙБЕР ЛАБ» були обрані найбільш характерні набори даних для кожного із розглянутих нормальних та аномальних станів (18 – 32 точок на один стан) (рис.2). Для більшої контрастності даних датасети пройшли логарифмічну нормалізацію.

Наочніше ці дані представлені у тривимірному вигляді (рис.3). Тривимірне представлення надає уявлення про те, як методи та метрики кластеризації дозволяють розділити набори точок на кластери у просторі. Наприклад, на відміну від двовимірного, тривимірне представлення наочно показує, що деякі стани перебувають на дуже великій відстані від усіх інших, що дозволяє не просто виділити їх в окремий кластер, а і створити чіткий еталон для ідентифікації такого типу аномального стану. Цей висновок можна віднести, наприклад, до

DDoS-атак та витоку даних (рис.4, 5). У разі подальшого збільшення деталізації представлення станів інформаційної системи в просторі станів (рис.6, 7) було виявлено, що досить чіткі еталони можна також створити для станів сканування портів та підбору паролів прямим перебором.

Для формування еталонів аномалій були запропоновані такі ознаки:

- межі кластеру за всіма вимірами;
- центр мас кластеру;
- середньо-квадратичне відхилення координат точок кластеру;
- коваріаційна матриця;
- власні вектори коваріаційної матриці.

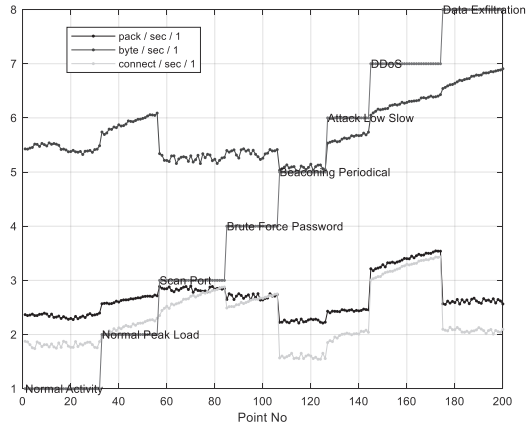

Рис. 2. Первинна статистика щодо станів корпоративної інформаційно-комунікаційної системи ТОВ «САЙБЕР ЛАБ» у вигляді функціональних залежностей

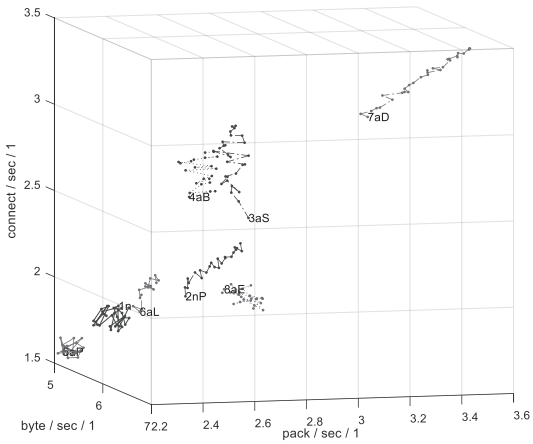

Рис. 3. Первинна статистика щодо станів корпоративної інформаційно-комунікаційної системи ТОВ «САЙБЕР ЛАБ» у тривимірному просторі факторів

