УДК 004.891

д.т.н., проф. **Лєнков С.В.** (ВІКНУ)
к.т.н., доц. **Джулій В.М.** (ХмНУ)
к.т.н. **Берназ А.М.** (ШСВ ЗСУ)
к.т.н., доц. **Муляр І.В.** (ХмНУ)
к.т.н., доц. **Пампуха І.В.** (ВІКНУ)

DOI: https://doi.org/10.17721/2519-481X/2023/78-12

# МЕТОД ПРОГНОЗУВАННЯ ВРАЗЛИВОСТЕЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ АНАЛІЗУ ДАНИХ ТЕМАТИЧНИХ ІНТЕРНЕТ-РЕСУРСІВ

***В роботі продено дослідження задачі прогнозування вразливостей інформаційної безпеки на основі проведеного аналізу даних тематичних інтернет-ресурсів. На тлі стрімкого розвитку інформаційних технологій відзначається, зростання активності різноманітності комп'ютерних атак, здійснюваних і запланованих із застосуванням сучасних новітніх технологій. Очевидною проблемою інформаційної безпеки суспільства, сьогодення стала шкідлива інформація, також необхідно зазначити, що злочинні та терористичні угрупування беруть на озброєння, дедалі частіше, засоби інформаційного впливу, розробляють та пишуть стратегії, спрямовані на залучення нових адептів та розширення сфери впливу через соціальні мережі. Аналіз проведеного дослідження поточного стану в області інформаційної безпеки показує, що темпи розвитку інформаційних та комп'ютерних технологій значно випереджають процес створення програмно-апаратного забезпечення в області інформаційної безпеки. Пріоритетними, в даній ситуації, є задача аналізу, класифікації, виявлення діючих механізмів та засобів проведення атак і загроз інформаційній безпеці системи, які можуть призвести до отримання несанкціонованого доступу до конфіденційних даних, порушення функціонування інформаційної системи, визначення заходів протидії атакам та загрозам, оцінка заданої шкоди, розробка нормативно-правової бази, механізмів захисту та критеріїв інформаційної безпеки системи протидії. На сьогодні не існує єдиного підходу до вирішення проблеми захищенності інформаційно-пошукових систем, стосовно предметних областей: розробниками програмно-апаратного захисту інформації пропонуються відповідні компоненти на вирішення конкретних задач; забезпечення надійного захисту інформаційних ресурсів потребує реалізації відповідних технічних та організаційних заходів в комплексі, що супроводжуються розробкою відповідної документації. Більшість сучасних програмно-апаратних систем виявлення комп'ютерних загроз та атак працюють із використанням підходів сигнатурного аналізу та фіксації інтернет-мережевих аномалій. Дані підходи мають недоліки, пов'язані із використанням потужних обчислювальних ресурсів на їх реалізацію, при виявленні нових комп'ютерних загроз мають низьку ефективність. Метод прогнозування вразливостей інформаційної безпеки на основі даних інтернет-ресурсів, заснований на нечіткому логічному виводу, семантичному та статистичному аналізі, відрізняється можливістю виявлення вразливостей та загроз до їх реалізації, дозволяє описувати закономірності інформаційного процесу наповнення тематичних ресурсів новими текстовими повідомленнями, що відображається на якості прогнозування.***

***Реалізований в інформаційно-аналітичній системі алгоритм прогнозування вразливостей та загроз безпеки інформації на основі аналізу потоку даних тематичних ресурсів дозволяє автоматизувати інформаційний процес виявлення нових вразливостей, загроз, надає фахівцям інформаційної безпеки можливість оцінити своєчасно ступінь захищеності ресурсів та при необхідності вжити відповідних заходів щодо нейтралізації можливих загроз та вразливостей, тим самим підвищити інформаційну безпеку обчислювальних комп'ютерних систем від реалізації нових мережевих комп'ютерних атак.***

***Ключові слова: інформаційна безпека, тематичні інтернет-ресурси соціальні мережі, джерела повідомлень, вразливості, атаки, інформаційна система.***

**Вступ.** На сучасному етапі на більшість сфер діяльності суспільства зростає вплив глобальних інформаційних технологій. Відзначаються, при цьому, високі темпи розвитку світових єдиних телекомунікаційного та інформаційного просторів, сформувалися в суспільстві нові соціальні групи, виявляється значний вплив на сформований історично спосіб життя людей [1,2]. На тлі стрімкого розвитку інформаційних технологій відзначається, зростання активності різноманітності комп'ютерних атак, здійснюваних і запланованих із застосуванням сучасних новітніх технологій. Очевидною проблемою інформаційної безпеки суспільства, сьогодення стала шкідлива інформація, також необхідно зазначити, що злочинні та терористичні угрупування беруть на озброєння, дедалі частіше, засоби інформаційного впливу, розробляють та пишуть стратегії, спрямовані на залучення нових адептів та розширення сфери впливу через соціальні мережі. Актуальними та пріоритетними на сучасному етапі є задачі аналізу, класифікації виявлення існуючих механізмів реалізації атак та загроз інформаційної безпеки, які можуть призвести до отримання несанкціонованого доступу до конфіденційної інформації, порушення функціонування інформаційних систем. Однією зі складових надійного забезпечення інформаційної безпеки держави є проведення аналізу, виявлення, моніторинг та активна протидія розповсюдженню шкідливої інформації в соціальних мережах [1,3-5].

