КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.32.1149
УДК 004.056:004.94:658.5

Савченко Тетяна Віталіївна
Кандидат технічних наук, доцент, доцент кафедри інформатики
Національний університет «Києво-Могилянська академія», Київ, Україна
ORCID: 0000-0002-8884-5360
tsavchenko@ukma.edu.ua

Бондар Ярослав Ігорович
Студент спеціальності «Кібербезпека та захист інформації»
Національний університет «Києво-Могилянська академія», Київ, Україна
ya.bondar@ukma.edu.ua

РИЗИКИ ТА ВРАЗЛИВОСТІ ЦИФРОВИХ ДВІЙНИКІВ ВИРОБНИЧИХ СИСТЕМ У ПРОЦЕСІ ОПТИМІЗАЦІЇ ЕНЕРГОСПОЖИВАННЯ

Анотація. Представлене наукове дослідження присвячене глибокому системному аналізу проблем інформаційної безпеки, що виникають під час впровадження, інтеграції та експлуатації технології цифрових двійників у сучасних кіберфізичних виробничих системах, орієнтованих на оптимізацію енергоспоживання. В умовах стрімкого переходу до Industry 4.0, який супроводжується тотальною цифровізацією та конвергенцією обчислювальних ресурсів із фізичними процесами, технологія цифрових двійників стає безальтернативним інструментом для досягнення високої енергоефективності, мінімізації втрат та забезпечення сталого розвитку промисловості. Однак, як доводить проведене дослідження, руйнування традиційного фізичного ізолювання мереж (Air Gap) та встановлення безперервного двостороннього потоку даних між сенсорним обладнанням операційних технологій та хмарними платформами інформаційних технологій створює безпрецедентну за своїми масштабами поверхню для кібератак. У статті здійснено фундаментальний огляд архітектурних вразливостей промислового Інтернету речей, проаналізовано слабкі місця ключових комунікаційних протоколів та розроблено розширену класифікацію загроз на базі адаптованої моделі STRIDE. Особливий акцент зроблено на загрозах цілісності даних, зокрема атаках типу ін’єкції хибних даних (FDI) та їх кумулятивному впливі на алгоритми оптимізації. У межах методологічної та практичної частини роботи побудовано репрезентативне локальне середовище за принципом Software-in-the-Loop, яке об’єднує Python-емулятор енергоспоживання, брокер повідомлень Eclipse Mosquitto та подієво-керовану платформу Node-RED. Шляхом практичного моделювання атак типу Spoofing та Tampering доведено, що компрометація вхідної телеметрії здатна повністю дезорієнтувати алгоритми ковзного середнього (SMA), приховуючи від операторів критичні перевантаження та блокуючи механізми аварійного скидання навантажень, що в реальних умовах гарантовано призводить до фізичного руйнування інфраструктури. Для протидії виявленим векторам загроз розроблено, імплементовано та верифіковано комплексну архітектуру багаторівневого захисту (Defense in Depth). Запропонований підхід поєднує криптографічне шифрування транспортного каналу за протоколом TLS із жорсткою аутентифікацією та прикладну валідацію цілісності корисного навантаження за допомогою цифрових підписів HMAC-SHA256 і перевірки часових міток. Оцінка ефективності продемонструвала повну нейтралізацію досліджуваних атак при збереженні функціональної стабільності алгоритмів керування та цілком прийнятних накладних витратах на мережеві затримки на рівні 15-20%.

Ключові слова: цифровий двійник; кібербезпека; оптимізація енергоспоживання; промисловий інтернет речей; ін’єкція хибних даних; багаторівневий захист; криптографічна аутентифікація; кіберфізичні системи.

ВСТУП

Сучасний етап розвитку промисловості характеризується глибокою цифровою трансформацією, яку прийнято називати Industry 4.0. Одним із ключових аспектів цього процесу є впровадження кіберфізичних систем, які об'єднують обчислювальні ресурси з фізичними процесами виробництва. У цьому контексті технологія цифрових двійників виступає як інноваційний підхід, що дозволяє створити віртуальну копію фізичного активу, процесу або системи для моніторингу, діагностики та прогнозування їхнього стану в режимі реального часу [1]. Особливо ця технологія набуває актуальності у сфері енергетичного менеджменту, де в умовах енергетичної кризи підприємства шукають шляхи для радикального підвищення енергоефективності. Цифрові двійники дозволяють вирішувати ці задачі шляхом точного моделювання споживання, балансування навантажень та оптимізації режимів роботи обладнання без ризику зупинки реального виробництва [2].

Постановка проблеми. Широке впровадження технології цифрових двійників несе в собі значні ризики у сфері інформаційної безпеки. На відміну від ізольованих систем керування, цифрові двійники вимагають постійного, двостороннього обміну величезними масивами даних між фізичним рівнем, де знаходяться сенсори та контролери, та цифровим рівнем, представленим хмарними платформами. Це призводить до розмивання периметра безпеки підприємства та появи нових вразливостей на стику інформаційних та операційних технологій.

Компрометація даних енергомоніторингу безпосередньо призводить до прийняття хибних рішень алгоритмами оптимізації, що, в свою чергу, загрожує не лише економічними збитками, але й фізичними аваріями, перевантаженням мереж та загрозою життю персоналу. Незважаючи на значну кількість публікацій, присвячених загальним питанням кібербезпеки промислового Інтернету речей, специфіка захисту цифрових двійників саме в контексті енергетичної оптимізації залишається недостатньо вивченою. Існуючі методи захисту часто не враховують вплив маніпуляцій даними на фізику технологічних процесів. Відповідно, дослідження ризиків та вразливостей цифрових двійників у цій сфері є актуальним завданням.

Аналіз останніх досліджень і публікацій. Проблематика інтеграції кіберфізичних систем, розгортання технологій цифрових двійників та забезпечення їхньої стійкості до зовнішніх і внутрішніх деструктивних впливів є предметом інтенсивних дискусій у сучасній світовій науковій спільноті. Аналіз фахових публікацій свідчить про наявність кількох домінантних підходів до розгляду архітектурних особливостей DT та пов'язаних із ними безпекових викликів.

Фундаментальні теоретичні аспекти архітектури, класифікації та функціонального призначення цифрових двійників надзвичайно детально розкрито у класичних і визнаних працях дослідників Fuller A. [2] та Tao F. [6]. У своїх роботах ці автори ґрунтовно акцентують увагу на абсолютній важливості точності математичного моделювання фізичних процесів та бездоганної синхронізації інформаційних потоків між фізичним та віртуальним світами. Проте у згаданих базових парадигмах питання кібербезпеки найчастіше розглядаються як вторинні функції або адміністративні накладні витрати, що неминуче створює глибокі архітектурні вразливості ще на ранніх етапах проєктування таких систем.

