Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра комп'ютерних систем та мереж
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

Магістр
(назва освітнього ступеня)

на тему: **Методи і засоби захисту серверів електронної пошти від спаму**

Виконав: студент(ка) 6 курсу, групи СІм-61
спеціальності 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

(підпис) Ільченко Д. О. (прізвище та ініціали)

Керівник (підпис) Жаровський Р.О. (прізвище та ініціали)

Нормоконтроль (підпис) Луцик Н.С. (прізвище та ініціали)

Завідувач кафедри (підпис) Осухівська Г.М. (прізвище та ініціали)

Рецензент (підпис) Мацюк О.В. (прізвище та ініціали)

Тернопіль
2021

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних систем та мереж
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
Осухівська Г.М.
(підпис) (прізвище та ініціали)
« » 2021 р.

З А В Д А Н Н Я

НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня *магістр*
(назва освітнього ступеня)

за спеціальністю *123 «Комп'ютерна інженерія»*
(шифр і назва спеціальності)

студенту *Ільченку Дмитру Олександровичу*
(прізвище, ім'я, по батькові)

1. Тема роботи *Методи і засоби захисту серверів електронної пошти від спаму*

Керівник роботи *Жаровський Руслан Олегович, к.т.н.*
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «*28*» *жовтня* *2021* року № *4.7-916*

2. Термін подання студентом завершеної роботи *.12.2021 р.*

3. Вихідні дані до роботи *Характеристики електронних повідомлень, протоколи передачі електронних повідомлень, технічні ознаки спаму, алгоритми фільтрації*

4. Зміст роботи (перелік питань, які потрібно розробити)
Вступ. 1. Аналіз проблем передачі електронних повідомлень
2. Алгоритми обробки і фільтрації електронних повідомлень
3. Інформаційне забезпечення системи контентної фільтрації електронної кореспонденції
4. Охорона праці та безпека в надзвичайних ситуаціях
Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
1. Актуальність і мета дослідження.
2. Задачі дослідження, об'єкт і предмет, наукова новизна і практична цінність дослідження.
3. Ознаки спаму
4. Алгоритми формальних методів фільтрації
5. Базова архітектура нейронної мережі ART
6.Архітектура системи фільтрації електронних повідомлень
7. Результати експериментального дослідження.
8. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності,			
основи охорони праці			

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	*Аналіз сучасних технічних проблем передачі електронної пошти*	*29.10.2021-5.11.2021*	*виконано*
2	*Визначення основних ознак спам повідомлень на базі яких можливо здійснювати їх розпізнавання*	*6.11.2021 – 15.11.2021*	*виконано*
3	*Розробка моделі електронного повідомлення у формі стійких словосполучень*	*16.11.2021 – 26.11.2021*	*виконано*
4	*Розробка програмного забезпечення фільтрації ЕП та її тестування*	*26.11.2021– 03.12.2021*	*виконано*
5	*Охорона праці та безпека в надзвичайних ситуаціях*	*04.12.2021-07.12.2021*	*виконано*
6	*Оформлення пояснювальної записки і графічного матеріалу*	*06.12.2021-09.12.2021*	*виконано*
7	*Попередній захист дипломної роботи магістра*	*15.12.2021*	*виконано*
8	*Захист дипломної роботи магістра*	*21.12.2021*	

Студент _____ *Ільченко Д. О.*
(підпис) (прізвище та ініціали)

Керівник роботи _____ *Жаровський Р.О.*
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Методи і засоби захисту серверів електронної пошти від спаму // Дипломна робота // Ільченко Дмитро Олександрович // ТНТУ, комп'ютерна інженерія, група СІм-61 // Тернопіль, 2021 // с. – 78, рис. – 32, табл. – 5, аркушів А1 – 8, додат. – 1 , бібліогр. – 21.

Ключові слова: спам, фільтрація, алгоритм, електронна пошта, Інтернет, автоматизація.

У дипломній роботі магістра проведений аналіз проблем передачі електронних повідомлень.

Досліджено методи і засоби побудови систем захисту серверів електронної пошти від поширення нелегітимної кореспонденції.

Розроблена векторна модель електронного повідомлення з виявленням стійких словосполучень.

Розроблено програмне забезпечення на основі поєднання формальних методів фільтрації електронних повідомлень та нейронної мережі ART.

Обґрунтовано застосування алгоритмів фільтрації електронних повідомлень на основі показників їх ефективності.

ABSTRACT

Methods and means of protection of e-mail servers from spam // Thesis // Ilchenko Dmitry Alexandrovich // TNTU, computer engineering, group CIm-61 // Ternopil, 2021 // p. - 78, fig. - 32, table. - 5, sheets A1 - 8, app. - 1, ref. - 21.

Keywords: spam, filtering, algorithm, e-mail, Internet, automation.

In the master's thesis the analysis of problems of transmission of electronic messages is carried out.

Methods and means of building systems for protecting e-mail servers from the spread of illegitimate correspondence have been studied.

A vector model of an electronic message with the detection of stable phrases has been developed.

Software based on a combination of formal methods of electronic message filtering and ART neural network has been developed.

The application of electronic message filtering algorithms based on their efficiency indicators is substantiated.

ЗМІСТ

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ І СКОРОЧЕНЬ

ІТ інформаційні технології

ЕП електронне повідомлення

ПЗ програмне забезпечення

RBL англ. Realtime Blackhole List чорний список реального часу

ART англ. Adaptive Resonance Theory нейронні мережі адаптивного резонансу

ВСТУП

Актуальність теми. Завдання фільтрації спаму в глобальних інформаційних мережах в даний час є дуже актуальним. Це пов’язано з бурхливим розвитком соціального спілкування з допомогою телекомунікаційних мереж. Спам-повідомлення різко зменшують корисну компоненту трафіку в мережах, призводять до перевантажень хостингових комп'ютерів, знижують доступність та цілісність інформації. Слабкий розвиток законодавчої бази щодо блокування спаму викликає необхідність розробки наукових методів аналізу спаму, а також впровадження інструментальних засобів виявлення та фільтрації небажаних повідомлень. Особливу увагу слід приділяти спаму в поштових повідомленнях, а також інтерактивних частинах сайтів мережі Інтернет.

ІТ-ринок пропонує різні засоби фільтрації вмісту інформаційного обміну каналами Інтернет. В даний час умовно виділяють три типи засобів, що забезпечують контроль використання Інтернет-ресурсів на корпоративному рівні:

– маршрутизатори, міжмережеві екрани, системи виявлення вторгнень, проксі-сервери тощо;

– антивірусне програмне забезпечення, які мають базові можливості контентної фільтрації;

– спеціалізовані засоби, розроблені безпосередньо для контролю використання інтернет – ресурсів, такі як системи моніторингу електронної пошти, засоби контролю веб-трафіку, антиспам – фільтри тощо.

Найбільш ефективними є багатокомпонентні антиспам системи, що поєднують у собі кілька методів детектування спаму. Починаючи з 2000-х років ведуться активні дослідження у галузі розробки програмних систем виявлення поштового спаму. Питання протидії спаму присвячені роботи Пола Греема (Paul Graham), Дж. Грем-Каммінга (John Graham-Cumming), Дж. Здзіарскі (Jonathan Zdziarski), У. Йеразуніса (William S. Yerazunis), Г. Робінсона (Gary Robinson), І. С. Ашманова та ін. Всі розроблені алгоритми та системи використовують базові принципи фільтрації

електронної пошти на основі класифікації текстової інформації, що міститься у листі.

Однак незважаючи на зусилля існуючих методів фільтрації спам повідомлення продовжують попадати в скриньки користувачів. У зв'язку з цим, розвиток багатокомпонентних антиспамових систем для фільтрації повідомлень в телекомунікаційних мережах є актуальним завданням та представляє науковий та практичний інтерес у галузі захисту інформації в комп’ютерних системах у частині забезпечення доступності та цілісності даних.

Мета кваліфікаційної роботи. Мета роботи полягає у підвищенні ефективності фільтрації спаму в електронних повідомленнях на основі алгоритмів класифікації текстової інформації в поєднанні з класичними методами обробки.

Задачі кваліфікаційної роботи:

1. Розробка концепції побудови системи фільтрації спаму в телекомунікаційних мережах.

2. Розробка багаторівневої архітектури системи фільтрації повідомлень, що включає:

- етап нормалізації повідомлення з його попередньою підготовкою для забезпечення фільтрації із застосуванням морфологічних прийомів;

- Модифікацію статистичних алгоритмів класифікації повідомлень на основі суміщення семантичного методу та класичних методів фільтрації.

3. Розробка моделі електронного поштового повідомлення, що враховує семантику контенту поштової кореспонденції.

4. Проведення тестування з оцінки продуктивності та експериментальне використання розробленої системи фільтрації повідомлень.

Відповідно до цілей та завдань дисертаційної роботи визначено її об'єкт та предмет.

Об’єкт дослідження: є процес фільтрації спаму в електронних повідомленнях.

Предмет дослідження: методи і засоби розпізнавання спаму на основі технічних характеристик і вмісту електронних повідомлень.

Методи дослідження: методи системного аналізу та дослідження операцій; теорія прийняття рішення; методи та засоби захисту інформації; теорія статистичних рішень.

Наукова новизна одержаних результатів:

1. семантичний метод фільтрації що використовує модифіковану нейронну мережу ART2a, в поєднанні з класичними методами фільтрації, що дало змогу підвищити ефективність обробки електронних повідомлень.

Практичне значення результатів кваліфікаційної роботи. Розроблено систему фільтрів електронної пошти, що дає можливість відфільтровувати близько 90% спаму.

Публікації. Результати дослідження апробовано на VIII науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (7-8 грудня 2021 року) у вигляді тез конференцій.

1. Ільченко Д.О, Жаровський Р.О. Методи фільтрації спаму в сучасних поштових системах. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі системи та технології» (7-8 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 110.

2. Ільченко Д.О, Жаровський Р.О. Семантичні методи фільтрації спаму. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі системи та технології» (7-8 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 111.

Структура роботи. До складу дипломної роботи магістра входить розрахунково-пояснювальна записка та графічний матеріал. Розрахунково-пояснювальна записка містить вступ, 4 розділи, загальні висновки, список використаної літератури і додатки. Обсяг роботи: розрахунково-пояснювальна записка – 78 арк. формату А4, графічна частина – 8 аркушів формату А1.

РОЗДІЛ 1
АНАЛІЗ ПРОБЛЕМ ПЕРЕДАЧІ ЕЛЕКТРОННИХ ПОВІДОМЛЕНЬ

Спам – це масове, анонімне, розсилання електронних повідомлень людям, які не висловили бажання її одержувати [20.]. Зазвичай спамом називають рекламні повідомлення. Але в спам повідомленнях можуть розсилатись і інші більш небезпечні речі, наприклад віруси. На даний час з розвитком соціальних мереж, різноманітних форумів і блогів, спам поширився і на них.

При розсиланні спаму адресу відправника зазвичай приховують або підмінюють для того щоб не було змоги виявити відправника. Крім того спам характеризується масовістю, тобто це не 10-100 повідомлень, а десятки тисяч.

1.1. Розвиток технічних методів розсилки і методів фільтрації спаму

Історично склалось так, що розвиток методів розповсюдження спаму напряму залежав від розвитку методів їх фільтрації.

Розсиланя спаму вручну

Вперше розсилку спам-реклами зробила компанія Digital Equipment Corporation в кінці 70-х років з повідомленням про вихід нової моделі комп’ютера. Кількість повідомлень була близько 400 проте на той час це було 15% усіх користувачів. Схожі розсилання можуть проводитись з реальної адреси яка чи які створені спеціально на поштових серверах.

Розсилання через некоректно налаштовані поштові сервери

В кінці 90-х років проблема спаму стала актуальною як ніколи. Так як для розсилки спаму почали використовувати недолік поштового сервера Sendmail який працював по замовчуванню як open relay (поштовий сервер, який дозволяє любому користувачеві відправити електронного листа на будь-яку адресу). З часом проблему усунули однак все таки залишилась можливість знайти open relay для відправки спаму. Пошук таких серверів здійснювали як спамери, так і анти-спамери. Одні для

розсилання спаму, інші для внесення таких поштових сервісів у базу даних чорного списку [21.].

Некоректно налаштований proxy - сервер

Пізніше для спам розсилки почали використовувати так звані socks і proxy сервери, до яких був можливий неавторизований доступ. Ці сервери призначені для збору трафіку організації і виходу через них в мережу Інтернет. Якщо сервер допускав неавторизоване з'єднання з любої IP адреси, то його використовували спамери для передачі SMTP трафіку. Також почали використовувати і відкриті http-proxy (зокрема порти 3128, 8080 і інші), які підтримують метод Connect. Для цього достатньо в команді Connect вказати ім'я сервера і порт для передачі поштового повідомлення. Як приклад таку ваду використовували в досить поширеному Web-сервері Apache.

На початку 2000 технології розсилки спаму отримали подальший істотний розвиток (для прикладу спам в 2008 році складав близько 90% трафіку електронної пошти). Оскільки спамери почали масово використовувати комп’ютерні віруси, зокрема поштові хробаки і троянські програми [15.], для створення мережі заражених комп’ютерів з яких в подальшому здійснювали масові розсилки спаму.

Створення спам-ботів.

Для цього використовують зазвичай два види вірусів це поштові хробаки або троянські програми. Використання поштових хробаків. Це тип вірусів, який зазвичай поширюється за допомогою електронної пошти. Після зараження комп'ютера вірус здійснює пошук на ньому поштових або IP адрес й розсилає себе за цими адресами.

Поштові хробаки автоматично підставляють випадкові адреси (зі знайдених на зараженому комп'ютері) у поле листа «From». У результаті сотні користувачів одержують повідомлення про те, що вони розсилають віруси і спам, хоча фактично у розсилках брали участь інфіковані комп'ютери, користувачі яких навіть не підозрювали про розсилку.

Троянські проксі.

Що стосується троянських вірусів то дані програми додатково могли здійснювати оновлення, маскуючись під легітимне ПЗ і автоматично копіювати себе на інші комп'ютери і т.д. Використання троянських програм дозволяє скачати список IP-адрес, адрес електронної пошти (провести збір інформації і оновити списки розсилки спамерів) і здійснити розсилку повідомлення.

Комп'ютери які заражені троянським вірусом або поштовим хробаком і які здійснюють розсилку спаму називають спам – ботами. Схема роботи спам – бота наведена на рисунку 1.1.

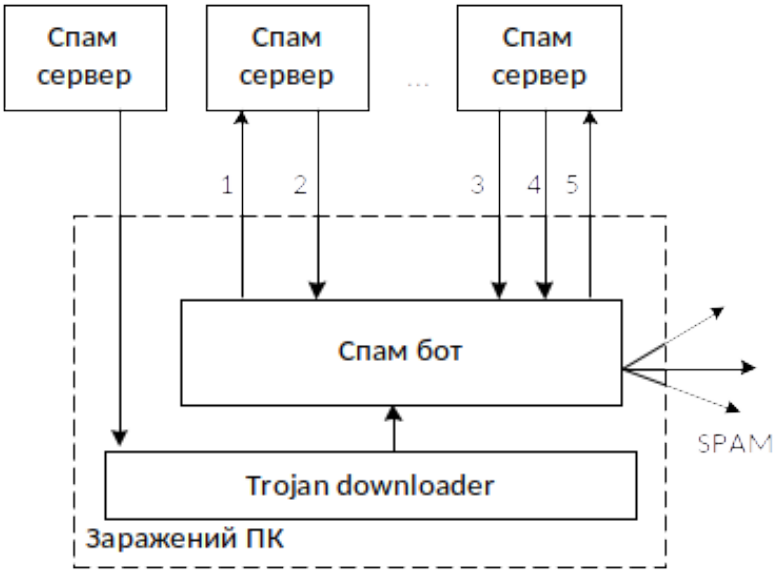

Рис. 1.1 Схема роботи спам-бота

Після встановлення спам-бот зв'язується з одним із серверів, що належать спамерам і передає звіт про встановлення(крок 1). Адреси цих серверів зазвичай задані у тілі спам-бота. У відповідь спам-бот отримує конфігурацію, в якій визначено перелік спам серверів з яких він буде отримувати дані для подальшого розсилання (крок 2). Після отримання цієї інформації він зв'язується із вказаним сервером і отримує список поштових адрес для розсилки, шаблони для формування листів та інші налаштування (крок 3-4). Після розсилки спам-бот надсилає звіт, в якому зазвичай міститься статистика та список адрес, за якими не вдалося надіслати пошту із зазначенням помилок. Для оновлення версій спам-бот на уражений

комп'ютер може встановлюватись Trojan-Downloader або спам-бот може мати функцією самооновлення.

З розвитком комп’ютерних мереж та масовим розповсюдженням мобільних засобів зв’язку кількість IP адрес з яких потенційно можлива розсилка спаму постійно зростає і класичні системи блокування спаму стають неефективними. Наприклад це RBL (Realtime Blackhole List) системи. Їхня робота базується на списках IP-адрес з яких раніше фіксували розсилання спаму, доступ до RBL здійснюється в режимі реального часу по DNS протоколу. Поштові сервери, що використовують RBL списки, в момент прийому чергового листа надсилають запит на RBL сервіс, який перевіряє IP-адресу відправника листа. Далі на підставі відповіді RBL сервісу, лист або приймають, або відкидають.

Однак є очевидний недолік - повідомлення приймається або відкидається лише на підставі IP-адреси відправника. Тому, якщо якийсь поштовий сервер попадає з якоїсь причини до чорного списку RBL, то пошта з цього сервера не буде прийматись всіма поштовими серверами, які використовують цей RBL. Відповідно разом спам повідомленнями легітимні повідомлення також будуть видалені або відфільтровані.

Недоліком RBL є невисока ефективність, неможливість самонавчання і висока ймовірність втрат легітимної пошти. Це сприяло розробці нових алгоритмів боротьби із спамом.

Не зважаючи на недоліки RBL-сервіси на даний час широко використовуються поштовими службами, Інтернет провайдерами та організаціями в яких є власні поштові сервіси. Так як цей вид фільтрації є самим простішим, незважаючи на його недоліки.

Для оцінки якості RBL використовується параметр який показує кількість спам листів, які проходять через поштовий сервер. Якщо в результаті роботи RBL кількість спаму зменшується, цей RBL-сервіс вважається "хорошим". Є й інша, характеристика - скільки легітимних листів було відфільтровано (проблема помилкових спрацювань).

