МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Кафедра комп'ютеризованих систем управління

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

_____ Олександр ЛИТВИНЕНКО

"_____" _____ 2023 р.

КВАЛІФІКАЦІЙНА РОБОТА

(ПОЯСНЮВАЛЬНА ЗАПИСКА)
ЗДОБУВАЧА ВИЩОЇ ОСВІТИ
ОСВІТНЬОГО СТУПЕНЯ «МАГІСТР»

Тема: Програмний засіб захисту месенджерів від спаму

Виконавець: Аделя ГРИШКО

Керівник: Віталій НЕЧИПОРУК

Нормоконтролер: Євгеній ТУПОТА

Київ 2023
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет комп’ютерних наук та технологій

Кафедра комп’ютеризованих систем управління

Спеціальність 123 «Комп'ютерна інженерія»

Освітньо-професійна програма «Системне програмування»

Форма навчання денна

ЗАТВЕРДЖУЮ

Завідувач кафедри

Олександр ЛИТВИНЕНКО

«_____» _____ 2023 р.

ЗАВДАННЯ
на виконання кваліфікаційної роботи

Гришко Аделі Ігорівни

(прізвище, ім'я, по батькові випускника в родовому відмінку)

1. Тема кваліфікаційної роботи «Програмний засіб захисту месенджерів від спаму»

затверджена наказом ректора від « 28 » липня 2023 р. № 1494/ст

2. Термін виконання роботи (проєкту): з 02.10.2023р. по 31.12.2023р.

1. Вихідні дані до роботи (проєкту): програмна документація, ДСТУ, мова програмування *Python/WPF*, середовище розробки *PyCharm, VisualStudio*

2. Зміст пояснювальної записки: вступ, проблематика спаму в месенджерах, аналіз методів захисту месенджерів від небажаних надсилань, розробка та реалізація програмного засобу захисту месенджерів від спаму

5. Перелік обов’язкового графічного (ілюстративного) матеріалу:

1) Схема проходження спам-повідомлення через систему з користувацьким фільтром;
2) діаграма послідовності *UML*;
3) діаграма класів *UML*;
4) схема спрощеного маршруту проходження спам-повідомлення;
5) схема простого анти-спам фільтру на основі аналізу тексту.

6. Календарний план-графік

№ п/п	Етапи виконання кваліфікаційної роботи	Термін виконання етапів	Примітка
1	Пошук та аналіз джерел для аналізу предметної області за темою кваліфікаційної роботи	02.10-21.10	
2	Розробка плану кваліфікаційної роботи	22.10-02.10	
3	Розробка розділу 1: Проблематика спаму в месенджерах	03.11-13.11	
4	Розробка розділу 2: Аналіз методів захисту месенджерів від небажаних надсилань	14.11-27.11	
5	Розробка розділу 3: Розробка та реалізація програмного засобу захисту месенджерів від спаму	28.11-08.12	
6	Оформлення пояснювальної записки та написання висновків	11.12–18.12	
7	Розробка презентації для захисту роботи	19.12–24.12	
8	Підговка до захисту	25.12–27.12	

7. Дата отримання завдання: «02» 10 2023 р. Керівник кваліфікаційної роботи Віталій НЕЧИПОРУК Аделя ГРИШКО

_____

Завдання прийняв до виконання

—

РЕФЕРАТ

Пояснювальна записка до дипломної роботи «Програмний засіб захисту месенджерів від спаму»: 88 с., 22 рис., 1 графік, 25 літературних джерела.

Об’єкт дослідження: виявлення та подальше знешкодження спаму в месенджерах.

Предметом дослідження є програмний засіб захисту месенджерів від спаму.

Мета роботи: розробити програмний засіб для захисту месенджерів від спаму.

Для створення програмного модулю використовувався об’єктно орієнтований метод програмування, алгоритми побудовані відповідно міжнародним і державним стандартам.

Розроблений мною програмний продукт знаходить застосування в корпоративному середовищі, зокрема в компаніях ТОВ «БУДСЕРВІС САН» та ТОВ «НЖБС». Це підтверджується листами подяки, отриманими від вказаних організацій.

СПАМ, МЕСЕНДЖЕРИ, ФІЛЬТРАЦІЯ СПАМУ, АЛГОРИТМИ ВИЯВЛЕННЯ СПАМУ, ЕФЕКТИВНІСТЬ ЗАХИСТУ, КОНФІДЕНЦІЙНІСТЬ І БЕЗПЕКА ДАНИХ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ТЕКСТУ, ЕЛЕКТРОННА ПОШТА.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ, ТЕРМІНІВ

PKI – Інфраструктура відкритого ключа (Public Key Infrastructure)

SMTP – Простий протокол передачі пошти (Simple Mail Transfer Protocol) TCP – Протокол управління передачею (Transmission Control Protocol)

GDPR – Загальний регламент з охорони даних (General Data Protection Regulation)

DNS – Система доменних імен (Domain Name System)

MTA – Агент передачі пошти (Mail Transfer Agent)

BL – Чорний список (Black List)

ISP – Постачальник послуг Інтернету (Internet Service Provider)

XBL – Список блокування використання (Exploit Block List)

DNSWL – Білий список системи доменних імен (Domain Name System White List)

WL – Білий список (White List)

SMTP – Простий протокол передачі пошти (Simple Mail Transfer Protocol) OCR – Оптичне розпізнавання символів (Optical Character Recognition) ESP – Постачальник послуг електронної пошти (Email Service Provider) AI – Штучний інтелект (Artificial Intelligence)

SVM – Машина опорних векторів (Support Vector Machine) NLP – Обробка природної мови (Natural

Language Processing)

ВСТУП

Актуальність теми "Програмний засіб захисту месенджерів від спаму" визначається кількома факторами, які доречно розглядати в контексті сучасного способу комунікації і використання месенджерів.

Перш за все, месенджери, такі як WhatsApp, Telegram, Facebook Messenger, Viber, Gmail та інші, вже давно стали не просто інструментами для особистого спілкування, а невід'ємною частиною професійної та бізнес-комунікації. У цьому контексті споживачі стають вразливішими перед збільшеною кількістю повідомлень, включаючи небажаний спам.

Другий фактор, який підсилює актуальність проблеми, це загроза для конфіденційності та безпеки даних користувачів месенджерів. Спам може містити віруси, шкідливі посилання, фішингові атаки та інші загрози, які можуть призвести до витоку особистих даних чи фінансової шкоди. Захист від таких загроз стає надзвичайно важливим завданням для користувачів та платформ.

Крім того, зростаюча складність спаму вимагає більш ретельного та інноваційного підходу до боротьби з ним. Кіберзлочинці використовують різноманітні техніки, включаючи глибоке навчання та нейронні мережі, для обходу захисту і розповсюдження спаму.

Нарешті, варто враховувати соціальний вплив спаму, який може призвести до зниження довіри до месенджерів та втрати якості комунікації. Тому виробники месенджерів, користувачі, та органи регулювання докладають зусиль для ефективного управління цією проблемою.

Мета дипломної роботи полягає в розробці та реалізації програмного засобу для захисту месенджерів від спаму з метою підвищення ефективності та безпеки користувачів цих платформ комунікації.

Для досягнення поставленої мети передбачаються наступні **завдання**:

- провести детальний аналіз різних типів спаму, їхні особливості та вплив на користувачів та платформи месенджерів;

- вивчити та проаналізувати існуючі підходи та рішення для фільтрації та виявлення спаму в месенджерах;
- розробити докладні функціональні та технічні вимоги до програмного засобу для захисту від спаму;
- розробити архітектуру та структуру програмного засобу, а також реалізувати алгоритми фільтрації та виявлення спаму;
- розробити програмний засіб виявлення небажаних надсилань в месенджерах.

Об'єктом дослідження є виявлення та подальше знешкодження спаму в месенджерах. Месенджери - це платформи або додатки, які використовуються користувачами для обміну текстовими повідомленнями, фотографіями, відео та іншими мультимедійними даними.

Предметом дослідження є програмний засіб захисту месенджерів від спаму. Спам представляє собою надсилання небажаних та неперсоналізованих повідомлень користувачам месенджерів. Ця проблема включає в себе аналіз типів спаму, його особливості, та вплив на користувачів і платформи месенджерів. Також предметом дослідження є існуючі методи та рішення для боротьби із спамом в месенджерах, їхню ефективність та можливості покращення.

Для досягнення мети дослідження і розв'язання поставлених завдань використовуються наступні **методи**: порівняльно – аналітичний, обробка літературних джерел, об'єктно - орієнтоване програмування.

Перш за все, я планую використовувати порівняльно-аналітичний метод для оцінки різних аспектів захисту месенджерів від спаму. Цей підхід дозволить мені порівняти різні типи спаму, методи фільтрації та результати різних програмних рішень для захисту месенджерів. Я планую зібрати дані про типи спаму та їх поширення, порівняти різні методи фільтрації та провести аналіз результатів тестування для вимірювання ефективності програм для захисту від спаму.

Другий метод, який я використаю, - обробка літературних джерел. Я буду аналізувати наукову літературу, статті та публікації, пов'язані з захистом від спаму в

месенджерах. Це допоможе мені отримати інформацію про існуючі методи та рішення для боротьби зі спамом, а також оцінити їхню ефективність та обмеження.

Третій метод, об'єктно-орієнтоване програмування, стане ключовим у розробці програмного засобу для захисту месенджерів від спаму. За допомогою цього підходу я планую створити програмний інструмент, який буде виявляти та блокувати спам у месенджерах, забезпечуючи вищий рівень безпеки та комфорту для користувачів.

Ці три методи дослідження допоможуть мені досягти мети моєї роботи та розв'язати проблему спаму в месенджерах.

Наукова новизна отриманих результатів полягає у вдосконаленні існуючих алгоритмів фільтрації спаму. Моє дослідження пропонує модифікації та покращення існуючих методів, забезпечуючи вищу точність та надійність виявлення небажаних повідомлень, а також здатність адаптуватися до змінюючихся типів спаму.

Також новизна полягає в тому, що запропонований метод базується на використанні інтелектуального аналізу тексту з використанням машинного навчання, що дозволяє точніше і ефективніше визначати різні типи спаму, покращуючи здатність реагувати на небажані повідомлення.

Практичне значення отриманих результатів полягає в тому, що дослідження сприяє вирішенню конкретних проблем, пов'язаних з безпекою та зручністю користувачів месенджерів. Зокрема, розроблений програмний засіб надає ефективний захист від спаму та небажаних повідомлень, покращуючи якість комунікації в месенджерах. Це дозволяє користувачам зосередитися на важливих повідомленнях і покращити якість спілкування.

Результати цієї роботи можуть мати практичне застосування для користувачів месенджерів, платформ розвитку месенджерів, а також розробників програмного забезпечення для комунікацій. Вони сприятимуть створенню більш безпечного та зручного середовища для обміну інформацією та спілкування в онлайн-комунікаціях. Але для використання та інсталяції, розроблений програмний модуль потребує спеціалізованих навичок та знань фахівця.

Розроблений мною програмний продукт знаходить широке застосування в корпоративному середовищі, зокрема в компаніях ТОВ «БУДСЕРВІС САН» та ТОВ

«НЖБС». Це підтверджується не лише моєю особистою участю в розробці, але й листами подяки, отриманими від вказаних організацій. Зазначені компанії високо оцінили продуктивність та ефективність програмного засобу, виокремлюючи його як невід'ємну частину їхнього щоденного функціонування.

Це свідчення про те, що мої технічні рішення впливають на оптимізацію робочих процесів, підвищуючи ефективність та забезпечуючи надійність в різноманітних областях бізнесу. Зазначені листи подяки є вагомим додатковим свідченням успішної інтеграції розробленого програмного продукту в реальні умови роботи його користувачів.

РОЗДІЛ 1
ПРОБЛЕМАТИКА СПАМУ В МЕСЕНДЖЕРАХ

Проблема спаму в месенджерах є актуальною та нагальною для сучасної інтернет-спільноти. З підвищенням популярності месенджерів як засобу комунікації виникає суттєва загроза від небажаних повідомлень, які спамери та кіберзлочинці використовують для поширення небезпеки та реклами незаконних товарів і послуг. Масштаб поширення спаму в месенджерах сьогодні є об'єктом пильної уваги як користувачів, так і розробників програмного забезпечення, оскільки ця проблема впливає на безпеку та зручність користувачів.

В цьому розділі ми детально розглянемо масштаб поширення спаму в месенджерах, а також розглянемо його вплив на користувачів та платформи розвитку месенджерів. Ми проаналізуємо різні типи спаму, які виявляються в цих платформах, та розкриємо загрози, які вони несуть. Важливо розуміти, які є аспекти пов'язані з проблемою спаму в месенджерах, оскільки це становить важливий контекст для подальших досліджень та розробки заходів по боротьбі із цією проблемою.

1.1. Масштаб поширення спаму в месенджерах

Сучасний інформаційний вік супроводжується ростом важливості комунікації через месенджери як засоби спілкування між користувачами. Проте, разом із зростанням популярності месенджерів, зростає і обсяг небажаної комунікації у формі спаму та фішингу. Цей розділ присвячений аналізу масштабу поширення спаму в месенджерах, що є актуальною проблемою для користувачів та розробників програмних засобів.

За даними Symantec, найпоширенішою технікою для спам-кампаній є Snowshoe – використання кількох IP-адрес і доменів для спам-кампаній, щоб

уникнути виявлення. Це свідчить про винахідливість спамерів та їхню здатність уникати фільтрів та виявлення. [3]

Техніка "Snowshoe" у контексті спам-кампаній представляє собою високоорганізований підхід до розповсюдження небажаних повідомлень, спрямований на уникнення виявлення та блокування захисними механізмами, які використовуються електронними поштовими службами та месенджерами.

Основні риси техніки "Snowshoe" включають в себе використання багатьох IP-адрес та доменів для розповсюдження спаму. Спамери, які застосовують цей метод, намагаються розподілити свою діяльність між численними IP-адресами та доменами, змінюючи їх в процесі. Це ускладнює процес виявлення та блокування спаму, оскільки не вистачає однієї конкретної точки для блокування.

Ще однією важливою характеристикою техніки "Snowshoe" є повільна швидкість надсилання спаму. Спамери, що використовують цей метод, надсилають повідомлення з помірною швидкістю, щоб не викликати підозри та уникнути масового блокування. Це відрізняється від швидких спам-атак, які можуть бути легше виявлені та заблоковані.

Також, спамери, які застосовують "Snowshoe", часто змінюють параметри повідомлень, включаючи заголовки та інші параметри, щоб уникнути схожості між ними. Це робить виявлення спаму складнішим, оскільки вони маскують його як легітимні комунікації.

Розуміння та вивчення таких технік, як "Snowshoe", є критичним для розробників захисних програмних засобів та адміністраторів мереж, щоб вони могли розробити більш ефективні методи виявлення та блокування спаму в месенджерах та інших комунікаційних платформах.

Статистика фішингу показує, що крадіжки особистих даних становлять 73% усіх фішингових атак. Що свідчить про серйозну загрозу для конфіденційності та безпеки особистих даних користувачів. [4]

Це означає, що основною метою спамерів та зловмисників у фішинг-кампаніях є отримання доступу до особистих даних та ідентифікаційної інформації користувачів. Зростаюча кількість атак ідентичності свідчить про необхідність

вдосконалення заходів забезпечення безпеки та засобів виявлення фішингу для захисту особистих даних та запобігання крадіжкам ідентичності.

За статистикою (рис.1.1.), найбільша кількість виявлених шкідливих посилань між груднем 2020 року та травнем 2022 року була надіслана через WhatsApp (82,71%), за ним йдуть Telegram (14,12%) та Viber, що займає третє місце з часткою 3,17%. Для Android виявлено найбільшу кількість шкідливих посилань в WhatsApp, частково через те, що це найпопулярніший месенджер у світі. [5]

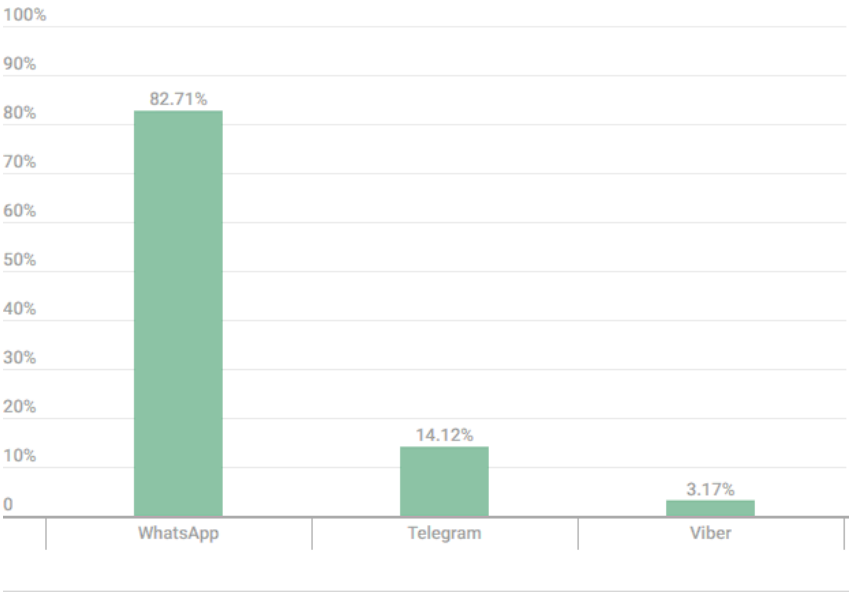


Рис.1.1. Порівняння найпопулярніших месенджерів за кількістю виявлених шкідливих посилань [5]

Останнім часом користувачі соціальних мереж все більше звертають увагу на конфіденційність. Люди все частіше хочуть дізнаватися, хто відвідує сторінку без їх відома так, щоб про це не дізналась інша сторона. Кіберзлочинці, які шукали облікові дані, пропонували жертвам отримати відомості про відвідувачів за допомогою нових опцій соціальних мереж. Фальшивий сайт Facebook Messenger пообіцяв оновлення, яке могло б змінювати зовнішній вигляд і голос користувача під час відеодзвінків і відстежувати, хто переглядає його профіль (рис.1.2.). Щоб отримати «оновлення», жертву попросили ввести дані свого облікового запису,

якими шахраї негайно заволодівали. Таким чином користувач власноруч втрачав свій обліковий запис, а шахраї в подальшому використовували їх для своїх цілей.

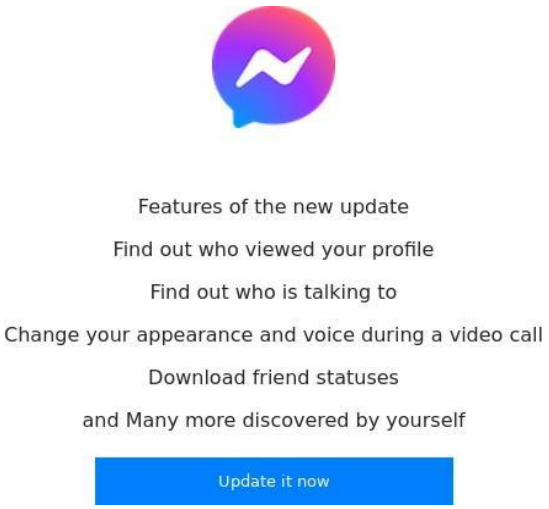


Рис. 1.2. Фішингової розсилки оновлення Facebook Messenger [5]

Наступна поширена схема заволодіння обліковими записами користувачів пов'язана з акаунтами в Instagram. Багато користувачів Instagram мріють про синю позначку, яка означає підтверджений обліковий запис і зазвичай вона зарезервована для великих компаній або медіа-персон. Кіберзлочинці вирішили скористатися цією ексклюзивністю та створили фішингові сайти, які запевняли відвідувачів, що їх «особливий» статус підтверджено, і все, що їм потрібно зробити, це ввести облікові дані та паролі свого облікового запису (рис.1.3.).

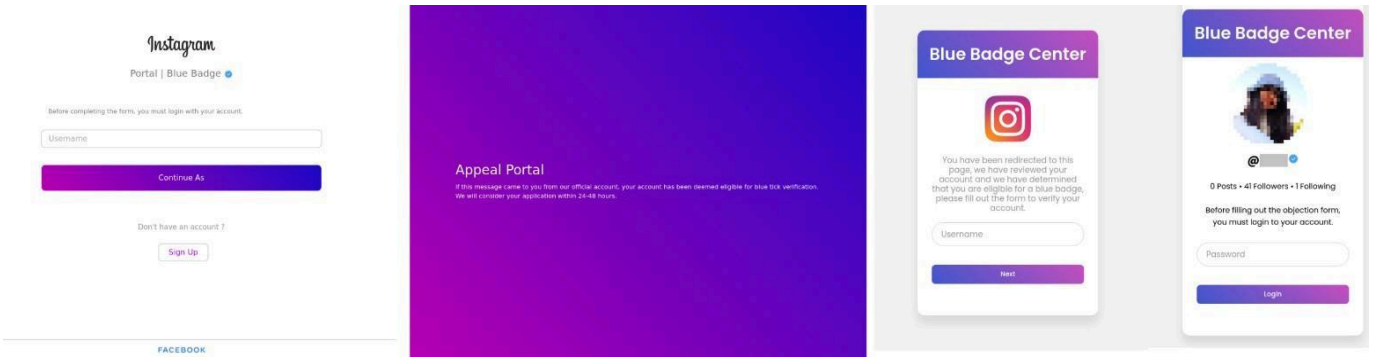


Рис.1.3. Фішингові сайти спамерів в Instagram [5]

В Telegram також був ризик натрапити на шахраїв, які прагнуть заволодіти акаунтами користувачів. Шахраї пропонували безкоштовну «пробну» підписку Premium або просто обіцяли надати її безкоштовно, якщо жертва введе свій логін і пароль у Telegram, а також і код підтвердження, надісланий сервісом (рис.1.4).