Друга значна група науковців та інституцій фокусується безпосередньо на розробці таксономії та класифікації кіберзагроз. Зокрема, у масштабних аналітичних звітах Національного інституту стандартів і технологій США та у ґрунтовних працях

Alcaraz C. [3] наведено вичерпний огляд потенційних векторів атак, спрямованих на складні IoT-системи [7]. Дослідник Alcaraz C. іде далі і пропонує специфічну адаптацію класичної моделі моделювання загроз STRIDE безпосередньо для парадигми цифрових двійників, обґрунтовано виділяючи унікальні ризики десинхронізації станів фізичного об’єкта та його віртуальної копії. Водночас слід визнати, що цінні результати цих досліджень мають переважно описовий та оглядовий характер, вони не пропонують готових прикладних алгоритмів або протоколів захисту, які можна було б імплементувати у специфічних галузях із жорсткими обмеженнями.

Надзвичайно важливу специфіку кібератак, спрямованих виключно на енергетичні системи та Smart Grids, досліджують автори на чолі з He Y. У їхніх публікаціях наведено математичне обґрунтування руйнівної природи атак типу ін’єкції хибних даних (FDI) [5]. Дослідники доводять, що кваліфікований зловмисник, маючи інформацію про топологію енергомережі, здатен математично точно сформувати такий вектор атаки, який буде визнаний технічно коректним і гарантовано пройде перевірку стандартними статистичними фільтрами помилок, що дозволить йому непомітно маніпулювати алгоритмами енергетичної оптимізації. Основною проблемою є те, що запропоновані авторами методи виявлення таких складних FDI-атак вимагають залучення колосальних обчислювальних потужностей, що є практично нереалізованим завданням на рівні сучасних промислових програмованих логічних контролерів або периферійних шлюзів.

Питання технологічної протидії ідентифікованим загрозам активно розглядаються у перспективних працях Suhail S. та Eckhart M. Зокрема Suhail S. аргументовано розглядає технологію блокчейн [4] та розподілених реєстрів як надзвичайно перспективний інструмент забезпечення криптографічної незмінності та цілісності даних. Водночас Eckhart M. справедливо вказує на фундаментальну проблему непередбачуваних і значних часових затримок при використанні консенсусних механізмів розподілених реєстрів у кіберфізичних системах, які вимагають реакції у режимі жорсткого реального часу [8].

Враховуючи стан наукової думки, можна констатувати наявність суттєвої прогалини: переважна більшість існуючих методів забезпечення кібербезпеки або зовсім не враховує критично жорсткі вимоги до часу реакції в енергетичних системах, або ж повністю ігнорує вплив кібернетичних маніпуляцій даними на базову фізику та термодинаміку технологічних процесів. Саме тому сфокусоване дослідження ризиків, вразливостей та методів захисту цифрових двійників [9] виключно в контексті оптимізації енергоспоживання набуває статусу особливо актуального та невідкладного науково-технічного завдання.

Мета статті. Метою роботи є систематизація і розширення наукових знань про специфічні загрози інформаційній безпеці цифрових двійників, а також розробка, впровадження та експериментальна верифікація ефективних технологічних підходів до мінімізації кіберризиків під час використання DT-технологій для оптимізації енергоспоживання у виробничих системах.

Для досягнення поставленої мети передбачено вирішення комплексу взаємопов’язаних завдань, що охоплюють системний аналіз технології цифрових двійників у контексті інтелектуального енергоменеджменту, дослідження багаторівневих інтеграційних архітектур і вразливостей протоколів передавання даних, розроблення адаптованої класифікації загроз для середовищ DT в енергетиці, експериментальне оцінювання впливу кібератак на алгоритми енергетичної оптимізації на базі віртуального середовища дослідження, а також формування та впровадження

комплексу криптографічних і логічних механізмів забезпечення інформаційної безпеки каналів комунікації та критичних потоків даних.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У сучасному кіберфізичному середовищі загрози інформаційній безпеці стають дедалі витонченішими, тому активне виявлення вразливостей цифрових двійників через цілеспрямоване моделювання кібератак є критично важливим компонентом стратегії захисту промислового підприємства. Цей процес виступає проактивним методом оцінювання захищеності систем енергоменеджменту, що передбачає імітацію деструктивних дій потенційного зловмисника з метою виявлення слабких місць в інтеграційних архітектурах та комунікаційних протоколах. Такий підхід, реалізований ізольовано від реального обладнання, дає змогу своєчасно ідентифікувати та усувати вразливості до моменту їх експлуатації реальними нападниками, що дозволяє уникнути фізичних аварій та непередбачуваних зупинок технологічних процесів. Експериментальне моделювання загроз сьогодні є невід'ємним і ключовим елементом формування надійних практик кіберзахисту для рішень парадигми Industry 4.0.

Практична перевірка захищеності розгорнутої моделі цифрового двійника охоплювала комплекс спеціальних заходів, спрямованих на реалістичне відтворення векторів атак на інформаційні потоки із застосуванням підходу Software-in-the-Loop. У ході експерименту було здійснено глибокий аналіз впливу скомпрометованої телеметрії на логіку прийняття рішень алгоритмами енергетичної оптимізації. Результати цього комплексного моделювання наочно демонструють наслідки успішної реалізації атак типу Spoofing та Tampering на цілісність даних. Отримані експериментальні дані забезпечують вичерпне розуміння архітектурних недоліків базових конфігурацій IIoT-мереж, дозволяють здійснити кількісну оцінку впливу хибних даних на показники енергоспоживання та формують міцний фундамент для впровадження стратегічних пропозицій з удосконалення інформаційної безпеки, зокрема шляхом розгортання криптографічних протоколів TLS та прикладних алгоритмів цифрового підпису HMAC-SHA256.

1. Архітектурний ландшафт та комунікаційні вразливості цифрових двійників.

Для виявлення локалізації вразливостей систем енергетичної оптимізації було здійснено деконструкцію типової архітектури розгортання цифрового двійника. Інтеграційна архітектура, яка демонструє взаємодію фізичних активів, периферійних обчислень та хмарних сервісів у середовищі Smart Factory, наведена на рис. 1. Промислові цифрові двійники характеризуються багаторівневою структурою, де об'єктивно виникає архітектурний конфлікт: IT-протоколи розроблялися з акцентом на конфіденційність, тоді як OT-протоколи пріоритетно забезпечують доступність та швидкість, утворюючи інтеграційні сірі зони безпеки.