Одним з таких методів є DNS-перевірка відповідності даних, що повідомляються в SMTP-сесії. Тобто при встановленні з'єднання надсилається команда EHLO аргументом якої є domain name. Фактично, клієнт може називати себе будь-яким іменем у командному рядку. Тому якщо сервера намагаються ідентифікувати клієнта, то вводиться механізм зворотного пошуку DNS, щоб визначити справжнє ім'я хоста клієнта відповідно до системи доменних імен за його IP-адресою. Як правило, з міркувань безпеки SMTP-сервери відмовляються підключатися до хостів, IP-адреса яких не перетворюється у відповідне ім'я хоста.

Інший метод це коли спам можна виявити за вмістом заголовків листів електронної пошти.

На даний час розроблено багато методів фільтрації спаму які ґрунтуються на його основних ознаках. Були розроблені алгоритми контекстної фільтрації електронної пошти і статистичних машинних методів аналізу текстів повідомлень, які на сьогоднішній день є найбільш ефективними.

Однак не зважаючи на всі зусилля розсилка спаму на сьогоднішній день має досить великий масштаб – щодня здійснюється розсилка мільярдів спам-повідомлень, що становить близько 40-70% від всього трафіку електронної пошти. Такі масштаби вимагають розробки нових заходів для боротьби з технологіями розсилок.

1.2. Ознаки спаму

Основні ознаки, які дозволяють віднести той або інший лист до категорії спам наведені на рисунку 1.2.

Також ознаками спаму можуть бути:

- відправка однакових повідомлень на велику кількість адрес
- наявність рекламного текстового повідомлення;
- багаторазове повторне надсилання листів;
- в тексті листа відсутні контакти осіб яким можна надати відповідь;

- простота і лаконічність тексту листа;
- підробка адреси отримувача.

Рис.1.2. Основні ознаки спаму

Як бачимо набір ознак досить великий. Наявність більшості перелічених ознак дозволяє віднести електронне повідомлення до категорії спам. Приклад такого електронного листа наведений на рисунку 1.3.

Рис.1.3. Приклад спам листа

Маємо класичний рекламний лист який містить лише одне посилання і коротенький рекламний текст. І має ряд технічних ознак наведених на рисунку 1.4.

Ідентифікатор повідомлення	<ZF3If1bGTw-Ct-Y9ADdP1acjDYQt7ufHg-89.163.255.108@ismtpd0003p1iad1.sendgrid.net>
Створено:	11 грудня 2021 р. о 20:58 (доставлено за 510 секунд)
Від:	PerfectGifts <yFTPrto-1acjDYQt7uf-noReply@evksx.southbeachmangofestival.com>
Кому:	Test_mail@aol.com
Тема:	Xmas 2021 gift guide
SPF:	PASS, IP-адреса 89.163.255.108 Докладніше

Рис.1.4 Заголовки спам листа

В даному листі вказана досить дивна адреса відправника PerfectGifts <yFTPrto-1acjDYQt7uf-noReply@evksx.southbeachmangofestival.com>. В полі «Кому» вказана адреса отримувача яка не відповідає моїй поштовій скриньці (поштовий сервер мав

бути gmail.com). Крім того якщо проглянути шлях пересилки листа можна бачити ознаки підміни адрес.

Як видно єдине що спамери не можуть – це зашифрувати текст листа чи вести якісь символи, оскільки повідомлення мають легко читатися. Структура, як бачимо з прикладу зазвичай проста текст і адреса переходу для послуги буде завжди автентичною. Розмір листа практично завжди невеликий. У випадку підробки адреси можливо використати функцію антиспуфінгу. Також завдяки схожим рекламним фразам в тексті листа можна застосувати машинні методи аналізу тексту для виявлення спаму.

1.3. Технологічні особливості розповсюдження спаму

Розвиток методів розсилки спаму привів до того, що спам-пошта має ряд технологічних ознак, а процес її створення та розповсюдження виглядає наступним чином (рисунок 1.5).

Рис.1.5. Основні етапи створення і розповсюдження спаму

Кожен з цих кроків може реалізовуватись незалежно один від одного, а також змінюватись в процесі розсилки. Розглянемо дані етапи більш детально.

1. Збір адрес отримувачів і їх верифікація.

Для розсилки спаму необхідно створити базу адрес потенційних одержувачів. Такі списки отримали назву E-Mail database. Адреси в них можуть мати додаткову інформацію (рис. 1.6):

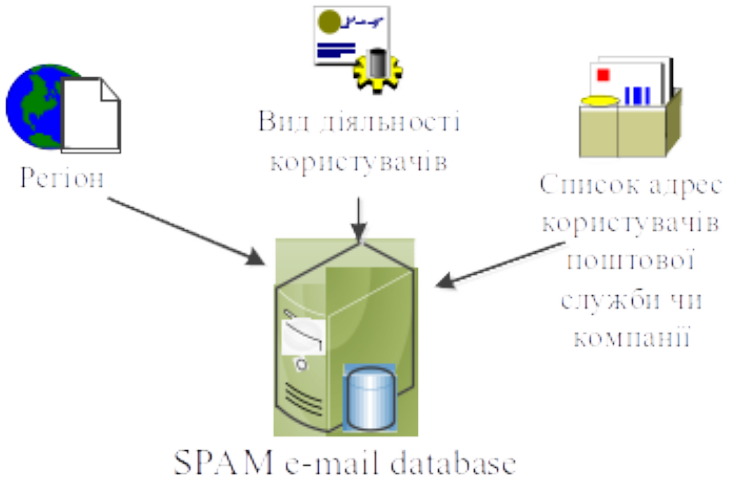

Рис.1.6. Формування списку розсилки

Для отримання списків адрес користуються наступними методами:

- Крадіжка персональних даних користувачів. Може здійснюватись з допомогою комп'ютерних вірусів. Також дані можуть бути отримані шляхом злому баз даних Інтернет провайдерів, Інтернет магазинів, форумів, сервісів і т.д.

- Сканування джерел інформації які знаходяться в загальному доступі. Тобто аналіз повідомлень на форумах, чатах. Також аналізують сайти оголошень. З допомогою сервісу Whois здійснюється пошук поєднань ознак імен скриньок електронної пошти слово1@слово2.слово3.... При цьому, вкінці завжди має бути домен верхнього рівня: gov,.com, . ua , .info).

- Купівля готових списків адрес від недобросовісних компаній і організацій.

- Підбір по шаблонах або словниках власних імен, або написання модулів які перевіряють існування тих чи інших імен поштової скриньки на серверах (наприклад oleh@, oleh_ua@, olehA@, Oleh1999@).

- Метод аналогій. Коли здійснюють пошук однакових імен скриньок на різних поштових серверах, наприклад oleh@ukr.net логічно пошукати чи є аналогічна скринька на сервері gmail.com або tntu.

В ідеальному варіанті додатково визначають супутню інформацію про користувача. Персональні дані, інформацію про відвідані сторінки. Логічно що спеціалізовані повідомлення для певної групи користувачів будуть ефективнішими.

Крадіжки здійснюють з допомогою вірусів з поштових книг користувачів, більшість таких адрес в яких є актуальними. Зважаючи на збільшення кількості вірусних атак, даний метод збору адрес і персональних даних буде вдосконалюватись і використовуватись як один з основних.

Наступний етап це верифікація або перевірка активності зібраних адрес. Так як необхідно перевірити чи дійсно працюють адреси потенційних одержувачів спаму. Як правило це здійснюється одним із способів:

- Пробне надсилання повідомлення. Для цього надсилають лист з текстом який гарантовано буде проходити через систему фільтрації спаму. Крім перевірки роботоздатності адреси також аналізують відповідь самого поштового сервера (пошта прийнята або не прийнята, затримки прийому, інші особливості роботи сервера), таким чином перевіряють кожну адресу майбутнього спам списку.

- Вставлення унікального посилання на Web-сторінку в текст спам-повідомлення. При відкритті листа інформація буде завантажена, а власник сайту дізнається про факт прочитання пробного листа і доступність адреси.

- Розміщення в тексті перевірочного листа інформації про відмову в підписці до якогось ресурсу. Це один з самих поширених методів коли надходить лист з текстом в якому йдеться про те що користувач підписався на розсилку і пропонується, якщо ви не хочете отримувати розсилку далі натисніть на посилання. Одержувач натискає на це посилання і його адреса позначається як активна. Даний метод дозволяє не тільки перевірити чи активна поштова адреса, але і перевірити активність одержувача.

Описані способи перевірки адрес не є єдиними. Спамери постійно розробляють нові методи. Однак ці методи не є достатньо надійними, тому, в spam database є багато неактивних адрес.

2. Створення "точок розсилки".

Розсилка спаму здійснюється в основному методами наведеними на рисунку1.7. Кожен з цих методів має свої переваги та недоліки.

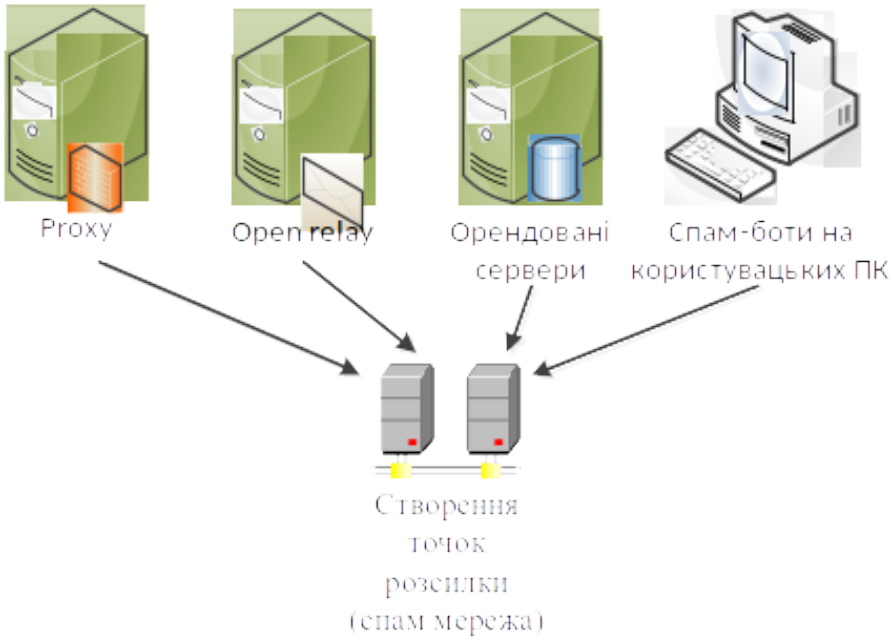

Рис. 1.7. Створення "точок розсилки"

Самим простим варіантом є використання відкритих спам сервісів та орендованих серверів. Проблемою в даному випадку є необхідність частої зміни IP адрес. Так як адреси таких серверів швидко попадають до чорних списків. Відповідно даний вид точок розсилки може бути використай лише у випадку коли поштовий сервер не використовує чорних списків RBL. Спамери постійно оновлюють список таких спам сервісів, однак постійне оновлення чорних RBL списків нівелює отримані результати.

Використання open relay і open proxy також на даний момент все менш ефективне. Провайдери більш професійно здійснюють адміністрування серверів і кількість помилково сконфігурованих постійно знижується.

Тому, на сьогоднішній день, найбільш популярним є створення спам-ботів, тобто встановлення на комп'ютерах користувачів троянських компонент для розсилки спаму. Таким чином створюється розподілена мережа яка постійно оновлюється і зростає завдяки збільшенню кількості пристроїв що підключені до мережі Інтернет і безграмотністю користувачів.

Причому розсилка може вестись як з окремого ПК так і з цілих корпоративних мереж. Як правило з комп'ютерів окремих користувачів надсилається відносно невелика частка повідомлень. Але при цьому в розсилці беруть участь тисячі таких ПК і кількість спам листів може вимірюватись сотнями тисяч.

Також спамери здійснюють контроль доставки повідомлень. Тобто якщо з однієї IP адреси лист був відхилений, то повторна відправка буде вестись вже з іншої IP адреси. Таким чином, як було сказано вище, фільтрація по RBL базах стає неефективною. Розподіленість розсилки реалізується одним із наступних методів:

- Розсилання троянських вірусів і поштових хробаків, піратське програмне забезпечення (патчі, сумнівні сайти, деякі ігри, генератори ключів, і т.д.).
- Використовуються вразливості в браузерах. Коли деякі з таких програм завдяки помилкам в свому коді дозволяють запустити інсталяцію прихованих компонентів що розміщені на веб сторінках. Після цього комп'ютер користувача буде відкрито для віддаленого доступ з боку спамерів. Такі програми поширюються в основному через взламані популярні сайти, призводить що до масового зараження комп'ютерів користувачів, які відвідували такі сайти.

Як бачимо найбільш ефективним методом є використання вірусів, які поширюються з використанням протоколів електронної пошти і використовують вразливості в мережевих сервісах операційних систем. Дана методика набуває все більшого поширення. Основною метою вірусних атак останніх років є отримання

віддаленого несанкціонованого доступу до комп'ютера, також збільшується кількість спроб використання уразливостей операційних систем. Комп'ютер підключений до мережі Internet, без встановленого брандмауера і пакетів оновлень може бути заражений вже за декілька хвилин.

Сучасні комп’ютерні віруси та інші шкідливі програми постійно вдосконалюються та модифікуються. Наприклад, з допомогою деяких вірусів можна приховати від провайдера несанкціоновану розсилку спаму, здійснити координацію розсилки спаму, здійснити оновлення баз даних адрес розсилки, запустити DDoS атаку для виведення з ладу серверів.

3. Програмне забезпечення для розсилки спаму.

Середня сучасна спам-розсилка має об'єм не менше декількох мільйонів повідомлень. Тому в ручному режимі здійснити дану розсилку є неможливою задачею. Адже при запуску розсилки поштові сервери відразу будуть здійснювати оновлення чорних списків на RBL сервісах і блокувати точки розсилки.

Щоб встигнути розіслати повідомлення до перенастроювання або оновлення бази антиспам-фільтрів, потрібно щоб це було здійснено за максимально короткий часовий проміжок, зазвичай це одна дві години. При цьому розсилка великої кількості E-Mail повідомлень вимагає залучення багатьох технічних ресурсів. Тому, як наслідок, були розроблені ряд програм які дозволяють:

- Генерувати динамічні текстові повідомлення.
- Підтримку open proxy і open relay.
- Формування і використання мережі спам-ботів.
- Надсилання звітів про актуальність бази даних адрес розсилки.
- Відслідковування включення точок розсилки до чорних списків.
- Перенаправлення листів в обхід блокованих IP адрес.
- Генерувати та змінювати заголовки листів.

Дані програми постійно вдосконалюються. Розробляють як декстопні так і онлайн сервіси, де зловмиснику надається можливість вибрати особливості функціонування в залежності від своїх потреб.

4. Пошук клієнтів.

Зазвичай клієнти, що хочуть здійснити розсилку спаму знаходять дані сервіси завдяки тим же спам повідомленням. При цьому використовують скриньки приманки з відключеними фільтрами. Проаналізувавши спам зловмисники використовують найбільш ефективні спам сервери.

Також існують додатки які практично цілком легально використовуються для розсилання повідомлень користувачам. Це реклама про заходи, події. Інформування користувачів про зміни в графіках роботи компаній, розсилки новин і т. д. Відмінність полягає в тому що користувачі надають згоду і самі надають свою адресу для розсилки.

5. Формування тексту листів.

На перших етапах розвитку спаму розсилання проводилося без модифікації тексту повідомлення. На сьогоднішній день розсилка однакових тестових спам-повідомлень втратила свою ефективність. Це призвело до появи перших фільтрів, які розраховували контрольну суму листа і порівнювали його з базою. В свою чергу були розроблені нові методики автоматичної модифікації тексту для утруднення його ідентифікації як спам:

- Персоналізація повідомлень. Наприклад, замість статичного тексту: Доброго дня! Фахівці нашої компанії вивчили ваш сайт і дійшли висновку, що наша фірма може вдосконалити його дизайн…» — легко можна отримати динамічний: «Доброго дня Univ! Фахівці нашої компанії вивчили ваш сайт www.tntu.edu.ua і дійшли висновку, що наша фірма може вдосконалити його дизайн…». Ці дані легко отримати з поштової адреси univ@tntu.edu.ua яка зазначена на сторінці університету.

– Внесення до тексту листа «білого шуму». Використовують пробіли, наприклад «Купіть» записують як «К у п і т ь». Або використовують транслітерацію чи спеціальні символи в словах.

– Перефразування - дана методика полягає в тому, що в ході розробки спам-листа можна створити десятки схожих за змістом фраз і потім автоматично зібрати листа. Наприклад, фрази «Купіть супертовар» та «Придбайте нашу продукцію» схожі за змістом, але абсолютно різні для фільтра Байєса або сигнатурного пошуку. Крім того, застосовують словники синонімів для випадкової заміни слів на близькі за змістом еквіваленти.

– Використання можливостей HTML. Для цього додають частини тексту мілким шрифтом і колір символів збігається з фоном листа. Наприклад (рис.1.8.) маємо код і відповідно текст в поштовому клієнті користувача

```
<font color=#FFFFFF>У попа була собака</font><br>
<b>Купуйте товари з супер знижками. Чорна п’ятниця настала</b><br>
<!— Він її любив —>
<b>Купивши один товар на інший знижка 70% !</b><br>
<font size=1 color=Gray>Вона з’їла кусок м’яса...</font>
<br>
<br><br><br><br><br><br><br><br><br><br><br><br><br><br>
<br><br><br><br>
Текст невидимка 1<br>
<table width=”5000" border=”0" cellspacing=”0"
cellpadding=”0">
<tr><td align=right> Текст неведимка 2
</table>
```

Купуйте товари з супер знижками. Чорна п’ятниця настала

Купивши один товар на інший знижка 70%

Вона з’їла кусок м’яса...

Рис. 1.8. HTML-спам

Перший рядок тексту не видно, оскільки колір символів збігається з кольором тла. Третій рядок теж не помітний, тому що є коментарем HTML. Нарешті, текст «З'їла вона кусок м'яса...» видно, але розмір шрифту та його колір майже непомітний на тлі основних написів. Текстовий рядок «Невидимий текст 1» є видимим з погляду HTML, але непомітний для користувача, оскільки перед ним стоїть багато тегів
 і він опиняється поза увагою. Текст «Невидимий текст 2» HTML теж бачить, а читача листа він непомітний, оскільки вирівнюється з правого краю таблиці шириною 5 тис. пікселів і виявляється поза увагою. Наведені методи вкрай прості в реалізації, але вони суттєво ускладнюють аналіз листа за допомогою сигнатур або

статистичних методів. Проте виявлення маскування тексту може використовуватися фільтром як додаткову ознаку.

- Заміна тексту зображенням. Статичні зображення без маскування - їх цілком достатньо для обходу статистичних фільтрів. Зокрема, можна порівняти сигнатури зображення з базою сигнатур подібних спам-картинок або застосувати алгоритми OCR, що дозволяють розпізнати текстові дані та передати текст для аналізу.

