ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО–НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КАФЕДРА КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: **«ВПРОВАДЖЕННЯ АЛГОРИТМІВ ФІЛЬТРАЦІЇ КОНТЕНТУ В СОЦІАЛЬНИХ МЕРЕЖАХ НА БАЗІ СЕРВЕРНОГО ОБЛАДНАННЯ»**

на здобуття освітнього ступеня магістр
за спеціальності 123 Комп'ютерна інженерія
(код,найменування спеціальності)
освітньо-професійної програми Комп'ютерні системи та мережі
(назва)

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Антон БЕРЕЗОВСЬКИЙ
(підпис) *(ім'я, ПРІЗВИЩЕ здобувача)*

Виконав: здобувач вищої освіти гр.КСДМ-62
Антон БЕРЕЗОВСЬКИЙ
(ім'я, ПРІЗВИЩЕ)

Керівник: Андрій ЛЕМЕШКО
доктор філософії, доцент *(ім'я, ПРІЗВИЩЕ)*

Рецензент: _____
науковий ступінь, вчене звання *(ім'я, ПРІЗВИЩЕ)*

Київ 2023

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут інформаційних технологій

Кафедра Комп'ютерної інженерії
Ступінь вищої освіти «Магістр»

Спеціальність 123 Комп'ютерна інженерія
Освітньо-професійна прорама Комп'ютерні системи та мережі

ЗАТВЕРДЖУЮ
Завідувач кафедрою Комп'ютерної інженерії
Наталія ЛАЩЕВСЬКА
(ім'я, ПРІЗВИЩЕ)
"____" _____2023 року

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Березовському Антону Юрійовичу
(прізвище, ім'я, по батькові здобувача)

1.Тема кваліфікаційної роботи: Впровадження алгоритмів фільтрації контенту в соціальних мережах на базі серверного обладнання
керівник роботи доктор філософії, доцент Лемешко А.В
(ім'я, ПРІЗВИЩЕ, науковий ступінь, вчене звання)
затверджені наказом Державного університету інформаційно-комунікаційних технологій від "19" 10 2023 р. №145
2. Строк подання кваліфікаційної роботи
3. Вихідні дані кваліфікаційної роботи:
3.1. Існуючі системи фільтрування доступу до сайтів.
3.2. Веб-платформи стосовно фільтрування.
3.3. Науково-технічна література.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити):
4.1. Основні відомості про мережу інтернет .
4.2. Методи управління Інтернет каналом.
4.3. Фільтрація контенту сайтів соціальних мереж за допомогою методів машинного навчання.

5.Перелік ілюстраційного матеріалу: *презентація*

6. Дата видачі завдання “19” жовтня 2023р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	**Примітка**
1.	Підбір технічної літератури	.2023р. .2023р.	Виконано
2.	Основні відомості про мережу інтернет	.2023р. .2023р.	Виконано
3.	Методи управління Інтернет каналом	.2023р. .2023р.	Виконано
4.	Фільтрація контенту сайтів соціальних мереж за допомогою методів машинного навчання	.2023р. .2023р.	Виконано
5.	Оформлення роботи, висновки	.2023р. .2023р.	Виконано
6.	Розробка демонстраційного матеріалу, доповідь	.2023р. .2023р.	Виконано

Здобувач вищої освіти Антон БЕРЕЗОВСЬКИЙ
(підпис) *(ім’я, ПРІЗВИЩЕ)*

Керівник кваліфікаційної роботи Андрій ЛЕМЕШКО
(підпис) *(ім’я, ПРІЗВИЩЕ)*

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття ступеня магістр: 84 стор., 17 рис., 23 джерел.

Мета роботи – розробка та впровадження алгоритмів фільтрації контенту в соціальних мережах на базі серверного обладнання.

Об'єкт дослідження – розробка та впровадження алгоритмів фільтрації контенту.

Предмет дослідження – алгоритми фільтрації контенту.

Короткий зміст роботи: В роботі розглянуто програмні та апаратні компоненти, з яких складаються комп'ютерні мережі, основні аспекти функціонування комп'ютерних мереж та основні причини затримок і втрат пакетів в процесі передачі. В роботі було ознайомлено з технологіями і видами керування Інтернет каналом.

У дипломній роботі представлено систему для фільтрації небажаних повідомлень, маркованих зображень та небажаних зображень зі стіни користувача. Для фільтрації текстових повідомлень система використовує текстові класифікатори та методи машинного навчання. Для фільтрації зображень з мітками було використано методи розпізнавання тексту.

