УДК 004.092

ІНТЕЛЕКТУАЛЬНА СИСТЕМА ВИЯВЛЕННЯ АНОМАЛІЙ ТА ІДЕНТИФІКАЦІЇ ПРИСТРОЇВ РОЗУМНИХ БУДИНКІВ ІЗ ЗАСТОСУВАННЯМ КОЛЕКТИВНОЇ КОМУНІКАЦІЇ

А. О. Нічепорук, А. А. Нічепорук, О. С. Савенко, А. Д. Казанцев
Хмельницький національний університет

Анотація. *В роботі запропоновано інтелектуальну систему виявлення аномалій та ідентифікації пристроїв розумних будинків із застосуванням колективної комунікації. Концепція роботи система заснована на отриманні виграшу від об'єднання розумних будинків в соціальну мережу в частині підвищення безпеки як окремо взятого розумного будинку, так і всієї соціальної мережі поєднаних розумних будинків. Ключовою особливістю системи є комунікація кластерів розумних будинків між собою для обміну інформації про наявні профілі розумних пристроїв в білих списках кожного кластеру.*

Ключові слова: *розумний будинок, мережевий трафік, профілювання поведінки, комунікація, аномальна поведінка,класифікація*

Вступ

Людство вже досить значний період часу отримує переваги від використання розумних пристроїв, поєднаних в мережу для покращення та автоматизації життєдіяльності. Розумні будинки, система охорони здоров'я, системи автоматизації та промисловий Інтернет речей (IIoT) є основними прикладами систем на базі використання мережі поєднаних розумних пристроїв. Наступною сходинкою в еволюції таких систем є їх об'єднання в соціальні мережі. Об'єднання, наприклад, розумних будинків у соціальну мережу утворює вищий ієрархічний рівень їхньої взаємодії, яка продукує нові можливості та переваги [1], зокрема в аспекті управління, зберігання та обробки інформації, підвищення рівня обслуговування кінцевих користувачів, попередження та колективне реагування на надзвичайні події, в економічному плані, в плані безпеки, тощо. Нажаль з точки зору безпеки функціонування таких мереж, об'єднання розумних будинків у соціальні мережі не позбавляє їх вразливостей, що були притаманні їх складовим компонентам на нижчому рівні.

Метою дослідження є отримання виграшу від об'єднання розумних будинків в соціальну мережу в частині підвищення безпеки як окремо взятого розумного будинку, так і всієї соціальної мережі поєднаних розумних будинків. Ми представляємо кожен розумний будинок як кластер, що є автоматизованою системою з набором розумних пристроїв, що поєднанні між собою та шлюзом, комунікація між якими здійснюється із залучення стека протоколів TCP/IP. Механізм взаємодії кластерів має характер кожен з кожним, тобто з одного кластера, через маршрутизатор з виходом в Інтернет, є можливість отримати / передати інформацію на / з всіх інших кластерів. Таким чином задача дослідження полягає у прийняті рішення щодо відсутності / наявності аномальної поведінки в комунікаційному середовищі розумного будинку шляхом моніторингу його мережевого трафіку та ідентифікації мережевих потоків даних від конкретних пристроїв, а також, в разі потреби, залучення інформації про аномальну активність із інших кластерів, що входять у соціальну мережу колективної комунікації.

1. Вразливості та атаки на IoT мережі

З року в рік зловмисники намагаються скомпрометувати та заволодіти приватною інформацією через здійснення атак на локальні та корпоративні мережі. З точки зору IoT мереж, специфіка їх роботи розкриває ще більше вразливостей та наявних вузьких місць для здійснення на них атак у порівнянні із «звичайними» мережами [2]. Найбільш вагомими причинами, за якими злочинці вибирають для атаки саме IoT-пристрої, зокрема розумні будинки, є їх постійна доступність в мережі Інтернет, обмежені обчислювальні можливості, що тим самим унеможливлює встановлення систем захисту безпосередньо на самих пристроях, вразливості пов'язані з авторизацією / аутентифікацією пристроїв в мережі, гетерогенність самих пристроїв та середовища їх взаємодії, вразливості у веб-інтерфесах, відсутність належної уваги зі сторони кінцевих користувачів (тобто «поставив і забув», що часто проявляється у залишенні стандартних логінів та паролів, відсутності перевірки оновлень, тощо). Сукупність цих факторів призводить до значного інтересу

серед зловмисників, які намагаються реалізувати все нові і нові кібер-атаки.

Основними видами атак на мережі Інтернету речей є DoS / DDoS атаки та Man in the middle (MITM) атаки [3, 4]. Спільною їх метою є захоплення керування пристроєм та використання його для власних цілей. Цими цілями може бути наприклад участь пристрою в ботнеті (ботнет Mirai [5]), читання, перехоплення, спотворення інформації, шляхом компрометації каналу зв'язку (наприклад атака ARP Spoofing), відключення пристрою від мережі, що тим самим призводить до порушення уставленого процесу роботи всієї системи (наприклад відключення камер спостережень, датчиків руху, тощо) [6, 7]. Окрім того простота та гнучкість введення нових розумних пристроїв, додатків та сервісів у систему Smart House робить їх схожими на будівельні блоки [8], не розуміння принципів функціонування яких, може призвести до появи невідомих вразливих місць та їх поширення у великих масштабах.

2. Огляд попередніх досліджень

На сьогоднішній день, серед науковців проблемі виявлення кібер-атак на IoT мережі приділяється значна увага. Існуючі системи виявлення вторгнень в IoT мережі можна розділити на декілька основних груп: системи виявлення вторгнень засновані на сигнатурах [9], системи виявлення вторгнень на основі правил та системи виявлення аномалій [10-17]. Одним із найбільш перспективних напрямків є виявлення аномалій в мережевому трафіку.

Метод виявлення аномалій полягає у визначені характеристик мережевого трафіку при відсутності сторонніх впливів та атак, з подальшим відстеженням відхилень від цих номінальних показників. Оскільки процес виявлення аномалій перевіряє відхилення від звичайного функціонування трафіку, а не підписи атак, то дана методика є здатною до виявлення атак нульового дня, що якісно, вирізняє її серед інших [11]. Відомі роботи по виявлення аномальної активності в мережевому трафіку перш за все відрізняються у виборі ознак або параметрів, які визначають різницю між нормальним і шкідливим профілем, способом їх представлення та методами їх обробки. Розглянемо детальніше деякі з них.