Інший варіант це динамічні зображення - відрізняється від попереднього тим, що в момент його генерації змінюється розмір зображення, використовувані шрифти, колір символів та колір фону. В цьому випадку єдиним методом аналізу залишається OCR-метод.

Динамічно формується зображення з анти-OCR-елементами - як випливає з назви, даний метод відрізняється від попереднього тим, що при формуванні зображення в нього включаються елементи, що ускладнюють пошук і розпізнавання тексту. У найпростішому випадку на зображення наносяться графічні «шуми» у вигляді точок і ліній. OCR-системи мають досить обмежені можливості і з такими зображеннями справляються погано.

1.4. Аналіз програмних систем захисту поштових сервісів

Існує значна кількість програмних засобів, як платних, так і безкоштовних, що здійснюють захист від несанкціонованих розсилок поштових сервісів.

Недоліком існуючих систем фільтрації є не лише низька якість розпізнавання спаму, але й блокування легітимних повідомлень

Однією з таких систем є система «Дозор – Джет». Дана система підходить до фільтрації спаму комплексно, використовуючи статистичні алгоритми та технологію фільтрації на основі ознак спаму. При цьому фільтрація здійснюється двома основними способами: за формальними ознаками і змістом текстової складової листів.

Система «Дозор-Джет» функціонує на UNIX-платформі та включає до свого складу підсистему архівування. Система «Дозор-Джет» використовується в організаціях, обсяг поштового трафіку яких досягає 5 гігабайт на день, а кількість поштових адрес перевищує 5000. Це вимагає застосування апаратних засобів, які здатні забезпечити високу продуктивність системи.

AntiSPAMer. Для здійснення фільтрації електронних поштових повідомлень передбачено можливість створення «чорних» та «білих» списків для повідомлень.

Перевірка повідомлень відбувається лише за заголовками, що дозволяє видаляти спам-розсилку на сервері. Система фільтрації здатна самостійно навчатись відповідно до заданих правил.

Недоліком даної системи фільтрації є необхідність тривалого навчання фільтра підвищення ефективності. Крім того, перевірка ефективності даної системи після початкової установки показала, що тільки 25% з 500 спам розсилок були ідентифіковані правильно.

Mailbox cleaner здійснює перевірку та блокування несанкціонованих електронних повідомлень на поштовому сервері. Перевірка повідомлень здійснюється за заголовками повідомлень. Класифіковані повідомлення розподіляються за різними категоріями, що дозволяє швидко зорієнтуватися при великому обсязі перевірки.

Налаштування системи фільтрації передбачає можливість вказання заголовків листів, які система фільтрації повинна вважати як спам повідомлення. Але за таких параметрів системи заблокованими можуть бути легітимні листи від користувачів, що некоректно заповнюють заголовні поля повідомлень.

Недоліком системи Mailbox cleaner є те, що використання даної системи фільтрації вимагає постійного контролю адміністратора безпеки для налаштування системи у зв'язку з тенденціями зміни спам повідомлень. Тестування даної системи дозволило здійснити ідентифікацію 60% спам розсилок.

Недоліком існуючих систем фільтрації є не лише низька якість розпізнавання спаму, а й блокування легітимних повідомлень

Основні результати, які отримано в даному розділі:

1. Проведено аналіз розвитку технічних методів розсилки і фільтрації спаму, основні ознаки спаму, що дало можливість виявити притаманні спам листам особливості і можливості використання методів їх фільтрації.

2. Досліджено особливості процесу формування спам розсилок, що дало можливість враховувати їх при організації роботи проектованих антиспамерних фільтрів.

3. Проаналізовані програмні засоби які використовуються для фільтрації спаму, що дало можливість визначити їх переваги, недоліки і визначити вимоги до проектованої системи фільтрації.

РОЗДІЛ 2
АЛГОРИТМИ ОБРОБКИ І ФІЛЬТРАЦІЇ ЕЛЕКТРОННИХ ПОВІДОМЛЕНЬ

Аналіз методів показав, що найперспективнішим напрямом досліджень у галузі фільтрації спаму є методи які здійснюють аналіз не формальних ознак спаму, а здійснюють аналіз самого повідомлення. Такими є методи на основі теореми Байєса, кластерний аналіз, нейромережеві методи.

Основними їх перевагами є: можливість аналізу даних в умовах неповноти, спотвореності та неточності інформації; робота у режимі реального часу; незалежність обсягу обчислень від кількості об'єктів виявлення та ідентифікації.

Знизити відсоток помилкових спрацьовувань пропонується за рахунок розробки дворівневої системи фільтрації, представленої на малюнку 2.1, яка складається з формального та інтелектуального фільтрів.

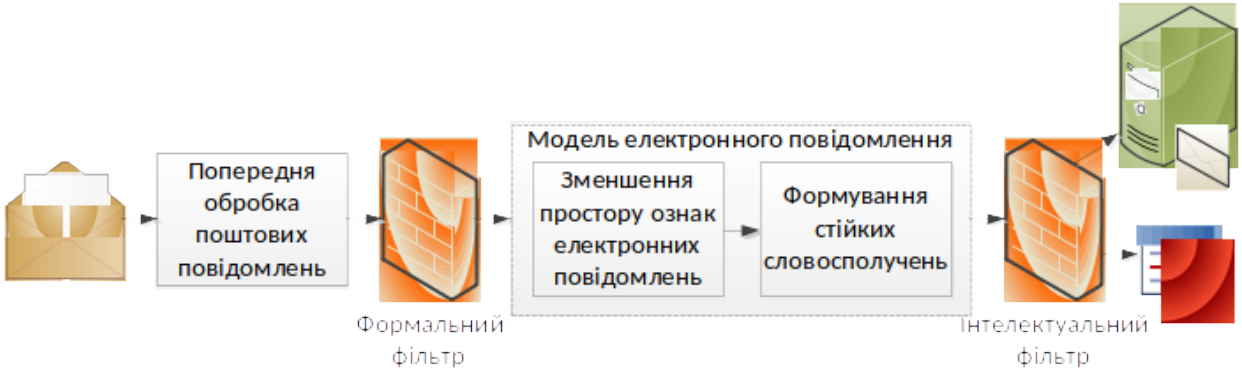

Рис. 2.1. Запропонована схема фільтрації електронних повідомлень

Попередня обробка повідомлень полягає у приведенні до стандартного типу кодування, видалення стоп-слів, гіперпосилань та знаків пунктуації.

Формальний фільтр використовує класичні методи фільтрації: адреси (IP, e-mail), що поділяють листи на дозволені та заборонені, і є базою даних ознак листів, що формується адміністратором.

Інтелектуальний фільтр здійснює семантичну класифікацію листів конкретного адресата електронної пошти. Даний фільтр потребує навчання класифікатора.

2.1. Алгоритми формальних методів фільтрації

Фільтрація по заголовку «From:»

Це один з найпростіших і поширених видів фільтрації. Організовується у вигляді так званих «білих» і «чорних» списків. Якщо поштова адреса із заголовка «From:» присутня в «білому» списку, то приймаємо повідомлення і припиняємо подальшу перевірку, інакше відсікаємо як спам. Також часто спам повідомлення містять порожній або не є поштовою адресою заголовок «From:». Даний факт є вагомою підставою для відсікання повідомлення як потенційного спаму.

Можна самостійно видаляти і додавати нові записи в «білий» і «чорний» список. Проте даний вид фільтрації не можна вважати ефективним. Річ у тому, що адреса, що міститься в полі «From:», може не бути дійсною адресою відправника. І при наступній розсилці спамер може запросто змінити свою адресу або взагалі генерувати фіктивні поштові адреси в процесі розсилки. Обробка заголовка «From:» проводиться для кожного повідомлення окремо. На рисунку 2.2 показано блок-схему роботи цього методу.

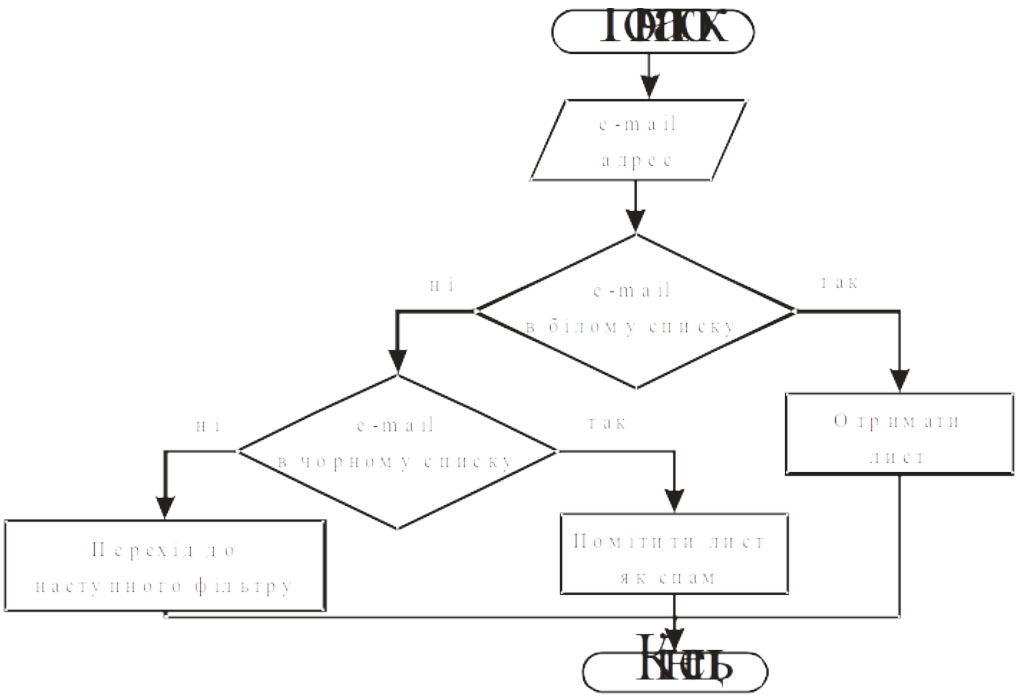

Рис. 2.2. Блок-схема обробки заголовка «From:»

Фільтрація по заголовку «To:»

Зустрічається, що спамери залишають порожнім заголовок «To:» або указують в ньому яку-небудь іншу адресу, що не є адресою одержувача. В даному випадку основна ставка робиться на психологічний момент: користувач обов'язково зацікавиться змістом листа, адресованого не йому, відкриє і прочитає. Природною реакцією програми-антиспамера на подібний заголовок є відсікання повідомлення як потенційного спама.

Блок-схема алгоритму обробки заголовка «To:» показана на рисунку 2.3.

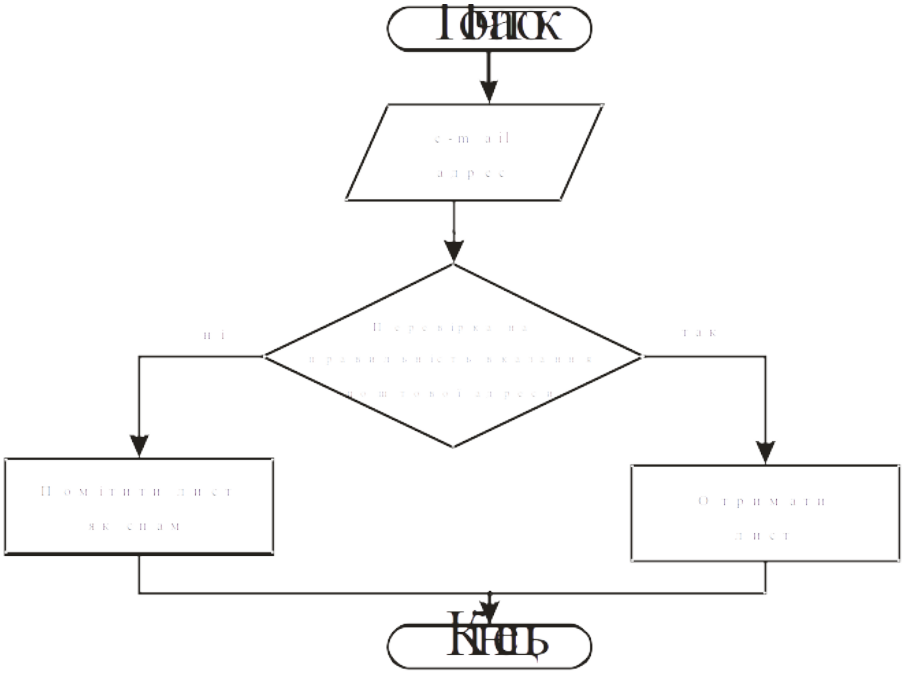

Рис. 2.3. Блок-схема обробки заголовка «To:»

Фільтрація по заголовку «Message-ID:»

Заголовок «Message-ID:» містить унікальний ідентифікатор поштового повідомлення. Звичайно цей ідентифікатор привласнюється повідомленню поштовим сервером у момент прийому повідомлення від поштового клієнта відправника. Якщо повідомленню не привласнений унікальний ідентифікатор, то поштовий сервер зобов'язаний привласнити його повідомленню, навіть якщо повідомлення вже знаходиться на поштовому сервері одержувача. Формат унікального ідентифікатора показаний на рисунку 2.4

Довільний набір символів	@	Доменне ім'я поштового сервера, що привласнив повідомленню Message-ID

Рис. 2.4. Формат заголовка Message-ID

Останнім часом серед спамерів велику популярність має метод розсилки повідомлень прямо на поштові сервери одержувачів (Mail Exchanger). Звичайно для розсилки використовується dial-up з'єднання з IP-адресою, що динамічно виділяється, тому додавання такої IP-адреси в базу джерел спаму не дасть ніякого результату, навпаки, подібними діями можна завдати шкоди іншим користувачам, що одержали цю IP-адресу, заблокувавши їх роботу.

Приклад заголовка спамерного повідомлення, посланого прямо на поштовий сервер адресата, показаний на рисунку 2.5

```
Received: from 12-225-19-197.client.attbi.com ([12.225.19.197])
          by mx.inf.tsu.ru (Lotus Domino Release 5.0.11)
          with SMTP id 2003031308304714:5451 ;
          Thu, 13 Mar 2003 08:30:47 +0600
From: Американский Деловой Центр 411-0232 <quryarm@mail.ru>
To: 7119 <7119@inf.tsu.ru>
Subject: Бизнес Английский для Вас ,Вашей Фирмы и Вашей Семьи
Date: Thu, 13 Mar 2003 08:30:48 +0600
Message-ID: OFAB261D12.B331D4AF-ONC6256CE8.000DCE72@inf.tsu.ru
```

Рис. 2.5. Приклад заголовка спамерного повідомлення, посланого прямо на поштовий сервер адресата

У разі спамерної розсилки такого роду унікальний ідентифікатор буде привласнений повідомленню поштовим сервером адресата. І заголовок «Message-ID:» міститиме після символу «@» доменне ім'я поштового сервера, обслуговуючого адресата.

Фільтрація по заголовку «Subject:»

Заголовок «Subject:» містить тему повідомлення. Майже завжди спамерні поштові повідомлення містять осмислені теми. Тому доцільно вести пошук за словами і словосполученнями, що зустрічаються в спам листах (див. рис. 2.6).

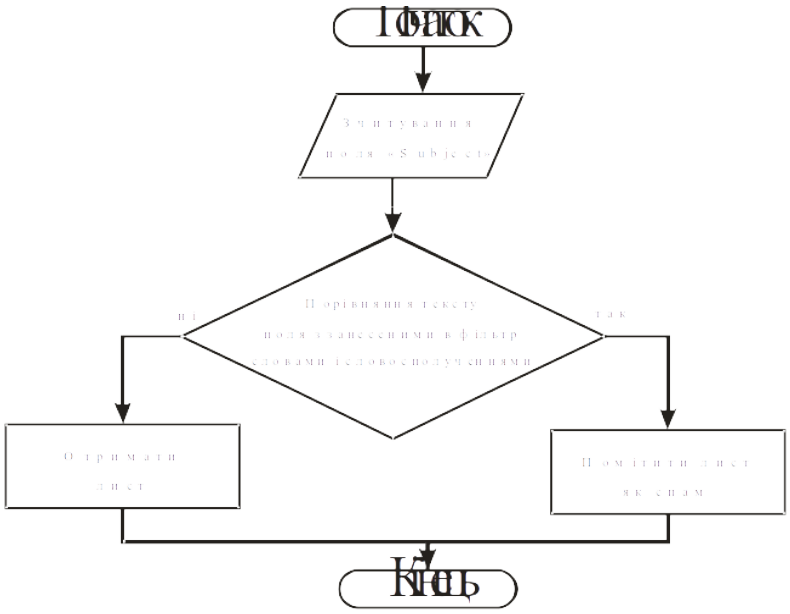

Рис. 2.6. Блок-схема обробки заголовка «Subject:»

Адміністратору надається можливість видаляти і додавати нові слова і словосполучення, а також визначати подальші дії над повідомленням, потенційним спамом, що є. Проте слід бути вкрай обережним у використовуванні даного виду фільтрації, оскільки при неправильному його використовуванні можливе відсікання корисних повідомлень.

Фільтрація по заголовках «Received:»

Заголовки «Received:» містять повну інформацію про шлях проходження листа від джерела до адресата. При прийомі повідомлення поштовий сервер додає в початок листа новий заголовок «Received:». Формат заголовка «Received:» показаний на рисунку 2.7.

from	DNS-ім'я передаючого, яке було заявлене їм при передачі повідомлення	Справжній IP і/або DNS-ім'я передаючого, виявлене приймаючим повідомлення сервером

by	DNS-ім'я сервера, що прийняв повідомлення	for	Поштова адреса одержувача повідомлення

Рис. 2.7. Формат заголовка «Received:»

На рисунку 2.7 показані лише обов'язкові складові заголовка, які корисні для виявлення спам повідомлень. Поштовий сервер може самостійно додавати будь-які інші поля і їх значення, ніякої стандартизації тут немає. Це зв'язано з тим, що первинно заголовки «Received:» не призначалися для машинної обробки, а лише служили додатковим джерелом інформації для людини. Зрозуміло, що людина в міру своєї компетенції може витягнути із заголовка корисну для себе інформацію, проте машинна обробка заголовка представляє додаткові труднощі.

Суть спам фільтрації по заголовках «Received:» полягає в дослідженні шляху проходження повідомлення. Цей шлях повинен бути послідовним і логічним.

