ВІДГУК
ОФІЦІЙНОГО ОПОНЕНТА

на дисертаційну роботу Адамова Олександра Семеновича на тему «Моделі і методи захисту кіберпростору на основі аналізу великих даних з використанням машинного навчання», подану на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти

Тема дисертаційної роботи пов'язана з розробкою моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування за рахунок введення обчислювальної надмірності в інфраструктуру кіберпростору. Інструментами виступають аналіз великих даних та машинне навчання, що є підвищує **актуальність роботи**, оскільки за даними компанії Gartner Inc. 2018 року глибоке навчання становитиме найближчим часом 80% стандартних засобів для вчених.

Для досягнення **мети дослідження** – істотне скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування за рахунок введення обчислювальної надмірності в інфраструктуру кіберпростору – автором поставлено та вирішено такі **задачі**: 1) Удосконалити структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів на основі використання дедуктивного аналізу обчислювальних систем. 2) Розробити сигнатурно-кубітні методи синтезу еталонних логічних схем malware-функціональностей і паралельного моделювання malware-driven великих даних для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці. 3) Розробити сигнатурно-кубітну модель активного online cyber security комп'ютингу для моніторингу вхідних потоків malware-даних і управління процесом видалення деструктивних компонентів. 4) Удосконалити засоби захисту кіберпростору шляхом логічного тестування і діагностування атак і шкідливих компонентів на основі використання алгоритмів машинного навчання. 5) Розробити метод атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів і метод перевірки поліморфних шкідливих програм на основі врахування контрольних сум Portable Executable секцій у виконуваних файлах і застосування апарату інтелектуального аналізу даних. 6) Виконати тестування і верифікацію роз-

роблених програмних засобів тестування, перевірки та діагностування шкідливих програм шляхом емуляції атак на основі існуючих malware бібліотек.

Результати дисертації отримано відповідно до планів держбюджетних НДР і господарських договорів, виконуваних на кафедрі АПОТ Харківського національного університету радіоелектроніки в період з 2007 року, у тому числі: 1) Прикладна держбюджетна НДР № 216«Енергозберігаючі інформаційні технології на основі паралельних обчислювальних процесів, безпровідних систем і мереж», 2007-2008, № ДР 0107U001540. 2) Договір про дружбу і співробітництво між ХНУРЕ та компанією «Aldec Inc» (USA) № 04 від 01.11.2011. 3) Фундаментальна держбюджетна НДР № 232«Теорія й проектування енергозберігаючих цифрових обчислювальних систем на кристалах, що моделюють і підсилюють функціональні можливості людини, 2009-2011, № ДР 0109U001646. 4) Фундаментальна держбюджетна НДР № 269«Мультипроцесорна система пошуку, розпізнавання та прийняття рішень для інформаційної комп'ютерної екосистеми», 2011-2013, № ДР 0111U002956. 5) Фундаментальна держбюджетна НДР № 258«Персональний віртуальний кіберкомп'ютер та інфраструктура аналізу кіберпростору», 2012-2014, № ДР 0112U000209. 6) Фундаментальна держбюджетна НДР № 297 "Кіберфізична система –«Розумне хмарне управління транспортом» (Cyber Physical System – Smart Cloud Traffic Control)", 2015-2017, № ДР 0115U-000712 від 04.03.2015. 7) Фундаментальна держбюджетна НДР № 316 "Cyber Physical System – Smart Cyber University", 2017-2019, № ДР 0117U0002524. 8) Проект SEIDA BAITSE "Baltic Academic IT Security Exchange", Blekinge Institute of Technology, Sweden; 2011-2014. 9) Проект 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR «Curricula Development for New Specialization: Master of Engineering in Microsystems Design (MastMEMS)» сумісно з університетом «Львівська політехніка», Київським національним університетом, Технічним університетом м. Лодзь (Польща), Ліонським університетом (Франція), Університетом м. Ільменау (Німеччина), Університетом м. Павія (Італія), 2012-2016. 10) Проект 544455-TEMPUS-1-2013-1-SE-TEMPUS-JPCR «Educating the Next Generation Experts in Cyber Security: the new EU-recognized Master's program (ENGENSEC)», 01 Dec 2013 – 30 Nov 2017.