У роботі [12] авторами було запропоновано фреймворк IOT-KEEPER, хостову систему, здатну виконувати онлайн класифікацію трафіку на мережевих шлюзах. Для представлення мережевого трафіку фреймворк використовує множину ознак, які не залежать від технології зв'язку розумних пристроїв, а залежить виключно від особливостей TCP / IP на кожному вузлі. Для ідентифікації різних типів діяльності розумних пристроїв використано нечітку кластеризацію C-means. Надалі фреймворк використовує ці характеристики для виявлення аномалій в процесі функціонування розумних пристроїв шляхом аналізу їх мережевого трафіку.

В [13] автори дослідили можливість використання коефіцієнта Херста для визначення рівня самоподібності мережевого трафіку, що впливає на здатність визначати типові робочі стани, а також на виявлення певних аномалій, таких як атака відмова в обслуговуванні, атака переповнення буфера та атака «людина посередині».

Автори роботи [14] використовують підхід глибокого аналізу пакетів, який залучає техніку бітових шаблонів. Корисне навантаження мережі розглядаються як послідовність байтів, що називається бітовим шаблоном, а вибір функцій виконується як накладання набору байт, що називається n-грамами. Коли відповідні біти збігаються з усіма позиціями, відбувається збіг між бітовим шаблоном і n-грамами.

У роботі [15] автори запропонували дворівневу модель виявлення аномальної активності для системи виявлення вторгнень у мережах Інтернету речей. Перший рівень запропонованої моделі проводить бінарну класифікацію, при якій мережевий трафік визначається як легітимний або як шкідливий. У випадку віднесення досліджуваного зразка мережевого трафіку до шкідливого, здійснюється перехід до другого рівня та визначення типу кібер-загрози, зокрема DDoS-HTTP, DDoS-TCP, DDoS-UDP, DoS-HTTP, DoS-TCP, DoS-TCP. Для проведення класифікації використано алгоритм машинного навчання Random Forest.

Підхід, що використовує приховану марківську модель для вивчення звичної діяльності користувачів у розумному будинку представлено у [16]. Представлений підхід використовує в якості спостереження інформацію, отриману від датчиків домашніх пристроїв, та здійснює навчання параметрів прихованих марківських моделей. Отримана модель використовується для виявлення аномальної активності в мережі розумного будинку.

Представлені роботи показали досить високий рівень ефективності виявлення аномальної поведінки у мережах Інтернету речей, проте вони перш за все орієнтуються на захист однієї окремої екосистеми розумного будинку, не беручи до уваги можливий виграш при забезпеченні безпеки в разі залучення багатьох кластерів розумних будинків.

3. Профілювання пристроїв розумних будинків

Незважаючи на ідентичність фізичного середовища та протоколу передачі інформації в TCP/IP мережах характеристики мережевого трафіка для розумних пристроїв та звичайних не інтелектуальних вузлів будуть відрізнятися. Така ситуація пояснюється природою та основним призначенням розумних пристроїв. Головна мета розумних пристроїв полягає у періодичному моніторингу процесів навколишнього середовища та обміну отриманою інформацією із іншими розумними пристроями або кінцевими споживачами (в ролі кінцевих споживачів виступають або інші розумні пристрої, або засоби опрацювання та/або відображення отриманої інформації на різних пристроях користувачів – бази даних, комп'ютери, смартфони, тощо). З огляду на специфіку роботи розумних пристроїв можна виділити наступні характеристики мережевого трафіку: період активності, період сну, розмір пакетів та кількість переданої інформації в межах сесії, частота та кількість DNS запитів [18].

Період активності відображає фазу життєвого циклу розумного пристрою, що проявляється у його активній мережевій взаємодії з іншими пристроями в мережі. Цією активністю виступає передача інформації про параметри фізичного середовища або синхронізуючі сигнали для підтримки комунікації із іншими учасниками M2M взаємодії. Відповідно такому періоду буде притаманний сплеск мережевої активності, що продукується розумним пристроєм. Огляд попередніх досліджень показав [18], що для значної кількості розумних пристроїв, зокрема розумних розеток Tp-link smart plug, зазвичай (близько 95%) цей період становить не більше 5 секунд.

Активність пристроїв змінюється періодами сну при яких, обмін пакетами в мережі за участю розумного пристрою та іншими учасниками в мережі відсутній. Для більшості розумних пристроїв такий період складає не більше 20 секунд.

Розмір пакетів та кількість переданої інформації в межах сесії за участю розумного пристрою теж може бути ключовим аспектом при профілюванні (розмежуванні поведінки) пристроїв розумних будинків. Зазвичай розмір порцій даних (packet size), що передаються розумними пристроями є невеликим. В той час обсяг переданої інформації в межах однієї сесії складає не більше 1 Кб.

Ще однією характеристикою, яка дозволяє розмежувати профілі розумних пристроїв є кількість та частота DNS запитів. З огляду на вузькоспеціалізований характер функціонування розумних пристрої частота та кількість DNS запитів є не високою. Найчастіше розумні пристрої оперюють доменними іменами фірм-виробників розумних пристрої, наприклад, колонка Amazon Echo здійснює DNS запити до softwareupdates.amazon.com, device-metrics-su, amazon.com, example.org, pindorama.amazon.com та pool.ntp.org [11]. В той час як смарт-лампа LiFX lightbulb комунікує тільки з двома доменами v2.broker.lifx.co та pool.ntp.org.

Таким чином розуміння особливостей взаємодії розумних пристроїв дозволяє виокремити набір ознак (або атрибутів), що можуть використанні для опису поведінки та характеристик розумного пристрою в розумному будинку.

4. Система виявлення аномалій та ідентифікації пристроїв

Імплементація системи виявлення аномалій та ідентифікації пристроїв в розумних будинках заснована на моніторингу мережевого трафіку та формуванні профілів розумних пристроїв, які присутні у мережі. На основі цього здійснюється формування білого списку дозволених профілів функціонування пристроїв у кластері. Ключовою особливістю системи є комунікація кластерів між собою для обміну інформації про наявні профілі розумних пристроїв. Узагальнена функціональна схема системи виявлення аномалій та ідентифікації пристроїв наведено на рис. 1. Розглянемо детальніше складові запропонованої системи.

Система виявлення аномалій та ідентифікації пристроїв розміщується у внутрішній мережі розумного будинку та складається з наступних модулів: моніторингу мережевого трафіку, виявлення аномалій у мережевій поведінці, ідентифікації та прийняття рішення, модуля перетворення в PDML, модуля отримання ознак, класифікації. Інформація про виявленні пристрої зберігається в білому списку профілів функціонування розумних пристроїв. Передбачається, що розумні будинки поєднані у соціальну мережу.