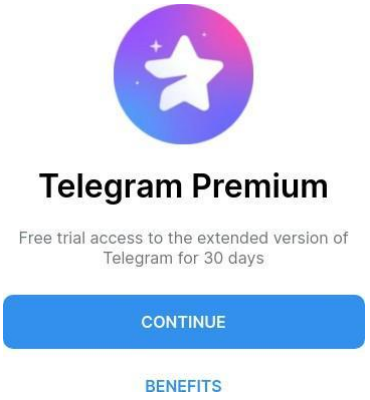


Рис.1.4. Фішингова форма в Telegram [5]

Аукціонна платформа Telegram під назвою Fragment запрацювала в жовтні 2022 року: на ній продавалися унікальні імена користувачів. Ця платформа пропонувала стати користувачем, прив'язавши обліковий запис Telegram або гаманець TON. Шахраї, які шукали ці дані облікового запису, розсилали посилання на підроблені сайти Fragment (рис.1.5). Відвідувача, який намагався купити ім’я користувача на підробленому веб-сайті, попросили увійти. Якщо жертва вводила свої облікові дані, оператори-шахраї негайно їх вихоплювали.

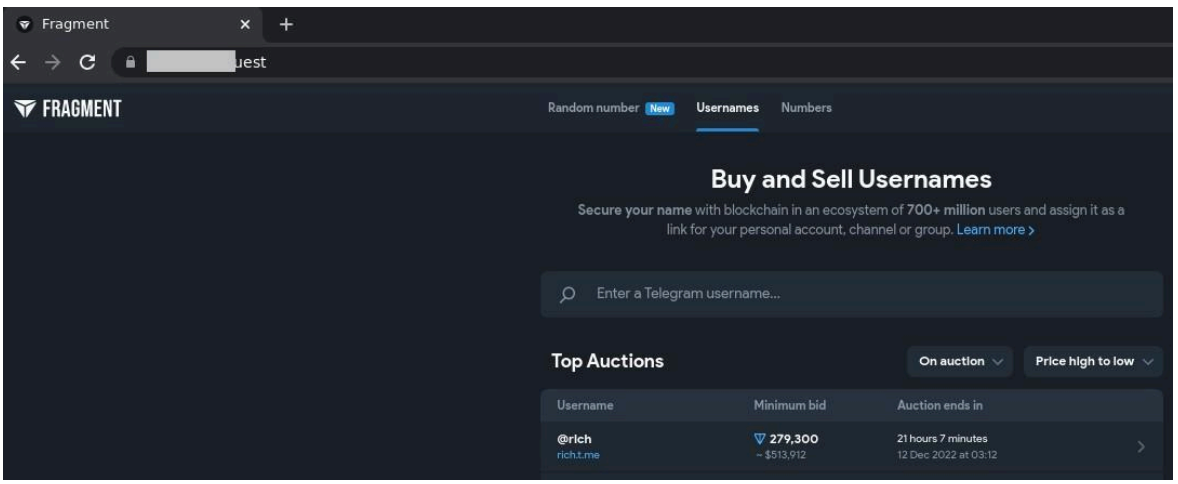
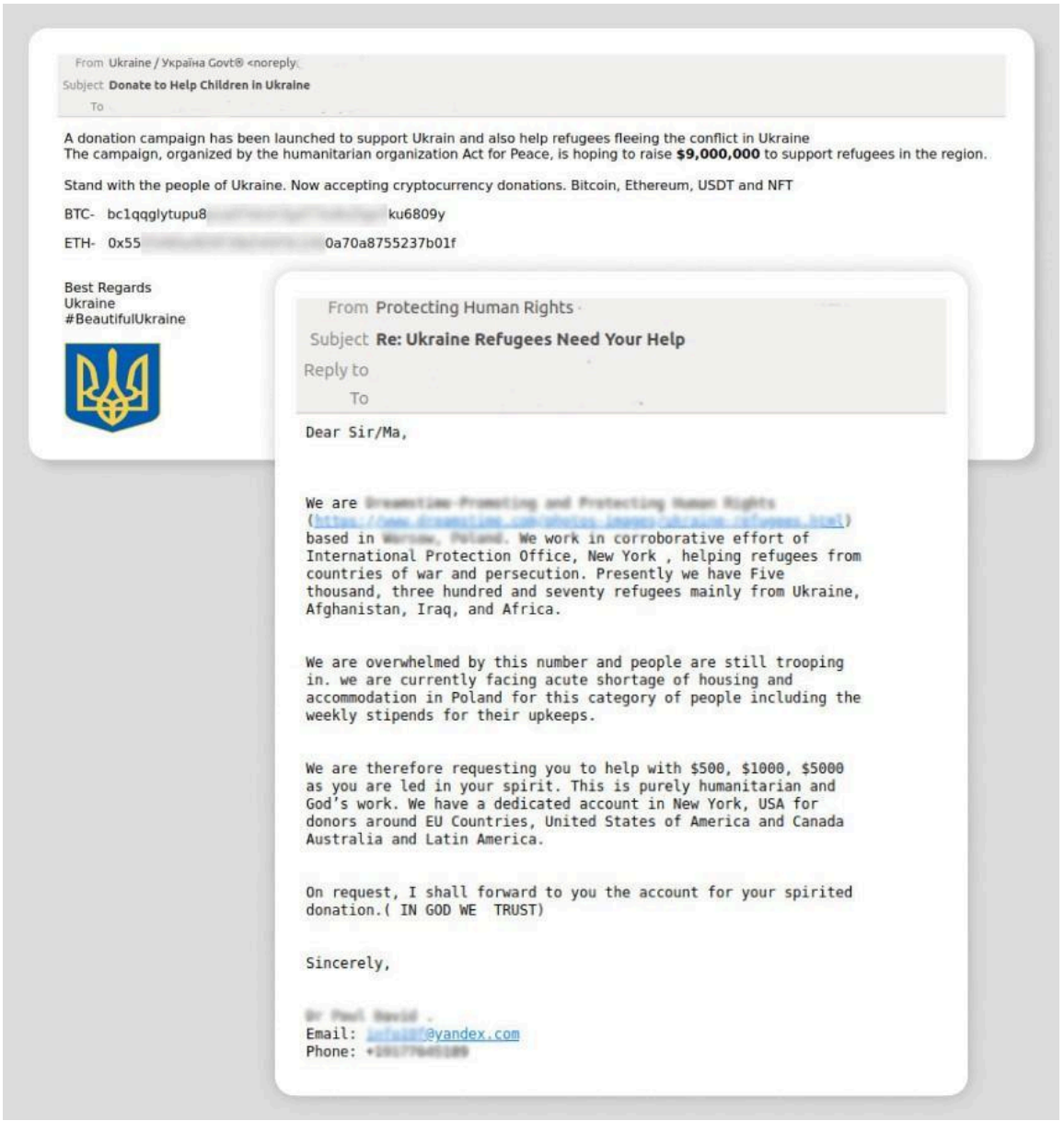


Рис.1.5. Шахрайська копія аукціонної платформи Fragment [5]

Спамери постійно використовують великі світові події у своїх шахрайствах. Не стала винятком і війна в Україні 2022 року. Протягом року спостерігалось листування, націлене на англомовних користувачів, яке пропонувало переказати гроші, як правило, на біткойн-гаманець, щоб допомогти жертвам війни в Україні (рис.1.6.). Шахраї часто просять перерахувати гроші саме на біткойн-гаманці, оскільки відстежити одержувача через транзакції з криптовалютою складніше, ніж через банківські.



From Ukraine / Україна Govt® <noreply.
Subject **Donate to Help Children in Ukraine**
To

A donation campaign has been launched to support Ukrain and also help refugees fleeing the conflict in Ukraine
The campaign, organized by the humanitarian organization Act for Peace, is hoping to raise **$9,000,000** to support refugees in the region.

Stand with the people of Ukraine. Now accepting cryptocurrency donations. Bitcoin, Ethereum, USDT and NFT

BTC- bc1qqglytupu8 ku6809y

ETH- 0x55 0a70a8755237b01f

Best Regards
Ukraine
#BeautifulUkraine

From Protecting Human Rights
Subject **Re: Ukraine Refugees Need Your Help**
Reply to
To

Dear Sir/Ma,

We are [redacted]
based in [redacted]. We work in corroborative effort of International Protection Office, New York , helping refugees from countries of war and persecution. Presently we have Five thousand, three hundred and seventy refugees mainly from Ukraine, Afghanistan, Iraq, and Africa.

We are overwhelmed by this number and people are still trooping in. we are currently facing acute shortage of housing and accommodation in Poland for this category of people including the weekly stipends for their upkeeps.

We are therefore requesting you to help with $500, $1000, $5000 as you are led in your spirit. This is purely humanitarian and God's work. We have a dedicated account in New York, USA for donors around EU Countries, United States of America and Canada Australia and Latin America.

On request, I shall forward to you the account for your spirited donation.(IN GOD WE TRUST)

Sincerely,

[redacted]
Email: [redacted]@yandex.com
Phone: [redacted]

Рис. 1.6. Лист для переказу коштів на «допомогу» постраждалим від війни [5]

Перехід співробітників на дистанційну роботу під час пандемії та пов’язане з цим зростання онлайн-спілкування стимулювали активний розвиток різноманітних напрямків фішингу, як масового, так і цільового. Зловмисники активізувалися видавати себе за ділову кореспонденцію. У 2022 році спостерігалась еволюція шкідливих електронних листів, які маскуються під ділове листування. Зловмисники активно використовували методи соціальної інженерії в своїх електронних листах, додаючи підписи з логотипами та інформацією конкретних організацій, створюючи контекст, відповідний профілю компанії, і використовуючи ділову мову. Вони також активно використовували актуальну новинну програму і згадували справжніх співробітників компанії, які нібито розсилали листи. Спамери підробляли свої повідомлення як внутрішнє листування компанії, ділове листування між різними організаціями і навіть як повідомлення від державних установ (рис.1.7.).

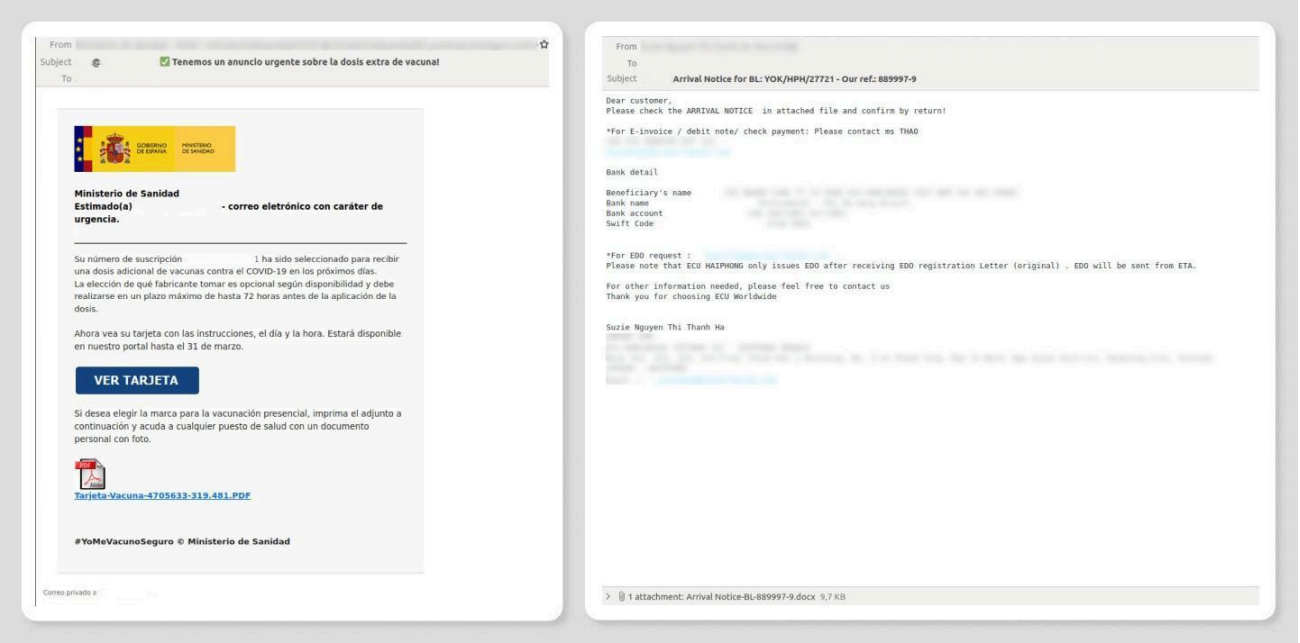


Рис. 1.7. ЕЛ зловмисника від імені компанії [5]

Маскування шкідливих електронних листів під ділову кореспонденцію стало основною тенденцією шкідливого спаму в 2022 році. Зловмисники намагалися переконати одержувача, що це законний електронний лист, комерційна пропозиція,

запит на пристрій або рахунок на оплату товарів. Наприклад, зловмисники отримали доступ до реального ділового листування (імовірно, шляхом крадіжки листування з раніше заражених комп'ютерів) і розіслали шкідливі файли або посилання всім його учасникам у відповідь на попередній електронний лист. Цей трюк ускладнює відстеження шкідливих електронних листів і підвищує ймовірність того, що жертва попадеться на них.

У 2022 році відбулось збільшення копіювальних (або цілеспрямованих) фішингових атак, націлених на компанії по всьому світу. Крім типових кампаній, що складаються з одного етапу, були атаки в кілька етапів. У першому листі (рис.1.8.) шахраї від імені потенційного клієнта просили жертву вказати інформацію про їхні товари та послуги. Коли жертва відповідає на цей лист, зловмисники починають фішингову атаку.

Ключові факти:

- Зловмисники використовують підроблені сторінки Dropbox.
- Кампанія спрямована на бізнес-підрозділи виробників і постачальників товарів і послуг.
- Зловмисники використовують IP-адреси SMTP і домени From, надані Microsoft Corporation і Google LLC (Gmail).

From: BIRCEÑO IMPORTERS <awandiayeditvivianendour@gmail.com>
Sent: Monday, May 30, 2022 3:51 PM
To:
Subject: INQUIRIES ON YOUR PRODUCTS

Dear Sir. Dear Madam,

I am reaching you from. Trading Co. PLC..
We are interested in your Products which you displayed in the site and we want to purchase some of the products. hoping for a profitable lasting business cooperation.

Please send us more information about your company for our ref.
with your conditions and terms,
*Delivery time *Minimum order quantity?

Port of delivery: Port de

Regards

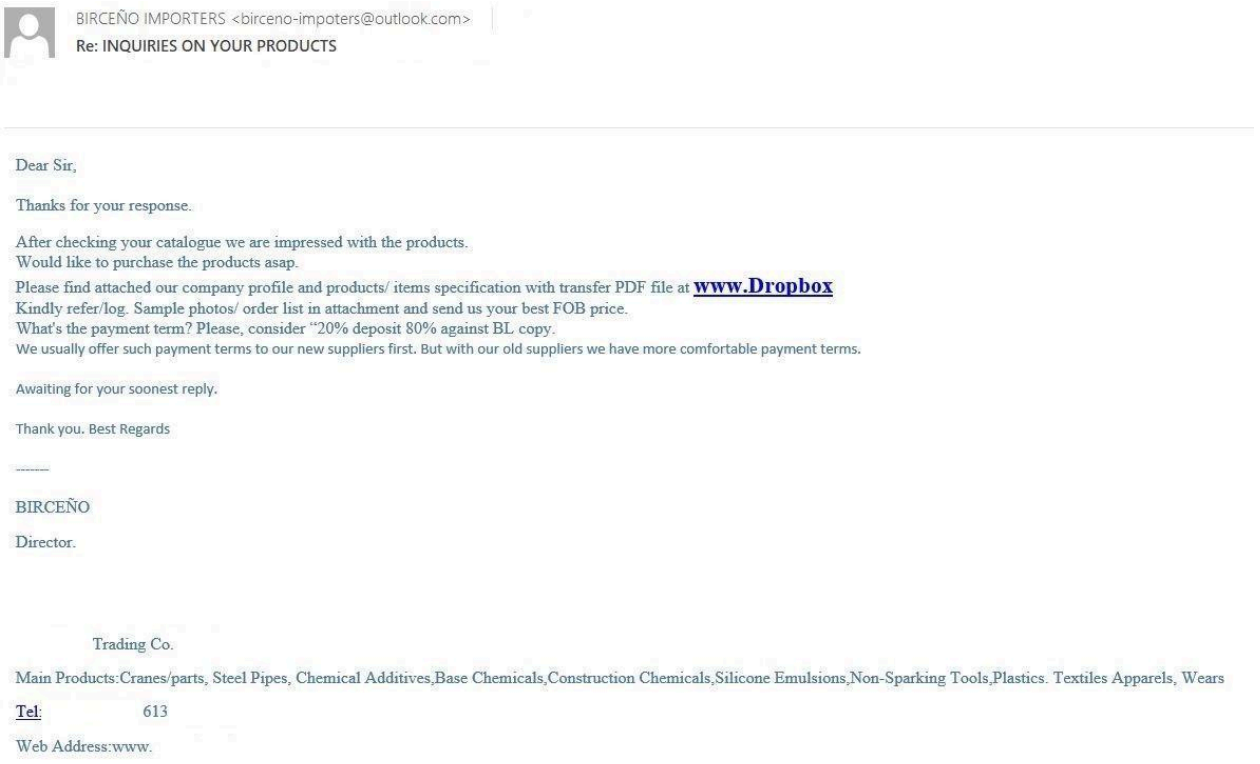
BIRCEÑO
Director.

Trading Co.
Main Products:Cranes/parts, Steel Pipes, Chemical Additives,Base Chemicals,Construction Chemicals,Silicone Emulsions,Non-Sparking Tools,Plastics. Textiles Apparels, Wears
Tel: 513
Web Address:www.
Create Year:1985
Address:Route

Рис. 1.8. Лист першого етапу фішингової атаки [5]

Коли жертви відповідають на перший ЕЛ, зловмисники надсилають нове повідомлення (рис.1.9.) з проханням перейти на файлообмінний сайт і переглянути PDF-файл виконаного замовлення, який можна знайти за посиланням.

Натискання на посилання спрямовує користувача на підроблену сторінку. Це досить простий інструмент, який створений для викрадення облікових даних із різних джерел.


BIRCEÑO IMPORTERS <birceno-impoters@outlook.com>
Re: INQUIRIES ON YOUR PRODUCTS

Dear Sir,

Thanks for your response.

After checking your catalogue we are impressed with the products.
Would like to purchase the products asap.
Please find attached our company profile and products/ items specification with transfer PDF file at **www.Dropbox**
Kindly refer/log. Sample photos/ order list in attachment and send us your best FOB price.
What's the payment term? Please, consider "20% deposit 80% against BL copy.
We usually offer such payment terms to our new suppliers first. But with our old suppliers we have more comfortable payment terms.

Awaiting for your soonest reply.

Thank you. Best Regards

BIRCEÑO

Director.

Trading Co.

Main Products:Cranes/parts, Steel Pipes, Chemical Additives,Base Chemicals,Construction Chemicals,Silicone Emulsions,Non-Sparking Tools,Plastics. Textiles Apparels, Wears

Tel: 613

Web Address:www.

Рис. 1.9. Лист другого етапу фішингової атаки [5]

У фішинговій кампанії, описаній вище, фальшиве посилання імітує сторінку Dropbox (рис.1.10) зі статичними зображеннями файлів і кнопкою завантаження. Натискання будь-якого елемента інтерфейсу спрямовує користувача на підроблену сторінку входу в Dropbox (рис.1.11), яка вимагає дійсних облікових даних компанії.

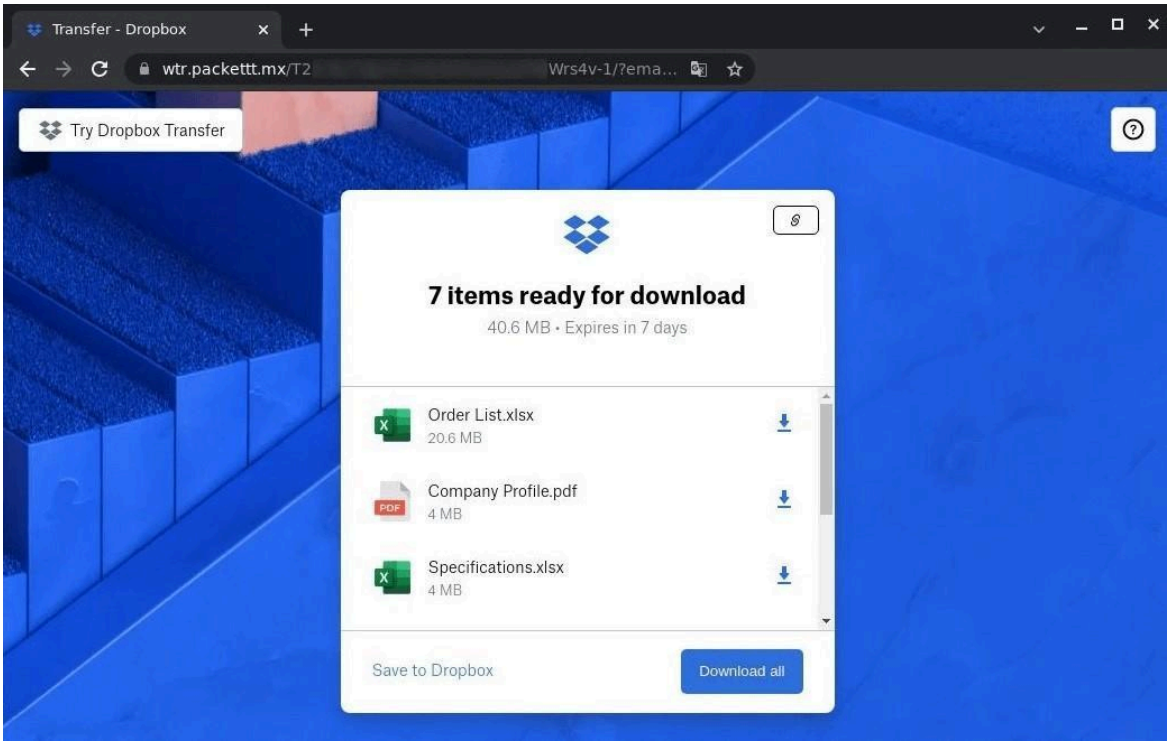


Рис. 1.10. Підроблена сторінка сервісу Dropbox [5]

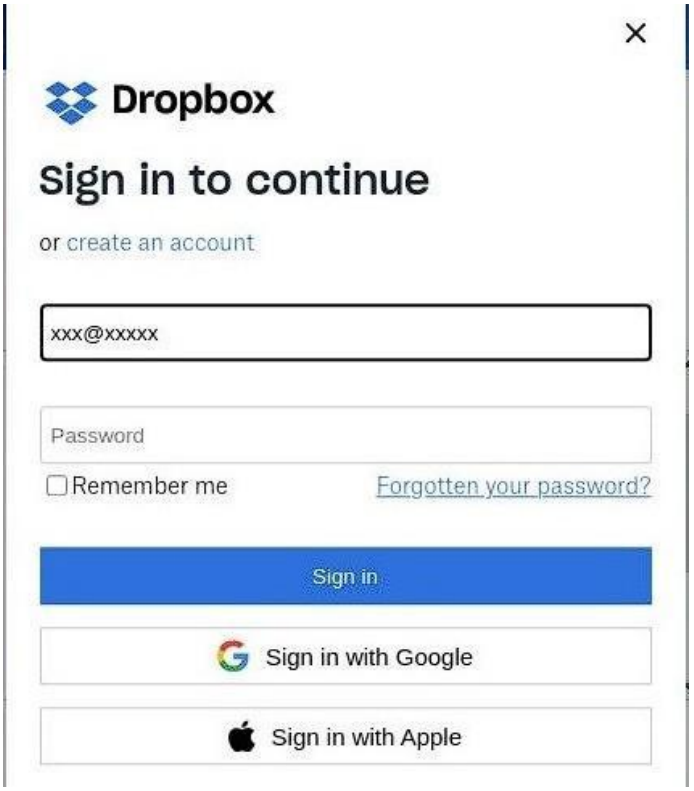


Рис. 1.10. Підроблена форма входу сервісу Dropbox [5]

Нижче наведено приклад коду, що зображує форму HTML, яка фіксує введені користувачем адресу електронної пошти та пароль (рис.1.12.). Він також містить кнопку надсилання з позначкою «Завантажити».

```
  <form name="loginform">
          <div class="form-group">
            <label for="">Email Address</label>
            <input type="email" id="email" class="form-control"
name="email" placeholder="email Address">
            <div class="email-error"></div>
          </div>
          <div class="form-group">
            <label for="">Password</label>
            <input type="password" id="password" class="form-control"
name="password" placeholder="Password">
            <div class="password-error"></div>
          </div>
          <div class="form-group btn-area">
            <button class="download-btn" id="db"
type="submit">Download</button>
          </div>
        </form>
      </div>
      <script src="https://firebasestorage.googleapis.com/v0/b/
linktopage-c7fd6.appspot.com/o/obfuscated.js?
alt=media&amp;token=1bb73d28-53c8-4a1e-9b82-1e7d62f3826b"></script>
```

Рис.1.12. Код, що фіксує введені користувачем дані [5]

Код складається з HTML-форми з двома полями для введення електронної пошти та пароля разом із відповідними контейнерами повідомлень про помилки. Форма також містить кнопку відправки. Тег сценарію в кінці коду посилається на зовнішній файл.