Характеристика основних розділів роботи.

Вступна частина містить обґрунтування актуальності розв'язуваних задач, визначення мети, об'єкту, предмету і завдань дослідження; наукову новизну і практичну значущість результатів дослідження, відомості про публікації та апробацію отриманих результатів.

Перший розділ присвячено розгляду існуючих моделей, методів і технологій захисту індивідуального сервіс-комп'ютингу. Визначаються переваги і недоліки найбільш затребуваних моделей і методів, опублікованих в матеріалах конференцій і наукових журналах. На основі проведеного аналізу сфор-

мульовано мету і задачі дослідження, що орієнтовані на усунення проблемних місць і недоліків існуючих моделей і методів у контексті їх реалізації в інфраструктурі захисту індивідуального сервіс-комп'ютингу.

У другому розділі розглянуті та вирішені питання, пов'язані з математичним апаратом інфраструктури захисного сервісу, а саме: удосконаленням структурно-логічних моделей і методів перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування шкідливих програм (malware).

У *третьому розділі* запропоновано унітарно кодовані кубітно-матричні моделі, структури даних, обчислювальні архітектури і методи паралельного логічного аналізу деструктивних кодів в кіберфізичному просторі. Введено кубітні векторні структури даних для опису параметрів змінних, що беруть участь у формуванні еталонних зразків (патернів) деструктивних вихідних кодів. Запропоновано паралельний сигнатурно-кубітний метод моделювання malware даних для визначення їх приналежності до існуючих деструктивних компонентів malware library, а також сигнатурно-кубітний метод синтезу еталонних логічних схем malware-функціональностей, який відрізняється від аналогів унітарним кодуванням сигнатур для кодів деструктивних компонентів і формуванням кубітних матриць. Подано сигнатурно-кубітний процесор активного online кіберфізичного cyber security комп'ютингу на основі моніторингу вхідних malware-даних і їх моделювання на еталонних логічних схемах malware-функціональностей для подальшого актюаторного управління процесом видалення деструктивних компонентів.

Четвертий розділ включає опис моделі загроз кіберпростору, а також методів діагностування кібератак на кіберпростір з використанням алгоритмів машинного навчання на основі великих даних: метод атрибутно-орієнтованого впізнання Інтернет посилань з використанням частотних шаблонів; метод детектування поліморфних шкідливих програм; метод пошуку криптопримітивів в троянських програмах-шифрувальниках.

П'ятий розділ містить практичну реалізацію компонентів інфраструктури Cyber Security. Пропонується хмарний сервіс для виявлення кібератак на основі аналізу великих даних, що включає три основних компоненти: 1) Сервер, керуючий потоком вхідних і вихідних даних про кібератаки. 2) Мультисканер з препроцесором для статичного аналізу. 3) Пісочниця (Sandbox), яка призначена для автоматизованого запуску шкідливого коду з метою проведення динамічного аналізу. Для реалізації сервісу створена інфраструктура на основі гипервізора VMWare ESXi.

Практична значущість отриманих результатів полягає у тестування і верифікації розроблених програмних засобів тестування, перевірки та діагностування шкідливих програм шляхом емуляції атак на основі існуючих

malware бібліотек. Результати дисертаційної роботи у вигляді моделей, методів та інфраструктури впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акти про впровадження від 20.05.2019, 21.05.2019); у науково-виробничу діяльність компанії Design & Test Lab (довідка від 18.05.2019), у навчальний процес Blekinge Institute of Technology (BTH), Karlskruna, Sweden (лист 'Statment of Reseach Results Impact on University Education Program' від 29.05.2019).

Висновок

Наукова новизна роботи визначається такими пунктами:

1) Удосконалено структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware.

2) Запропоновано нові методи синтезу еталонних логічних схем malware-функціональностей, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці.

3) Запропонована нова модель активного online cyber security комп'ютингу, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкодію процесів моніторингу вхідних потоків malware-даних і управління видаленням деструктивних компонентів.