Фундаментальним є фізичний рівень, який об'єднує реальні енергоємні виробничі активи, трансформатори, розподільчі щити та мережі сенсорів (лічильники, аналізатори якості електроенергії). Головна проблема безпеки цього рівня полягає в критично малих обчислювальних ресурсах кінцевих пристроїв, що унеможливлює використання складної асиметричної криптографії чи сучасних агентів антивірусного захисту. Здебільшого такі пристрої функціонують на застарілих прошивках, акумулюючи відомі вразливості протягом тривалого часу.

На рівні периферійних обчислень та зв'язку дані з цеху проходять через Edge-шлюзи, де відбувається агрегація телеметрії та конвертація застарілих протоколів.

Використання традиційних промислових стандартів, таких як Modbus TCP, PROFINET чи EtherCAT, становить суттєву загрозу, оскільки вони переважно передають дані у відкритому вигляді і не підтримують механізмів криптографічної аутентифікації джерела. Це формує передумови для реалізації атак Man-in-the-Middle та несанкціонованої підміни даних у разі отримання зловмисником доступу до мережевого сегмента.

Для передавання даних від Edge-рівня до хмарних платформ, де розгортається логіка цифрового двійника (із застосуванням баз даних часових рядів типу InfluxDB та інтеграційних шин Apache Kafka), застосовуються IT-орієнтовані протоколи, серед яких домінує MQTT. Попри те, що MQTT є гнучким і стійким до обривів зв'язку завдяки архітектурі Publish/Subscribe, його базова конфігурація часто залишається незахищеною. Інтеграція цих рівнів остаточно нівелює класичну модель фізично ізольованої мережі (Air Gap), створюючи наскрізний інформаційний тунель від виробничого цеху до глобальної мережі. Ситуація додатково ускладнюється явищем тіньових інфраструктур, коли персонал самовільно розгортає бездротові комунікаційні модулі в обхід корпоративних брандмауерів, формуючи неконтрольований плацдарм для реалізації кібератак.

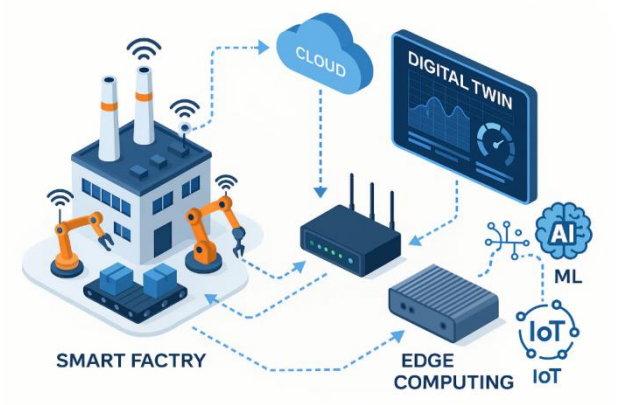


Рис. 1. Типова архітектура розгортання цифрового двійника у середовищі Smart Factory [10]

2. Класифікація загроз та специфіка атак на енергетичні алгоритми.

Адаптація методології моделювання загроз STRIDE до контексту кіберфізичних систем оптимізації енергоспоживання дозволила систематизувати специфічні вектори атак. Встановлено, що ключовим пріоритетом в енергетичних системах є не конфіденційність, а цілісність та доступність даних у режимі реального часу.

Однією з найбільш деструктивних категорій визнано атаки на цілісність даних, серед яких особливе місце посідає ін'єкція хибних даних. У сценарії такої атаки зловмисник маніпулює телеметрією, впроваджуючи до системи підроблені, штучно занижені показники споживання потужного обладнання. Цифровий двійник, аналізуючи ці фальсифіковані дані, ідентифікує наявність резерву потужності та генерує дозвіл на запуск додаткового обладнання. Фізичним наслідком такої логічної помилки є спрацювання апаратного захисту на підстанції та повне знеструмлення підприємства.

Не менш критичними є загрози доступності (DoS/DDoS), які спрямовані на перевантаження шлюзів або брокерів повідомлень MQTT. Оскільки енергетична система вимагає реакції за мілісекунди, штучна затримка даних фактично

унеможливлює їх використання та призводить до втрати синхронізації цифрового двійника з фізичним об'єктом. Схема такої розсинхронізації наведена на рис. 2. У стані десинхронізації алгоритми оптимізації або безпідставно зупиняють виробництво, що призводить до хибнопозитивного результату, або не реагують на реальну аварійну ситуацію. Додатковим вектором виступають атаки типу Battery Draining на бездротові сенсори, коли безперервний потік мережевих запитів виснажує елемент живлення за лічені дні, позбавляючи систему критичного джерела даних.

Окрему небезпеку становлять атаки на низхідний потік даних (Downstream Attacks) від цифрового двійника до актуаторів, зокрема несанкціоноване управління (Unauthorized Actuation). Зловмисник, перехопивши канал керування, здатен непомітно підмінити легітимний керуючий сигнал на деструктивний, що провокує аварійний режим роботи обладнання. Ця атака супроводжується передачею фальсифікованих зворотних підтверджень: на інтерфейсі оператора відображається штатна робота системи, що створює хибну ілюзію нормального функціонування та позбавляє персонал ситуаційної обізнаності у критичний момент.

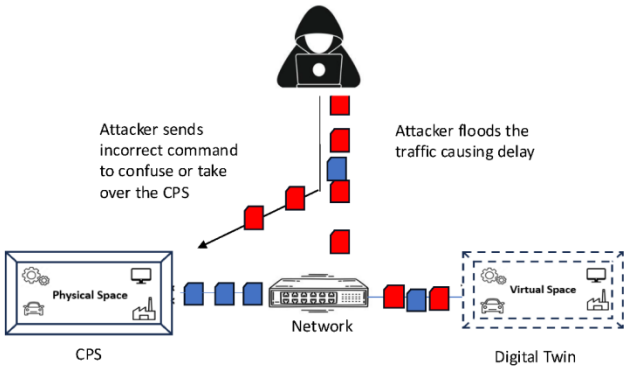


Рис. 2. Схема атаки на доступність та розсинхронізацію даних між CPS та цифровим двійником [11]

Окрім прямих мережевих втручань, використання відкритих програмних рішень, таких як середовище Node-RED та пов'язані з ним npm-пакети, формує вразливості ланцюжка постачання (Supply Chain Attacks). Шкідливий код, інжектований у вразливу бібліотеку, стає частиною моделі цифрового двійника і отримує легітимний доступ до внутрішньої інфраструктури підприємства. Механізм реалізації такої атаки проілюстровано на рис. 3.

