Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0824U001783

Особливі позначки: відкрита

Дата реєстрації: 02-05-2024

Статус: Наказ про видачу диплома

Реквізити наказу МОН / наказу закладу: Наказ НТУ "ХПІ" №1112СТ від 12.07.2024

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Горносталь Олексій Андрійович
2. Oleksii A. Hornostal

Кваліфікація:

Ідентифікатор ORCHID ID: 0000-0001-5820-9999

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 123

Назва наукової спеціальності: Комп’ютерна інженерія

Галузь / галузі знань: інформаційні технології

Освітньо-наукова програма зі спеціальності: ОП 28983 Комп’ютерна інженерія

Дата захисту: 20-06-2024

Спеціальність за освітою: 123 Комп'ютерна інженерія

Місце роботи здобувача: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): ДФ 64.050.138-5697

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 50.37.23, 20.55, 20.55.01

Тема дисертації:

1. Ансамблевий метод ідентифікації стану комп'ютерних систем
2. Ensemble method of computer system state identification

Реферат:

1. Дисертаційна робота присвячена вирішенню актуальної науково-прикладної задачі вдосконалення, розробки та впровадження методів ідентифікацій стану комп’ютерних систем з метою покращення їх ефективності за рахунок використання ансамблевих методів машинного навчання. Метою дисертаційної роботи є підвищення якості ідентифікації стану комп’ютерних систем шляхом розробки та удосконалення методів розпізнавання аномалій та зловживань. Об’єкт дослідження – процес виявлення вторгнень у комп’ютерні системи в умовах зовнішніх впливів. Предмет дослідження – методи ідентифікації стану комп’ютерних систем на основі технології машинного навчання з використанням ансамблевих мета-алгоритмів. За результатами дослідження отримано такі наукові результати: 1. Отримав подальший розвиток метод ідентифікації стану комп’ютерної системи на основі дерев рішень та мета-алгоритму беггінг за

рахунок вибору оптимальних гіперпараметрів налаштування класифікатора та використання процедури попередньої обробки даних, яка сфокусована на видаленні аномальних даних та зменшенні статистичної залежності між ознаками, що дозволило підвищити якість ідентифікації стану КС. 2. Отримав подальший розвиток ансамблевий метод ідентифікації стану комп'ютерної системи завдяки використанню багатошарового перцептрону у якості базової моделі ансамблю та вибору оптимальних гіперпараметрів налаштування класифікатора, що дозволило підвищити якість його функціонування. 3. Удосконалено ансамблевий метод ідентифікації стану комп'ютерної системи на основі гомогенного мета-алгоритму беггінг за рахунок розробки спеціальної процедури зменшення кількості базових класифікаторів та їх ранжування під час зваженого голосування, що дозволило зменшити час роботи ансамблю та підвищити якість класифікації стану КС. 4. Вперше запропоновано метод ідентифікації стану комп'ютерної системи, який відрізняється від відомих методів використанням гетерогенного мета-алгоритму беггінг та включає триетапний процес підбору базових моделей класифікатора на основі технології Pasting, що дозволило підвищити ефективність ідентифікації стану КС. Практичне значення отриманих результатів полягає в наступному: - сформовано програмну модель попередньої обробки даних, яка сфокусована на видаленні аномальних даних та зменшенні статистичної залежності між ознаками, що дозволяє збільшити швидкість розпізнавання до 1,62 разів, зменшити час навчання моделі до 24,76 разів, а також підвищити якість класифікації; - розроблено метод ідентифікації стану комп'ютерної системи, який включає сформовану процедуру попередньої обробки даних, процес вибору алгоритму формування вхідних даних та побудову беггінг-класифікатора з налаштуванням його гіперпараметрів, що дозволило підвищити якість класифікації: значення AUC-ROC класифікатору на навчальній вибірці зростає на 11%, а на тестовій вибірці – на 3%; - реалізовано програмну модель ансамблевого класифікатору на основі багатошарового перцептрону у якості базового класифікатора та процедури підбору оптимальних налаштувань його параметрів, а саме: алгоритм формування вибірок даних, кількості базових класифікаторів, функцію оптимізації ваг нейронних мереж, розмірів першого та другого прихованих шарів та функцію активації, що дозволило підвищити значення точності класифікації на 4,67%; - розроблено програмне забезпечення, яке виконує обрізку ансамблю на основі максимізації абсолютної точності базових класифікаторів та класифікує за допомогою зваженого голосування з використанням вагових коефіцієнтів на основі функції логарифмічних втрат, що дозволило підвищити показники якості класифікації беггінг-ансамблю, а саме значення метрики F1- Score – на 2,4%; - запропоновано метод формування гетерогенного ансамблю, який включає відбір базових класифікаторів, навчання на їх основі однорідних беггінг-ансамблів, створення комбінаційних груп (пулів) із базових класифікаторів та формування гетерогенного ансамблю за допомогою процедури Pasting, що дозволило підвищити якість класифікації, а саме збільшити показник F1-Score моделі при роботі на тестових даних на 9,5% у порівнянні зі стандартним однорідним беггінг-ансамблем на основі дерев рішень та на 2% у порівнянні з максимальним значенням серед однорідних ансамблів. За результатами дослідження підтверджено теоретичну та практичну цінність, проведено дослідження їх ефективності та сформовано практичні рекомендації, щодо їх застосування.

