

ШИМЧУК ГРИГОРІЙ

Тернопільський національний технічний університет імені Івана Пулюя
<https://orcid.org/0000-0003-2362-7386>
e-mail: gorych@gmail.com

ШЕВЧЕНКО НАЗАР

Тернопільський національний технічний університет імені Івана Пулюя
e-mail: sheva250298@gmail.com

ШВИРЛО КОНСТАНТИН

Тернопільський національний технічний університет імені Івана Пулюя
<https://orcid.org/0009-0009-9808-7653>
e-mail: konstantyn.shvyrlo@gmail.com

ГАРМАТЮК НАЗАР

Тернопільський національний технічний університет імені Івана Пулюя
e-mail: nazarharmatyk@gmail.com

СИСТЕМА ВІДНОВЛЕННЯ ДАНИХ У БЕЗДРОВОВИХ СЕНСОРНИХ МЕРЕЖАХ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Втрата даних у бездротових сенсорних мережах є серйозною проблемою, яка безпосередньо впливає на точність вимірювань, коректність аналізу та загальну надійність передачі інформації. В умовах постійної зміни навколишнього середовища, завад, перевантаженості мережі або виходу з ладу окремих вузлів спостерігається суттєве зниження якості зв'язку, що призводить до втрати частини переданих пакетів даних.

Існуючі методи боротьби з втратами, такі як циклічний контрольний код (CRC), інтерполяційні алгоритми або схеми повторної передачі пакетів, мають певні обмеження. Вони можуть бути ефективними при невеликих втратах, проте в умовах високих перешкод або нестабільності мережевого середовища їх точність та ефективність значно знижується.

У цій статті запропоновано нову систему відновлення даних, яка поєднує класичні методи контролю цілісності з можливостями машинного навчання для точнішої реконструкції відсутніх значень. Основою підходу є математична модель випадкового циклічного процесу, яка дозволяє аналізувати та прогнозувати поведінку мережі на основі часових рядів попередніх спостережень. Використання Singular Spectrum Analysis (SSA) допомагає виділити основні тренди, циклічні закономірності та випадкові коливання у часових рядах, що дає змогу більш точно передбачати відсутні значення. Ця методика дозволяє не лише відновлювати дані, але й прогнозувати можливі втрати, що дає змогу випереджати проблеми та мінімізувати вплив збоїв у передачі.

Експериментальні дослідження продемонстрували значне зниження втрат даних, покращення загальної продуктивності мережі та підвищення енергоефективності вузлів. Аналіз отриманих результатів показав, що запропонована система дозволяє зменшити середню затримку передачі інформації, що є важливим фактором для критично важливих застосувань, таких як дистанційний моніторинг або системи безпеки. Порівняння з традиційними методами підтвердило переваги нового підходу, особливо в умовах високого рівня шумів, перевантаження мережі та обмеженого енергоресурсу вузлів.

У майбутніх дослідженнях планується розширити модель, додавши можливість врахування різних рівнів шуму та зовнішніх факторів, що впливають на якість передачі даних, а також розглянути можливість покращення відновлення даних за допомогою гібридних підходів, що поєднують машинне навчання з оптимізаційними алгоритмами.

Ключові слова: бездротові сенсорні мережі, відновлення даних, машинне навчання, Singular Spectrum Analysis, прогнозування, CRC, оптимізація маршрутизації.

SHYMCHUK GRIGORII**SHEVCHENKO NAZAR****SHVYRLO KONSTANTIN****GARMATYUK NAZAR**

Ternopil Ivan Puluj National Technical University

MACHINE LEARNING-BASED DATA RECOVERY SYSTEM FOR WIRELESS SENSOR NETWORKS

Data loss in wireless sensor networks (WSNs) is a serious issue that directly affects measurement accuracy, analytical correctness, and overall reliability of data transmission. In a constantly changing environment, the presence of interference, network congestion, or failure of individual nodes significantly degrades the quality of communication, leading to the loss of transmitted data packets.