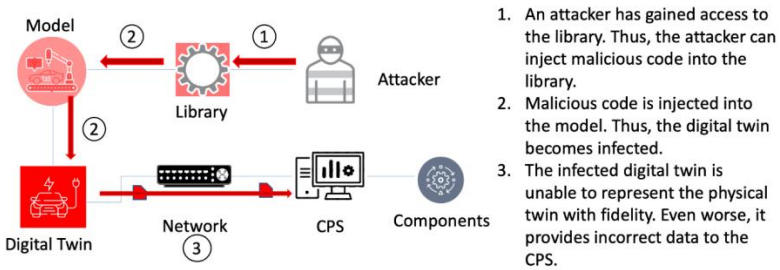


Рис. 3. Реалізація атаки на ланцюжок поставок через компрометацію програмних бібліотек моделі [12]

3. Розгортання експериментального середовища Software-in-the-Loop.

Для емпіричного підтвердження теоретичних висновків та безпечного моделювання деструктивних сценаріїв було розроблено локальне експериментальне середовище за парадигмою Software-in-the-Loop. Цей підхід забезпечує програмну емуляцію фізичних активів, ізолюючи процес дослідження від ризиків виникнення реальних техногенних аварій. Архітектура середовища побудована за трирівневою моделлю промислового IIoT. Схема інформаційних потоків експериментального середовища, де взаємодія між компонентами здійснюється виключно за протоколом MQTT через брокер повідомлень Eclipse Mosquitto, наведена на рис. 4. Запропонована конфігурація дозволяє гнучко відтворювати мережеві вразливості, перехоплювати та модифікувати трафік для аналізу.

Рис. 4. Схема інформаційних потоків експериментального середовища

На нижньому рівні (генерації даних) розгорнуто програмний емулятор енергоспоживання, реалізований мовою Python. Відмову від використання статичних наборів даних (датасетів) обґрунтовано необхідністю дослідження динамічної реакції системи управління в режимі реального часу. Математична модель емулятора базується на суперпозиції декількох сигналів: базова синусоїда імітує добовий виробничий цикл із ранковими та вечірніми піками, а накладений білий шум генерує природні стохастичні флуктуації, характерні для показань реальних вимірювальних приладів. Базова активна потужність $P_{raw}(t)$ у момент часу t описується рівнянням:

$$P_{raw}(t) = P_{base} + A \cdot \sin(\omega t) + \varepsilon(t) \quad (1)$$

де P_{base} – базове навантаження, A – амплітуда коливань, ω – кутова частота, $\varepsilon(t)$ – випадкова величина (шум). Відповідно, сила струму $I(t)$ розраховується як:

$$I(t) = \frac{P_{raw}(t) \cdot 1000}{U(t)} \quad (2)$$

За допомогою бібліотеки paho-mqtt розроблений скрипт із секундним інтервалом публікує JSON-пакети зі значеннями активної потужності, напруги та струму.

На середньому рівні, який виконує роль цифрового двійника та системи прийняття рішень, розгорнуто платформу Node-RED. Вибір цього середовища зумовлений його орієнтованістю на подієво-керовану (event-driven) архітектуру та

наявністю потужних інструментів для потокової обробки телеметрії. Практична реалізація логіки обробки даних у середовищі Node-RED проілюстрована на рис. 5. Базовий алгоритм оптимізації ґрунтується на обчисленні простого ковзного середнього (SMA), що ефективно згладжує імпульсні шуми та виділяє стійкі тренди споживання. Згладжене значення потужності SMA_n для вікна з n вимірювань розраховується за формулою:

$$SMA_n = \frac{1}{n}\sum_{i=0}^{n-1} P(t-i) \tag{3}$$

де $P(t-i)$ – значення спожитої потужності у попередні моменти часу.

Логіка управління функціонує за жорстким пороговим принципом: якщо обчислене згладжене значення потужності перевищує встановлений ліміт у 80 кВт, цифровий двійник генерує команду на відключення непріоритетного навантаження. Математично він реалізує функцію керуючого впливу $L(t)$, яка визначає ліміт потужності для фізичного обладнання:

$$L(t) = \begin{cases} 50\%, & \text{якщо } SMA_n > 80 \\ 100\%, & \text{якщо } SMA_n \leq 80 \end{cases} \tag{4}$$

Оптимізована потужність $P(t)$ після втручання цифрового двійника становить:

$$P(t) = P_{raw}(t) \cdot \frac{L(t)}{100} \tag{5}$$

Відповідно до цієї логіки, Node-RED публікує у зворотний топік MQTT команду на примусове зниження споживання. Фундаментальна вразливість цієї системи була свідомо закладена на архітектурному рівні під час проєктування середовища дослідження: алгоритм керування приймає рішення, повністю довіряючи будь-яким вхідним даним без попередньої перевірки їхнього походження чи криптографічної цілісності.

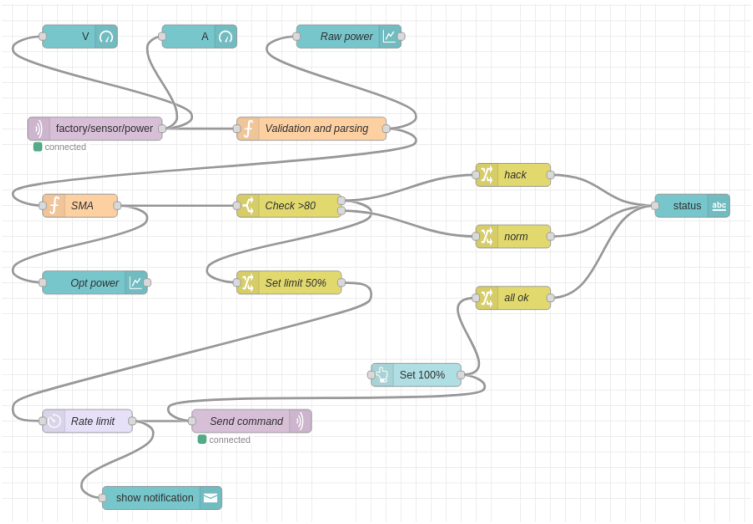


Рис. 5. Реалізація логіки обробки даних у середовищі Node-RED

4. Симуляція кібератак на алгоритми оптимізації.