2. The dissertation work is devoted to the solution of the actual scientific and applied problem of improvement, development and implementation of methods for identifying the state of computer systems with the aim of improving their efficiency due to the use of ensemble methods of machine learning. The purpose of the dissertation is to improve the quality of identification of the state of computer systems by developing and improving methods for recognizing anomalies and abuses. The object of research is the process of detecting intrusions into computer systems under conditions of external influences. The subject of research is methods of identifying the state of computer systems based on machine learning technology using ensemble meta-algorithms. The following scientific results were obtained within this area: 1. The computer system state identification method based on decision trees and the bagging meta-algorithm was further developed due to the selection of optimal hyperparameters of the classifier setting and the use of a data pre-processing procedure, which is focused on removing anomalous data and reducing the statistical dependence between features, which allowed to improve the quality of state identification computer systems. 2. The ensemble method of identifying the state of the computer

system was further developed due to the use of a multilayer perceptron as the basic model of the ensemble and the selection of optimal hyperparameters for the classifier setting, which made it possible to improve the quality of its functioning. 3. The ensemble method for identifying the state of a computer system based on the homogeneous meta-algorithm of bagging has been improved by developing a special procedure for reducing the number of basic classifiers and their ranking during weighted voting, which made it possible to reduce the time of the ensemble and improve the quality of classification of the state of the CS. 4. For the first time, the method for identifying the state of a computer system was proposed, which differs from known methods by using a heterogeneous bagging meta-algorithm and includes a three-stage process for selecting basic classification models based on the Pasting technology, which made it possible to increase the efficiency of identifying the state of the computer system. The practical significance of the obtained results includes the following achievements: - a software model for data pre-processing focused on removing anomalous data and reducing the correlation of features was formed, which allows to increase the recognition speed up to 1.62 times, reduce the training time up to 24.76 times, and also improve the quality of their classification; - a method for identifying the state of the computer system was developed, which includes the established data preprocessing procedure, the process of selecting the input data generation algorithm, and the construction of a bagging classifier with the adjustment of its hyperparameters, which made it possible to improve the quality of classification: the AUC-ROC value of the classifier on the training sample increases by 11% , and on the test sample – by 3%; - the software model of an ensemble classifier based on a multilayer perceptron as a basic classifier and a procedure for selecting the optimal settings of its parameters has been implemented, which made it possible to increase the value of classification accuracy by 4.67%; - the software that performs ensemble pruning based on the maximization of the absolute accuracy of the base classifiers and classification using weighted voting based on the logarithmic loss function was developed, which allowed to improve the quality indicators of the bagging ensemble classification, namely the value of the F1-Score metric up to 2.4%; - the method for forming a heterogeneous ensemble, which includes the selection of basic classifiers, learning homogeneous bagging ensembles based on them, creating combination groups (pools) from basic classifiers and forming a heterogeneous ensemble using the Pasting procedure was developed, which made it possible to improve the quality of classification, namely to increase the F1-Score of models when working on test data by 9.5% compared to the standard homogeneous bagging ensemble based on decision trees and by 2% compared to the maximum value among homogeneous ensembles. According to the results of the research, the theoretical and practical value was confirmed, a study of their effectiveness was conducted and practical recommendations were formed regarding their application.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- 1. О. А. Горносталь та С. Ю. Гавриленко, "Розробка адаптивних шаблонів фіксації аномальної поведінки комп'ютерної системи", Зб. наукових праць Системи обробки інформації, Харків.: ХУ ПС, 2016, Вип. 3(140), с.11-14. (Б)
- 2. O. Hornostal and S. Gavrylenko, V. Chelak, and V. Vassilev, "Development of a method for identification the state of a computer system using fuzzy cluster analysis", Advanced Information Systems, Kharkiv, 2020, vol. 4, no. 2, pp. 8-11. (Б)
- 3. O. Hornostal and S. Gavrylenko, "Development of a method for identification of the state of computer systems based on bagging classifiers", Advanced Information Systems, 2021, vol. 5, no. 4, pp. 5–9. (Б)

