УДК 004.94 DOI https://doi.org/10.35546/kntu2078-4481.2025.1.2.11

О. В. КОЗЛОВСЬКИЙ
аспірант кафедри програмних засобів і технологій
Херсонський національний технічний університет
ORCID: 0009-0006-1864-1107

М. В. ЖАРІКОВА
доктор технічних наук,
професор кафедри програмних засобів і технологій
Херсонський національний технічний університет
ORCID: 0000-0001-6144-480X

РОЗРОБКА МОДЕЛІ БЕЗПЕКИ ДЛЯ БАГАТОАГЕНТНОЇ МЕРЕЖІ В КІБЕРФІЗИЧНІЙ СИСТЕМІ

У цьому дослідженні аналізуються сучасні кіберфізичні системи, що складаються з багатьох агентів, організованих у мережу. Однією з основних загроз для таких систем є атаки на фізичному рівні, зокрема перехоплення та підробка даних сенсорів, що може спричинити дестабілізацію або компрометацію системи. Кожен агент виконує роль основної структурної одиниці мережі, містить сенсори, актуатори та панель керування, що забезпечують взаємодію з фізичним середовищем та прийняття рішень у режимі реального часу. Ефективність кожного агента визначається набором ключових метрик, які дозволяють розраховувати його коефіцієнт корисності в рамках мережі. Основними параметрами для розрахунку є безпека, якість обслуговування та витрати. Ці параметри формуються на основі фізичних даних, отриманих із сенсорів, що забезпечує їхню адаптивність до змін у середовищі.

Головною метою дослідження є розробка моделі безпеки для фізичного рівня системи, яка використовує Марківський процес прийняття рішень для оцінки ефективності мережі та вибору оптимальної стратегії безпеки. Оптимізація стратегії триває до досягнення найвищого показника ефективності, який обчислюється як зважена сума індивідуальних коефіцієнтів корисності агентів. Вагові коефіцієнти визначаються відповідно до конкретних вимог користувачів та умов експлуатації, що забезпечує гнучкість моделі для різних сценаріїв використання. Застосування методу навчання з підкріпленням в свою чергу дозволить системі аналізувати як індивідуальні, так і колективні потреби мережі та збільшувати ефективність на основі отриманих знань. Запропонована модель сприяє підвищенню рівня захисту та стабільності мережі в кіберфізичній системі, забезпечуючи баланс між якістю обслуговування, витратами та безпекою.

Ключові слова: *кіберфізична система, агент, конфігурація передачі сигналу, безпека фізичного рівня системи, мережа.*

O. V. KOZLOVSKYI
Postgraduate Student at the Department of Software and Technologies
Kherson National Technical University
ORCID: 0009-0006-1864-1107

M. V. ZHARIKOVA
Doctor of Technical Sciences,
Professor at the Department of Software and Technologies
Kherson National Technical University
ORCID: 0000-0001-6144-480X

DEVELOPMENT OF A SECURITY MODEL FOR A MULTI-AGENT NETWORK IN A CYBER-PHYSICAL SYSTEM

In this article, the modern cyber-physical systems composed of multiple agents organized into a network, was analyzed. One of the main threats to such systems is physical-layer attacks, including signal interception and sensor data manipulation, which can lead to system destabilization or compromise. Each agent serves as a fundamental structural unit of the network, incorporating sensors, actuators, and a control panel to interact with the physical environment and make real-time decisions. The efficiency of each agent is determined by a set of key metrics that allow for calculating its utility coefficient within the network. The primary parameters for this calculation include security, quality of service, and costs. These parameters are derived from physical data obtained from sensors, ensuring their adaptability to environmental changes.

The main goal of this research is to develop a security model for the physical layer of the system, utilizing a Markov decision process to assess network efficiency and select the optimal security strategy. The optimization process continues

until the highest efficiency score is achieved, calculated as a weighted sum of the individual utility values of the agents. The weighting coefficients are determined according to specific user requirements and operating conditions, ensuring the model's flexibility for various usage scenarios. The application of reinforcement learning will enable the system to analyze both individual and collective network needs, improving its efficiency based on acquired knowledge. The proposed model enhances the security and stability of the network within a cyber-physical system while maintaining a balance between service quality, costs, and security.

Key words: *cyber-physical system, agent, transmission policy, physical layer security, network.*

Постановка проблеми

Кіберфізичні системи (CPS) – це поєднання цифрових (ПЗ, мережеві технології) та фізичних компонентів у замкненій системі управління, що забезпечують виконання різних задач. Вони оперують у циклічному режимі за допомогою панелі керування та актуаторів та збирають дані з сенсорів. Прикладом такої системи може слугувати «розумний будинок», що має операційні вимоги та задачі, які можна вирішити за допомогою різних сервісів комунікації між компонентами, зокрема сервісу наднадійної комунікації з низькою затримкою (URLLC) або сервісу масового зв'язку (mMTC) [1]. Завдяки цим сервісам компоненти CPS можуть дистанційно керуватися та адаптуватися до змінних умов середовища у будь-який момент.

Очікується, що подібні системи стануть ще більш адаптивними за рахунок впровадження "когнітивних" здібностей для сервісів зв'язку з появою шостого покоління бездротових комунікацій 6G [1]. Такі здібності головним чином будуть використовуватися для оптимізації доступних мережевих та обчислювальних ресурсів. Інтеграція подібного сервісу забезпечить вищу швидкість передачі даних, меншу затримку та розширені можливості для штучного інтелекту. Таку систему називають когнітивною CPS.