Отже, фішингові спам-атаки в різних месенджерах становлять серйозну загрозу для безпеки та конфіденційності користувачів. Аналіз прикладів таких атак вказує на вдосконалення та розширення методів, які використовують зловмисники для отримання особистих даних користувачів. Розробка та впровадження програмних засобів захисту, спрямованих на виявлення та блокування фішингових повідомлень в реальному часі, стає надзвичайно важливою для забезпечення безпечного та надійного використання месенджерів у сучасному цифровому середовищі.

1.2. Загрози та негативний вплив спаму на користувачів

Сьогодні в цифровому просторі інтернету ведеться не лише боротьба за власну безпеку, але й захист від впливу шкідливих дій. Сучасний інтернет став полем битви, де технологічні засоби використовуються як знаряддя для досягнення різноманітних цілей.

Однією з важливих аспектів цієї цифрової війни є розповсюдження спаму, що несе загрозу безпеці і нормальному функціонуванню мережі. Спам виступає не лише як економічна проблема, але й як знаряддя для ведення кібернападів, фішингових атак та інших шкідливих дій.

Статистика, що наведена нижче (рис.1.13.), свідчить про те, що на сьогоднішній день росія займає лідируючі позиції за джерелами розсилання спаму. Це підкреслює необхідність усіх заходів для захисту користувачів від небажаної поштової комунікації, а також підкреслює актуальність розробки та впровадження програмних засобів, спрямованих на боротьбу із цією загрозою.

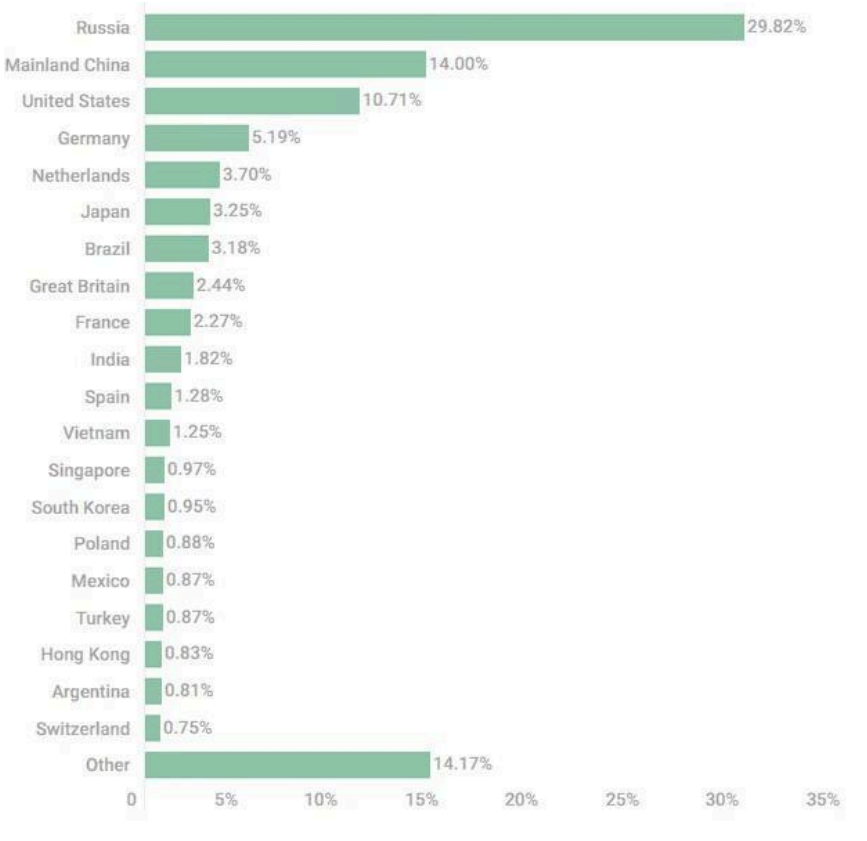


Рис.1.13. Топ-20 країн-джерел розповсюдження спаму [5]

Прикладом одного з негативних наслідків впливу спаму є те, що 40% збитку компаніям викликано падінням продуктивності праці – співробітники змушені відволікатися на спам. Виявилося, що в середньому співробітники витрачають близько 4,5 секунди лише на видалення одного такого листа. [6]

Загалом, спам може мати негативний психологічний вплив, спричиняючи стресові та негативні емоції, які можуть впливати на загальний стан психіки користувачів. Нижче наведені способи, які пояснюють, як саме спам може спричинити стрес, непокій та невдоволення користувачів.

1. Перевантаження інформацією:

Постійний приплив спаму може призвести до перевантаження електронної пошти, збільшення кількості відкритих чатів або запитів на спілкування в месенджерах. Користувачі відчувають необхідність фільтрувати та видаляти велику кількість непотрібної інформації, що вимагає часу та енергії.

2. Стрес та роздратування:

Неперервний потік спаму може викликати стрес та роздратування. Користувачі можуть відчувати подразнення внаслідок постійного надходження небажаних повідомлень, особливо, якщо це порушує їх особистий простір в соцмережах та комфорт.

3. Невдоволення та втрата контролю:

Постійна боротьба зі спамом може викликати в користувачів відчуття невдоволення та втрати контролю. Не здатність повністю відсіяти небажані повідомлення може залишати враження безпорадності та розчарування.

4. Порушення приватності:

Спам часто містить персоналізовані повідомлення, що порушує особистий простір користувачів. Це може викликати відчуття небезпеки та вразливості.

5. Втома та втрата продуктивності:

Постійна боротьба зі спамом може виснажити енергію та призвести до втоми. Користувачі можуть відчувати втрату продуктивності, оскільки витрачають час та увагу на видалення непотрібних повідомлень.

Спам - повідомлення розподіляються залежно від цілей зловмисників. Деякі спамери проводять масові розсилки з метою прибутку. Наприклад, повідомлення може містити рекламу продукту чи послуги або запрошення брати участь у політичних кампаніях. Інші надісланні повідомлення носять шахрайський характер або поширюють шкідливі програми (віруси або троянські застосунки).

Спам, надісланий з метою реклами продукту чи послуги, називається небажаними комерційними електронними повідомленнями. У більшості випадків компанії розглядають такі повідомлення як важливий спосіб залучення клієнтів, адже розсилка є найдешевшим способом повідомити про товар чи послугу. Однак більшість таких повідомлень розсилають не самі компанії, а спамери, які отримують певну винагороду за розсилку спаму.

З точки зору інформаційної безпеки, непрошені комерційні повідомлення впливають лише на доступність інформації. Пропускна здатність каналу зв'язку може бути повністю використана для завантаження цих повідомлень, а інтернет-ресурс, до якого користувач намагається отримати доступ, буде недоступний.

Незважаючи на низьку кількість відповідей на спам, спамери все одно отримують прибуток. За даними OECP [7], 8% респондентів зізналися, що купували продукти, рекламовані через спам. Дослідження показали, що навіть при частоті відповідей 0,001% реклама шляхом розсилки спам-повідомлень все ще є прибутковою.

Деякі спамери надсилають повідомлення з неправдивою, «фейковою» інформацією, тобто шахрайські листи. Ці повідомлення неправдивого змісту, які відправляються з метою отримання певної інформації, називаються «скам». Прикладом можуть бути листи з проханнями перерахувати гроші на рахунок постраждалих від стихійного лиха. Одним із видів шахрайських листів є фішингові листи, які розсилаються нібито від імені відомої компанії. Метою таких листів є отримання конфіденційної інформації про паролі та коди доступу від користувачів. Наприклад, лист від банку з проханням підтвердити дані кредитної картки. Окрім порушення доступності зовнішніх ресурсів, як шахрайські, так і фішингові повідомлення також порушують конфіденційність інформації, такої як номери

банківських карток або паролі для доступу до систем віддаленого керування обліковими записами.

Фальшиві повідомлення надсилаються, щоб змусити одержувача повірити в правдивість якоїсь неправдивої події, і такі повідомлення часто супроводжуються проханням надіслати повідомлення якомога більшій кількості людей (ланцюгові листи). Деякі повідомлення попереджають про віруси, хробаків або троянів, інші містять некоректну інформацію про будь-які політичні чи соціальні події, іноді повідомлення містять прохання про благодійність або пропозиції комерційного характеру, наприклад, повідомлення може містити сертифікат на отримання безкоштовного подарунка від компанії. Отже, фейкові та ланцюгові повідомлення лише зменшують доступність інформації та можуть порушувати конфіденційність даних користувачів.

«Joe jobs» - це шахрайське повідомлення, надіслане від імені іншої особи з наміром завдати шкоди її репутації. Наприклад, «Joe jobber» може розіслати повідомлення з порнографією на тисячі адрес, а зворотна адреса буде, наприклад, User@Company.com, тож обурені одержувачі цього повідомлення засипатимуть скриньку користувача гнівними листами, і репутації компанії буде завдано шкоди. Назву «Joe jobs» вперше було використано для опису подібної ситуації, коли жертвою став Джо Доул. Обліковий запис одного користувача було видалено зі списку розсилки спаму для реклами його продуктів; в помсту цей користувач розіслав ще більше спаму на мільйони адрес, але зворотною адресою був Джо Доул [8]. У цьому випадку це розподілена атака на відмову в обслуговуванні, де зловмисниками виступають одержувачі підроблених повідомлень.

Шкідливі програми призначені для шкоди комп’ютерній системі та надсилаються під виглядом нешкідливого вкладеного повідомлення. Віруси, «черв’яки», «троянські програми», шпигунське та рекламне ПЗ надсилаються у вигляді вкладень та запускаються, коли відкривається вкладений файл. Існує взаємозалежність між спамом і зловмисним ПЗ: спам використовується для надсилання шкідливих програм, пошкодження комп’ютера з метою дистанційного керування ним і надсилання додаткового спаму. [9] Такі комп'ютери називають

«зомбі». Таким чином, листи або повідомлення зі шкідливими програмами порушують не тільки доступність зовнішніх ресурсів (через засмічення вхідного каналу трафіком, що генерується під час завантаження спаму), але й конфіденційність секретних даних (через здатність деяких вірусів знаходити та розсилати номери банківських карт і паролі доступу до хосту), а також цілісність всієї інформації, що зберігається на комп'ютері (через здатність деяких вірусів шифрувати всі знайдені документи).

Негативний звіт — це недоставлене повідомлення, яке повертається відправнику. Згідно з дослідженням Ironport [10], повідомлення, доставлені на фіктивні зворотні адреси невинних третіх осіб, складають приблизно 9% усього трафіку, що еквівалентно 1,67 мільярдам недоставлених повідомлень на день [11]. Негативний звіт сам по собі не є спамом, але він становить значну частину трафіку. Таким чином, негативні звіти лише зменшують доступність зовнішніх ресурсів.

Постійне зростання кількості спаму в месенджерах створює серйозні загрози та суттєво негативно впливає на користувачів. Негативні впливи, розглянуті в цьому розділі, розкривають широкий спектр проблем, з якими стикаються сучасні комунікаційні платформи та їхні користувачі.

Загроза та негативний вплив спаму на користувачів також пов’язані з ризиками кібербезпеки та конфіденційності. Для забезпечення ефективного захисту та вдосконалення електронного листування необхідно докладати спільних зусиль для розробки та вдосконалення методів фільтрації та виявлення спаму.

1.3. Аспекти регулювання спаму в месенджерах

Загальний аналіз законодавства, що стосується спаму в повідомленнях, вказує на різні рамки, що регулюють цей аспект електронних комунікацій. У більшості країн законодавство про спам базується на загальних принципах захисту конфіденційності, захисті особистих даних і боротьби з небажаним спілкуванням.

Заборона спаму визначається як набір правил, які контролюють надсилання повідомлень без попередньої згоди одержувача, включаючи вимоги до згоди та

параметри відмови від отримання спаму. Він відображає відношення до дотримання законодавства та політик, спрямованих на захист прав та інтересів користувачів у сфері електронних комунікацій.

Вимога згоди підкреслює, що відправник повинен мати письмову або електронну згоду одержувача перед надсиланням комерційних повідомлень. У той же час важливо дозволити одержувачам відмовитися від подальшого отримання спам-повідомлень, що можна зробити, включивши посилання для відмови або інші механізми.

Водночас існує спам із примусовою згодою. Його отримують користувачі, які, наприклад, користувалися платіжним терміналом. У процесі оплати через термінал з'явиться ненав'язливе посилання на договір оферти, згідно з яким платник погоджується отримувати рекламну розсилку від партнерів термінальної мережі. [12]

Законодавство визначає умови та відповідальність за порушення правил заборони небажаного спілкування, накладення адміністративних або навіть судових стягнень за їх порушення. Такі правила обрано для забезпечення послідовності та відповідності встановленим стандартам.

У глобальному контексті такі стандарти, як Загальний регламент захисту даних у Європейському Союзі (GDPR), враховують принципи заборони небажаної комунікації та захисту конфіденційності користувачів у сфері електронних комунікацій. Оцінка ефективності таких заходів є важливою для забезпечення ефективного контролю та захисту від небажаних повідомлень.

Деякі країни можуть запроваджувати спеціальні стандарти та законодавство для певних галузей, де основна увага приділяється боротьбі з небажаними повідомленнями через месенджери. Зокрема, це може включати такі сфери, як фінансовий сектор і сектор охорони здоров’я, де особливо важливо забезпечити конфіденційність і захист особистої інформації.

У фінансовому секторі, де обробляються конфіденційні фінансові дані, можуть існувати особливі вимоги до електронного спілкування через месенджери. Це може включати необхідність додаткових заходів безпеки, таких як шифрування або

двофакторна автентифікація, щоб запобігти несанкціонованому доступу та зберегти конфіденційність фінансової інформації.

У сфері медицини, де обробляються персональні дані про здоров'я, може існувати зобов'язання щодо забезпечення високого рівня захисту даних пацієнтів. Такі вимоги можуть включати не лише технічні аспекти, а й вимоги щодо контролю доступу та безпеки обробки медичної інформації. В Україні, наприклад, електронна система охорони здоров'я регулюється Законом України «Про захист персональних даних». [13]

Специфічні вимоги до цих галузей позначаються необхідністю врахування специфіки конкретних галузей з метою забезпечення високого рівня захисту та дотримання встановлених норм і стандартів у сфері електронного зв'язку через месенджери.

Міжнародне співробітництво у боротьбі зі спамом може стати ключовим елементом ефективного контролю над небажаним спілкуванням через месенджери. Деяке законодавство може містити спеціальні положення, що регулюють міжнародні аспекти боротьби зі спамом, особливо якщо месенджери використовуються у великому міжнародному масштабі.

Одним із можливих аспектів міжнародної співпраці може стати обмін інформацією та досвідом між країнами з метою розробки ефективних стратегій і методів боротьби зі спамом у месенджерах. Такий обмін досвідом дозволяє країнам впроваджувати кращі практики, сприяючи створенню єдиної міжнародної стратегії зменшення спаму та захисту користувачів.

Крім того, можливе укладання міжнародних угод або конвенцій, які визначають загальні стандарти та вимоги щодо боротьби зі спамом у месенджерах. Такі угоди можуть визначати правові рамки діяльності країн у сфері обміну інформацією, взаємодії та співпраці для ефективного управління спамом на міжнародному рівні.

Такий підхід сприяє створенню глобальної системи, спрямованої на зниження поширення спаму в месенджерах і захист користувачів на міжнародному рівні.

Співпраця країн у цьому напрямку може стати важливим інструментом забезпечення ефективності заходів контролю та регулювання спаму у світі.

1.4. Висновки до розділу

Розглядаючи різні аспекти проблеми спаму в месенджерах та враховуючи його зростаючий масштаб, серйозні загрози та негативний вплив на користувачів, а також розглянувши аспекти регулювання, можна зробити кілька висновків.

Зростання кількості користувачів месенджерів відзначається не лише збільшенням зручності взаємодії, але й розширенням простору для поширення спаму. Величезний потік небажаних повідомлень створює серйозні труднощі в управлінні та забезпеченні безпеки використання месенджерів.

У дослідженні виявлено, що еволюція спаму в месенджерах супроводжується не лише кількісним зростанням, але і стратегічною різноманітністю тактик та технік. Із збільшенням обсягу користувацької активності в цьому середовищі, спамери активно адаптуються, використовуючи різні маніпуляції та штучний інтелект для обхіду захисту та фільтрації.

Спам у месенджерах породжує низку негативних наслідків для користувачів, охоплюючи від марної витрати часу на його усунення та неприємностей до роздратованості користувачів. Від постійних спроб зламу особистих облікових записів до серйозної загрози конфіденційності та безпеки цілих компаній, спам стає важливою проблемою, яка вимагає комплексного підходу до її вирішення.

Регулювання спаму в месенджерах вимагає інтегрованої взаємодії між законодавством, політиками та технічними засобами. Аналіз законодавства, розробленого розробниками месенджерів, та роль технологій у боротьбі зі спамом вказує на необхідність єдиної та ефективної стратегії в цьому напрямі.

Загальна картина проблеми спаму в месенджерах є непередбачуваною та вимагає системних рішень. Проаналізувавши виявлені аспекти, вироблено висновок, що необхідно активно рухатися вперед, досліджувати різноманітні методи боротьби з цією проблемою та оцінювати їх, з метою визначення найбільш ефективних стратегій у протистоянні спаму в месенджерах.

Однією з перспективних тенденцій є вдосконалення технологій машинного навчання та інтелектуальних алгоритмів для розпізнавання та фільтрації спаму. Динамічне адаптивне навчання моделей може ефективно протидіяти змінам у структурі спам-повідомлень, що робить їх менш передбачуваними для спамерів.

У кінцевому висновку важливо підкреслити, що спільні зусилля стають вирішальним чинником у перемозі над спамом в месенджерах. Однак виклики та нові варіації спаму вимагають постійного вдосконалення та апгрейдів в сфері технологій та політик, спрямованих на забезпечення безпеки та комфорту користувачів.

Додатково, важливо вдосконалювати системи звітності та взаємодії користувачів у процесі виявлення та блокування спаму. Створення зручних та ефективних механізмів для швидкого та легкого повідомлення про спам дозволить оперативно реагувати на нові загрози та негайно впроваджувати заходи для їхнього усунення.

В цілому, висновки дослідження вказують на необхідність поєднання традиційних методів з інноваційними підходами для ефективного управління та боротьби зі спамом в месенджерах. Посилення технічних, правових та соціальних аспектів гарантує створення більш стійкої та динамічної системи захисту від небажаного контенту, що максимально враховує різноманітні виклики сучасного спаму в цифровому середовищі.

РОЗДІЛ 2
АНАЛІЗ МЕТОДІВ ЗАХИСТУ МЕСЕНДЖЕРІВ ВІД НЕБАЖАНИХ НАДСИЛАНЬ

Методи фільтрації спаму можна використовувати на різних етапах доставки повідомлення його отримувачу. Вони можуть брати участь у фазі роботи поштового клієнта, поштового вузла за межами поштового провайдера відправника та одержувача. На цих наведених етапах можна застосовувати профілактичні заходи. Ідеальним було б якомога швидше зупинити спам, щоб не витрачати такі ресурси, як пропускна здатність і час завантаження. Однак найпоширеніші методи боротьби зі спамом застосовуються на стороні одержувача повідомлень.

Комплексний захист від спаму складається з наступних етапів:

- аналіз відправника повідомлення;
- використання фільтрів;
- аналіз змісту повідомлення.

Технічно ці етапи базуються на основних підходах фільтрації спаму [14]:

- фільтрація за формальними ознаками повідомлення (за способом надсилання та дизайном);
- фільтрація за змістом (методи семантичної

фільтрації). Формальні методи включають:

- фільтрація за списками (поштові адреси, IP-адреси);
- фільтрація за формальними ознаками листа (наявність великої кількості відправників, відсутність адресата, формат, розмір тощо).