Модуль ідентифікації та прийняття рішення організовує роботу системи в трьох режимах: моніторинг мережевого трафіку та виявлення аномалій; пошук профілю пристроїв у кластері; пошук профілю в інших кластерах (рис. 1). Нехай задано множину розумних пристроїв $D=\{d_1,d_2,...,d_n\}$, що представлені у білому списку та підключених у кластері C, де $C=\{c_1,c_2,...,c_m\}$, m – кількість кластерів. Вважатимемо, що у кожному кластері всі розумні пристрої представлені у білому списку, тобто білий список було згенеровано одразу після підключення всіх пристроїв. Окрім того передбачається, що розумні пристрої працювали в звичай-

ному режимі, без виконання операцій зміни прошивки, переконфігурування чи інших подібних операції. Також нехай кожний кластер має свою множину розумних пристроїв, і відповідно, свій білий список, який може відрізнятись від існуючих на інших кластерах.

В першому режимі здійснюється відстеження аномальної поведінки в мережевому трафіку. У випадку, якщо було виявлено відхилення від уставленого функціонування здійснюється перехід до другого режиму, в якому здійснюється класифікація мережевого трафіку та ідентифікація розумних пристроїв. Нехай задано структурований набір даних трафіку, тоді результатом роботи системи для заданого потоку IP пакетів (сесій) є співставлення його із множиною D у кластері c_j та визначення d_i, для якого поведінка заданого потоку пакетів є найбільш близькою. Якщо в результаті класифікації вдалося віднести підозрілий профіль до одного із профілів розумних пристроїв d_i, то в такому випадку система повертається в перший (штатний) режим роботи.

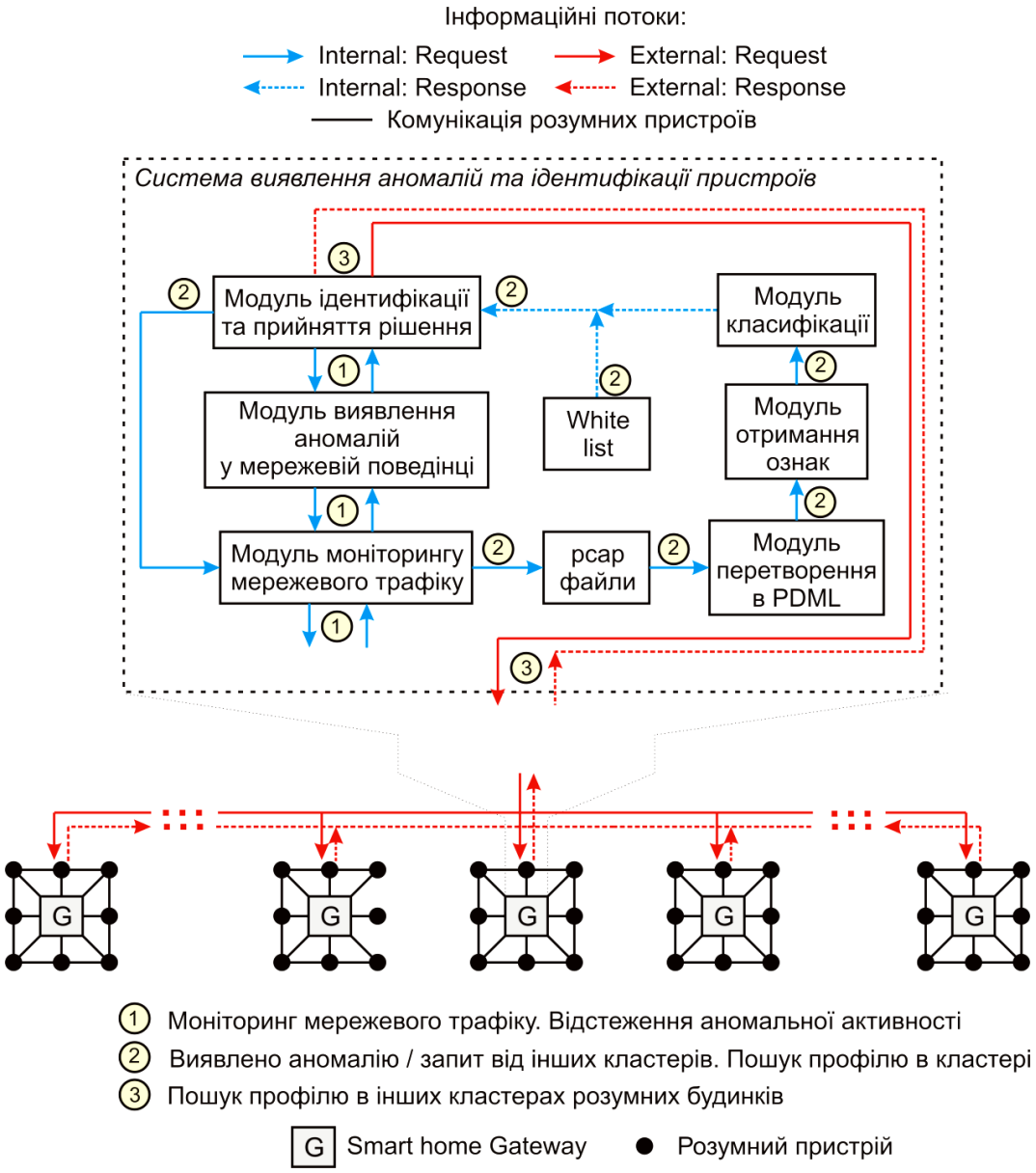

Рис. 1. Система виявлення аномалій та ідентифікації пристроїв розумних будинків на основі колективної комунікації

В протилежному випадку, якщо у кластері c_j задану послідовність пакетів не віднесено до одного із профілів d_i (не перевищено порогове значення), послідовність пакетів позначається як «невідома послідовність пакетів» та здійснюється перехід до третього режиму. Даний режим передбачає запит до інших кластерів $c_k, 1 < k \le m-1$ на предмет зіставлення профілю

послідовності пакетів, отриманих у кластері c_j, із власними білими списками.