Existing methods for mitigating data loss, such as cyclic redundancy check (CRC), interpolation algorithms, or packet retransmission schemes, have certain limitations. While they may be effective for minor losses, their accuracy and efficiency significantly decrease in environments with high interference or network instability.

This paper proposes a new data recovery system that combines classical integrity control methods with machine learning capabilities for more accurate reconstruction of missing values. The core of the approach is a mathematical model of a random cyclic process, which enables the analysis and forecasting of network behavior based on time series of previous observations. The use of Singular Spectrum Analysis (SSA) helps identify key trends, cyclic patterns, and random fluctuations in time series, allowing for more precise prediction of missing values. This methodology not only facilitates data recovery but also enables proactive loss prediction, minimizing the impact of transmission failures.

Experimental studies have demonstrated a significant reduction in data loss, improved overall network performance, and increased energy efficiency of nodes. The analysis of obtained results indicates that the proposed system reduces the average data transmission delay, which is a crucial factor for mission-critical applications such as remote monitoring and security systems. Comparisons with traditional methods confirm the advantages of this approach, particularly in high-noise conditions, network congestion, and limited node energy resources.

Future research aims to extend the model by incorporating different levels of noise and external factors affecting data transmission quality. Additionally, hybrid approaches that integrate machine learning with optimization algorithms will be explored to enhance data recovery efficiency.

Keywords: wireless sensor networks, data recovery, machine learning, Singular Spectrum Analysis, forecasting, CRC, routing optimization.

Стаття надійшла до редакції / Received 22.04.2025

Прийнята до друку / Accepted 16.04.2025

Постановка проблеми у загальному вигляді

та її зв'язок із важливими науковими чи практичними завданнями

Бездротові сенсорні мережі (WSN) відіграють ключову роль у багатьох сферах, зокрема в екологічному моніторингу, промисловій автоматизації, розумних містах та охороні здоров'я. Вони складаються з великої кількості сенсорних вузлів, які передають дані у реальному часі. Однак одним із найбільших викликів WSN є проблема втрати даних, яка може виникати через нестабільність зв'язку, перевантаженість мережі, зовнішні перешкоди або обмежені енергетичні ресурси вузлів.

Втрати даних можуть суттєво впливати на точність вимірювань та коректність аналізу отриманої інформації. У критичних застосуваннях, таких як моніторинг стану здоров'я пацієнтів або контроль промислових процесів, навіть невеликі втрати можуть призвести до серйозних наслідків. Традиційні методи компенсації втрат, такі як повторна передача пакетів або використання циклічного контрольного коду (CRC), є ефективними лише у випадках незначних втрат. Проте в умовах високого рівня шумів або обмежених ресурсів вони не забезпечують необхідної точності та ефективності.

Сучасні підходи до відновлення даних використовують методи прогнозування та машинного навчання. Вони дозволяють не лише відновлювати втрачені значення, але й передбачати можливі втрати, що дає змогу мінімізувати їхній вплив на функціонування мережі. Особливо перспективним є використання Singular Spectrum Analysis (SSA), який дає змогу виявляти закономірності у часових рядах та відновлювати втрачену інформацію.

Таким чином, існує потреба у розробці нової системи відновлення даних для WSN, яка поєднуватиме класичні методи контролю цілісності, моделі прогнозу та машинне навчання, що дозволить значно підвищити ефективність роботи мережі навіть за умов високих перешкод та обмежених ресурсів.

Аналіз досліджень та публікацій

Останні дослідження в галузі бездротових сенсорних мереж (WSN) підтверджують, що проблема втрати даних залишається серйозною, особливо коли умови передачі нестабільні. Щоб більше інформації передається через такі мережі, то важливішим стає питання її цілісності. Саме тому необхідні ефективні методи відновлення втрачених даних, які зможуть гнучко адаптуватися до змін у мережі та мінімізувати негативний вплив втрат на якість переданої інформації.

Традиційно для вирішення цієї проблеми використовуються такі методи, як циклічний контрольний код (CRC), механізми корекції помилок (FEC) і повторна передача пакетів (ARQ). Дослідження [1] показують, що CRC добре справляється з виявленням помилок, однак не здатний коригувати значні втрати даних. Водночас ARQ-схеми, наприклад, Hybrid ARQ (HARQ), хоч і зменшують вплив втрат, створюють додаткове навантаження на мережу і споживають більше енергії [2].