Таким чином, постає задача визначення заходів протидії атакам та загрозам, усунення вразливостей, оцінки заданої можливої шкоди, розробка нормативно-правової бази, механізмів захисту та критеріїв безпеки. Важливість даних проблем пов'язана з наступними основними факторами: зростанням різноманітності та кількості засобів комп’ютерної техніки та сфер людської діяльності їх застосування; високим рівнем довіри до інформаційно-пошукових систем обробки та управління даними; зростанням числа користувачів інформаційного простору взаємодії; накопиченням великих об’ємів різнотипної інформації, інтенсивним обміном потоком даних в мережі між користувачами, з використанням широкого спектра механізмів доступу до конфіденційних ресурсів, інформаційних процесів; промисловим шпигунством та конкурентною боротьбою у сфері інформаційних послуг суспільства; недостатньою кількістю, на сучасному етапі, фахівців високої кваліфікації в області інформаційної безпеки, ринковими відношеннями в області розробки програмного забезпечення, обслуговування, розповсюдження, виробництва обчислювальної комп'ютерної техніки для реалізації інформаційної безпеки; різноманіттям атак, загроз і різнотипних каналів отримання несанкціонованого доступу до конфіденційних ресурсів та диференціацією негативних наслідків [ 6-8,13].

**Аналіз останніх досліджень та постановка задачі.** Аналіз проведеного дослідження поточного стану в області інформаційної безпеки показує, що темпи розвитку інформаційних та комп'ютерних технологій значно випереджають процес створення програмно-апаратного забезпечення в області інформаційної безпеки. Пріоритетними, в даній ситуації, є задача аналізу, класифікації, виявлення діючих механізмів та засобів проведення атак і загроз інформаційній безпеці системи, які можуть призвести до отримання несанкціонованого доступу до конфіденційних даних, порушення функціонування інформаційної системи, визначення заходів протидії атакам та загрозам, оцінка заданої шкоди, розробка нормативно-правової бази, механізмів захисту та критеріїв інформаційної безпеки системи протидії [2, 3,9,10,13,14]. На сьогодні не існує єдиного підходу до вирішення проблеми захищенності інформаційно-пошукових систем, стосовно предметних областей: розробниками програмно-апаратного захисту інформації пропонуються відповідні компоненти на вирішення конкретних задач; забезпечення надійного захисту інформаційних ресурсів потребує реалізації відповідних технічних та організаційних заходів в комплексі, що супроводжуються розробкою відповідної документації [2,6,8]. Результати аналізу проведеного дослідження вказують на необхідність вирішення наступних задач для забезпечення інформаційної безпеки: формування основ для опису процесів реалізації та виникнення атак, загроз,

вразливостей інформаційної безпеки системи в умовах невизначеності та непередбачуваності їх прояву; розробка відповідних засобів забезпечення захисту конфіденційної інформації на основі проведеного дослідження та класифікації вразливостей, загроз; визначення загальних підходів до створення інформаційних систем забезпечення захисту конфіденційних даних, механізмів управління захистом на різних рівнях діяльності суспільства [2,7,11]. Одним із підходів вирішення наведених задач є застосування існуючих систем виявлення комп'ютерних атак, для захисту інформації. В аналітичних оглядах компаній, у сфері інтернет-технологій та захисту інформації наводяться висновки, що в останні роки на інформаційно-пошукові системи, про зростання кількості загроз, а також трансформації засобів, які використовуються нелигітимними кореспондентами, у повноцінну інформаційну зброю [3,4,6,11,13]. Більшість сучасних програмно-апаратних систем виявлення комп'ютерних загроз та атак працюють із використанням підходів сигнатурного аналізу та фіксації інтернет-мережевих аномалій. Дані підходи мають недоліки, пов'язані із використанням потужних обчислювальних ресурсів на їх реалізацію, а також, при цьому, при виявленні нових комп'ютерних загроз мають низьку ефективність [6,10].

Основними джерелами надходження знань про вразливості та атаки інформаційної безпеки є бази даних та знань, створювані державними, українськими та зарубіжними комерційними структурами. Наповнення інформаційних баз даних здійснюється із залученням дослідних авторитетних центрів експертним шляхом. Разом з тим, інформація, що міститься в базах даних та знань вразливостей та загроз не є повною. Актуальним залишається задача виявлення доступних інформаційних ресурсів, про комп'ютерні загрози, віруси, вразливості, а також можливість доступу до результатів досліджень компаній з виявлення загроз інформаційної безпеки систем протидії [3,8,10,13].

Одним із джерел надходження інформації про вразливості та загрози інформаційної безпеки є інтернет-ресурси (інформаційні соціальні ресурси, також анонімні, які відносяться до інформаційної безпеки), обумовлено популярністю спеціалізованих інтернет-ресурсів, хто цікавиться відповідними предметними областями.

Основна складність виявлення та протидії поширенню шкідливої інформації в соціальних мережах безпосередньо слідує із використанням на сучасному етапі тенденцій розвитку інформаційно - технологічної сфери, а саме: збільшення швидкості поширення шкідливої інформації в соціальних мережах; швидкості виникнення нових джерел поширення шкідливої інформації; збільшення об’єму інформації, що містить шкідливі повідомлення; швидкості тиражування повідомлень в мережі; кількості сценаріїв привернення уваги аудиторії; рівня гетерогенності даних. Таким чином, розглянуті тенденції поширення шкідливої інформації в Інтернет мережах, зумовлюють необхідність підвищення ефективності протидії та виявлення в соціальних мережах шкідливої інформації, враховуючи також при цьому, обґрунтованість та оперативність [1,3,5,7,14].

За своєю архітектурою соціальні мережі є багатокомпонентними рішеннями, в архітектурі мережі знаходяться: компоненти, які здійснюють обробку контенту; компоненти, які забезпечують функції маркетингу, адміністрування, зберігання даних. Соціальні мережі не містять окремого компонента виявлення та протидії поширенню шкідливій інформації [4].