Бездротові комунікаційні системи через свою природу є вразливими до масштабних атак, які можуть виникати на фізичному рівні передачі даних. Наприклад, такі атаки можуть збирати або фальсифікувати дані сенсорів через пристрої перехоплення сигналів, що може призвести до нестабільної роботи системи або її компрометації. Таким чином безпека фізичного рівня системи є ключовим фактором надійності в CPS.

Аналіз останніх досліджень і публікацій

Дослідники Д. Вонґ та Б. Бей у своїй роботі [2] надають критерії результативності стратегії безпеки для фізичного рівня системи (PLS). Стратегією або політикою безпеки в даному контексті є набір правил та механізмів для захисту комунікацій та стабільності системи на фізичному рівні. Одним зі способів захисту передачі даних є розподіл ресурсів на передавальному вузлі (роутер). Інші стратегії безпеки включають в себе різні методи обробки сигналів, такі як направлення сигналу та попереднє кодування. Перший метод дозволяє спрямовувати сигнал в потрібному напрямку для зменшення можливості перехоплення, а другий забезпечує зміну фази та амплітуди сигналу. Оскільки сигнал попередньо кодується відповідно до каналу приймача, то продуктивність будь-якого іншого вузла чи каналу, що розташований на іншій позиції, буде знижуватися.

Р. Шобер та І. Коллінс у своєму дослідженні [3] пропонують механізм захисту, що аналізує параметри каналу передавача та обирає антену (або групу антен), яка забезпечує найкраще співвідношення сигнал/перешкода для приймача і одночасно мінімізує потужність сигналу, що доходить до зловмисника. Потім обирається оптимальний вузол, який може містити в собі декілька таких передавачів, що зменшує ризик перехоплення або перешкод під час передачі даних. Цей метод застосовується у розгалужених мережах, що складаються з приймальних та передавальних антен (MIMO). Проблема полягає в тому, що такі механізми засновуються на виборі статичних стратегій захисту фізичного рівня системи, що робить систему вразливою у динамічному середовищі.

Інші дослідники пропонують систему захисту на базі ігрової теорії [4], щоб моделювати взаємодію між зловмисником та системою. В такій системі використовується прогнозуюча модель для передбачення майбутніх станів системи, що дозволяє аналізувати можливі загрози та приймати рішення на основі прогнозованих змін.

Варто згадати і про технологію неортогонального множинного доступу (NOMA), коли декілька компонентів можуть передавати сигнал на одній частоті [5]. Ця технологія використовується для створення моделі перехоплення, що описує можливість прослуховування зловмисником та дозволяє оптимізувати коефіцієнти розподілу потужності, тобто потужності яку передавач надає кожному компоненту в бездротовій мережі (антена, маршрутизатор).

Дослідник Ф. Ґебрі у своїй роботі пропонує механізм підсилення сигналу для двосторонньої когнітивної радіомережі [6], в якій два пристрої можуть передавати і отримувати дані, використовуючи спектр, який не заблокований іншими легітимними пристроями. Один пристрій може передавати дані, а інший – здійснювати їх обробку та надавати зворотній зв'язок для подальшої комунікації. Залежно від умов, система може змінювати параметри передачі сигналу, зменшуючи перешкоди і покращуючи ефективність мережі.

У дослідженні В. Янґа описується підхід до спільної оптимізації доступу до мережі та потужності передачі сигналу для вторинних передавачів [7], щоб максимізувати безпеку всіх вузлів мережі. Це стосується когнітивних радіомереж, де вторинні пристрої мають обмежений доступ до частот, що вже використовуються основними користувачами спектру (діапазону частот).

Вищезгадані дослідження розглядають ефективність конкретного механізму за стабільних параметрів каналу, в той час як дане дослідження запропонує механізм адаптивного вибору PLS стратегії та конфігурації передачі сигналу на основі розрахунку коефіцієнтів корисності наявних стратегій і враховуючи динамічну природу фізичного середовища.

Формулювання мети дослідження

Метою даного дослідження є розробка механізму вибору найкращої стратегії безпеки для фізичного рівня та конфігурації передачі сигналу в кіберфізичній системі на основі розрахунку коефіцієнтів корисності кожного агента в мережі.

Викладення основного матеріалу дослідження

Кіберфізичні системи можна розглядати як особливу форму розподіленої системи з інтегрованими замкнутими контурами, що складаються з вузлів (агентів), для виконання спеціалізованих завдань. У цій роботі ми розглянемо особливий тип кіберфізичної системи, у якій всі агенти обмінюються своїми станами та інформацією через бездротові канали. На рисунку 1 зображено модель CPS у вигляді мережі агентів, де кожен агент містить панель керування, сенсор (датчик) та актуатор (виконавчий механізм).