На рис. 2.1. зображено методи SPAM – фільтрації. [14]

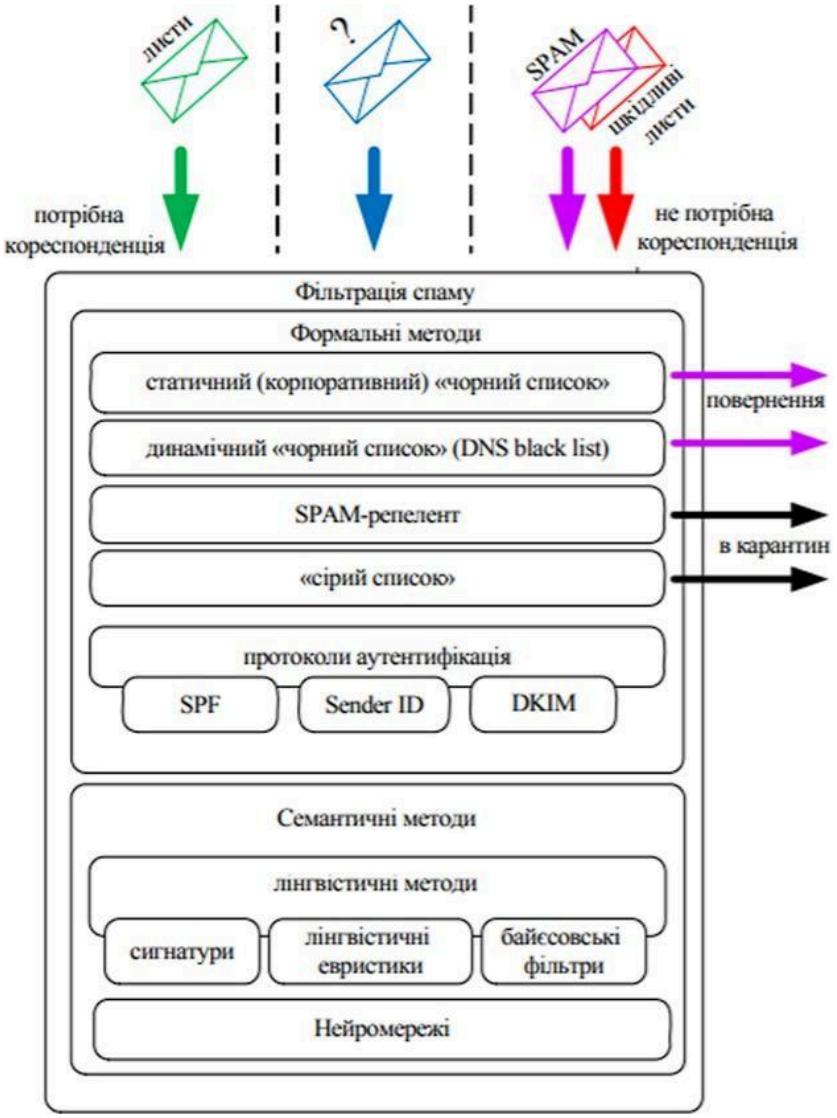


Рис. 2.1 Методи SPAM – фільтрації. [14]

2.1. Аналіз методів захисту месенджерів від спаму

Шляхи розсилки спаму можуть відрізнятися. Методи боротьби зі спамом можна розділити на короткострокові, середньострокові та довгострокові. Механізми фільтрації та блокування вважаються короткостроковими методами, оскільки їх застосування може бути обмежене інфраструктурою місцевої поштової організації з незначними змінами. Деякі методи на основі DNS, які впливають на структуру та вміст записів, можуть працювати протягом місяців, а іноді і років. Інфраструктура відкритих ключів (PKI - Public key infrastructure) і методи на основі ресурсів можуть бути довгостроковими через значні зміни та розширення інфраструктури.

В даний час існує досить велика кількість методів вирішення завдань ідентифікації спаму, які поділяються на кілька груп, що зображено на рисунку 2.2:

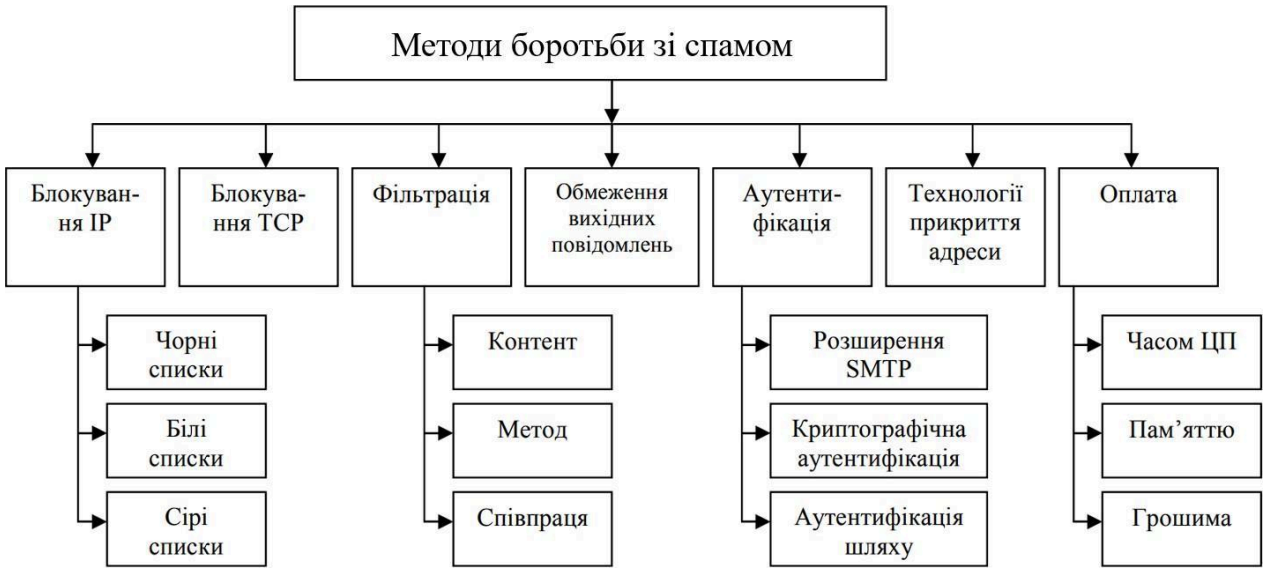


Рис. 2.2. Технічні засоби боротьби зі спамом в месенджерах

2.1.1. Метод блокування IP-адрес відправника

Як тільки користувач ініціює з'єднання SMTP, на мережевому або транспортному рівні встановлюється з'єднання TCP/IP із сервером SMTP. IP-адресу хоста відправника легко визначити, і це буде першою інформацією про користувача, яку отримує SMTP сервер. Якщо IP-адреса схожа на адресу користувача, який раніше надсилав спам, підключення може бути відхилено («чорний список»). Іноді в чорний список потрапляє цілий діапазон адрес, наприклад, адреси певного домену або провайдера.

Якщо ж IP-адреса належить довіреному користувачу, з'єднання встановлюється («білий список»). Термін «сірий список» описується як те, що будь-яка IP-адреса є частиною інформації і вона використовується для прийняття або відхилення підключення. Кожна спроба спочатку відхиляється, а дані (адреса відправника, адреса отримувача та тема надісланого повідомлення) цієї початкової невдалої спроби зберігаються. Якщо клієнт намагається повторно виконати невдалу

спробу в спеціальному тимчасовому вікні SMTP, сервер приймає це повідомлення як таке, що відповідає збереженим параметрам. Сірий список базується на припущенні, що більшість джерел спаму не надсилають повідомлення повторно, вважаючи, що поштовий сервер є недоступним.

Блокування IP-адреси легко реалізувати та не потребує багато ресурсів, оскільки рішення (прийняти/відхилити запит) приймаються на початковому етапі діалогового вікна SMTP. Але є й певні недоліки IP-блокування:

1. Це марна трата ресурсів, якщо IP-адреса відправника спам повідомлення підроблена. Підробка IP-адреси є проблемою, властивою протоколам TCP/IP. Проте підробка IP не є серйозною проблемою, оскільки: SMTP-з'єднання базуються на TCP-з'єднаннях із початковим 3-стороннім «рукостисканням»; блокування IP-адреси — не єдиний засіб боротьби зі спамом, який використовують сучасні MTA. На практиці рідко трапляються випадки, коли підробка IP-адреси організовується з метою розповсюдження спаму.

2. IP-блокування працює логічно, але воно може викликати помилки першого і другого роду.

Чорні списки (BL) можуть відрізнятися за багатьма параметрами. Деякі організації мають приватні BL і не надають їх Інтернет-користувачам. Деякі ж організації надають приватний доступ безкоштовно або за гроші в режимі реального часу, оскільки спамери змінюють хости, з яких відбувається надсилання. База даних може містити адреси вже виявлених спамерів або нелегальних серверів-посередників (XBL). Приклади таких баз даних є: Spamhaus block list (SBL), Arbitrary black hole list (ABL), чорний список системи доменних імен реального часу (Domain name system real time black list - DNSRBL), відкрита ретрансляційна база даних (Open Relay Database - ORDB), список користувачів віддаленого доступу MAPS (MAPS Dial-up user list) і система раннього попередження про спам (Spam Prevention Early Warning System - SPEWS). Spamhaus надає доступ до XBL списків, які можуть допомогти боротися зі спам-атаками через відкриті проксі-сервери, а також блокувати хробаків або віруси за допомогою вбудованих механізмів блокування спаму та інших типів троянів, які використовують шахраї.

Фільтрація за допомогою BL, мабуть, є одним з найпоширеніших методів на сьогоднішній день. BL поділяються на два види:

- статичні;

- динамічний.

Статичні BL – це списки IP-адрес і доменів. При виявленні листа, надісланого з однієї з цих адрес або доменів, сервер блокує це повідомлення. Програми фільтрації спам-повідомлень дозволяють вам заповнювати ці списки новими адресами, з яких зазвичай поширюється виявлений спам, і зберігати адреси в свій BL на ваш розсуд.

Статичне відновлення BL виконується періодично і частіше за все вручну. Ця обставина є недоліком, оскільки вимагає тимчасових ресурсів та адміністратора цих списків.

Динамічні BL майже кардинально відрізняються від статичних BL. DNSRBL — це мережева служба, що надається провайдером. Ці провайдери відстежують IP- адреси (і іноді доменні імена), які зазнали атаки зловмисників. [15]

Поштові фільтри, які підтримують використання динамічних BL, надсилають запит до постачальника BL, який містить адресу відправника, а також адреси, через які лист пройшов на шляху до одержувача. Якщо під час виконання запитів виявиться, що адреса знаходиться в BL постачальника послуг, то висока ймовірність того, що лист є спамом. Зв’язок із динамічними BL здійснюється через службу DNS, щоб перевірити, чи IP-адреси не зареєстровано в списках спамерів.

Такий метод «пасток» (SPAM-traps) зазвичай використовується для постійного оновлення BL. Сервер має спеціальний агент передачі пошти (MTA), який записує:

- IP-адреси комп'ютерів, які розсилають спам;
- скарги користувачів на небажані листи;
- автоматичне оновлення списку результатів інших фільтрів (наприклад, статистичного фільтра Байєса). [15]

Наступними недоліками BL можна вважати те, що BL очевидно не може містити абсолютно всі шахрайські адреси. Спамери, як правило, використовують IP протягом короткого часу, як приклад близько 2 годин, BL може не оновлюватися так

часто, що призведе до помилкових відмов. BL часто можуть помилково містити адреси або навіть діапазони адрес, що належать ISP або ESP, це відбувається оскільки деякі щахраї можуть використовувати інфраструктуру провайдера для своїх шахрайських цілей. Адреси електронної пошти тисяч або навіть мільйонів користувачів в цей час можуть бути заблоковані до того моменту поки адміністрація компанії-провайдера дізнається про цей факт і вирішить проблему. Чорні списки DNS призводять до збільшення трафіку та роблять DNS більш критичним ресурсом, оскільки він стає більш вразливим з точки зору цілісності та надійності.

Також так само, як і BL, можна використовувати білі списки (WL — Білий список), що зберігаються на сайті або доступні для спільного користування клієнтів. Якщо ці списки публікуються через DNS, вони називаються білими списками системи доменних імен (DNSWL). Плюсом є те, що WL не потрібно оновлювати в реальному часі. Але, на жаль, WL без додаткових заходів не є ефективним способом боротьби зі спамом, оскільки рівень помилок II типу буде дуже високим. WL рекомендовано використовувати в поєднанні з іншими методами боротьби. Оскільки повідомлення від хостів, перелічених у WL, не вимагають перевірки іншими методами захисту від спаму, WL слід вважати методом боротьби зі спамом I рівня.

Сірі списки використовуються для боротьби зі спамом на поштових серверах. Однак і тут є певні недоліки, які знижують ефективність цих списків.

По-перше, для кожного надісланого листа зберігається IP-адреса та невелика кількість інформації з повідомлення. Таким чином, команда для дііï SMTP не може бути відхилена або прийнята, доки не будуть отримані всі необхідні дані. Отже робимо висновок, що цей метод вимагає трохи більше часу для прийняття рішення на відміну від WL і BL. Також, необхідно місце для зберігання отриманих даних про надіслане повідомлення.

По-друге, сірі списки збільшують обсяги поштового трафіку, оскільки надіслане повідомлення пересилається.

По-третє, списки не вирішують проблеми зловживань з боку звичайних користувачів.

По-четверте, поштові системи, що мають великі обсяги надсилань, частіше складаються з кількох серверів відправки, які можуть по черзі пересилати відхилені повідомлення. Оскільки всі ці сервери відправки мають різні IP-адреси, повідомлення може бути втрачено.

2.1.2. Метод фільтрування надісланих повідомлень

Методи фільтрації спаму ділять отримані повідомлення на 2 категорії: спам і звичайна пошта (також відома як «ham»). Фільтри можуть перевіряти лист на наявність ключових слів або шукати певну спеціальну структуру листа, також можлива перевірка на використовувану мову (наприклад, для обраного фільтру задаємо параметр, що англійська вважається підозрілою). Фільтри спам повідомлень можуть використовуватися як клієнтом ESP, самим ESP, так і одержувачем повідомлення безпосередньо.

Методи фільтрації спам повідомлень постійно використовуються і можуть відрізнятися в залежності від методу і об'єкта перевірки. Наприклад, одні фільтри перевіряють лише тему або зміст повідомлення, інші перевіряють ці обидва пункти. Фільтрація спаму – це завдання, яке включає 2 фази: «навчання» та «класифікація». Таким чином, сфера машинного навчання може надати ресурси для забезпечення функціонування необхідних алгоритмів.

Спам-фільтри бувають розподіленими, тобто такими, що включають в свою роботу багато серверів, які аналізують та згодом надають інформацію про виявлені спам-повідомлення. Сумісні системи на основі підпису описані нижче.

Для того, щоб такі фільтри були ефективними в розпізнавані спам повідомлень, необхідно виконати деякі умови:

- оскільки зловмисники часто змінюють зміст та структуру своїх спам-повідомлень, фільтри необхідно постійно перебудовувати та міняти налаштування;
- оскільки організації або користувачі використовують різну термінологію, фільтри повинні бути налаштовані індивідуально під користувача;
- спам-фільтр має забезпечувати надійність і відмовостійкість.

Спамери завжди активно розробляють різні способи боротьби з спам-фільтрами, і одним із таких методів може бути ніби випадкове розділення слів на частини і навмисні орфографічні помилки, щоб фільтри не виявляли «заборонені» слова. Крім того, фільтри мають адаптуватись для вирішення проблем, враховуючи, що спам і «ham» все частіше набувають схожості.

Оскільки спам все більше стає схожим на звичайні «ham» повідомлення, спам-фільтрам все важче стає визначати його, а отже вони стають менш ефективними. Алгоритми фільтрації звісно зменшують ймовірність того, що спам потрапить до користувача. Але небезпека існує в тому, що спамери обирають шлях надсилати більше повідомлень в надії, що це допоможе їм обійти спам-фільтри. Тому фільтри посилюють проблему споживання ресурсів, спричинену спамом.

Варто зазначити, що існують цілі системи фільтрації спаму, які використовують різні методи та підсумовують частини результатів цих окремих методів в загальний результат. Наприклад, «SpamAssassin» - це організація, що використовує систему фільтрації, яка складається з: загального аналізу тексту, Байєсівської фільтрації та комбіновану фільтрацію баз даних, після виконання цих методів система призначає бали кожному повідомленню для визначення ймовірності того, що це спам. Кожен поштовий сервер встановлює та використовує власні обмеження, щоб відрізнити спам від звичайного повідомлення.

Якщо фільтр спаму виконується за правилами, вони можуть бути створені вручну або автоматично. Просте правило зазвичай стосується як теми повідомлення, так і його змісту. Наприклад, якщо в темі повідомлення є «СУПЕР ПРОПОЗИЦІЯ», а в самому повідомленні — «Шановні пані та панове», це спам. Більш детальне обговорення параметрів правил фільтрів представлено в роботах Коена [16] і Кроуфорда [17]. Основний недолік застосування таких правил полягає в тому, що зловмисник може їх обійти, трохи змінивши написання слів у повідомленні. Наприклад у слові «пропозиція» замість літер «і» зловмисник пише «1», таким чином фільтр вже не розпізнає цього слова «ПРОПОЗИЦ1Я».

За допомогою фільтру спаму на основі підпису повідомлення стискається до розміру підпису, наприклад, за допомогою хеш-функції. Однак для ефективності

цього методу важлива його стійкість до невеликих змін у повідомленнях, наприклад, згадування імені в привітанні, а частота оновлення бази підписів є досить важливою, оскільки спам-повідомлення змінюються день у день. Загальною процедурою викриття спаму серед звичайних повідомлень є створення підпису отриманого повідомлення та порівняння його з підписами, які вже є в базі даних.

VipuFs Razor — це розподілена уніфікована мережа виявлення спаму. Поштові клієнти підключаються до каталогу спаму, що постійно оновлюється користувачами. Спам ідентифікується за статистичними та випадковими підписами, які ефективно розпізнають зміну вмісту повідомлення. Якщо спам пройшов фільтр, користувач може створити новий підпис самостійно після його знаходження та відправити його в каталог. Якщо подібні підписи надходять від інших користувачів, то довіра системи до вибору цього користувача підвищується, в протилежному ж випадку система поступово перестає вірити оцінкам даного користувача.

Damiani пропонує фільтр спаму на стороні сервера, який працює в одноранговій мережі. Для кожного оскарженого повідомлення розраховується 256-бітний дайджест, що є абсолютно стійким до певних змін в повідомленні. Два повідомлення вважаються однаковими, якщо їхні дайджести відрізняються не більше ніж на 74 біти. Автори цього методу використовують трирівневу структуру фільтру, яка складається з:

- рівня безпосередньо користувача;
- рівня з поштовим сервером;
- рівня, де збираються скарги.

Підключені до цієї системи сервери обмінюються інформацією із серверами вищого рівня, які також обмінюються інформацією з іншими своїми рівнями.

Zhou, як і Damiani також використовує однорангову мережу. Але замість підсумків створюється набір контрольних сум кожного повідомлення, що розповсюджується через загальну розширену систему децентралізованого розташування та маршрутизації об’єктів (DOLR - Decentralized object location and routing system).

2.1.3. Блокування TCP

Блокування TCP — один із методів фільтрації спаму в месенджерах, який призначений для ефективного виявлення та обмеження передачі небажаних повідомлень. Цей метод заснований на моніторингу та аналізі TCP-трафіку, який виникає під час взаємодії між користувачами та сервером обміну повідомленнями. Нижче наведено основні аспекти блокування TCP у фільтрації спаму.

Блокування TCP дозволяє аналізувати структуру та характеристики трафіку, який виникає під час обміну повідомленнями. Використовуючи алгоритми машинного навчання та евристичні методи, можна визначити типові характеристики спам-повідомлень.

Цей метод дає змогу виявляти шаблони спаму. Це може включати аналіз текстового вмісту повідомлень, частоти надсилання та інших ознак, які вказують на небажані повідомлення.

Блокування TCP дозволяє виявляти зміни в характеристиках трафіку, які можуть вказувати на нові методи або технології, які використовують спамери. Це дозволяє швидко налаштувати правила фільтрації.

Програму для блокування TCP можна інтегрувати в загальну систему безпеки месенджера, включаючи моніторинг вразливостей і запобігання атакам.

Ефективне використання блокування TCP передбачає динамічне оновлення системи, яка може адаптувати правила фільтрації для врахування змін у поведінці спамерів і еволюції їхніх методів.

Враховуючи зазначені аспекти, TCP-блокування стає ефективним інструментом для запобігання поширенню спаму в месенджерах та забезпечення безпеки та комфорту користувачів.

2.1.4. Гібридні спам-фільтри

«Мішок слів» («Bag of words») — це модель, що дозволяє з текстових даних витягувати необхідні для неї функції. Цей алгоритм використовує для слів числове представлення. Таке представлення пізніше вводиться в будь-який інший алгоритм для подальшого аналізу повідомлення на наявність спаму. [18]

Алгоритм називається «Bag of words», тому що в ньому втрачається порядок слів або структура речень. Важливим є лише наявність або присутність слова.

Ця модель розглядається так - є великий мішок, спочатку порожній, і є словник. Слова вибираються одне за одним і складаються в мішок, додаючи їх частоту появи, а потім найпоширеніші вибираються як символи для проходження за вибраним алгоритмом. Таким чином, алгоритм підтримує ідею, що з подібних слів складаються подібні документи.

«DNS lookup» — це такий фільтр, який надає можливість ретельно відстежувати веб-контент, до якого мають доступ співробітники, гості та клієнти організації, що підключені до дротової або Wi-Fi мережі.[19]

DNS-фільтр спрацьовує коли користувач отримує доступ в Інтернет мережу. Для того, щоб користувач міг відвідати веб-сторінку чи веб-сайт, його URL-адресу за допомогою пошуку DNS спочатку пеертворюють на унікальну IP-адресу. Потім надсилається запит на сервер DNS, щоб знайти IP-адресу, яка пов’язана з початковою оригінальною URL-адресою веб-сторінки або веб-сайту. Якщо знайдено, вона буде надіслано в браузер. Після цього браузер виконує з’єднання з URL-адресою сайту, і він відображається для користувача. [19]

Фільтр сервера DNS робить дкілька перевірок, щоб зрозуміти, чи потрібно надавати IP-адресу певному ресурсу. Якщо цей ресурс не дотримується попередньо визначентх політик, замість надання IP-адреси користувач перенаправляється на локальну сторінку блокування та отримує попередження, що йому заборонено доступ до цього веб-ресурсу.