Протидія поширенню шкідливої інформації у соціальних мережах є важливим елементом інформаційної безпеки особистості, суспільства, держави, проте більшість систем, на теперішній час не враховує простір функціональності системи виявлення та протидії поширенню шкідливою інформації. При розробці методу протидії поширенню шкідливої інформації в Інтернет мережі, необхідно: в повній мірі, враховувати кількість повідомлень на сторінці, характеристики джерела, зворотній зв'язок від джерела та аудиторії повідомлень; підтримувати дві стадії роботи системи протидії: налаштування, експлуатація; ранжувати контрзаходи з урахуванням коефіцієнтів складності [2,5,10,13].

Задача підвищення ефективності методів виявлення нових вразливостей та загроз конфіденційним даним інформаційних систем на основі розробки комплексів програм та алгоритмів є актуальною, дозволить здійснювати аналіз та виявлення інформаційних джерел, які містять інформацію про вразливості, шкідливе програмне забезпечення, комп'ютерні атаки. Обґрунтована можливість проведення аналізу тематичних інтернет-ресурсів як джерела виявлення вразливостей та загроз інформаційній безпеці [6,9,11].

**Алгоритм фільтрації потоку тематичних повідомлень та статистичного аналізу поширенню шкідливої інформації в соціальних мережах**

Проведений порівняльний аналіз досліджень в області протидії та виявлення шкідливої інформації в соціальних мережах дозволив визначити загальні вимоги до системи протидії, в основу реалізації, покладено модельно-методичний аппарат [2,4,10,11,13].

На підставі описаних особливостей функціонування тематичних інтернет-ресурсів, методів семантичної фільтрації текстових повідомлень та послідовність проведення аналізу створюваних учасниками форуму повідомлень в період проведення аналізу тематичних повідомлень може бути представлена наведеним алгоритмом на рис. 1.

Запропонований алгоритм фільтрації потоку повідомлень та статистичного аналізу передбачає фільтрацію тематичних повідомлень, що не відносяться до заданої предметної області, яка задана відповідною онтологією, а також підрахунок кількості текстових повідомлень, що пройшли етап фільтрації потоку даних, та визначення середнього рейтингу авторів текстових повідомлень.

Вхідними параметрами алгоритму фільтрації потоку повідомлень та статистичного аналізу інформаційної безпеки є: $D_\tau$ – множина текстових повідомлень тематичних інтернет-ресурсів, створених в період проведення аналізу потоку даних; $O$ – онтологія предметної області вразливостей та загроз інформаційної безпеки конфіденційних даних.

Основні кроки алгоритму проведення аналізу потоку текстових повідомлень наступні:

1. Обнулення значень $K_\tau$ – кількості тематичних повідомлень про вразливості та загрози інформаційної безпеки конфіденційних даних та $A_\tau$ – середнього рейтингу авторів тематичних повідомлень створених у період часу проведення аналізу $\tau$;

2. Обчислення для кожного текстового повідомлення коефіцієнта $k_{Ont}$ – близькості до термінів предметної області $O$ заданої онтології;

3. Додавання тематичних повідомлень множини $D_\tau$, для яких виконується нерівність $k_{Ont} > 0$, до бази даних прецедентів для їх подальшого використання для формування відповідних звітів прогнозування вразливостей та загроз інформаційної безпеки конфіденційних даних;

4. Обчислення $K_\tau$ – кількості повідомлень множини $D_\tau$, для яких виконується нерівність $k_{Ont} > 0$;

5. Обчислення $A_\tau$ – середнього рейтингу авторів тематичних повідомлень множини $D_\tau$, для яких виконується нерівність $k_{Ont} > 0$.

Результатом роботи запропонованого алгоритму аналізу потоку текстових повідомлень є визначення статистичних показників, що характеризують потік тематичних повідомлень в період проведення аналізу потоку даних: $K_\tau$ – кількість текстових повідомлень, що містять терміни вразливостей та загроз інформаційної безпеки з онтології конфіденційним даним; $A_\tau$ – середній рейтинг авторів тематичних повідомлень, що містять терміни вразливостей та загроз інформаційної безпеки з онтології конфіденційним даним; поповнення бази даних прецедентів текстовими повідомленнями, що містять терміни вразливостей та загроз інформаційної безпеки з онтології конфіденційним даним.

Отриманні результати застосування запропонованого алгоритму аналізу потоку текстових повідомлень можуть бути використані як значення вхідних параметрів у системі логічного нечіткого виводу та при формуванні звітів прогнозування вразливостей та загроз інформаційній безпеці організації.

Запропоноаваний алгоритм аналізу тематичних інтернет-ресурсів потоку текстових повідомлень дозволяє обчислювати статистичні параметри, здійснювати семантичну фільтрацію текстових повідомлень, а також результати алгоритму можуть бути використанні для побудови системи логічного нечіткого виводу для прогнозування подій предметної області для якої проводиться аналіз.