Після проведення класифікації та ідентифікації на кожному із кластерів c_k, результат надсилається у кластер c_j, в якому модуль ідентифікації та прийняття рішення робить остаточний висновок. Для формування висновку модуль ідентифікації та прийняття рішення у кластері c_j із результатів відповідей всіх кластерів c_k утворює список, в якому кожен елемент представляє собою або тип пристрою в знайденому кластері $d_i^{c_k}$ або «невідома послідовність пакетів». Результат роботи системи $\phi(c_j)$ в кластері c_j для сформованого списку обчислюється як:

$$\phi(c_j)=\frac{\sum_{k=1}^{k\le m-1} h_{c_k}\cdot w_{c_k}}{\sum_{k=1}^{k\le m-1} u_{c_k}\cdot w_{c_k}}, \qquad (1)$$

де m – кількість кластерів;

h_{c_k} – результат кластеру c_k в якому послідовність пакетів з кластеру c_j визначена як тип пристрою, що міститься в білому списку кластера c_k, $h_{c_k}=\{x \mid x\in 0\vee 1\}$;

u_{c_k} – результат кластеру c_k в якому послідовність пакетів з кластеру c_j визначена як «невідома послідовність пакетів», тобто профіль поведінки відсутній в білому списку кластера c_k, $u_{c_k}=\{x \mid x\in 0\vee 1\}$;

w_{c_k} – ваговий коефіцієнт важливості результату.

Ваговий коефіцієнт важливості результату для кластеру c_k, обчислюється як відношення кількості розумних пристроїв в кластері c_k до загальної кількості всіх розумних пристроїв у всіх кластерах:

$$w_{c_k}=\frac{p_{c_k}}{\sum_{i=1}^{i\le m} p_{c_i}}, \qquad (2)$$

де p_{c_k} – кількість розумних пристроїв у білому списку кластера c_k; p_{c_i} – кількість розумних пристроїв у білому списку кластера c_i.

Якщо результат роботи системи $\phi(c_j)$ більший за одиницю, то послідовність пакетів, що утворюють підозрілий профіль, вдалося віднести до одного із профілів розумних пристроїв d_i в одному із кластерів у мережі:

$$\phi(c_j)=\begin{cases} d_i^{c_k} & if\ \phi(c_j)\ge 1 \\ \text{"unknown"} & if\ \phi(c_j)<1 \end{cases}, where\ c_j, c_k \in C. \qquad (3)$$

Модуль моніторингу мережевого трафіку використовується для сканування трафіку, що генерується розумними пристроями та отримання вхідних даних у вигляді послідовності TCP пакетів. Дані зберігаються у вигляді pcap файлів, кожен з яких представляє собою множину TCP-сесій. Початок та кінець сесії визначались прапорами SYN та FIN відповідно. В рамках кожної сесії пакети згруповані на основі чотирьох полів заголовка:

$$\langle src\_ip, src\_port, dst\_ip, dst\_port\rangle, \qquad (4)$$

де src_ip – IP адреса джерела, src_port – номер порту джерела, dst_ip– IP адреса призначення, dst_port – номер порту призначення.

Модуль моніторингу трафіку залучається як для відстеження аномальної активності (в складі модуля виявлення аномалій) в режимі реального часу так і для отримання «сирих даних» мережевого трафіку, які в подальшому будуть опрацьовуватись для підготовки до процесу класифікації мережевого трафіку.

Модуль виявлення аномалій використовується для відстеження критичних характеристики мережі у реальному часі. Роботу цього модуля можна представити у вигляді тригера, який формує сигнал тривоги у випадку наявності підозрілих характеристик. Цими характеристиками виступають обсяг трафіку, використання смуги пропускання, збільшення кількості підключень на один/ з одного порту TCP, збільшення кількості підключень на одну/ з однієї IP-адреси. Значення зазначених вище характеристик отримується на основі статистичної оцінки роботи мережі на протязі часу t_a від моменту початкового налаштування та підключення всіх розумних пристроїв у мережу.

З метою перевірки наявності профілю у білому списку, що описує послідовності пакетів мережевого трафіку характерну для кожного із розумних пристроїв d_i, залучається *модуль класифікації*, в основі якого, використано алгоритм Random Forest [19]. Для ототожнення розумних пристроїв з потоком мережевих пакетів використано MAC адресу пристроїв в заголовках пакетів. З метою отримання ознак для проведення класифікації (тобто профілю поведінки потоку мере-

жевих пакетів) весь фіксований період моніторингу трафіку *T* розбивався на часові інтервали t_i, кожний з яких складався з набору сесій *s*. Якщо сесія *s* розпочиналась в межах часового інтервалу t_i, але закінчувалась за межами цього інтервалу, то така сесія відносилась до інтервалу t_i. Якщо сесія *s* закінчувалась до закінчення інтервалу часу t_i, час до наступної сесії відносився до часового інтервалу t_i.

Далі в межах кожного інтервалу t_i отримується набір ознак $F = \langle f_1, f_2, ..., f_{10} \rangle$, які описують поведінку потоку пакетів з огляду на специфіку функціонування розумних пристроїв (табл. 1). Слід відзначити, що з огляду на те, що корисне навантаження пакетів (payload) є зашифрованим, жодна із ознак не враховує це поле.

Таблиця 1

Ознаки, що використовуються для опису поведінки потоку пакетів

№	Ознака
1	Період активності
2	Період сну
3	Кількість унікальних DNS-запитів
4	Частота DNS-запитів
5	Середній розмір пакетів
6	Максимальний розмір пакетів
7	Відношення кількості переданих до кількості прийнятих байт інформації
8	Сума кількості переданих та кількості прийнятих байт інформації
9	Загальна кількість переданих байт
10	Час між першим та останнім пакетом
11	NTP інтервал

Для цього весь зібраний мережевий трафік конвертується в Packet Description Markup Language (PDML) за допомогою *модуля перетворення в PDML* [20]. PDML представляє поля заголовків пакетів у XML форматі, що дозволяє отримати доступ до всіх атрибутів пакетів, які використовуються в якості складових ознак.

В результаті роботи класифікатора для кожного часового інтервалу t_i, який представлений вектором ознак *F*, отримується вектор ймовірностей, кожен елемент якого, визначає належність потоку пакетів до розумного пристрою d_i у кластері c_j, $r^{c_j}(t_i) = \left\langle r_{d_1}^{c_j}, r_{d_2}^{c_j}, ..., r_{d_n}^{c_j} \right\rangle$, так, що $\sum_{i=1,j=1}^{n,m} r_{d_i}^{c_j} = 1$, де *m* – кількість кластерів в соціальній мережі, *n* – кількість розумних пристроїв в c_j кластері.