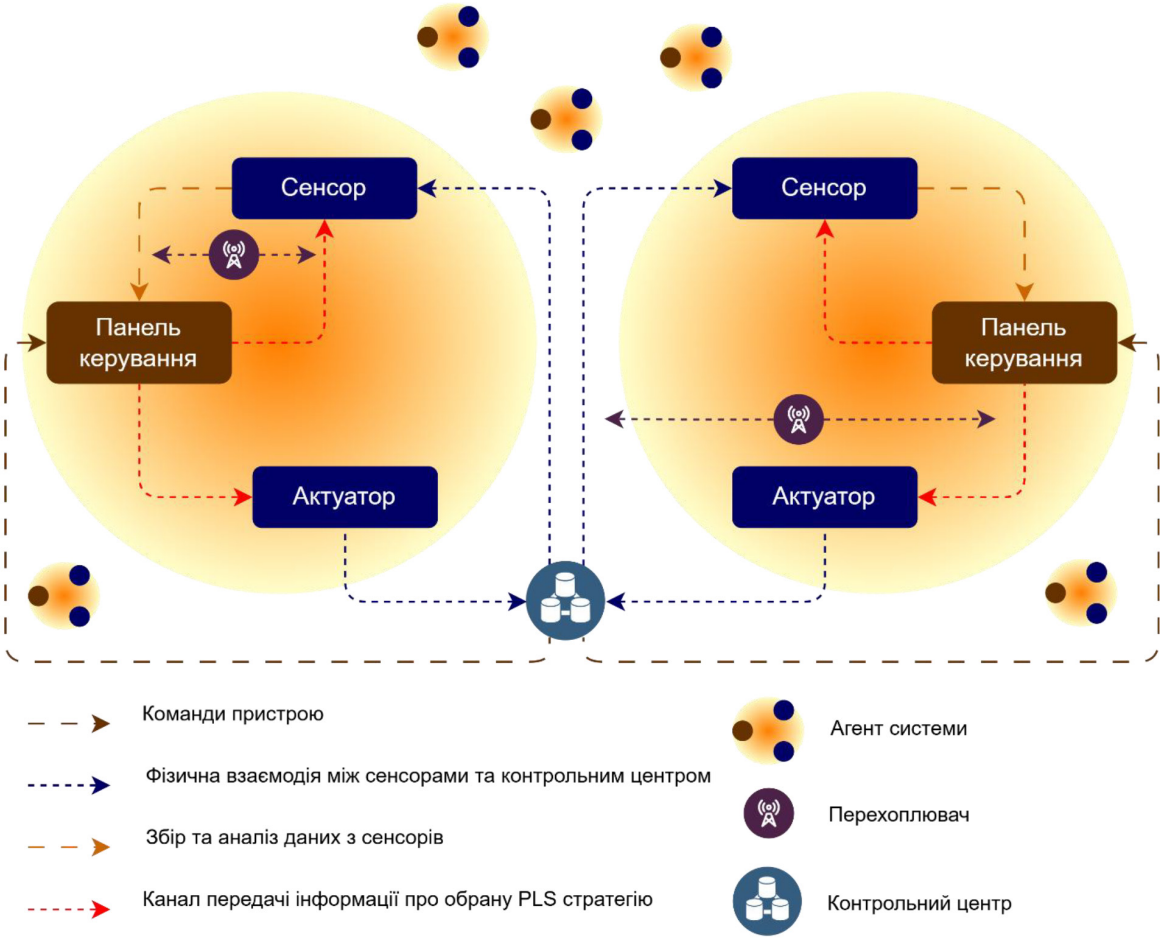

Рис. 1. Модель CPS у вигляді мережі агентів з урахуванням можливих атак

Зазвичай CPS система складається з декількох агентів, які взаємодіють між собою, наприклад промислові мережі або система «транспорт-до-всього» (V2X), де агентами є транспортні засоби, які обмінюються даними про трафік, погодні умови та потенційні небезпеки [8].

Когнітивні CPS мають здатність локально (на периферії мережі) аналізувати середовище та приймати рішення в обмеженому обсязі. У даній роботі ми розглянемо систему, де агенти визначають коефіцієнти корисності обраних стратегій на основі даних, що дозволяє їй аналізувати середовище та приймати рішення в реальному часі. У багатьох розподілених системах середовище також постійно змінюється (наприклад, додаються нові агенти, оновлюються моделі атак, збільшується рівень шуму). Тому CPS повинні адаптуватися до змінних умов середовища, щоб забезпечити безперервну стабільність мережі, а також відповідати вимогам цільового застосування. Ця важлива еволюція буде можливою, якщо агенти когнітивних CPS зможуть адаптувати свої операційні параметри:

технологію передачі сигналу (LoRa), потужність передачі, PLS стратегію тощо, щоб одночасно забезпечувати бажану продуктивність системи та безпеку [9].

У науковій літературі запропоновано безліч механізмів захисту, що інтегрують різні можливості підвищення безпеки на фізичному рівні, враховуючи різні сценарії атак та архітектури систем. Проте у більшості вищезгаданих робіт моделюється середовище з незмінними умовами передачі даних та статичною моделлю атак. Використання лише однієї стратегії безпеки для агентів CPS може стати вузьким місцем для загальної безпеки та надійності мережі. Для вирішення цієї проблеми ми опишемо адаптивний механізм прийняття рішень, який обиратиме найкращу PLS стратегію (шифрування, направлення сигналу) з доступної множини стратегій та оптимальну конфігурацію передачі сигналу. Ми визначимо контрольний центр, який буде відстежувати діяльність агентів у мережі. Він здійснюватиме аналіз і прийматиме рішення щодо оптимальних стратегій безпеки, адаптуючи їх до поточних умов. Панель управління збиратиме інформацію про зовнішні фактори та зміни навколишнього середовища з датчиків, оцінюватиме потенційні ризики та стан комунікаційних ресурсів. Вона також буде налаштовувати параметри фізичного рівня всієї системи відповідно до потреб агентів. Сенсори та актуатори в свою чергу отримуватимуть команди від контрольного центру для оновлення параметрів передачі сигналу. Замість того, щоб використовувати одну фіксовану стратегію безпеки, система адаптивно обиратиме найкращу PLS політику в кожен дискретний момент часу, що підвищить ефективність та стабільність роботи мережі.