**Метод прогнозування вразливостей інформаційної безпеки на основі аналізу даних тематичних інтернет-ресурсів**

Результати проведеного дослідження вказують на важливість при забезпеченні захисту інформаційно - обчислювальних систем задачі щодо підтримки в актуальному стані моделі загроз інформаційної безпеки конфіденційних даних [2]. Фахівцю, який забезпечує безпеку даних інформаційно - обчислювальних систем, необхідно своєчасно приймати адекватні рішення щодо необхідності перегляду моделі інформаційної безпеки конфіденційних даних, у разі виявлення вразливостей або виникнення загроз. Існуючі, на теперішній час, методи підтримки прийняття рішень надають наступні можливості: визначати критеріальні оцінки параметрів та ранжувати критерії, значущими для заданої задачі (надає можливість оцінити надані варіанти рішень); формалізувати процес на основі наявних даних знаходження рішення (процес генерації варіантів розв'язку); використовувати формальні процедури прийнятих рішень прогнозування наслідків; використовувати під час прийняття колективних рішень формалізовані процедури узгодження; вибирати кращий варіант, що призводить до розв'язання поставленої задачі [3,6,8,10,11,13]. Основні задачі, які вирішуються методами підтримки прийняття рішення: вибір альтернативи; генерація варіантів рішення (альтернатив). Критерій підтримки прийняття рішень - функція, що виражає переваги особи, яка приймає відповідне рішення, визнає правило, за яким вибирається оптимальний чи прийнятний варіант рішення задачі. Існує безліч критеріїв підтримки прийняття рішень, що використовуються в залежності від умов поставленої задачі [4,8,12]. Нечіткі множини застосовуються при вирішенні поставленої задачі при необхідності описувати нечіткі знання та поняття, а також в подальшому проводити операції з цими знаннями і поняттями та формувати нечіткі виводи.

Обґрунтованість застосування нечітких моделей при вирішенні поставленої задачі пов'язана зі значним ступенем присутності невизначеності, по причині складності предметної області та неповноти наданої інформації, а також наявністю відповідних відомостей про систему якісного характеру [8-10]. Основною перевагою використання нечітких систем є їх універсальність, будь-яку безперервну функцію можна представити із заданою точністю нечіткою моделлю. Інформаційні системи, що побудовані на логічній нечіткій логіці, дозволяють синтезувати модель об'єкта предметної області, на основі евристичної інформації, а також інформації отриманої експертним шляхом або в результаті проведення експерименту. До недоліків логічних нечітких систем відносять низьку швидкість при великій кількості керуючих правил їх роботи, відсутність, на теперішній час, алгоритмів, що дозволяють здійснювати синтез стійких моделей [3-12].

Побудова логічних нечітких систем при вирішенні певних відповідних задач, на відміну від використання класичних методів, нерідко передбачає введення суб'єктивного характеру додаткових аксіом. У зв'язку з цим, процес створення логічних нечітких моделей притаманні елементи творчості [8, 12,13]. Як правило, методи логічного нечіткого виводу застосовуються для вирішення задач, пов'язаних, насамперед з апроксимацією функцій, класифікацією та розпізнаванням образів, управлінням та моделюванням нелінійними об'єктами, прийняття адекватних рішень в умовах невизначеності [12, 13].

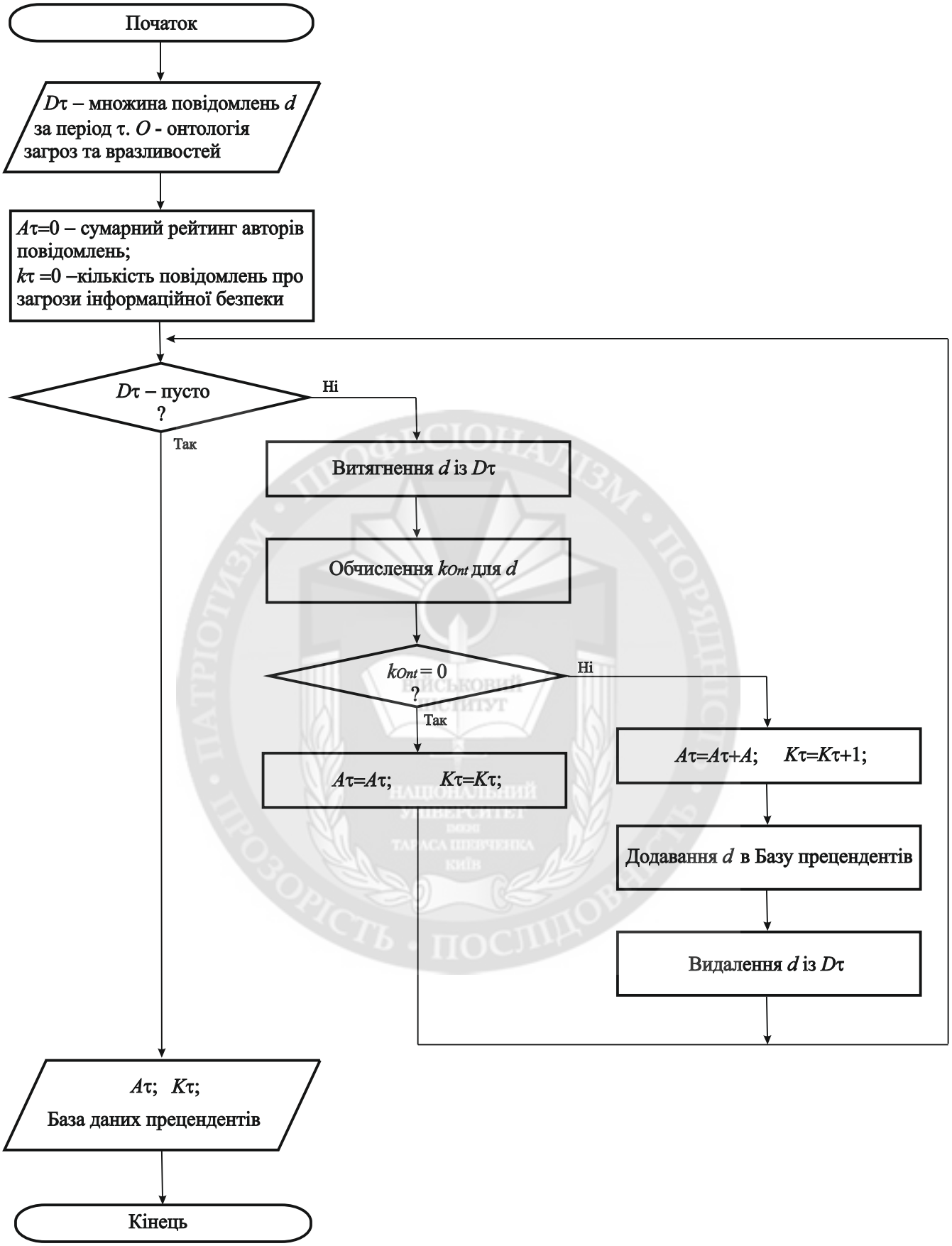


Рисунок 1 . Алгоритм фільтрації потоку тематичних повідомлень та статистичного аналізу

Центральне місце в системах логічного нечіткого моделювання займає нечіткий вивід, який є відповідною процедурою або алгоритмом для отримання логічних нечітких виводів, ґрунтуючись на застосуванні операцій нечіткої логіки та нечітких передумов.