Послідовність пакетів вважається віднесеною до одного із розумних пристроїв d_i, якщо $r_{d_i}^{c_j} \geq \delta$, де δ – порогове значення ймовірності, яка визначає належність до класу розумного пристрою d_i у кластері c_j.

Остаточний вибір розумного пристрою d_i у кластері c_j (або позначення зібраного профілю як «невідома послідовність пакетів») для всього періоду моніторингу *T* визначається як модальне значення зі всіх інтервалів t_i. Таким чином *білий список профілів*, що міститься у кожному кластері, фактично представляє собою натреновану модель класифікатора та містить промаркований набір значень ознак при заданому періоду моніторингу *T* та заданими часовими інтервалами t_i.

Таким чином на основі застосування колективної комунікації для обміну інформацією про пошук підозрілого профілю в білих списках інших кластерів мережі здійснюється прийняття рішення щодо відсутності / наявності аномальної поведінки в комунікаційному середовищі розумного будинку. Знання цієї інформації дозволяє користувачу або адміністратору мережі виконати превентивні дії по блокуванню трафіку мережі або відключення пристрою від мережі, який продукує аномальну активність.

5. Експериментальні дослідження

Для проведення експерименту, що визначав ефективність запропонованої системи виявлення аномалій та ідентифікації пристроїв розумних будинків із застосуванням колективної комунікації, було розгорнуто соціальну мережу, яка складалась із двох кластерів. Окрім маршрутизатора та RPi в кожному кластері, було задіяно сім розумних пристроїв, чотири з яких, були розміщені в першому кластері, решта три в другому (табл. 2). Для збору мережевого трафіку, що генерувався розумними пристроями було використано утиліту Tshark [21], яка була встановлена на RPi під управлінням Raspberry Pi OS.

Процес збору мережевого трафіку тривав 2 тижні кожного дня із заданими інтервалами. Для автоматизації збору мережевого трафіку було написано скрипт, планування запуском якого було реалізовано за допомогою утиліти Cron Job. Для ототожнення розумних пристроїв з потоком мережевих пакетів використано MAC адресу пристроїв в заголовках пакетів, що дозволило розмежувати трафік розумних пристрої один від одного. Після отримання всього масиву мережевого трафіку, було проведено його конвертацію в PDML та реалізовано видобування ознак. В ре-

зультаті було отримано множину векторів ознак, що описували поведінку послідовності пакетів для періоду моніторингу *T*. Зазначені дії повторювались для обох кластерів.

Таблиця 2

Розподіл розумних пристроїв між кластерами

№	Мітка	Розумний пристрій	Кластер	MAC-адреса
1	A	Смарт колонка Amazon echo	1	44:65:0d:62:a4:d7
2	B	Розумна розетка Belkin WeMo Switch	1	ec:1a:59:a3:f1:4b
3	C	Датчик руху Belkin WeMo Motion	1	ec:1a:59:d3:c0:17
4	D	Розумна лампочка Philips Hue Light Bulb	1	00:17:88:28:45:81
5	E	Датчик газу NEST Smoke Sensor	2	18:b4:30:35:f4:c3
6	F	Розумна камера TP-Link Camera	2	f4:f2:6d:97:d8:10
7	G	Розумні ваги Withings Scale	2	00:24:e4:14:47:bd

При проведенні експерименту ми опустили моделювання першого режиму роботи системи (моніторинг мережевого трафіку для виявлення аномальної активності, див. рис.1). Це пов'язано з тим, що даний режим може бути легко реалізований за допомогою системи попередження вторгнень, наприклад Snort [22], шляхом написання правил, які відстежуватимуть появу аномальної активності. Для активації другого режиму був реалізований скрипт на Python, який запускав пошук профілю в кластері (тобто другий режим системи), і у випадку відсутності такого профілю, переводив систему у третій режим – пошук профілю в інших кластерах соціальної мережі.

Перший експеримент передбачав визначення ефективності ідентифікації профілів розумних пристроїв, які присутні у як у білому списку профілів так і поза цим списком локально на одному кластері (тобто перевірка ефективності другого режиму роботи системи в кластері 1).

Для визначення ефективності було використано стандартну міру оцінки класифікатора, що визначає відношення вірно класифікованих об'єктів.

Весь обсяг даних було поділено на дві частини: навчальна (три чверті) та тестова вибірка (одна чверть). Тестова вибірка включала в себе дані мережевого трафіку від всіх чотирьох розумних пристроїв. Навчальна вибірка використовувалась для формування білого списку профілів поведінки, тобто навчання класифікатора. З метою імітації аномальної активності, всього було проведено 4 серії експериментів, в кожній з яких було здійснено навчання класифікатора, на даних, що включали тільки три з чотирьох розумних пристроїв. Таким чином для перевірки ефективності ідентифікації пристроїв було використано тестову вибірку, яка містила профілі всіх розумних пристроїв у кластері. В кожному із експериментів порогове значення ймовірності δ, яка визначає належність до класу розумного пристрою d_i у кластері c_j було обрано експериментальним шляхом на рівні 0,56. Усереднені результати чотирьох серії експериментів наведено на рис 2.

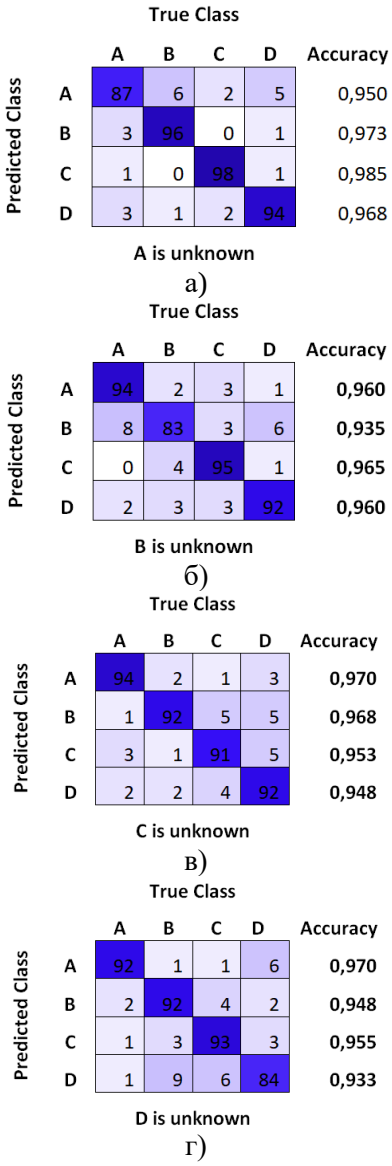

Рис. 2. Матриця помилок для класифікатора для класифікатора у кластері 1, поза білим списком: а)Amazon echo; б) Belkin WeMo Switch; в) Belkin WeMo Motion; г) Philips Hue Light Bulb