У зв'язку з цим дослідники все частіше звертаються до прогнозних моделей для відновлення даних. Наприклад, методи Singular Spectrum Analysis (SSA) дозволяють розкласти часові ряди на трендові та випадкові компоненти, що значно підвищує точність прогнозування втрачених значень [3]. Додатково застосовуються методи машинного навчання (ML), зокрема глибокі нейронні мережі (DNN) і регресійні моделі, що дає змогу адаптувати алгоритми відновлення до конкретних умов мережі [4].

Останні дослідження показують, що найкращі результати дають гібридні підходи, які поєднують статистичний аналіз і штучний інтелект. Наприклад, у роботах [5] розглядається комбінація машинного навчання та SSA для прогнозування і відновлення даних. Інші дослідники, такі як [6], пропонують використання байесових оцінок для підвищення точності прогнозування втрачених значень.

Однак, попри всі досягнення у використанні машинного навчання для відновлення даних, деякі методи мають серйозні обмеження, зокрема через високу обчислювальну складність. Алгоритми глибокого навчання потребують значних обчислювальних ресурсів, що не завжди підходить для сенсорних вузлів із низьким рівнем енергоспоживання [7]. Тому перспективним напрямом досліджень є розробка більш енергоефективних алгоритмів, які можуть забезпечити високу точність відновлення без значного навантаження на ресурси мережі.

Формулювання цілей статті

Метою роботи є: розробка ефективної системи відновлення втрачених даних у бездротових сенсорних мережах (WSN) шляхом поєднання класичних методів контролю цілісності з алгоритмами

машинного навчання та моделями прогнозу. Запропонований підхід спрямований на підвищення точності відновлення даних, зменшення затримок у їх передачі та покращення енергоефективності мережі.

Виклад основного матеріалу

Система відновлення даних у WSN починається з отримання інформації від сенсорів, які реєструють втрати пакетів та якість сигналу в мережі. Зібрані дані проходять етап попередньої обробки, під час якого відбувається фільтрація, нормалізація та усунення шумів, що можуть впливати на точність аналізу. Після цього виконується виявлення втрат даних, де аналізується наявність прогалин у переданій інформації та визначаються відсутні або пошкоджені пакети.

Наступним кроком є прогнозування відсутніх даних, під час якого використовуються математичні моделі або алгоритми машинного навчання для реконструкції неповної інформації. Потім механізм відновлення, який може включати циклічний контрольний код (CRC) або методи глибокого навчання, перевіряє правильність відновленої інформації та вносить коригування. Після цього відновлені дані передаються до вузлів призначення, де вони використовуються відповідно до потреб системи.

На етапі оцінки продуктивності аналізується ефективність відновлення даних, враховуючи точність, швидкість обробки та вплив на загальну продуктивність мережі. Отримані результати допомагають визначити можливі покращення, що реалізуються через механізм зворотного зв'язку. Зворотний зв'язок дозволяє оптимізувати параметри роботи системи, адаптуючи її до змін у середовищі передавання даних. Завершальним етапом є надсилання остаточно відновленої інформації кінцевому отримувачу, який отримує коректні та узгоджені дані для подальшого використання.

Структурна схема процесу відновлення втрачених даних приведена на рис. 1.



Рис.1 Структурна схема процесу відновлення втрачених даних у бездротовій сенсорній мережі

Запропонована система відновлення даних у бездротових сенсорних мережах (WSN) базується на гібридному підході, який поєднує класичні методи контролю цілісності даних (наприклад, CRC), математичне моделювання часових рядів та алгоритми машинного навчання. Основна ідея системи полягає у своєчасному виявленні втрат даних, прогнозуванні відсутніх значень та їх відновленні з високою точністю. Це дозволяє значно зменшити вплив втрат на якість інформації та підвищити стабільність роботи мережі.

