ВІДГУК

офіційного опонента про дисертаційну роботу Мартовицького Віталія Олександровича " Моделі та метод виявлення аномалій функціонування комп'ютерних систем на основі технології машинного навчання", подану на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп’ютерні системи та компоненти

1. Актуальність теми дисертаційної роботи та її зв'язок з науковими програмами, планами, темами

В останні роки відбувається значне збільшення обсягів інформації, що накопичується, зберігається та оброблюється за допомогою розподілених комп’ютерних систем (РКС). При цьому концентрування в єдиних базах даних інформації різного призначення та різної належності і різке розширення кола користувачів, що мають безпосередній доступ до ресурсів РКС, породжують проблему забезпечення їх захисту від різного роду вторгнень. Зростання складності апаратно-програмних засобів та існуючі недоліки сучасних інформаційних технологій призводять до постійного збільшення методів зламу захисту і, як наслідок, до вторгнення в комп’ютерну систему з метою порушення її нормального функціонування. Для запобігання цьому створюються системи виявлення вторгнень, які є невід’ємною частиною будь-якої сучасної системи безпеки, а світова тенденція свідчить про те, що виявлення вторгнень, стане обов’язковою функцією операційної системи та вже застосовується в різному програмному забезпеченні. Один з підходів щодо виявлення аномального стану РКС базується на використанні неперервного моніторингу, який передбачає збирання метрик системи, візуалізацію даних і попередження операторів про кібернетичні впливи на комп’ютерну систему. Тому вибір об’єкту дисертаційного дослідження – процесу моніторингу стану інформаційного та комунікаційного середовища розподілених комп’ютерних систем є актуальним та доцільним.

Варто відзначити, що процес моніторингу має забезпечити можливість блокування різних видів кібератак за максимально короткий час. Тому вибір предмету дисертаційного дослідження – методи і алгоритми моніторингу в розподілених комп’ютерних системах із застосуванням технологій інтелектуального аналізу даних, та визначення мети дослідження – покращення показників виявлення аномалій функціонування РКС в умовах кібернетичних впливів зовнішнього та внутрішнього середовищ шляхом побудови моделей і методів на основі технологій інтелектуального аналізу даних, є обґрунтованими та відповідають темі дисертаційної роботи.

Тема досліджень та одержані результати безпосередньо пов’язані з науково-дослідними роботами, виконаними у Харківському національному університеті радіоелектроніки: «Створення науково-методичних основ забезпечення живучості мережевих систем обміну інформацією в умовах зовнішнього впливу потужного НВЧ випромінювання» (№ ДР 0117U003916), «Методи, системи та засоби криптографічного захисту інформації з гарантованим рівнем стійкості та підвищеною швидкодією» (№ДР 0115U002431), «Створення науково-методичних основ забезпечення живучості мережевих систем обміну інформацією в умовах зовнішнього впливу потужного НВЧ випромінювання» (№ ДР 0118U000832).

Таким чином, усе сказане обумовлює актуальність теми дисертаційної роботи Мартовицького В.О. і наукову новизну поставлених в ній задач досліджень.

2. Наукова новизна результатів роботи

У результаті виконання дисертаційної роботи набув подальшого розвитку науковий напрям, пов’язаний із розробленням систем виявлення аномалій функціонування РКС.

Виходячи з того, що нові наукові результати - це нові знання в певній галузі фундаментальних чи прикладних наук, можна вважати основними науковими результатами дисертації таке:

- вперше запропоновану модель класифікації стану системи, яка ґрунтується на структурному представленні показників функціональності розподілених комп’ютерних систем, що дозволяє виділити множину станів залежно від функціональних завдань, розмежувати процеси цільового функціонування системи та інтерфейсні процеси взаємодії з мережевою інфраструктурою та використовувати їх у методах інтелектуального аналізу для виявлення аномалій функціонування розподілених комп’ютерних систем;

- удосконалену мультиагентну модель системи збирання і зберігання інформації, що побудована на основі агентів, метою яких є надання користувачеві або інформаційній системі більш високого рівня інформації про стан мережевої інфраструктури, отриманої в результаті збирання та інтелектуального оброблення параметрів, що дозволило зменшити навантаження на мережу за рахунок застосування запропонованого протоколу обміну інформацією між агентами;

- удосконалений метод класифікації стану мережі на основі статистичних параметрів за рахунок рівномірної вибірки об'єктів із поверненням для формування навчальних вибірок, що дозволяє адаптувати процес навчання ансамблю класифікаторів до розмірів навчальної вибірки.

