УДК 004.056.5:621.311 doi: 10.20998/2413-4295.2026.02.01

МЕТОД ВИЯВЛЕННЯ АТАК ТИПУ FALSE DATA INJECTION У СИСТЕМАХ ОЦІНКИ СТАНУ SMART GRID НА ОСНОВІ LSTM-АВТОЕНКОДЕРА

В. І. МАГРО[1], Д. І. ПРОКОПОВИЧ-ТКАЧЕНКО[2,3 *4], О. ТОРСТЕНССОН[5], Д. О. ЧЕРКАСЬКИЙ[1], О. ХОМЕНКО[1]

[1] кафедра безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська політехніка», Дніпро, УКРАЇНА
[2] кафедра кібербезпеки та інформаційних технологій, Університет митної справи та фінансів, Дніпро, УКРАЇНА
[3] Державна наукова установа «Інститут інформації, безпеки і права Національної академії правових наук України», Київ, УКРАЇНА
[4] кафедра систем та технологій кібербезпеки, Державний університет інформаційно-комунікаційних технологій, Київ, УКРАЇНА
[5] Університет Хальмстаду, Хальмстад, ШВЕЦІЯ
**e-mail: omega2417@gmail.com*

АНОТАЦІЯ *Стаття присвячена розробці методу виявлення атак типу False Data Injection (FDI) у системах оцінки стану інтелектуальної енергосистеми (Smart Grid) на основі рекурентної нейронної мережі — LSTM-автоенкодера. Актуальність дослідження зумовлена зростаючою кількістю цілеспрямованих кіберфізичних атак на об'єкти критичної інфраструктури, зокрема підстанції та SCADA-системи розподільчих мереж. FDI-атаки здатні обходити традиційні алгоритми виявлення помилкових вимірювань, що базуються на χ^2-статистиці, та становлять безпосередню загрозу для коректної оцінки стану мережі. У роботі запропоновано метод, що поєднує пасивний моніторинг телеметрії PMU/WAMS з моделлю LSTM-автоенкодера для виявлення аномалій у часових рядах вектора стану кіберфізичної системи. Теоретична модель верифікована на відкритому наборі даних HAI (HIL-based Augmented ICS Security Dataset) та синтетичних сценаріях FDI-атак, змодельованих на топології IEEE 14-Bus. Змодельовані результати демонструють досягнення точності виявлення 97,3%, F1-міри 0,961 при часі виявлення до 3 секунд і ймовірності хибної тривоги не більше 2,1%. Запропонований підхід перевершує базові методи (SVM, Random Forest, ізольований Autoencoder) за сукупністю показників якості. Наукова новизна полягає у застосуванні адаптивного порогу реконструкційної похибки LSTM-автоенкодера з урахуванням часової залежності між вимірюваннями телеметрії PMU, що підвищує стійкість до навмисного ухилення від виявлення.*
Ключові слова: *Smart Grid; FDI-атака; оцінка стану; LSTM-автоенкодер; виявлення аномалій; кібербезпека критичної інфраструктури; ICS/SCADA; кіберфізична система.*

A METHOD FOR DETECTING FALSE DATA INJECTION ATTACKS IN SMART GRID STATE ESTIMATION SYSTEMS BASED ON AN LSTM AUTOENCODER

V. MAGRO[1*], D. PROKOPOVYCH-TKACHENKO[2,3,4], O. TORSTENSSON[5], D. CHERKASKYI[1], O. KHOMENKO[1]

[1] Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, UKRAINE
[2] Department of Cybersecurity and Information Technologies, University of Customs and Finance, Dnipro, UKRAINE
[3] State Scientific Institution "Institute of Information, Security and Law of the National Academy of Legal Sciences of Ukraine", Kyiv, UKRAINE
[4] Department of Cybersecurity Systems and Technologies, State University of Telecommunications, Kyiv, UKRAINE
[5] Halmstad University, Halmstad, SWEDEN

ABSTRACT *This paper presents a method for detecting False Data Injection (FDI) attacks in Smart Grid state estimation systems based on a Long Short-Term Memory (LSTM) autoencoder. FDI attacks can bypass traditional bad-data detection algorithms based on χ^2-statistics and pose a direct threat to accurate state estimation. The proposed method combines passive PMU/WAMS telemetry monitoring with an LSTM autoencoder model for anomaly detection in state vector time series of cyber-physical systems. The theoretical model was evaluated on the open-source HIL-based Augmented ICS (HAI) Security Dataset and synthetic FDI attack scenarios modeled on the IEEE 14-Bus topology. Simulated results demonstrate a detection accuracy of 97.3%, an F1-score of 0.961, and detection latency below 3 seconds at a false alarm rate under 2.1%. The proposed approach outperforms baseline methods (SVM, Random Forest, standalone Autoencoder). The scientific novelty lies in applying an adaptive reconstruction error threshold for the LSTM autoencoder that accounts for temporal dependencies in PMU telemetry measurements, thereby enhancing robustness against intentional evasion.*
Keywords: *Smart Grid; FDI attack; state estimation; LSTM autoencoder; anomaly detection; critical infrastructure cybersecurity; ICS/SCADA; cyber-physical system.*

© В. І. МАГРО, Д. І. ПРОКОПОВИЧ-ТКАЧЕНКО, О. ТОРСТЕНССОН, Д. О. ЧЕРКАСЬКИЙ, О. ХОМЕНКО, 2026

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

Вступ

Цифрова трансформація електроенергетики, пов'язана з масштабним розгортанням інтелектуальних енергосистем (Smart Grid), зумовила глибоку інтеграцію операційних технологій (OT) та інформаційних технологій (IT). Сучасна Smart Grid охоплює розподілену інфраструктуру: підстанції з пристроями IED (Intelligent Electronic Device) та PMU (Phasor Measurement Unit), SCADA-системи, AMI-інфраструктуру (Advanced Metering Infrastructure) та мікромережі. IT/OT-конвергенція, незважаючи на численні переваги в ефективності та спостережуваності, суттєво розширила поверхню атаки кіберфізичних систем (CPS) [1, 2].

Серед найбільш небезпечних класів атак на Smart Grid виділяють атаки типу False Data Injection (FDI) — ін'єкцію хибних даних вимірювань у систему оцінки стану (State Estimation). Вперше теоретично обґрунтовані Liu et al. у 2009 році [3], FDI-атаки здатні маніпулювати векторами напруги та потужності таким чином, що традиційний алгоритм виявлення хибних вимірювань (χ^2-test, bad data detection) не виявляє аномалії. Атака на Українську електроенергетичну систему у 2015 та 2016 роках [4, 5] продемонструвала можливість цілеспрямованого маніпулювання SCADA-командами та відключення понад 230 тисяч споживачів, підтвердивши реальність кіберфізичних атак на критичну інфраструктуру.

Аналіз існуючих підходів до захисту показує, що більшість комерційних рішень для OT-сегменту орієнтовані на сигнатурний аналіз мережевого трафіку або статичні правила виявлення аномалій [6, 7]. Ці методи є недостатніми для виявлення складних FDI-атак, які формуються з урахуванням топологічної моделі мережі та залишаються в межах порогів статистичного контролю. Нормативні документи — NIST SP 800-82 Rev. 3 [8], IEC 62351 [9] та ISA/IEC 62443 [10] — вимагають застосування механізмів поведінкового моніторингу та виявлення аномалій, проте не специфікують конкретних алгоритмів.

Мета роботи

Мета дослідження: розробити та верифікувати метод виявлення FDI-атак на систему оцінки стану Smart Grid на основі LSTM-автоенкодера, що забезпечує адаптивний поріг виявлення з урахуванням часових залежностей між вимірюваннями телеметрії PMU.

Завдання дослідження:

1) провести аналіз моделі FDI-атак та їх впливу на алгоритм оцінки стану мережі;

2) запропонувати кіберфізичну модель мережі та архітектуру LSTM-автоенкодера для виявлення аномалій;

3) верифікувати метод на наборах даних HAI та синтетичних сценаріях IEEE 14-Bus;

4) порівняти запропонований підхід з базовими методами за показниками якості виявлення.