Основним критерієм ефективності компонента мережі є так звана «корисність». Це означає, що для визначення продуктивності агента у CPS необхідно розрахувати середнє арифметичне між трьома взаємопов'язаними параметрами: безпека, якість обслуговування та витрати [10]. Вага кожного параметру визначається операційними вимогами, що висуваються користувачами або пристроями до системи. Оскільки існують різні експлуатаційні вимоги до застосування систем CPS (рої дронів, розумні енергомережі), ми можемо налаштовувати ваги для досягнення найкращих результатів у реальних умовах (використовуючи дані про середовище, отримані з датчиків) та враховуючи наявні ресурси.

Хоча для кожного агента можна вибрати окрему оптимальну стратегію, інтеграція цих стратегій без урахування взаємодії між агентами та спільних витрат ресурсів може призвести до неефективності роботи мережі. У таких випадках важливо враховувати колективні рішення, де агенти приймають рішення не лише з огляду на власні вимоги, а й на вплив своїх дій на всю систему. Це дозволяє досягти оптимальних результатів для мережі в цілому.

Замість того, щоб оптимізувати продуктивність кожного агента окремо, пропонується інтегрувати результати (корисність) всіх агентів у правильних пропорціях, щоб отримати коефіцієнт мережевої корисності, який відображатиме загальний вплив агентів на продуктивність мережі. Цей підхід дозволяє покращити ефективність і забезпечити стабільність роботи мережі CPS.

Підсумовуючи, ми запропонуємо адаптивний механізм вибору стратегії PLS та конфігурації передачі сигналу для системи безпеки фізичного рівня, заснований на алгоритмі Марківського процесу прийняття рішень (MDP). Механізм дозволить максимізувати як індивідуальні, так і спільні винагороди (числове значення, що характеризує якість виконання агентом задачі). Винагорода окремих агентів обчислюється на основі коефіцієнту корисності, що є середнім арифметичним параметрів якості обслуговування, безпеки та витрат. Крім того, цей механізм забезпечить адаптацію стратегії в реальному часі залежно від змін у мережевих умовах та зовнішньому середовищі. Таким чином, система зможе ефективно реагувати на непередбачувані зміни та оптимізувати роботу мережі.

Розрахунок коефіцієнту корисності агентів у мережі

У нашій моделі ми припускаємо що мережа CPS складається з I агентів, де i – це індекс агента. Через різні вимоги та умови навколишнього середовища, кожен агент може мати різні набори доступних PLS стратегій та конфігурацій передачі сигналу, які позначаються як $\mathcal{P}_i$ та $\mathcal{R}_i$ для i-го агента відповідно. Ці множини можна визначити як $\mathcal{P}_i = \left\{P_i^1, P_i^2, \ldots, P_i^{N_i}\right\}$ та $\mathcal{R}_i = \left\{R_i^1, R_i^2, \ldots, R_i^{M_i}\right\}$, де N_i – число доступних PLS стратегій, а M_i – число доступних конфігурацій передачі для i-го агента. Якщо ми розглядаємо часові інтервали в яких відбувається передача даних, то t – це індекс часового інтервалу, що дорівнює $t = 1, 2, \ldots, T$. Оскільки умови в динамічному середовищі змінюються, то і кількість часових інтервалів, в яких приймаються рішення щодо вибору стратегій, збільшується. Тому необхідно визначати яка конфігурація була обрана для кожного часового інтервалу t з множин $\mathcal{P}_i$ та $\mathcal{R}_i$ для кожного агента, де $i = 1, 2, \ldots, I$. За цих умов $\mathcal{K}$ та $\mathcal{L}$ це множини обраних PLS стратегій та конфігурацій для кожного агента в момент часу t, де $\mathcal{K} = \left\{P_1^{k_1}[t], P_2^{k_2}[t], \ldots, P_I^{k_I}[t]\right\}$ та $\mathcal{L} = \left\{R_1^{l_1}[t], R_2^{l_2}[t], \ldots, R_I^{l_I}[t]\right\}$. У цих рівняннях k_i та l_i є індексами обраних PLS стратегій та конфігурацій для кожного агента, що задовольняють наступні обмеження: $k_i \le N_i$, $l_i \le M_i$, $\forall_i \in \{1, 2, \ldots, I\}$.

На рисунку 2 графічно зображено модель розрахунку коефіцієнту корисності агентів у мережі та вибору оптимальної стратегії.