Зустрічаються спам повідомлення, в які додані фіктивні заголовки «Received:». Це робиться для приховування істинного джерела спама і перенаправлення скарг користувачів на ні в чому не повинний вузол. Приклад фіктивних заголовків показаний на рисунку 2.8.

```
Received: from va-nrrws-ubr-a-024-196-179-061.charterva.net
([24.196.179.61])
          by mx.inf.tsu.ru (Lotus Domino Release 5.0.11)
          with SMTP id 2003031208003149:4586 ;
          Wed, 12 Mar 2003 08:00:31 +0600
Received: from wetwetwet.com (2544 [102.50.246.26])
        by  swol.de (8.12.1/8.12.1) with ESMTP id 31657
        for <7119@inf.tsu.ru>; Tue, 11 Mar 2003 18:57:19 -0800
Received: from ubn.cscoms.com ([242.15.193.141])
        by aug.com (8.9.3/8.9.3) with SMTP id 11746
        for <7119@inf.tsu.ru>; Tue, 11 Mar 2003 18:57:14 -0800
```

Рис. 2.8. Приклад фіктивних заголовків «Received:»

Шлях передачі повідомлення можна прослідити, послідовно читаючи вміст заголовків «Received:» від низу до верху.

З прикладу видно, що джерелом повідомлення є вузол ubn.cscoms.com, що має IP-адресу 242.15.193.141. Це повідомлення було прийнято вузлом aug.com. Заголовок «Received:» також був доданий поштовим сервером, що працює на вузлі aug.com.

Далі вузол aug.com повинен передати повідомлення або поштовому сервера адресата, або іншому поштовому сервера. Але в другому знизу заголовку «Received:» видно, що при отриманні повідомлення вузлом swol.de, вузол aug.com називається вже як wetwetwet.com і має IP-адресу 102.50.246.26.

Потім відбувається ще одне дивне перетворення: поштовий сервер адресата mx.inf.tsu.ru одержує повідомлення вже не від swol.de, а від va-nrrws-ubr-a-024-196-179-061.charterva.net, має IP-адресу 24.196.179.61.

Єдиним справжнім заголовком «Received:» в даному прикладі є перший зверху, оскільки він був доданий поштовим сервером адресата mx.inf.tsu.ru. А джерелом спам повідомлення є va-nrrws-ubr-a-024-196-179-061.charterva.net.

Для остаточної впевненості в підробці заголовків слід перевірити, чи існують доменні імена, присутні в двох фіктивних заголовках і чи відповідають їм IP-адреси, що містяться в цих же заголовках. Для цього можна скористатися програмою «ping», сервісом «whois» [18.]. У разі автоматизованої перевірки у складі антиспамерного програмного комплексу слід посилати запит на DNS-сервер.

Фільтрація по заголовках «Priority:» і «X-Priority:»

Заголовок «Priority:» абсолютно вільний заголовок, тобто його вміст не регламентується. Більшість поштових клієнтів його просто ігнорує. Рідко використовується спамерами.

Заголовок «X-Priority:» використовується поштовими клієнтами для графічного відображення терміновості повідомлення. Наприклад, в поштовому клієнті «The Bat» нові непрочитані повідомлення, що мають пріоритет «5 (Normal)» відображаються у вигляді поштового конверта жовтого кольору, а повідомлення з пріоритетом «7 (High)» відображаються червоним кольором. Привласнюючи повідомленню щонайвищий пріоритет, спамери сподіваються привернути увагу користувача незвичайним відображенням цього повідомлення.

Перевірка існування відправника повідомлення

Для здійснення даної перевірки необхідно взяти адресу відправника з поля «From:» і спробувати відправити йому у відповідь повідомлення, але не через проміжний поштовий сервер, а прямо на поштовий сервер відправника. Якщо

поштова адреса, вказана в полі «From:» існує на сервері, то сервер запитає тіло поштового повідомлення, інакше поверне помилку про відсутність вказаної адреси. Після цього необхідно просто відключитися від поштового сервера. Блок-схема алгоритму перевірки існування відправника показана на рисунку 2.9.

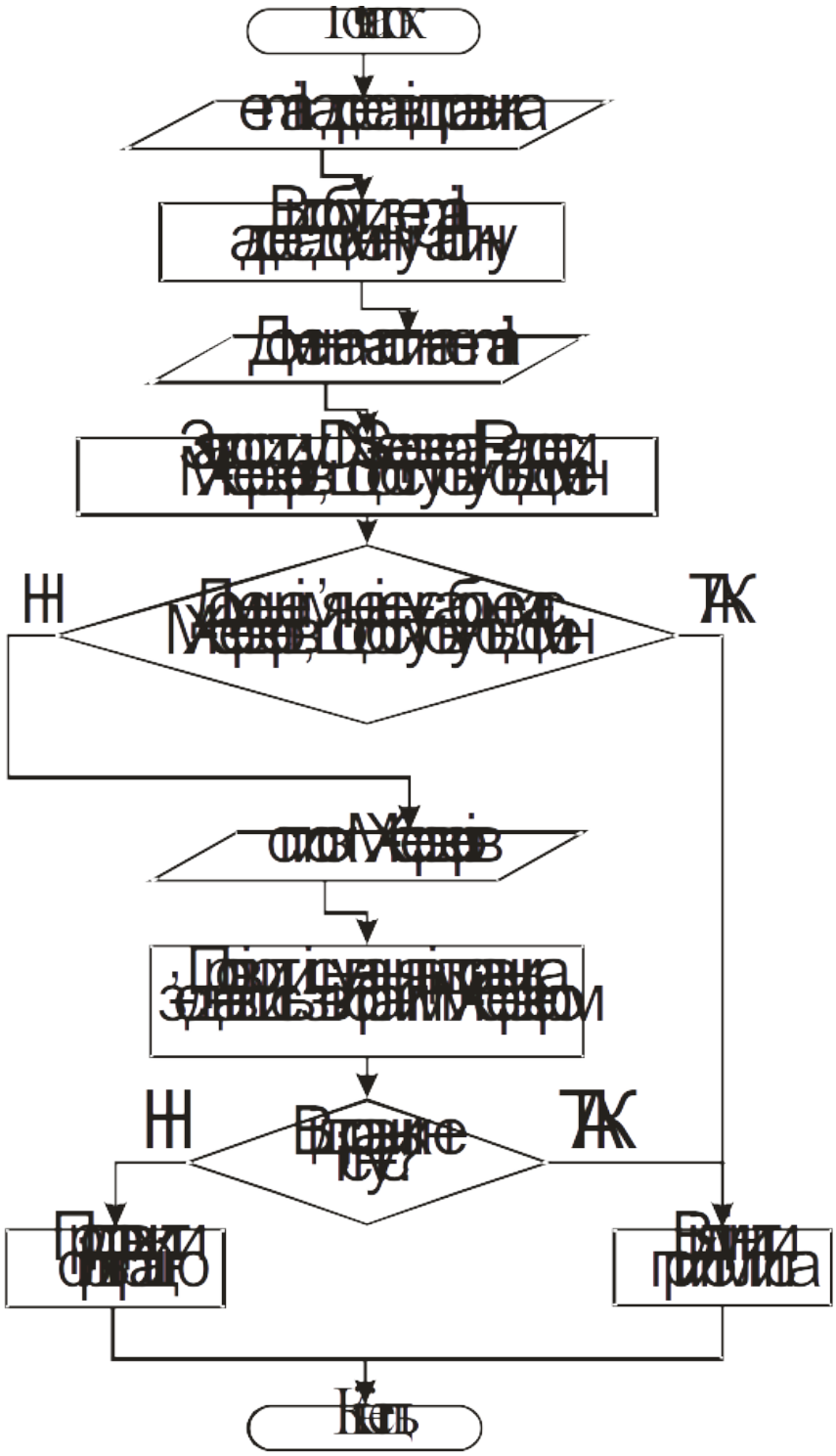

Рис. 2.9. Блок-схема алгоритму перевірки існування відправника

Необхідно відзначити, що цей метод перевірки часто використовується спамерами для виявлення нових поштових адрес. Спеціалізована програма перевіряє наявність адресатів необхідного поштового сервера, послідовно перебираючи їх. Потенційні імена адресатів беруться із спеціального словника імен. Таким методом можна одержати біля 80% адрес користувачів поштового сервера.

Приклад, що демонструє роботу метода перевірки існування відправника повідомлення, показаний на рисунку 2.10.

```
220 info.tsu.ru ESMTP Sendmail 8.9.3/8.9.3/TSU; Mon, 17 Mar 2003 20:17:15
+0600 (TSK)
HELO inetclub.tomica.ru
250 info.tsu.ru Hello [217.106.32.193], pleased to meet you
MAIL FROM:yefim@post.tomica.ru
250 yefim@post.tomica.ru... Sender ok
RCPT TO:spam_sender@tsu.ru
550 spam_sender@tsu.ru... User unknown
QUIT
221 info.tsu.ru closing connection
```

Рис. 2.10. Приклад роботи методу перевірки існування відправника

Даний приклад був одержаний за допомогою програми «Telnet». Скористатися нею можна, виконавши в командному рядку команду «telnet». Як параметри можна відразу вказати DNS-ім'я або IP-адресу вузла і порт, до якого проводиться підключення. В прикладі використовувався вузол info.tsu.ru, що є MX-сервером для домена tsu.ru і всіх його субдоменів. Підключення проводилося на порт 25. На цьому порту сервер працює по протоколу SMTP (Simple Mail Transfer Protocol).

Далі, користуючись командами протоколу SMTP, ініціювалася передача повідомлення. В прикладі команди, посилані користувачем на сервер, виділені жирним шрифтом. Після спроби вказівки як одержувач spam_sender@tsu.ru сервер повернув відповідь, що такого користувача не існує: «550 spam_sender@tsu.ru... User unknown». Після цього на сервер посилалася команда завершення сеансу роботи з сервером «QUIT».

Якщо як одержувач вказати «7119@inf.tsu.ru», то відповідь сервера на наш запит буде наступною: «250 7119@inf.tsu.ru... Recipient ok», оскільки даний користувач обслуговується MX-сервером info.tsu.ru.

Посилка у відповідь листа спамеру з серверною помилкою «адресат не знайдений»

Даний метод слід використовувати тільки тоді, коли достовірно відомо, що повідомлення є спамом і поштова адреса відправника існує.

Одержавши у відповідь лист з серверною помилкою «адресат не знайдений», програма, що розсилає спам листи, повинна виключити відповідну поштову адресу з своєї бази поштових адрес. Шаблон для такого повідомлення показаний на рисунку 2.11.

```
Received: from доменн.имя.почтов.серв.пользов [IP-адрес почтов.серв.пользов]
by доменное.имя.пользов (8.11.6/8.11.6) id ID_сообщения;
День_недели, День Месяц Год Час:Мин:Сек +Часов (Часов.пояс)
Received: from localhost (localhost)
by доменн.имя.почтов.серв.пользов (8.11.6/8.11.6) id ID_сообщения;
День_недели, День Месяц Год Час:Мин:Сек +Часов (Часов.пояс)
Date: День_недели, День Месяц Год Час:Мин:Сек +Часов (Часов.пояс)
From: Postmaster@доменн.имя.почтов.серв.пользов
To: адрес_спамера@доменное.имя.почтового.сервера.спамера
Message-Id: <ID_почтового_сообщения@доменн.имя.почтов.серв.пользов>
MIME-Version: 1.0
Subject: DELIVERY FAILURE: 550 < адрес_пользов@доменн.имя.почтов.серв.пользов
>... User unknown

Your message

Subject: Тема из спамового сообщения

was not delivered to:

адрес_пользов@доменн.имя.почтов.серв.пользов

because:

550 5.1.1 < адрес_пользов@доменн.имя.почтов.серв.пользов > user unknown
```

Рис. 2.11. Шаблон повідомлення, посланого спамеру

Повідомлення про помилку відсилається прямо на MX-сервер, обслуговуючий спамера. Але для більшої достовірності повідомлення слід підроблювати заголовки «Received:», подібно тому, як це роблять спамери. Важливо також пам'ятати, що

MX-серверу спамера буде достовірно відомий DNS і IP-адреса користувача. Тому цю адресу необхідно включити в логічний ланцюжок заголовків «Received:».

2.2. Модель текстового контенту електронних листів

Модель представлення текстового документа є важливою складовою для його комп'ютерної обробки. Вибір моделі визначає ефективність виділення змістового змісту та структури електронного повідомлення.

Під моделлю тексту розуміється його наближений опис, виражений з допомогою математичної символіки. Модель завжди простіша самого тексту, відображає лише його основні властивості.

Текст електронного поштового повідомлення можна подати у вигляді термів (слів), що дозволяє подати кожен документ у вигляді вектора у просторі ознак. Графічне відображення документа D з трьох термів t_1, t_2, t_3 представлено на рисунку 2.12.

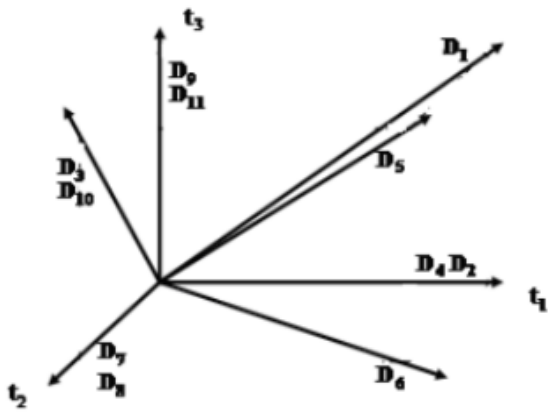

Рис.2.12. Векторне представлення документа

Тривимірне відображення документа, що складається з трьох термів, може бути поширене і на N-мірні документи, де N – кількість термів у документі.

В рамках векторної моделі електронний лист описується вектором у деякому просторі ознак, у якому кожному використовуваному в повідомленні терму ставиться у відповідність його вага (значимість).

У матричному вигляді лист можна представити як матрицю стовпець S_i. Її елементами є ваги відповідних термів у повідомленні.

Основною перевагою векторної моделі є можливість використання алгоритмів класифікації заснованих на аналізі статистичних характеристик, а також можливість порівнювати вектор у векторному просторі ознак. Завдання перетворення тексту на вектор у просторі ознак вимагає визначення координати ознак.

2.3. Векторна модель електронного повідомлення і завдання класифікації

При використанні векторної моделі зміст листа можна описати за допомогою термів *t*, безліч яких утворюватиме тезаурус $T\{t_1,\ldots.t_q\}$ певного класу *k*. Як терми використовуються слова, що становлять зміст повідомлення. Аналіз простору ознак $S(p_i)$ представлених на рисунку 2.13, дозволив вибрати як ознаку вагу терма.

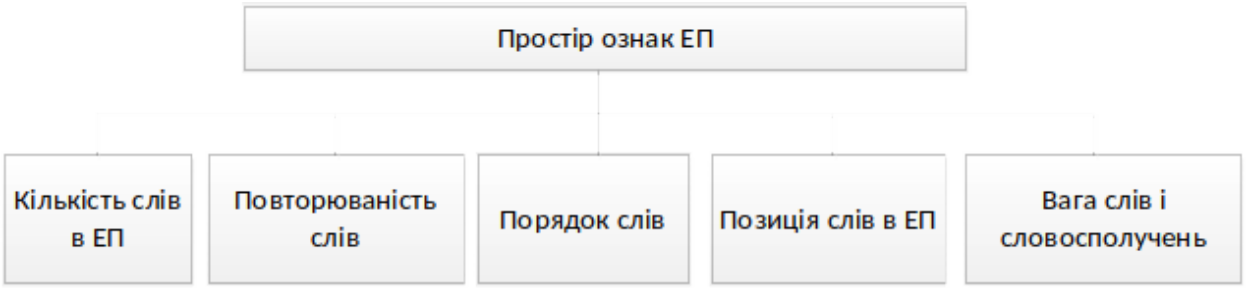

Рис.2.13. Простір ознак електронних повідомлень

Тоді модель ЕП можна подати у вигляді 2.1

$$S(p_i)=(t_j, w(t_j)), \tag{2.1}$$

де t - j -ий терм у повідомленні;

p_i - простір ознак, що визначають повідомлення;

$w(t_j)$ – вага терму в повідомленні після видалення стоп-слів.

Якщо одне електронне повідомлення можна представити у вигляді 2.1, то всю колекцію повідомлень певного класу L_k набуде вигляду 2.2

$$L_k=(T_k, w(t_j)),$$

(2.2)

де T_k - k -тий тезаурус повідомлення класу k;

$w(t_j)$ – вага терму в повідомленні.

Підхід до побудови векторного представлення тексту при використанні логічної міри значущості термів має наступний недолік: при оцінці близькості векторів з використанням даної міри значущості термів результати можуть бути не завжди однозначними. Крім того, на результат класифікації також може впливати зміна кількості термів у повідомленні.

Позбавитись від цього недоліку можливо за умови зміни способу визначення ваги терму. Отже, одним з основних завдань при роботі з текстовим вмістом ЕП стає визначення вагового коефіцієнта w_{iq}, що визначає вагу відповідного терма t_j в i-му документі.

Одним з методів визначення ваги терму в тексті є LTC. Даний підхід полягає у використанні логарифму входження в текст частоти слова замість просто частоти слова. Даний метод дозволяє виділяти середньо і низькочастотні терми.

$$w_{ij}=\frac{\log(f_{ji}+1)\log\left(\frac{N}{n_j}\right)}{\sqrt{\sum_{l=1}^{M}\left[\log(f_{li}+1)\log\left(\frac{N}{n_l}\right)\right]^2}}.$$

(2.3)

Таким чином, використання LTC ваги дозволяє скоротити ефект великих відмінностей у частотах, що робить використання даного підходу найбільш прийнятним. Повідомлення, які формують навчальну вибірку, можна представити як матриці, стовпцями якої будуть листи, а рядками терми, які містяться у листах:

$$L_k = \begin{bmatrix} w_{11} & w_{21} & \dots & w_{j1} \\ w_{12} & w_{22} & \dots & w_{j2} \\ \vdots & \vdots & \vdots & \vdots \\ w_{1i} & w_{2i} & \dots & w_{ji} \\ \vdots & \vdots & \vdots & \vdots \\ w_{1N} & w_{2N} & \dots & w_{MN} \end{bmatrix},$$

(2.4)

де $w_{it_j} = L_{tc_{it_i}}, j=1,\dots,M, i=1,\dots,N$.

Отримана матриця ознак ЕП має розмірність обробка якої вимагає великих обчислювальних ресурсів.

Крім того, згідно із законами Ципфа [11.], слова, що зустрічаються в тексті найчастіше, є малоінформативними, тобто не мають вирішального змістового значення, що стає основою зниження розмірності матриці ознак за рахунок позбавлення від малоінформативних термів без втрати змістового змісту ЕП.

Також проведений аналіз частотності термів у повідомленнях класу легітим і спам (рис. 2.14) показав що:

- Існують терми, що зустрічаються в одному класі, але не зустрічаються в іншому класі.
- Існують терми, які найчастіше зустрічаються у певному класі, що говорить про можливість даних термів характеризувати той чи інший клас;
- Існують терми, яким важко визначити приналежність до того чи іншого класу, оскільки частота попадань даних термів у двох класах практично не відрізняється.