Експериментальне моделювання передбачало реалізацію векторів атак від імені внутрішнього порушника, який здобув несанкціонований доступ до локальної виробничої мережі (наприклад, через компрометацію бездротового сенсора), проте не володіє адміністративними привілеями на рівні серверної інфраструктури.

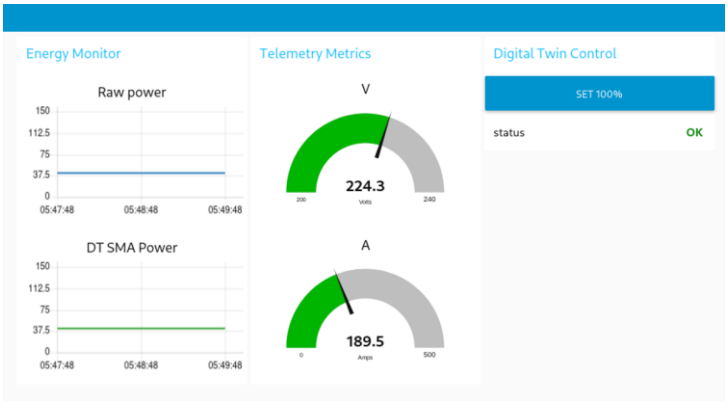
Перший сценарій дослідження (Spoofing-атака, спрямована на порушення доступності та цілісності) базувався на експлуатації відсутності механізмів аутентифікації клієнтів у базовій конфігурації брокера повідомлень Mosquitto. У ході експерименту шкідливий скрипт здійснив підключення до брокера, використовуючи ідентифікатор легітимного сенсора PLC-01-Main, після чого ініціював флуд фальсифікованих пакетів даних із критичними значеннями у цільовий топік. Це спровокувало Race Condition в обчислювальному ядрі платформи Node-RED. Чергування валідних показників зі сфабрикованими миттєво дестабілізувало роботу алгоритму SMA. Візуалізація аномалій телеметрії під час реалізації даного вектора атаки наведена на рис.6. На інтерфейсі оператора спостерігалися хаотичні коливання графіків від номінальних до критичних значень, що змусило систему оптимізації безперервно і безпідставно генерувати команди на перемикання обладнання. У реальних промислових умовах така аномальна поведінка алгоритмів неминуче призводить до прискореного механічного та електричного зносу комутаційної апаратури.


```
[ATTACK] Sent fake data: 120.0 kW
[ATTACK] Sent fake data: 120.0 kW
[ATTACK] Sent fake data: 5.0 kW
[ATTACK] Sent fake data: 10.5 kW
[ATTACK] Sent fake data: 120.0 kW
[ATTACK] Sent fake data: 5.0 kW
[ATTACK] Sent fake data: 120.0 kW
[ATTACK] Sent fake data: 150.0 kW
[ATTACK] Sent fake data: 150.0 kW
[ATTACK] Sent fake data: 5.0 kW
[ATTACK] Sent fake data: 5.0 kW
[ATTACK] Sent fake data: 120.0 kW
[ATTACK] Sent fake data: 150.0 kW
[ATTACK] Sent fake data: 10.5 kW
[ATTACK] Sent fake data: 10.5 kW
[ATTACK] Sent fake data: 10.5 kW
[ATTACK] Sent fake data: 150.0 kW
[ATTACK] Sent fake data: 5.0 kW
[ATTACK] Sent fake data: 150.0 kW
[ATTACK] Sent fake data: 10.5 kW
[ATTACK] Sent fake data: 150.0 kW
```

Рис. 6. Візуалізація аномалій телеметрії під час Spoofing атаки

Другий сценарій (Tampering) був розроблений для перевірки здатності цифрового двійника розпізнавати приховані маніпуляції та оцінки їхнього безпосереднього впливу на алгоритм оптимізації. Із застосуванням методу логічного проксіювання (еквівалента ARP Spoofing на прикладному рівні), атакуючий скрипт-перехоплювач успішно втрутився в інформаційний потік. Під час експерименту програмний емулятор було виведено на режим реального пікового навантаження із фактичною активною потужністю 83.2 кВт. Водночас скрипт зловмисника здійснював перехоплення пакетів у проміжному топіку, динамічно підміняв значення поля power_kw у корисному навантаженні JSON на стабільний показник 42.5 кВт і пересилав модифіковані дані у цільовий топік цифрового двійника.

Наслідки такої ін’єкції хибних даних виявилися критичними для логіки прийняття рішень. Алгоритм SMA цифрового двійника обробляв сфальсифікований стабільний потік даних (42.5 кВт), який знаходився в межах штатного режиму роботи. Відповідно, логічний компаратор, налаштований на відсікання значень понад 80 кВт, не спрацював, і життєво важлива команда на аварійне зниження навантаження не була згенерована. Відображення сфальсифікованого стану системи на панелі оператора проілюстровано на рис. 7. Інтерфейс Node-RED Dashboard демонстрував ідеально нормальний стан системи, створюючи хибну ілюзію безпеки, тоді як фізичний об’єкт де-факто перебував у стані критичного перевантаження. Проведений експеримент емпірично довів, що успішна компрометація вхідної телеметрії повністю нівелює ефективність аналітичних алгоритмів цифрового двійника, перетворюючи його з інструмента оптимізації на приховане джерело техногенної загрози.




```
[MITM] Real: 111.53kW  >  FAKE: 42.5kW
[MITM] Real: 107.08kW  >  FAKE: 42.5kW
[MITM] Real: 115.49kW  >  FAKE: 42.5kW
[MITM] Real: 103.78kW  >  FAKE: 42.5kW
[MITM] Real: 115.03kW  >  FAKE: 42.5kW
[MITM] Real: 116.35kW  >  FAKE: 42.5kW
[MITM] Real: 104.34kW  >  FAKE: 42.5kW
[MITM] Real: 109.27kW  >  FAKE: 42.5kW
[MITM] Real: 100.33kW  >  FAKE: 42.5kW
[MITM] Real: 100.71kW  >  FAKE: 42.5kW
[MITM] Real: 111.82kW  >  FAKE: 42.5kW
[MITM] Real: 105.63kW  >  FAKE: 42.5kW
[MITM] Real: 111.00kW  >  FAKE: 42.5kW
[MITM] Real: 101.91kW  >  FAKE: 42.5kW
[MITM] Real: 101.87kW  >  FAKE: 42.5kW
[MITM] Real: 98.76kW  >  FAKE: 42.5kW
[MITM] Real: 101.45kW  >  FAKE: 42.5kW
```

Рис. 7. Відображення сфальсифікованого стану системи на панелі оператора

5. Проєктування та оцінка ефективності комплексного механізму захисту.