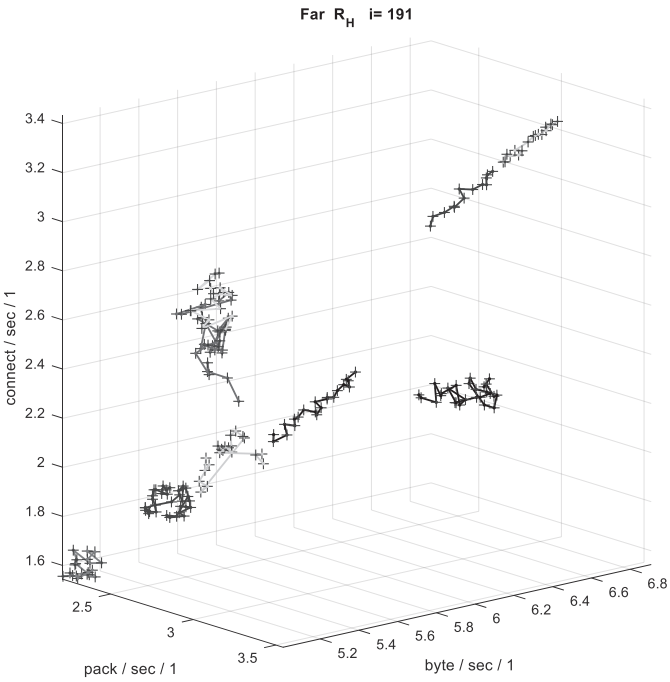

Рис. 4. Результати кластеризації методом Farthest. Метрика Хемінга

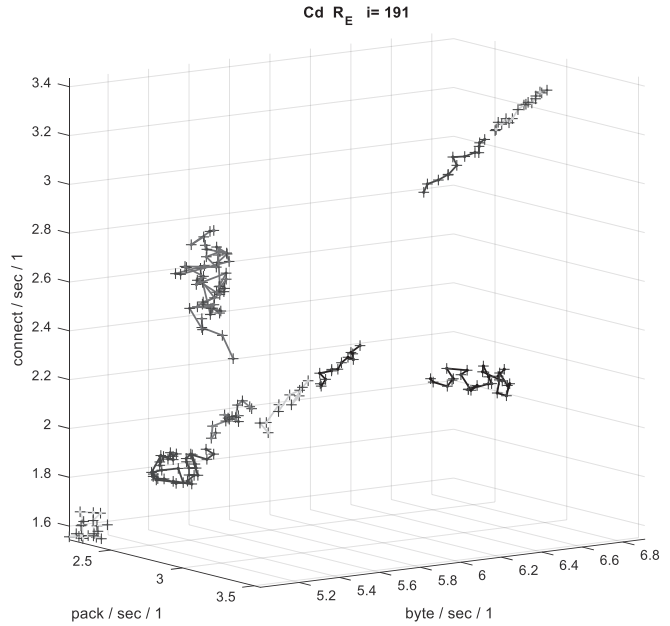

Рис. 5. Результати кластеризації методом Centroid. Метрика Евкліда

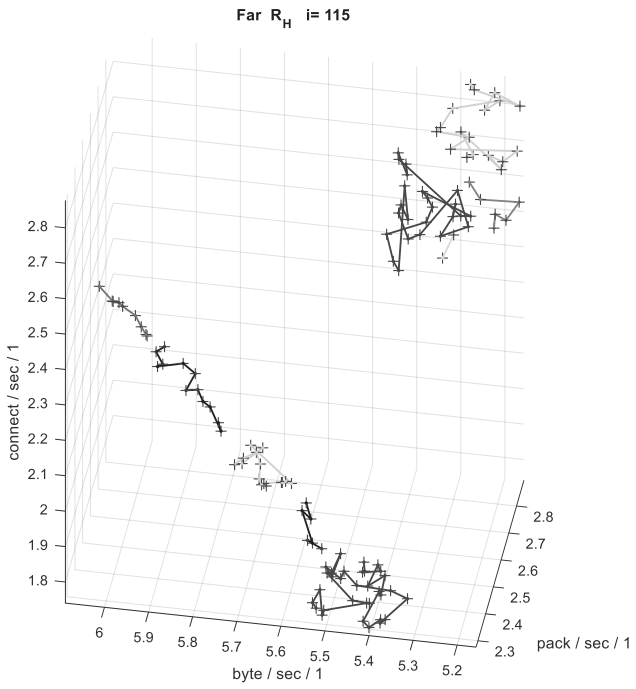

Рис. 6. Результати кластеризації методом Farthest. Метрика Хемінга

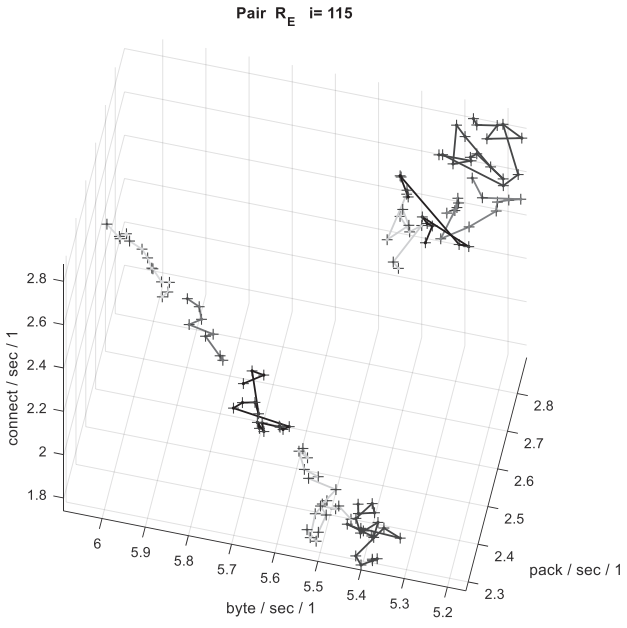

Рис. 7. Результати кластеризації методом PairMean. Метрика Евкліда

Визначення рівня запасу функціональної стійкості системи.

