ВІДГУК

ОФІЦІЙНОГО ОПОНЕНТА

на дисертаційну роботу Адамова Олександра Семеновича

«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

Актуальність роботи.

Розробка моделей і методів захисту кіберпростору є практично затребуваною задачею. Сьогодення характеризується зростанням кількості кібератак на державні і корпоративні структури, тобто комп'ютерні мережі, сервери та обчислювальні сервіси. Надійність цифрової системи з точки зору кібербезпеки визначається захищеністю найбільш уразливого її компонента: користувача, програмного або апаратного сервісу. Тут, зокрема, актуальними виявляються питання кіберрозвідки і запуску атаки на обчислювальну систему за допомогою експлуатації виявленої технічної вразливості або за допомогою методів соціальної інженерії. Вразливість цифрового сервісу, системи або людини передбачає несправність, використовуючи яку можна навмисно порушити цілісність, доступність і конфіденційність інформації, з якою оперують зазначені суб'єкти.

У дисертаційній роботі досліджуються питання, пов'язані з мінімізацією проміжку часу між моментом запуску атаки на кіберпростір і моментом її діагностування, протягом якого обчислювальний сервіс залишається скомпрометованим, що одночасно дозволяє поліпшити якість сервісу шляхом забезпечення доступності, цілісності і конфіденціійності оброблюваної інформації на період атаки; мінімізацією витрат на відновлення працездатності сервісу і фінансових втрат від його простою за рахунок введення мінімально необхідної надмірності в інфраструктуру діагностування.

Метою дослідження є істотне зменшення 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.

Для досягнення поставленої мети автором вирішено такі задачі, пов'язані з **розробкою**: *сигнатурно-кубітних методів* 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; *сигнатурно-кубітної моделі* активного online cyber security комп'ютингу для моніторингу вхідних потоків malware-даних і управління процесом видалення деструктивних компонентів; *методу 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;

удосконаленням *структурно-логічних моделей і методів* перевірки кіберпростору для тестування і діагностування шкідливих компонентів на основі використання дедуктивного аналізу обчислювальних систем; *засобів захисту кіберпростору* шляхом логічного тестування і діагностування атак і шкідливих компонентів на основі використання алгоритмів машинного навчання;

а також **здійсненням практичної реалізації** окремих сервісів – верифікацією розроблених програмних засобів тестування / перевірки / діагностування шкідливих програм шляхом емуляції атак на основі існуючих malware бібліотек.

В роботі зазначено, що розробка основних положень дисертації здійснювалась відповідно до планів держбюджетних науково-дослідних робіт і міжнародних договорів, виконуваних на кафедрі Автоматизації проектування обчислювальної техніки ХНУРЕ 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 роботи.

Дисертаційна робота складається зі вступу, 5 розділів, 37 підрозділів, висновків, списку використаних джерел зі 160 назв, 4 додатків.

У **вступній частині** описано мотивацію виконання дослідження, актуальність науково-практичної задачі, що розв'язується; сформульовано мету, об'єкт і задачі дослідження; сукупність наукових результатів, що виносяться на захист; визначено наукову новизну та практичну значущість отриманих результатів; наведено відомості про їх апробацію та реалізацію, характеристику публікацій.

У **першому розділі** дисертації викладено аналіз існуючих публікацій в області створення моделей, методів і технологій захисту індивідуального сервіс-комп'ютингу. Визначаються переваги і недоліки найбільш затребуваних моделей і методів, опублікованих в спеціальній літературі: матеріалах конференцій і наукових журналах. На основі проведеного аналізу сформульовано мету і задачі дослідження, орієнтовані на усунення проблемних місць і недоліків існуючих моделей і методів у частині їх реалізації в інфраструктурі захисту індивідуального сервіс-комп'ютингу.