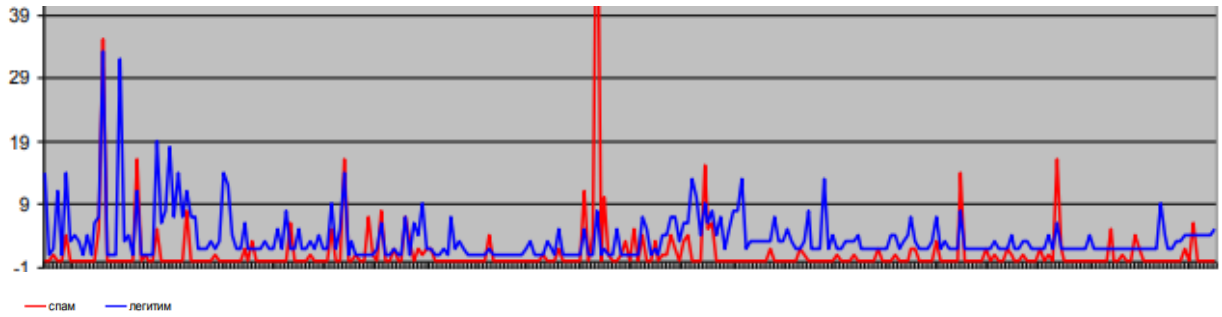

Рис.2.14. Кількість термів за класами (spam/legitim) типового поштового електронного повідомлення

Отримані висновки свідчать про необхідність скорочення простору ознак з метою вилучення з малоінформативних термів і виділення термів здатних характеризувати той або інший клас. Для скорочення простору ознак в задачах класифікації використовуються наступні підходи:

- скорочення просторів ознак безпосередньо для кожного класу;
- скорочення просторів ознак для всіх листів навчальної вибірки без врахування належності тому чи іншому класу.

Для скорочення простору ознак використано підхід, що ґрунтується на тому, що для кожної терми у повідомленнях певного класу вираховується величина $RF_{t_j}^k$, яка характеризує вага терм для певного класу k:

$$RF_{t_j}^k = \log_2\left(2 + \frac{a_i}{max\ (1, b_i)}\right), \tag{2.5}$$

де a_i - кількість ЕП, які містять t_j терм і відносяться до класу k,

b_i - кількість ЕП, які містять t_j терм і не відносяться до класу k.

Терми, вага яких $RF_{t_j}^k \leq 1{,}5$, виключаються в даному повідомленні класу k.

2.4. Виявлення стійких словосполучень в тексті листа

Під стійким словосполученням розуміється комбінація двох і більше термів, що мають тенденцію до спільного використання.

Алгоритм формування стійкого словосполучення полягає в наступному:

1) виділення ваг термів для відповідного класу k (spam/legitim);

2) розрахунок близькості термів та прийняття рішення щодо формування стійкого словосполучення;

3) підтвердження смислової значущості сталого словосполучення.

Таким чином завдання фільтрації вхідної поштової кореспонденції можна подати у вигляді етапів, показаних на малюнку 2.15.

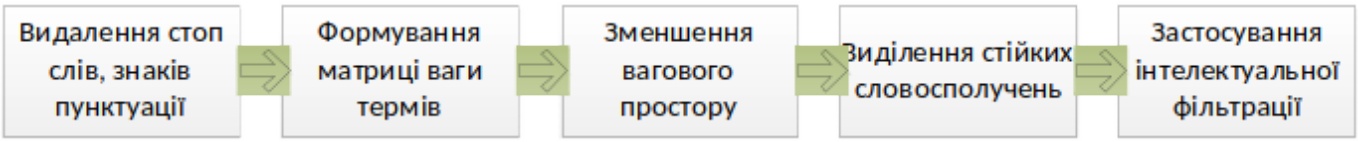

Рис.2.15. Послідовність вирішення задачі фільтрації ЕП

Нехай нам на пошту прийшло наступне повідомлення

Від: Lib@mail.ua
Відправлено: 9 грудня 2021 р. 17:04
Кому: kaf@mail.ua
Вкладення: Інформаційні ресурси.doc
Шановні завідувачі кафедр! З метою інформованості про інформаційні ресурси наукової бібліотеки, а також для успішного проходження акредитації у 2022 р. освітніх програм, розсилаю Вам довідку про стан фонду бібліотеки вузу.
З повагою, Іванов Іван Іванович, зам. директора Наукової бібліотеки тел.: (xxxx) xxx-xx-xx; внутр. xx-xx

Рис. 2.16. Приклад легітимного ЕП

У таблиці 2.1 наведено терми та розраховане значення загальної ваги RF.

Таблиця 2.1

Розраховані значення RF термів у повідомленні

Терми у повідомленні	RF
внутр	3
акредитації	1
2022	4,45943
інформаційні	1
освітніх	2
програм	2,32193
наукової	3,32193
розсилаю	1
бібліотеки	3,32193
довідку	1
шановні	2,39232
стані	1
завідувачі	2
кафедрами	2,32193

вузу	1
метою	2
повагою	2,48543
поінформованості	1
Іванов	2,80735
інформаційних	2
Іван	3,16993
ресурси	1
Іванович	4,24793
також	1,80735
директора	3,58496
успішного	1
проходження	1

Відповідно до прийнятого порогу $RF_{t_j}^{k} \leq 1{,}5$ то такі терми виключаються у цьому повідомленні. Аналіз ваги термів у повідомленні після скорочення показав, що існують терми які мають більшу класифікаційну вагу і отже можуть більшою мірою вплинути на результат фільтрації, а деякі меншу, що говорить про необхідність виділення термів, найбільш важливих для детектування повідомлень. Тому необхідно застосувати методику виділення ключових термів з текстового вмісту повідомлення, що відображають специфіку ЕП і дозволяють виділити терми, що мають найбільш виражену класифікаційну вагу.

У той самий час, не всі терми, які є у повідомленні, відображають його тематику. Службові слова не несуть смислового навантаження, а використовуються для зв'язку слів у реченнях. Самостійні слова, частота використання яких у тексті невелика, також не відображають тематику тексту. У зв'язку з цим необхідно виділити терми, здатні відображати зміст повідомлення. Вилучення таких термів можна змоделювати через процедури виділення ключових слів тексту.

Існують два основні підходи до автоматичного виділення термінів.

Перший підхід відноситься до галузі статичної обробки природної мови - обчислення різних мір асоціативного зв'язку, які визначають, чи є взаємна поява лексичних одиниць випадковою, або вона статично значуща.

Другий підхід спирається на семантичну близькість термів. Що передбачає визначення семантичного зв’язку термів у повідомленнях.

Нехай

$$D_i = \{d_{jq}\},\ j = 1, \ldots, N$$

характеристика зв'язку між термами в i -му повідомленні, де d_{jq} – ступінь смислової близькості j-го та q-го термів.

Як міра близькості між термами в повідомленні використано відстань Дайса. Даний статистична міра дозволить об'єднати терми у стійкі (ключові) словосполучення, що характеризують семантичний зміст повідомлень.

Близькість D і частота $f(t_1, t_2)$ спільного використання термів є причиною для знаходження стійких словосполучень.

Міра Дайса D розраховується з виразу:

$$D(t_1, t_2) = \log_2\left(\frac{2f(t_1, t_2)}{f(t_1) + f(t_2)}\right),$$

де $f(t_1)$ і $f(t_2)$ - частота входжень термів t_1 і t_2 в повідомленні;

$f(t_1, t_2)$ – частота спільних входжень термів t_1 і t_2 в повідомленні.

Для завдання фільтрації електронних поштових повідомлень у даній роботі пропонується формувати стійке словосполучення, якщо значення D_{jq}дорівнює або вище, ніж у сусідніх парах термів.

Таким чином, модель ЕП у формі стійких словосполучень дозволяє без втрати змісту забезпечити інтелектуальну класифікацію поштової електронної кореспонденції.

Висновки до розділу 2

1. Запропонована багатоетапна система фільтрації, що дозволить знизити відсоток помилкових спрацьовувань.

2. Використання попередньої обробки листів дозволить знизити навантаження на сервіс категоризації тексту.

РОЗДІЛ 3
ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ СИСТЕМИ КОНТЕНТНОЇ ФІЛЬТРАЦІЇ ЕЛЕКТРОННОЇ КОРЕСПОНДЕНЦІЇ

3.1. Розробка архітектури спам фільтра

Архітектура програмного засобу – це його будова, тобто представлення програмної системи, що складається з деякої сукупності взаємодіючих підсистем. Як підсистеми будуть виступати програмні модулі, так як програмний комплекс має модульну структуру.

При розробці був використаний метод низхідного проектування, який полягає в тому, що спочатку будується модульна структура програми у вигляді дерева та модулі проектуються по черзі, починаючи з модуля самого верхнього (головного) рівня, а перехід до програмування будь-якого іншого модуля здійснюється лише у разі, якщо вже запрограмований модуль, який до нього звертається. Після того, як усі модулі програми запрограмовані, проводиться їхнє почергове тестування та налагодження в такому ж (низхідному) порядку. Метод низхідного проектування іноді називають функціональною декомпозицією.

Для вибору архітектури розроблюваного програмного забезпечення було проаналізована архітектура існуючих програм. В результаті було розроблена наступна модульна архітектура програми (рис. 3.1.).

Класифікація повідомлень відбувається на стороні поштового сервера за допомогою модуля формальної фільтрації та модуля контент-фільтрації (модуля фільтрації за вмістом).

Спершу вхідний потік потрапляє на модуль фільтрації по формальним ознакам. Алгоритми фільтрації цього модуля описані в 2.1.

Модуль навчання запитує каталог зі спамом та каталог з легітимною поштою, список адрес до яких належить ця пошта. Навчальна база розділена по окремих напрямках інтересів, які задає користувач (або якщо це організація то окремих відділах), що дозволяє при ідентифікації поштової кореспонденції по e-mail адресі

одержувача звертатися до бази, що характеризує інтереси конкретного користувача. Це дозволяє точніше визначати вагу термів.

Модуль попередньої обробки ЕС призначений для приведення тексту повідомлення до єдиного регістру та кодування, видалення стоп-слів, знаків пунктуації, розбиття повідомлень на терми.

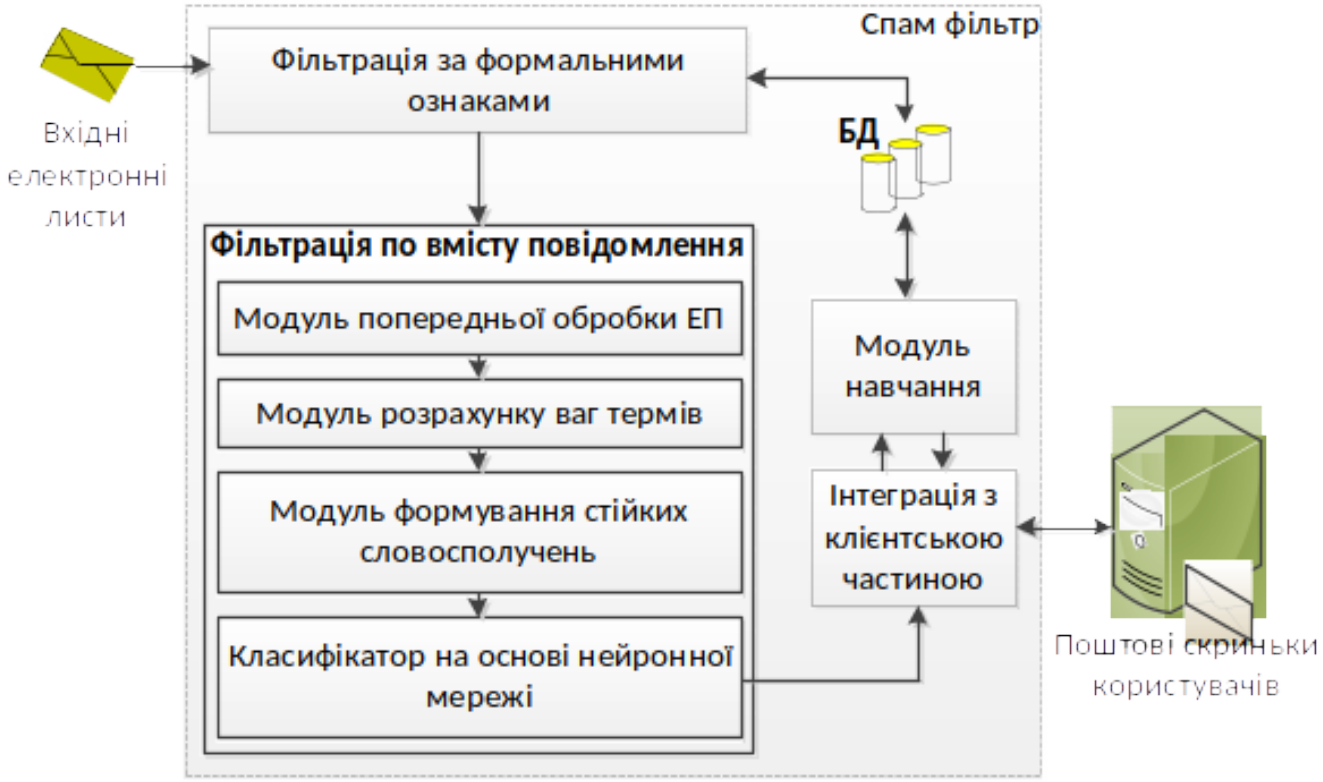

Рис. 3.1. Архітектура програми

Модуль розрахунку ваг термів дозволяє в попередньо підготовленому повідомленні здійснити розрахунок частот термів у повідомленні та в класі та визначити вагу терму в кожному повідомленні.

Модуль формування стійких словосполучень дозволяє визначити частоту входжень пов’язаних термів, підтвердити смислову значущість термів, тим самим виділити терми, що характеризують зміст електронного поштового повідомлення.

Класифікатор на основі нейронної мережі здійснює виявлення ознак спаму в повідомленні.

Модуль інтеграції з клієнтською частиною відповідає за взаємодію системи фільтрації з користувачем.

Модуль навчання проводить навчання системи фільтрації та дозволяє користувачеві вибрати та підготувати листи для навчання.

Прототип системи фільтрації, що розробляється, складається з проекту e-mail_filtering, в який входить 9 модулів, представлений на малюнку 4.2

Модулі, що реалізують основні функції, викликаються з основного модуля при виборі певного пункту меню. Допоміжні модулі використовують у процесі роботи основних модулів.

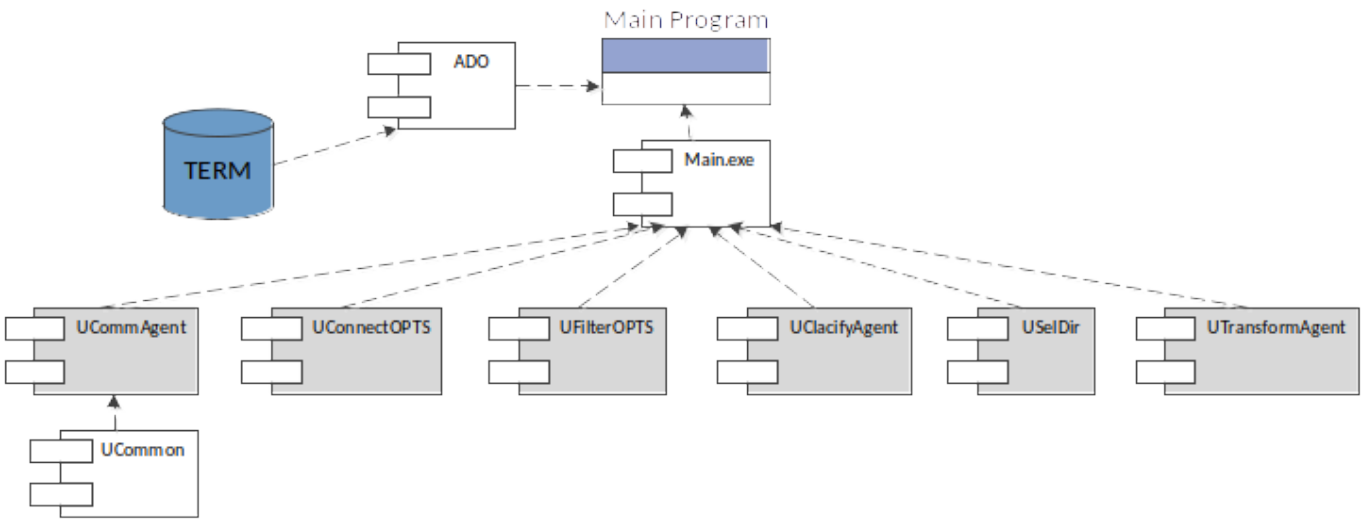

Рис 3.2. Діаграма компонентів програмного проекту системи контентної фільтрації

Дам коротку характеристику модулів програми.

UMain. Модуль служить для реалізації головного вікна програми, створення та анімації значка у системному дереві. Зберігає компонент Proxy-сервера, який забезпечує взаємодію між POP-сервером електронної пошти та клієнтською програмою доставки пошти.

UClassifyAgent. Визначає належність листа до відповідної категорії. Реалізовано методи навчання системи.

UCommon. Модуль, в якому зберігаються описи типів, константи, процедури та функції, загальні для всієї системи. У ньому реалізовано завантаження та збереження налаштувань.

UCommAgent. Модуль реалізує взаємодію програми та бази даних, де зберігається інформація про нейрони системи.

UConnectOpts. Надає інтерфейс користувача для налаштування підключення до сервера електронної пошти та управління базою даних MySQL.

UFilterOpts. Модуль реалізує діалогове вікно, що надає інтерфейс користувача для налаштування самої програми. У ньому задається поріг чутливості системи, особливості обробки повідомлень.

USelDir. Модуль, що реалізує діалог для вибору каталогу.

UTransformAgent. Модуль реалізує агента перетворення, який перетворює вхідне повідомлення у вигляд, зручний для аналізу агента класифікації.

3.2. Розробка алгоритмів нейромережевого класифікатора

Оскільки задача фільтрації вхідних електронних поштових повідомлень є задачею класифікації, то з метою вирішення цієї задачі були проаналізовані методи класифікації та зроблено висновок, що найбільш ефективним для фільтрації електронних повідомлень є нейронні мережі.