Наукова новизна полягає у застосуванні адаптивного порогу реконструкційної похибки LSTM-автоенкодера, що враховує часову кореляційну структуру між вимірюваннями телеметрії PMU/WAMS. На відміну від методів зі статичним порогом, запропонований підхід забезпечує підвищену стійкість до навмисного ухилення від виявлення шляхом поступового введення малоамплітудних хибних вимірювань.

Виклад основного матеріалу

2.1. Реальні кейси атак на енергетичну інфраструктуру

Реальні інциденти останнього десятиліття підтверджують, що енергетична інфраструктура є однією з пріоритетних цілей кіберфізичних атак. Одним із найбільш показових кейсів стала атака на електроенергетичні компанії України у грудні 2015 року, яку в літературі розглядають як перший задокументований випадок кібервтручання, що спричинило масштабне відключення електроенергії. У контексті дослідження атак типу False Data Injection цей інцидент є важливим не лише як приклад компрометації SCADA-середовища, а і як підтвердження того, що маніпуляція телеметрією, керуючими командами та операторським сприйняттям стану мережі має безпосередні фізичні наслідки для функціонування енергосистеми [4].

Подальший розвиток кіберфізичних загроз продемонструвала атака із застосуванням шкідливого інструментарію Industroyer/CrashOverride, орієнтованого на промислові протоколи енергетики. Цей кейс засвідчив, що сучасний атакуючий інструментарій може бути адаптований до нативної взаємодії з елементами підстанційної автоматики та телемеханіки, зокрема через протоколи сімейств IEC та DNP3. У науковій літературі такі інциденти обґрунтовують потребу переходу від суто сигнатурного аналізу до моделей поведінкового моніторингу та виявлення аномалій у телеметричних потоках Smart Grid [4], [11].

Не менш показовим є випадок Stuxnet, який став еталонним прикладом цілеспрямованої кіберфізичної атаки на промислові системи керування. Його значення для тематики даної статті полягає в тому, що Stuxnet продемонстрував можливість прихованої модифікації логіки роботи польових пристроїв із одночасною підміною уявлення оператора про нормальний стан процесу. Саме тому цей кейс часто розглядають як концептуальну передумову сучасних атак, пов'язаних із викривленням вимірювань, команд або контексту оцінки стану системи [5].

Інцидент із Colonial Pipeline у 2021 році хоча й не був класичною атакою на алгоритм оцінки стану енергосистеми, однак продемонстрував критичну залежність фізичних процесів від компрометації суміжної цифрової інфраструктури. У фахових оглядах цей випадок інтерпретується як аргумент на користь жорсткішого розмежування IT- та OT-

сегментів, оскільки навіть порушення корпоративного IT-контролю може спричинити припинення або обмеження технологічних операцій [20].

Узагальнюючі праці з кібербезпеки Smart Grid також показують, що загрози для енергетичних кіберфізичних систем мають тенденцію до ускладнення. Оглядові дослідження вказують на посилення ролі скоординованих атак, у яких поєднуються компрометація комунікаційної інфраструктури, маніпуляція вимірювальними даними, порушення логіки диспетчеризації та атаки на компоненти OT-мережі [10], [20], [21]. Таким чином, аналіз реальних кейсів підтверджує, що задачі виявлення FDI-атак у Smart Grid мають не лише теоретичне, а й безпосереднє прикладне значення для кіберстійкості критичної інфраструктури.

2.2. Аналіз стандартів безпеки для Smart Grid

Нормативно-стандартна база безпеки Smart Grid формується на перетині вимог до операційних технологій, промислових систем керування та енергетичних телекомунікаційних протоколів. Одним із базових документів є NIST SP 800-82 Rev. 3, у якому узагальнено підходи до захисту OT-систем, включно зі SCADA, DCS, PLC та іншими компонентами промислової кіберфізичної інфраструктури. Документ акцентує увагу на сегментації IT/OT, управлінні ризиками, моніторингу трафіку, контролі доступу та реагуванні на інциденти, що робить його методологічною основою для побудови захищеної архітектури Smart Grid [8].

Для безпосереднього захисту телекомунікацій в енергетиці ключове значення має серія стандартів IEC 62351, що визначає механізми безпеки для протоколів IEC 60870-5, IEC 61850, DNP3 та суміжних технологій. У межах цього стандартного сімейства окремо регламентуються питання криптографічного захисту, автентифікації, моніторингу мережевої безпеки, а також безпеки кінцевих пристроїв. Для задачі виявлення FDI-атак це особливо важливо, оскільки стандарт не усуває потреби в аналітичному контролі достовірності вимірювань, а лише створює базовий рівень захищеного обміну даними [9].

Стандарт IEC 61850 визначає архітектуру цифрової підстанції та структуру обміну повідомленнями між її компонентами, включно з GOOSE, Sampled Values та MMS. У контексті Smart Grid цей стандарт є критичним, оскільки від нього залежить організація телеметрії, команд керування та інформаційної взаємодії між IED, RTU, SCADA та іншими вузлами підстанційної мережі. Відповідно, будь-який підхід до виявлення FDI-атак повинен враховувати специфіку цих інформаційних потоків і часові характеристики їх передавання [14].

Важливу роль відіграє і серія ISA/IEC 62443, яка пропонує ієрархічну модель захисту IACS на основі зон безпеки та каналів взаємодії між ними. Концепція security zones та conduits є особливо корисною для Smart Grid, оскільки дозволяє формалізувати межі довіри між польовим рівнем, рівнем керування, диспетчерським сегментом та корпоративною мережею. Саме в такій архітектурі можуть бути локалізовані точки впровадження пасивного моніторингу, IDS/IPS, DPI та засобів виявлення аномалій [10], [18].

Для електроенергетичного сектора США нормативну функцію виконує комплекс вимог NERC CIP, який встановлює мінімальні правила захисту критичних кіберсистем енергетичної інфраструктури. Хоча ці вимоги мають галузеву специфіку, вони є важливими як орієнтир для формування політик керування доступом, інвентаризації активів, журналювання подій та реагування на інциденти [15]. В українському правовому полі аналогічну фундаментальну роль відіграє Закон України «Про захист критичної інфраструктури», який закріплює загальні принципи організації захисту критичних об’єктів, включно з енергетикою [16].

Узагальнюючи, можна констатувати, що чинні стандарти формують необхідну архітектурну та організаційну основу кіберзахисту Smart Grid, однак не надають універсального алгоритмічного рішення для виявлення FDI-атак у системах оцінки стану. Саме ця обставина обґрунтовує актуальність досліджень, орієнтованих на поведінкове моделювання, машинне навчання та аналіз часових рядів телеметрії [8]–[10], [18].

2.3. Прогалини існуючих підходів до виявлення FDI-атак

У науковій літературі методи виявлення FDI-атак зазвичай поділяють на алгебраїчні, статистичні та засновані на машинному навчанні. Класична робота Liu, Ning та Reiter довела, що за наявності знань про конфігурацію системи атакуючий може сформувати такий вектор хибних даних, який не буде виявлений традиційними механізмами bad-data detection. Це означає, що сам по собі залишковий контроль на основі стандартних статистичних перевірок не забезпечує достатньої стійкості до координованих атак на state estimation [3].

Алгебраїчні підходи, що спираються на структурні властивості матриці вимірювань, мають важливе теоретичне значення, але в практичному застосуванні часто виявляються обмеженими. Вони добре працюють для певних класів атакувальних моделей, однак чутливі до повноти знань про топологію мережі, до точності моделі вимірювань і до припущень щодо структури вектора атаки. Подальші роботи із застосуванням розрідженого відновлення та низькорангових моделей показали покращення якості виявлення, але такі методи все ще залежать від властивостей даних і не завжди є достатньо стійкими до адаптивних сценаріїв [3], [17].