У загальному вигляді структура системи логічного нечіткого виводу та послідовність реалізованих системою етапів представлена на рис. 2.

Продукційне правило для системи нечітко логічного виводу, відповідно до існуючих на теперішній час методик побудови бази знань правил логічної нечіткої системи, представляється наступним чином (1):

$$ЯКЩО\left(u_1 \in A_1\right) I \ldots I \left(u_n \in A_n\right) ТО \left(y \in Q_j\right), \quad (1)$$

де $u_1,\ldots,u_n$ – нечіткі змінні логічної нечіткої системи з $n$ входами; $A_1,\ldots,A_n$ – нечіткі множини, що відповідають нечітким змінним $u_1,\ldots,u_n$; $y$ – нечітка вихідна логічна змінна; $Q_j$ – нечітка множина, що відповідає нечіткій логічній змінній $y$.


Рисунок 2. Структура системи нечіткого логічного виводу

Фазифікацією вхідних змінних називається - процес перетворення чітких значень вхідних параметрів у відповідні їм нечіткі множини. В залежності від виду функцій приналежності, реалізуються процеси фазифікації наступні: гаусова, трикутна, одноелементна [12]. В результаті одноелементної, наприклад, фазифікації чіткого числа $u_i$ для $i$ – го входу нечіткої системи створюється $\tilde{A}_i$ – нечітка множина з функцією приналежності «Сінглтон»:

$$\mu_{\tilde{A}_i}(x) = \begin{cases} 1, & x = u_i \\ 0, & інакше \end{cases} \quad (2)$$

Значення коефіцієнтів ступенів приналежності підумов нечітких логічних продукцій обчислюються як результат перетину нечітких множин системи $\tilde{A}_i$, отриманих шляхом фазифікації вхідних параметрів $u_i$ та нечітких множин системи $A_i$ із відповідних правил бази знань нечітких продукцій. При перетині нечітких множин системи застосовується, так звана Т-норма. Її часним випадком є операція отримання значення мінімуму (3):

$$\tilde{A}_i\left(u_i\right) = \tilde{A}_i\left(u_i\right) \wedge A_i\left(u_i\right), \quad (3)$$

де $A_i$ - нечітка множина системи, яка визначена для $i$ - ї підумови деякого продукційного правила (1); $\tilde{A}_i$ – нечітка множина системи, яка отримана в результаті фазифікації чіткого

значення змінної для *i* - го входу системи; $\tilde{A}_i$ – нечітка множина логічної системи, що відповідає *i* – й умові деякого продукційного правила бази даних.

Процедури агрегування умов, акумулювання та активізації підзаключень правил нечітких продукцій логічної системи, а також операція дефазифікації залежить від вибору відповідного алгоритму нечіткого логічного виводу [12].

На теперішній час найбільш затребувані алгоритми нечіткого логічного виводу Ларсена, Такагі-Сугено, Цукамото та Мамдані. Найбільшою популярністю при вирішенні прикладних завдач використуються алгоритми Мамдані і Такаги-Сугено [12].

Аналіз проведених досліджень оцінки ефективності наведених алгоритмів нечіткого логічного виводу показав, що їх застосування залежить від специфіки задачі, яку необхідно вирішувати з їх використанням. Застосування алгоритму Мамдані дозволяє, при цьому, уникнути великих об'ємів обчислювальних операцій. З урахуванням даної особливості пов'язана його популярність при вирішенні практичних задач нечіткого логічного моделювання.

Таким чином, беручи до уваги, що алгоритм Мамдані для вирішення нечіткої задачі використовує менші обчислювальні ресурси і реалізації нечіткого виводу, то для вирішення задачі прогнозування вразливостей та загроз інформаційної безпеки конфіденційних даних обрано алгоритм Мамдані.

Метод прогнозування вразливостей та загроз інформаційної безпеки включає наступні етапи:

1. Етап формування правил логічних нечітких продукцій у вигляді :

$$ЯКЩО\left(u_1 \in A_1\right) I \ldots I\left(u_n \in A_n\right) TO\left(y \in Q_j\right)$$

2. Етап фазифікації вхідних параметрів за формулою:

$$\mu_{\tilde{A}_i}(x) = \begin{cases} 1, & x = u_i \\ 0, & інакше \end{cases}$$

3. Етап обчислення коефіцієнтів ступенів приналежності підумов відповідно до правил логічних нечітких продукцій за формулою:

$$\tilde{A}_i\left(u_i\right) = \tilde{A}_i\left(u_i\right) \wedge A_i\left(u_i\right),$$

4. Етап агрегування умов, відповідно до правил нечітких логічних продукцій. Визначення значень коефіцієнтів ступенів приналежності передумов кожного продукційного правила. При перетині нечітких логічних множин використовується метод Т-норма, часним випадком є операція мінімуму:

$$a_j = \tilde{A}_1\left(u_1\right) \wedge \tilde{A}_2\left(u_2\right) \wedge \ldots \wedge \tilde{A}_n\left(u_n\right),$$

де $a_j$ - ступінь приналежності передумови для *j* - го правила; $\tilde{A}_1\left(u_1\right) \wedge \tilde{A}_2\left(u_2\right) \wedge \ldots \wedge \tilde{A}_n\left(u_n\right)$ - нечіткі логічні множини для *n* підумов *j*-го правила. При цьому використовуються і вважаються активними для подальших розрахунків ті продукційні правила, для яких значення коефіцієнтів ступенів приналежності передумов не є нулем.

