УДК 004.056.53

МЕТОД ВИЯВЛЕННЯ ТА НЕЙТРАЛІЗАЦІЇ КОМПРОМЕТОВАНИХ ВУЗЛІВ У РОЗПОДІЛЕНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ ІЗ ЗАБЕЗПЕЧЕННЯМ ЦІЛІСНОСТІ ДАНИХ

Олексій Німич, Ігор Макєєв

У сучасних розподілених інформаційних системах, зокрема в середовищах Інтернету речей (IoT), що характеризуються значною кількістю взаємопов'язаних вузлів та інтенсивним обміном даними, забезпечення кібербезпеки та цілісності інформації є однією з ключових задач. Особливої актуальності набуває проблема компрометації окремих вузлів, яка може призводити до спотворення даних, порушення коректності обчислень і зниження загальної надійності системи. Традиційні підходи до захисту не завжди забезпечують своєчасне виявлення аномальної поведінки вузлів, що зумовлює необхідність розроблення ефективних методів детекції та нейтралізації загроз.

У статті запропоновано метод виявлення та нейтралізації компрометованих вузлів у розподілених інформаційних системах, який базується на аналізі відхилень параметрів, що надходять від окремих вузлів, від узагальнених характеристик системи. Метод передбачає формування агрегованих показників, зокрема середніх значень, та визначення порогових критеріїв, що дозволяють ідентифікувати аномальні відхилення. У разі перевищення встановленого порогу вузол класифікується як потенційно компрометований.

Запропонований підхід включає механізм ізоляції підозрілих вузлів та коригування процесу оброблення даних з урахуванням виключення їхнього впливу на результати. Це дозволяє підвищити достовірність обчислень і забезпечити збереження цілісності інформації навіть у разі часткової компрометації системи. Окрему увагу приділено простоті реалізації методу, що забезпечує можливість його застосування в системах з обмеженими обчислювальними ресурсами.

Отримані результати демонструють, що запропонований метод дозволяє ефективно виявляти аномальну поведінку вузлів і зменшувати негативний вплив кіберзагроз на функціонування системи. Запропонований підхід може бути використаний як складова комплексних систем забезпечення кібербезпеки розподілених інформаційних середовищ.

Ключові слова: *компрометовані вузли, виявлення аномалій, цілісність даних, нейтралізація загроз, розподілені інформаційні системи, Інтернет речей.*

Вступ

Сучасні розподілені інформаційні системи характеризуються наявністю великої кількості взаємопов'язаних вузлів, що здійснюють обмін даними в реальному часі та спільно виконують обчислювальні задачі [1, 2]. Зростання складності та масштабованості таких систем зумовлює підвищення рівня кіберзагроз і розширення поверхні атак. Однією з небезпечних загроз є компрометація окремих вузлів, яка може виникати внаслідок несанкціонованого доступу, впровадження шкідливого програмного забезпечення або порушення протоколів безпеки [3, 7]. Компрометовані вузли здатні передавати спотворені або недостовірні дані, що призводить до зниження достовірності обчислень, порушення цілісності інформації та дестабілізації функціонування всієї системи [6]. Особливо критичною ця проблема є для промислових розподілених інформаційних систем, зокрема систем Інтернету речей (IoT) [6], що використовуються на об'єктах критичної інфраструктури, де рішення приймаються на основі агрегованих даних з багатьох джерел.

Традиційні підходи до забезпечення кібербезпеки, як правило, орієнтовані на захист периметра системи або централізований контроль доступу, що не завжди є ефективним у розподілених середовищах із динамічною структурою [3, 7]. У зв'язку з цим зростає потреба у розробленні методів, здатних забезпечувати виявлення аномальної поведінки вузлів та нейтралізацію їхнього впливу на систему в реальному часі.