Результати проведених експериментів показали досить високі результати, зокрема найвищий показник ефективності (0,985) було отримано в першому експерименті, в білому списку якого, була відсутня розумна колонка Amazon echo. В цьому експерименті 98% сесій мережевого трафіку, що відповідали датчику руху Belkin WeMo Motion, були визначені як такі, що дійсно йому належать (рис. 2а значення true positive для класу C). Найнижчий показник true positive в першому експерименті, як і очікувалось, було для класу A (87%), тобто коли система намагалась передбачити дані Amazon echo (які були вістуні у білому списку). В решті експериментів середня значення ефективності роботи системи склала для другого експерименту 0,955, для третього – 0,959, для четвертого – 0,951. Також слід відзначити, що для проведення експериментів було значення T становило 20 с., а значення t_i – 5 с.

В другому експерименті було перевірено ефективність роботи всієї системи, тобто послідовне виконання другого та третього режимів. Метою експерименту було перевірка чи зможе система ідентифікувати пристрій, якщо цей пристрій відсутній в одному кластері, але є в інших кластерах. З цією метою було задіяно два кластери та створено білі списки профілів пристроїв, що присутні в ньому (здійснено навчання класифікатора на даних для навчання, що складались із розумних пристроїв, які присутні в цьому кластері). Було проведено три серії експериментів, які передбачали ідентифікацію профілів пристроїв NEST Smoke Sensor, TP-Link Camera та Withings Scale (E, F, G відповідно) в кластері 1, білий список якого не містив зазначені розумні пристрої. Результати проведених експериментів представлені у таблиці 3.

Таблиця 3

Ефективність виявлення розумних пристроїв E, F та G в кластері 1

	Розумні пристрої			Середнє
	E	F	G	
Кількість сесій	932	1420	502	
Ефективність, %	98,20	97,36	96,08	**97,21**
Сесії, що були віднесені до одного із класу в кластері 1 (False Positives), %	6,32	5,84	5,67	**5.94**

В даному експерименті рівень False Positives визначав кількість сесій, що належали тестовому розумному пристрою (E, F та G), і які були віднесені до одного із пристроїв промаркованих як A, B, C та D. Згідно із результатами проведеного експерименту загальна ефективність роботи системи склала 97,21% із середнім рівнем похибок першого роду 5,94%.

Висновок

В роботі запропоновано новий підхід в організації безпеки розумних будинків, що передбачає їх об'єднання у кластери. Для реалізації цього підходу представлено систему виявлення аномалій та ідентифікації пристроїв розумних будинків із застосуванням колективної комунікації. Виявлення аномалій та ідентифікації пристроїв в кожному із розумних будинках заснована на моніторингу мережевого трафіку та формуванні профілів розумних пристроїв, які присутні у мережі. Профілі складаються із набору ознак, що описують поведінку розумних пристроїв в мережі, зокрема період активності пристрою та період його сну. На основі цього здійснюється формування білого списку дозволених профілів функціонування пристроїв у кластері. З метою перевірки наявності профілю у білому списку кластера використано алгоритм Random Forest. У випадку відсутності досліджуваного профілю у білому списку кластеру здійснюється запит до інших кластерів, що утворюють соціальну мережу, на предмет зіставлення профілю послідовності пакетів, отриманих у кластері, із власними білими списками. Для оцінки ефективності запропонованої системи було проведено ряд експериментальних досліджень. Результати проведених експериментів показали загальну ефективність роботи системи на рівні 97,21% із середнім рівнем похибок першого роду 5,94%.

Список використаної літератури

1. Asadullah, M. Social Networks of Things for Smart Home s Using Fuzzy Logic [Text] / M. Asadullah, S. Abbas, N. Naz, et al. // International Journal of Computer Science and Network Security. – 2018. – Vol. 18. – No. 2. – P. 168–173.

2. Wheelus, C. IoT Network Security: Threats, Risks, and a Data-Driven Defense Framework [Text] / C. Wheelus, X. Zhu // Cyber Security and Privacy in IoT. – 2020. – P. 259–285. doi: 10.3390/iot1020016

3. Нічепорук, А. О. Метод виявлення DDoS атак на IoT мережі [Текст] / А. О. Нічепорук, А. А. Нічепорук, О. В. Фегир, А. Д. Казанцев, Ю. О. Нічепорук // Вісник Хмельницького національного університету. Серія: Технічні науки. – 2020. – № 1. – С. 156–164.

4. Нічепорук, А. О. Метод виокремлення фрагментів бот-мереж на основі аналізу мережевого трафіку [Текст] / А. О. Нічепорук, А. А. Ні-

чепорук, Ю. О. Нічепорук, А. Д. Казанцев // Вісник Хмельницького національного університету. Серія: Технічні науки. – 2020. – № 2. – С. 141–150.

5. McAfee, Inc., McAfee Labs Threats Report: April 2017, [Електронний ресурс]. – Режим доступу: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-mar-2017.pdf

6. Džaferović, E. DoS and DDoS vulnerability of IoT: A review [Text] / E. Džaferović, A. Sokol, A. A. Almisreb, et al. // Sustainable Engineering and Innovation. – 2019. – P. 43–48. doi: 10.37868/sei.v1i1.36

7. Ali, I. Internet of Things Security, Device Authentication and Access Control: A Review [Text] / I. Ali, S. Sabir, Z. Ullah // International Journal of Computer Science and Information Security. – 2016. – Vol. 14. – P. 456–466.

8. Nicheporuk, A. An android malware detection method based on CNN mixed-data model [Text] / A. Nicheporuk, O. Savenko, A. Nicheporuk, et al. // CEUR Workshop Proceedings. – 2020. – Vol. 2732. – P. 198–213.

9. Савенко, О. С. Формування сигнатури поведінки програми на основі трасування API викликів [Текст] / О. С. Савенко, А. О. Нічепорук, А. А. Нічепорук, Ю. О. Нічепорук // Електротехнічні та комп'ютерні системи. – 2018. – №29(105). – С. 67–77.