5. Етап активізації нечітких виводів у правилах логічних нечітких продукцій. Здійснюється, дана операція, із застосуванням операції мінімуму. Для вихідних параметрів визначаються «усічені» функції приналежності, розглядаються лише активні правила логічних нечітких продукцій.

$$\overline{Q}_i\left(y\right) = a_j \wedge Q_i\left(y\right),$$

де $a_j$ - значення коефіцієнта ступеня приналежності передумови $j$ - го правила продукції, $Q_i(y)$ - нечітка множина виводів $j$ - го продукційного правила, $\overline{Q}_i(y)$ - «усічена» нечітка множина виводів $j$ - го продукційного правила.

6. Етап акумуляції виводів правил логічних нечітких продукцій. Здійснюється об'єднанням знайдених «усічених» логічних функцій приналежності та отриманням для вихідного параметру підсумкової логічної нечіткої множини. Для об'єднання логічних нечітких множин застосовується метод S-норма, окремим випадком застосування якого є операція максимуму:

$$\overline{Q}(y) = \overline{Q}_1(y) \vee \overline{Q}_2(y) \vee ... \vee \overline{Q}_j,$$

де $\overline{Q}(y)$ – логічна нечітка множина, що відповідає результату роботи логічної нечіткої системи; $\overline{Q}_1(y) \vee \overline{Q}_2(y) \vee ... \vee \overline{Q}_j$ – «усічені» нечіткі логічні множини, що відповідають виводам продукційним активним правилам.

7. Етап дефазифікації. Отриманий нечіткий результат логічного виводу приводиться до чіткого представлення, із застосуванням методу центра ваги.

$$y = \frac{\sum_{j=1}^{R} b_j \int \mu_{\overline{Q}_j}(y) dy}{\sum_{j=1}^{R} \int \mu_{\overline{Q}_j}(y) dy}, \tag{4}$$

де $y$ - чітке значення результату виходу логічної нечіткої системи; $b_j$ - центри функцій приналежності відповідних термів онтології вихідної нечіткої змінної $y$ для $j$ - го правила продукції; $R$ – кількість правил логічних нечітких продукцій; $\int \mu_{\overline{Q}_j}(y) dy$ – величина площі під усіченою нечіткою множиною $\overline{Q}_j$ для $j$-го правила продукції.

Для прискоренного проведення обчислень застосовується дискретна форма:

$$y = \frac{\sum_{j=1}^{R} a_j b_j}{\sum_{j=1}^{R} a_j} \tag{5}$$

При побудові системи логічного нечіткого виводу виникнення вразливостей та загроз інформаційної безпеки конфіденційних даних на основі проведеного аналізу потоку текстових повідомлень тематичних інтернет-ресурсів, як вхідними параметрами можуть виступати наступні показники статистичного аналізу - середній рейтинг авторів текстових повідомлень, створених у період часу проведення аналізу, частота виникнення нових текстових повідомлень, що містять терміни вразливостей та загроз. Частота появи текстових повідомлень вимірюється в одиницях на добу, середній рівень рейтингу авторів тематичних повідомлень – в одиницях, ймовірність виникнення вразливості чи загрози – у відсотках.

Запропонована база знань правил нечітких продукцій та функції приналежності для системи логічного нечіткого вививоду про виникнення вразливостей та загроз інформаційної безпеки конфіденційних даних, ґрунтується на проведенні аналізу потоку даних тематичних форумів інтернет-ресурсів, відрізняється від наявних, можливістю адаптивного опису закономірностей процесу наповнення інтернет - форумів новими текстовими повідомленнями, шляхом застосування додаткових вхідних параметрів системи логічного нечіткого виводу та модифікації функцій приналежності, що дозволяє покращити якість прогнозування можливих вразливостей та загроз.

Запропоновано алгоритм проведення аналізу потоку текстових повідомлень тематичних форумів інтернет-ресурсів, що відрізняється від наявних, можливістю здійснювати: обчислення статистичних параметрів, семантичну фільтрацію текстових повідомлень для побудови системи логічного нечіткого виводу для прогнозування подій під час проведення дослідження заданої предметної області.

**Висновки**. Вирішена задача, полягає в підвищенні ефективності засобів та методів виявлення вразливостей та загроз інформаційної безпеки конфіденційних даних на основі розробки інформаційно-аналітичної системи та алгоритмів для проведення дослідження потоку повідомлень тематичних форумів інтернет-ресурсів.

Метод прогнозування вразливостей та загроз безпеки інформації, заснований на логічному нечіткому виводу, семантичному та статистичному аналізі, відрізняється від аналогів можливістю виявлення вразливостей та загроз до їх безпосередньої реалізації, а також гнучко дозволяє описувати закономірності процесу наповнення тематичних форумів інтернет-ресурсів новими текстовими повідомленнями, що в результаті сприяє покращенню якості прогнозування загроз. Алгоритм проведення дослідження потоку текстових повідомлень тематичних форумів інтернет-ресурсів, заснований на семантичному та статистичному аналізі, відрізняється від наявних можливістю обчислювати статистичні параметри, здійснювати семантичну фільтрацію текстових повідомлень, для прогнозування подій системи нечіткого логічного виводу.