«Простий протокол передачі пошти на основі сервера» (Simple Mail Transfer Protocol (SMTP) server based) — це відносно простий текстовий протокол, у якому з’єднання завжди ініціюється відправником. SMTP — це синхронний протокол, який складається з серії команд, які надсилає клієнт, і з відповідей від сервера.

Відправником зазвичай є поштовий клієнт або поштовий сервер кінцевого користувача.[20]

SMTP був розроблений як транспортний протокол і протокол доставки, тому системи, які користуються SMTP, повинні завжди бути запущеними. SMTP часто виконує функцію передачі повідомлень між користувачами електронної пошти, які не мають можливості виконувати роль сервера.

Система захисту поштового сервера SMTP включає в себе захист збору адрес, чорні та сірі списки, аналіз заголовків та вкладень,. Крім того, він використовує алгоритм масової перевірки, що удосконалюється з кожним роком. Хороша система безпеки поштового сервера може обробляти сотні або навіть тисячі електронних листів за секунду без значного збільшення завантаження мережі.[20]

«На основі зображень» (Image based) – це такий фільтр, що перевіряє, чи містять вкладені зображення спам. Він поділяється на групи:

- метод на основі заголовків - суть цього способу полягає в аналізуванні заголовків електронних листів. Їх структура складається з декількох полів, які містять необхідну інформаційну область. Деякі елементи заголовків також використовуються, але як частина машинного навчання;
- контент-орієнтовані методи (на основі вмісту) - контент-орієнтовані стратегії з використанням виділення ознак і аналізу контенту зображення. Цей вид спам- фільтрів застосовується для аналізування поверхні зображення та вивчення його сутності. Наприклад цей фільтр цікавлять такі характеристики, як затінення, краї поверхні зображень тощо;
- методи на основі OCR – ці методи застосовують інструмент оптичного розпізнавання символів для того, щоб вилучити текст, вбудований в надіслане зображення. Це електронна або навіть механічна версія перевірених шаблонів друкованого, машинописного, чи навіть рукописного вмісту в машинно-кодований.

Класифікація спаму в текстових повідомленнях все ще є досить ефективною та успішною. Однак спамери продовжують винаходити способи, які допомагають уникнути блокування. Наприклад, включивши зображення або кольоровий текст у тіло повідомлення, як показано на малюнку 2.3.

Рис. 2.3. Кольорове спам-повідомлення [21]

Цей приклад показує, що смислове навантаження несе не тільки текст, а й кольори. Звідси випливає, що окрім класифікації безпосередньо тексту, є можливість обрати для додавання додаткові ознаки, які можна побачити в повідомленні. Якщо потрібно обробити надіслане в повідомленні зображення, фільтру необхідно розпізнати всі символи на ньому та скласти їх словами. Саме це і робить OCR. Після того, як фільтр розпізнав їх, до кожної літери/слова/речення прив’язується відповідна колірна характеристика, і ці дані повинні бути класифіковані, щоб визначити, чи є повідомлення спамом [21].

Наступний метод «Спільні фільтри» (Collaborative filters) — фільтр, який застосовує антиспамовий підхід на основі спільноти, збираючи інформацію відтисяч або навіть мільйонів користувачів електронної пошти. Користувачі таких систем можуть ідентифікувати вхідну пошту як звичайну або спам, і ці мітки передаються до центральної бази даних. Коли певна кількість користувачів позначає один і той самий лист як спам, цей фільтр автоматично блокує цьому листу доступ до інших скриньок спільноти.[22]

Коли система фільтрації спільного вмісту має велику базу активних користувачів, вона може швидко зупинити хвилю спаму, іноді за лічені хвилини. Але є можливий недолік цього методу і це те, що група спамерів може зібратися у великій кількості та згодом почати видавати себе за нормальних звичайних користувачів системи, і якщо вони так активізуються, вони матимуть змогу спотворювати результати, позначаючи спам-повідомлення як нормальні, не спамові.

«Система викликів/відповідей» (Challenge/Response System) – метод, який застосовуючи систему викликів/відповідей, блокує спам пошту, перед відправкою користувач змушений виконати завдання, перш ніж повідомлення буде доставлено. Коли відправник захоче надіслати комусь, хто використовує систему викликів/відповідей, перед відправкою він отримає лист, в якому буде завдання відвідати веб-сайт і ввести код, який там відображається, у спеціальну форму. Якщо він впорається з цим завданням, його електронний лист (і вся майбутня електронна пошта) буде доставлено отримувачу. Якщо відправник не виконує данне завдання через певний час, надсилання повідомлення буде скасовано.

Цей метод для блокування спаму є діючим, оскільки зазвичай лише людина може відповісти на прохання пройти завдання. Це спрацьовує тому, що спамери зазвичай покладаються на автоматизовані програми надсилання повідомлень, щоб була можливість надсилати тисячі й мільйони електронних листів одночасно, і дуже рідко спамери перевіряють, які повідомлення надходять у відповідь на їх розсилку. Але навіть якщо вони побачать повідомлення, не будуть відповідати на нього, оскільки це створює ризик стати викритим, а цього вони звісно уникають.

Однак є і мінуси - метод викликів/відповідей може блокувати підписані новини або ваші узгоджені акційні пропозиції, оскільки сьогодні такі повідомлення частіше всього відправляються автоматизованими програмами розсилки клієнтам. Також недоліком є й те, що деякі члени організацій можуть не витрачати час на виконання виклику або можуть банально не зрозуміти чи не побачити, не відреагувати вчасно на зміст такого повідомлення, а це означає, що їх електронні листи не дійдуть до одержувача. Також завжди слід розуміти, що присутня невелика ймовірність того, що і відправник, і одержувач застосували метод виклику/відповіді,

тоді їх програми захисту від спаму, на жаль, будуть продовжувати кидати виклики і блокувати одне одного, а в такому випадку електронна пошта опиняється в циклі недоступності.[22]

2.1.5. Машинне навчання як основа фільтру блокування спаму

Такі лідери світового ринку, як Google, для фільтрації спаму працюють зі штучним інтелектом TensorFlow, що спрямований на блокування небажаних надсилань. Більше ніж 100 мільйонів зловмисних повідомлень щодня блокуються, перш ніж зловмисники зможуть завдати шкоди цільовим компаніям або окремим особам.

Спам – це не тільки просто незручність, це загроза безпеці та конфіденційності даних користувача. AI забезпечує додаткові заходи безпеки, такі як брандмауери та виявлення шкідливих програм, щоб не допустити несанкціонованого витоку даних. Якщо ж користувачі ігноруватимуть оновлення програмного забезпечення, це призведе до того, що лінії захисту (наприклад, брандмауер) можуть ставати менш ефективними. Фільтрація небажаних повідомлень з використанням AI може вдосконалювати заходи безпеки організацій і запобігати витоку інформації, покращуючи рівні безпеки і конфіденційності.

Фільтрація спаму з використанням штучного інтелекту, дає змогу IT-спеціалістам вчасно проводити технічне обслуговування. Листи приходять у папку "Вхідні" з відповідною швидкістю. Іноді кількість спам-повідомлень стає більшою за кількість релевантних електронних листів, і для більшості людей спам надсилань надто багато, щоб відсіяти їх або мати час їх для обробки. AI працює зі швидкістю, яка перевищує наші когнітивні та емоційні можливості.

Коли AI відфільтровує спам повідомлення, він усуває більше технологічного навантаження, ніж дратівливий безлад у вхідних. Для компаній це економить мережевий простір і фінанси.

Баєсівські фільтри — це фільтри, які аналізують як спам, так і хороші повідомлення, щоб обчислити ймовірність появи різних небажаних слів в обох випадках, а потім створити списки таких слів.[23]

Простими словами принцип роботи можна описати так, що якщо певне слово ніколи не з’являється у спам-повідомленнях, але часто зустрічається у звичайних отриманих листах, то ймовірність того, що це слово буде визначено як спам, майже нульова. Наприклад, приходить багато звичайних електронних листів, які в своєму вмісті мають слово «звіт». З цієї причини зменшується ймовірність того, що повідомлення, які містять слово «звіт», будуть позначені як спам. Або ж якщо лист часто надходить зі словом «акція» і визначається як спам, то такі листи в майбутньому будуть блокуватися, оскільки це вважається як спам.

Такі фільтри розглядають в тілі повідомлення:

- текст в самому вмісті повідомлення;
- текст заголовку повідомлення (в тому числі поля відправника);
- HTML та CSS-код (форматування тексту, кольори і тому подібне);
- відслідковується пари слів, а також частота і місця їх з’явлень в тілі повідомлення.

Коли отримується новий електронний лист, байєсівський фільтр на основі цих пунктів аналізує отриманий лист та розраховує ймовірність того, що повідомлення є спамом.

Використовуючи цю автоадаптивну техніку, байєсівські фільтри можуть навчатися на власних даних, а також на рішеннях користувачів (якщо клієнт власноруч виправлятиме неправильно оцінений лист).

Як і характеристики спаму, характеристики звичайної електронної пошти такі ж важливі для процесу байєсівської фільтрації,. Оскільки байєсівські фільтри підлаштовуються під кожного користувача, спамерам їх важче обійти. Ці фільтри можуть адаптуватися майже до будь-яких хитрощів, що застосовують зловмисники, коли намагаються обійти обмеження.

Незалежно від того, наскільки ефективним є байєсівський фільтр, окреме слово або характеристика, які часто фігурують у звичайній пошті, можуть бути

достатньо важливими, щоб запобігти класифікації повідомлення, яке містить їх, як спам. Таким чином, якщо спамери знайдуть спосіб дізнатися такі слова у звичайній пошті, вони зможуть включити одне з них у спам і розіслати зловмисні листи навіть якщо ваша пошта захищена за допомогою добре навченого байєсовського фільтра.

Машинне навчання інформує AI під час сканування вхідної електронної пошти. AI шукає повідомлення, які викликають підозру у спамі, наприклад, якщо вони містять:

- IP-адреси та URL-адреси, які можуть бути шкідливими або зловмисними;
- ключові слова, які здаються AI підозрілими;
- шкідливі ненадійні вкладення;
- заміна символів (наприклад використання цифр замість літер і навпаки), часті однакові синтаксичні або орфографічні помилки;
- використання спеціальних символів та/або емодзі в надмірній кількості.

Використовуючи бази даних із величезною кількістю посилань, IP та URL-адрес, AI може сканувати та аналізувати вміст надісланих повідомлень на наявність підозрілої активності за лічені хвилини або навіть секунди. Перевірка та аналіз можуть підтвердити, що посилання містить фальшиві сторінки входу або підроблені форми введення особистих даних, корпоративні підписи співробітників.

Чим більше AI аналізує, навчається та вдосконалює свої навички й знання, тим точніше він визначає спам серед нормальних повідомлень. AI автоматизує майже всі процеси, які в більшості фільтрів здійснюються вручну користувачем, наприклад створення білих і чорних списків.

Окрім аналізування ключових слів, вмісту та заголовків вхідних повідомлень, AI також використовує наступні методи фільтрації спаму:

- AI порівнює нові вхідні повідомлення з уже наявними листами, що зберігаються на його серверах, іншими словами можна сказати - метод подібності;
- AI оцінює нові вхідні листи на основі шаблонів легітимних і нелегітимних небажаних спам-повідомлень;

- AI класифікує окремі повідомлення, після чого він порівнює потенційні спам-повідомлення із спеціалізованими категоріями. Цей метод є адаптивним, оскільки він проводить коригування спеціалізованих категорій даних.

Чим складніші алгоритми використовує AI, тим більше він стає стійкішимі адаптивним до нових загроз і хитрощів зі сторони спамерів. Одним з прикладів, до якого AI має адаптуватись є те, що зміст спам-повідомлень постійно міняється в залежності від світових тенденцій і актуальних міжнародних подій. Яскравим прикладом були часи пандемії, коли зловмисники використали цю подію і масову паніку зі сторони користувачів для розповсюдження зловмисних повідомлень. Сьогодні ж зловмисники використали іншу інфоподію – війну в Україні, при чому це торкнулось не лише жителів нашої країни, а й користувачів з Європи та Америки. Такі подібні інформаційні події міжнародного масштабу викликають коливання в наборі даних машинного навчання і подальше відхилення, але AI можна навчити враховувати ці події і як наслідок коливання в наборах даних.

Машинне навчання є одним із напрямів технологій, що розвивається найшвидше, але навіть незважаючи на те, що цей метод часто називають «машинним навчанням», воно не стосується чогось одного, це узагальнюючий термін, який описує сукупність різних концепцій і технік.

2.2. Оцінка ефективності та недоліків існуючих рішень

З метою створення найбільш об'єктивного образу, буде ретельно розглянуто кожен метод і визначено його переваги та недоліки. Оцінка буде зосереджена не лише на загальній ефективності, але й на впливі на користувачів, легітимних відправників, та загальною придатністю до реальних умов використання.

Висвітлення ефективних методів фільтрації спаму є ключовим аспектом розробки ефективної системи захисту від небажаної комунікації. Ретельний аналіз дозволить визначити оптимальний метод, який забезпечить якість фільтрації, при цьому мінімізуючи негативний вплив на легітимних користувачів.

Проведення оцінки та порівняння методів є важливим кроком у напрямку підвищення якості і надійності систем захисту в месенджерах, що в свою чергу сприятиме поліпшенню користувацького досвіду та врахуванню сучасних викликів у сфері кібербезпеки.

Отже, почнемо з методу блокування IP-адрес відправників.

Переваги:

Блокування певних джерел: цей метод дозволяє відфільтрувати спам із певних IP-адрес відправників, що ефективно для блокування відомих джерел спаму.

Простота впровадження: Реалізація цього методу є відносно простою та швидкою.

Недоліки:

Проблема з динамічними IP-адресами: у деяких випадках відправники можуть використовувати динамічні IP-адреси, що ускладнює ефективність блокування.

Можливість блокувати законних відправників: іноді законні користувачі можуть випадково заблокувати себе за допомогою спільних IP-адрес або VPN.

Обмеження:

Динамічні IP-адреси: багато користувачів використовують динамічні IP-адреси, які можуть змінюватися кожного разу, коли вони підключаються до мережі. Це ускладнює постійне блокування спамерів, які можуть легко змінити свої IP-адреси.

Можливість блокувати законних відправників: якщо спамери використовують IP-адреси, які відповідають законним користувачам або компаніям, цей метод може призвести до блокування законних відправників.

Використання спільних IP-адрес: у деяких випадках група користувачів може використовувати одну IP-адресу (наприклад, через NAT), і блокування може вплинути на всю групу.

Відсутність ідентифікації конкретного користувача: блокування IP-адреси не дозволяє ідентифікувати конкретного користувача, тому у випадку спільної IP-адреси може статися неправильне блокування.

Ці обмеження підкреслюють важливість використання інших методів фільтрації, щоб уникнути негативних наслідків для законних користувачів.

Рекомендації:

Розгляд динамічних IP-адрес: важливо регулярно оновлювати список заблокованих IP-адрес, беручи до уваги можливі зміни в динаміці IP-адрес.

Наступний метод - метод чорних, сірих та білих списків:

Переваги:

Точна фільтрація: Списки дозволених (білих) та заборонених (чорних) відправників можуть забезпечити точну фільтрацію спаму та легітимних повідомлень.

Гнучкість налаштувань: Користувач може налаштовувати власні списки, враховуючи свої унікальні вимоги та пріоритети.

Зменшення кількості ложних сигналів: Точне управління списками може допомогти уникнути блокування легітимних відправників.

Недоліки:

Підтримка списків: Постійне оновлення списків вимагає значних ресурсів та уваги.

Можливість обходу: Деякі спамери можуть знайти способи обійти фільтрацію, наприклад, змінюючи відправницькі IP-адреси.

Рекомендації:

Регулярне оновлення списків: Забезпечення актуальності списків є критично важливим для ефективної роботи методу.

Обмеження:

Підтримка списків: Регулярне оновлення та управління списками може бути витратними за ресурсами.

Можливість обходу: Деякі винахідливі спамери можуть знаходити шляхи обходу фільтрації шляхом зміни IP-адрес або використання інших хитрощів.

Людський фактор: Налаштування списків може вимагати знань та уваги від користувача.

Ложні сигнали: Несприятливі умови можуть призвести до помилкового блокування чи пропуску повідомлень.

Метод чорних, сірих та білих списків є ефективним інструментом, але важливо розуміти його обмеження та вживати заходи для оптимізації його функціонування.

Метод фільтрування надісланих повідомлень за допомогою фільтрів, що перевіряють вміст, заголовки, вкладені зображення листа:

Переваги:

Висока точність: Аналіз контенту повідомлення дозволяє виявити характеристики спаму з високою точністю.

Адаптивність: Фільтри можуть навчатися на основі нових методів або характеристик, що поліпшує їхню ефективність з часом.

Можливість виявлення нових форм спаму: Здатність аналізувати зміст дозволяє виявляти нові форми спаму.

Недоліки:

Ресурсозатратність: Аналіз контенту може вимагати значних обчислювальних ресурсів, особливо при великому обсязі повідомлень.

Ложні сигнали: Вища точність також може вести до збільшення кількості ложних сигналів, що може призводити до блокування легітимних повідомлень.

Рекомендації:

Оптимізація ресурсів: Ефективне використання обчислювальних ресурсів та оптимізація фільтрів можуть покращити продуктивність методу.

Постійне навчання: Апаратне навчання фільтрів на нових прикладах спаму може допомогти виявляти нові види небажаних повідомлень.

Обмеження:

Ресурсозатратність: Обчислювальні ресурси можуть стати обмежувальним фактором, особливо для великих обсягів повідомлень.

Ложні сигнали: Точний аналіз контенту може викликати ложні сигнали, що призводить до блокування легітимних повідомлень.

Невідомі форми спаму: Нові форми спаму, які не мають характерних ознак, можуть проходити мимо фільтрів.

Залежність від якості аналізатора: Ефективність методу може залежати від якості використовуваного аналізатора, його можливостей в розпізнаванні характеристик спаму.

Метод "Блокування TCP":

Переваги:

Виявлення типових характеристик спаму: Аналіз структури та характеристик TCP-трафіку дозволяє виявляти типові ознаки спам-повідомлень.

Виявлення шаблонів спаму: Метод може аналізувати текстовий вміст повідомлень, частоту надсилання та інші ознаки для виявлення шаблонів спаму.

Адаптація до нових методів спамерів: Можливість виявлення змін у характеристиках трафіку дозволяє швидко адаптувати систему до нових методів спамерів.

Недоліки:

Неідентифікація конкретного користувача: Блокування TCP не дозволяє ідентифікувати конкретного користувача, що може призвести до невірного блокування у випадку спільного використання IP.

Врахування динамічних IP: Користувачі, які використовують динамічні IP-адреси, можуть обходити блокування.

Рекомендації:

Динамічне оновлення правил: Регулярне оновлення списку блокованих IP дозволяє враховувати зміни у динаміці IP-адрес.

Інтеграція в систему безпеки: Блокування TCP може бути ефективним, якщо інтегроване в загальну систему безпеки месенджера.

Обмеження:

Неідентифікація конкретного користувача: Без можливості ідентифікувати конкретного користувача, може виникнути проблема з невірним блокуванням у випадку спільного використання IP.

Врахування динамічних IP: Користувачі, які використовують динамічні IP-адреси, можуть обходити блокування, оновлюючи їхні IP.

Підвищена ресурсозатратність: Збільшення об'єму аналізу трафіку може призводити до підвищення ресурсозатрат, особливо при великому обсязі повідомлень.

Метод гібридних спам-фільтрів:

Переваги:

Комбінує різні методи: Гібридний підхід, що об'єднує кілька методів, може бути ефективним у виявленні різноманітних типів спаму.

Використання статистики та ключових слів: Модель «Bag of words» використовує статистику входження слів та ключових виразів, що може покращити точність виявлення.

Недоліки:

Піддатливість до нових методів спаму: Гібридні фільтри можуть бути менш ефективними при з'яві нових методів обхіду фільтрів.

Складність реалізації: Побудова та підтримка гібридних систем може вимагати значних зусиль та ресурсів.

Рекомендації:

Систематичне оновлення: Постійне оновлення моделей та алгоритмів для виявлення нових видів спаму.

Тестування на реальних даних: Важливо тестувати ефективність на реальних даних для підтвердження точності.

Цей метод може бути ефективним, але йому притаманні деякі обмеження, які важливо враховувати при використанні.

Обмеження:

Складність побудови та обслуговування: Розгортання та підтримка гібридних фільтрів може вимагати значних зусиль та ресурсів, особливо при використанні різних методів.

Неефективність проти нових методів спаму: Гібридні фільтри можуть виявлятися менш ефективними, коли стикаються з новими методами обходу фільтрів, оскільки їх моделі можуть бути непідготовленими до нових викликів.

Необхідність постійного оновлення: Систематичне оновлення та адаптація гібридних систем є критично важливими для збереження їхньої ефективності з часом.

Залежність від якості вихідних даних: Ефективність гібридних фільтрів може залежати від якості вихідних даних та навчальних наборів.

Вимоги до обчислювальних ресурсів: Залежно від складності методів, гібридні фільтри можуть вимагати значних обчислювальних ресурсів для швидкого та точного аналізу повідомлень.

Розглядаючи ці обмеження, важливо збалансувати вигоди та недоліки при впровадженні гібридних фільтрів у систему фільтрації спаму.

Метод "Блокування спаму на основі машинного навчання":

Переваги:

Адаптивність: Системи машинного навчання можуть адаптуватися до нових форм і методів спаму, роблячи їх ефективними проти різноманітних загроз.

Висока точність: Завдяки навчанню на великих обсягах даних, системи машинного навчання можуть досягати високої точності у виявленні спаму.

Аналіз контексту: Машинне навчання може враховувати контекст повідомлень, що поліпшує виявлення спаму на основі смислового змісту.

Недоліки:

Потреба у великій кількості данних: Ефективне функціонування систем машинного навчання вимагає значних обсягів даних для навчання та підтримки.