Основні завдання, які вирішує система:

- Виявлення та ідентифікація відсутніх або пошкоджених пакетів даних.
- Використання математичного аналізу для прогнозування значень втрачених даних.
- Застосування алгоритмів машинного навчання для адаптації до змінних умов мережі.
- Оптимізація енергоспоживання для мінімізації обчислювальних витрат.

Нехай $X(t)$ – це часовий ряд значень, отриманих від сенсорних вузлів у WSN. Втрата даних означає, що у певний момент часу t значення $X(t)$ відсутнє або пошкоджене. Завдання полягає у

побудові оцінки $\hat{X}(t)$, яка є максимально точною апроксимацією реального значення $X(t)$, на основі попередніх та наступних вимірювань.

Математично проблему можна представити як:

$$\hat{X}(t) = f(X(t-1), X(t-2), \dots, X(t-n)) + \varepsilon \quad (1)$$

де f – функція прогнозування, що може бути реалізована за допомогою SSA або нейронної мережі, а ε – похибка оцінки.

Основні підходи для прогнозування:

- Методи часових рядів – Singular Spectrum Analysis (SSA), регресійні моделі.
- Методи машинного навчання – рекурентні нейронні мережі (RNN), глибоке навчання.

Запропонована система складається з таких основних компонентів:

1. Модуль збору даних – отримує значення з сенсорних вузлів та передає їх для аналізу.
2. Модуль виявлення втрат – аналізує отримані дані та ідентифікує відсутні значення.
3. Модуль прогнозування – виконує оцінку відсутніх значень на основі аналізу часових рядів та машинного навчання.
4. Модуль перевірки цілісності – застосовує CRC або інші методи для перевірки достовірності відновлених даних.
5. Модуль оптимізації енергоспоживання – адаптує алгоритми відновлення до ресурсних обмежень WSN.
6. Система управління мережею – приймає рішення щодо передачі відновлених даних та корекції маршрутизації.

Основним підходом у прогнозуванні є Singular Spectrum Analysis (SSA), який дозволяє представити часовий ряд у вигляді суми складових:

$$X(t) = S(t) + T(t) + R(t) \quad (2)$$

де $S(t)$ – основний тренд, $T(t)$ – періодичні циклічні компоненти, $R(t)$ – стохастичний шум.

Метод SSA дозволяє відновити трендову і циклічну компоненти, що дає змогу точніше оцінити втрачені значення. Для адаптивного підходу також використовується машинне навчання: нейронні мережі, які навчаються на історичних даних і прогнозують відсутні значення з урахуванням особливостей поточного мережевого середовища.

Завдяки запропонованій системі відновлення даних WSN може забезпечити високу надійність передачі інформації, що є критично важливим для додатків у розумних містах, промисловості, медичних системах та екологічному моніторингу.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

У статті розглянуто актуальну проблему втрати даних у бездротових сенсорних мережах (WSN), що суттєво впливає на надійність і точність функціонування таких систем. Запропонована система відновлення даних базується на гібридному підході, який поєднує класичні методи контролю цілісності з алгоритмами машинного навчання та аналізом часових рядів, зокрема методом Singular Spectrum Analysis (SSA). Такий підхід дозволяє не лише точно реконструювати втрачені значення, але й передбачати можливі втрати заздалегідь, що сприяє зменшенню негативного впливу збоїв на роботу мережі.

Результати експериментальних досліджень підтвердили ефективність запропонованої системи. Було досягнуто суттєвого зниження рівня втрат даних, зменшення середньої затримки передачі інформації, а також підвищення енергоефективності вузлів. Порівняння з традиційними методами показало явну перевагу нового підходу, особливо в умовах складного мережевого середовища з високим рівнем шуму та обмеженими ресурсами.

Таким чином, запропонована система є перспективним рішенням для підвищення надійності бездротових сенсорних мереж у критичних застосуваннях, таких як моніторинг довкілля, охорона здоров'я, розумні міста та промислові об'єкти. Подальші дослідження спрямовуватимуться на удосконалення моделі шляхом урахування змін зовнішніх факторів, підвищення адаптивності до умов мережі та застосування гібридних алгоритмів для ще більш ефективного відновлення втрачених даних.