Реалізований в інформаційно-аналітичній системі метод прогнозування вразливостей та загроз безпеки інформації на основі дослідження потоку даних тематичних ресурсів дозволяє автоматизувати інформаційний процес виявлення нових вразливостей, загроз, надає фахівцям інформаційної безпеки можливість оцінити своєчасно ступінь захищеності ресурсів та при необхідності вжити відповідних заходів щодо нейтралізації можливих загроз та вразливостей, тим самим підвищити інформаційну безпеку обчислювальних комп'ютерних систем від реалізації нових мережевих комп'ютерних атак.

ЛІТЕРАТУРА:

1. Ленков, С.В. Модель безпеки поширення забороненої інформації в інформаційно-телекомунікаційних мережах / С.В. Ленков, В.М. Джулій, В.С. Орленко, О.В. Сєлюков, А.В. Атаманюк // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2020. – Вип. №68. – С. 53-64.

2. Джулій, В.М. Модель потоку текстових повідомлень тематичних інтернет-ресурсів системи прогнозування інформаційної безпеки / В. Джулій, Н. Петляк, Ю. Хмельницький, О. Пахар // Вісник Хмельницького національного університету. Технічні науки. – 2022. – № 5. – С. 294-300.

3. Lienkov, S., Podlipaiev, V., Tolok, I., Lisitsky I., .Lytvynenko, N., Kuznichenko, S. The Information and Analytical Using of Non-Structured Information Resources CEUR Workshop Proceedingsthis link is disabled, 2021, 3126, стр. 81–87.

4. Соціальні мережі – реальні загрози віртуального світу. [Електронний ресурс]. – Режим доступу : http://ogo.ua/ articles/view/011–02–23/26490.htm.

5. Ленков, С.В. Методы и средства защиты информации. В 2-х томах /С.В. Ленков, Д.А. Перегудов, В.А. Хорошко –К: Арий, 2008.–464с

6. Остапов С. Е. Технології захисту інформації: навчальний посібник / С.Е. Остапов, С.П. Євсеєв, О.Г. Король – Харків : Вид-во ХНЕУ, 2016. – 476 с.

7. Лєнков, С.В. Аналіз існуючих методів та алгоритмів виявлення атак в бездротових мережах передачі даних / С.В. Лєнков, В.М. Джулій, Н.М. Берназ, С.О. Божук // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2017. – Вип. № 56. – С.124-132

8. Джулій, В.М. Інформаційно-ознакова модель шкідливої інформації в соціальних мережах/ І.В. Муляр, В.М. Джулій, В. М. Пічура, О.О Зацепіна – Вимірювальна та обчислювальна техніка в технологічних процесах № 3 (2022)-73–78с.

9. Джулій, В.М. Модель потоку текстових повідомлень тематичних інтернет-ресурсів системи прогнозування інформаційної безпеки/ В.М. Джулій, Ю.В. Хмельницький, Н.С. Петляк, О.В. Пахар– Вісник Хмельницького національного університету. Технічні науки. 2022. № 5. С. 294-300с.

10. Джулій, В.М., Кльоц Ю.П., Муляр І.В., Жилевич М.Л., Джулій А.В. Контроль додатків інтернет-трафіка комп'ютерних мереж методами машинного навчання. Вісник Хмельницького національного університету. Технічні науки. 2021. № 5. С. 22-26.

11. Джулій, В.М. Метод класифікації додатків трафіка комп'ютерних мереж на основі машинного навчання в умовах невизначеності / В.М. Джулій, О.В. Мірошніченко, Л.В. Солодєєва // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2022. – Вип. №74. – С. 73-82.

12. Лавров, Є. А. Математичні методи дослідження операцій : підручник / Є. А. Лавров, Л. П. Перхун, В. В. Шендрик – Суми : Сумський державний університет, 2017. – 212 с.

13. Гончар С. Ф. Оцінювання ризиків кібербезпеки інформаційних систем об'єктів критичної інфраструктури : монографія. / С. Ф. Гончар. – Київ, 2019. – 175 с.

14. Yemchuk L. Organizational Network Analysis as a Tool for Leadership Assessment in Software Development Team. Zhylinska O.; Chornyi A.; Dzhuliy V. – Institute of Electrical and Electronics Engineers (30 September 2020); INSPEC Accession Number: 20008165; DOI: 10.1109/ACIT49673.2020.

REFERENCES:

1. Lenkov, S.V. (2020), Model bezpeky poshyrennia zaboronenoi informatsii v informatsiino-telekomunikatsiinykh merezhakh / S.V. Lenkov, V.M. Dzhulii, V.S. ORLENKO, O.V. Sieliukov, A.V. Atamaniuk // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU. – №68. – pp. 53-64.

2. Dzhulii, V.M. (2022.), Model potoku tekstovykh povidomlen tematychnykh internet-resursiv systemy prohnozuvannia informatsiinoi bezpeky / V. Dzhulii, N. Petliak, Yu. Khmelnytskyi, O. Pakhar // Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. – 2022. – № 5. – pp. 294-300.

3. Lienkov, S., Podlipaiev, V., Tolok, I., Lisitsky I., .Lytvynenko, N., Kuznichenko, S. The Information and Analytical Using of Non-Structured Information Resources CEUR Workshop Proceedingsthis link is disabled, 2021, 3126, стр. 81–87.

4. Cotsialni merezhi – realni zahrozy virtualnoho svitu. [Elektronnyi resurs]. – Rezhym dostupu : http://ogo.ua/ articles/view/011– 02–23/26490.htm

5. Lenkov, S.V. (2008), Metodyy sredstva zashchyты ynformatsyy. V 2-kh tomakh / S.V. Lenkov, D.A. Perehudov, V.A. Khoroshko –K: Aryi–464s.