З метою нейтралізації виявлених векторів атак розроблено та програмно реалізовано комплексну архітектуру багаторівневого захисту, яка охоплює транспортний та прикладний рівні інформаційної екосистеми цифрового двійника.

Базову конфігурацію брокера повідомлень MQTT, що передбачала передавання даних у відкритому вигляді, було суттєво модернізовано. Впроваджено інфраструктуру відкритих ключів (PKI): згенеровано самопідписаний кореневий сертифікат, сертифікати сервера та відповідні ключі шифрування. У конфігураційному файлі mosquitto.conf активовано підтримку протоколу TLS, а обмін даними переведено на захищений мережевий порт 8883. Аналіз перехопленого комунікаційного трафіку за допомогою програмного засобу Wireshark підтвердив, що всі пакети надійно зашифровані та ідентифікуються як «Encrypted Application Data». Результат впровадження шифрування TLS для захисту каналу MQTT наведено на рис. 8. Це повністю унеможливлює пасивне перехоплення трафіку та базові атаки типу Man-in-the-Middle. Додатково застосовано політику жорсткого розмежування прав: заборонено анонімний доступ (allow_anonymous false) та налаштовано списки контролю доступу, згідно з якими програмний емулятор отримав права виключно на публікацію в топік телеметрії, а цифровий двійник лише на підписку.

Оскільки криптографічний захист транспортного рівня забезпечує виключно безпеку каналу зв’язку та не запобігає загрозам компрометації кінцевого вузла або крадіжки його облікових даних, логіку безпеки системи було посилено на прикладному рівні. До структури JSON-повідомлень, що формуються в емуляторі, інтегровано механізм генерації цифрового підпису на базі Hash-base Message Authentication Code. Перед відправленням даних Python-алгоритм формує конкатенований рядок, що містить поточні значення спожитої потужності та часову мітку (timestamp). Із використанням попередньо розподіленого секретного симетричного ключа обчислюється геш-значення за стандартом SHA-256, яке інтегрується у вихідний пакет (поле signature).

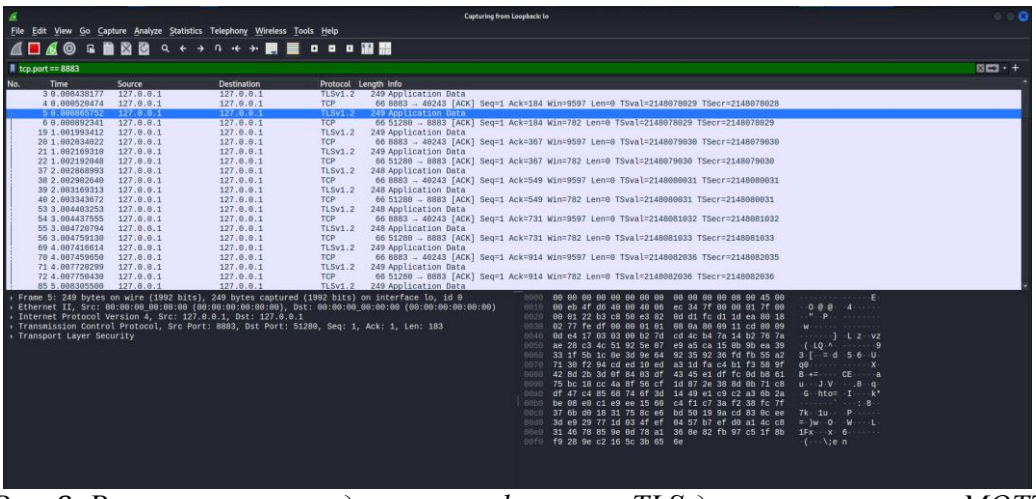


Рис. 8. Результат впровадження шифрування TLS для захисту каналу MQTT

На боці платформи Node-RED реалізовано спеціалізований функціональний вузол перевірки безпеки Security Check. Цей модуль виконує процедуру зворотної верифікації: приймає вхідний пакет, вилучає заявлений цифровий підпис, самостійно обчислює геш-значення отриманих метрик із використанням ідентичного секретного

ключа та виконує порівняння результатів. У разі спроби зловмисника модифікувати значення телеметрії (як було продемонстровано у другому сценарії атак), розрахований геш не збігається з переданим. Це однозначно свідчить про порушення цілісності даних, внаслідок чого пакет автоматично відкидається системою як скомпрометований. Крім того, впроваджено механізм валідації часових міток: пакети, мережева затримка яких перевищує встановлений поріг у 5 секунд, ігноруються. Це надійно захищає систему від Replay Attacks, унеможливлюючи деструктивне використання перехоплених легітимних, але застарілих пакетів.

6. Аналіз ефективності впроваджених рішень.

Повторне експериментальне моделювання всього спектра досліджуваних векторів атак на захищений локальний стенд емпірично підтвердило повну ефективність розроблених контрзаходів. Спроби реалізації Spoofing-атак успішно блокувалися на транспортному рівні брокером повідомлень внаслідок неможливості проходження несанкціонованим клієнтом процедури TLS-аутентифікації. Водночас усі пакети телеметрії, модифіковані в процесі імітації атак типу Man-in-the-Middle, миттєво ідентифікувалися та відкидалися логічним вузлом перевірки на прикладному рівні. Схема логічної верифікації цілісності даних із застосуванням криптографічних підписів у середовищі Node-RED наведена на рис. 9. Завдяки роботі цього вузла алгоритм енергетичної оптимізації продовжував коректно функціонувати, спираючись виключно на підтверджені та валідні масиви даних.