У процесі вирішення задачі фільтрації електронних поштових повідомлень було визначено, що не завжди за змістом електронного повідомлення можна визначити, чи текст повідомлення є «новим», чи це повідомлення є модифікованим варіантом «старого» повідомлення. Така ситуація знайшла свій розвиток теорії адаптивного резонансу нейронних мереж типу ART, і отримала назву проблеми стабільності-гнучкості. Тобто коли сприйняття нового образу гнучке – адаптоване до нової інформації, та при цьому стабільне – не руйнує пам'ять про старі образи.

В даний час існують такі види нейронних мереж типу ART:

1) ART 1 - нейронна мережа, призначена для обробки двійкових векторів;

2) ART 2 і ART 2a – нейронні мережі, що дозволяють працювати як з двійковими, так і з аналоговими векторами;

3) ART 3 – нейронна мережа, призначена для моделювання часових, хімічних та біологічних процесів;

4) ARTMAP – комбінація двох нейронних мереж;

5) FuzzyART – гібридна мережа створена на основі нечіткої логіки та ART мереж.

Базова архітектура мереж ART складається з трьох видів нейронів: вхідних нейронів (шари F0 і F1), що шару розпізнавання F2 і нейрона керування R.

Шар F0 - вхідний шар приймає вхідні вектори для подальшої обробки.

Шар F1 – інтерфейсний шар здійснює додаткову обробку вхідного образу та передає вхідний вектор для класифікації шару F2.

Шар F2 – шар розпізнавання, призначений для формування нейрона (у процесі навчання) відповідального за певний зразок відповідного класу та визначення нейрона, що має максимальний резонанс.

Нейрон управління – нейрон, який відповідає за прийняття рішення про результати класифікації на основі визначення міри подібності вхідного вектора та образу, що зберігається у пам'яті нейронної мережі.

Базова архітектура нейронної мережі ART представлена на рис. 3.3.

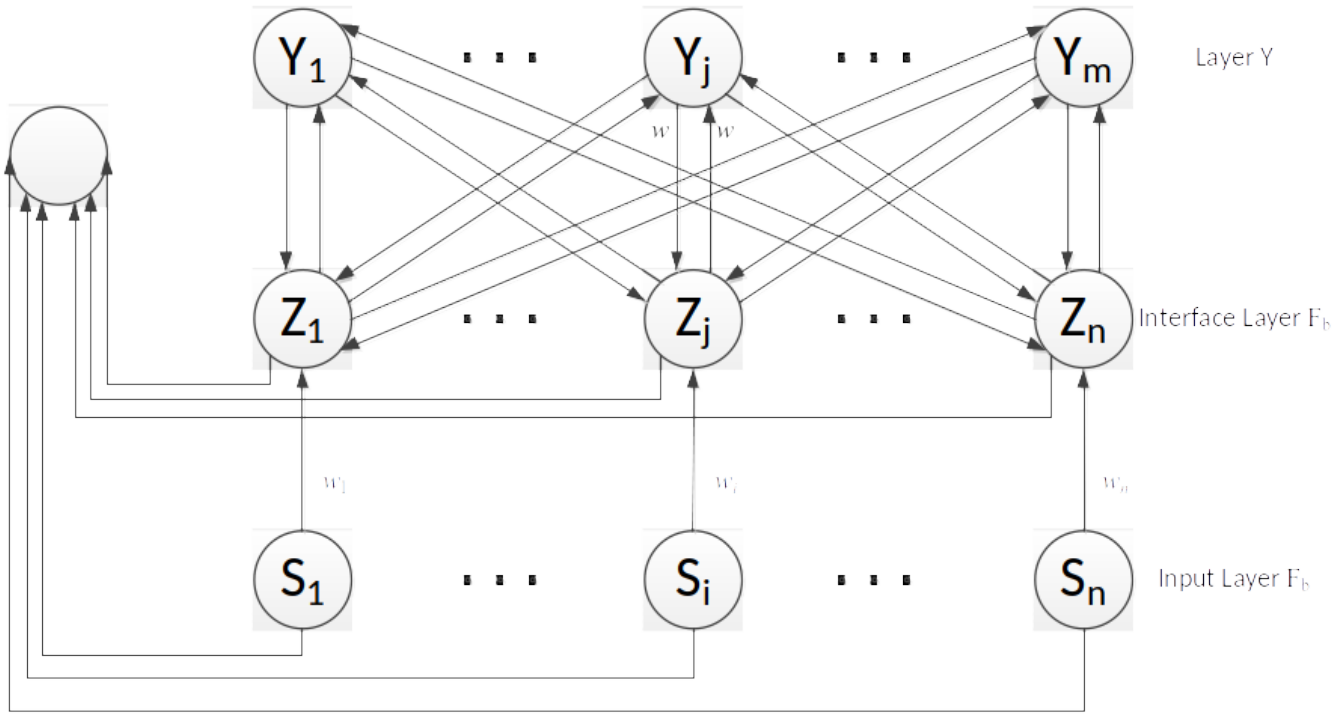

Рис.3.3. Основні компоненти нейромережевого класифікатора ЕП

Кожен нейрон шару F_1 пов'язаний з кожним нейроном шару F_2 та нейроном управління R.

Нейронна мережа ART2a поєднує шар F0 і F1 за рахунок чого швидкість навчання у даної нейронної мережі вище ніж у ART2, також обробка вхідного вектора відбувається в два-три рази швидше. Таким чином, ART2 може використовуватися для тих завдань при вирішенні яких застосування ART2 не дає бажані результати, але нейронна мережа ART2 може ефективно замінювати ART2a в більшості розв'язуваних завдань.

Алгоритм нейронної мережі, у загальному вигляді, складається з наступних кроків:

Стадія попередньої обробки вхідного вектора (вхідний шар)

- Вхідний вектор S_i проходить стадію попередньої обробки та подається на вхід нейронної мережі.

Для кожного вектора S_i з навчальної вибірки обчислюється:

$$I = X' \cdot X'',$$

де

$$X' = \frac{x}{\|x\|},$$

$$X'' = \begin{cases} x_i, \text{якщо } x_i \geq \theta \\ 0, \text{в інших випадках} \end{cases},$$

θ приймає значення $0<\theta \leq 1/\sqrt{M}$ де M – кількість елементів в векторі.

Шар активації F2

Для кожного j-го нейрону шару F_2 виконується перевірка

$$T_j = \begin{cases} \alpha \sum I x_i, \text{якщо } j \text{ нейрон не задіяний} \\ I \cdot z_j, \text{якщо } j \text{ нейрон задіяний} \end{cases}$$

Спочатку всі нейрони шару F_2 незадіяні. Параметр α є константою $\alpha \leq 1/\sqrt{M}$. причому z_j- вага зв'язку від нейрона вхідного шару до нейрона шару F_2. Крім того, згідно з методологією нейронної мережі ART2a значення α має бути достатньо

малим, щоб у тому випадку, якщо $z_j = I$ для деяких векторів, то при пред'явленні цього вектора нейронної мережі *j* нейрон повинен бути обраний.

Вибір зразка з найбільшим значенням відповідності

Знаходиться реакція кожного нейрона шару F_2 із вхідним вектором. Потім вибирається нейрон із максимальною реакцією

$$T_j = max\,(T_j).$$

Якщо нейронів з максимальною реакцією більше одного, то вибирається нейрон випадковим чином.

Порівняння з порогом (нейрон управління)

Поріг *p* показує, наскільки повинен вхідний сигнал збігатися з одним із збережених зразків, щоб вони вважалися схожими

$$T_j^{max} \geq p.$$

Близьке до одиниці значення порогу вимагає майже повної відповідності вхідного образу та образу який зберігається в пам'яті нейронної мережі. Якщо значення відношення менше встановленого порогу, то вважається, що вхідний вектор відрізняється від образу, що зберігається в нейронній мережі, і відбувається пошук іншого нейрона. Якщо вхідний вектор відрізняється від усіх образів, він розглядається як новий зразок.

Навчання

Після представлення нейронної мережі вхідного вектора z_j змінюється, так що

$$z_j^{¿} = \begin{cases} I, \text{якщо нейрон не задіяний} \\ \beta\varphi + (1-\beta)\, z_j, \text{якщо нейрон задіяний} \end{cases},$$

де $\varphi = \begin{cases} I, \text{якщо } z_j > \theta \\ 0, \text{в інших випадках} \end{cases}$. Параметр β приймає значення від 0 до 1.

Для адаптації нейронної мережі ART2a до вирішення задачі ідентифікації електронних поштових повідомлень до її алгоритму роботи були внесені наступні зміни:

1) Змінено попередню обробку вхідного вектора (стадія попередньої обробки розроблена для того, щоб вхідний вектор відповідав основній вимозі нейронної мережі, що пред'являється до вхідних сигналів: вхідний вектор повинен залишатися

незмінним без можливості скидання своїх параметрів до того, як внутрішній шар розпізнавання не стане активним і не почне з ним працювати).

Відповідно до методології нейронної мережі ART2a стадія попередньої обробки використовується для зменшення шуму у вхідному векторі, тобто виділення малоінформативних елементів та відповідно скорочення їх кількості. Для вирішення задачі фільтрації електронних повідомлень дана стадія була замінена запропонованою у розділі 2 методикою формування стійких словосполучень.

2) Змінено структуру нейронної мережі ART2a додаванням додаткового нейрона управління.

Для виключення помилкового розпізнавання легітимного ЕП в нейронну мережу в доповнення до нейрона управління R, що забезпечує обчислення скалярного твору векторів, введений додатковий нейрон управління $R_{дод}$ для визначення міри подібності по коефіцієнту Жаккара виду:

$$K_J = \frac{c}{a+b-c},$$

де a – кількість термів у вхідному повідомленні;

b – кількість еталонних термів, що зберігається в основі,

c – кількість загальних термів, що зустрічаються в 1-му та 2-му повідомленні.

Структура нейронної мережі з введеним додатковим керуючим нейроном набуде вигляду, представленого на рис. 3.4.

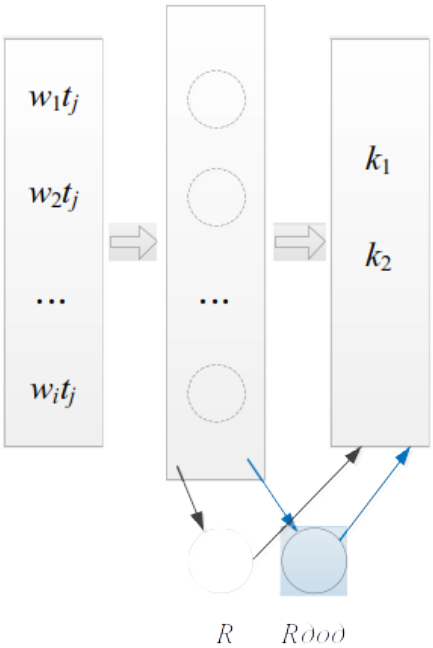

Рис. 3.4. Структура модифікованого класифікатора ART 2a

Внесені зміни не суперечать основним теоремам нейронної мережі АРТ:

1 Після досягнення стабільного стану навчання пред'явлення одного з навчальних векторів відразу призводитиме до правильної класифікації без фази пошуку, на основі прямого доступу.

2 Процес пошуку стійкий.

3 Процес навчання стійкий.

4 Процес навчання скінченний. Навчання буде завершене для заданого набору образів за кінцеве число ітерацій, при цьому подальше пред'явлення цих образів не викликає циклічних змін значень ваг.

Модифікований алгоритм нейромережевої класифікації прийме вид, представлений малюнку 3.5

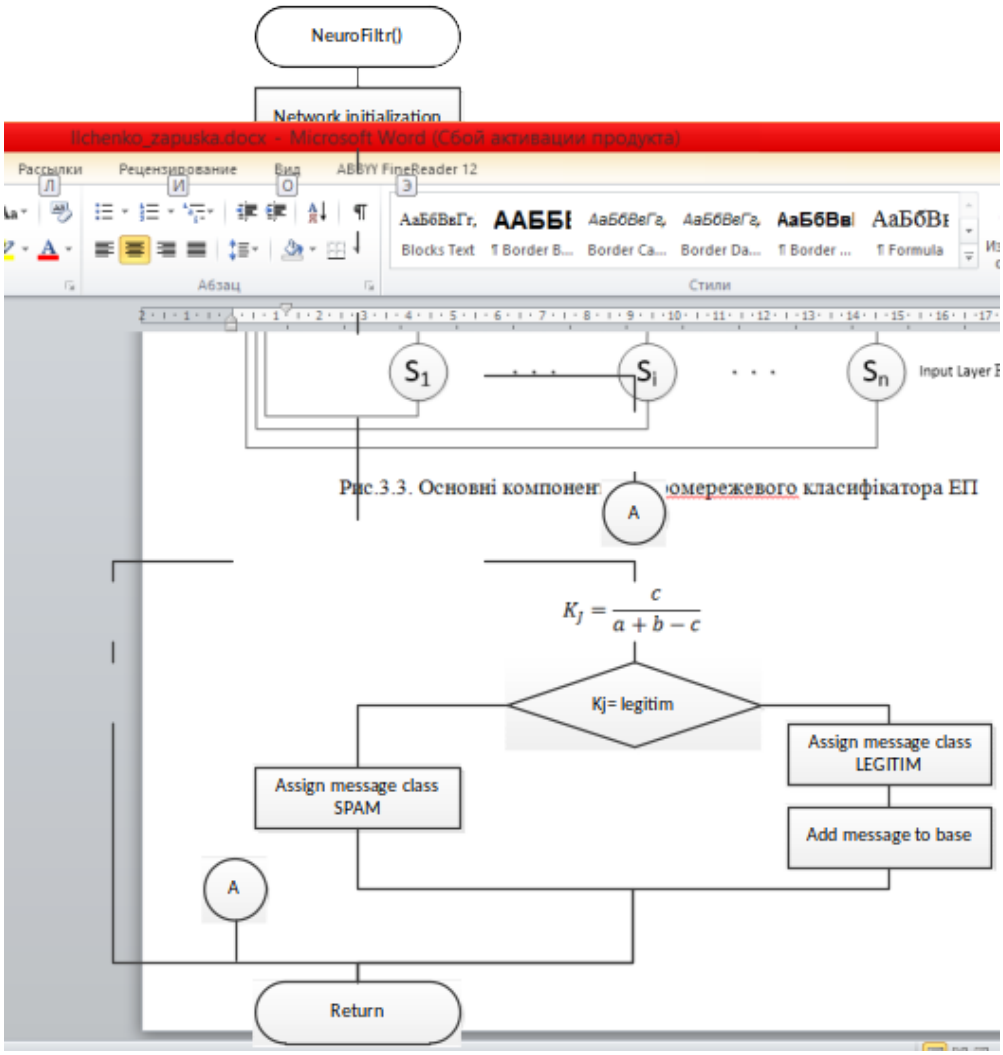

Рис. 3.5 Укрупнений алгоритм нейромережевої класифікації

Функціональна схема будується з метою однозначного розуміння всіх функцій, виконуваних цією системою. Найчастіше функціональна специфікація формулюється природною мовою з допомогою спеціальних об'єктів і тверджень, які безпосередньо описують функції програмного засобу.

Розроблена функціональна схема ПС представлена малюнку 4.14. ПС підтримує режим адміністрування, що дозволяє проводити навчання та тестування системи фільтрації.

Розроблена програмна система (ПС) дозволяє вирішувати такі завдання: фільтрація вхідного потоку електронних повідомлень, налаштування фільтрації електронних повідомлень, налаштування підключень фільтра, навчання фільтра на спамних та легітимних повідомленнях, тестування налаштованого та навченого

фільтра, ведення бази термів та легітимних повідомлень, введення та оновлення інформації в базі даних, формування вихідних.

3.3. Дослідження роботи системи контентної фільтрації електронної пошти

Предмет дослідження дозволяє виконати експеримент на основі реальних електронних повідомлень, частину яких можна віднести до несанкціонованих розсилок.

Завдання експериментального дослідження:

1) перевірка дослідно-експериментальним шляхом ефективності використання запропонованої моделі та методу нейромережевої фільтрації спаму;

2) розробка рекомендацій для подальших досліджень.

Експериментальний набір даних складається з електронних повідомлень, що являють собою спам-розсилку та легітимних листів. Вибірка спам тематики складається із листів отриманих протягом одного місяця після попереднього вимкнення фільтрації. Легітимні листи отримано із загальнодоступних ресурсів та листів отриманих від певних користувачів.

В основу обґрунтування порогу відповідності вхідного ЕП певному класу покладено метод теорії статистичних рішень для завдання перевірки двоальтернативної гіпотези H_0 та H_1, що виражають припущення про легітимність ЕП або наявність спаму.

Змінні, що використовуються:

N_{sp} – кількість об'єктів, які належать до класу спам;

N_l – кількість об'єктів, які належать до класу легітимних повідомлень;

FN_{sp} – кількість спам листів, класифікованих як легітимний лист;

FP_l – кількість легітимних листів, класифікованих як спам.

Кількість об'єктів, що відносяться до класу спам, і кількість об'єктів, що належать до класу легітимних повідомлень, у сумі повинні відповідати загальній кількості об'єктів в експериментальній вибірці, тобто.

$$N = N_l + N_{sp}.$$

Нехай висунуто гіпотезу *H0* про те, що програмна система правильно класифікує легітимні розсилки або розпізнає спам.

Тоді, кількість хибних пропусків FN_{sp} і кількість хибних виявлень *FPl* визначать кількість правильно класифікованих легітимних ЕП TP_l та вірних виявлень спаму TN_{sp} виду:

$$TP_l = N_l - FP_l,$$

$$TN_{sp} = N_{sp} sp - FN_{sp}.$$

Звідси помилка першого α (ймовірність відкинути нульову гіпотезу, коли вона хибна, тобто ймовірність прийняти рішення про легітимність повідомлення, коли воно спам) та другого роду β (ймовірність відкинути нульову гіпотезу, коли вона справедлива, тобто ймовірність відкинути рішення про легітимність повідомлення, коли воно є легітимним) будуть визначатися залежностями:

$$\alpha = \frac{FN_{sp}}{N_{sp}}, \ \beta = \frac{FP_l}{N_l}.$$

Ці величини характеризують якість розпізнавання, так як не залежить від кількості об'єктів у тестовому наборі.

На основі характеристик TP та TN можна розрахувати міру повноти та точності. Для найбільш наочного уявлення щодо часто оперують не абсолютними показниками, а відносними (частками) вираженими у відсотках. Міра повноти (precision) оцінює частку вірного

розпізнавання щодо всіх об'єктів певного класу. Міра точності (recall) оцінює частку вірних виявлень щодо всіх об'єктів. Ці заходи розраховують за такими формулами:

$$precision = \frac{TP}{TP + FP_l} \cdot 100\%,$$

$$recall = \frac{TP}{TP + FN_{sp}} \cdot 100\%.$$

Використовуючи таку залежність можна визначити частку хибно класифікованих об'єктів FPR відповідних класів:

$$FPR = \frac{FP_l}{TN + FP_l} \cdot 100\%.$$

Зведена оцінка якості класифікації (F-міра), що залежить від повноти та точності, визначається залежністю:

$$F = \frac{2}{1/precision + 1/recall}.$$

Проведення імітаційного експерименту дозволило вирішити такі завдання:

1 Оцінити налаштування та навчання класифікатора.