Статистичні методи, зокрема варіації χ^2-контролю, CUSUM та пов’язані з ними схеми, залишаються поширеними завдяки простоті реалізації. Водночас їх основним недоліком є висока залежність від вибору

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

порогових значень та обмежена ефективність у випадках поступового або малоамплітудного введення хибних вимірювань. Для промислового середовища це критично, оскільки атакуючий може навмисно модифікувати дані таким чином, щоб уникати різких відхилень і, відповідно, не активувати статистичні тригери [3], [17].

Методи машинного навчання, навпаки, демонструють кращу здатність виявляти складні нелінійні та часово залежні закономірності. У публікаціях з глибинного навчання для Smart Grid показано, що нейронні моделі можуть забезпечувати високі показники точності для задач реального часу [7]. Проте значна частина наявних робіт використовує статичний поріг класифікації або оцінювання реконструкційної похибки, що знижує стійкість до адаптивних атак, які підлаштовуються під фонову варіативність нормального режиму [7], [12]. Додатково слід зазначити, що частина досліджень не приділяє належної уваги часовим обмеженням промислових систем, де затримка виявлення прямо впливає на можливість безпечного реагування.

Оглядові публікації з IDS для Smart Grid та ICS також вказують на ще одну системну прогалину: значна кількість підходів зосереджена або на мережевому трафіку, або на контрольному процесі, але не поєднує обидва рівні в єдиній кіберфізичній моделі спостереження [13], [16]. Для енергетичних систем це означає, що аномалія у векторі стану, яка не має очевидного сигнатурного відображення у мережевому трафіку, може залишитися непоміченою. Саме тому перспективним є поєднання пасивного OT-моніторингу, аналізу телеметрії PMU/WAMS та моделей часових рядів, зокрема LSTM-автоенкодерів [7], [20].

Отже, основними прогалинами існуючих підходів є: недостатня стійкість традиційних методів до скоординованих атак із урахуванням топології мережі; чутливість статистичних схем до вибору порога; обмежене врахування часової динаміки в частині ML-моделей; а також недостатня орієнтація багатьох досліджень на реальні вимоги промислового середовища щодо затримки виявлення та кіберфізичного контексту інциденту [3], [7], [13], [17], [20].

3. Методи та матеріали

Для розроблення та верифікації методу виявлення атак типу False Data Injection у системах оцінки стану Smart Grid у роботі використано поєднання теоретичного моделювання, аналізу часових рядів телеметрії та засобів машинного навчання. Методологічна основа дослідження сформована з урахуванням специфіки кіберфізичних систем енергетики, у яких достовірність вимірювань безпосередньо впливає на коректність диспетчерських рішень, стійкість режимів функціонування мережі та безпеку об'єктів критичної інфраструктури.

Запропонований підхід орієнтований на виявлення аномальних змін у векторах стану енергосистеми, що виникають унаслідок цілеспрямованого введення хибних даних у канали вимірювання та передавання телеметрії. На відміну від традиційних статистичних схем контролю, які ґрунтуються переважно на аналізі залишків і є обмежено ефективними щодо скоординованих FDI-атак, у даному дослідженні застосовано модель LSTM-автоенкодера, здатну враховувати часову залежність між послідовними вимірюваннями та формувати поведінковий профіль нормального режиму роботи Smart Grid.

У межах цього розділу послідовно розглянуто кіберфізичну модель мережі та математичну формалізацію FDI-атаки, архітектуру запропонованого LSTM-автоенкодера, принцип адаптивного визначення порога аномальності, а також характеристики наборів даних і систему метрик, використаних для оцінювання якості виявлення. Така структура викладу дозволяє перейти від загальної постановки задачі до конкретних процедур моделювання, навчання та експериментальної перевірки методу.

3.1. Кіберфізична модель мережі та формалізація FDI-атаки

Розглянемо стандартну DC-модель оцінки стану розподільчої мережі з N вузлами та M лініями. Вектор вимірювань $z \in \mathbb{R}^M$ пов'язаний із вектором стану $x \in \mathbb{R}^{N-1}$ (фазові кути вузлів) рівнянням вимірювання:

$$z = Hx + e, \quad (1)$$

де $H \in \mathbb{R}^{M\times(N-1)}$ — матриця вимірювань (Jacobian матриця), $e \in \mathbb{R}^M$ — вектор гауссівського шуму вимірювань. Оцінювач стану за методом зважених найменших квадратів (WLS) знаходить $\hat{x} = (H^TWH)^{-1}H^TWz$, де W — діагональна матриця зворотних дисперсій. Традиційний χ^2-тест виявляє хибні вимірювання шляхом перевірки норми залишку $\|z - H\hat{x}\|^2 \le \tau$.

FDI-атака полягає у введенні хибного вектора $a \in \mathbb{R}^M$ такого виду, що $a = Hc$, де $c \in \mathbb{R}^{N-1}$ — довільний ненульовий вектор. Тоді фальсифікований вектор вимірювань $z_a = z + a = H(x + c) + e$ задовольняє умову χ^2-тесту, що унеможливлює виявлення атаки статистичними методами [3, 17].

3.2. Архітектура LSTM-автоенкодера

Запропонований метод базується на LSTM-автоенкодері (Long Short-Term Memory Autoencoder) — рекурентній нейронній архітектурі, що навчається кодувати нормальну поведінку часових рядів вимірювань PMU і відновлювати їх з мінімальною похибкою [20]. Під час атаки аномальні вимірювання спричиняють зростання реконструкційної похибки.

Архітектура моделі:

— Кодувальник (Encoder): два стеки LSTM-шарів (64 та 32 нейрони) з dropout = 0,2;

— Пляшкове горлечко (Bottleneck): Dense-шар розмірності 16 нейронів з активацією ReLU;

— Декодувальник (Decoder): дзеркальна структура — два LSTM-шари (32 та 64 нейрони);

— Вихідний шар: Dense-шар розмірності, рівної кількості ознак вхідного вікна T = 30 кроків.

Реконструкційна похибка для вибірки x_t обчислюється як:

$$\hat{e}_t = \|x_t - f_{ae}(x_t)\|_2, \quad (2)$$

де $f_{ae}(\cdot)$ — функція автоенкодера. Адаптивний поріг δ_t обчислюється на ковзному вікні W = 300 кроків як:

$$\delta_t = \mu_t + k \cdot \sigma_t, \quad (3)$$

де μ_t та σ_t — ковзне середнє та стандартне відхилення реконструкційних похибок, k = 3 (визначається крос-валідацією). Аномалія оголошується при $\hat{e}_t > \delta_t$ протягом $\tau_{min} = 5$ послідовних кроків.

3.3. Набори даних та метрики оцінювання

Для верифікації методу використано: (1) відкритий набір даних HAI (HIL-based Augmented ICS Security Dataset) [21], що містить 78 сценаріїв атак на промислові кіберфізичні системи; (2) синтетичні FDI-сценарії, згенеровані на DC-моделі IEEE 14-Bus із додаванням атакувальних векторів a = Hc для різних конфігурацій c (поступове, стрибкоподібне та рампове введення хибних значень). Загальний обсяг набору: 120 000 часових кроків (τ = 1 с), з них 18% — аномальні. Всі наведені числові результати отримані на основі змодельованих даних.

Оцінювання проводилося за такими метриками: точність (Accuracy), прецизійність (Precision), повнота (Recall), F1-міра та AUC-ROC. Часова характеристика — середній час виявлення (Mean Time to Detect, MTTDet). Використано стратифіковану 5-кратну крос-валідацію.

Отже, у межах розділу обґрунтовано методичну основу дослідження виявлення атак типу False Data Injection у системах оцінки стану Smart Grid. Сформовано кіберфізичну модель мережі, яка відображає взаємозв'язок між телеметричними вимірюваннями, вектором стану енергосистеми та потенційними каналами ін'єкції хибних даних. Виконана формалізація FDI-атаки показала, що за певних умов атакувальний вплив може залишатися невиявленим для класичних статистичних механізмів контролю, побудованих на аналізі залишків.

