В І Д Г У К

офіційного опонента
на дисертаційну роботу Карпачева Ігоря Ігоровича
«Інформаційна технологія забезпечення функціональної безпеки мобільних пристроїв», поданої на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології

Актуальність теми дисертації

Мобільні пристрої у наш час є необхідним елементом як складних інформаційних систем, так і щоденного життя громадянина. З використанням мобільних застосунків реалізуються різноманітні фінансові та соціальні послуги, виконується зберігання та обмін інформацією, здійснюється керування Інтернет-речами, відбувається управління в кіберфізичних системах. Фактично там, де раніше використовували комп'ютер, тепер все частіше використовують мобільний пристрій. Тому вимоги до надійності мобільних пристроїв значно зросли в останні роки.

У дисертації автором розроблені методи динамічного виявлення та блокування потенційно небезпечних додатків та модель прав доступу при взаємодії ОС Android із програмними додатками. Розроблені моделі та методи доведені до практичної реалізації у вигляді програмного комплексу. Експериментальне дослідження, яке проведено, доводить ефективність розроблених методів та моделей.

Привертає увагу застосування методів біоінформатики для порівняння послідовностей викликів API функцій. Поведінковий аналіз діяльності мобільного застосунку з метою виявлення шкідливих дій надає можливість виявляти безпосередньо дії застосунку, що можуть бути шкідливими, замість відшукування фрагментів зловмисного коду у програмах, що запускаються на виконання. Застосування методів вирівнювання послідовностей, що використовують для аналізу геномів, надало можливість автору дисертаційного дослідження підвищити ефективність виявлення шкідливих мобільних застосунків у порівнянні з широко відомими. При цьому спостерігається не тільки збільшення частки правильно визначених шкідливих мобільних застосунків, але й зменшення частки хибно визначених шкідливих мобільних застосунків, що важливо для зменшення кількості хибних попереджень користувача про шкідливі мобільні застосунки.

Таким чином, тему дисертаційної роботи І. І. Карпачева слід характеризувати як актуальну, спрямовану на розв'язання наукового завдання підвищення функціональної безпеки мобільних пристроїв.

Наукова новизна, **обґрунтованість і достовірність наукових положень, висновків і рекомендацій, сформульованих у дисертації**

Наукові положення, отримані особисто здобувачем, та їх новизна полягає в тому, що:

– *вперше запропонована* класифікація типів небезпечних додатків, яка, на відміну від існуючих, базується на групуванні дозволів на використання API функцій за ступенем потенційних впливів, що дає можливість оцінити застосунки за рівнем небезпеки для користувача при прийнятті рішень на їх використання;

– *вперше розроблена* модель прав доступу при взаємодії ОС Android з програмними застосунками, яка, на відміну від існуючих, встановлює відношення між групами, дозволами та API функціями, що надає можливість використовувати псевдосимволи функцій для прискорення ідентифікації застосунків;

– *вдосконалено* метод динамічного виявлення потенційно небезпечних мобільних застосунків за рахунок використання сигнатур функціональних ланцюжків застосунків, що будуються та порівнюються під час виконання застосунків, що дозволяє оцінити ризик при їх ідентифікації;

– *набула подальшого розвитку* інформаційна технологія забезпечення функціональної безпеки мобільних пристроїв, яка, на відміну від існуючих, надає можливість здійснювати динамічне управління процесами використання застосунків ОС Android за рахунок визначення вектору атаки.

Наукові результати дисертації, висновки й рекомендації достатньо обґрунтовані, оскільки вони отримані шляхом коректного використання математичного апарату теорії ймовірностей та математичної статистики, теорії множин, системного аналізу, методів біоінформатики. Достовірність отриманих результатів підтверджується коректним використанням математичних методів математичної статистики, експериментальними дослідженнями, апробацією отриманих практичних результатів на науково-технічних конференціях, а також використанням результатів дисертаційного дослідження при розробці захищеної системи електронного голосування «Mobile-RADA» для Чернігівського регіону.

Наукове та практичне значення отриманих у дисертації результатів полягає у розробці інформаційної технології забезпечення функціональної безпеки мобільних застосунків має практичне втілення у вигляді програмного комплексу, до складу якого входять програмний засіб AMalDetector, що призначений для моніторингу процесу виконання мобільних застосунків та підтримки прийняття рішення щодо шкідливості мобільного застосунку в умовах операційної системи Android шляхом комунікації з серверним ядром аналізатора CMMD, та програмний засіб CMMD, що призначений для виявлення схожості та ідентичності аналізованої та шаблонної сигнатур послідовностей викликів API функцій з використанням методів глобального та локального вирівнювання послідовностей.