У даному дослідженні ми припускаємо, що агенти є незалежними, і кожен з них містить передавальний та приймальний вузол. Ми визначимо ці вузли як a_i та b_i для i-го агента відповідно. Кількість антен у передавальному вузлі i-го агента позначається як W_i. Ми припускаємо, що кожен приймальний вузол оснащений однією антеною. Також вважається, що перехоплювач (зловмисник) має лише одну антену. Коефіцієнт згасання сигналу

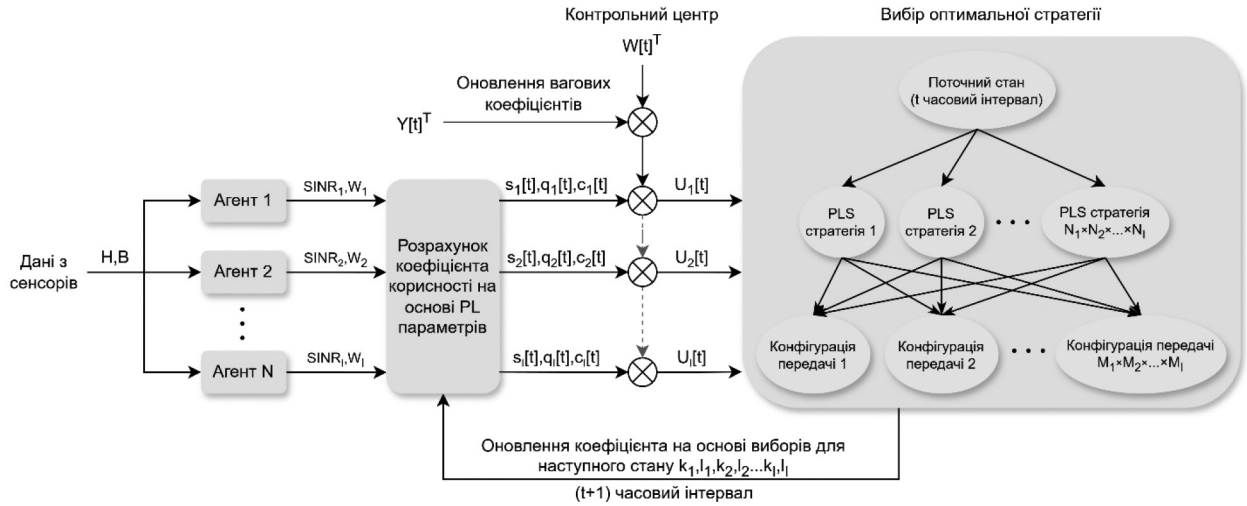

Рис. 2. Модель розрахунку коефіцієнту корисності агентів у мережі

між вузлами, що описує як сигнал послаблюється під впливом середовища, позначається як c_i та d_j. Коефіцієнт можна розрахувати за наступною формулою: $h_{c_i d_j} = [h_{c_i d_j}(1), h_{c_i d_j}(2), \ldots, h_{c_i d_j}(W_i)]^T$, де $c \in \{a, b\}$, $d \in \{a, b\}$, $i, j \in \{1, 2, \ldots, I\}$ та $i \neq j$.

Розгляньмо чотири основні PLS стратегії, а саме: штучна інтерференція для підканалів в системах із динамічним розподілом частот (SC-AN), перешкоди в реальному часі під час прийому та передачі сигналів (FD-AI), направлення сигналу з одночасним зашумленням інших каналів (AN) та формування променя (BF). Всі ці стратегії спрямовані на створення перешкод зловмиснику, забезпечуючи додатковий рівень захисту даних шляхом зниження ефективності атак на канали зв'язку [11].

SC-AN стратегія застосовує штучну інтерференцію до певних підканалів у багатоканальній системі, щоб зашкодити потенційному зловмиснику, ускладнюючи його здатність перехоплювати та розшифровувати важливий сигнал. Кожен підканал може використовувати свою частоту для передачі даних, що дозволяє зменшити рівень інтерференції між каналами та підвищити ефективність передачі.

FD-AI стратегія створює інтерференцію в реальному часі під час прийому та передачі сигналів. Головною метою є ускладнення перехоплення сигналу зловмисником. При цьому легітимний користувач має вбудовані механізми для компенсації цієї інтерференції, що дозволяє йому отримувати правильний сигнал.

AN стратегія передбачає, що сигнал легітимного користувача спрямовується таким чином, щоб мінімізувати шум у напрямку отримувача, одночасно генеруючи максимальні перешкоди для зловмисників по інших каналах.

BF стратегія дозволяє направляти сигнал з антени в певному напрямку для покращення якості передачі і прийому сигналу. Це досягається шляхом коригування фази та амплітуди сигналів, що надходять від декількох антен, щоб створити напрямлений пучок сигналу (промінь).

У мережі потужність передавання сигналу для всіх агентів позначається як R^s. Відношення сигналу до шуму та інтерференції (SINR) у вузлі b_i та у часовому інтервалі t визначається за формулою [3]:

$$SINR_i^b[t] = \frac{\left| v_i^H h_{a_i b_i} \right|^2 R^s}{\gamma_i^{k_i}(Y_i[t] + N_{0i})},$$

де γ_i – це сума перешкод, які виникають через передачу сигналів іншими агентами в мережі, N_{0i} – це дисперсія (розсіювання) шуму на приймачі i-го агента, що враховує фоновий шум, який може бути викликаний тепловими ефектами чи електромагнітними перешкодами, $\gamma_i^{k_i}$ – коефіцієнт помилки при усуненні інтерференції, що показує наскільки ефективно вибрана PLS стратегія забезпечує якість зв'язку. Вектор передавальної інформації, тобто вектор, який визначає, як антенна система передавача повинна формувати сигнал, щоб максимізувати його прийом на приймачі, для i-го агента визначається як $v_i = h_{a_i, b_i}$. Інтерференція від інших агентів розраховується за формулою:

$$Y_i[t] = \sum_{j \neq i} (Y_{ji}^s[t] + Y_{ji}^c[t]),$$

де γ_i^c – інтерференція, яку отримує агент через сигнали, що передаються іншими агентами, і яка може впливати на якість прийому повідомлення, Y_i^a – інтерференція, що виникає від сигналів безпеки (наприклад, від штучного шуму), переданих іншими агентами. Сигнал безпеки – це спеціально створений або модифікований сигнал, що