10. Spadaccino, P. Intrusion Detection Systems for IoT: opportunities and challenges offered by Edge Comuting [Text] / P. Spadaccino, F. Cuomo // arXiv preprint arXiv:2012.01174. – 2020.

11. Sivanathan, A. IoT Behavioral Monitoring via Network Traffic Analysis. Doctor of Philosophy thesis [Text] / A. Sivanathan. – Sydney, 2020. – 180 p.

12. Hafeez, I. IoT-KEEPER:Detecting malicious IoT network activity using online traffic analysisat the edge [Text] / I. Hafeez, M. Antikainen, A. Y. Ding, et al. // IEEE Transactions on Network and Service Management. – 2020. – P. 45–59. doi: 10.1109/TNSM.2020.2966951

13. Dymora, P. Anomaly detection in iot communication network based on spectral analysis and hurst exponent [Text] / P. Dymora, M. Mazurek // Applied Sciences. 2019. – Vol. 9. – No. 5319. doi: 10.3390/app9245319

14. Summerville, D. H. Ultra-lightweight deep packet anomaly detection for internet of things devices [Text] / D. H. Summerville, K. M. Zach and Y. Chen // Proceedings of the 2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC). – 2018. – P. 1–8. doi: 10.1109/PCCC.2015.7410342

15. Ullah I. A Two-Level Flow-Based Anomalous Activity Detection System for IoT Networks [Text] / I. Ullah, Q. H. Mahmoud // Electronics. – 2020. – Vol. 9. – No. 530. doi: 10.3390/electronics9030530

16. Ramapatruni, S. Anomaly detection models for smart home security [Text] / S. Ramapatruni, S. N. Narayanan, S. Mittal, et al. // Proceedings of 2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS). – 2019. – P. 19–24.

17. Yamauchi, M. Anomaly Detection in Smart Home Operation From User Behaviors and Home Conditions [Text] / M. Yamauchi, Y. Ohsita, M. Murata, et al. // Transaction and consumer electronics. – 2020. – Vol. 66. – P. 183–192. doi: 10.1109/ICCE.2019.8661976.

18. Sivanathan, A. Characterizing and classifying IoT traffic in smart cities and campuses [Text] / D. Sivanathan, H. Sherratt, H. Gharakheili, et al. // Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). – 2017. – P. 559–564. doi: 10.1109/INFCOMW.2017.8116438.

19. Meidan, Y. Detection of Unauthorized IoT Devices Using Machine Learning Techniques [Text] / Y. Meidan, M. Bohadana, A. Shabtai, et al. // arXiv preprint arXiv:1709.04647. – 2017.

20. Anthi, E. A Supervised Intrusion Detection System for Smart Home IoT Devices [Text] / E. Anthi, L. Williams, M. Słowińska, et al. // IEEE Internet of Things Journal . – 2019. – P. 9042–9053. doi: 10.1109/JIOT.2019.2926365

21. Tshark: The Wireshark Network Analyzer. [Електронний ресурс]. – Режим доступу: https://www.wireshark.org/docs/man-pages/tshark.html

22. Snort. [Електронний ресурс]. – Режим доступу: https://www.snort.org/

References

1. Asadullah, M., Abbas, S., Naz, N., et al. (2018) "Social Networks of Things for Smart Home s Using Fuzzy Logic", *International Journal of Computer Science and Network Security*, 168–173.

2. Wheelus, C., Zhu, X. (2020) "IoT Network Security: Threats, Risks, and a Data-Driven Defense Framework", *Cyber Security and Privacy in IoT*, 259–285.

3. Nicheporuk, A. O., Nicheporuk, A. A., Fegur, O. V., Kazantsev, A. D., Nicheporuk, Y. O. (2020) "Method of detecting DDoS attacks on IoT networks" [Metod vyyavlennya DDoS atak na IoT

merezhi] *Scientific Journal Herald of Khmelnytskyi National University*, 1, 156–164 [in Ukrainian].

4. Nicheporuk, A. O., Nicheporuk, A. A., Nicheporuk, Y. O., Kazantsev, A. D. (2020) "Method of detecting fragments of botnets based on the analysis of network traffic" [Metod vyokremlennya frahmentiv bot-merezh na osnovi analizu merezhe-voho trafiku] *Scientific Journal Herald of Khmelnytskyi National University*, 2, 141–150 [in Ukrainian].

5. McAfee, Inc., McAfee Labs Threats Report: April 2017, available at: https://www.mcafee.com /enterprise/en-us/assets/ reports/rp-quarterly-threats-mar-2017.pdf

6. Džaferović, E., Sokol, A., Almisreb, A. A. (2019) "DoS and DDoS vulnerability of IoT: A review" *Sustainable Engineering and Innovation*, 43–48.

7. Ali, I., Sabir, S., Ullah, Z. (2016) "Internet of Things Security, Device Authentication and Access Control: A Review" *International Journal of Computer Science and Information Security*, 14, 456–466.

8. Nicheporuk, A., Savenko O., Nicheporuk, A., Nicheporuk, Y. (2020) "An android malware detection method based on CNN mixed-data model" *CEUR Workshop Proceedings*, 2732, 198–213.

9. Savenko, O. S., Nicheporuk, A. O., Nicheporuk, A. A., Nicheporuk, Y. O. (2018) "Forming of the program's behavioral signature based on the API call tracing" [Formuvannya syhnatury povedinky prohramy na osnovi trasuvannya ARI vyklykiv] *Electrotechnic and Computer Systems*, 29(105), 67-77 [in Ukrainian].

10. Spadaccino, P., Cuomo, F. (2020) "Intrusion Detection Systems for IoT: opportunities and challenges offered by Edge Comuting" *arXiv preprint arXiv:2012.01174*.

11. Sivanathan, A. (2020) "IoT Behavioral Monitoring via Network Traffic Analysis" Doctor of Philosophy thesis. Sydney: The University of New South Wales.

12. Hafeez, I., Antikainen, M., Ding, A. Y., et al. (2020) "IoT-KEEPER: Detecting malicious IoT network activity using online traffic analysis at the edge" *IEEE Transactions on Network and Service Management*, 45–59.

13. Dymora, P., Mazurek, M. (2019) "Anomaly detection in iot communication network based on spectral analysis and hurst exponent" *Applied Sciences*, 9 (5319).

14. Summerville, D. H., Zach, K. M., Chen, Y. (2018) "Ultra-lightweight deep packet anomaly detection for internet of things devices" *Proceedings of the 2015 IEEE 34th International Performance Computing and Communications Conference*, Nanjing, China, 1–8.