Повнота викладення здобувачем основних результатів дисертаційної роботи в публікаціях

Основний зміст дисертації повною мірою представлено у 10 наукових працях, з них 1 стаття у іноземному періодичному виданні країни ЄС, 4 статті у фахових виданнях України категорії «Б», 2 статті у фахових виданнях з переліку до 2020 року. Результати наукового дослідження достатньо обговорювались на трьох міжнародних наукових конференціях.

Оцінка мови, стилю та оформлення дисертації й автореферату

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків та чотирьох додатків. Загальний обсяг дисертації складає 141 сторінка тексту, з яких 110 сторінок основного тексту, 142 найменувань літературних джерел, 52 рисунків та 9 таблиць. У додатках наведено документи, що підтверджують впровадження у розробку інформаційних технологій результатів досліджень дисертаційної роботи.

Автореферат відповідає змісту дисертації, написаний українською мовою, з використанням сучасної української наукової термінології. Опубліковані в авторефераті положення співпадають з основними положеннями дисертаційної роботи.

Зауваження щодо оформлення дисертаційної роботи та термінології.

1. Терміни «застосунок» і «додаток» автором використовуються як синоніми терміну «застосунок», а це не так на сьогоднішній день. Термін «додаток» використовують для позначення невеличких прикладних програм, що розширюють функції основного програмного застосунку, які англійською називають «add-on», а «застосунок» використовують для означення широкого класу прикладних програм, які англійською називають «application».
2. Використання терміну «ланцюжок» в математичному описі не є коректним. У даному випадку слід спиратись на поняття «послідовність».
3. На стор. 28 помилково вказаний стандарт ISO 61508, оскільки по тексту дисертації мова йде про стандарт IEC 61508. Останній визначає стандарт для електричних та електронних технологій і не стосується програмного забезпечення.
4. Поняття «середня довжина безперервної послідовності», що використовується на стор. 81 дисертації, не зрозуміле. На мою думку, цей термін слід формулювати як «середня довжина фрагменту послідовності, що співпадає при порівнянні з шаблоном»
5. Назва рисунку 1.12 не відповідає його змісту, оскільки вразливості, зловмисний код і т.п. не є типами кіберзагроз.
6. У блок-схемі, наведеній на рис. 2.17 (стор. 79) використовуються змінні I та I_{cr}, які не описані в тексті дисертації. Напевно, це коефіцієнт ідентичності, який автор вводить пізніше на стор.83.
7. На рис. 2.19 у першій послідовності пропущений символ B після першого символу A.

8. У таблиці 4.2 (стор.119) наведене значення показника «Повнота», яке співпадає зі значенням «TPR». Показник «Повнота» у тексті автореферату до таблиці не роз'яснений. Напевно, це показник Recall = TPR, що використовується для розрахунку F-score. Оскільки його значення співпадає з TPR, його можна вилучити з таблиці.
9. Має місце дублювання інформації у висновках. Кількість пунктів у висновках було б добре скоротити до 8.

Зауваження до автореферату дисертації

1. На стор. 4 вказано помилково значення показника FPR 0.02%, оскільки в експерименті, результати якого вказані на стор. 15, зафіксовано значення цього показника 0.2%.
2. На сторінці 5 перехоплення зловмисниками конфіденційних даних віднесено до функціональної безпеки. А по тексту дисертації така дія відноситься до інформаційної безпеки.
3. На сторінці 8 використовується необґрунтовано термін «вектор» для послідовності. Його слід замінити тут на множину або послідовність.
4. На стор. 10 у C_2 пропущений символ B, оскільки на рисунку 4 використовується послідовність символів A, B, K, C, W, B.

