# Облікова картка дисертації

## I. Загальні відомості

Image not found or type unknow

**Державний обліковий номер:** 0824U000903

**Особливі позначки:** відкрита

**Дата реєстрації:** 13-02-2024

**Статус:** Наказ про видачу диплома

**Реквізити наказу МОН / наказу закладу:** Наказ Чернівецького національного університету імені Юрія Федьковича від 06 березня 2024 р. №76 Про видачу диплома доктора філософії

## II. Відомості про здобувача

**Власне Прізвище Ім'я По-батькові:**

1. Кириченко Оксана Леонідівна

2. Oksana L. Kyrychenko

**Кваліфікація:**

**Ідентифікатор ORCID ID:** 0000-0003-0282-9958

**Вид дисертації:** доктор філософії

**Аспірантура/Докторантура:** так

**Шифр наукової спеціальності:** 121

**Назва наукової спеціальності:** Інженерія програмного забезпечення

**Галузь / галузі знань:**

**Освітньо-наукова програма зі спеціальності:** Інженерія програмного забезпечення

**Дата захисту:** 19-02-2024

**Спеціальність за освітою:** Математик

**Місце роботи здобувача:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, 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 76.051.044 (ID 4519)

**Повне найменування юридичної особи:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, 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:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

## V. Відомості про дисертацію

**Мова дисертації:** Українська

**Коди тематичних рубрик:** 20.54.07

**Тема дисертації:**

1. Дослідження статистичних характеристик складних мереж методами інтелектуального аналізу даних

2. The Study of Statistical Characteristics of Complex Networks by Methods of Intelligent Data Analysis

**Реферат/анотація:**

1. Дисертаційна робота присвячена дослідженню статистичних характеристик складних мереж та кластерної структури веб-простору з використанням методів інтелектуального аналізу даних, зокрема розробці інформаційної технології для кластеризації даних великого розміру, які були зібрані й оброблені спеціально створеним програмним забезпеченням. Також вивчено стохастичні матриці, які завдяки особливостям своїх спектральних властивостей є основним математичним об'єктом при дослідженні кластерної структури веб-простору. Дисертація складається зі вступу, чотирьох розділів, висновків, переліку використаних джерел та додатків. У вступі обґрунтовано актуальність дослідження, сформульовано мету, завдання, предмет, об'єкт та методи дослідження, вказано наукову новизну, теоретичне та практичне значення отриманих результатів, проаналізовано зв'язок роботи з науковими темами. Зазначено особистий внесок здобувача, а також

наведено відомості про апробацію та публікації основних результатів дисертації. Перший розділ містить ключові відомості з теорії складних мереж, опис основних напрямів досліджень та завдання, якими займається теорія складних мереж. Проведено огляд та опис основних моделей (Ердоша–Рені, Уаттса–Строґаца, Барабаші–Альберт). Розглянуто та проаналізовано приклади реальних складних мереж та їх особливості Здійснено класифікацію та огляд методів однієї з важливих технік інтелектуального аналізу складних мереж – кластерного аналізу. У другому розділі дисертаційного дослідження описано концепцію кроулінгу як одного із засобів збирання інформації, проведено огляд існуючих програмних засобів для збирання інформації у веб-просторі. Другий розділ має прикладне значення, основним його результатом є розроблене спеціалізоване програмне забезпечення – кроулер з вбудованим аналітичним модулем для інтелектуальної обробки інформації. Третій розділ присвячений дослідженню освітніх сегментів веб-простору (українського (edu.ua), ізраїльського (ac.il) та польського (edu.pl), інформація про які була зібрана та оброблена за допомогою самостійно розробленої інформаційної технології. Застосування даної розробки дозволило отримати статистичні характеристики та кластерну структуру вказаних вище сегментів веб-простору та здійснити порівняльний аналіз. У четвертому розділі розглянуто питання кластеризації в графі на основі матриці суміжності. Основним об'єктом дослідження є стохастична матриця, що задає ймовірності переходу на графі та визначається із матриці суміжності. У розділі проаналізовано спектральні властивості стохастичної матриці із врахуванням кластерної структури графу. Основні теоретичні результати розділу наступні: доведено факт збіжності власних значень матриці P за умов, накладених на елементи матриці суміжності A (теорема 4.3.1); встановлений факт про асимптотичну еквівалентність спектрів матриць та дозволяє використовувати стохастичну матрицю із незалежними елементами замість відповідної стохастичної матриці P, елементи якої не є незалежними (лема 4.4.1); розглянуто частинний підхід до оцінки розподілу елементів матриці P за умови показникового розподілу елементів матриці A (лемах 4.5.1 та 4.5.2). Розроблено новий алгоритм перевірки належності елементів (вершин графу) до одного кластеру; побудовано критерій оцінки оптимальної кількості кластерів k_opt. На основі методу Монте-Карло проведено порівняння розробленого методу оцінки кількості кластерів із деякими класичними алгоритмами, в результаті чого встановлено ряд його переваг та недоліків. У висновках підсумовано основні результати дисертаційного дослідження. У додатках подано наукові публікації, відомості про апробацію результатів дисертації. Теоретичне значення. Результати теоретичних досліджень, а саме розвитку теорії графових досліджень, сформульовані та доведені леми і теореми, можуть використовуватися для подальших досліджень у цій галузі, а також у навчальних курсах кафедр математичних проблем управління і кібернетики та програмного забезпечення комп'ютерних систем Чернівецького національного університету імені Юрія Федьковича (та інших ЗВО), пов'язаних з інтелектуальним аналізом даних, методичних розробках, навчальних посібниках для освітнього процесу та науково-дослідної роботи студентів, аспірантів. Практичне значення. Розроблені у дисертаційній роботі кроулер, інформаційна технологія та метод визначення оптимальної кількості кластерів можуть в подальшому використовуватися для практичного дослідження складних мереж. Запропоновані підходи, результати використовуються компаніями «Квант Азимут» та «Qlicks B.V.». Ключові слова: модель (математична, економічна), моделювання, динаміка, інтелектуальний аналіз даних, кластеризація, k-means, інформаційна система, інформаційна технологія, інтелектуальна система, програмне забезпечення, тестування програмного забезпечення, рівні тестування програмного забезпечення, специфікація вимог до програмного забезпечення, функціональні та нефункціональні вимоги до програмного забезпечення, статистичні методи.