Залежність від навчання: Моделі машинного навчання можуть бути вразливі до маніпуляцій та обхідних методів, якщо не навчатися на відповідній інформації.

Схильність до фальшивих позитивів та негативів: Навіть з великою точністю, існує ризик невірного класифікування повідомлень як спаму або навпаки.

Обмеження:

Великі витрати обчислювальних ресурсів: Застосування методів машинного навчання може вимагати значних обчислювальних потужностей, особливо при навчанні та вдосконаленні моделей.

Необхідність постійного оновлення: Машинне навчання потребує постійного оновлення для врахування змін у характеристиках спаму та нових загроз.

Залежність від якості даних: Якість навчальних даних може впливати на ефективність системи машинного навчання. Неправильно анотовані чи ненадійні дані можуть призвести до недостатньої продуктивності.

Ризик фільтрації легітимних повідомлень: Недостатнє навчання може призвести до включення легітимних повідомлень до категорії спаму або навпаки.

Оцінка ефективності та обмежень цього методу є важливою для забезпечення надійного фільтрування спаму в електронній пошті.

Після докладного аналізу та порівняння різних методів боротьби із спамом можна зробити висновок, що найефективнішим методом є "Метод блокування спаму на основі машинного навчання." Цей метод відзначається високою ефективністю завдяки використанню алгоритмів машинного навчання та штучного інтелекту для виявлення та блокування спаму.

"Метод блокування спаму на основі машинного навчання" дозволяє аналізувати широкий спектр параметрів та ознак повідомлень, враховуючи їхню структуру, вміст, та частоту відправлення. Такий підхід дозволяє ефективно розпізнавати нові форми спаму та адаптуватися до змін в методах атак спамерів.

Високий рівень ефективності "Методу блокування спаму на основі машинного навчання" пояснюється його здатністю до самонавчання та автоматичного вдосконалення відповідно до змін у характеристиках спаму. Додатково, цей метод дозволяє враховувати особливості поведінки користувачів та адаптувати фільтрацію до індивідуальних потреб.

Таким чином, враховуючи комплексність та гнучкість "Методу блокування спаму на основі машинного навчання," можна стверджувати, що цей підхід є оптимальним для забезпечення високої якості фільтрації спаму в електронній пошті та інших комунікаційних платформах.

2.3. Огляд передових технологій у галузі захисту месенджерів від спаму

У сучасному інформаційному суспільстві проблема спаму є актуальною, особливо в контексті месенджерів та електронної пошти. Щоб подолати цю проблему, вчені та інженери впроваджують інноваційні технології та методи боротьби з небажаними повідомленнями. Розглянемо кілька перспективних напрямків у цьому плані.

Штучний інтелект (ШІ) і машинне навчання (МН) використовуються для аналізу текстового та візуального вмісту повідомлень. Це дає змогу ідентифікувати спам на основі семантичних і мовних елементів, а також графічного вмісту. Глибинне навчання використовується для аналізу структурних елементів повідомлень для виявлення складних атак спаму.

У сучасну цифрову епоху, коли електронний зв’язок стає невід’ємною частиною нашого повсякденного життя, проблема спаму набуває нових масштабів і вимагає інноваційних рішень. Серед таких рішень ключове місце посідає використання ШІ і МН – двох технологічних і наукових напрямків, що стрімко розвиваються.

ШІ використовується для покращення процесу фільтрації спаму шляхом аналізу семантики та вмісту повідомлень. Алгоритми на основі ШІ намагаються розпізнавати та класифікувати повідомлення на основі їх семантичного навантаження. Це дозволяє більш ефективно виявляти спам-повідомлення, які можуть мати різну форму та структуру.

МН виявилося надзвичайно корисним у боротьбі зі спамом, оскільки воно здатне адаптуватися до нових сценаріїв і методів атак. Алгоритми МН можуть навчатися на великих обсягах даних, враховуючи мінливі характеристики спаму. Це дозволяє системі фільтрації бути ефективнішою при виявленні та блокуванні нових типів спаму.

Глибинне навчання, що є підгалуззю МН, фокусується на аналізі структури повідомлень. Він використовується для виявлення та класифікації спаму з урахуванням складності та властивостей структурних елементів повідомлень.

Останні досягнення в області ШІ та МН сприяють створенню більш точних і надійних систем фільтрації. Використання нейронних мереж, файлів моделей і

вдосконалених методів аналізу даних робить ці технології надзвичайно потужними для виявлення та блокування спаму.

Оскільки ШІ та МН швидко розвиваються, можна очікувати подальших удосконалень систем фільтрації спаму. Перспективи включають більш точне виявлення аномальної поведінки, розширення можливостей графічного аналізу контенту та розробку гібридних систем, які поєднують різні аспекти ШІ та МН для оптимальної ефективності.

Усе це вказує на те, що роль штучного інтелекту та машинного навчання у боротьбі зі спамом лише зростатиме та забезпечуватиме нам безпечніше та ефективніше середовище електронного спілкування.

Системи аналізу поведінки користувачів (САПК) використовують алгоритми для виявлення аномальних моделей, які можуть вказувати на активність спаму. САПК є передовими інструментами в цьому відношенні, що забезпечує новий рівень ефективності та адаптивності.

Однією з ключових функцій САПК є аналіз змін у звичній поведінці користувача. Створюючи унікальний «профіль» для кожного користувача, система може виявляти невідповідності та аномалії, які можуть вказувати на спроби розсилання спаму. Наприклад, зміна звичайного часу активності або надмірна кількість невдалих спроб авторизації можуть бути індикатором активності спаму.

Системи аналізу поведінки користувачів використовують методи машинного навчання для автоматичного виявлення моделей і аномалій у поведінці. Алгоритми, навчені на великих обсягах даних, можуть виявляти найменші відмінності, які може бути важко виявити за допомогою традиційних методів фільтрації.

У порівнянні з традиційними методами САПК демонструє високий ступінь адаптивності. Він може виявляти нові сценарії та методи атак і швидко адаптуватися до змін у середовищі спаму. Це особливо важливо, оскільки спамери постійно змінюють свої методи, намагаючись уникнути виявлення.

Системи аналізу поведінки користувачів можуть бути вдосконалені елементами ШІ. Використання алгоритмів ШІ дозволяє системі автоматично

навчатися та вдосконалюватися з часом, щоб забезпечити найвищу точність виявлення спаму.

Системи аналізу поведінки користувачів також враховують контекстне розуміння поведінки. Вони враховують конкретний контекст використання платформи чи сервісу, коли визначають, чи є певні дії спробами спаму чи правомірними.

Хоча САПК є потужним інструментом у боротьбі зі спамом, він також має свої недоліки. Найважливішим з них є необхідність великої кількості даних для ефективного функціонування алгоритмів машинного навчання. Вони також можуть пропустити нові, раніше не виявлені методи атаки, якщо вони налаштовані погано.

Системи аналізу поведінки користувачів є важливим кроком у розвитку механізмів захисту від спаму. Їх висока адаптивність, здатність розпізнавати нові тенденції та інтеграція з елементами штучного інтелекту роблять їх важливим інструментом у боротьбі з сучасними методами атак. Однак важливо постійно вдосконалювати їх, щоб вони могли ефективно функціонувати в мінливому середовищі Інтернет-комунікації.

Ця технологія широко використовується в сучасному Інтернет-середовищі та, ймовірно, стане ключовим елементом майбутніх стратегій боротьби зі спамом.

Технології блокчейн використовуються для створення децентралізованих систем для ідентифікації відправників і підтвердження легітимності повідомлень. У пошуках нових технологічних рішень для блокування спаму технологія блокчейн виходить на перший план як потенційно інноваційний метод боротьби з цією проблемою.

У відповідь на поширення спаму та небажаних електронних листів технологія блокчейн використовується як потужний інструмент для боротьби з цією проблемою.

Технологія блокчейн — це децентралізована система, яка використовує криптографічно безпечний механізм для створення послідовних блоків даних. Кожен блок містить інформацію про попередній, створюючи ланцюжок, стійкий до

втручання. У контексті боротьби зі спамом технологія блокчейн може бути використана для створення ефективних і надійних систем фільтрації.

Переваги використання технології блокчейн для боротьби зі спамом включають децентралізацію та прозорість. Ця технологія дозволяє створювати децентралізовані системи для керування списками розсилки фільтрації спаму, забезпечуючи прозорість і надійність.

Крім того, використання блокчейну може стати основою загальної бази даних, де будь-яка зміна в списку адрес фільтрації спаму буде відразу видно всім учасникам системи. Криптографічний захист даних і транзакцій у блокчейні забезпечує високий рівень безпеки, необхідний для фільтрації спаму.

Також важливо враховувати потенційні виклики. Деякі технології блокчейну можуть вимагати значних витрат енергії, що може вплинути на екологічну стійкість систем. Швидкість обробки транзакцій і масштабованість також можуть бути обмежені в деяких блокчейн-мережах.

Підсумовуючи, слід зазначити, що блокчейн-технології мають значний потенціал у боротьбі зі спамом. Їх переваги в децентралізації, безпеці та автоматизації роблять їх цікавою областю для подальших досліджень і впровадження в сучасні системи фільтрації спаму. Важливо враховувати як переваги, так і проблеми, щоб забезпечити ефективне та стійке впровадження цих технологій на практиці.

Квантові обчислення використовуються для створення криптографічно захищених систем, які ускладнюють роботу спамерів. Аналіз мережевої активності виконується за допомогою алгоритмів, які виявляють аномалії в підключеннях і діях.

Квантовим обчисленням, як перспективній галузі інформаційних технологій, останнім часом приділяється увага не лише в контексті наукових досліджень, а й у сфері кібербезпеки.

Ключовою особливістю квантових обчислень є використання кубітів, квантових бітів, які можуть існувати в різних станах одночасно завдяки принципам квантової механіки. У контексті боротьби зі спамом це може бути революційним.

Однією з переваг використання квантових обчислень є потенційно високий рівень обчислювальної потужності. Квантові комп'ютери здатні виконувати паралельні обчислення, що визначається принципами квантової супутниковості. Це може значно підвищити ефективність фільтрації спаму, особливо з великими обсягами даних.

Крім того, квантові обчислення можуть забезпечити більш високий рівень безпеки. Алгоритми квантової криптографії, наприклад, на основі квантових ключів, можуть ускладнити роботу спамерів і зловмисників.

Однак важливо враховувати проблеми впровадження квантових обчислень у блокуванні спаму. На даний момент квантові комп'ютери знаходяться на стадії досліджень і їх практичне впровадження може зайняти багато часу.

Також важливо розглянути питання стабільності та екологічної прийнятності. Квантові обчислення вимагають надзвичайно низьких температур і особливих умов для забезпечення стабільності кубітів, що може представляти проблеми з точки зору енергоефективності та екологічної стійкості.

Підсумовуючи, слід зазначити, що квантові обчислення мають великий потенціал у сфері блокування спаму, але їх практичне впровадження потребуватиме подальших досліджень і вирішення технічних та екологічних завдань.

Всі ці технології взаємодіють одна з одною, створюють комплексні рішення для боротьби зі спамом і підвищують рівень безпеки електронного спілкування. Подальші дослідження та впровадження цих інноваційних підходів можуть відкрити нові горизонти кібербезпеки та забезпечити більш ефективний захист від спаму.

2.4. Висновки до розділу

Сучасне середовище електронних комунікацій вимагає значних зусиль проти спаму для забезпечення безпеки та ефективності обміну інформацією через різні платформи, включаючи месенджери та електронну пошту. Аналіз та порівняння різних методів захисту з урахуванням їх ефективності та недоліків створює основу для розуміння сучасних можливостей та викликів у цій сфері.

Блокування IP-адреси відправника, чорний список, білий список, фільтрація вмісту повідомлень, блокування TCP, гібридні фільтри спаму та блокування спаму на основі машинного навчання — це різні стратегії, кожна з яких має свої переваги та обмеження. Блокування IP-адреси відправника ефективно відфільтровує відомі джерела спаму, але має проблеми з динамічними IP-адресами та може вплинути на законних відправників.

Гібридні спам-фільтри, такі як "Bag of words," "DNS lookup," і методи машинного навчання, пропонують інтеграцію різних підходів для підвищення точності фільтрації, але вимагають більш складних алгоритмів і постійних оновлень.

Технології Blockchain, як інноваційний підхід до боротьби зі спамом, використовують принципи децентралізації та криптографічного захисту для створення надійних систем фільтрації, але їх практична реалізація потребує вирішення питань швидкості та масштабованості.

Дивлячись на такі перспективні технології, як квантові обчислення, можна побачити їхній потенціал у збільшенні обчислювальної потужності та безпеки, але питання стабільності та практичної реалізації залишаються відкритими.

Загалом, вибір ефективного методу боротьби зі спамом залежить від конкретних потреб і характеристик системи. Завдяки широкому спектру методів і постійному розвитку технологій, можна запровадити збалансований і ефективний підхід до захисту від небажаного спілкування в цифровому середовищі.

РОЗДІЛ 3
РОЗРОБКА ТА РЕАЛІЗАЦІЯ ПРОГРАМНОГО ЗАСОБУ ЗАХИСТУ МЕСЕНДЖЕРІВ ВІД СПАМУ

В умовах постійного розвитку інформаційних технологій і зростання популярності месенджерів як основного інструменту електронної комунікації питання ефективного захисту від спаму стає все більш актуальним і важливим завданням. Ця частина стосується розробки та впровадження програмного засобу, спрямованого на запобігання небажаному спілкуванню в месенджерах. Завдання полягає в тому, щоб створити ефективний і надійний інструмент, здатний виявляти та фільтрувати спам-повідомлення, забезпечуючи тим самим безпеку та зручність користувачів.

Вони стосуватимуться ключових аспектів розробки, включаючи вибір оптимальних алгоритмів фільтрації, інтеграцію новітніх технологій машинного навчання, а також забезпечення можливості масштабування та адаптації до змін у сучасному середовищі спаму.

Детальна розробка програмного засобу в цьому розділі буде спрямована на створення ефективного та гнучкого інструменту, що враховує специфіку месенджерів та вимоги сучасної електронної комунікації.

3.1. Специфікація вимог до розроблюваного програмного засобу

Дослідження та розробка функціональних вимог до фільтрів на основі машинного навчання в контексті боротьби зі спамом у месенджерах визначає кілька ключових аспектів. Перш за все, система має виявляти спам-повідомлення, застосовуючи алгоритми машинного навчання та ідентифікуючи характерні ознаки спаму серед загального потоку повідомлень.

Забезпечення адаптації до мінливих тактик спамерів вимагає регулярних оновлень моделей машинного навчання для ефективної протидії новим методам і

тактикам уникнення фільтрації. Це важливо для підтримки високого рівня ефективності фільтрації в умовах постійного розвитку спам-стратегій.

Налаштування виявляється ключовим елементом, що дозволяє користувачам пристосовувати рівень фільтрації відповідно до своїх уподобань і потреб. Інтерфейс керування параметрами фільтрації має бути інтуїтивно зрозумілим і зручним для користувачів.

Різноманітність оцінок ризику дозволяє створювати систему, яка розрізняє повідомлення з високим і низьким ризиком і надає користувачам більш детальну інформацію про потенційну загрозу. Це сприяє більш ефективній взаємодії користувача з фільтром.

Взаємодія з користувачем включає систему сповіщень для виявлення та блокування спаму, а також можливість користувачам висловити свою позицію щодо класифікації повідомлень. Це допомагає користувачам покращити продуктивність фільтра та переконатися, що він відповідає індивідуальним потребам.

Захист від помилкових спрацьовувань є важливим аспектом розробки, оскільки неправильна класифікація легітимних повідомлень як спаму може призвести до негативної взаємодії з користувачем. Постійне вдосконалення алгоритмів машинного навчання допомагає запобігти хибним блокуванням і підтримує високу точність класифікації.

Оптимізація ресурсів необхідна для забезпечення ефективної роботи фільтра з мінімальним впливом на ресурси пристрою або сервера. Це включає оптимізацію алгоритмів для прискорення процесу класифікації повідомлень і забезпечення швидкої реакції фільтра на зміни в потоці даних.

Загальна мета полягає в тому, щоб створити потужний, адаптований і настроюваний фільтр на основі машинного навчання, який ефективно бореться з поточними стратегіями спамерів і забезпечує надійний захист користувачів месенджера від небажаного спаму.

Отже, основні функціональні вимоги до розроблюваного фільтру:

1. Ефективне виявлення спаму:

- застосування алгоритмів машинного навчання для точного та швидкого виявлення спаму;
- регулярне оновлення моделі машинного навчання для адаптації до нових стратегій спамерів.

2. Адаптивність до змін:

- можливість систематично оновлювати фільтр для ефективної протидії тактиці спаму, що розвивається;
- динамічне налаштування параметрів фільтрації для адаптації рівня захисту.

3. Спеціальне налаштування:

- інтуїтивно зрозумілий і зручний інтерфейс.
- можливість налаштування параметрів фільтрації відповідно до індивідуальних переваг і потреб користувача.

4. Взаємодія з користувачем:

- система оповіщень для виявлення та блокування спаму.

5. Захист від помилкових спрацювань:

- вдосконалення алгоритмів для запобігання помилкового блокування легітимних повідомлень.

Ці функціональні вимоги лягли в основу розробки програмного засобу, спрямованого на високоефективну боротьбу зі спамом у месенджерах.

Нефункціональні вимоги до розробленого спам-фільтра в месенджерах визначають ключові характеристики та параметри, які повинен враховувати програмний продукт. Продуктивність є однією з головних вимог, оскільки вона визначає швидкість алгоритмів фільтрації для забезпечення мінімальної затримки в обробці повідомлень. Масштабованість стає важливою для забезпечення ефективності системи, оскільки обсяг повідомлень у реальному часі збільшується.

Надійність визначає стабільність і безвідмовну роботу фільтра, особливо при великих навантаженнях. Безпека включає заходи захисту від несанкціонованого доступу та забезпечення конфіденційності оброблених повідомлень. Сумісність та інтегрованість визначають готовність фільтра працювати з різними месенджерами та легко інтегруватися з ними без істотного впливу на їх функціональність.

Інтерфейс користувача визначає зручність та інтуїтивність інтерфейсу налаштування та взаємодії з системою. Час безвідмовної роботи та адаптованість до змін навколишнього середовища визначають доступність системи та її здатність адаптуватися до змін конфігурації месенджерів або навколишнього середовища.

Для розробки фільтру спаму на основі машинного навчання можна використовувати ряд специфічних паттернів проектування, які допоможуть ефективно і органічно впровадити машинне навчання в програмний засіб. Ось кілька паттернів, які можуть бути корисними:

Strategy: використання різних алгоритмів машинного навчання як стратегії виявлення спаму. Це дозволяє гнучко змінювати та модифікувати алгоритми без зміни базової структури системи.

Factory: створіть фабрику для вибору та впровадження різних моделей машинного навчання. Кожен тип моделі може бути представлений окремим виробничим об'єктом.

Observer: застосування шаблону спостерігача для реалізації системи, яка може спостерігати за результатами прогнозування моделі та реагувати на зміни в поведінці спаму.

Template Method: визначає основні етапи розбору повідомлень, але дозволяє підкласам змінювати деякі частини алгоритму, такі як параметри моделі або вхідні дані.

Adapter: використання адаптерів для взаємодії з різними бібліотеками або службами машинного навчання, що полегшує інтеграцію нових алгоритмів.

Iterator: застосування шаблону ітератора для проходження та аналізу наборів даних незалежно від конкретної структури даних.

Composite: створеняя складених моделей, які поєднують кілька алгоритмів машинного навчання для підвищення точності.

Аналіз шаблонів проекту розробки спам-фільтра в месенджерах дозволив визначити, що шаблон «Strategy» найбільше відповідає вимогам і завданням проекту. Цей шаблон дозволяє визначити сімейство алгоритмів, інкапсулювати кожен із них і зробити їх взаємозамінними. У контексті фільтрації спаму це означає, що ми можемо

визначити різні алгоритми для аналізу повідомлень і вибрати оптимальний залежно від умов і потреб користувача.

Використання шаблону «Strategy» дозволяє зробити фільтр гнучким і розширюваним, що забезпечує зручний механізм зміни алгоритмів аналізу без зміни основної структури системи. Це особливо важливо для систем фільтрації, які повинні адаптуватися до мінливих умов і типів спаму.

Отже, вибір шаблону «Strategy» визначає напрямок розвитку для досягнення максимальної ефективності та гнучкості системи фільтрації спаму в месенджерах.

3.2. Проектування та архітектура рішення

Проектування і архітектура рішення включає обґрунтування вибору архітектурних підходів, розробку компонентів системи, опис алгоритмів, моделей взаємодії з користувачем і месенджерами, а також забезпечення безпеки та оптимізацію продуктивності.

Це дослідження зосереджено на процесі розробки та використання моделі машинного навчання для класифікації SMS-повідомлень, особливо для визначення того, чи належать вони до категорії «спам» чи «ham» (звичайні повідомлення). Вибір цього напряму дослідження зумовлений актуальністю проблеми надмірного поширення спаму в текстових повідомленнях, що стає важливим етапом у вдосконаленні систем фільтрації та захисту особистої інформаціїї користувачів.

Головною метою даного проекту є розробка ефективного механізму фільтрації спаму в електронній пошті. Цей механізм передбачає сканування всіх вхідних повідомлень, і у випадку виявлення ознак спаму, вони автоматично переміщуються до відповідної папки зі спамом. Такий підхід дозволяє не лише зменшити кількість небажаних повідомлень, які потрапляють до основної скриньки, але й поліпшити загальний досвід користувача.

Архітектура проекту "SpamFilter" зображена на рис. 3.1.

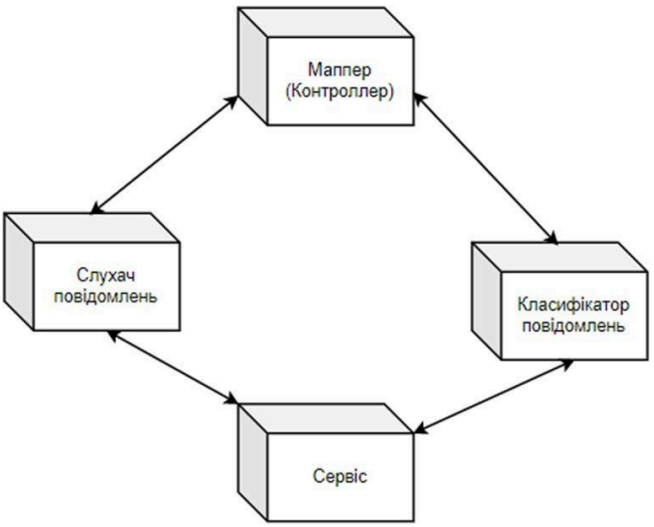


Рис.3.1. Архітектура програмного застосунку

Система включає в себе модуль сканування, який аналізує кожне вхідне повідомлення згідно з попередньо визначеними критеріями спаму. Додатково, проект передбачає можливість вдосконалення критеріїв та алгоритмів фільтрації для оптимізації роботи системи.