Аналіз сучасних досліджень показує, що для вирішення зазначеної проблеми широко застосовуються методи виявлення аномалій, статистичного аналізу та машинного навчання, спрямовані на ідентифікацію відхилень у поведінці вузлів [4, 5]. Зокрема, використовуються підходи, що базуються на кластеризації, методах найближчих сусідів, байєсівських моделях, а також нейромережевих алгоритмах, здатних виявляти складні нелінійні залежності у даних [4, 5]. Значна увага приділяється побудові моделей, які дозволяють автоматично визначати аномальні патерни та прогнозувати потенційні загрози на основі історичних даних [4].

Водночас існуючі підходи часто мають високу обчислювальну складність або потребують значних обсягів навчальних вибірок, що обмежує їх застосування в системах з обмеженими ресурсами, зокрема у промислових середовищах Інтернету речей. Крім того, такі методи можуть бути чутливими до якості вхідних даних та не забезпечувати достатньої стійкості до

цілеспрямованих атак, зокрема у випадках компрометації окремих вузлів. У багатьох випадках складність налаштування моделей і необхідність їх постійного перенавчання знижують практичну ефективність зазначених підходів у динамічних розподілених середовищах.

Таким чином, актуальною є задача розроблення простих та ефективних методів виявлення і нейтралізації компрометованих вузлів, які забезпечують збереження цілісності даних та підвищення стійкості розподілених інформаційних систем до кіберзагроз.

Метою даної роботи є розроблення методу виявлення та нейтралізації компрометованих вузлів у розподілених інформаційних системах на основі аналізу відхилень їхніх параметрів від узагальнених характеристик системи з подальшою ізоляцією підозрілих елементів.

Постановка задачі

Розглядається розподілена інформаційна система, що складається з множини вузлів

$$S = \{n_1, n_2, \dots, n_N\},$$

які здійснюють обмін даними та беруть участь у спільному процесі оброблення інформації.

Припускається, що частина вузлів може бути компрометованою та передавати спотворені або недостовірні дані. Позначимо множину таких вузлів як

$$C \subset S,$$

причому її склад заздалегідь невідомий.

Задача полягає у розробленні методу виявлення компрометованих вузлів на основі аналізу відхилень їхніх даних від узагальнених характеристик системи з подальшою нейтралізацією їхнього впливу та забезпеченням цілісності даних.

Метод виявлення та нейтралізації компрометованих вузлів

Запропонований метод ґрунтується на статистичному аналізі відхилень параметрів, що надходять від вузлів розподіленої системи, від узагальнених характеристик її стану. Метод передбачає формування агрегованої оцінки даних, визначення індивідуальних відхилень для кожного вузла та їх порівняння з пороговим значенням. Адаптивні властивості методу проявляються у динамічному визначенні порогового значення на основі поточних статистичних характеристик даних, що дозволяє враховувати зміну умов функціонування системи. Основна ідея полягає у виявленні вузлів, поведінка яких істотно відрізняється від типової, з подальшою ізоляцією таких вузлів та коригуванням процесу оброблення даних для забезпечення цілісності інформації..

Нехай розподілена система складається з Nвузлів, кожен з яких передає значення деякого параметра x_i, де $i = 1,2, \dots, N$. Для оцінювання загального стану системи визначається агреговане значення параметра у вигляді середнього:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

Величина $\bar{x}$, визначена за формулою (1), використовується як еталонна характеристика нормального функціонування системи.

Для кожного вузла обчислюється абсолютне відхилення його значення від середнього:

$$D_i = | x_i - \bar{x} | \qquad (2)$$

де D_iхарактеризує ступінь відхилення параметра вузла n_iвід типового значення. Чим більше значення D_i, тим більшою є ймовірність аномальної або компрометованої поведінки вузла.

Для підвищення стійкості методу до випадкових коливань даних вводиться оцінка дисперсії або середньоквадратичного відхилення:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \bar{x})^2} \qquad (3)$$

Отримане значення σ, визначене за формулою (3), дозволяє оцінити природний рівень варіації даних у системі.

На основі цього вводиться порогове значення T, яке визначається як кратне стандартного відхилення:

$$T = k \cdot \sigma \qquad (4)$$

де k— емпіричний коефіцієнт (зазвичай $k \in [2; 3]$).