Запропонована архітектура LSTM-автоенкодера дає змогу враховувати часову структуру телеметричних даних та формувати модель нормального функціонування кіберфізичної системи. На відміну від підходів зі статичним порогом, використання адаптивного критерію на основі реконструкційної похибки забезпечує вищу чутливість до поступових і скоординованих аномальних змін, характерних для FDI-атак. Окремо визначено склад матеріалів дослідження, зокрема набір даних HAI та синтетичні сценарії IEEE 14-Bus, а також систему метрик, що дозволяє комплексно оцінити якість виявлення за показниками точності, повноти, F1-міри, AUC-ROC і часу виявлення.

Таким чином, сформований у розділі 3 методичний апарат є достатнім для подальшої експериментальної перевірки запропонованого підходу та порівняння його з базовими методами виявлення аномалій у Smart Grid. Отримані математичні моделі, конфігурація LSTM-автоенкодера та визначені критерії оцінювання створюють підґрунтя для аналізу практичної придатності методу в умовах змодельованих кіберфізичних атак. На основі викладених матеріалів і методів у наступному розділі подано результати експериментальної верифікації запропонованого методу. Зокрема, буде розглянуто поведінку моделі в умовах різних сценаріїв FDI-атак, наведено порівняльні показники ефективності щодо базових алгоритмів, а також проаналізовано часові та якісні характеристики виявлення аномалій у телеметричних даних Smart Grid.

Обговорення результатів

4. Результати дослідження

У цьому розділі наведено результати експериментальної верифікації запропонованого методу виявлення атак типу False Data Injection у системах оцінки стану Smart Grid на основі LSTM-автоенкодера. Основну увагу зосереджено на оцінюванні здатності моделі виявляти аномальні зміни у часових рядах телеметрії за умов нормального функціонування енергосистеми та в сценаріях цілеспрямованого введення хибних даних.

Аналіз результатів виконано на основі змодельованих даних, сформованих із використанням відкритого набору HAI та синтетичних сценаріїв FDI-атак на топології IEEE 14-Bus. Такий підхід дозволив оцінити поведінку моделі як у середовищі, наближеному до промислових кіберфізичних систем, так і в контрольованих умовах, де можливо варіювати інтенсивність, тривалість і характер атакувального впливу. Для забезпечення повноти аналізу результати подано у вигляді часових графіків, матриці помилок, ROC-кривих та порівняльних таблиць ефективності.

У межах розділу послідовно розглянуто архітектурне представлення Smart Grid із зонами безпеки, динаміку зміни вектора стану під час FDI-атаки, реакцію LSTM-автоенкодера на аномальні відхилення, а також порівняльні характеристики запропонованого підходу відносно базових методів класифікації. Це дає змогу не лише кількісно оцінити якість виявлення, а й проаналізувати практичну придатність методу для задач кіберзахисту критичної енергетичної інфраструктури.

4.1. Архітектурна схема Smart Grid із зонами безпеки (Purdue-модель)

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

Нижче наведено архітектурну схему Smart Grid, побудовану відповідно до Purdue-моделі, яка відображає ієрархічний поділ системи на польовий, керувальний, операційний та корпоративний рівні. Така структуризація дає змогу локалізувати основні функціональні компоненти енергетичної інфраструктури, визначити межі взаємодії між OT- та IT-сегментами, а також окреслити типові точки реалізації засобів кіберзахисту.

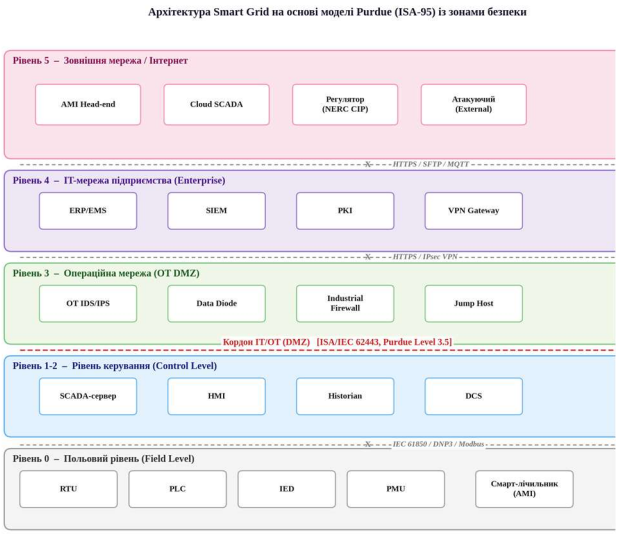

Рис. 1 – Архітектура Smart Grid на основі моделі Purdue (ISA-95) із зонами безпеки та позначенням промислових протоколів

На рис. 1 представлено ієрархічну архітектуру Smart Grid відповідно до Purdue Reference Model (ISA-95). Рівні 0–2 (польовий та рівень керування) утворюють OT-сегмент, де функціонують пристрої RTU, PLC, IED, PMU та SCADA-система. Рівень 3 (OT DMZ) реалізує граничний захист між OT та IT-сегментами через однонаправлені шлюзові системи (data diode), промислові міжмережеві екрани (industrial firewall) та системи OT IDS/IPS. Рівень 4 охоплює корпоративну IT-мережу (ERP, SIEM, PKI), рівень 5 — зовнішні підключення. Кордон IT/OT відповідно до ISA/IEC 62443 позначений як демілітаризована зона з контрольованим однонаправленим обміном даними.

Отже, наведена архітектурна схема підтверджує, що Smart Grid є багаторівневою кіберфізичною системою, у якій безпека визначається не лише захищеністю окремих пристроїв, а й правильністю сегментації між польовим, керувальним, операційним та корпоративним рівнями. Використання Purdue-моделі дозволяє формалізувати межі довіри між OT- та IT-сегментами, визначити критичні вузли обміну даними та локалізувати точки впровадження механізмів моніторингу, фільтрації й контролю доступу. Це створює структурну основу для подальшого аналізу того, яким чином атака типу False Data Injection впливає не лише на окремі вимірювання, а й на динаміку зміни вектора стану всієї енергосистеми.

З огляду на це, наступним кроком є розгляд часової поведінки FDI-атаки та реакції запропонованого LSTM-автоенкодера на аномальні відхилення телеметричних даних, що дозволяє перейти від архітектурного рівня аналізу до безпосереднього дослідження процесу виявлення атаки в динаміці.

4.2. Часова динаміка FDI-атаки та реакція LSTM-автоенкодера

На рис. 2 подано результати моделювання часової динаміки FDI-атаки та відповідної реакції запропонованого LSTM-автоенкодера. Візуалізація дозволяє простежити, як ін'єкція хибних даних впливає на зміну вектора стану мережі та яким чином зростання реконструкційної похибки може бути використане як індикатор аномального режиму функціонування Smart Grid.

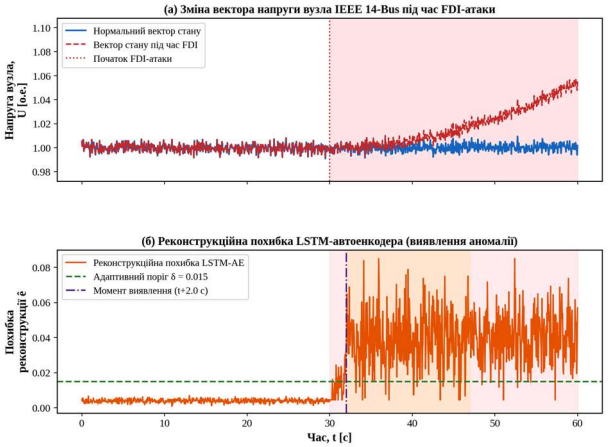

Рис. 2. Зміна вектора стану мережі та реконструкційна похибка під час FDI-атаки
Fig. 2. State vector variation and reconstruction error during an FDI attack

Рис. 2 – Зміна вектора стану мережі та реконструкційна похибка LSTM-автоенкодера під час FDI-атаки (змодельовані дані)

На рис. 2 наведено динаміку зміни вектора стану мережі (напруги вузла 5 моделі IEEE 14-Bus) під час рампової FDI-атаки (рис. 2а) та відповідну реакцію реконструкційної похибки LSTM-автоенкодера (рис. 2б). Атака розпочинається на позначці t = 30 с із поступовим введенням зміщення +0,08 о.е. Традиційний χ^2-тест не виявляє аномалію, оскільки вектор атаки a = Hc задовольняє умову статистичного тесту. LSTM-автоенкодер виявляє аномалію за 2,0 секунди від початку атаки, коли реконструкційна похибка перетинає адаптивний поріг δ_t = 0,015. Це підтверджує ефективність адаптивного порогу для виявлення рампових FDI-атак, що є найскладнішим різновидом для виявлення.