2 Визначити об'єкти класифікації, що найбільш ймовірно викликають помилки роботи класифікатора з метою корегування або розширення навчальної вибірки.

3 Навчання класифікатора з різними параметрами та оцінка результатів з метою визначення найкращих параметрів.

4 Виявлення ознак які найкраще характеризують певні класи.

У ході експерименту досліджено кілька версій системи фільтрації, представлені в таблиці 3.1. Для кожної версії систем фільтрації змінювався поріг на відповідність класу S_n. Експериментальна вибірка ЕП для оцінки ефективності прототипу системи фільтрації складалася з легітимних повідомлень та спам-розсилки. Тематика повідомлень експериментальної вибірки представлена в таблиці 3.2. Всього досліджено 908 ЕП (424 легітимних повідомлень та 484 спам-повідомлень) та здійснено 13 запусків прототипу запропонованої системи фільтрації ЕП. Поріг відповідності S_n змінювався в діапазоні від 0,4 до 0,9.

Таблиця 3.1

Варіанти побудови класифікатора

Назва	Модель ЕП	Вага	Метод скорочення простору знаків	Виділення стійких словосполучень	Алгоритм класифікації
Met1	Векторна	Tf-idf	RF	+	нейрон.МережаArt
Met2	Векторна	Ltc	IG	+	нейрон.МережаArt
Met3	Векторна	Ltc	RF	-	нейрон.МережаArt
Met4	Векторна	Ltc	IG	+	нейрон.МережаArt
Met5	Векторна	Tf-idf	RF	+	нейрон.МережаArt

Таблиця 3.2

Тематика повідомлень

Вид повідомлення	Тематика повідомлень
Спам повідомлення	“Порожні” повідомлення, що містять лише посилання або вкладення.
Спам повідомлення	Реклама товарів
Спам повідомлення	Реклама послуг (юридичних, бухгалтерських, будівельних, освітніх, туристичних, медичних та ін.)
Спам повідомлення	Запрошення на курси, пропозиції схем «відмивання» грошей («нігерійські» листи)
Легітимні повідомлення	Ділова листування (вільна форма)
Легітимні повідомлення	Ділове листування (накази, розпорядження, звіти тощо)
Легітимні повідомлення	Запрошення на участь у грантах, конференціях, виставках тощо.

Таким чином, запропонована методика експерименту дозволяє оцінити ефективність використання запропонованої моделі ЕП на основі стійких словосполучень та методу класифікації, заснованого на нейронній мережі адаптивного резонансу.

У таблиці 3.3 представлені показники ефективності версій та порівняльні результати оцінки запропонованого фільтра легітимних ЕП за різних значень порога (таблиця 3.4).

Таблиця 3.3

Показники ефективності версій системи

Показники ефективності	Назва версій				
	Met1	Met2	Met3	Met4	Met5
помилка I роду (%)	13	17	19	16	15
помилка II роду (%)	9	12	14	7	3
міра повноти (%)	96	88	85	91	98
міра точності (%)	90	87	83	93	96

F-міра (%)	80	79,9	80	81	83

Таблиця 3.4

Показники ефективності за зміни значення порогу

Показники ефективності	Значення порогу			
	Sn=0,4	*Sn*=0,7	*Sn*=0,8	*Sn*=0,9
помилка I роду (%)	13	17	19	16
помилка II роду (%)	9	12	14	7
міра повноти (%)	96	88	85	91
міра точності (%)	90	87	83	93
F-міра (%)	80	79,9	80	81

Результати проведеного експерименту представлені як діаграми рисунках 3.6-3.8

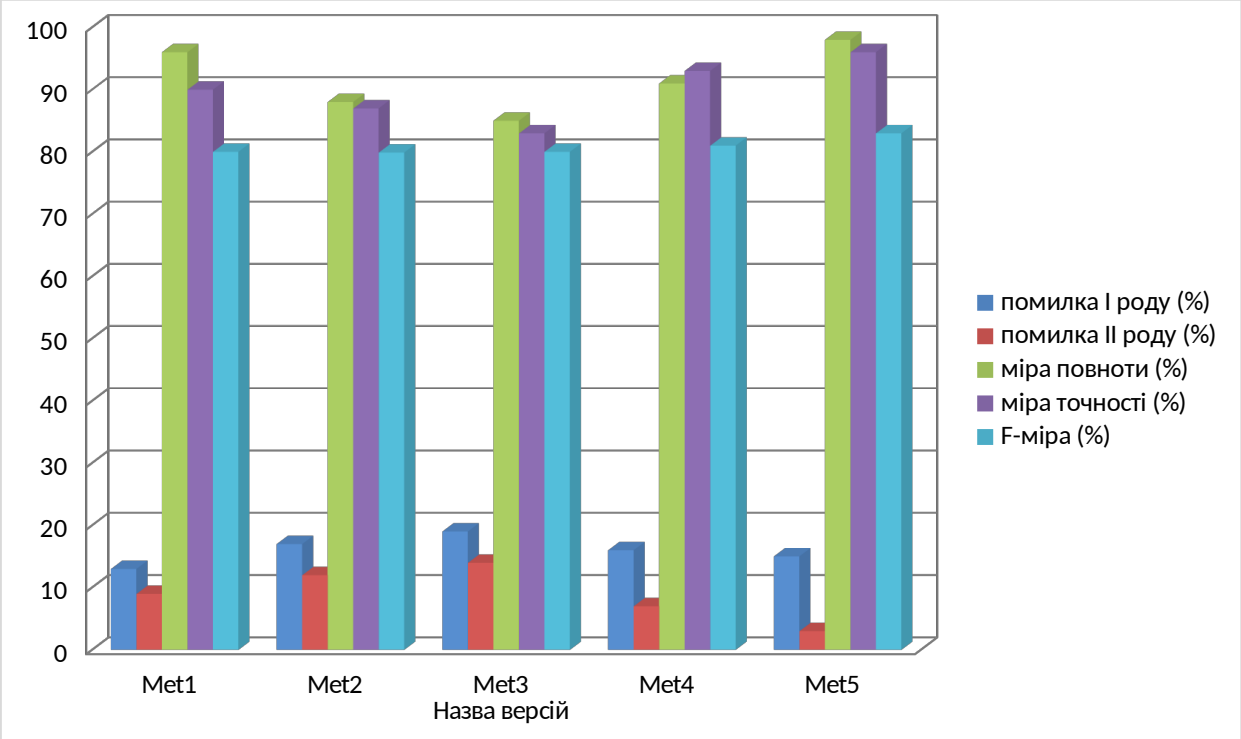

Рис. 3.6 Значення критеріїв ефективності різних версій

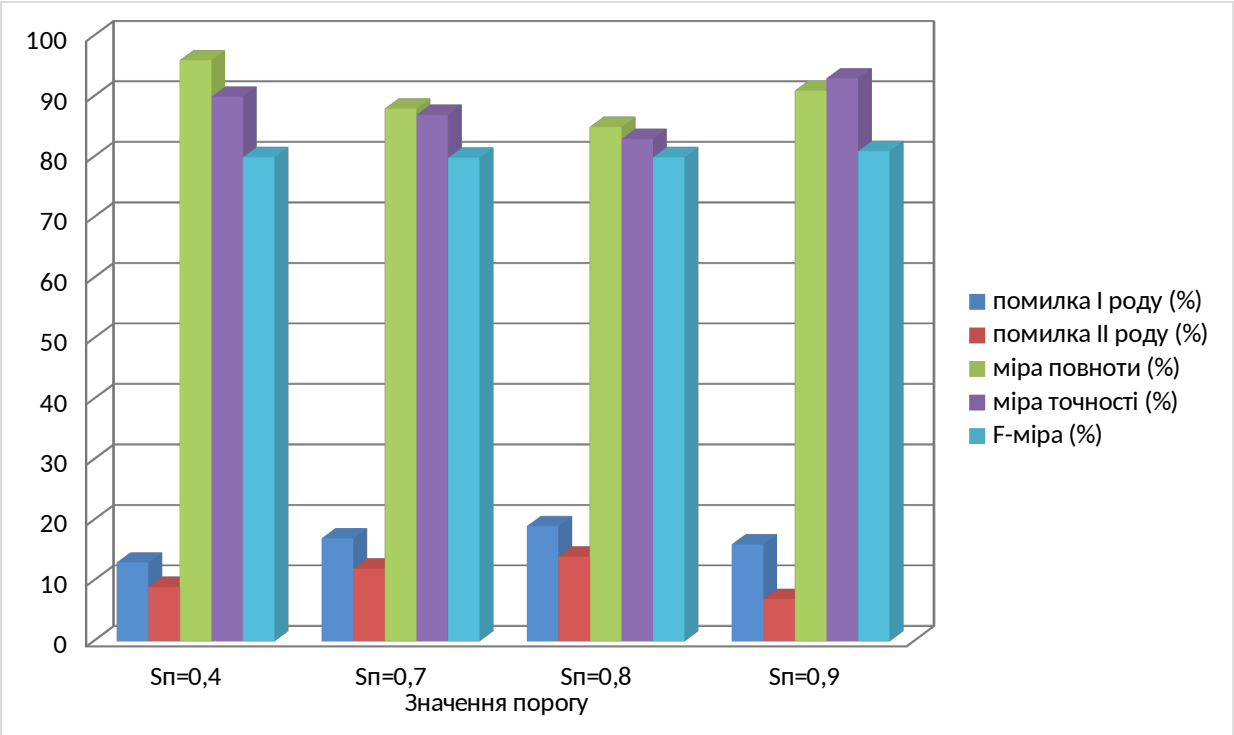

Рис. 3.7 Діаграма результатів імітаційного експерименту

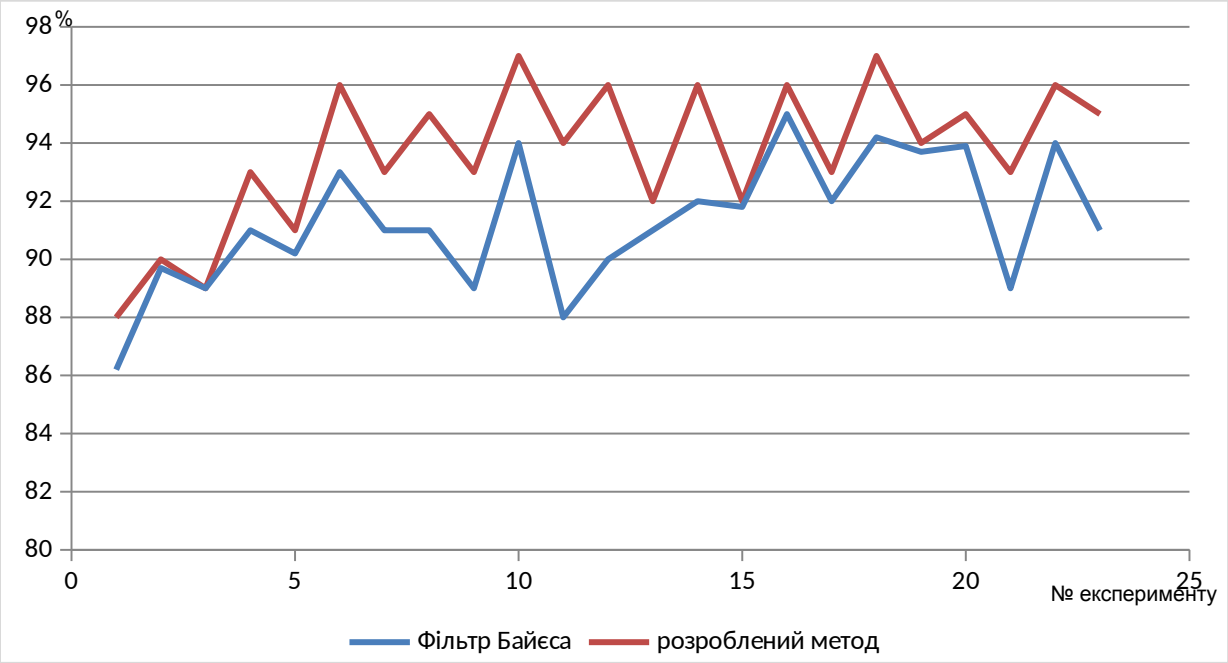

Рис. 3.8 Порівняльна оцінка ефективно сті запропонованих технічних рішень

Як видно з результатів імітаційного експерименту, найефективніша версія met5. Аналіз результатів досліджень met5 показав, що при зміні порога відповідності Sn змінюються показники якості фільтрації ЕП. При встановленні порога Sn=0,4 число легітимних повідомлень прийнятих за спам становить 7%, а

число спам-повідомлень прийнятих за легітимні становить 15%. При збільшенні порога Sn до 0,7 знижується рівень помилки 2 роду до 3%, проте рівень помилки 1 роду становить 17% і при подальшому збільшенні порога Sn продовжує зростати, що свідчить про високу вимогливість нейронної мережі (при встановленому порозі, близькому до одиниці, нейронна мережа вимагає майже повної відповідності вхідного повідомлення та прототипу що зберігається в базі). Установка порога $Sn=0{,}8$ показує найкращі результати: помилка 2 роду прямує 0 і становить 0,001, помилка 1 роду – 0,07.

Частка спаму, виявлена запропонованою системою фільтрації, вище, ніж у байесовського фільтра, при ймовірності помилкового спрацьовування трохи більше 0,05.

Таким чином, результати експерименту підтверджують досягнення поставленої мети дослідження і свідчать про ефективність розробленого прототипу системи фільтрації ЕП при порозі $Sn=0{,}8$.

Висновки до розділу.

1. Запропоновано методику та розроблено алгоритми контентної фільтрації електронної кореспонденції поштових сервісів на основі нейромережевого класифікатора АРТ2а.

2. Запропоновано використанням додаткового нейрона для перевірки повідомлень, що ідентифікуються як несанкціоновані повідомлення, мірою схожості векторів Жаккара.

3. Запропоновано прототип системи захисту поштових сервісів, що базується на дворівневій фільтрації електронних поштових повідомлень.

4. Результати експериментальних досліджень запропонованого прототипу системи захисту поштових сервісів свідчать про підвищення достовірності ідентифікації поштової кореспонденції помилки класифікації легітимних повідомлень до 0,1%, а помилки класифікації спам-розсилок до 7%

РОЗДІЛ 4
ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Усі дослідження антиспамерних методів фільтрації електронних повідомлень проводились з дотриманням правил та норм охорони праці і вимог техніки безпеки.

Сервери, в яких використовуються різного роду системи охолодження, а також робота інших системних компонентів є потенційними джерелами цілого ряду звуків, що містять як коливання, які можна почути, так і коливання ультразвукового діапазону. Цей шум справляє негативний вплив на функціональний стан користувачів.

Вимірювання шуму на робочих місцях і санітарні норми наведені в даних документах:

- ДСТУ 2867-94 (Шум. Методи оцінювання виробничого шумового навантаження. Загальні вимоги),
- ДСН 3.3.6.037-99 (Державні санітарні норми виробничого шуму, ультразвуку та інфразвуку).

Згідно з ДСТУ 2867-94 шум у приміщенні, де виконують роботу, пов'язану з створенням нових програм, викладацькою роботою, творчістю, не повинен перевищувати 40 дБ. Висококваліфікована розумова робота, що вимагає зосередженості, може проводитись у приміщеннях, де рівень шуму не перевищує 55 дБ. Сумарний вплив численних джерел шуму у приміщенні у результаті багаторазового відбиття звукових хвиль може значно перевищити енергію прямого звука від тих же джерел. Шум від окремих приладів не повинен перевищувати фоновий більше ніж на 5 дБ.

Для боротьби з шумом в серверній використані так звані «тихі» системи охолодження тобто вентилятори великому діаметру і з спеціальною формою крильчатки. Сервери встановлені в спеціальних шафах які додатково поглинають шуми і вібрації. Мережеве обладнання змонтоване в окремій серверній, що забезпечує в приміщенні де виконується наукова робота рівень шуму в районі 40-

45 дБ, який створюється роботою офісної техніки і ЕОМ. Тобто по рівню шумового навантаження приміщення відповідає вимогам охорони праці.

Також при роботі за ЕОМ необхідно особливу увагу звертати на правильне освітлення. Неправильне освітлення (пряма та відбита від екранів близькість, вуалюючі відбиття, несприятливий розподіл яскравості в полі зору, невірна орієнтація робочого місця відносно світлових отворів) призводить до негативних фізіологічних впливів на користувачів ЕОМ. Погана якість символів, що представлені на екрані, також може викликати зоровий дискомфорт, бути стресовим фактором та ін.

Вимоги до освітлення для візуального сприймання користувачами інформації з двох різних носіїв (з екрана ЕОМ та паперового носія) різні. Надто низький рівень освітленості погіршує сприймання інформації при читанні документів, а надто високий призводить до зменшення контрасту зображення знаків на екрані. Відношення яскравості екрана ЕОМ до яскравості оточуючих його поверхонь не перевищує у робочій зоні 3:1.

Наближено можна вважати, що при 10%-ному зменшенні освітленості працездатність знижується на 1%. Коли за характером роботи вимагається комбінація цих двох носіїв інформації, освітленість можна варіювати від 300 до 700 лк, причому чим рідшою є зміна полів зору в процесі роботи (з екрана на документ та навпаки), тим вищим може бути рівень освітленості. 300-500 лк — оптимальна освітленість робочих приміщень для роботи з ЕОМ. Стрибки яскравості при зміні полів зору мають бути мінімальними, тобто інтенсивність освітлення поверхні, де знаходяться рукописи та документи, не повинна перевищувати яскравості екрана дисплея.

Освітлення повинно відповідати нормальним рівням за ДБН В.2.5-28:2018 Природне і штучне освітлення.

Приміщення в якому виконувалась дипломна робота забезпечене природнім і штучним освітленням. При роботі за ЕОМ обрано місце, щоб в поле зору не потрапляли вікна або освітлювальні прилади. Крім того шкідливо коли вікна знаходяться за спиною оскільки на моніторі з’являється відбиття світла. Завдяки

наявним жалюзі можна регулювати світловий потік і захистити робоче місце від попадання прямих сонячних променів. Адже вікна приміщення орієнтовані на південний схід.