Зауваження щодо змісту дисертації

1. Метою роботи вказано покращення функціональної безпеки. Проте, в тексті дисертації обґрунтовується підвищення ефективності визначення шкідливих застосунків. Тому формулювання «покращення» варто замінити на «підвищення».
2. Інформація, наведена у розділі 1, недостатньо чітко структурована. Не розкриті означення ключових термінів, що використовуються в дисертації, таких як «функціональна безпека», «вектор атаки». Відсутній порівняльний аналіз існуючих аналогів систем, що забезпечують функціональну безпеку, у вигляді таблиці.
3. Модель прав доступу демонструється на дещо застарілому датасеті проекту Android Malgenome Project 2015 року і не даються роз'яснення як побудувати аналогічну модель прав доступу на нових датасетах. Представлення моделі недостатньо абстраговано для використання в інших операційних системах.
4. Ряд зауважень до математичного опису моделей та методів. Позначання множин, послідовностей і векторів не завжди коректні. Наприклад, вираз $C^q = (c_i^q)$ (стор. 80 дисертації) не є математично коректним. Якщо використовувати вектори, то кількість компонентів мала б бути однаковою для всіх векторів, але ж автор пише про порівняння послідовностей різної довжини. Тому автору слід було обмежитись послідовностями, які математично описують з використанням виключно фігурних дужок.

5. У таблиці 2.5 використовуються псевдосимволи P та K, які не наведені у таблиці псевдосимволів 2.4. тому не ясно, виклик яких API функцій вони символізують.
6. Коефіцієнт ідентичності I (формула (2.7)) варто не множити на 100, щоб потім не ділити при обчисленні ризику виявлення шкідливого застосунку. Потрібно уникати надлишкових обчислень. Не обґрунтовано у розрахунку коефіцієнту ідентичності, чому ділиться на максимальну з двох довжин порівнюваних послідовностей. Щодо коефіцієнту подібності (формула (2.9)), аналогічне зауваження.
7. У формулах коефіцієнтів ідентичності та подібності використовується максимальне зі значень довжин порівнюваних послідовностей. Це означає, що при збільшенні довжини послідовності, в якій відшукується шаблон послідовності шкідливого мобільного застосунку, значення цих коефіцієнтів буде значно зменшуватись, що нелогічно. На мою думку, тут слід розглянути мінімальне зі значень довжин порівнюваних послідовностей.
8. Для оцінювання ризику виявлення шкідливого застосунку (формула(2.8)) запропонована евристична формула, яка не виведена з позицій теорії ймовірності. По-друге, запропонований вираз не оцінює ризик виявлення шкідливого застосунку, оскільки він буде малим при великому значення ідентичності і, навпаки, буде близьким до одиниці при маленьких значеннях ідентичності. Такий вираз можна використовувати для оцінювання рівня впевненості, що застосунок є дійсно шкідливим, проте не для оцінювання ризику виявлення шкідливого мобільного застосунку.
9. У висновках до розділу 2 автор стверджує, що введено поняття вектору атаки, проте таке поняття в тексті розділу 2 чітко не сформульовано. Далі, у висновках, автор пише, що визначені групи дозволів полегшують аналіз послідовностей викликів API функцій, проте метою розробки нових методів є збільшення швидкості аналізу послідовностей. Обґрунтування чи експериментальне доведення ефекту «полегшення» не наведено в тексті дисертації.
10. На рис.3.13 наведено попередження про одну дію мобільного застосунку. На мою думку, слід інформувати про всю виявлену послідовність викликів API, яка визначена як шкідлива.
11. Є ряд зауважень до оформлення списку літературних джерел. Наприклад, відшукати публікацію, яка вказана у списку літературних джерел за номером 37, не вдається, оскільки не вказаний автор публікації і doi не відповідає вказаній публікації.

Вказані зауваження не мають принципового значення та не зменшують наукової цінності дисертаційної роботи.

Загальні висновки

На підставі вивчення тексту дисертаційної роботи та супровідних матеріалів можна зробити висновок, що за рівнем наукових і практичних результатів, ступенем їх впровадження, відповідністю паспорту спеціальності, дисертація Карпачева Ігоря Ігоровича є завершеною науковою працею, яка повністю відповідає вимогам пп. 9, 11 "Порядку присудження наукових ступенів" до дисертацій на здобуття наукового ступеня кандидата технічних наук та Наказу МОН України №40 від 12.01.2017 р. «Про затвердження вимог до оформлення дисертації». Вважаю, що Карпачев Ігор Ігорович, заслуговує на присудження йому наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент:
професор кафедри інформатики та програмної інженерії
Національного технічного університету України
«Київський політехнічний інститут
імені Ігоря Сікорського»,
д.т.н., професор І.В. Стеценко

Підпис Стеценко І.В. засвідчую.
Вчений секретар КПІ ім. Ігоря Сікорського,
к.т.н., доцент В.В. Холявко