6. Ostapov, S. E. (2016) Tekhnolohii zakhystu informatsii: navchalnyi posibnyk / S.E. Ostapov, S.P. Yevseiev, O.H. Korol–Kharkiv : Vyd-vo KhNEU. – 476 s.

7. Lenkov, S.V. (2017), AnalIz Isnuyuchih metodiv ta algoritmiv viyavlennya atak v bezdrotovih merezhah peredachI danih / S.V. Lenkov, V.M. Dzhuliy, N.M. Bernaz, S.O. Bozhuk // Zbirnik naukovih prats Viyskovogo Institutu Kiyivskogo natsionalnogo universitetu imeni Tarasa Shevchenka. – K.: VIKNU. – Vip. No 56. – p.124-132

8. Dzhulii, V.M. Informatsiino-oznakova model shkidlyvoi informatsii v sotsialnykh merezhakh/ I.V. Muliar, V.M. Dzhulii, V. M. Pichura, O.O Zatsepina – Vymiriuvalna ta obchysliuvalna tekhnika v tekhnolohichnykh protsesakh № 3 (2022)-73–78s.

9. Dzhulii, V.M. Model potoku tekstovykh povidomlen tematychnykh internet-resursiv systemy prohnozuvannia informatsiinoi bezpeky/ V.M. Dzhulii, Yu.V. Khmelnytskyi, N.S. Petliak, O.V. Pakhar– Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. 2022. № 5. S. 294-300s.

10. Dzhulii V.M., Klots Yu.P., Muliar I.V., Zhylevych M.L., Dzhulii A.V. (2021), Kontrol dodatkiv internet-trafika kompiuternykh merezh metodamy mashynnoho navchannia. Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. – Khmelnytskyi. – №5. – pp. 22–26.

11. Dzhulii, V.M. (2022), Metod klasyfikatsii dodatkiv trafika kompiuternykh merezh na osnovi mashynnoho navchannia v umovakh nevyznachenosti / V.M. Dzhulii, O.V. Miroshnichenko, L.V. Solodieieva // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU. – Vyp. №74. – pp. 73-82.

12. Lavrov, Ye. A. (2017.), Matematychni metody doslidzhennia operatsii : pidruchnyk / Ye. A. Lavrov, L. P. Perkhun, V. V. Shendryk – Sumy : Sumskyi derzhavnyi universytet, – 212 p

13. Гончар С. Ф. Оцінювання ризиків кібербезпеки інформаційних систем об'єктів критичної інфраструктури : монографія. / С. Ф. Гончар. – Київ, 2019. – 175 с.

14. Yemchuk L. Organizational Network Analysis as a Tool for Leadership Assessment in Software Development Team. Zhylinska O.; Chornyi A.; Dzhuliy V. – Institute of Electrical and Electronics Engineers (30 September 2020); INSPEC Accession Number: 20008165; DOI: 10.1109/ACIT49673.2020.

D.Sci. of Techn., prof. **Lienkov S.V.,** Ph.D. **Dzhuliy V.M.,** Ph.D. **Bernaz A.M.,** PhD. **Muliar' I.V.,** PhD **Pampukha I.V.**

## METHOD OF FORECASTING INFORMATION SECURITY VULNERABILITIES BASED ON DATA ANALYSIS OF THEMATIC INTERNET RESOURCES

***In the paper, a study of the task of predicting information security vulnerabilities is carried out based on the analysis of the data of thematic Internet resources. Against the backdrop of the rapid development of information technology, there has been an increase in the activity of a variety of computer attacks carried out and planned using modern latest technologies. Harmful information has become an obvious problem in the information security of society today, it should also be noted that criminal and terrorist groups are increasingly adopting means of information influence, developing and writing strategies aimed at attracting new adherents and expanding the sphere of influence through social networks. The analysis of the conducted research of the current state in the field of information security shows that the pace of development of information and computer technologies is significantly ahead of the process of creating software and hardware in the field of information security. The priority in this situation is the task of analysis, classification, identification of active mechanisms and means of attacks and threats to the information security of the system, which can lead to unauthorized access to confidential data, disruption of the functioning of the information system, determination of countermeasures against attacks and threats, assessment of the given damage, development of the legal framework, protection mechanisms and information security criteria of the countermeasure system. Today, there is no single approach to solving the problem of security of information and search systems, in relation to subject areas: developers of hardware and software protection of information offer appropriate components for solving specific problems; ensuring reliable protection of information resources requires the implementation of appropriate technical and organizational measures in a complex, accompanied by the development of appropriate documentation. Most of the modern software and hardware systems for detecting computer threats and attacks work using the approaches of signature analysis and fixing of Internet network anomalies. These approaches have disadvantages associated with the use of powerful computing resources for their implementation, and have low efficiency when detecting new computer threats. The method of predicting information security vulnerabilities based on data from Internet resources, based on fuzzy inference, semantic and statistical analysis, is distinguished by the ability to identify vulnerabilities and threats to their implementation, allows you to describe the patterns of the information process of filling thematic resources with new text messages, which affects the quality of forecasting . The algorithm for forecasting vulnerabilities and threats to information security implemented in the information and analytical system, based on the analysis of the data flow of thematic resources, allows automating the information process of detecting new vulnerabilities and threats, provides information security specialists with the opportunity to assess the degree of security of resources in a timely manner and, if necessary, take appropriate measures to neutralize possible threats and vulnerabilities, thereby increasing the information security of computing computer systems against the implementation of new network computer attacks.***

***Keywords: information security, thematic Internet resources, social networks, sources of messages, vulnerabilities, attacks, information system.***