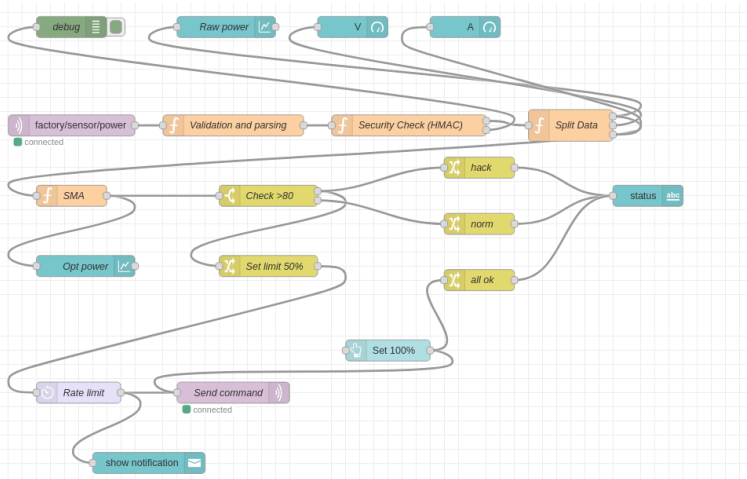


Рис. 9. Схема логічної верифікації цілісності даних у Node-RED

Кількісна оцінка продуктивності захищеної системи продемонструвала, що впровадження процесів інкапсуляції TLS та обчислення криптографічних гешів за стандартом HMAC-SHA256 призвело до прогнозованого збільшення мережевого навантаження та загального часу затримки від сенсора до обчислювального ядра на 15–20%. Однак, зважаючи на специфіку задач порогової оптимізації та енергетичного балансування, де інтервал дискретизації традиційно вимірюється секундами, виявлені накладні витрати обчислювальних та мережевих ресурсів класифікуються як цілком прийнятні. Запропонований комплексний механізм безпеки не чинить критичного впливу на загальну швидкодію та операційну стабільність цифрового двійника, гарантуючи при цьому необхідний рівень стійкості до кібернетичних втручань у потоки телеметрії.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У межах представленої роботи було здійснено глибоке дослідження проблеми забезпечення інформаційної та операційної безпеки цифрових двійників, які інтегруються у виробничі кіберфізичні системи для виконання критично важливих завдань оптимізації енергоспоживання. Проведений комплексний аналіз доводить, що архітектурна суть цифрових двійників – необхідність безперервного, автоматичного, двостороннього обміну телеметричними даними між рівнем сенсорів і хмарними аналітичними платформами – неминуче призводить до розмивання традиційного периметра безпеки і формує колосальну поверхню для кібератак. В умовах енергетичного менеджменту пріоритет тріади CIA радикально зміщується від забезпечення конфіденційності до гарантування абсолютної цілісності та доступності даних, оскільки компрометація метрик споживання миттєво призводить до прийняття алгоритмами руйнівних фізичних рішень.

Завдяки розгорнутому за парадигмою Software-in-the-Loop експериментальному середовищу, яке об’єднало Python-емулятор, брокер Mosquitto MQTT та логічне ядро Node-RED, було наочно продемонстровано фатальні наслідки експлуатації базових комунікаційних протоколів IIoT. Практичне моделювання атак типу Spoofing та Tampering довело, що зловмисник, вдаючись до логічного проксіювання та ін’єкції хибних даних, здатен повністю дезорієнтувати алгоритми оптимізації на базі SMA. Підміняючи показники критичного перевантаження 83.2 кВт на фіктивні нормальні значення 42.5 кВт, атака блокує спрацювання захисних механізмів скидання навантаження та створює для оператора небезпечну ілюзію штатного функціонування, що в умовах реального підприємства загрожує повномасштабним блекаутом та руйнуванням обладнання.

Результати дослідження підтверджують, що єдиним ефективним підходом до нейтралізації таких комплексних загроз є відмова від парадигми абсолютної довіри до телеметрії та перехід до концепції багаторівневого захисту. Практична імплементація транспортного шифрування TLS 1.3 із суворою аутентифікацією повністю усунула ризики пасивного перехоплення та неавторизованого підключення. Водночас вирішальним фактором забезпечення стійкості цифрового двійника до тонких маніпуляцій стало впровадження механізмів криптографічної перевірки цілісності корисного навантаження на прикладному рівні за допомогою цифрових підписів HMAC-SHA256 та валідації часових міток. Оцінка ефективності цього архітектурного рішення підтвердила здатність системи миттєво відхиляти фальсифіковані пакети, зберігаючи стабільність контуру енергоменеджменту. Доведено, що неминучі накладні витрати обчислювальних ресурсів на виконання криптографічних операцій (збільшення затримки на 15-20%) є цілком прийнятним компромісом для забезпечення надійності систем енергетичної оптимізації.

Запропонована архітектура багаторівневого захисту формує надійний фундамент, проте вимагає подальшого розширення. По-перше, надзвичайно перспективним напрямом є інтеграція методів фізично-обґрунтованого виявлення аномалій (Physics-based Anomaly Detection), які дозволять цифровому двійнику крос-валідувати енергетичні метрики з термодинамічними та кінематичними законами роботи обладнання, виявляючи навіть ті атаки, що пройшли криптографічну перевірку. По-друге, критично важливим є дослідження впливу технологій блокчейну на забезпечення аудиту незмінності історичних даних для систем машинного навчання та предиктивного обслуговування, за умови розробки алгоритмів подолання консенсусних

затримок. Таким чином, розроблена комплексна система захисту, що поєднує мережеву безпеку та прикладну валідацію даних, забезпечує стійкість цифрового двійника до основних векторів кібератак, зберігаючи при цьому необхідну функціональність алгоритмів енергетичної оптимізації.

ПОДЯКА

This work was supported by a grant from the Simons Foundation International (SFI-PD-Ukraine-00014577; O.G.).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. IT-Enterprise. (2024). *Digital twin: Technology transforming manufacturing.* https://www.it.ua/knowledge-base/technology-innovation/cifrovoj-dvojnik-digital-twin
2. Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access, 8*, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358
3. Alcaraz, C., & Lopez, J. (2022). Digital twin: A comprehensive survey of security threats. *IEEE Communications Surveys & Tutorials, 24*(3), 1475–1503. https://doi.org/10.1109/COMST.2022.3171465
4. Suhail, S., Hussain, R., Jurdak, R., Oracevic, A., Salah, K., Hong, C. S., & Matulevičius, R. (2022). Blockchain-based digital twins: Research trends, issues, and future challenges. *ACM Computing Surveys, 54*(11s), 1–34. https://doi.org/10.1145/3517189
5. He, Y., Mendis, G. J., & Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid, 8*, 2505–2516. https://doi.org/10.1109/TSG.2017.2703842
6. Tao, F., Zhang, M., & Nee, A. Y. C. (2019). *Digital twin driven smart manufacturing.* Academic Press. https://doi.org/10.1016/C2018-0-02206-9
7. National Institute of Standards and Technology. (2019). *Considerations for managing Internet of Things (IoT) cybersecurity and privacy risks* (NIST IR 8228). https://doi.org/10.6028/NIST.IR.8228
8. Eckhart, M., & Ekelhart, A. (2018). Towards security-aware virtual environments for digital twins. In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security* (pp. 61–72). https://doi.org/10.1145/3198458.3198464
9. Savchenko, T., Lutska, N., Vlasenko, L., Sashnova, M., Zahorulko, A., Minenko, S., Ibaiev, E., & Tytarenko, N. (2025). Risk analysis and cybersecurity enhancement of digital twins in dairy production. *Technology Audit and Production Reserves, 2*(2(82)), 37–49. https://doi.org/10.15587/2706-5448.2025.325422
10. InHandGO. (2024). *Revolutionizing Industry 4.0: How digital twins are powering the future of smart manufacturing.* https://inhandgo.com/blogs/articles/revolutionizing-industry-4-0-how-digital-twins-are-powering-the-future-of-smart-manufacturing
11. Suleiman, R., Maradapu, V. V. S., Wei, Y., & Wang, C. (2025). Blockchain for security in digital twins. *Future Internet, 17*(9), 385. https://doi.org/10.3390/fi17090385
12. TXOne Networks. (2023). *Digital twins: The benefits and challenges of revolutionary technology in automotive industries.* https://www.txone.com/blog/digital-twins-benefits-and-challenges-revolutionary-technology-in-automotive-industries/