Вузол вважається потенційно компрометованим, якщо його відхилення перевищує встановлений поріг:

$$D_i > T \qquad (5)$$

Умова (5) є критерієм виявлення аномальної поведінки вузлів. У разі її виконання вузол n_iкласифікується як підозрілий та підлягає подальшій обробці.

Після ідентифікації компрометованих вузлів реалізується механізм їх нейтралізації, який полягає у виключенні їхніх даних із процесу агрегування. Для цього формується нова множина довірених вузлів:

$$S' = S \setminus C \qquad (6)$$

де C— множина вузлів, що задовольняють умову (5).

Оновлене агреговане значення обчислюється лише на основі довірених вузлів:

$$\bar{x}' = \frac{1}{|S'|}\sum_{n_i \in S'} x_i \qquad (7)$$

де $|S'|$— кількість вузлів після виключення підозрілих.

Таким чином, запропонований метод дозволяє не лише виявляти компрометовані вузли на основі аналізу їхніх відхилень, але й зменшувати їхній вплив на результати оброблення даних шляхом адаптивного коригування складу вузлів, що беруть участь у формуванні агрегованих характеристик системи.

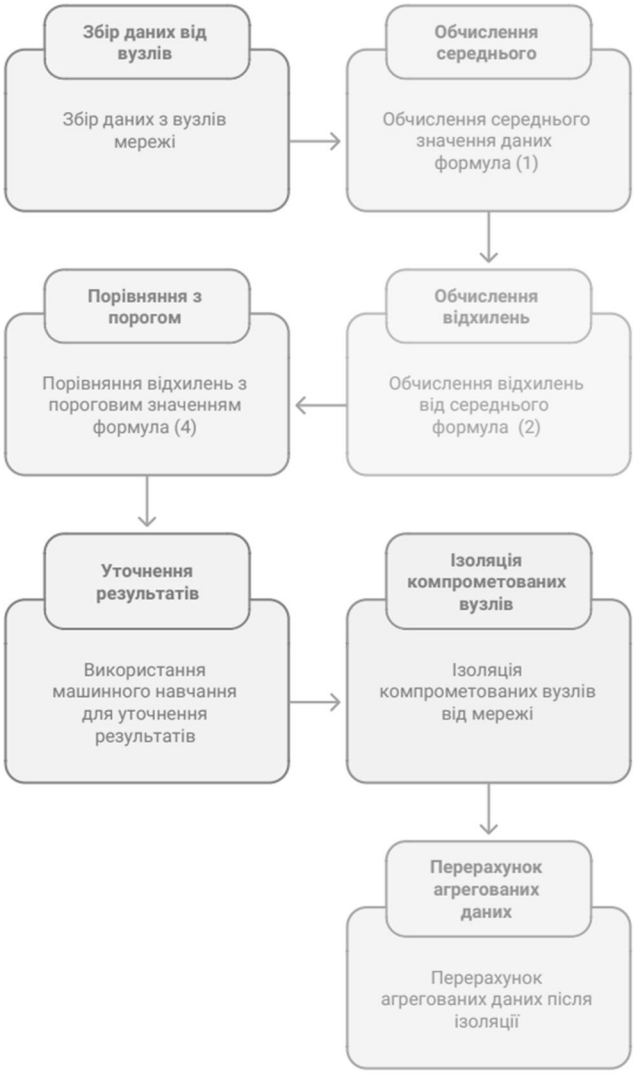

Рис. 1 – Структурна схема методу виявлення та нейтралізації компрометованих вузлів у розподілених інформаційних системах

Як показано на рис. 1, запропонований метод реалізується у вигляді послідовності етапів, що включають статистичний аналіз даних, порогову оцінку відхилень та уточнення результатів із використанням методів машинного навчання.