Для визначення рівня небезпечності кластера використана оцінка запасу функціональної стійкості на основі відстаней поточного кластера до раніше ідентифікованих кластерів. Тобто, маємо ситуацію, коли виявлений новий кластер невідомого походження, але водночас ми можемо знайти відстані від цього кластера до всіх ідентифікованих кластерів:

Dp_i – відстань від невідомого до відомого позитивного кластера з номером i;

Dn_i – відстань від невідомого до відомого негативного кластера з номером i.

Відстань може вимірюватися між такими величинами:

- центри кластерів (центроїди);
- найближчі та найдальші точки кластерів.

Визначимо **агреговану позитивність** кластера, що досліджується,

$$P_a = \sum_{i=1}^{m_p} \frac{1}{(1+Dp_i)} - \sum_{i=1}^{m_n} \frac{1}{(1+Dn_i)}$$

де m_p – кількість позитивно ідентифікованих кластерів, m_n – кількість негативно ідентифікованих кластерів.

Вираз для агрегованої позитивності кластера може змінювати значення від m_p до -m_n. Позитивні значення відповідають стану безпечності кластера. Негативні – стану небезпечності.

Рівень ФС ІС в умовах стану, що досліджується, знаходимо за допомогою виразу

$$FS = 0.5(1 + th(P_a)).$$

ФС може приймати значенні від 0 до 1. Для корисних ефектів та вхідних ресурсів межа ФС встановлюється на рівні 0.6. Але для ФС на основі метрик безпеки, межу ФС встановлюємо на рівні 0.5. ЗФС знаходимо як

$$zFS = 0.5\ th(P_a).$$

Алгоритм визначення ЗФС ІС

1. Вибір метрик та методів кластеризації.

2. Перевірка ефективності пар метод + метрика на модельних та на реальних даних.

3. Виявлення станів, для яких можливе створення еталонів.

4. Визначення кількості кластерів.

5. Кластеризація. Виявлення підозрілих кластерів.

6. Детальне вивчення підозрілих кластерів.

7. Визначення рівня ЗФС в умовах стану, що досліджується.

8. Висновки щодо аномалій та небезпеки.

Висновки

Робота присвячена створенню моделі запасу функціональної стійкості інформаційної системи в умовах інцидентів нульового дня.

Актуальність досліджень визначена зростанням кількості інцидентів функціональної стійкості інформаційних систем, зокрема зростанням інцидентів нульового дня.

В роботі запропонована модель системи забезпечення функціональної стійкості інформаційної системи на основі гібридної адитивно-мінімізуючої згортки з параметром керування рівнем взаємозамінності підсистем від повної взаємозамінності до повної невзаємозамінності.

За критерії якості використані показники функціональної стійкості, запасу функціональної стійкості, корисного ефекту та ефективності системи.

Виявлення інцидентів нульового дня виконується шляхом визначення аномалій поведінки системи за допомогою множини методів кластеризації.

Для кластеризації станів системи використаний набір агломеративних методів кластеризації в зваженій та незваженій формах (найближчого сусіда, найдальшого сусіда, попарного середнього, центроїдний метод, метод Варда та його модифікації), а також різні метрики відстані (Чебишева, Хемінга, Евкліда та квадрат метрики Евкліда).

Всі методи і метрики використовуються для виконання однієї задачі багато разів різними способами. Якщо хоч одна перевірка виявить кластер, який буде визначений як аномальний, то такий кластер обов'язково буде досліджений додатково і більш детально.

Підхід був апробований на модельних, реальних навчальних та на контрольних даних.

Для визначення рівня небезпечності кластера використана оцінка запасу функціональної стійкості на основі відстаней поточного кластера до раніше ідентифікованих кластерів.

Запропоновано алгоритм визначення запасу функціональної стійкості інформаційної системи.

Напрями подальших досліджень – розширення набору метрик і методів кластеризації.