- 4. О. Горносталь та С. Гавриленко, "Метод ідентифікації стану комп’ютерної системи на основі ансамблевих класифікаторів з покращеною процедурою голосування", Системи управління, навігації та зв’язку. Збірник наукових праць, 2023, т. 3, вип. 73, с. 79-85. (Б)
- 5. O. Hornostal and S. Gavrylenko, "Application of heterogeneous ensembles in problems of computer system state identification", Advanced Information Systems, 2023, vol. 7, no. 4, pp. 5–12. (Б)
- 6. О. А. Горносталь та С. Ю. Гавриленко, "Аналіз ефективності фільтрації несприятливого мережевого трафіку з використанням комплексних систем", Інформатика, управління та штучний інтелект. Матеріали другої науково-технічної конференції студентів, магістрів та аспірантів, Харків, 2015, с. 13.
- 7. О. А. Горносталь та С. Ю. Гавриленко, "Виявлення аномальної поведінки комп'ютерних систем за допомогою контрольних карт Шухарта та карт кумулятивних сум", Матеріали міжнародної конференції «Проблеми науково-технічного та правового забезпечення кібербезпеки у сучасному світі», Харків, 2016, с.14-15.
- 8. O. Hornostal, V. Chelak, S. Gavrylenko and, S. Gornostal, "Intrusion detection in computer systems", Proceedings of the symposium "Metrology and metrology assurance", Sozopol, Bulgaria, 2016, pp. 342-347.
- 9. O. Hornostal, S. Gavrylenko, and V. Chelak, "Development of a heuristic scanner for an antivirus program on the basis of the Mamdani fuzzy logic method", Proceedings of the 28th International Scientific Symposium Metrology and Metrology Assurance, Sozopol, Bulgaria, 2018, pp.129-133.
- 10. O. Hornostal, V. Chelak, S. Gavrylenko, and S. Gornostal, "Identification of the computer system state based on multidimensional discriminant analysis", in Proceedings of the 29th International Scientific Symposium Metrology and Metrology Assurance, Sozopol, Bulgaria, 2019, pp. 192-196. (Scopus, Bulgaria)
- 11. O. Hornostal, and S. Gavrilenko, "Identification of Anomalies in the Behavior of a Computer System using Fuzzy Cluster Analysis", Proceedings of the 7th International Informatics, management and artificial intelligence, Kharkiv, 2019, p. 21.
- 12. O. Hornostal, V. Chelak, and S. Gavrylenko, "Research of Intelligent Data Analysis Methods for Identification of Computer System State", in Proceedings of the 30th International Scientific Symposium Metrology and Metrology Assurance (MMA), Sozopol, Bulgaria, 2020, pp. 1-5. (Scopus, Bulgaria)
- 13. O. Hornostal, S. Gavrylenko and V. Chelak, "Ensemble approach based on bagging and boosting for identification the computer system state", in Proceedings of the 31th International Scientific Symposium Metrology and Metrology Assurance, Sozopol, Bulgaria, 2021, pp. 1-7. (Scopus, Bulgaria)
- 14. О. А. Горносталь та С. Ю. Гавриленко, "Дослідження методів підвищення ефективності роботи беггінг-класифікаторів у задачах ідентифікації стану комп’ютерних систем", Матеріали VIII міжнародної науково-технічної конференції “Інформатика, управління та штучний інтелект" (ІУШІ-2021), Харків, 2021.
- 15. О. А. Горносталь та С. Ю. Гавриленко, "Дослідження беггінг-алгоритмів для ідентифікації стану комп’ютерної системи", Матеріали IV Всеукраїнської науково-практичної інтернет-конференції студентів, аспірантів та молодих вчених за тематикою «Сучасні комп’ютерні системи та мережі в управлінні»: збірка наукових праць, під редакцією Г.О. Райко, Херсон, 2021, с. 27-28.
- 16. О. А. Горносталь, та С. Ю. Гавриленко, "Дослідження та вдосконалення методів підвищення точності роботи bagging-ансамблів для классифікації стану комп’ютерних систем", на дев'ятій міжнародній науково-технічної конференції "Інформатика, Управління та Штучний Інтелект" (ІУШІ-2022), Харків - Краматорськ, 2022, с. 29.
- 17. O. Hornostal, S. Gavrylenko, and V. Chelak, "Construction Method of Fuzzy Decision Trees for Identification the Computer System State", in Proceedings of the 32th International Scientific Symposium Metrology and Metrology Assurance, Sozopol, Bulgaria, 2022, pp. 1-5. (Scopus, Bulgaria)
- 18. O. Hornostal, S. Gavrylenko, and V. Chelak, "Research of Methods of Identifying the Computer Systems State based on Bagging Classifiers", in IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek), Kharkiv, Ukraine, 2022, pp. 1-6. doi: 10.1109/KhPIWeek57572.2022.9916439. (Scopus, Ukraine)
- 19. O. Hornostal and S. Gavrylenko, "Study of Methods for Improving the Meta-Algorithm of the Bagging Classifier", 2023 IEEE 4th KhPI Week on Advanced Technology (KhPIWeek), Kharkiv, Ukraine, 2023, pp. 1-6.