2. The dissertation deals with the study of statistical characteristics of complex networks and the cluster structure of the web space using methods of intelligent data analysis, in particular, the development of information technology for the clustering of large data, which were collected and processed by specially created software. In addition, stochastic matrices have been studied, which, due to their specific spectral properties, are the main mathematical object in the study of the cluster structure of the web space. The dissertation contains an introduction, four chapters, conclusions, a list of literature, and appendices. The introduction substantiates the relevance of the research topic, formulates the goal, task, subject, object and research methods; indicates the

scientific novelty, theoretical and practical significance of the obtained results; presents and analyzes the link between the current research and scientific topics. The personal contribution of the candidate, as well as information about the approval and publication of the main results of the research are shown. Chapter 1 contains key information on the theory of complex networks, a description of the main areas of research and tasks that the theory of complex networks deals with. An overview and description of the main models (the Erdős–Rényi, Watts-Strogatz, Barabási–Albert models) are provided. The methods of cluster analysis, an important technique of intellectual analysis of complex networks, are classified and reviewed. Chapter 2 of the dissertation describes the concept of crawling as one of the means of gathering information, and provides an overview of existing software tools for collecting information in the web space. Chapter 2 has a practical value; its main result is the developed specialized software – a crawler with a built-in analytical module for intelligent information processing. Chapter 3 deals with the study of following educational segments of the web space: Ukrainian (edu.ua), Israeli (ac.il) and Polish (edu.pl). The information on these segments was collected and processed using a personally developed information technology. The application of this development enables to obtain the statistical characteristics and cluster structure of the above-mentioned segments of the web space and to carry out a comparative analysis. Chapter 4 deals with the issue of clustering in a graph based on the adjacency matrix. The main object of research is the stochastic matrix, which specifies the transition probabilities on the graph and is determined from the adjacency matrix. In this chapter, the spectral properties of the stochastic matrix are analyzed, taking into account the cluster structure of the graph. The main theoretical results of this chapter are follows: the fact of the convergence of the eigenvalues of the matrix P under the conditions imposed on the elements of the adjacency matrix A is proved (theorem 4.3.1); the established fact about the asymptotic equivalence of the spectra of the matrices and allows using a stochastic matrix with independent elements instead of the corresponding stochastic matrix P, the elements of which are not independent (lemma 4.4.1); a partial approach to the estimation of the distribution of the elements of matrix P is considered under the condition of the indicative distribution of matrix A elements (lemmas 4.5.1 and 4.5.2). A new algorithm has been developed for checking whether elements (graph vertices) belong to one cluster. a criterion for estimating the optimal number of clusters k_opt is built. . On the basis of the Monte Carlo method, a comparison of the developed method for estimating the number of clusters with some classical algorithms was carried out. The main results of the dissertation research are summarized in the conclusions. The appendices present scientific publications, information on the approval of the results of the dissertation. Theoretical significance. The results of theoretical research, namely the development of the theory of graph research, formulated and proven lemmas and theorems, can be used for further research in this field. They can also be applied in the educational courses in Yuriy Fedkovych Chernivtsi National University, related to teaching aids for the educational process and research. Practical significance. The crawler, information technology, and method for determining the optimal number of clusters developed in the dissertation can be used for further practical research of complex networks. The proposed approaches are used by the companies "Kvant Azimuth" and "Qlicks B.V.". Keywords: model (mathematical, economic), simulation, dynamics, intelligent data analysis, clustering, k-means, information system, information technology, intelligent system, software, software testing, software testing levels, specification of software requirements, functional and non-functional software requirements, statistical methods.