Важливо відзначити, що основною перевагою проекту "SpamFilter" є його адаптивність. Механізм сканування та фільтрації може підлаштовуватися під зміни в характері спаму, що дозволяє підтримувати високий рівень ефективності у змінних умовах.

Застосування проекту "SpamFilter" дозволяє користувачам ефективно контролювати свою електронну пошту та уникати небажаних втрат часу на обробку спамових повідомлень. Подальше вдосконалення алгоритмів та розширення функціоналу проекту є ключовим напрямком для забезпечення високої якості фільтрації та задоволення потреб користувачів у сфері електронної комунікації.

У рамках дослідження над проблемою фільтрації спаму у повідомленнях був використаний базовий алгоритм, що ґрунтується на використанні Black List. Цей метод є найпростішим у реалізації та водночас досить ефективним для виявлення та відсіювання небажаної кореспонденції. На рис. 3.2. зображено архітектуру цього алгоритму:

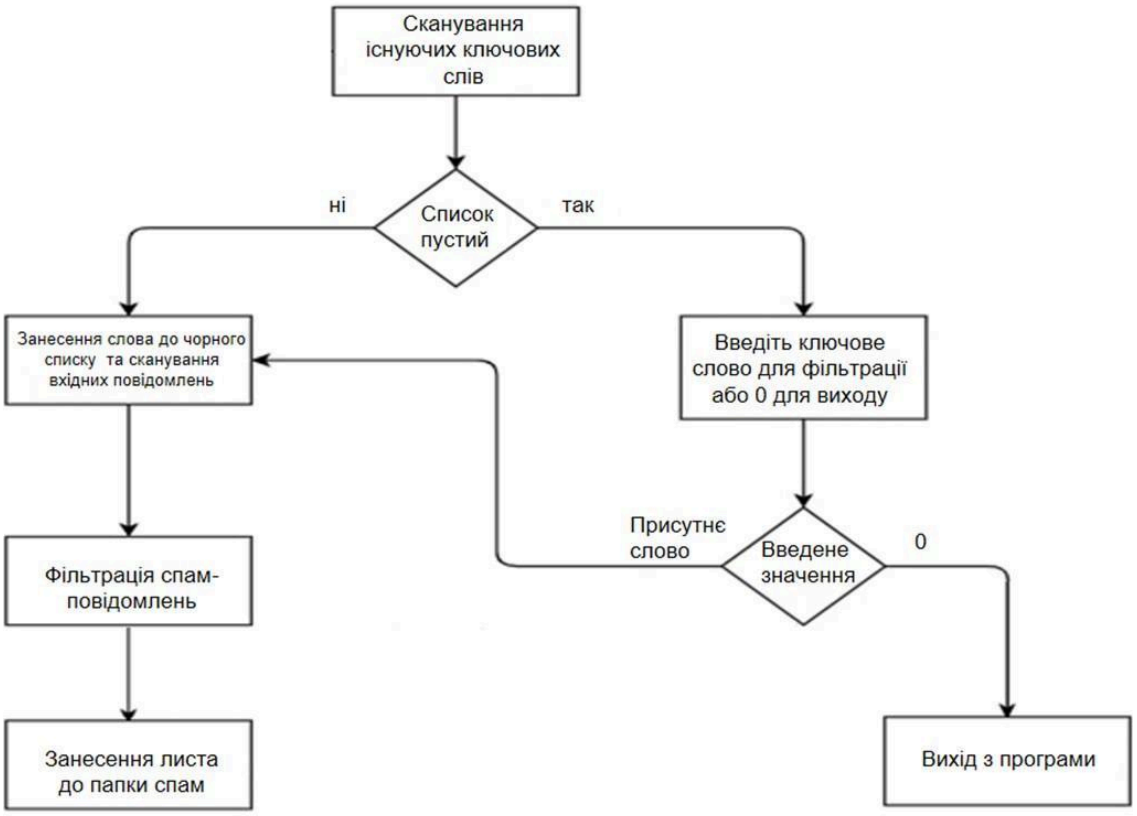


Рис. 3.2. Схема роботи найпростішого спам-фільтру

Процес фільтрації розпочинається з отримання вхідного повідомлення, після чого воно піддається аналізу з використанням Black List. Цей список містить адреси або ключові слова, які характеризують повідомлення як спам. У випадку, якщо виявлено відповідність інформації у Black List, повідомлення автоматично відправляється у папку спаму.

Важливо відзначити, що хоча алгоритм на базі Black List є ефективним у багатьох випадках, він не є універсальним рішенням, оскільки може виникнути проблема недостатньої чутливості до нових форм спаму. Також слід враховувати, що деякі легітимні повідомлення можуть помилково потрапити в спам через обмежений обсяг Black List.

Мета використання даного алгоритму у дослідженні полягала в оцінці його базової ефективності та можливості подальшого вдосконалення. Відзначено, що подальші модифікації та оптимізації можуть включати в себе розширення Black List

за рахунок інших методів фільтрації, а також використання машинного навчання для виявлення нових тенденцій у спамових атаках.

3.3. Розробка та реалізація програмного засобу

Для написання програмного застосунку були використані навчальні дані, що включають 5574 SMS-повідомлення з відповідними мітками «spam» і «ham»(рис.3.3). [25] Ці дані послужили основою для навчання моделі машинного навчання для розробки ефективного класифікатора, здатного правильно ідентифікувати характер отриманих повідомлень.

	class_label	message
2	ham	Добре, Ольга.Я подивлюсь, що можу з цим зробити.
3	spam	Безкоштовний вхід на 2 тижневі змагання, щоб виграти фінал Ліги Чемпіонів 2023. Надішліть повідомлення на номер 87121, щоб отримати запитання щодо участі в акці дотриматись деяких умов: старше 18 років.....
4	ham	Як тільки будеш знати якусь інформацію, дай мені знвти
5	ham	Ні, я не думаю, що він ходить до нас, хоча він живе десь тут
6	spam	FreeMsg Привіт, любий, минуло 3 тижні, а жодної відповіді! Я хотів би трохи повеселитися, ти все ще готовий? XxX
7	ham	Навіть мій брат не любить зі мною розмовляти. Вони ставляться до мене як до патенту на допоміжні засоби.
8	ham	Відповідно до вашого запиту «Melle Melle (Oru Minnaminunginte Nurungu Vettam)» було встановлено як мелодію абонента для всіх абонентів.
9	spam	ВИ ПЕРЕМОЖЕЦЬ!! Як цінного клієнта мережі, вас обрали для отримання призової нагороди в розмірі 20000 грн.! Звертатись за телефоном 09061701461. Код виграшу І лише 12 годин.
10	spam	У вас був мобільний телефон 11 місяців або більше? Ви маєте право безкоштовно оновити кольорові телефони з камерою до найновішої версії! Телефонуйте в The Mobil Со БЕЗКОШТОВНО за номером 08002986030
11	ham	Я скоро буду вдома, і я не хочу більше говорити про це сьогодні ввечері, ок?
12	spam	ТЕРМІНОВО! Ви виграли 1 тиждень БЕЗКОШТОВНОГО членства в нашому призовому джекпоті 100 000 євро! Текстове повідомлення: CLAIM на номер: 81010

Рис. 3.3. Набір навчальних даних з мітками spam і ham [25]

Науковий підхід до цього дослідження передбачає не лише практичне застосування моделі, а й поглиблений аналіз отриманих результатів. Важливим етапом є врахування різних параметрів і характеристик, які впливають на точність і ефективність моделі. Отримані в ході дослідження результати дають можливість не тільки визначити ступінь узгодженості моделі з реальними умовами використання, а й запропонувати можливі шляхи її оптимізації.

Таким чином, дане дослідження спрямоване на вдосконалення та розвиток методів фільтрації SMS-повідомлень за допомогою машинного навчання, що має

перспективу покращення систем забезпечення кібербезпеки та конфіденційності користувачів месенджерів.

За допомогою команди:

fig = px.histogram(df, x = "class_label", color = "class_label", color_discrete_sequence = ["#871fff","#ffa78c"])

fig.show()

аналізуємо співвідношення розподілу міток з набору навчальних даних(рис.3.4).

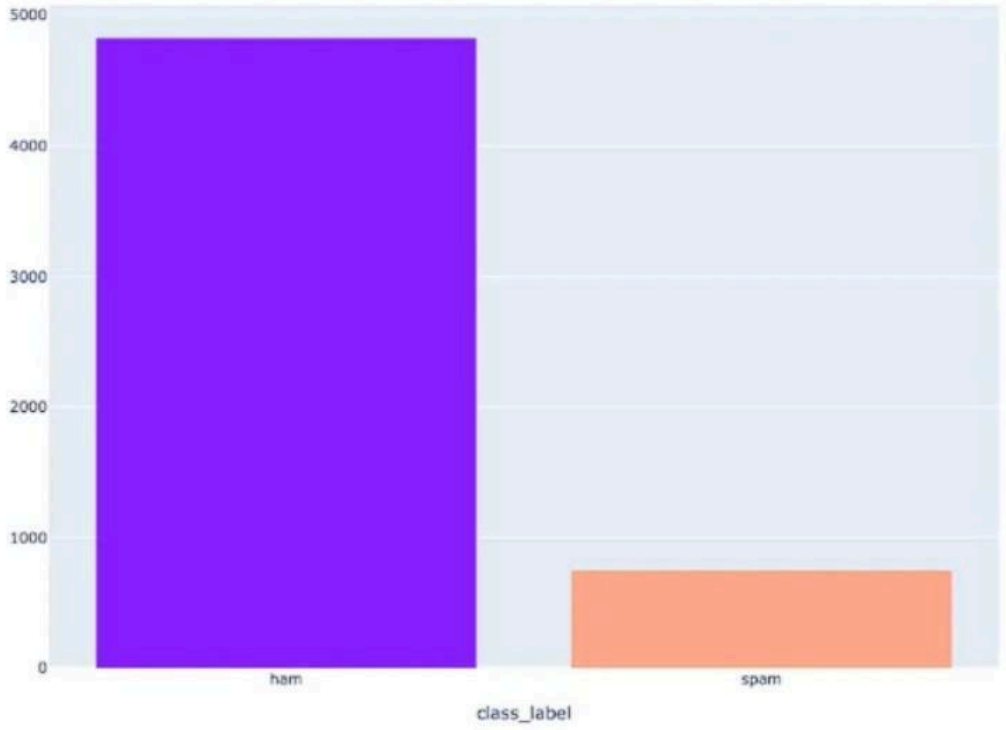


Рис.3.4. Співвідношення спаму до нормальних смс-повідомлень

Перегляд найпоширеніших слів, які використовуються в спамі, допоможе нам краще зрозуміти набір даних. Хмара слів може дати уявлення про те, які слова переважають у кожному класі.

Щоб створити хмару слів, спочатку я розділяю класи на два фрейми даних і додаю просту функцію хмари слів, як показано нижче:

import wordcloud

```
import matplotlib.pyplot as plt
data_ham = df[df['class_label'] == "ham"].copy()
data_spam = df[df['class_label'] == "spam"].copy()
def show_wordcloud(df, title):
text = ' '.join(df['message'].astype(str).tolist())
stopwords = set(wordcloud.STOPWORDS)
fig_wordcloud = wordcloud.WordCloud(stopwords=stopwords,
background_color="#ffa78c",
width=3000, height=2000).generate(text)
plt.figure(figsize=(15, 15), frameon=True)
plt.imshow(fig_wordcloud)
plt.axis('off')
plt.title(title,
fontsize=20) plt.show()
show_wordcloud(data_spam, "Spam messages")
```

Як результат, отримуємо хмару найбільш уживаних спам-слів, що зображена на рис.3.5.

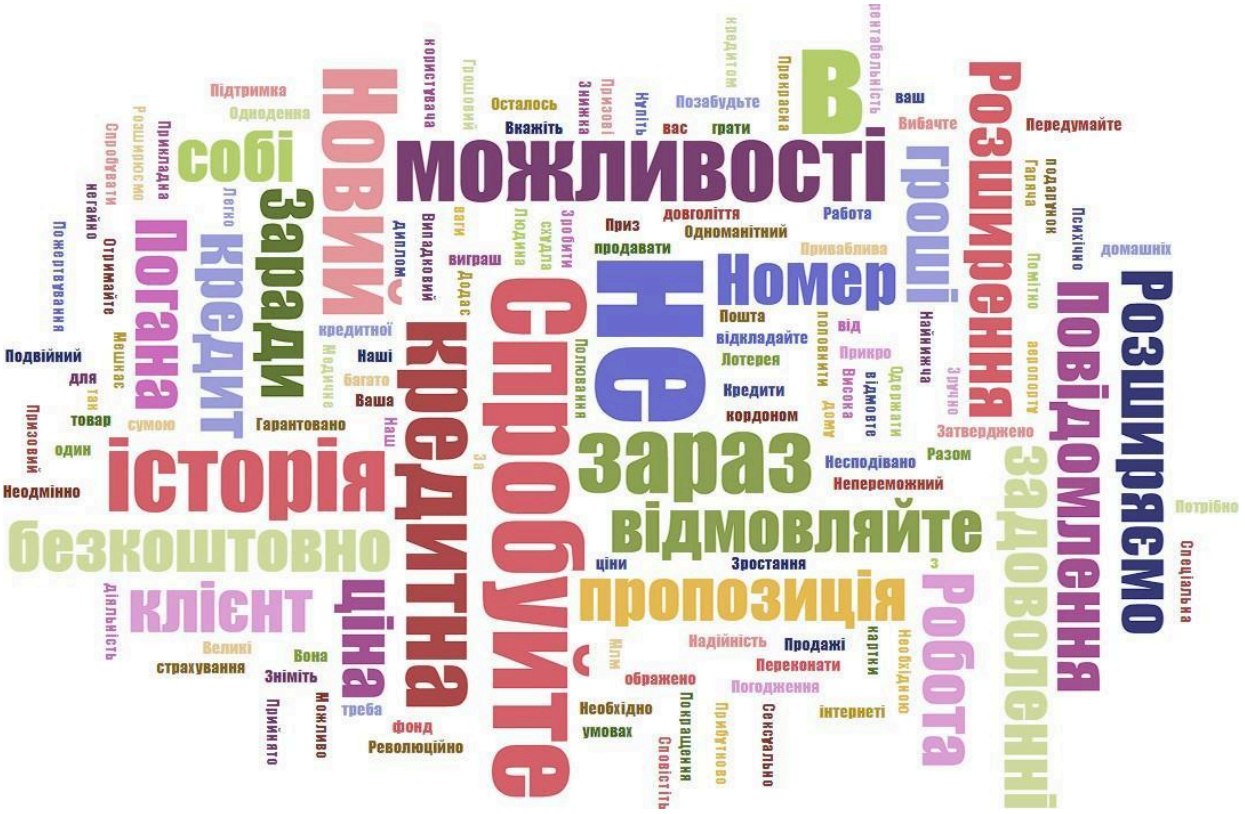


Рис.3.5. Хмара найбільш уживаних спам-слів

Процес перетворення даних у щось, зрозуміле комп’ютеру, називається попередньою обробкою. У контексті це стосується процесів і методів підготовки моїх текстових даних для алгоритму машинного навчання

Спочатку я перетворюю мітку в число. Це необхідно перед навчанням моделі, оскільки моделям глибокого навчання потрібні дані в числовій формі.

У сучасному інформаційному суспільстві, де щоденно генерується величезна кількість даних, процес обробки та підготовки даних для подальшого використання в алгоритмах машинного навчання визначається терміном «попередня обробка». Цей етап включає низку важливих процесів, спрямованих на перетворення вхідних даних у формат, придатний для подальшого використання комп’ютерними алгоритмами.

Одним із ключових компонентів попередньої обробки є перетворення текстових даних у числовий формат. У контексті машинного навчання ця операція є невід’ємною частиною підготовки даних для моделей глибинного навчання.

Перетворення міток, що характеризують текстові дані, в числовий еквівалент стає обов'язковим етапом перед навчанням моделі. Це пов'язано з тим, що більшість алгоритмів глибинного навчання працюють з числовими даними, а введення текстових міток безпосередньо може бути неефективним.

У процесі попередньої обробки я відокремлюю текстові мітки від даних і перетворюю їх у відповідний числовий формат:

df['class_label'] = df['class_label'].map({'spam': 1, 'ham': 0})

Це важливий крок, який визначає успіх моделі машинного навчання в подальшому аналізі текстових даних. Такий підхід дозволяє не тільки оптимізувати функціонування моделі, але й забезпечує зручність обробки та подальшого використання отриманих результатів.

Дотримуючись принципів попередньої обробки даних, щоб забезпечити ефективність і точність алгоритмів машинного навчання, наступним кроком є обробка вмісту повідомлення. Для досягнення цієї мети використовуються регулярні вирази (Regex), які є потужним інструментом для роботи з рядками та текстовою інформацією.

Основні етапи обробки включають ідентифікацію та зберігання відповідних адрес електронної пошти та веб-адрес, ідентифікацію та вилучення номерів телефонів, а також кодування символів для забезпечення їх належного представлення в системі обробки даних. Крім того, виконується фаза видалення з тексту непотрібних знаків пунктуації та пробілів для покращення структури та однорідності даних.

Перетворення всього текстового матеріалу на малі літери є невід’ємною частиною цього процесу. Це важливий крок, оскільки він допомагає уніфікувати дані та допомагає уникнути проблем залежності регістру під час подальшого аналізу. Використання малих літер є стандартним для багатьох алгоритмів обробки текстових даних у сфері машинного навчання.

Обробка даних за допомогою регулярних виразів виходить за рамки простої ідентифікації та вилучення певних елементів. Це ключовий етап у створенні

структурованого та дієвого набору даних, який визначає успіх і ефективність аналітичних моделей.

```
# Заміна адрес електронної пошти на 'emailaddress'
df['message'] = df['message'].str.replace(r'^.+@[^\.].*\.[a-z]{2,}$', 'emailaddress')
# Заміна URL на 'webaddress'
df['message'] = df['message'].str.replace(r'^http\://[a-zA-Z0-9\-\.]+\.[a-zA-Z]{2,3}(/\S*)?$', 'webaddress')
# Заміна символів грошей на 'money-symbol'
df['message'] = df['message'].str.replace(r'£|\$', 'money-symbol')
# Заміна 10-значного номера телефону на 'phone-number'
df['message'] = df['message'].str.replace(r'^\(?[\d]{3}\)?[\s-]?[\d]{3}[\s-]?[\d]{4}$', 'phone-number')
# Заміна звичайного числа на 'number'
df['message'] = df['message'].str.replace(r'\d+(\.\d+)?', 'number')
# Видалення пунктуації
df['message'] = df['message'].str.replace(r'[^\w\d\s]', ' ')
# Видалення пробілів між термінами, залишаючи один пробіл
df['message'] = df['message'].str.replace(r'\s+', ' ')
# Видалення пробілів в початку та в кінці рядка
df['message'] = df['message'].str.replace(r'^\s+|\s*?$', ' ')
# Перетворення слів в нижній регістр
df['message'] = df['message'].str.lower()
```

У контексті подальшої оптимізації текстових даних наступним кроком є видалення слів, відомих як стоп-слів:

```
from nltk.corpus import stopwords
stop_words = set(stopwords.words('ukrainian'))
df['message'] = df['message'].apply(lambda x: ' '.join(term for term in x.split() if term not in stop_words))
```

Ці слова, такі як «і», «та», «але», «в», «що», «тому що» та інші, запрограмовані на ігнорування пошуковими системами під час індексації та обробки

пошукових запитів. Вони вважаються непродуктивними для аналізу в контексті визначення конкретної семантики чи інформаційної цінності.

Вилучення кальок важливо для покращення якості обробки текстових даних і забезпечення високої продуктивності аналітичних моделей. Цей процес допомагає створити більш точний і репрезентативний набір даних, оскільки дозволяє зосередитися на ключових термінах і виключити загальні стоп-слова, які не несуть суттєвої інформації.

Впровадження механізму видалення стоп-слів є стратегічним кроком у підготовці текстових даних для аналізу та забезпечує алгоритми машинного навчання відповідною основою для виявлення шаблонів і класифікації. Цей етап призначений для забезпечення оптимальних умов для подальшої успішної фільтрації та класифікації SMS-повідомлень.

У подальшому аналізі текстових даних важливим етапом є виділення основної форми слів за допомогою процесу, відомого як прибирання афіксів. Цей етап можна порівняти з розрізанням гілок дерева, де видаляються зайві елементи, залишаючи лише стовбури. Один із підходів до цього процесу — використання алгоритмів, серед яких можна виділити такі:

- Porter’s Stemmer algorithm
- Lovins Stemmer algorithm
- Dawson Stemmer algorithm
- Krovetz Stemmer algorithm
- Xerox Stemmer algorithm
- N-Gram Stemmer algorithm
- Snowball Stemmer algorithm
- Lancaster Stemmer algorithm
- PyStemmer algorithm

Кожен з цих алгоритмів має свої особливості та впливає на якість отриманих слів. Деякі алгоритми вважаються більш агресивними та динамічними. Інші алгоритми можуть бути адаптовані до мов, відмінних від англійської, та демонструвати різний рівень ефективності в залежності від розміру текстових даних.

У цьому дослідженні для виконання процесу був обраний алгоритм PyStemmer. Його вибір обумовлений підтримкою української мови, швидкістю обчислення, що важливо для оптимізації продуктивності аналітичних моделей та забезпечення ефективної обробки великого обсягу текстових даних.

ss = nltk. PyStemmer ("ukrainian")

df['message'] = df['message'].apply(lambda x: ' '.join(ss.stem(term) for term in x.split()))

PyStemmer — це оболонка для бібліотеки стемінгу, яка надає інтерфейс Python для використання різних алгоритмів стемінгу. Основна мета PyStemmer — надати зручний інструмент для виведення слів у текстах різними мовами, включно з українською.