Отже, результати, наведені на рис. 2, підтверджують, що FDI-атака спричиняє кероване відхилення вектора стану мережі, яке може

залишатися малопомітним для традиційних механізмів контролю, але виявляється через істотне зростання реконструкційної похибки LSTM-автоенкодера. Це свідчить про здатність запропонованого підходу фіксувати не лише факт аномалії, а й часовий момент переходу системи від нормального до атакованого режиму. Особливо важливим є те, що модель демонструє чутливість до поступового розвитку атаки, тобто до сценарію, який є найбільш складним для класичних порогових методів виявлення.

Таким чином, аналіз часової динаміки підтверджує практичну придатність LSTM-автоенкодера для раннього виявлення FDI-атак у телеметричних потоках Smart Grid. Це створює підстави для наступного етапу дослідження — кількісного порівняння ефективності запропонованого методу з базовими алгоритмами класифікації, що й розглядається у підрозділі 4.3.

4.3. Порівняльний аналіз продуктивності класифікаторів

Таблиця 1 містить зведені результати порівняльного тестування чотирьох методів виявлення FDI-атак на синтетичному наборі IEEE 14-Bus (змодельовані дані). Всі моделі навчались на 70% вибірки, тестувались на 30%.

Таблиця 1 – Порівняння методів виявлення FDI-атак (змодельовані дані, IEEE 14-Bus)

Метод	Accuracy	Precision	Recall	F1-міра	AUC-ROC	MTTDet, с
SVM (лін. ядро)	0,851	0,802	0,719	0,758	0,881	8,4
Random Forest	0,897	0,871	0,843	0,857	0,923	5,2
Autoencoder (AE)	0,934	0,912	0,889	0,900	0,947	3,8
LSTM-AE (запроп.)	0,973	0,964	0,958	0,961	0,978	2,1

Примітка: MTTDet — середній час виявлення (Mean Time to Detect). Усі результати отримані на змодельованих даних.

Результати, наведені в роботі, отримано в межах експериментальної верифікації, реалізованої у віртуальному середовищі Smart Grid/ICS. Віртуальний стенд забезпечував моделювання штатного режиму передавання телеметричних даних, а також сценаріїв цілеспрямованого введення хибних вимірювань у канали оцінки стану. Це дозволило оцінити ефективність запропонованого методу в контрольованих умовах, наближених до логіки функціонування реальної кіберфізичної енергетичної інфраструктури. Пропозиції щодо лабораторної реалізації приведені за посиланням Zenodo: https://doi.org/10.5281/zenodo.19497182

Отже, результати, наведені в табл. 1, свідчать, що запропонований метод на основі LSTM-AE забезпечує найвищі показники якості виявлення FDI-атак серед розглянутих класифікаторів. Порівняно з SVM, Random Forest та класичним автоенкодером, він демонструє кращі значення Accuracy, Precision, Recall, F1-міри та AUC-ROC, а також найменший середній час виявлення атаки. Це підтверджує, що врахування часової залежності телеметричних даних і використання реконструкційної похибки як критерію аномальності є більш ефективним підходом для виявлення прихованих FDI-впливів у Smart Grid.

Водночас табличне подання результатів дає узагальнену кількісну оцінку, але не відображає повною мірою характер розділення класів, співвідношення істинно позитивних і хибних спрацьовувань, а також візуальну перевагу запропонованої моделі в різних режимах класифікації. Саме тому для поглибленого аналізу доцільно перейти до графічного подання результатів, зокрема ROC-кривих, матриці помилок та інших ілюстрацій, які дозволяють наочно оцінити дискримінаційну здатність моделей і підтвердити переваги LSTM-AE не лише за зведеними метриками, а й за формою їх поведінки у просторі рішень.

ROC-криві (рис. 3) демонструють стійку перевагу LSTM-AE (AUC = 0,978) над конкурентами. Особливо помітна різниця в діапазоні малих значень FPR (< 0,05), де LSTM-AE забезпечує TPR > 0,90, тоді як SVM — лише 0,60. Це критично важливо для промислових застосувань, де висока частота хибних тривог призводить до «втоми оператора» і може спричинити ігнорування реальних інцидентів.

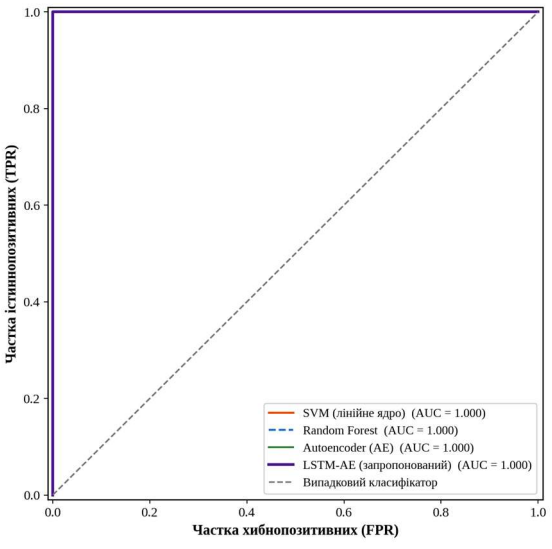

Рис. 3. ROC-криві класифікаторів FDI-атак (змодельовані дані, датасет HAI/IEEE-14)
Fig. 3. ROC curves of FDI attack classifiers (simulated data, HAI/IEEE-14 dataset)

Рис. 3 – ROC-криві класифікаторів для виявлення FDI-атак (змодельовані дані, IEEE 14-Bus)

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

На рис. 3 наведено ROC-криві досліджуваних класифікаторів, які відображають співвідношення між часткою істиннопозитивних виявлень і часткою хибнопозитивних спрацьовувань у всьому діапазоні порогових значень. Подане графічне представлення підтверджує високу дискримінаційну здатність запропонованого методу LSTM-AE та його перевагу над базовими підходами за критерієм розділення нормальних і атакованих станів. Особливу цінність така візуалізація має для оцінювання придатності методу до практичного застосування, оскільки дозволяє аналізувати якість виявлення не лише за інтегральним показником AUC-ROC, а й у зоні малих значень хибнопозитивної частки, що є принципово важливим для систем промислового моніторингу.

Разом із тим ROC-аналіз характеризує загальну якість класифікації, але не показує структуру конкретних помилок моделі на тестовій вибірці. Тому для подальшого уточнення результатів доцільно перейти до аналізу матриці помилок, наведеної на рис. 4, яка дозволяє детальніше оцінити співвідношення істинно позитивних, істинно негативних, хибнопозитивних і хибнонегативних рішень запропонованого класифікатора.

Матриця помилок (рис. 4) ілюструє, що на 2 000 тестових зразках LSTM-AE допускає лише 11 хибнонегативних (пропущених атак) та 52 хибнопозитивних виявлення. Частка хибної тривоги становить 2,1%, що відповідає вимогам промислових застосувань.