У **другому розділі** наведено 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. Визначено компоненти блокчейн технології, які використовуються для створення надійної інфраструктури захисту даних, складеної з ненадійних елементів. Інфраструктура захисних сервісів створюється разом з кіберсистемою (КС) і супроводжує останню протягом всього життєвого циклу, обслуговуючи всі наступні модифікації КС, і сама постійно підвищує свій інтелект шляхом поповнення історії та бібліотек конструктивних і деструктивних компонентів. Функція мети представлена підвищенням ефективності сервісного обслуговування на основі стандартів тестування, граничного сканування і спеціальних технологій діагностування та відновлення невразливості КС, яка визначається мінімальним значенням рівня вразливості, часу відновлення працездатності і нефункціональної програмно-апаратної

надмірності. Запропоновано вдосконалені методи синтезу тестів для функціональностей, заданих матричними формами опису поведінки компонентів КС, які відрізняються паралелізмом векторних операцій над таблицями, що дає можливість істотно (x2) підвищити швидкодію обчислювальних процедур. Процес-моделі і методи синтезу тестів для функціональностей і діагностування ФН можуть бути використані як вбудовані компоненти інфраструктури сервісного обслуговування КС із застосуванням стандартів тестопридатності.

Третій розділ присвячено розробці нових методів 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. Пропонуються унітарна кодовані кубітно-матричні моделі, структури даних, обчислювальні архітектури і методи паралельного логічного аналізу деструктивних кодів в кіберфізичному просторі. Вводяться кубітні векторні структури даних для опису параметрів змінних, що беруть участь у формуванні еталонних зразків (патернів) деструктивних вихідних кодів. Пропонується паралельний сигнатурно-кубітний метод моделювання malware даних для визначення їх приналежності до існуючих деструктивних компонентів malware library. Пропонується сигнатурно-кубітний метод синтезу еталонних логічних схем malware-функціональностей, який відрізняється від аналогів унітарним кодуванням сигнатур для кодів деструктивних компонентів і формуванням кубітних матриць. Вводиться сигнатурно-кубітний процесор активного online кіберфізичного cyber security комп'ютингу (CSC) на основі моніторингу вхідних malware-даних і їх моделювання на еталонних логічних схемах malware-функціональностей з метою подальшого актюаторного управління процесом видалення деструктивних компонентів.

У **четвертому розділі** запропоновано модель загроз кіберпростору, а також методи діагностування кібератак на кіберпростір з використанням

алгоритмів машинного навчання на основі великих даних, що дозволяють виявити загрози, запропоновані у моделі загроз кіберпростору. Пропонуються 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: 1) Метод атрибутно-орієнтованого впізнання інтернет посилань з використанням частотних шаблонів, що може провести оцінку атрибутів та відрізнити доброякісні, фішинг та шкідливі посилання. 2) 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. 3) Метод пошуку криптопримітивів у троянських програмах-шифрувальниках. Завдяки цьому методу можливо відповісти, які алгоритми шифрування використовував здирник та чи можливо дешифрувати зашифровані файли користувача або організації.

У **п'ятому розділі** виконана практична реалізація компонентів інфраструктури Cyber Security. Запропоновано хмарний сервіс для виявлення кібератак на основі аналізу великих даних, який 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.

Таким чином, дисертаційне дослідження містить компоненти у вигляді моделей, методів та інфраструктури: 1) *Нові* методи синтезу еталонних логічних схем malware-функціональностей. 2) *Нова* модель активного online cyber security комп'ютингу. 3) *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. 4) *Удосконалені* структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів. 5) *Удосконалені* засоби захисту кіберпростору на основі використання алгоритмів машинного навчання.

Практична значущість результатів досліджень полягає у 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 компоненти інфраструктури впроваджені у навчальний процес Харківського національного університету радіоелектроніки; у науково-виробничу діяльність компанії Design & Test Lab, у навчальний процес Blekinge Institute of Technology (BTH), Karlskruna, Sweden, про що свідчать відповідні довідки та акти.

На основі викладеного вище можна зробити такі висновки.

1. **Наукову повизну** роботи визначають:

– 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;

– вперше запропоновані 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;

– вперше розроблена 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;

– вперше запропонований 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;

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

2. **Практичне значення** отриманих результатів полягає у:

– 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 його 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.