Штучне освітлення у приміщені реалізовано у вигляді комбінованої системи освітлення з використанням люмінесцентних джерел світла у світильниках загального освітлення, які слід розташовані над робочими поверхнями у рівномірно-прямокутному порядку. Для запобігання засвітленню екранів ЕОМ прямими світловими потоками лінії світильників розташовані з достатнім бічним зміщенням відносно робочих місць, а також паралельно до вікон.

На робочому місці забезпечена рівномірна освітленість за допомогою переважно відбитого або розсіяного світлорозподілу. Світлових відблисків з клавіатури, екрана та від інших частин ЕОМ у напрямку очей користувача немає. Дискомфорт від відбиття світла знижується при збільшенні яскравості екрана та зниженні рівня навколишнього освітлення.

Пульсація освітленості люмінесцентних ламп, що використовуються, відповідно до технічної документації 10%, що відповідає діючим вимогам.

Інформація, яку одержує користувач, генерується на екрані, а комфортність її сприймання залежить від чіткості символів. При обговоренні проблеми дискомфорту або негативних наслідків для здоров'я та ефективності роботи на ЕОМ слід враховувати ряд параметрів. Ці параметри поділені на три групи, пов'язані з мигтінням, структурою та яскравістю символів, що представляються на екрані.

На робочих місцях користувачів використовуються сучасні ноутбуки та ПК обладнані рідкокристалічними моніторами. Завдяки використанню IPS матриць з частотою оновлення не менше 60 Гц очі практично не втомлюються. Крім того IPS матриці забезпечують хорошу кольоропередачу, а матове покриття екрану виключає відблиск сторонніх джерел світла. Роздільна здатність моніторів 1920×1080 або 1280×720 точок що забезпечує високу чіткість зображення. Ще одною важливою перевагою даних моніторів це практично відсутнє опромінення користувача.

Отже в даному підрозділі розглянуто вплив середовища на працездатність та здоров’я користувачів комп’ютерів. Як висновок можна сказати, що робоче місце

яке використовувалось для написання даного наукового дослідження відповідає вимогам з охорони праці.

Однак необхідно не забувати що надмірна робота з ПК може привезти до порушення роботи організму користувача. Тому необхідно дотримуватись вимог щодо планування робочого часу за ЕОМ.

4.2. Оцінка стійкості роботи об’єкту економіки до впливу ударної хвилі ядерного вибуху

Основні вражаючі фактори ядерного вибуху - це ударна хвиля, світлове випромінювання, проникаюча радіація, радіоактивне зараження місцевості, електромагнітний імпульс.

Ударна хвиля - основний вражаючий фактор ядерного вибуху. Більшість руйнувань і ушкоджень споруд, будинків, а також поразки людей, як правило, обумовлені її впливом. Джерело її виникнення - величезне тиск, що утворюється в центрі вибуху і досягає у перші миті мільярдів атмосфер. Передня межа стисненого шару повітря називається фронтом ударної хвилі. Ступінь поразки ударною хвилею людей і різних об'єктів залежить від потужності і виду вибуху, а також від відстані, на якому стався вибух, рельєфу місцевості і положення об'єктів на ній.

Швидкість руху і відстань, на яку поширюється ударна хвиля, залежать від потужності ядерного вибуху. Зі збільшенням відстані від місця вибуху швидкість швидко падає. Так, при вибуху боєприпасів потужністю 20 кт ударна хвиля проходить 1 км за 2 с; 2 км - за 5 с, 3 км - за 8 с. За цей час людина після спалаху може укритися й тим зменшити ймовірність ураження ударною хвилею або взагалі уникнути поразки.

Стійкість роботи об'єкту економіки – це здатність його в надзвичайних ситуаціях випускати продукцію у запланованому обсязі, необхідної номенклатури і відповідної якості, а у випадку впливу на об'єкт вражаючих факторів, стихійних лих та виробничих аварій – в мінімально короткі строки відновити своє виробництво.

Більш підготовленими до стійкої роботи будуть ті об'єкти економіки, які реально оцінять фактори, їх несприятливий вплив на виробництво і розроблять

відповідні заходи. Завчасне проведення організаційних, агрохімічних, агротехнічних, інженерно-технічних, ветеринарно-санітарних, лісотехнічних, лісогосподарських, меліоративних та інших заходів максимально знизить результати впливу вражаючих факторів мирного і воєнного часу на людей, сільськогосподарських тварин і створить сприятливі умови для швидкої ліквідації наслідків надзвичайної ситуації.

Для розробки заходів підвищення і забезпечення стійкості роботи об'єктів у надзвичайних ситуаціях необхідно оцінити стійкість об'єкту проти випливу вражаючих факторів.

Вихідними даними для проведення розрахунків стійкості об'єкта до ураження є:

- максимальні значення параметрів можливих вражаючих факторів,
- характеристики елементів об'єкта.

Дія ударної хвилі на об'єкт характеризується складним комплексом навантажень:

- надмірним тиском,
- тиском відбивання,
- тиском швидкісного напору,
- тиском затікання.

Все це буде залежати від виду і потужності вибуху, відстані до об'єкта, конструкції й розмірів елементів об'єкта, орієнтації відносно вибуху, розміщення будівель і споруд, рельєфу місцевості. Врахувати їх разом для кожного об'єкта неможливо.

Тому опір конструкцій до дії вибухової хвилі прийнято характеризувати надмірним тиском у фронті ударної хвилі який призводить до слабких, середніх і сильних руйнувань.

Для оцінювання стійкості об’єкта до дії повітряної ударної хвилі необхідно провести розрахунок значення надлишкового тиску який є основним критерієм оцінки стійкості об’єкта до дії ударної хвилі (УХ).

Критерієм стійкості об’єкта до дії УХ є граничне значення надлишкового тиску, за якого елементи об’єкта зберігаються або отримують слабкі та середні руйнування. Це значення надлишкового тиску називають границею стійкості об’єкта до УХ і позначають $\Delta P_{фгран.}$

Стійкість об’єкта оцінюють для екстремальних умов.

Умови стійкості об’єкта такі:

- якщо $\Delta P_{фmax.} \geq \Delta P_{фгран.}$ – об’єкт нестійкий,
- якщо $\Delta P_{фmax.} < \Delta P_{фгран.}$ – об’єкт стійкий до дії УХ

Методика оцінювання стійкості об’єкта до дії УХ включає:

- розрахунок максимального значення надлишкового тиску УХ, що очікується в районі об’єкту $\Delta P_{фmax.}$;
- розрахунок границі стійкості об’єкту до дії УХ, $\Delta P_{фгран.}$ Спочатку з технічної документації виділяють основні елементи об'єкта і їх характеристики. Потім визначається межа (границя) стійкості кожного з основних елементів об'єкта. Границею стійкості елемента є надмірний тиск, при якому елемент дістане середню ступінь зруйнувань. Якщо надмірний тиск, при якому елемент отримує середні руйнування, визначений не одним значенням, а діапазоном (наприклад, 20...30 кПа), то за границю стійкості приймають нижню межу діапазону (у прикладі 20 кПа). За границю стійкості об'єкта в цілому приймають границю стійкості найбільш слабкого елемента об'єкта;
- аналіз результатів оцінювання: висновок – чи стійкий об'єкт чи ні. Які з елементів найменш стійкі. До якої величини доцільно підвищувати стійкість об'єкта;
- визначення заходів щодо підвищення стійкості об'єкту.

Висновки.

Таким чином в даному підпункті розглянуто питання оцінки стійкості об’єктів економіки до дії ударної хвилі ядерного вибуху. На основі результатів оцінки стійкості об'єкта роблять висновки і заходи по кожному елементу і об'єкту в цілому. Такими заходами можуть бути:

- укріплення несучих конструкцій та перекриттів будівель установленням додаткових колон, ферм, контрфорсів або підкосів;

- розміщення обладнання на нижніх поверхах будівель або в підвалах, надійне закріплення на фундаменті, установлення захисних кожухів або ковпаків;
- прокладання кабельних мереж та трубопроводів під землею;
- створення резервних запасів обладнання, апаратури, матеріалів для відновлення виробництва.

ВИСНОВКИ

В результаті виконання дипломної роботи досягнуто поставлену мету і розв’язано поставлені задачі:

1. Проведений аналіз методів розсилки і головних ознак спаму дозволив визначити критерії по яким ефективно можна розпізнати спам розсилки.

2. Аналіз технічних особливостей розповсюдження спаму показав, що найбільш перспективним є розвиток методів фільтрації електронних повідомлень на основі моделей які відображають семантику легітимної поштової кореспонденції.

3. Визначені основні ознаки електронних поштових повідомлень, які необхідні для класифікації електронних розсилок.

4. Розроблено модель електронного повідомлення у формі стійких словосполучень, яка дозволяє без втрати змістової інформації забезпечити класифікацію легітимної електронної кореспонденції.

5. Застосування міри ваги термів дозволило уникнути великих відмінностей у частотах фіксації термів. Визначення і видалення термів з малою інформативною вагою і виділенням стійких словосполучень, дозволяють посилити смисловий зміст термів і скоротити простір ознак на 25%.

6. Запропоновано методику та розроблено алгоритми контентної фільтрації електронної кореспонденції поштових сервісів на основі нейромережевого класифікатора.

7. Запропоновано прототип системи захисту поштових сервісів, що базується на дворівневій фільтрації електронних поштових повідомлень, що відрізняється попередньою підготовкою повідомлень до нейромережевої класифікації та забезпечує контентну фільтрацію легітимної кореспонденції.

8. Результати експериментальних досліджень запропонованої системи захисту поштових сервісів свідчать про підвищення достовірності ідентифікації поштової кореспонденції з мінімальними помилками класифікації легітимних повідомлень до 0,1%, а помилки класифікації спам-розсилок до 7%.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. An introduction to machine learning with scikit-learn [Електронний ресурс] Режим доступу: https://scikit-learn.org/stable/tutorial/basic/tutorial.html

2. Desktop macOS Version Market Share Worldwide. [Електронний ресурс] Режим доступу: http://gs.statcounter.com/os-version-market-share/macos/ desktop/ worldwide #monthly-201902-201902-bar

3. Desktop Operating System Market Share Worldwide. [Електронний ресурс] Режим доступу: http://gs.statcounter.com/os-market-share/desktop/worldwide/#monthly-201902-201902-bar

4. Desktop Windows Version Market Share Worldwide. [Електронний ресурс] Режим доступу: http://gs.statcounter.com/os-version-market-share/ windows/desktop /worldwide#monthly-201902-201902-bar

5. Does your PC have a good rep?. [Електронний ресурс] Режим доступу https://www.cnet.com/news/does-your-pc-have-a-good-rep-to-send-e-mail-it-better/

6. Eliminate Spam. [Електронний ресурс] Режим доступу http://awildduck.com/?p=277

7. Garreta R. Learning scikit-learn: Machine Learning in Python / Raúl Garreta Packt Publishing 2013 108 ст.

8. Lutz M. Learning Python, 5th Edition / Mark Lutz O'Reilly Media 2013 -648 ст.

9. NLTK Workbook [Електронний ресурс] Режим доступу: http://www.nltk.org/book/ch06.html

10. Richert W. Building Machine Learning Systems with Python / Willi Richert Packt Publishing 2013 290 ст.

11. Закон Ципфа. [Електронний ресурс]. – Режим доступу: https://www.wolfram.com/language/11/text-and-language-processing/zipfs-law.html.ru

12. Ільченко Д.О, Жаровський Р.О. Методи фільтрації спаму в сучасних поштових системах. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі системи та технології» (7-8 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 110.

13. Ільченко Д.О, Жаровський Р.О. Семантичні методи фільтрації спаму. Матеріали IX науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі системи та технології» (7-8 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 111.

14. Классификация с помощью мешка слов. Руководство [Електронний ресурс] Режим доступу: http://datareview.info/article/klassifikatsiya-tekstov-s-pomoshhyu-meshka-slov-rukovodstvo/

15. Конспект лекцій з дисципліни «Захист інформації у комп'ютерних системах» для студентів спеціальності 123 “Комп'ютерна інженерія”/ укл. Жаровський Р.О. – Тернопіль: ТНТУ ім. І. Пулюя, 2019 р. - 268 с.

16. Мезенцева, Е.М. Защита компьютерных сетей. Веб программирование многомодульного спам фильтра / Е.М. Мезенцева, В.Н. Тарасов // Программная инженерия. - 2012,- № 4.- С. 27-32

17. Методы автоматической классификации текста. [Електронний ресурс] Режим доступу: http://www.swsys.ru/index.php?page=article&id=425

18. Сервіс Whois. [Електронний ресурс]. – Режим доступу: https://www.ukraine.com.ua/uk/domains/whois-service/

19. Синтаксический анализ в NLTK. [Електронний ресурс] Режим доступу: https://habr.com/ru/post/340574/

20. Спам и фишинг в 2018 [Електронний ресурс] Режим доступу: https://securelist.ru/spam-and-phishing-in-2018/93453/

21. Способи поширення спаму. [Електронний ресурс]. – Режим доступу: http://www.refine.org.ua/pageid-5411-1.html – Дата доступу: 01.12.2021.

Додаток А.

Тези конференцій

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

IX НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

8–9 грудня 2021 року

ТЕРНОПІЛЬ
2021

УДК 004.658.114

Д.О. Ільченко, Р.О. Жаровський канд. техн. наук
(Тернопільський національний технічний університет ім. Івана Пулюя, Україна)

МЕТОДИ ФІЛЬТРАЦІЇ СПАМУ В СУЧАСНИХ ПОШТОВИХ СИСТЕМАХ

UDC 004.658.114

D. Ilchenko, R. Zharovskyi Ph.D.

SPAM FILTERING METHODS IN MODERN MAIL SYSTEMS

Спам – масове розсилання кореспонденції рекламного чи іншого характеру людям, які не висловили бажання її одержувати. Передусім термін «спам» стосується рекламних електронних листів.

Розсилання спаму це одна з основних проблем яка існує в сучасному інформаційному суспільстві оскільки призводить до витрат часу на обробку вхідної кореспонденції як з боку поштових серверів так і збоку користувачів електронної пошти. За даними компанії Cisco Systems в рекордний потік спаму був зафіксований у 2016 році і становив 65 % в загальному трафіку електронної пошти.

Для боротьби з даною проблемою використовують різні методи фільтрації. Даний вид програмного забезпечення може використовуватись як на стороні сервера так і на стороні клієнта. Для створення даних фільтрів використовується два основних підходи (рис. 1).

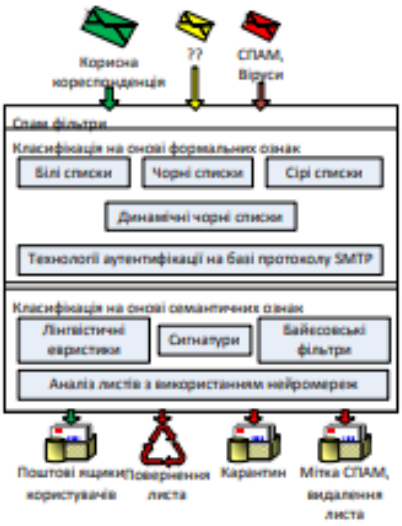

Рисунок 1. Методи фільтрації спаму

Перший підхід класифікує відправника листа як спамера, не заглядаючи в текст листа. Для визначення застосовуються різні методи. В основному даний вид фільтрів працює тільки на сервері, який безпосередньо приймає пошту. Другий полягає у аналізі тексту листа на основі якого робиться висновок, спам це чи ні.

Якщо лист класифікований як спам, він може бути позначений, переміщений в іншу папку або навіть видалений. Такі фільтри можуть працювати як на сервері, так і на комп'ютері користувача. Кожен з даних методів має свої переваги і недоліки. Тому логічним є розробка системи фільтрації спаму яка б поєднала сильні сторони зазначених вище методів.

УДК 004.658.114
Д.О. Ільченко, Р.О. Жаровський канд. техн. наук
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

СЕМАНТИЧНІ МЕТОДИ ФІЛЬТРАЦІЇ СПАМУ

UDC 004.658.114
D. Ilchenko, R. Zharovskyi Ph.D.

SEMANTIC METHODS OF SPAM FILTRATION

Задача фільтрації спаму, являє собою задачу класифікації – визначення належності об'єкта (електронного повідомлення) до одного з заздалегідь виділених класів (спам і «не спам») на підставі аналізу сукупності ознак, що характеризують електронне повідомлення.

Семантичні методи передбачають розпізнавання спам листів за змістом (словосполучення, евристики, статистика) або розпізнавання за зразками листів (за сигнатурами). Більшість семантичних методів це методи з попереднім навчанням. Тому необхідно провести їх початкове навчання, тобто задати базу нормальних і спам листів.

Теорема Байєса покладена в основі багатьох сучасних систем, призначених для роботи в умовах невизначеності. Такі системи дають ймовірнісну оцінку, тому звичайно не заміняють експерта, а лише забезпечують підтримку прийняття рішення.

Віднесення листа до спаму або корисних листів виконується з врахуванням заданого програмістом, адміністратором, користувачем параметру «спамерності» електронного листа. Після ухвалення рішення щодо класифікації листа в базі даних обновляються ймовірнісні бази для слів, які входять до нього, тобто з кожним новим листом фільтр вдосконалюється. В основі фільтра лежить список ознак, за якими проводиться аналіз повідомлення і обчислюється умовна ймовірність спамності за кожного ознакою. Загальна ймовірність спаму повідомлення визначається за одним з методів:

1) об'єднуються всі ймовірності за теоремою Байєса;

2) ймовірності комбінуються і перевіряються на скільки отримана множина схожа з випадковою (метод Фішера).

Основним недоліком Байєсівського фільтра є припущення, що події, які відповідають наявності того чи іншого слова в електронному листі або повідомленні, є незалежними по відношенню один до одного, тобто всі слова статистично незалежні. Максимальний результат, досягнутий байєсовськими фільтрами складає близько 95% відфільтрованого спаму.

Існують модифікації, які дозволяють збільшити ефективність фільтра:

- метод градуйованої фільтрації «спаму», який забезпечує підвищення якості оцінок даних за рахунок врахування додаткових параметрів – кількості листів, в яких зустрічалися слова певної категорії, використання слів, що вперше зустрілися в листі і не існували до цього в базі, частоти використання слів у листах певної категорії,;
- побудова фільтра на основі багатошарового персептрона, що дозволяє враховувати семантичні зв'язки автоматично.

Перевага нейромережевого підходу перед Байєсовським класифікатором полягає в тому, що не робиться ніяких попередніх припущень про характер небажаних повідомлень, а семантичні зв'язки враховуються автоматично. Малодослідженим залишається питання використання нейромереж, що добре зарекомендували себе в задачах розпізнавання образів, окремим випадком яких є фільтрація спаму.

Таким чином, вважаю розвиток методів спам фільтрації буде в напрямку ймовірнісних методів і штучного інтелекту.