використовується для створення перешкод потенційним зловмисникам, не впливаючи при цьому на легітимних користувачів мережі. Загалом, інтерференцію від передачі повідомлення можна виразити через наступну формулу:

$$Y_{ji}^{s}[t] = \left| v_i^H h_{a_j b_i} \right|^2 R^s.$$

Для стратегії SC-AN значення $Y_{ji}^{c}[t]$, що позначає інтереренцію, яка виникає від сигналу безпеки *j*-го агента в певний момент часу, розраховується за формулою $Y_{ji}^{c,SC-AN}[t] = \left| h_{a_j b_i} \right|^2 R_j^{l_j}[t]$, де $R_j^{l_j}[t]$ виражає обрану конфігурацію передачі сигналу *j*-го агента. Для стратегії FD-AI відповідно: $Y_{ji}^{c,FD-AI}[t] = \left| h_{b_j b_i} \right|^2 R_j^{l_j}[t]$. Для стратегії AN значення $Y_{ji}^{c}[t]$ вираховується наступним чином: $Y_{ji}^{c,AN}[t] = \left| \alpha_{a_j b_j} * h_{a_j b_i} \right|^2 R_j^{l_j}[t]$, де $\alpha_{a_j b_j} = \dfrac{\mathcal{NS}(h_{a_j b_j})}{h_{a_j b_j}}$, $\mathcal{NS}(h_{a_j b_j})$ – нормалізація сигналу $h_{a_j b_j}$. Для стратегії формування променя (BF) формула виглядатиме як: $Y_{ji}^{c,B}[t] = 0$. Подібним чином можна виразити SINR для зв'язку від передавача i -го агента до перехоплювача за допомогою формули [6]:

$$SINR_i^e[t] = \frac{\left| v_i^H h_{a_i e} \right|^2 R^s}{(Y_{ie}[t] + N_{0e})}.$$

В цьому випадку інтерференція буде розраховуватися за формулою:

$$Y_{ie}[t] = \sum_{j \neq i} Y_{je}^{s}[t] + \sum_{i=1}^{I} Y_{ie}^{c}[t],$$

де $Y_{je}^{s}[t] = \left| v_j h_{a_j e} \right|^2 R^s$. Отже, крім інтерференції, викликаної сигналами безпеки від інших агентів, сигнал безпеки самого агента також впливає на зниження ефективності роботи перехоплювача.

Для того аби визначити оптимальну стратегію для захисту агентів або системи загалом необхідно розрахувати параметри корисності. Замість того щоб визначати точні значення параметру безпеки для кожного конкретного випадку, ми оцінимо загальний рівень безпеки в середовищі.

Спочатку необхідно змоделювати позицію перехоплювача як двовимірну випадкову величину, що підпорядковується нормальному розподілу Гауса. Далі обчислюється секретна ємність для кожного *i*-го агента, використовуючи різницю між ємністю каналу передачі даних агента C_B^i і ємністю каналу між агентом і перехоплювачем $C_E^i(x, y)$, де (x, y) – це координати перехоплювача на фізичній поверхні *S*. Секретна ємність – це міра здатності каналу зв'язку забезпечити безпеку даних, що передаються. Вона залежить від різниці між сигналом, який отримує легітимний приймач, та сигналом, який зловмисник може перехопити $C_{sec}^i(x, y) = \max\{0, (C_B^i - C_E^i(x, y))\}$.

Наступним етапом є агрегування результатів по всій поверхні *S*, зважуючи значення секретної ємності та ймовірність присутності перехоплювача в кожній точці. Враховуючи ці дані можна обчислити відносний рівень безпеки агента за формулою $s_i[t] = \dfrac{\Omega_i[t]}{\max_{j \in \{1,2,\ldots,I\}}[\Omega_j[t]]}$.

Для кожного агента у мережі його параметр якості обслуговування вимірюється через рівень SINR на його приймачі, тобто через коефіцієнт, що характеризує співвідношення між корисним сигналом, що передається, та інтерференцією, що може погіршити якість сигналу. Для визначення цього слід поділити рівень SINR на приймачі *i*-го агента в момент часу *t* на максимальний рівень SINR серед усіх агентів *j* на момент часу *t* [10]. Отже якість обслуговування це відношення $q_i[t] = \dfrac{SINR_i^b[t]}{\max_{j \in \{1,2,\ldots,I\}}[SINR_j^b[t]]}$.

Параметр витрат визначає наскільки ефективно агент використовує апаратні та енергетичні ресурси, що впливає зокрема на час роботи системи та споживану енергію. Основними факторами, що визначають витрати роботи агента є кількість активних антен $W_i[t]$ та потужність передачі сигналу $R_i^{l_i}[t]$. Таким чином витрата ресурсів визначається за наступною формулою $c_i[t] = \dfrac{W_i[t] + R_i^{l_i}[t]}{\max_{j \in \{1,2,\ldots,I\}}[W_j[t] + R_j^{l_j}[t]]}$.