Література

1. Kadel, R., Paudel, K., Guruge, D. B., & Halder, S. J. (2020). Opportunities and challenges for error control schemes for wireless sensor networks: A review. *Electronics*, 9(3), 504.
2. Ahmed, A., Al-Dweik, A., Iraqi, Y., Mukhtar, H., Naeem, M., & Hossain, E. (2021). Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey. *IEEE Communications Surveys & Tutorials*, 23(4), 2711-2752.
3. Lupenko, S., Lytvynenko, I., Nazarevych, O., Shymchuk, G., & Hotovych, V. (2021, December). Approach to gas consumption process forecasting on the basis of a mathematical model in the form of a random cyclic process. In *Proceedings of the International Conference „Advanced applied energy and information technologies 2021”*, 2021 (pp. 213-219). TNTU, Zhytomyr «Publishing house „Book-Druk “» LLC.
4. W. Njima, M. Chafii, A. Nimr and G. Fettweis, «Deep Learning Based Data Recovery for

Localization» in IEEE Access, vol. 8, pp. 175741-175752, 2020, doi: 10.1109/ACCESS.2020.3026615

5. P. Joshi and A. S. Raghuvanshi, «Hybrid Approaches to Address Various Challenges in Wireless Sensor Network for IoT Applications: Opportunities and Open Problems», IJCNA, Vol. 8, No. 3, p. 151, 2021, doi: 10.22247/ijcna/2021/209186.

6. H. Alamro et al., «Modeling of Bayesian-Based Optimized Transfer Learning Model for Cyber-Attack Detection in Internet of Things Assisted Resource Constrained Systems» in IEEE Access, vol. 12, pp. 177298-177311, 2024, doi: 10.1109/ACCESS.2024.3482876

7. X. Liu, J. Li, Z. Dong and F. Xiong, «Joint Design of Energy-Efficient Clustering and Data Recovery for Wireless Sensor Networks» in IEEE Access, vol. 5, pp. 3646-3656, 2017, doi: 10.1109/ACCESS.2017.2660770

References

1. Kadel, R., Paudel, K., Guruge, D. B., & Halder, S. J. (2020). Opportunities and challenges for error control schemes for wireless sensor networks: A review. *Electronics*, 9(3), 504.

2. Ahmed, A., Al-Dweik, A., Iraqi, Y., Mukhtar, H., Naeem, M., & Hossain, E. (2021). Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey. *IEEE Communications Surveys & Tutorials*, 23(4), 2711-2752.

3. Lupenko, S., Lytvynenko, I., Nazarevych, O., Shymchuk, G., & Hotovych, V. (2021, December). Approach to gas consumption process forecasting on the basis of a mathematical model in the form of a random cyclic process. In *Proceedings of the International Conference „Advanced applied energy and information technologies 2021”*, 2021 (pp. 213-219). TNTU, Zhytomyr «Publishing house „Book-Druk “» LLC.

4. W. Njima, M. Chafii, A. Nimr and G. Fettweis, «Deep Learning Based Data Recovery for Localization» in IEEE Access, vol. 8, pp. 175741-175752, 2020, doi: 10.1109/ACCESS.2020.3026615

5. P. Joshi and A. S. Raghuvanshi, «Hybrid Approaches to Address Various Challenges in Wireless Sensor Network for IoT Applications: Opportunities and Open Problems», IJCNA, Vol. 8, No. 3, p. 151, 2021, doi: 10.22247/ijcna/2021/209186.

6. H. Alamro et al., «Modeling of Bayesian-Based Optimized Transfer Learning Model for Cyber-Attack Detection in Internet of Things Assisted Resource Constrained Systems» in IEEE Access, vol. 12, pp. 177298-177311, 2024, doi: 10.1109/ACCESS.2024.3482876

7. X. Liu, J. Li, Z. Dong and F. Xiong, «Joint Design of Energy-Efficient Clustering and Data Recovery for Wireless Sensor Networks» in IEEE Access, vol. 5, pp. 3646-3656, 2017, doi: 10.1109/ACCESS.2017.2660770