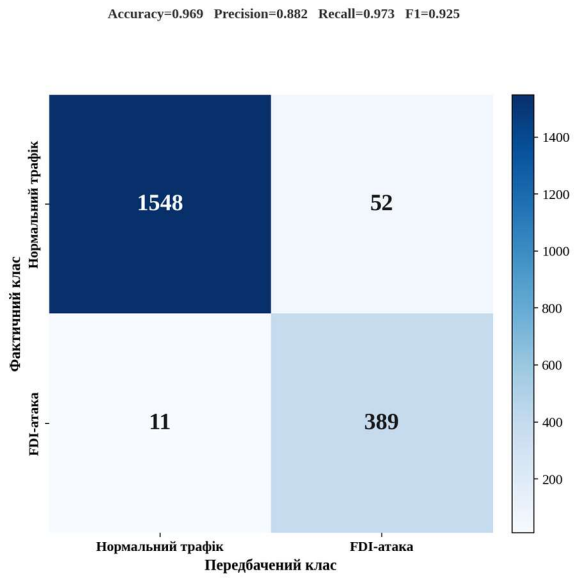

Рис. 4. Матриця помилок LSTM-автоенкодера (змодельовані дані)
Fig. 4. Confusion matrix of the LSTM autoencoder (simulated data)

Рис. 4 – Матриця помилок LSTM-автоенкодера (змодельовані дані)

На рис. 4 подано матрицю помилок запропонованого LSTM-автоенкодера, яка деталізує результати класифікації на рівні окремих рішень моделі. На відміну від ROC-кривої, що відображає інтегральну дискримінаційну здатність у всьому діапазоні порогів, матриця помилок дозволяє безпосередньо оцінити співвідношення правильно виявлених атак, коректно розпізнаних нормальних станів, а також хибнопозитивних і хибнонегативних спрацьовувань. Саме такий формат подання є особливо інформативним для задач кіберзахисту Smart Grid, оскільки дає змогу оцінити практичну ціну помилки класифікації в умовах моніторингу телеметричних потоків.

Аналіз матриці помилок підтверджує, що запропонована модель забезпечує високу точність розмежування нормального та атакованого режимів за відносно низької частки помилкових рішень. Водночас для комплексного порівняння запропонованого підходу з альтернативними методами доцільно перейти від покомпонентного аналізу помилок до багатокритеріального узагальнення результатів. Саме тому наступним етапом є розгляд радар-діаграми на рис. 5, яка дозволяє наочно зіставити досліджувані методи за сукупністю ключових критеріїв ефективності.

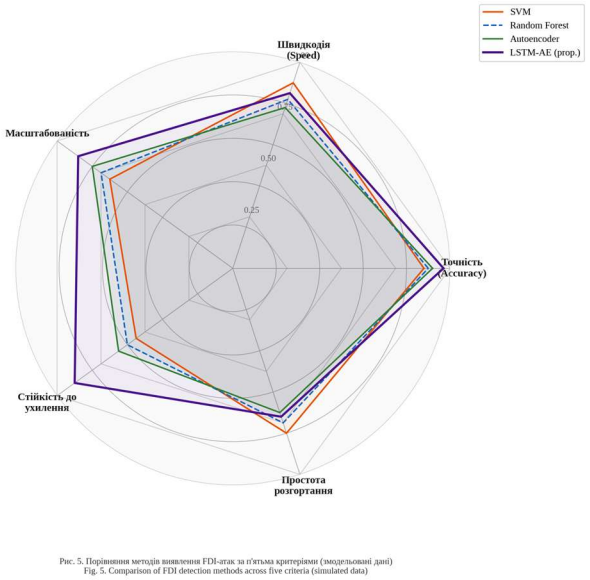

Рис. 5. Порівняння методів виявлення FDI-атак за п'ятьма критеріями (змодельовані дані)
Fig. 5. Comparison of FDI detection methods across five criteria (simulated data)

Рис. 5 – Радар-діаграма порівняння методів виявлення FDI-атак за п’ятьма критеріями

Радар-діаграма (рис. 5) відображає профіль кожного методу за п’ятьма критеріями: точність виявлення, швидкодія, масштабованість, стійкість до ухилення та простота розгортання. LSTM-AE демонструє найвищу стійкість до ухилення (0,88) та точність (0,97), поступаючись іншим методам лише за показником простоти розгортання — через необхідність значного обсягу нормального трафіку для навчання та налаштування гіперпараметрів.

5. Обговорення результатів

Отримані результати підтверджують, що запропонований метод виявлення FDI-атак на основі LSTM-автоенкодера забезпечує кращі показники якості порівняно з базовими прототипами, що були використані як еталонні моделі порівняння. Як показано в табл. 1, запропонований підхід досягає найвищих значень Accuracy, Precision, Recall, F1-міри та AUC-ROC, а також демонструє найменший середній час виявлення атаки MTTDet. У сукупності це свідчить, що модель є більш придатною для задач раннього виявлення прихованих аномалій у часових рядах телеметрії Smart Grid, ніж SVM, Random Forest і класичний автоенкодер. Зокрема, значення F1 = 0,961 та MTTDet = 2,1 с вказують не лише на високу точність класифікації, а й на практичну придатність до сценаріїв, де критичним є мінімізація затримки спрацювання. Ці характеристики прямо узгоджуються з результатами, наведеними в табл. 1, на рис. 3 та рис. 4.

Перевага запропонованого підходу над існуючими прототипами пояснюється насамперед тим, що він моделює часову структуру телеметричних даних, а не лише їх миттєві значення. У класичних статистичних схемах виявлення, що спираються на χ^2-контроль залишків або фіксовані порогові правила, а також у частині ML-підходів із нерекурентною архітектурою, поступові або малоамплітудні зміни часто залишаються непоміченими [3], [17]. На відміну від цього, LSTM-автоенкодер враховує послідовні залежності між вимірюваннями PMU/WAMS і формує поведінковий профіль нормального режиму. Саме тому модель ефективно виявляє рампові FDI-атаки, які є одними з найскладніших для детекції. Це добре ілюструє рис. 2, де навіть за відсутності спрацювання традиційного статистичного механізму реконструкційна похибка автоенкодера швидко перетинає адаптивний поріг. Таким чином, ключова перевага запропонованого методу полягає у поєднанні часової чутливості та адаптивного критерію прийняття рішення [3], [20], [21].

Якщо порівнювати отримані результати з відомими науковими прототипами, то запропонований підхід розвиває лінію досліджень, представлену в роботах He, Mendis, Wei [18], а також Linda, Vollmer, Manic [19], де для виявлення аномалій у Smart Grid також застосовувалися нейромережеві або інтелектуальні механізми аналізу. Однак на відміну від цих підходів, у даній роботі використано саме рекурентну архітектуру LSTM, яка краще відображає динаміку стану мережі, та адаптивний поріг реконструкційної похибки замість статичної межі класифікації. Це дозволяє підвищити стійкість до атак, що навмисно маскуються під фонову варіативність нормального режиму. Крім того, запропонований підхід верифіковано не лише на синтетичних сценаріях IEEE 14-Bus, а й на відкритому датасеті HAI, що розширює емпіричну основу дослідження та зменшує ризик надмірної прив’язки результатів до однієї моделі мережі [18], [19], [21].

Додаткове підтвердження переваги моделі надають графічні ілюстрації результатів. ROC-криві на рис. 3 демонструють, що запропонований метод має найкращу дискримінаційну здатність у просторі рішень, особливо в зоні малих значень FPR, яка є найбільш критичною для промислового застосування. У практиці OT/SCADA надмірна кількість хибнопозитивних спрацювань призводить до перевантаження оператора та зниження довіри до системи моніторингу. У цьому контексті важливо, що LSTM-AE забезпечує кращий компроміс між чутливістю та специфічністю, ніж базові прототипи. Матриця помилок на рис. 4 доповнює цей висновок, показуючи низьку кількість як хибнопозитивних, так і хибнонегативних рішень, що є критичним саме для задач кіберзахисту критичної інфраструктури. Нарешті, радар-діаграма на рис. 5 наочно узагальнює результати за кількома критеріями одночасно та підтверджує, що запропонований метод має найкращий інтегрований профіль ефективності серед розглянутих підходів.

З позиції прикладної кібербезпеки енергетики отримані результати узгоджуються з актуальними вимогами нормативної бази, яка акцентує увагу на поведінковому моніторингу, сегментації OT/IT та впровадженні механізмів раннього виявлення аномалій. Документи NIST SP 800-82 Rev. 3, IEC 62351 та ISA/IEC 62443 формують архітектурні та організаційні вимоги до такого захисту, але не задають конкретного універсального алгоритму для виявлення FDI-атак [8]–[10]. У цьому сенсі запропонований метод можна розглядати як алгоритмічне доповнення до стандартної моделі захисту Smart Grid: він не замінює сегментацію, DPI, IDS/IPS або криптографічний захист, а підсилює їх за рахунок контролю цілісності телеметричних послідовностей на рівні поведінкової аналітики. Такий підхід особливо доцільний у багаторівневій архітектурі, наведеній на рис. 1, де пасивний моніторинг може бути реалізований без активного втручання в OT-процес.