Корисність агента визначається з урахуванням критеріїв корисності та їх ваг w_i в момент часу *t*, що визначають наскільки важливі для агента ці критерії. Ваги задовольняють умову нормалізації $w_{i,s}[t] + w_{i,q}[t] + w_{i,c}[t] = 1$ та визначаються за формулою:

$$w_i[t] = \gamma_s[t]\gamma_q[t]\gamma_c[t] \circ \begin{bmatrix} w_{i,s}[t] \\ w_{i,q}[t] \\ w_{i,c}[t] \end{bmatrix}.$$

Також необхідно враховувати коефіцієнт переходу між стратегіями $\delta_i[t]^{k_i \to k_i'}$. Якщо стратегія для агента не змінюється, то коефіцієнт переходу буде дорівнювати 1, а якщо змінюється, тоді буде < 1, що в свою чергу негативно вплине на коефіцієнт корисності. Коефіцієнт також буде зменшуватися, якщо витрати на зв'язок збільшуються, тому, замість використання параметра $c_i[t]$, використовується різниця $(1 - c_i[t])$, що дозволяє враховувати негативний вплив витрат. Формула розрахунку коефіцієнту корисності агента з врахуванням описаних аспектів виглядає так:

$$U_i^{k_i,l_i}[t] = \delta_i[t]^{k_i \to k_i'} * \left(w_i[t]^T \times \begin{bmatrix} s_i[t] \\ q_i[t] \\ (1 - c_i[t]) \end{bmatrix} \right).$$

Останнім етапом є вибір оптимальної PLS стратегії та конфігурації передачі сигналу в CPS із застосуванням алгоритму Марківського процесу прийняття рішень. Марківський процес прийняття рішень – це математична модель для опису процесів прийняття рішень у стохастичному середовищі. У MDP агент приймає рішення на основі поточного стану і вибирає дію, яка визначає перехід до нового стану, що має певну ймовірність. MDP використовується для оптимізації рішень, де мета агента – максимізувати загальну винагороду за допомогою вибору оптимальних дій на основі поточного стану.

Кожен агент в CPS має певний стан в момент часу t, що відповідає обраній PLS стратегії та певній конфігурації передачі. Дії між станами полягають у зміні або продовженні поточної конфігурації для кожного часового інтервалу, враховуючи обчислені винагороди, засновані на значеннях корисності. При індивідуальному рішенні агент має N_i можливих дій на першому етапі та M_i дій на другому. При спільному рішенні кількість можливих дій значно збільшується: $N_1 \times N_2 \times \ldots \times N_I$ – на першому етапі та $M_1 \times M_2 \times \ldots \times M_I$ – на другому.

Далі відбувається оцінка корисності для кожного стану, що буде визначати винагороду. Винагорода – це числовий показник, який оцінює, наскільки добре агент виконав дію в певному стані, застосовуючи певну стратегію та конфігурацію. Вона допомагає агенту навчитися вибирати найкращі дії для досягнення максимальної ефективності. Агент порівнює корисність різних станів та вибирає найкращий варіант. Після кількох ітерацій алгоритм знаходить оптимальну PLS стратегію і конфігурацію передачі. Якщо один агент виконує перехід від однієї стратегії до іншої, то інші агенти це враховують у своїх розрахунках, що забезпечує адаптивність в багатоагентній мережі.

Застосовуючи метод навчання з підкріпленням агент навчається на основі рішень та винагород, вибираючи найкращі стратегії та конфігурації для забезпечення високої якості і безпеки при мінімальних витратах. Цей метод дозволяє агентам в CPS навчатися через досвід та оптимізувати свої стратегії, щоб отримати найкращі результати в умовах обмежених ресурсів та динамічно змінюваного середовища.

Висновки

В даному дослідженні було запропоновано механізм вибору оптимальних стратегій безпеки на фізичному рівні (PLS) для агентів кіберфізичної системи. В основі механізму лежить розрахунок величини корисності агента, як комбінації трьох основних критеріїв та їх ваг: безпеки, якості обслуговування та витрат. Ці критерії базуються на даних фізичного середовища, отриманих з сенсорів. За допомогою алгоритму MDP приймається рішення щодо вибору оптимальної стратегії та конфігурації для кожного агента у певному часовому інтервалі. Ітерації відбуваються до тих пір, поки обрана стратегія не буде мати найвищий коефіцієнт корисності, враховуючи динамічні зміни в середовищі. Застосування методу навчання з підкріпленням в свою чергу дозволить системі навчатися на прийнятих рішеннях та підвищувати ефективність системи в цілому.

Список в икористаної літератури

1. Khan A.A., Rehmani M.H., Rachedi A. Cognitive-Radio-Based Internet of Things: Applications, Architectures, Spectrum Related Functionalities, and Future Research Directions. *IEEE Wireless Communications*. 2017. Vol. 24, no. 3. P. 17–25. URL: https://doi.org/10.1109/mwc.2017.1600404

2. A Survey of Optimization Approaches for Wireless Physical Layer Security / D. Wang et al. *IEEE Communications Surveys & Tutorials*. 2019. Vol. 21, no. 2. P. 1878–1911. URL: https://doi.org/10.1109/comst.2018.2883144

3. Transmit Antenna Selection for Security Enhancement in MIMO Wiretap Channels / N.Yang et al. *IEEE Transactions on Communications*. 2013. Vol. 61, no. 1. P. 144–154. URL: https://doi.org/10.1109/tcomm.2012.12.110670