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023

Tetiana Savchenko
Candidate of Technical Sciences, Associate Professor,
Associate Professor of the Department of Informatics
National University of Kyiv-Mohyla Academy, Kyiv, Ukraine
ORCID: 0000-0002-8884-5360
tsavchenko@ukma.edu.ua

Yaroslav Bondar
Bachelor's Student in Cybersecurity and Information Protection
National University of Kyiv-Mohyla Academy, Kyiv, Ukraine
ya.bondar@ukma.edu.ua

RISKS AND VULNERABILITIES OF DIGITAL TWINS IN MANUFACTURING SYSTEMS DURING ENERGY CONSUMPTION OPTIMIZATION

Abstract. The presented scientific research is dedicated to an in-depth systemic analysis of information security issues arising during the implementation, integration, and operation of Digital Twin technology in modern cyber-physical manufacturing systems focused on energy consumption optimization. In the context of the rapid transition to Industry 4.0, accompanied by total digitalization and the convergence of computing resources with physical processes, Digital Twin technology has become an indispensable tool for achieving high energy efficiency, minimizing losses, and ensuring the sustainable development of the industry. However, as this research demonstrates, the breakdown of traditional physical network isolation (Air Gap) and the establishment of a continuous bidirectional data flow between the sensor equipment of operational technology and information technology cloud platforms create a large attack surface. The article provides a fundamental overview of the architectural vulnerabilities of the Industrial Internet of Things, analyzes the weaknesses of key communication protocols, and develops an extended threat classification based on an adapted STRIDE model. Special emphasis is placed on threats to data integrity, particularly False Data Injection (FDI) attacks and their cumulative impact on optimization algorithms. Within the methodological and practical part of the work, a representative local environment was built using the Software-in-the-Loop principle, which integrates a Python energy consumption emulator, the Eclipse Mosquitto message broker, and the event-driven Node-RED platform. Through the practical modeling of Spoofing and Tampering attacks, it was demonstrated that the compromise of input telemetry can completely disorient simple moving average (SMA) algorithms, hiding critical overloads from operators and blocking emergency load-shedding mechanisms, which under real conditions leads to the physical destruction of infrastructure. To counter the identified threat vectors, a comprehensive Defense in Depth architecture was developed, implemented, and verified. The proposed approach combines cryptographic encryption of the transport channel using the TLS protocol with strict authentication and application-level payload integrity validation using HMAC-SHA256 digital signatures and timestamp verification. The efficiency evaluation demonstrated the complete neutralization of the studied attacks while maintaining the functional stability of the control algorithms and resulting in entirely acceptable overheads on network delays at a level of 15-20%.

Keywords: digital twin; cybersecurity; energy consumption optimization; industrial internet of things; false data injection; defense in depth; cryptographic authentication; cyber-physical systems.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. IT-Enterprise. (2024). *Digital twin: Technology transforming manufacturing.* https://www.it.ua/knowledge-base/technology-innovation/cifrovoj-dvojnik-digital-twin
2. Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access, 8*, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358
3. Alcaraz, C., & Lopez, J. (2022). Digital twin: A comprehensive survey of security threats. *IEEE Communications Surveys & Tutorials, 24*(3), 1475–1503. https://doi.org/10.1109/COMST.2022.3171465

4. Suhail, S., Hussain, R., Jurdak, R., Oracevic, A., Salah, K., Hong, C. S., & Matulevičius, R. (2022). Blockchain-based digital twins: Research trends, issues, and future challenges. *ACM Computing Surveys, 54*(11s), 1–34. https://doi.org/10.1145/3517189
5. He, Y., Mendis, G. J., & Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid, 8*, 2505–2516. https://doi.org/10.1109/TSG.2017.2703842
6. Tao, F., Zhang, M., & Nee, A. Y. C. (2019). *Digital twin driven smart manufacturing.* Academic Press. https://doi.org/10.1016/C2018-0-02206-9
7. National Institute of Standards and Technology. (2019). *Considerations for managing Internet of Things (IoT) cybersecurity and privacy risks* (NIST IR 8228). https://doi.org/10.6028/NIST.IR.8228
8. Eckhart, M., & Ekelhart, A. (2018). Towards security-aware virtual environments for digital twins. In *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security* (pp. 61–72). https://doi.org/10.1145/3198458.3198464
9. Savchenko, T., Lutska, N., Vlasenko, L., Sashnova, M., Zahorulko, A., Minenko, S., Ibaiev, E., & Tytarenko, N. (2025). Risk analysis and cybersecurity enhancement of digital twins in dairy production. *Technology Audit and Production Reserves, 2*(2(82)), 37–49. https://doi.org/10.15587/2706-5448.2025.325422
10. InHandGO. (2024). *Revolutionizing Industry 4.0: How digital twins are powering the future of smart manufacturing.* https://inhandgo.com/blogs/articles/revolutionizing-industry-4-0-how-digital-twins-are-powering-the-future-of-smart-manufacturing
11. Suleiman, R., Maradapu, V. V. S., Wei, Y., & Wang, C. (2025). Blockchain for security in digital twins. *Future Internet, 17*(9), 385. https://doi.org/10.3390/fi17090385
12. TXOne Networks. (2023). *Digital twins: The benefits and challenges of revolutionary technology in automotive industries.* https://www.txone.com/blog/digital-twins-benefits-and-challenges-revolutionary-technology-in-automotive-industries/

Отримано редакцією журналу / Received: 20.01.26
Прорецензовано / Revised: 15.02.26
Схвалено до друку / Accepted: 26.03.26