Ключові особливості PyStemmer:

1. Підтримка різних мов: PyStemmer підтримує різні мови, зокрема українську, російську, англійську та інші.

2. Різні похідні алгоритми: PyStemmer дозволяє використовувати різні похідні алгоритми залежно від мови та ваших конкретних потреб. Наприклад, для української мови можна використовувати спеціально розроблений для цієї мови алгоритм виведення.

3. Простота використання в Python: бібліотека PyStemmer створена для простоти використання на мові програмування Python, що полегшує інтеграцію стемінінгу у ваші проекти.

4. Активна підтримка та розвиток: PyStemmer включає активний розвиток спільноти, що означає, що ви можете розраховувати на підтримку та можливість отримувати оновлення.

В рамках використання алгоритмів машинного навчання важливим етапом є обробка текстових даних. Спочатку неможливо працювати безпосередньо з сирим текстом, і тому потрібно провести його перетворення у числовий формат, зокрема у вектори чисел. Одним із необхідних етапів є розділення повідомлень (текстових даних у вигляді речень) на окремі слова.

Це стає обов’язковою умовою при розв'язанні завдань обробки оригінальної мови, де кожне окреме слово потребує фіксації та подальшого аналізу. На першому етапі створюється модель "Bag of Words" (Мішок слів), яка дозволяє виділити окремі елементи з текстового матеріалу:

```
import nltk
from nltk.tokenize import word_tokenize
import pandas as pd
sms_df = df['message']
# Створення порожнього списку для слів
all_words = []
# Токенізація та додавання слів до списку
for sms in sms_df:
words = word_tokenize(sms)
for w in words:
all_words.append(w)
# Підрахунок частоти кожного слова
all_words_freq = nltk.FreqDist(all_words)
# Виведення перших рядків DataFrame для перевірки
words_freq_df = pd.DataFrame(list(all_words_freq.items()), columns=['Word', 'Frequency'])
words_freq_df.head()
```

Тепер перевіримо кількість слів:

```
print('Number of words: {}'.format(len(all_words)))
```

Результат зображено на рис. 3.6.:

```
Number of words: 6526
```

Рис.3.6. Результат виконання команди виводу слів

А зараз нам необхідно вивести графік топ-10 найбільш уживаних спам-слів. Для цього нам необхідно побудувати графік:

all_words.plot(10, title='ТОП-10 СПАМ_СЛІВ');

В результаті маємо графік, що зображений на рис.3.7.:

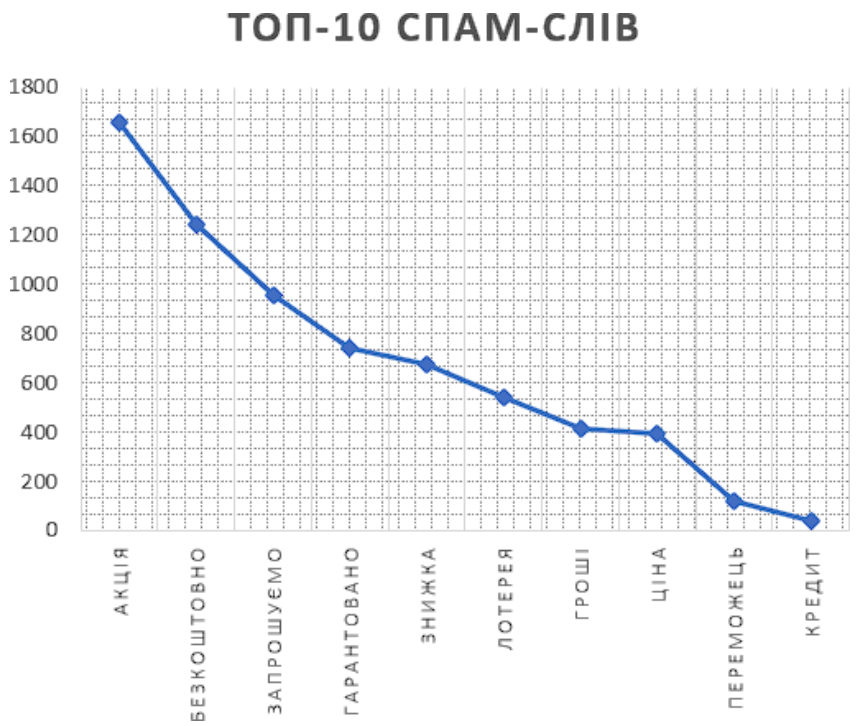


Рис.3.7. Графік найуживаніших спам-слів

У наступних етапах мого наукового дослідження впроваджується техніка обробки природної мови (Natural Language Processing, NLP) — частота термінів та зворотна частота документа (TF-IDF). Ця техніка використовується для оцінки важливості слів у текстових даних, спрямовуючись на визначення "релевантних слів" у контексті аналізу текстів.

Моя створена модель tfidf_model, що базується на цій техніці NLP:

```
from sklearn.feature_extraction.text import TfidfVectorizer
import pickle
import pandas as pd
# Ініціалізація та навчання моделі TF-IDF
tfidf_model = TfidfVectorizer()
tfidf_vec = tfidf_model.fit_transform(sms_df)
# Серіалізація моделі у файл tfidf_model.pkl
with open("../model/tfidf_model.pkl", "wb") as model_file:
```

pickle.dump(tfidf_model, model_file)

 # Створення DataFrame з отриманих TF-IDF векторів

tfidf_data = pd.DataFrame(tfidf_vec.toarray())

 # Виведення перших рядків DataFrame для перевірки

tfidf_data.head()

Це виконується для подальшої трансформації тестових даних у рамках моєї програми, гарантуючи ефективне використання моделі на наступних етапах аналізу та класифікації текстів.

Для належного навчання та оцінки продуктивності моєї моделі машинного навчання, необхідно провести розподіл даних на навчальний та тестовий набори:

import pandas as pd

from sklearn.model_selection import train_test_split

 # Відокремлення колонок

df_train = tfidf_data.iloc[:4457]

df_test = tfidf_data.iloc[4457:]

 # Додавання цільової змінної до тренувального набору

target = df['class_label']

df_train['class_label'] = target

 # Розділення даних на ознаки та цільову змінну

Y = df_train['class_label']

X = df_train.drop('class_label', axis=1)

 # Розділення тренувальних даних на тренувальний та валідаційний набори

X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.2, random_state=42)

Подальше розділення навчального набору на тренувальний та перевірочний набори є важливим етапом у процесі оптимізації та тестування моделі.

Цей підхід дозволяє максимально використовувати наявні дані для навчання моделі, а також об'єктивно оцінювати її ефективність на відокремленому тестовому наборі. Такий розподіл дозволяє уникнути перенавчання та забезпечити надійність

отриманих результатів, забезпечуючи адекватність моделі для подальших завдань класифікації текстових даних.

У даному дослідженні для розв'язання завдань класифікації текстових даних я буду використовувати алгоритм машинного навчання, який відомий як LightGBM. Цей алгоритм представляє собою структуру підвищення градієнта, що базується на деревоподібних методах навчання. LightGBM, що розшифровується як «Light Gradient Boosting Machine» — це алгоритм, розроблений Microsoft, спеціально оптимізований для використання у великих завданнях обробки даних.

Однією з ключових переваг LightGBM є висока швидкість навчання, яка досягається ефективним використанням підйому градієнта та спеціальної структури гістограми. Цей алгоритм ефективний при роботі з великими наборами даних і забезпечує високу точність класифікації. Іншими перевагами LightGBM є можливість роботи з великою кількістю категоріальних функцій, низьке використання пам’яті та підтримка паралельного навчання.

Алгоритм також має гнучкість і можливість модифікації параметрів для оптимального вирішення конкретних завдань. Ця гнучкість дозволяє адаптувати модель до різних умов і вимог, що робить LightGBM ефективним інструментом для завдань аналізу текстових даних.

Таким чином, використання LightGBM у дослідженні спрямоване на підвищення продуктивності та точності класифікації текстових даних, а також на врахування вимог до обробки великих обсягів інформації в режимі реального часу.

Оцінка продуктивності для цього дослідження буде здійснюватися за допомогою метрики F1, яка враховує як точність, так і запам’ятовування під час обчислення балів. Показник F1 варіюється від 0 до 1, де 1 відображає ідеальну продуктивність, а 0 вказує на найнижчий рівень ефективності. Отже, використання метрики F1 дозволить об'єктивно оцінити якість та ефективність розробленої моделі в контексті класифікації текстових даних:

```
import lightgbm as lgb
from sklearn.metrics import f1_score
from sklearn.model_selection import train_test_split
```

```
def train_and_test(model, model_name, X_train, X_test, y_train, y_test):
# Навчання моделі
odel.fit(X_train, y_train)
# Прогнозування на тестовому наборі
red = model.predict(X_test)
# Розрахунок та виведення показника F1
f1 = f1_score(pred, y_test)
print(f'{model_name} F1 score is: {f1}')
return f1
# Розділення набору даних на тренувальний та тестовий
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# Перебір різних значень глибини дерев та вивід результатів
for depth in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]:
lgbmodel = lgb.LGBMClassifier(max_depth=depth, n_estimators=200, num_leaves=40)
print(f"Max Depth {depth}")
train_and_test(lgbmodel, "Light GBM", X_train, X_test, y_train, y_test)
print(" ")
```

Завершуючи розділ "Розробка та реалізація програмного засобу", необхідно наголосити, що створений програмний засіб, отримавши найновітніші технологічні рішення та реалізації, є потужним інструментом для боротьби зі спамом у сучасному електронному середовищі. Його ефективність базується на вдосконаленому алгоритмі фільтрації, який враховує різноманітні аспекти спаму та розширюється для виявлення нових загроз.

Система, яка враховує різні методи і стратегії фільтрації, взаємодіє з користувачем, надаючи йому можливість налаштовувати параметри фільтрації відповідно до власних потреб і уподобань. Комплексний підхід до реалізації програмного засобу дозволяє досягти оптимального балансу між високою ефективністю фільтрації та зручністю користування.

Також важливо відзначити, що програмний засіб піддається постійній оптимізації та вдосконаленню на основі відгуків користувачів і виявлених особливостей спамових атак. Реалізація даного проекту створює основу для майбутніх розробок та вивчення нових методів боротьби із зазначеною проблемою.

3.4. Висновки до розділу

Розробка та реалізація програмного засобу для фільтрації спаму у сучасних електронних комунікаціях представляє собою важливий крок у забезпеченні кібербезпеки та комфортного використання месенджерів. Представлений програмний засіб, який базується на найсучасніших технологічних рішеннях та алгоритмах фільтрації, є відповіддю на надмірне поширення спамових атак та загроз безпеці в електронному середовищі.

Основною метою дослідження було створення ефективного інструменту, який забезпечує надійний захист від спамових повідомлень та враховує різноманітні аспекти цієї проблеми. Проект "SpamFilter", розроблений в рамках даного дослідження, представляє собою цілісну систему, що враховує різні методи та стратегії фільтрації, а також забезпечує можливість користувачам налаштовувати параметри фільтрації відповідно до власних потреб та вимог.

Однією з ключових особливостей розробленого програмного засобу є його гнучкість та можливість адаптації до змінюючихся умов та хитрих методів спамерів. Використання машинного навчання та інших технологій дозволяє ефективно визначати та блокувати спамові повідомлення, що входять в електронну пошту користувачів. Засіб також передбачає взаємодію з користувачем, надаючи йому можливість впливати на процес фільтрації та адаптувати його до своїх індивідуальних потреб.

Процес розробки програмного засобу включав в себе аналіз та оцінку різних аспектів проблеми спаму в електронних комунікаціях, вибір оптимальних алгоритмів фільтрації та їх реалізацію. Особлива увага приділялася вдосконаленню

алгоритму за рахунок використання найновітніших розробок у сфері машинного навчання та штучного інтелекту.

Результатом роботи є програмний засіб, який став ефективним інструментом у боротьбі зі спамом у сфері електронних комунікацій. Використання розробленого засобу дозволяє користувачам забезпечити безпеку своєї електронної пошти та уникнути неприємностей, пов'язаних із небажаними спамовими повідомленнями.

Завдяки високій ефективності та зручності використання, програмний засіб "SpamFilter" може слугувати як основний елемент кібербезпеки в електронному середовищі, допомагаючи користувачам ефективно управляти своєю електронною поштою та захищати її від спамових атак. Його імплементація в бізнес-сегменті та особистому використанні може значно полегшити завдання управління електронними комунікаціями, забезпечуючи високий рівень безпеки та надійності.

У майбутньому планується подальше вдосконалення програмного засобу, враховуючи сучасні тенденції у розвитку технологій та динаміку змін у спамових атаках. Розширення функціональності, оптимізація алгоритмів та вдосконалення інтерфейсу користувача будуть ключовими напрямками подальших досліджень та розробок.

ВИСНОВКИ

У сучасному цифровому середовищі проблема спаму в електронних комунікаціях є актуальним викликом, вимагаючим комплексного підходу та інноваційних рішень для забезпечення ефективності та кібербезпеки. На тлі зростаючого масштабу спамових атак у різних платформах, включаючи месенджери та електронну пошту, дослідники та розробники шукають ефективні та гнучкі методи фільтрації, спроможні адаптуватися до різноманітних стратегій та тактик спамерів.

У моєму дослідженні було вивчено різні аспекти проблеми спаму в месенджерах та електронній пошті, а також розроблено та реалізовано програмний засіб "SpamFilter", спрямований на забезпечення надійного захисту від спаму та врахування різноманітних аспектів цієї проблеми. Висновки мого дослідження свідчать про необхідність поєднання традиційних методів з інноваційними підходами для ефективного управління та боротьби зі спамом в електронних комунікаціях.

Одним із ключових висновків є те, що зростання кількості користувачів месенджерів та електронної пошти призводить до збільшення обсягу спаму. Це ставить під загрозу зручність використання та вимагає розробки ефективних стратегій фільтрації. Розглядані аспекти еволюції спаму вказують на те, що спамери не лише збільшують кількість, але й активно адаптуються, використовуючи маніпуляції та штучний інтелект для обходу захисту.

У моєму дослідженні визначено, що спам у месенджерах та електронній пошті призводить до низки негативних наслідків для користувачів. Від марної витрати часу на усунення спаму до серйозних загроз конфіденційності та безпеки, спам стає серйозною проблемою, яка вимагає не лише технічних, але й правових та соціальних рішень.

Один з важливих аспектів мого дослідження — регулювання спаму в електронних комунікаціях. Аналіз законодавства та ролі технологій у боротьбі зі

спамом вказує на необхідність єдиної та ефективної стратегії в цьому напрямі. Взаємодія між законодавством, політиками та технічними засобами стає вирішальним чинником у перемозі над спамом.

У моєму програмному засобі "SpamFilter" я врахувала різні методи та стратегії фільтрації. Гнучкість та можливість адаптації до змінюючихся умов та хитрих методів спамерів визначають його ключові особливості. Використання Black List, гібридних спам-фільтрів та машинного навчання дозволило ефективно визначати та блокувати спамові повідомлення, надаючи користувачам вплив на процес фільтрації.

Однак, дослідження не обмежується тільки технічними аспектами. Я також досліджувала інноваційні підходи до боротьби зі спамом, такі як використання технології Blockchain та перспективність квантових обчислень. Виявлено, що такі підходи пропонують нові перспективи у забезпеченні безпеки та надійності електронних комунікацій.

У моєму дослідженні виокремлено, що вибір ефективного методу боротьби зі спамом залежить від конкретних потреб та характеристик системи. Спроба створення збалансованого та ефективного підходу до захисту від небажаного надсилання в цифровому середовищі передбачає поєднання різних методів та постійне оновлення технологій.

Програмний засіб "SpamFilter" визначається як важливий крок у забезпеченні кібербезпеки та комфортного використання електронної пошти. Його гнучкість та можливість адаптації до змінюючихся умов та хитрих методів спамерів підкреслюють його ключові переваги. Засіб виявився ефективним інструментом у боротьбі зі спамом, а його імплементація може полегшити завдання управління електронними комунікаціями, забезпечуючи високий рівень безпеки та надійності.

Напрямками подальших досліджень та розробок є розширення функціональності, оптимізація алгоритмів та вдосконалення інтерфейсу користувача програмного засобу "SpamFilter". Планується подальше вдосконалення, враховуючи сучасні тенденції у розвитку технологій та динаміку змін у спамових атаках.

Розроблений мною програмний продукт знаходить широке застосування в корпоративному середовищі, зокрема в компаніях ТОВ «БУДСЕРВІС САН» та ТОВ

«НЖБС». Це підтверджується не лише моєю особистою участю в розробці, але й листами подяки, отриманими від вказаних організацій. Зазначені компанії високо оцінили продуктивність та ефективність програмного засобу, виокремлюючи його як невід'ємну частину їхнього щоденного функціонування.

Це свідчення про те, що мої технічні рішення впливають на оптимізацію робочих процесів, підвищуючи ефективність та забезпечуючи надійність в різноманітних областях бізнесу. Зазначені листи подяки є вагомим додатковим свідченням успішної інтеграції розробленого програмного продукту в реальні умови роботи його користувачів.

У підсумку, дане дослідження вносить важливий внесок у розвиток систем захисту електронних комунікацій від спамових атак, а розроблений програмний засіб виявляється практично важливим інструментом у сфері кібербезпеки.

СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ ДЖЕРЕЛ

1. ПОЛОЖЕННЯ ПРО ДИПЛОМНІ РОБОТИ (ПРОЕКТИ) ВИПУСКНИКІВ НАЦІОНАЛЬНОГО АВІАЦІЙНОГО УНІВЕРСИТЕТУ. Київ: НАУ, 2017.
2. ДОКУМЕНТАЦІЯ. ЗВІТИ У СФЕРІ НАУКИ І ТЕХНІКИ. Структура і правила оформлення. ДСТУ 3008-95. Київ.
3. Statista: Regular sending and receiving of emails in the world [Електронний ресурс]. – Режим доступу: https://www.statista.com/statistics/506315/sending-and-receiving-emails-in-the-world/
4. Symantec Enterprise Blogs: Security Threat Intelligence [Електронний ресурс]. – Режим доступу: https://symantec-enterprise-blogs.security.com/blogs/threat-intelligence/covid-19-outbreak-prompts-opportunistic-wave-malicious-email-campaigns
5. DataProt: Spam Statistics [Електронний ресурс]. – Режим доступу: https://dataprot.net/statistics/spam-statistics/
6. Securelist: Spam, Phishing, and Scam Report 2022 [Електронний ресурс]. – Режим доступу: https://securelist.com/spam-phishing-scam-report-2022/108692/#:~:text=In%202022%2C%20an%20average%20of,to%2046.16%25%20in%20the%20fourth.
7. Ferris Research: The Global Economic Impact of Spam, 2005. Report #409, p. 172–173.
8. OECD: 2004a, Background Paper for the OECD Workshop on Spam, p. 40–41.
9. Doll, J.: n.d., Spam Attack [Електронний ресурс]. – Режим доступу: http://www.joes.com/spammed.html, p. 3.
10. OECD: 2004b, Report of the 2nd OECD Workshop on Spam, p. 17–18
11. Ironport: 2006, Internet Email Traffic Emergency: Spam Bounce Messages are Compromising Networks [Електронний ресурс]. – Режим доступу: http://www.ironport.com/bouncereport/, p. 3.
12. Evett, D.: 2006, Spam Statistics 2006 [Електронний ресурс]. – Режим доступу: http://spam–filterreview.toptenreviews.com/spam–statistics.html, p. 5–6.

13. Slaidik: Як перестати надсилати спам і почати займатися месенджер-маркетингом [Електронний ресурс]. – Режим доступу: http://slaidik.com.ua/yak-perestati-nadsilati-spam-i-pochati-zajmatisya-mesendzher-marketingom/

14. UMJ: Персональні дані в системі охорони здоров'я: аналіз законодавства [Електронний ресурс]. – Режим доступу: https://umj.com.ua/uk/publikatsia-135215-personalni-dani-v-sistemi-ohoroni-zdorov-ya-analiz-zakonodavstva

15. BizMaster: Принципи роботи чорних списків [Електронний ресурс]. – Режим доступу: https://www.bizmaster.xyz/2020/04/pryncypy-roboty-chornyh-spyskiv-i-yak-do-nyh-ne-potrapyty.html

16. Cohen, W.: 1996, Learning rules that classify e–mail, Papers from the AAAI Spring Symposium on Machine Learning in Information Access, p. 18–25.

17. Crawford, E., Kay, J., and McCreath, E.: 2001, Automatic induction of rules for e–mail classification, Proceedings of the Sixth Australasian Document Computing Symposium, p. 34–35.

18. Heartbeat: Spam Filtering Using Bag-of-Words [Електронний ресурс]. – Режим доступу: https://heartbeat.fritz.ai/spam-filtering-using-bag-of-words-1c5484ff07f1.

19. SpamTitan: DNS Server Filter [Електронний ресурс]. – Режим доступу: https://www.spamtitan.com/dns-server-filter/.

20. ProHoster Blog: Захист поштового SMTP-сервера [Електронний ресурс]. – Режим доступу: https://prohoster.info/blog/zashhita-ot-ddos-atak/zashhita-pochtovogo-smtpservera.

21. Andriy Yerokhin, Oleg Zolotukhin. Fuzzy Probabilistic Neural Network in Document Classification Tasks. Information Extraction and Processing. 2018. No. 46, P. 68-71. DOI: http://dx.doi.org/10.15407/vidbir2018.46.068.

22. Connecting Up: Ten Spam-Filtering Methods Explained [Електронний ресурс]. – Режим доступу: https://www.connectingup.org/learn/articles/ten-spam-filtering-methodsexplained.

23. Lifewire: What Is Bayesian Spam Filtering? [Електронний ресурс]. – Режим доступу: https://www.lifewire.com/bayesian-spam-filtering-1164096.

24. UNITE.AI: Що таке машинне навчання? [Електронний ресурс]. – Режим доступу: https://www.unite.ai/uk/%D1%89%D0%BE-%D1%82%D0%B0%D0%BA%D0%B5-%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%B5-%D0%BD%D0%B0%D0%B2%D1%87%D0%B0%D0%BD%D0%BD%D1%8F/

25. Kaggle: SMS Spam Collection Dataset [Електронний ресурс]. – Режим доступу: https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset/data