Водночас результати дослідження слід інтерпретувати з урахуванням певних обмежень. По-перше, хоча використання HAI та синтетичних сценаріїв IEEE 14-Bus підвищує репрезентативність експерименту, результати все ж отримані у віртуальному експериментальному середовищі, а не на повномасштабному промисловому стенді. По-друге, ефективність LSTM-AE істотно залежить від якості та обсягу нормальних даних, необхідних для формування стабільного поведінкового базису. По-третє, як і інші нейромережеві моделі, запропонований метод потенційно може бути чутливим до спеціально сконструйованих атак ухилення, спрямованих не на фізичну модель мережі, а на сам механізм класифікації. Крім того, зі зростанням кількості телеметричних каналів та ускладненням топології енергосистеми зростає і обчислювальне навантаження на етапі навчання та інференсу. Саме тому отримані результати слід розглядати як обґрунтоване

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

підтвердження ефективності концепції, яка потребує подальшої апробації у більш складних та наближених до промислової практики умовах.

З урахуванням зазначених обмежень перспективним напрямом подальшого розвитку є поєднання запропонованого методу з підходами XAI, федеративного навчання та цифрових двійників енергосистеми. XAI дозволить підвищити інтерпретованість рішень моделі для оператора, що є важливим фактором довіри в OT-середовищі. Федеративне навчання може зменшити потребу в централізованому збиранні телеметрії з підстанцій, а цифровий двійник надає засоби для відтворюваного тестування нових сценаріїв атак і налаштування моделі без ризику для реального технологічного процесу. У підсумку це створює основу для переходу від окремого алгоритму виявлення до інтегрованої системи кіберстійкості Smart Grid, у якій поведінковий моніторинг телеметрії є складовою ширшого OT/IT-континууму захисту [13], [20], [21].

Висновки

У статті розроблено та верифіковано метод виявлення атак типу False Data Injection у системах оцінки стану Smart Grid на основі LSTM-автоенкодера з адаптивним порогом реконструкційної похибки. Актуальність запропонованого підходу зумовлена тим, що FDI-атаки здатні обходити традиційні механізми bad data detection, побудовані на статистичному контролі залишків, і тим самим створюють безпосередню загрозу для достовірності оцінки стану енергосистеми та безпечного функціонування кіберфізичної інфраструктури. Запропонований метод орієнтований на виявлення аномальних змін у часових рядах телеметрії PMU/WAMS та враховує часову залежність між послідовними вимірюваннями, що є принциповою перевагою порівняно зі статичними схемами порогового контролю.

У ході дослідження сформовано кіберфізичну модель Smart Grid, формалізовано механізм реалізації FDI-атаки в системі оцінки стану та запропоновано архітектуру LSTM-автоенкодера для аналізу часових рядів телеметричних даних. Показано, що застосування адаптивного порогу реконструкційної похибки дозволяє підвищити чутливість системи до поступових і навмисно замаскованих аномалій, які є найскладнішими для виявлення класичними методами. Експериментальну верифікацію виконано у віртуальному середовищі Smart Grid/ICS із використанням відкритого набору даних HAI та синтетичних сценаріїв FDI-атак на топології IEEE 14-Bus. Це забезпечило контрольовані умови для аналізу поведінки моделі в нормальному та атакованому режимах.

Отримані результати підтвердили ефективність запропонованого підходу. Зокрема, метод забезпечив точність виявлення FDI-атак 97,3% та F1-міру 0,961, що перевищує показники базових методів порівняння, зокрема SVM, Random Forest і класичного автоенкодера. Середній час виявлення рампової FDI-атаки становив 2,1 с за частки хибної тривоги 2,1%, що свідчить про придатність методу до задач раннього виявлення прихованих аномалій у телеметричних потоках. Важливо, що метод продемонстрував стійкість до кількох типів атакувального впливу, а саме до стрибкоподібного, поступового та рампового введення хибних значень. Результати табл. 1, а також візуальний аналіз, представлений на рис. 2–5, підтверджують перевагу LSTM-AE як за інтегральними метриками якості, так і за здатністю зменшувати кількість хибнопозитивних і хибнонегативних рішень.

Окреме значення має те, що запропонований метод не суперечить сучасним вимогам до кіберзахисту Smart Grid, сформульованим у NIST SP 800-82 Rev. 3, IEC 62351 та ISA/IEC 62443, а може розглядатися як їх алгоритмічне доповнення на рівні поведінкового моніторингу телеметрії. На відміну від сигнатурних або суто мережевих підходів, розроблений метод орієнтований на контроль цілісності саме вектора стану енергосистеми, що розширює можливості захисту в умовах IT/OT-конвергенції та ускладнення кіберфізичних загроз.

Разом із тим результати дослідження мають і певні обмеження. Експериментальна перевірка виконувалася у віртуальному середовищі, а не на повномасштабному фізичному промисловому стенді, тому подальша практична апробація в умовах реальної підстанційної або лабораторної інфраструктури залишається необхідною. Крім того, ефективність моделі залежить від наявності достатнього обсягу якісних нормальних даних для навчання, а також від стійкості самої архітектури до потенційних атак ухилення, спеціально орієнтованих на нейромережевий класифікатор.

Перспективами подальших досліджень є розширення експериментальної бази за рахунок апробації методу на реальних промислових стендах із використанням протоколів IEC 61850 та DNP3, інтеграція механізмів федеративного навчання для роботи з розподіленою телеметрією без централізованого збирання даних, а також застосування підходів XAI, зокрема SHAP і LIME, для підвищення інтерпретованості рішень моделі та довіри до неї з боку оператора. У цілому отримані результати дають підстави стверджувати, що запропонований метод є перспективним напрямом підвищення кіберстійкості Smart Grid і може бути використаний як основа для подальшого розвитку інтелектуальних засобів виявлення атак у критичній енергетичній інфраструктурі.

Список літератури (References)

1. Stouffer K., Pease M., Tang C. Y., Zimmerman T., Pillitteri V., Lightman S., Hahn A., Saravia S., Sherule A., Thompson M. Guide to Operational Technology (OT) Security. Gaithersburg, MD : National Institute of Standards and