Література

1. 2026 Data Breach Investigation Report. Public Sector snapshot. Verizon business. https://www.verizon.com/business/resources/reports/2026-dbir-public-sector-snapshot.pdf.
2. 2026 Data Breach Investigations Report https://www.verizon.com/business/resources/reports/dbir/.
3. Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey // ACM Computing Surveys. – 2009. – Vol. 41, No. 3. – Article 15. – P. 1–58. – https://doi.org/10.1145/1541880.1541882
4. Pang G., Shen C., Cao L., van den Hengel A. Deep Learning for Anomaly Detection: A Review // ACM Computing Surveys. – 2021/2022. – Vol. 54, No. 2. – Article 38. – P. 1–38. – https://doi.org/10.1145/3439950
5. Liu F.T., Ting K.M., Zhou Z.-H. Isolation Forest // Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (ICDM 2008). – Pisa, Italy, 15–19 December 2008. – P. 413–422. – https://doi.org/10.1109/ICDM.2008.17
6. Liu F.T., Ting K.M., Zhou Z.-H. Isolation-Based Anomaly Detection // ACM Transactions on Knowledge Discovery from Data. – 2012. – Vol. 6, No. 1. – Article 3. – P. 1–39. – https://doi.org/10.1145/2133360.2133363
7. Burbeck K., Nadjm-Tehrani S. Adaptive Real-Time Anomaly Detection with Incremental Clustering // Information Security Technical Report. – 2007. – Vol. 12, No. 1. – P. 56–67. – https://doi.org/10.1016/j.istr.2007.02.004
8. Song J., Takakura H., Okabe Y., Kwon Y. Unsupervised Anomaly Detection Based on Clustering and Multiple One-Class SVM // IEICE Transactions on Communications. – 2009. – Vol. E92-B, No. 6. – P. 1981–1990. – https://doi.org/10.1587/transcom.E92.B.1981
9. Tang C., Xiang Y., Wang Y., Qian J., Qiang B. Detection and Classification of Anomaly Intrusion Using Hierarchy Clustering and SVM

// Security and Communication Networks. – 2016. – Vol. 9. – P. 3401–3411. – https://doi.org/10.1002/sec.1547
10. Shin G., Kim D., Kim S., Han M. Unknown Attack Detection: Combining Relabeling and Hybrid Intrusion Detection // Computers, Materials & Continua. – 2021. – Vol. 68, No. 3. – P. 3289–3303. – https://doi.org/10.32604/cmc.2021.017502
11. Nguyen T.-L., Kao H., Nguyen T.-T., Horng M.-F., Shieh C.-S. Unknown DDoS Attack Detection with Fuzzy C-Means Clustering and Spatial Location Constraint Prototype Loss // Computers, Materials & Continua. – 2024. – Vol. 78, No. 2. – P. 2181–2205. – https://doi.org/10.32604/cmc.2024.047387
12. Cevallos M. J.F., Rizzardi A., Sicari S., Coen-Porisini A. HERO: From High-Dimensional Network Traffic to zERO-Day Attack Detection // Computer Networks. – 2025. – Vol. 265. – Article 111264. – https://doi.org/10.1016/j.comnet.2025.111264

Дата першого надходження до видання: 06.07.2026
Внутрішня рецензія отримана: 27.07.2026
Зовнішня рецензія отримана: 29.07.2026
Дата прийняття статті до друку: 10.08.2026
Дата публікації: 29.08.2026

Про авторів:

Ісмагілов Артур Ільясович,
аспірант
Ismahilov Artur,
PhD student
https://orcid.org/0009-0002-1332-8147

Шевченко Аліна Віталіївна,
кандидат технічних наук
Shevchenko Alina,
PhD (computer science)
http://orcid.org/0000-0003-3793-9364

Федоренко Руслан Миколайович,
кандидат економічних наук,
старший дослідник,
Fedorenko Ruslan,
PhD in Economics, senior researcher
https://orcid.org/0000-0001-9433-5458
E-mail: r_fedorenko@ukr.net

Ігнатенко Петро Петрович,
кандидат технічних наук,
старший науковий співробітник
Ignatenko Petro,
PhD (computer science),
senior researcher
https://orcid.org/0000-0002-2204-1554
E-mail: ignat@isofts.kiev.ua

Місце роботи авторів:

Інститут програмних систем
Національної академії наук України,
03187, м. Київ,
проспект Академіка Глушкова, 40,
корпус 5.
The Institute of Software Systems of the
National Academy of Sciences of Ukraine,
Kyiv, 03187, Ukraine,
Akademika Glushkova Avenue, 40,
building 5
https://iss.nas.gov.ua/