На початковому етапі здійснюється збір даних від вузлів системи та формування узагальненої характеристики у вигляді середнього значення. Далі обчислюються відхилення параметрів окремих вузлів від середнього значення та виконується їх порівняння з пороговим значенням. У разі перевищення порогу вузол класифікується як потенційно компрометований. Для підвищення точності виявлення на етапі уточнення може застосовуватися додатковий аналіз із використанням методів машинного навчання. На завершальному етапі здійснюється ізоляція компрометованих вузлів та оновлення агрегованих даних з урахуванням виключення їхнього впливу на результати оброблення. Нейтралізація компрометованих вузлів у запропонованому методі реалізується шляхом обмеження їхнього впливу на результати оброблення даних. Зокрема, вузли, які задовольняють умову (5), виключаються з процесу агрегування даних або їхній внесок у загальний результат ігнорується. Це дозволяє запобігти впливу спотворених значень на узагальнені характеристики системи та забезпечити збереження цілісності інформації.

Оцінювання ефективності запропонованого методу може бути здійснено шляхом імітаційного моделювання функціонування розподіленої інформаційної системи в умовах наявності компрометованих вузлів. У межах такого моделювання вузли системи генерують дані відповідно до заданих статистичних характеристик, тоді як частина вузлів формує спотворені значення, що імітують аномальну поведінку.

Ефективність методу доцільно оцінювати за такими показниками, як точність виявлення компрометованих вузлів, частота хибних спрацьовувань та ступінь впливу аномальних даних на результати агрегування. Використання статистичного підходу дозволяє забезпечити швидке виявлення вузлів із істотними відхиленнями, що є характерною ознакою компрометації.

Застосування адаптивного порогового значення, визначеного відповідно до формули (4), забезпечує стійкість методу до змін умов функціонування системи та дозволяє знизити ймовірність хибних спрацьовувань. Додаткове використання методів машинного навчання на етапі уточнення результатів може підвищити точність класифікації вузлів у випадках, коли відхилення є незначними.

Таким чином, запропонований метод може бути ефективно застосований у розподілених

інформаційних системах та середовищах Інтернету речей для забезпечення цілісності даних і зменшення впливу компрометованих вузлів на результати оброблення.

Висновки

У роботі запропоновано метод виявлення та нейтралізації компрометованих вузлів у розподілених інформаційних системах на основі аналізу відхилень параметрів від узагальнених характеристик системи. Метод базується на поєднанні статистичного підходу до виявлення аномалій із можливістю використання елементів машинного навчання для уточнення результатів класифікації.

Запропонований підхід дозволяє ефективно ідентифікувати вузли з аномальною поведінкою та обмежувати їхній вплив на результати оброблення даних шляхом виключення з процесу агрегування. Це забезпечує підвищення цілісності інформації та стійкості системи до внутрішніх і зовнішніх кіберзагроз.

Перевагами методу є простота реалізації, низька обчислювальна складність та наявність адаптивних властивостей, що проявляються у динамічному визначенні порогового значення залежно від статистичних характеристик даних. Це робить можливим застосування методу в системах з обмеженими обчислювальними ресурсами, зокрема у розподілених середовищах та середовищах Інтернету речей.

Перевірка ефективності запропонованого методу може бути здійснена шляхом імітаційного моделювання, у межах якого формуються масиви даних, що відображають нормальну та аномальну поведінку вузлів. Зокрема, можуть використовуватися як синтетичні дані із заданими статистичними характеристиками, так і реальні набори даних із розподілених систем або середовищ Інтернету речей.

Подальші дослідження доцільно спрямувати на проведення імітаційного моделювання для кількісної оцінки ефективності методу, зокрема аналізу точності виявлення компрометованих вузлів, рівня хибних спрацьовувань та впливу аномальних даних на результати агрегування. Крім того, перспективним є розширення методу за рахунок використання більш складних моделей машинного навчання та врахування динаміки поведінки вузлів у часі.

Список літератури

[1] Atzori L., Iera A., Morabito G. The Internet of Things: A survey. Computer Networks. 2010. Vol. 54, No. 15. P. 2787–2805.

[2] Tanenbaum A. S., Van Steen M. Distributed Systems: Principles and Paradigms. 2nd ed. Prentice Hall, 2007.