3. Ступінь обґрунтованості наукових положень дисертації та їх достовірність

Наукові положення, викладені в дисертаційній роботі, є достатньо обґрунтованими за рахунок використання апробованих математичних методів та елементів теорій.

Достовірність основних наукових результатів роботи підтверджується наведеною в розд. 2-4 системою формальних методик і перетворень, що не містить принципових помилок, а також рядом прикладів і збіжністю результатів експериментальних досліджень, отриманих під час реалізації відповідних засобів з теоретичними і практичними результатами.

4. Цінність дисертаційної роботи для науки

Цінність дисертації полягає в тому, що в ній запропоновано нове рішення важливої науково-технічної задачі в теорії побудови та використання систем виявлення кібернетичних впливів на комп'ютерні мережі. Змістовний аспект запропонованого рішення, який спрямований на розширення класу моделей і методів ідентифікації аномальних станів, що забезпечують розширення функціональних можливостей сучасних систем виявлення вторгнень, не був відомий раніше.

5. Практична корисність роботи

Практична корисність роботи обумовлена тім, що використання запропонованих в ній моделей, формальних методів, конкретних рішень і рекомендацій дозволяє створювати більш досконалі, порівняно з відомими, програмні засоби виявлення аномальних станів розподілених комп'ютерних систем. Запропонована система моніторингу може працювати паралельно з уже розгорнутими засобами моніторингу, заміщаючи їх на деяких рівнях, що дозволяє змінювати і розширювати набір доступних функцій цих систем.

Основні результати дисертаційних досліджень використано:

– при розробці перспективних систем передачі даних в Центральному конструкторському бюро «ПРОТОН»;

– у Харківському національному університеті радіоелектроніки на кафедрі електронних обчислювальних машин в процесі проведення лабораторних робіт з дисципліни «Технології виявлення загроз в комп'ютерних мережах».

6. Оцінка змісту дисертації, її завершеності й оформлення

Побудова дисертації відповідає прийнятим для наукового дослідження вимогам. Дисертація складається з анотації, списку скорочень, вступу,

чотирьох розділів, загальних висновків, списку використаних джерел і додатків.

У вступі обґрунтовано актуальність теми дисертаційної роботи, показано зв'язок роботи з науковими темами, сформульовано мету та задачі досліджень, викладено наукову новизну та практичне значення одержаних результатів, зазначено особистий внесок здобувача та наведено відомості про впровадження, апробації, структуру роботи.

У першому розділі проведено аналіз сучасного стану підходів щодо виявлення аномалій в комп'ютерних системах. Розглянуто особливості побудови сучасних систем моніторингу стану розподілених комп'ютерних систем. Проаналізовано методи виявлення атак, які реалізовані в сучасних системах виявлення атак і недостатньо опрацьовані в частині формальної моделі атаки.

Другий розділ присвячено розробці моделі моніторингу аномалій. Досліджено архітектуру та модель функціонування розподілених комп'ютерних систем. Визначено параметри моніторингу інфраструктури і додатків у розподілених комп'ютерних системах та наведено структуру моделі моніторингу для виявлення аномальних подій, що дозволяє забезпечити виявлення аномалій функціонування РКС в умовах кібернетичних впливів зовнішнього та внутрішнього середовищ. Розроблена модель дозволяє здійснити моніторинг РКС не тільки як єдиної обчислювальної системи, а й усіх її компонентів окремо, що дає можливість всебічно оцінити стан системи в цілому. При цьому автор пропонує модель системи збирання і зберігання даних, що забезпечує роботу з множиною різнорідних джерел, шляхом їх інтегрування з метою отримання більш повної інформації. Система, заснована на мультиагентному підході, дозволяє не припиняти обробку запитів при виконанні модифікацій набору і структур, які використовуються в базах даних.