**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:** Теоретичне узагальнення і вирішення важливої наукової проблеми

**Публікації:**

- 1. Kyrychenko O., Ostapov S., Kanovsky I. Investigation of the certain internet domain statistical characteristics / Статистичні характеристики деяких зон інтернету та їх дослідження. Eastern-European Journal of Enterprise Technologies. 2013. Vol. 6, no. 12(66). P. 91–96.
- 2. Кириченко О.Л., Малик І.В., Остапов С.Е. Стохастичні моделі в задачах штучного інтелекту. Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки. 2021. № 2. С. 53–57.
- 3. Kyrychenko O. Information technology for statistical cluster analysis of information in complex networks. Computer Systems and Information Technologies. 2022. No 4. P. 47–51.
- 4. Кириченко О. Особливості архітектури програмного забезпечення для збору та аналізу статистичної інформації в глобальній мережі. Information Technology: Computer Science, Software Engineering and Cyber Security. 2023. № 2. С. 107–112.


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

**Соціально-економічна спрямованість:**

**Охоронні документи на ОПІВ:**

**Впровадження результатів дисертації:** Впроваджено

**Зв'язок з науковими темами:** 011U007046, 0121U109232

## VI. Відомості про наукового керівника/керівників (консультанта)

### Власне Прізвище Ім'я По-батькові:

1. Остапов Сергій Едуардович

2. Serhii E. Ostapov

**Кваліфікація:** д.ф.-м.н., професор, 01.04.10

**Ідентифікатор ORCID ID:** Не застосовується

**Додаткова інформація:**

**Повне найменування юридичної особи:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, 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. Криворучко Олена Володимирівна

2. Olena Kryvoruchko

**Кваліфікація:** д. т. н., професор, 05.13.22

**Ідентифікатор ORCID ID:** 0000-0002-7661-9227

**Додаткова інформація:**

**Повне найменування юридичної особи:** Державний торговельно-економічний університет

**Код за ЄДРПОУ:** 44470624

**Місцезнаходження:** вул. Кіото, буд. 19, Київ, 02156, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

**Власне Прізвище Ім'я По-батькові:**

1. Говорущенко Тетяна Олександрівна

2. Tetiana Hovorushchenko

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

**Ідентифікатор ORCID ID:** 0000-0002-7942-1857

**Додаткова інформація:**

**Повне найменування юридичної особи:** Хмельницький національний університет

**Код за ЄДРПОУ:** 02071234

**Місцезнаходження:** вул. Інститутська, буд. 11, Хмельницький, Хмельницький р-н., 29016, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

**Рецензенти**

**Власне Прізвище Ім'я По-батькові:**

1. Ушенко Юрій Олександрович

2. Yurii O. Ushenko

**Кваліфікація:** д. ф.-м. н., професор, 01.04.05

**Ідентифікатор ORCID ID:** 0000-0003-1767-1882

**Додаткова інформація:**

**Повне найменування юридичної особи:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

**Власне Прізвище Ім'я По-батькові:**

1. Угрин Дмитро Ілліч

2. Dmytro Uhryn

**Кваліфікація:** д. т. н., доц., 05.13.06

**Ідентифікатор ORCID ID:** 0000-0003-4858-4511

**Додаткова інформація:**

**Повне найменування юридичної особи:** Чернівецький національний університет імені Юрія Федьковича

**Код за ЄДРПОУ:** 02071240

**Місцезнаходження:** вул. Коцюбинського, буд. 2, Чернівці, 58012, Україна

**Форма власності:** Державна

**Сфера управління:** Міністерство освіти і науки України

**Ідентифікатор ROR:**

## VIII. Заключні відомості

| | |
|---|---|
| **Власне Прізвище Ім'я По-батькові голови ради** | Григорків Василь Степанович |
| **Власне Прізвище Ім'я По-батькові головуючого на засіданні** | Григорків Василь Степанович |
| **Відповідальний за підготовку облікових документів** | Кириченко Оксана Леонідівна |
| **Реєстратор** | УкрІНТЕІ |
| **Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності** | Юрченко Тетяна Анатоліївна |