4) Запропоновано нові методи виявлення кіберзагроз, які здатні навчатися на великих даних, з метою виявлення кібератак, що можуть бути реалізовані у вигляді Інтернет посилань, поліморфних шкідливих програм (polymorphic malware) та троянських програм-шифрувальників, що займаються здирництвом, а саме:

метод атрибутно-орієнтованого впізнання Інтернет посилань з використанням частотних шаблонів, що може провести оцінку атрибутів та відрізнити доброякісні, фішинг та шкідливі посилання;

метод детектування поліморфних шкідливих програм, що дозволяє виявляти поліморфні шкідливі програми за допомогою аналізу хешів PE секцій та пошуку схожих секцій у бібліотеці шкідливого коду (після підтвердження детектування PE файла нові хеші секцій додаються до бібліотеки);

метод пошуку криптопримітивів у троянських програмах-шифрувальниках, що дозволяє виявити алгоритми шифрування, використовувані здирником, та дає можливість дешифрувати зашифровані файли користувача чи організації.

5) Удосконалено засоби захисту кіберпростору, які відрізняються використанням моделей і методів сигнатурно-логічного тестування атак, пошуку криптопримітивів у троянських програмах-шифрувальниках на основі використання алгоритмів машинного навчання, що дає можливість істотно скоротити час відновлення працездатності обчислювальної структури.

Практичне значення отриманих результатів полягає у розробці моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, а саме:

– тестуванні, верифікації і впровадженні розроблених програмних засобів перевірки, діагностування шкідливих програм і атак, що дає можливість виконувати їх моделювання із залученням існуючих додатків і malware бібліотек;

– програмній реалізації методу атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів, який відрізняється застосуванням апарату інтелектуального аналізу даних, що дає можливість визначати вірогідну оцінку небезпеки URL-адреси на основі атрибутів;

– програмній реалізації методу перевірки поліморфних шкідливих програм, який відрізняється інваріантністю до детермінізму сигнатур в коді і урахуванням тільки контрольних сум Portable Executable (PE) секцій у виконуваних файлах, що дає можливість поліпшити продуктивність процедур діагностування деструктивних компонентів.

Окремі моделі, методи та елементи інфраструктури реалізовані у вигляді програмних додатків і пройшли апробацію у навчальному процесі Харківського національного університету радіоелектроніки, университету Блекінге (Швеція), а також методологічному та технологічному забезпеченні компанії«Design & Test Lab».

Обґрунтованість теоретичних положень та наукових результатів підтверджується виконаними експериментами, тестуванням і верифікацією розроблених моделей і методів, точністю детектування і класифікацією прикладів загроз нульового дня, серед яких нові версії кріптолокерів і складних загроз.

Порівняльний аналіз змісту дисертаційної роботи та опублікованих робіт здобувача показав, що результати наукових досліджень, а саме: удосконалені структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів; нові методи синтезу еталонних логічних схем malware-функціональностей; нова модель активного online cyber security комп'ютингу; нові методи виявлення кіберзагроз, які здатні навчатися на великих даних (метод атрибутно-орієнтованого впізнання Інтернет посилань з використанням частотних шаблонів; метод детектування поліморфних шкідливих програм; метод пошуку криптопримітивів у троянських програмах-шифрувальниках); а також удосконалені засоби захисту кі-

берпростору відображені у 31 друкованій праці. Серед них: 3 розділи у закордонних монографіях (з них 1 входить до наукометричної бази Scopus), 7 статей (з них 5 – у наукових журналах, включених до«Переліку наукових фахових видань України»; 2 статті в міжнародних наукових журналах за кордоном; 4 статті входять до міжнародних наукометричних баз), а також у 21 міжнародній науковій конференції (з них 13 за кордоном, 12 входять до наукометричної бази Scopus). Здобувач має 13 публікацій у наукометричній базі Scopus та індекс Хірша h=3.

Автореферат відображає зміст дисертаційної роботи.

Зауваження по дисертаційній роботі:

1) Автор впроваджує поняття активного online cyber security комп'ютингу на основі сигнатурного-кубітного подання даних. При цьому не дається пояснення, чим це відрізняється від класичного обчислювального процесу.

2) Автор пропонує вдосконалені засоби захисту кіберпростору на основі використання методів сигнатурно-логічного тестування атак і алгоритмів машинного навчання. При цьому не наводяться розрахунки за часом відновлення працездатності обчислювальних систем, а також витрати на створення систем захисту.