(Scopus, Ukraine)

- 20. О. А. Горносталь та С. Ю. Гавриленко, "Метод підвищення якості ансамблевого класифікатору за рахунок диверсифікації базових моделей", XXIII Міжнародна науково-технічна конференція Проблеми інформатики та моделювання, Харків, 2023, с. 33-34.

Наукова (науково-технічна) продукція: методи, теорії, гіпотези; програмні продукти, програмно-технологічна документація

Соціально-економічна спрямованість: підвищення захисту інформації в комп'ютерних системах та мережах

Охоронні документи на ОПІВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами: 0122U200526

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Гавриленко Світлана Юріївна

2. Svitlana Y. Gavrylenko

Кваліфікація: д. т. н., професор, 05.13.05

Ідентифікатор ORCHID ID: 0000-0002-6919-0055

Додаткова інформація: https://www.scopus.com/authid/detail.uri?authorId=57189042150; https://scholar.google.com.ua/citations?user=4Vn1dBkAAAAJ&hl=ua; https://www.webofscience.com/wos/author/record/2299382

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Мелешко Єлизавета Владиславівна

2. Yelyzaveta V. Meleshko

Кваліфікація: д. т. н., професор, 05.13.05

Ідентифікатор ORCHID ID: 0000-0001-8791-0063

Додаткова інформація: https://www.scopus.com/authid/detail.uri?authorId=57212031323; https://www.webofscience.com/wos/author/record/AAD-6538-2022; https://scholar.google.com.ua/citations?user=hz93GDsAAAAJ

Повне найменування юридичної особи: Центральноукраїнський національний технічний університет

Код за ЄДРПОУ: 02070950

Місцезнаходження: просп. Університетський, буд. 8, Кропивницький, Кропівницький р-н., 25006, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Трубчанінова Карина Артурівна
2. Karyna A. Trubchaninova

Кваліфікація: д. т. н., професор, 05.13.05

Ідентифікатор ORCHID ID: 0000-0003-2078-2647

Додаткова інформація: https://www.scopus.com/authid/detail.uri?authorId=57208109791; https://www.webofscience.com/wos/author/record/2514056; https://scholar.google.com.ua/citations?user=tFdGngkAAAAJ&hl=uk

Повне найменування юридичної особи: Український державний університет залізничного транспорту

Код за ЄДРПОУ: 01116472

Місцезнаходження: майдан Фейєрбаха, буд. 7, Харків, Харківський р-н., 61050, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Поворознюк Анатолій Іванович
2. Anatolii I. Povoroznyuk

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCHID ID: 0000-0003-2499-2350

Додаткова інформація: https://www.scopus.com/authid/detail.uri?authorId=55225664000; https://www.webofscience.com/wos/author/record/2204478; https://scholar.google.com/citations?user=g6S23QsAAAAJ

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Кучук Георгій Анатолійович
2. Heorhii A. Kuchuk

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCHID ID: 0000-0002-2862-438X

Додаткова інформація: https://www.scopus.com/authid/detail.uri?authorId=57057781300; https://www.webofscience.com/wos/author/record/2485726; https://scholar.google.com.ua/citations?user=gHejYRUAAAAJ

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради	Заковоротний Олександр Юрійович
Власне Прізвище Ім'я По-батькові головуючого на засіданні	Заковоротний Олександр Юрійович
Відповідальний за підготовку облікових документів	Зайцев Юрій Іванович
Реєстратор	УкрІНТЕІ

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності

Юрченко Тетяна Анатоліївна