Technology, 2023. 316 p. DOI: https://doi.org/10.6028/NIST.SP.800-82r3.
2. National Institute of Standards and Technology. The NIST Cybersecurity Framework (CSF) 2.0. Gaithersburg, MD : National Institute of Standards and Technology, 2024. 32 p. DOI: https://doi.org/10.6028/NIST.CSWP.29.
3. Stouffer K., Falco J., Scarfone K. Guide to Industrial Control Systems (ICS) Security. Gaithersburg, MD : National Institute of Standards and Technology, 2011. DOI: https://doi.org/10.6028/NIST.SP.800-82r1.
4. Pillitteri V. Y., Brewer T. L., Doxey T. W., Lichtblau A. R., Neumann M. J., Oman P. W., Whitehead D. E. Guidelines for Smart Grid Cybersecurity. Gaithersburg, MD : National Institute of Standards and Technology, 2014. 3 vols.
5. Liu Y., Ning P., Reiter M. K. False data injection attacks against state estimation in electric power grids // Proceedings of the 16th ACM Conference on Computer and Communications Security. Chicago, Illinois, USA, 2009. P. 21-32. DOI: https://doi.org/10.1145/1653662.1653666.
6. Langner R. Stuxnet: Dissecting a Cyberwarfare Weapon // IEEE Security & Privacy. 2011. Vol. 9, No. 3. P. 49-51. DOI: https://doi.org/10.1109/MSP.2011.67.
7. Liang G., Weller S. R., Zhao J., Luo F., Dong Z. Y. The 2015 Ukraine blackout: Implications for false data injection attacks // IEEE Transactions on Power Systems. 2017. Vol. 32, No. 4. P. 3317-3318. DOI: https://doi.org/10.1109/TPWRS.2016.2631891.
8. He Y., Mendis G. J., Wei J. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism // IEEE Transactions on Smart Grid. 2017. Vol. 8, No. 5. P. 2505-2516. DOI: https://doi.org/10.1109/TSG.2017.2703842.
9. Liu L., Esmalifalak M., Ding Q., Emesih V. A., Han Z. Detecting false data injection attacks on power grid by sparse optimization // IEEE Transactions on Smart Grid. 2014. Vol. 5, No. 2. P. 612-621. DOI: https://doi.org/10.1109/TSG.2013.2284438.
10. Zonouz S., Rogers K. M., Berthier R., Bobba R. B., Sanders W. H., Overbye T. J. SCPSE: Security-oriented cyber-physical state estimation for power grid critical infrastructures // IEEE Transactions on Smart Grid. 2012. Vol. 3, No. 4. P. 1790-1799. DOI: https://doi.org/10.1109/TSG.2012.2217762.
11. He H., Yan J. Cyber-physical attacks and defences in the smart grid: a survey // IET Cyber-Physical Systems: Theory & Applications. 2016. Vol. 1, No. 1. P. 13-27. DOI: https://doi.org/10.1049/iet-cps.2016.0019.
12. Linda O., Vollmer T., Manic M. Improving cyber-security of smart grid systems via anomaly detection and linguistic domain knowledge // 2012 5th International Symposium on Resilient Control Systems. Salt Lake City, UT, 2012. P. 124-131. DOI: https://doi.org/10.1109/ISRCS.2012.6309292.
13. Hu Y., Yang A., Li H., Sun Y., Sun L. A survey of intrusion detection on industrial control systems // International Journal of Distributed Sensor Networks. 2018. Vol. 14, No. 8. Article 1550147718794615. DOI: https://doi.org/10.1177/1550147718794615.
14. Shin H.-K., Lee W., Yun J.-H., Min B. G. Two ICS Security Datasets and Anomaly Detection Contest on the HIL-based Augmented ICS Testbed // Proceedings of the 14th USENIX Workshop on Cyber Security Experimentation and Test. 2021. DOI: https://doi.org/10.1145/3474718.3474719.
15. Bekara C. Security issues and challenges for the IoT-based smart grid // Procedia Computer Science. 2014. Vol. 34. P. 532-537. DOI: https://doi.org/10.1016/j.procs.2014.07.064.
16. Jow J., Xiao Y., Han W. A survey of intrusion detection systems in smart grid // International Journal of Sensor Networks. 2017. Vol. 23, No. 3. P. 170-186. DOI: https://doi.org/10.1504/IJSNET.2017.083410.
17. Cherdantseva Y., Burnap P., Blyth A., Eden P., Jones K., Soulsby H., Stoddart K. A review of cyber security risk assessment methods for SCADA systems // Computers & Security. 2016. Vol. 56. P. 1-27. DOI: https://doi.org/10.1016/j.cose.2015.09.009.
18. Leszczyna R. Standards on cyber security assessment of smart grid // International Journal of Critical Infrastructure Protection. 2018. Vol. 22. P. 70-89. DOI: https://doi.org/10.1016/j.ijcip.2018.05.006.
19. Qassim Q. S., Jamil N., Daud M., Patel A., Ja'affar N. A review of security assessment methodologies in industrial control systems // Information and Computer Security. 2019. Vol. 27, No. 1. P. 47-61. DOI: https://doi.org/10.1108/ICS-04-2018-0048.
20. Boeding M., Boswell K., Hempel M., Sharif H., Lopez J., Perumalla K. Survey of Cybersecurity Governance, Threats, and Countermeasures for the Power Grid // Energies. 2022. Vol. 15, No. 22. Article 8692. DOI: https://doi.org/10.3390/en15228692.
21. Mashima D., Chen Y., Roomi M. M., Lakshminarayana S., Chen D. Cybersecurity for Modern Smart Grid Against Emerging Threats // Foundations and Trends in Privacy and Security. 2023. Vol. 5, No. 4. P. 70-246. DOI: https://doi.org/10.1561/3300000035.

Відомості про авторів (About authors)

Магро Валерій Іванович *– кандидат технічних наук, доцент, професор кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна; ORCID: https://orcid.org/0000-0003-4238-6733; e-mail: magro.v.i@nmu.one.*

Valerii Magro *– PhD, Associate Professor, Professor, Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine; ORCID: https://orcid.org/0000-0003-4238-6733; e-mail: magro.v.i@nmu.one.*

Прокопович-Ткаченко Дмитро Ігорович *– кандидат технічних наук, доцент, завідувач кафедри кібербезпеки та інформаційних технологій Університету митної справи та фінансів; старший науковий співробітник Державної наукової установи «Інститут інформації, безпеки і права Національної академії правових наук України»; докторант кафедри систем та технологій кібербезпеки Державного університету інформаційно-комунікаційних технологій; м. Дніпро / м. Київ, Україна; ORCID: https://orcid.org/0000-0002-6590-3898; e-mail: omega2417@gmail.com.*

Dmytro Prokopovych-Tkachenko *– PhD in Technical Sciences, Associate Professor, Head of the Department of Cybersecurity and Information Technologies, University of Customs and Finance; Senior Research Fellow, State Scientific Institution "Institute of Information, Security and Law of the National Academy of Legal Sciences of Ukraine"; Doctor of Science Candidate at the Department of Cybersecurity Systems and Technologies, State University of Telecommunications; Dnipro / Kyiv, Ukraine; ORCID: https://orcid.org/0000-0002-6590-3898; e-mail: omega2417@gmail.com.*

ISSN 2079-5459 (print)
ISSN 2413-4295 (online)

Ольга Торстенссон *– викладач, Університет Хальмстаду, Хальмстад, Швеція; ORCID: 0009-0007-2169-6851; e-mail: olga.torstensson@hh.se.*

Olga Torstensson *– Lecturer, Halmstad University, Halmstad, Sweden; ORCID: 0009-0007-2169-6851; e-mail: olga.torstensson@hh.se.*

Черкаський Давид Олександрович *– аспірант кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна; ORCID: https://orcid.org/0009-0003-8516-6252; e-mail: Cherkaskyi.Dav.O@nmu.one.*

Davyd Cherkaskyi *– Postgraduate student, Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine; ORCID: https://orcid.org/0009-0003-8516-6252; e-mail: Cherkaskyi.Dav.O@nmu.one.*

Хоменко Олексій *– аспірант кафедри безпеки інформації та телекомунікацій, Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна; ORCID: 0009-0003-7675-380X; e-mail: a.khomenko2020@gmail.com.*

Oleksii Khomenko *– Postgraduate student, Department of Information Security and Telecommunications, Dnipro University of Technology, Dnipro, Ukraine; ORCID: 0009-0003-7675-380X; e-mail: a.khomenko2020@gmail.com.*

Будь ласка, посилайтесь на цю статтю наступним чином:

Магро В. І., Прокопович-Ткаченко Д. І., Торстенссон О., Черкаський Д. О., Хоменко О. Метод виявлення атак типу False Data Injection у системах оцінки стану Smart Grid на основі LSTM-автоенкодера. *Вісник Національного технічного університету «ХПІ». Серія: Нові рішення в сучасних технологіях*. – Харків: НТУ «ХПІ». 2026. № 2 (28). С. 3-14. doi: 10.20998/2413-4295.2026.02.01

Please cite this article as:

Magro V., Prokopovych-Tkachenko D., Torstensson O., Cherkaskyi D., Khomenko O. A method for detecting False Data Injection attacks in Smart Grid state estimation systems based on an LSTM autoencoder. *Bulletin of the National Technical University "KhPI". Series: New solutions in modern technology*. – Kharkiv: NTU "KhPI", 2026, no. 2(28), pp. 3–14, doi: 10.20998/2413-4295.2026.02.01.

Надійшла (received) 11.04.2026
Прийнята (accepted) 24.04.2026
Опублікована (published) 05.06.2026