[3] Stallings W. Cryptography and Network Security: Principles and Practice. 7th ed. Pearson, 2017.

[4] Chandola V., Banerjee A., Kumar V. Anomaly Detection: A Survey. ACM Computing Surveys. 2009. Vol. 41, No. 3.

[5] Liao Y., Vemuri V. R. Use of k-nearest neighbor classifier for intrusion detection. Computers & Security. 2002. Vol. 21, No. 5. P. 439–448.

[6] Xu L. D., He W., Li S. Internet of Things in industries: A survey. IEEE Transactions on Industrial Informatics. 2014. Vol. 10, No. 4. P. 2233–2243.

[7] Bishop M. Computer Security: Art and Science. Addison-Wesley, 2003.

[8] Blaze M., Feigenbaum J., Lacy J. Decentralized Trust Management. Proceedings of the IEEE Symposium on Security and Privacy. 1996. P. 164–173.

[9] National Institute of Standards and Technology. Security and Privacy Controls for Information Systems and Organizations (NIST SP 800-53). NIST, 2020.

[10] International Organization for Standardization. ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection. ISO, 2022.

[11] Zhang Y., Kasahara S., Shen Y., Jiang X., Wan J. Smart contract-based access control for the Internet of Things. IEEE Internet of Things Journal. 2018. Vol. 6, No. 2. P. 1594–1605.

[12] Nguyen T. T., Reddi V. J. Deep reinforcement learning for cyber security. IEEE Transactions on Neural Networks and Learning Systems. 2021. Vol. 34, No. 2. P. 377–388.

METHOD FOR DETECTING AND MITIGATING COMPROMISED NODES IN DISTRIBUTED INFORMATION SYSTEMS WITH DATA INTEGRITY ASSURANCE

Modern distributed information systems, particularly Internet of Things (IoT) environments, are characterized by a large number of interconnected nodes and intensive data exchange. Ensuring cybersecurity and data integrity in such systems is one of the key challenges. Of particular importance is the problem of node compromise, which may lead to data distortion, incorrect computations, and a decrease in overall system reliability. Traditional security approaches do not always ensure timely detection of anomalous node behavior, which necessitates the development of effective methods for threat detection and mitigation.

This paper proposes a method for detecting and mitigating compromised nodes in distributed information systems based on the analysis of deviations of node-generated parameters from aggregated system characteristics. The method involves the formation of aggregate indicators, including average values, and the determination of threshold criteria for identifying anomalous deviations. If the threshold is exceeded, a node is classified as potentially compromised.

The proposed approach includes a mechanism for isolating suspicious nodes and adjusting the data processing procedure by excluding their influence on the results. This makes it possible to improve the reliability of computations and ensure data integrity even in the presence of partially compromised system components. Special attention is paid to the simplicity of implementation, which allows the method to be applied in systems with limited computational resources.

The obtained results demonstrate that the proposed method effectively detects anomalous node behavior and reduces the negative impact of cyber threats on system operation. The approach can be used as a component of comprehensive cybersecurity solutions for distributed information environments.

Keywords: compromised nodes, anomaly detection, data integrity, threat mitigation, distributed information systems, Internet of Things.

Німич Олексій Віталійович, аспірант кафедри кібербезпеки, Державного некомерційного підприємства «Державний університет «Київський авіаційний інститут», м.Київ, Україна.

Oleksii Nimych, PhD student of the Department of Cybersecurity of the State non-commercial company state university «Kyiv aviation institute», Kyiv, Ukraine.

E-mail: 5356349@stud.kai.edu.ua

ORCID ID: 0000-0003-1759-7088

Ігор Генрихович Макєєв, аспірант кафедри кібербезпеки, Державного некомерційного підприємства «Державний університет «Київський авіаційний інститут», м.Київ, Україна.

Ihor Makieiev, PhD student of the Department of Cybersecurity of the State non-commercial company state university «Kyiv aviation institute», Kyiv, Ukraine.

E-mail: 8390988@stud.kai.edu.ua

ORCID ID: 0009-0009-8679-5652