У третьому розділі розроблюється метод класифікації аномалій стану мережі на основі статистичних параметрів. Наведено результати порівняльного аналізу запропонованого та відомих алгоритмів класифікації. Для отримання цих результатів використовувалися дані чемпіонату з машинного навчання KDD 2009 і дані отримані під час моніторингу мережевої інфраструктури навчального дата-центра, розгорнутого на основі мережевої файлової системи Lustre.

Збільшення кількості інформації, що обробляється комп'ютерними системами, а також економія на кількості обслуговуючого персоналу потребують використання ефективних засобів моніторингу інформаційних ресурсів. Результатом цього є зростання кількості параметрів, які повинна

відстежувати така система моніторингу. За рахунок великих потоків даних від різних датчиків зростає ймовірність пропуску адміністратором системи негативних змін контрольованих параметрів комп'ютерної системи. Для запобігання цьому в системи моніторингу починають впроваджувати засоби автоматизованого експертного аналізу даних, заснованого на машинному навчанні.

У розділі проведено оцінювання інформативності параметрів контролю мережевої інфраструктури з метою забезпечення більш ефективного інтелектуального аналізу. З використанням ознак лише максимальної важливості побудовано алгоритм, який забезпечує майже таку середню помилку за метрикою MSE як алгоритм XGBClassifier, але з більшою швидкістю роботи.

У четвертому розділі для оцінювання стану РКС за допомогою розробленої мультиагентної системи моніторингу була запропонована методика моніторингу.

Дана методика визначає умови і порядок оцінювання стану РКС на основі аналізу відомостей про стан, що не відповідають нормальному функціонуванню РКС внаслідок кібернетичних впливів.

Розроблено архітектуру системи моніторингу з використанням автономних програмних агентів. Архітектура передбачає динамічне формування ієрархічної структури, вузлом якої може виступати будь-яка сутність, що визначається джерелом даних або сенсором. Таким чином, стосовно моніторингу РКС можуть існувати метрики грід, кластерів, обчислювальних вузлів і завдань.

У висновках стисло сформульовано основні наукові та практичні результати дисертаційної роботи.

У додатках містяться акти впровадження результатів дисертаційної роботи.

Таким чином, усі положення, винесені на захист, висвітлені в тексті дисертації. Зміст дисертаційної роботи відповідає її назві. Дисертація написана науковою мовою та оформлена відповідно до існуючих нормативних документів.

7. Рекомендації щодо використання результатів дисертації

Запропоновані в роботі моделі, методи та алгоритми можуть бути використані для побудови високоефективної системи моніторингу стану інформаційного та комунікаційного середовища розподілених комп'ютерних систем, яка є складовою комплексної системи захисту інформаційних ресурсів від кібератак.

8. Повнота викладення основних результатів дисертації

Основні результати дисертаційної роботи достатньо повно відображені в 13 наукових працях, з них: 6 статей у наукових фахових виданнях України, включених до міжнародних наукометричних баз (з них 2 внесені до міжнародної наукометричної бази даних SCOPUS), 7 публікацій у матеріалах міжнародних наукових конференцій (з них 1 внесена до міжнародної наукометричної бази даних SCOPUS). Для спільних наукових статей автором зазначено його особистий внесок.

9. Автореферат дисертації

Автореферат дисертації за своїм змістом повністю відповідає дисертаційній роботі.

10. Зауваження щодо змісту і оформлення дисертації

1. На теперішній час найуживанішими є поняття «комп'ютер» та похідні від нього поняття «комп'ютерні засоби» і «комп'ютерні системи». Однак у роботі поруч з цими поняттями вживаються застарілі поняття «ЕОМ» і «обчислювальні засоби».

2. Некоректно вживаються як синоніми такі поняття: «система безпеки інформації», «система інформаційної безпеки», «система захисту».

3. Автор наводить перелік методів захисту (захисних механізмів), що використовуються для захисту комп'ютерних систем від несанкціонованого втручання в процеси їх функціонування і несанкціонованого доступу до інформації, а далі відзначає, що «… перераховані механізми захисту можуть застосовуватися в конкретних технічних засобах та системах захисту в різних комбінаціях». При цьому як один із захисних механізмів вказано «страхування ризиків». Однак цей механізм не може застосовуватися в конкретних технічних засобах.