3) Стилістично перший розділ побудований на основі тезового перерахування існуючих технологій. Бажано було б навести більш розгорнуті характеристики щодо аналізованих моделей і методів.

4) Потребують суттєвого доопрацювання наведені автором на стор. 50 визначення штучного інтелекту, машинного навчання і data science.

5) Не зовсім зрозуміло, як метрики оцінювання процесів і явищ, наведені в розділі 2, використовуються для захисту кіберпростору.

6) У розділі 3 вводяться кубітні структури даних і алгоритми їх аналізу. Потребує пояснення, що дає кубічна технологія для підвищення швидкодії обробки великих даних.

7) Автор в розділі 5 акцентує увагу на різноманітті розроблених архітектур і призводить значну кількість графіків ефективності впроваджуваних методів, але при цьому відсутні аналітичні формули та оцінки, за якими будуються графіки.

Враховуючи викладене вище, можна зробити такий *висновок*: в дисертаційній роботі вирішено важливу науково-практичну задачу введення в інфраструктуру комп'ютингового простору програмної надмірності у формі моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування.

Отримано такі важливі **наукові результати**:

1) Удосконалені структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware.

2) Нові методи синтезу еталонних логічних схем malware-функціональностей, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці.

3) Нова модель активного online cyber security комп'ютингу, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкодію процесів моніторингу вхідних потоків malware-даних і управління видаленням деструктивних компонентів.

4) Нові методи виявлення кіберзагроз, які здатні навчатися на великих даних, з метою виявлення кібератак, що можуть бути реалізовані у вигляді Інтернет посилань, поліморфних шкідливих програм (polymorphic malware) та троянських програм-шифрувальників, що займаються здирництвом: метод атрибутно-орієнтованого впізнання Інтернет посилань з використанням частотних шаблонів, що може провести оцінку атрибутів та відрізнити доброякісні, фішинг та шкідливі посилання; метод детектування поліморфних шкідливих програм, що дозволяє виявляти поліморфні шкідливі програми за допомогою аналізу хешів PE секцій та пошуку схожих секцій у бібліотеці шкідливого коду (після підтвердження детектування PE файла нові хеші секцій додаються до бібліотеки); метод пошуку криптопримітивів у троянських програмах-шифрувальниках, що дозволяє виявити алгоритми шифрування, використовувані здирником, та дає можливість дешифрувати зашифровані файли користувача чи організації.

5) Удосконалені засоби захисту кіберпростору, які відрізняються використанням моделей і методів сигнатурно-логічного тестування атак, пошуку криптопримітивів у троянських програмах-шифрувальниках на основі використання алгоритмів машинного навчання, що дає можливість істотно скоротити час відновлення працездатності обчислювальної структури.

6) Практично реалізовано окремі моделі, методи та елементи інфраструктури у вигляді програмних додатків, що пройшли апробацію та впроваджені у навчальний процес впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акти від 20.05.2019, 21.05.2019), Blekinge Institute of Technology (BTH), Karlskruna, Sweden (лист від 29.05.2019), а також у науково-виробничу діяльність компанії Design & Test Lab (довідка від 18.05.2019).

Дисертаційна робота відповідає спеціальності 05.13.05 – комп'ютерні системи та компоненти, задовольняє вимогам пунктів 9, 11-14 "Порядку при-

судження наукових ступенів", затвердженого постановою Кабінету Міністрів України від 19 серпня 2015 № 656, а також вимогам Департаменту атестації кадрів МОН України до дисертацій на здобуття наукового ступеня кандидата технічних наук, а дисертант Адамов Олександр Семенович заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти.

Офіційний опонент:
начальник відділу математичного
моделювання та дослідження
ядерно-фізичних процесів і систем
Національного наукового центру
"Харківський фізико-технічний інститут"
доктор технічних наук, професор Хажмурадов М. А.

Підпис проф. Хажмурадова М.А. засвідчую:
Вчений секретар _____

ЗАСВІДЧУЮ
Учений секретар
ННЦ ХФТІ
05 « 09 2019 р.

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ Національний науковий центр "Харківський фізико-технічний інститут" №14312223