3. Отримані наукові висновки та положення дисертації є обґрунтованими і достовірними. Достовірність наукових висновків підтверджується проведеними експериментами, тестуванням і верифікацією, точністю детектування і класифікацією прикладів загроз нульового дня, серед яких нові версії кріптолокерів і складних загроз (Advanced Persistent Threats – APT). Результати дисертації у складі 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).

4. Автором опубліковано 31 друкована праця: 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 відповідає змісту дисертаційної роботи та містить опис основних наукових і практичних результатів, отриманих автором.

Зауваження по дисертаційній роботі Адамова Олександра Семеновича:

1) У розділі 2 виконано дуже ретельний аналіз щодо блокчейн комп'ютингу. Це можна віднести як до переваг, так і до недоліків.

2) З метою аналізу суттєвості деструктивності (уразливості і проникнення) для стану кіберсистеми автор пропонує математичний апарат інфраструктури захищеного сервісу, що містить метрику, алгебру, структури даних і моделі оцінювання якості взаємодії компонентів у кіберпросторі, необхідні при створенні ефективних двигунів для обчислювальних процедур

аналізу даних в процесах тестування, а також апарат булевих похідних для синтезу тестів та дедуктивний метод для пошуку вразливостей в кіберсистемі. Не достатньо розкрито, чим вони відрізняються від відомих методів технічної діагностики комп'ютерних систем?

3) Необхідне пояснення, як використовується метод дедуктивного паралельного аналізу для перевірки та діагностики шкідливих програм.

4) Автор не пояснює, як було реалізовано верифікацію методу перевірки поліморфних шкідливих програм, з урахуванням тільки контрольних сум Portable Executable в виконуваних файлах для підвищення продуктивності процедур діагностування деструктивних компонентів.

5) На стр. 154 подано таблицю сигнатурного стиснення деструктивних компонентів. Неясно, чим відрізняється сигнатурний аналіз від хешированих даних і чи можна тут застосувати хеш-функції.

6) Науково-практична задача формулюється, виходячи із задоволення потреб ринку, пов'язаних з часом, витратами і якістю. У той же час автор формулює її як: «...введення в інфраструктуру комп'ютингового простору програмної надмірності ...» (і далі по тексту). Даний пункт потребує зміни акцентів.

7) Неясно, що дає апарат кубічного числення (розділ 2) для синтезу логічних функцій моделювання деструктивних компонентів системи.

8) Сутність розділу 4 дуже стисло наведена у авторефераті, проте сам розділ у дисертації містить аналіз численних прикладів, що підкріплені таблицями та рисунками.

9) Подекуди у тексті зустрічаються терміни «кібернетична система» та «кіберфізична система» з однаковою абревіатурою. Не зрозуміло, що мав на увазі автор.

Незважаючи на зазначені зауваження, дисертаційна робота є завершеним науковим дослідженням, в якому поставлено за мету 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.

Роботу виконано на високому теоретичному рівні з використанням математичних методів і сучасних засобів обчислювальної техніки. Дисертаційна робота відповідає спеціальності 05.13.05 – комп'ютерні системи та компоненти. Актуальність вибраної теми, достовірність і обґрунтованість висновків, новизна досліджень, значення отриманих результатів для науки і практики свідчать про те, що дисертаційна робота «Моделі і методи захисту кіберпростору на основі аналізу великих даних з використанням машинного навчання», задовольняє вимогам пп. 9, 11-14 "Порядку присудження наукових ступенів", затвердженого постановою Кабінету Міністрів України від 19 серпня 2015 № 656, а здобувач Адамов Олександр Семенович заслуговує на присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти.

Офіційний опонент:
професор кафедри спеціалізованих
комп'ютерних систем
Українського державного університету
залізничного транспорту
доктор технічних наук, професор Мірошник М.А.

Особистий підпис
засвідчую _____ 20__ р.
Завідуючий канцелярією
УкрДУЗТ

Мірошник М.А.
Олександровська