15. Ullah I., Mahmoud, Q. H. (2020) "A Two-Level Flow-Based Anomalous Activity Detection System for IoT Networks" *Electronics*, 9 (530).

16. Ramapatruni, S., Narayanan, S. N., Mittal, S. (2019) "Anomaly detection models for smart home security" *Proceedings of 2019 IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, Washington, DC, USA, 19–24.

17. Yamauchi, M., Ohsita, Y., Murata, M. (2020) "Anomaly Detection in Smart Home Operation From User Behaviors and Home Conditions" *Transaction and consumer electronics*, 66, 183-192.

18. Sivanathan, A., Sherratt, H., Gharakheili, H. (2017) "Characterizing and classifying IoT traffic in smart cities and campuses" *Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 559–564.

19. Meidan, Y., Bohadana, M., Shabtai, A., et al. (2017) "Detection of Unauthorized IoT Devices Using Machine Learning Techniques" *arXiv preprint arXiv:1709.04647*.

20. Anthi, E., Williams, L., Słowińska, M. (2019) "A Supervised Intrusion Detection System for Smart Home IoT Devices" *IEEE Internet of Things Journal*, 9042–9053.

21. Tshark: The Wireshark Network Analyzer, available at: https://www.wireshark.org/docs/man-pages/tshark.html

22. Snort, available at: https://www.snort.org/

AN INTELLIGENT SYSTEM FOR DETECTING ANOMALIES AND IDENTIFYING SMART HOME DEVICES BASED ON THE COLLECTIVE COMMUNICATION

A. O. Nicheporuk, A. A. Nicheporuk, O. S. Savenko, A. D. Kazantsev
Khmelnytskyi National University

Abstract. *The fourth industrial revolution put new processes on the rails of automation in industry, healthcare, home and other areas of human life through the mass integration of the concept of the Internet of Things into these areas. At the same time, these processes are not bypassed by home automation systems or*

smart homes. A smart home is defined as a system of interconnected sensors, actuators, and other devices that are networked together with computer systems and controlled by appropriate software. This connection allows to collect, share and analyze data, which helps to increase comfort, automation and control over the parameters of the house. However, along with the obvious advantages and conveniences of rooting home automation systems, this concept leaves a number of potential security bottlenecks for attackers. Data collected by smart devices is always point of interest to hackers and hijackers of confidential information. Third-party access to data collected by smart devices can lead to a variety of emergencies, the degree of danger of which will depend solely on the wish of the owner of the intercepted data.

The paper proposes an intelligent system for detecting anomalies and identifying smart home devices based on the collective communication of smart homes. The concept of the system is based on the benefits of combining smart homes into a social network in terms of improving the security of both a single smart home and the entire social network of combined smart homes. Detection of anomalies and identification of devices in each of the smart homes is based on monitoring network traffic and forming profiles of smart devices that are present in the network. Profiles consist of a set of features that describe the behavior of smart devices on the network, including the period of activity of the device and the period of its sleep. Based on this, a whitelist of allowed profiles of devices operation in the cluster is formed. To verify the presence of a profile in the whitelist the Random Forest algorithm was used. A key feature of the system is the communication of smart home clusters with each other to exchange information about the available smart device profiles in the whitelists of each cluster.

Keywords: *smart home, network traffic, behavior profiling, communication, abnormal behavior, classification.*

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОБНАРУЖЕНИЯ АНОМАЛИЙ И ИДЕНТИФИКАЦИИ УСТРОЙСТВ УМНЫХ ДОМОВ НА ОСНОВЕ КОЛЛЕКТИВНОЙ КОММУНИКАЦИИ

А. О. Ничепорук, А. А. Ничепорук, О. С. Савенко, А. Д. Казанцев
Хмельницкий национальный университет

Аннотация. *В работе предложено интеллектуальную систему обнаружения аномалий и идентификации устройств умных домов с применением коллективной коммуникации. Концепция работы система основана на получении выигрыша от объединения умных домов в социальную сеть в части повышения безопасности как отдельно взятого умного дома, так и всей социальной сети объединенных умных домов. Ключевой особенностью системы является коммуникация кластеров умных домов между собой для обмена информацией об имеющихся профилях умных устройств в белых списках каждого кластера.*

Ключевые слова: *умный дом, сетевой траффик, профилирование поведения, коммуникация, аномальное поведение, классификация*

Отримано 17.03.2021

Нічепорук Андрій Олександрович, Хмельницький національний університет, кандидат технічних наук, доцент, доцент кафедри комп'ютерної інженерії та системного програмування. Вул. Інститутська, 11, Хмельницький, Україна, E-mail: andrey.nicheporuk@gmail.com, тел. +38 096 4687613

Andrii Nicheporuk, Khmelnytskyi National University, PhD, Docent, Associate Professor at the Department of computer engineering and system programming, Institutska str., 11, Khmelnytskyi, Ukraine

ORCID ID: 0000-0002-7230-9475

Нічепорук Анастасія Андріївна, Хмельницький національний університет, аспірант кафедри комп'ютерної інженерії та системного програмування. Вул. Інститутська, 11, Хмельницький, Україна, E-mail: eldess06@gmail.com, тел. +38 098 4812570

Anastasia Nicheporuk, Khmelnytskyi National University, PhD Student at the Department of computer engineering and system programming, Institutska str., 11, Khmelnytskyi, Ukraine

ORCID ID: 0000-0001-5366-5792

Савенко Олег Станіславович, Хмельницький національний університет, доктор технічних наук, професор, декан факультету програмування та комп'ютерних і телекомунікаційних систем Хмельницького національного університету. Вул. Інститутська, 11, Хмельницький, Україна, E-mail: savenko_oleg_st@ukr.net, тел. +38 067 9075315

Oleg Savenko, Khmelnytskyi National University, Dr. of Science, Professor, Dean of the Faculty of Programming and Computer and Telecommunication Systems, Institutska str., 11, Khmelnytskyi, Ukraine

ORCID ID: 0000-0002-4104-745X

Казанцев Андрій Дмитрович, Хмельницький національний університет, аспірант кафедри комп'ютерної інженерії та системного програмування. Вул. Інститутська, 11, Хмельницький, Україна, E-mail: andreykaololo@gmail.com, тел. +38 098 3470288

Andrii Kazantsev, Khmelnytskyi National University, PhD Student at the Department of computer engineering and system programming, Institutska str., 11, Khmelnytskyi, Ukraine

ORCID ID: 0000-0002-2355-5493