4. Автор невдало вводить скорочення МАС (мультиагентна система), оскільки далі користується загальноприйнятим у криптографії скороченням MAC (media access control).

5. Розглядаючи в підрозділі 1.3.1 критерії порівняння методів виявлення атак, автор відзначає, що «… в огляді не розглянуто такі важливі критерії як повнота і точність методу, тому що ці характеристики рідко наводяться в публікаціях». Це не є переконливим для того, щоб не розглядати важливі критерії або ці критерії не є важливими.

6. У роботі відсутнє обґрунтування вибору класифікаторів (kNN, наївний класифікатор Байеса, дерева класифікації, SVM), використаних для проведення експерименту.

7. Постановка задач досліджень була б більш обґрунтованою, якби наочно (у вигляді діаграм та графіків) був обґрунтований вибір методу виявлення аномалій з використанням інтелектуального аналізу. Це також сприяло б кращому висвітленню зроблених дисертантом висновків щодо обґрунтованості можливих напрямків розвитку відповідного методу класифікації стану функціонування комп'ютерних систем.

8. Частину тексту підрозділів 2.1 і 3.2 подано як огляд відомих архітектур розподілених комп'ютерних систем і методів класифікації стану, тому доцільно було б цей огляд навести в розділі 1.

9. Автор часто використовує якісні значення замість кількісних. Наприклад, велика кількість параметрів стану компонентів мережі, навчальна вибірка невеликого розміру, велика складність алгоритму, метрика зберігає лише набір даних за відносно невеликий проміжок часу і організує ефективний доступ до них.

10. Другий науковий результаті дисертації полягає в тому, що «…удосконалено мультиагентну модель системи збирання і зберігання інформації, що побудована на основі агентів, метою яких є надання користувачеві або інформаційній системі більш високого рівня інформації про стан мережевої інфраструктури, отриманої в результаті збирання та інтелектуального оброблення параметрів». Автор стверджує, що це дозволило зменшити навантаження на мережу за рахунок застосування запропонованого протоколу обміну інформацією між агентами (підрозділ 2.3.3). Проте, кількісної оцінки відповідних результатів у дисертації не наведено.

11. При розробці методу класифікації стану мережі на основі статистичних параметрів для виявлення аномалій в інформаційній структурі розподіленої комп'ютерної системи автором використовується ансамбль класифікаторів для визначення стану мережі розподілених комп'ютерних систем. Але відсутня оцінка обчислювальної складності реалізації даного методу і тому складно перевірити можливість моніторингу стану мережі в реальному часі.

12. У табл. 3.7 і 3.8 стовпчик з назвою «Кількість випробувань» повинен мати назву «Класифікатори». Крім того, не вказано, в яких одиницях подано «Кількість вірних рішень», «Кількість невірних рішень», «Помилки I роду» і «Помилки II роду». Це абсолютні значення чи відносні?

11. Загальна оцінка дисертації

Оцінюючи роботу в цілому, вважаю, що в дисертації отримано нове рішення важливої науково-технічної задачі в теорії побудови та

використання систем виявлення кібернетичних впливів на комп'ютерні мережі. Дисертація є завершеною науково-дослідною роботою. Вважаю, що за актуальністю вибраної теми, обсягом і рівнем виконаних теоретичних і експериментальних досліджень, достовірністю і обгрунтованістю висновків, новизною досліджень, значенням отриманих результатів для науки і практики дисертаційна робота задовольняє вимогам п. 9, 11, 12, 13 «Порядку присудження наукових ступенів», затвердженого Постановою КМУ від 19 серпня 2015 року № 656, а її автор, Мартовицький Віталій Олександрович, заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – комп'ютерні системи та компоненти.

Офіційний опонент
завідувач кафедри захисту інформації
Вінницького національного
технічного університету,
д.т.н., професор В.А. Лужецький

Підпис Лужецького В.А
ПОСВІДЧУЮ
Зав. канцелярією