4. Xu Z., Zhu Q. A cyber-physical game framework for secure and resilient multi-agent autonomous systems. *2015 54th IEEE Conference on Decision and Control (CDC)*, Osaka, 15–18 December 2015. 2015. URL: https://doi.org/10.1109/cdc.2015.7403026

5. Dynamic power allocation scheme with clustering based on physical layer security / T. Liu et al. *IET Communications*. 2018. Vol. 12, no. 20. P. 2546–2551. URL: https://doi.org/10.1049/iet-com.2018.5544

6. Energy Efficiency Analysis of Cooperative Jamming in Cognitive Radio Networks with Secrecy Constraints / F.Gabry et al. *IEEE Wireless Communications Letters*. 2015. Vol. 4, no. 4. P. 437–440. URL: https://doi.org/10.1109/lwc.2015.2432802

7. Energy-efficient optimization for physical layer security in large-scale random CRNs / X. Xu et al. *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*, Nanjing, China, 15–17 October 2015. 2015. URL: https://doi.org/10.1109/wcsp.2015.7341314

8. A Survey on Security and Privacy of 5G Technologies: Potential Solutions, Recent Advancements, and Future Directions / R. Khan et al. *IEEE Communications Surveys & Tutorials*. 2020. Vol. 22, no. 1. P. 196–248. URL: https://doi.org/10.1109/comst.2019.2933899

9. A Physical Layer Security Framework for Cognitive Cyber-Physical Systems / O. A. Topal et al. *IEEE Wireless Communications*. 2020. Vol. 27, no. 4. P. 32–39. URL: https://doi.org/10.1109/mwc.01.1900543

10. Joint optimization of QoS and security for differentiated applications in heterogeneous networks / Z. M. Fadlullah et al. *IEEE Wireless Communications*. 2016. Vol. 23, no. 3. P. 74–81. URL: https://doi.org/10.1109/mwc.2016.7498077

11. Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey / A. Mukherjee et al. *IEEE Communications Surveys & Tutorials*. 2014. Vol. 16, no. 3. P. 1550–1573. URL: https://doi.org/10.1109/surv.2014.012314.00178

References

1. Khan, A. A., Rehmani, M. H., & Rachedi, A. (2017). Cognitive-Radio-Based Internet of Things: Applications, Architectures, Spectrum Related Functionalities, and Future Research Directions. *IEEE Wireless Communications*, 24(3), 17–25. https://doi.org/10.1109/mwc.2017.1600404

2. Wang, D., Bai, B., Zhao, W., & Han, Z. (2019). A Survey of Optimization Approaches for Wireless Physical Layer Security. *IEEE Communications Surveys & Tutorials*, 21(2), 1878–1911. https://doi.org/10.1109/comst.2018.2883144

3. Yang, N., Yeoh, P. L., Elkashlan, M., Schober, R., & Collings, I. B. (2013). Transmit Antenna Selection for Security Enhancement in MIMO Wiretap Channels. *IEEE Transactions on Communications*, 61(1), 144–154. https://doi.org/10.1109/tcomm.2012.12.110670

4. Xu, Z., & Zhu, Q. (2015). A cyber-physical game framework for secure and resilient multi-agent autonomous systems. *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE. https://doi.org/10.1109/cdc.2015.7403026

5. Liu, T., Han, S., Meng, W., Li, C., & Peng, M. (2018). Dynamic power allocation scheme with clustering based on physical layer security. *IET Communications*, 12(20), 2546–2551. https://doi.org/10.1049/iet-com.2018.5544

6. Gabry, F., Zappone, A., Thobaben, R., Jorswieck, E. A., & Skoglund, M. (2015). Energy Efficiency Analysis of Cooperative Jamming in Cognitive Radio Networks with Secrecy Constraints. *IEEE Wireless Communications Letters*, 4(4), 437–440. https://doi.org/10.1109/lwc.2015.2432802

7. Xu, X., Cai, Y., Yang, W., Yang, W., Hu, J., & Yin, T. (2015). Energy-efficient optimization for physical layer security in large-scale random CRNs. *У 2015 International Conference on Wireless Communications & Signal Processing (WCSP)*. IEEE. https://doi.org/10.1109/wcsp.2015.7341314

8. Khan, R., Kumar, P., Jayakody, D. N. K., & Liyanage, M. (2020). A Survey on Security and Privacy of 5G Technologies: Potential Solutions, Recent Advancements, and Future Directions. *IEEE Communications Surveys & Tutorials*, 22(1), 196–248. https://doi.org/10.1109/comst.2019.2933899

9. Topal, O. A., Demir, M. O., Liang, Z., Pusane, A. E., Dartmann, G., Ascheid, G., & Kur, G. K. (2020). A Physical Layer Security Framework for Cognitive Cyber-Physical Systems. *IEEE Wireless Communications*, 27(4), 32–39. https://doi.org/10.1109/mwc.01.1900543

10. Fadlullah, Z. M., Wei, C., Shi, Z., & Kato, N. (2016). Joint optimization of QoS and security for differentiated applications in heterogeneous networks. *IEEE Wireless Communications*, 23(3), 74–81. https://doi.org/10.1109/mwc.2016.7498077

11. Mukherjee, A., Fakoorian, S. A. A., Huang, J., & Swindlehurst, A. L. (2014). Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey. *IEEE Communications Surveys & Tutorials*, 16(3), 1550–1573. https://doi.org/10.1109/surv.2014.012314.00178