КЛЮЧОВІ СЛОВА: ІНТЕРНЕТ, ФІЛЬТРУВАННЯ ДОСТУПУ, ВМІСТ, СОЦІАЛЬНА СТІНА, МЕТОДИ МАШИННОГО НАВЧАННЯ, ФІЛЬТРАЦІЯ ПОВІДОМЛЕНЬ, ФІЛЬТРАЦІЯ МІЧЕНИХ ЗОБРАЖЕНЬ, ФІЛЬТРАЦІЯ ЗОБРАЖЕНЬ

ABSTRACT

The text part of the qualification work for obtaining a master's degree: 84 pages, 17 figures, 23 sources.

The purpose of the work is development and implementation of content filtering algorithms in social networks based on server equipment.

The object of research is development and implementation of content filtering algorithms.

The subject of research is content filtering algorithms.

Summary of the work: The paper examines the software and hardware components that make up computer networks, the main aspects of the functioning of computer networks, and the main causes of delays and packet losses during transmission. In the work, the technologies and types of Internet channel management were introduced.

The thesis presents a system for filtering unwanted messages, tagged images and unwanted images from the user's wall. The system uses text classifiers and machine learning methods to filter text messages. Text recognition methods were used to filter images with labels.

KEYWORDS: INTERNET, ACCESS FILTERING, CONTENT, SOCIAL WALL, MACHINE LEARNING METHODS, MESSAGE FILTERING, TAGGED IMAGE FILTERING, IMAGE FILTERING

ЗМІСТ

ВСТУП

З появою Інтернету життя людей істотно змінилося. Скільки усього може подарувати всесвітня павутина. Багато користувачів Інтернету, переважно молодь, вже не можуть собі уявити, що б вони робили без мережі. Зараз навіть існують бізнесмени, що заробляють гроші в Інтернеті.

Для побудови соціальної мережі або соціальних зв'язків між людьми ми використовуємо соціальні мережі, такі як Facebook, Twitter, додатки тощо. Використовуючи ці медіа, користувачі можуть ділитися своїми поглядами та думками про певні речі. Багато людей використовують свої медіа в особистих інтересах, розвагах, на ринку акцій або в ділових цілях. На сьогоднішній день безпека користувачів є основною проблемою для соціальних мереж. Соціальні мережі в Інтернеті надають невелику підтримку щодо фільтрації контенту. У цій дипломній роботі запропоновано систему, яка забезпечує безпеку щодо шкідливого контенту, який розміщується на сайтах соціальних мереж. Для фільтрації контенту, який може бути небажаними повідомленнями, маркованими зображеннями або вульгарними зображеннями, ми запропонували трирівневу архітектуру. Користувач також може використовувати функцію автоматичного блокування.

У сучасному житті соціальні мережі відіграють дуже важливу роль. Люди проводять більшу частину свого часу в соціальних мережах, спілкуючись і ділячись своїми ідеями. Використовуючи ці медіа, люди можуть ділитися своєю інформацією або обмінюватися різними типами контенту, такими як зображення, відео, текстові або аудіо повідомлення. Багато людей коментують цей спільний контент. Люди отримують зворотній зв'язок на будь-який матеріал, яким вони поділилися на стіні. Іноді це може бути позитивна відповідь, негативна відповідь або пропозиції, які є дуже корисними для покращення. За даними Facebook, користувачі створюють 90 біт контенту щомісяця; щомісяця вони обмінюються більш ніж 30 мільярдами біт контенту (веб-посилання, новини, записи в блогах, нотатки, фотоальбоми тощо).

Користувачі можуть розміщувати в соціальних мережах будь-який тип контенту. Прикладами можуть бути небажані текстові повідомлення, брендовані фотографії, непристойні, порнографічні зображення, особисті пустотливі коментарі тощо. Інші користувачі можуть бачити ці дописи та коментувати їх. Соціальний імідж користувача може постраждати в результаті такого повідомлення. Як наслідок, захист стіни такого користувача є критично важливим. До певного моменту Facebook забезпечує захист. Лише обрана група людей на Facebook має доступ до стін інших людей, таких як друзі, друзі друзів або створені групи друзів. Користувач має можливість заблокувати зображення свого профілю. Однак, оскільки фільтрація на основі контенту не підтримується, такі небажані повідомлення не можуть бути попереджені. Метою даної системи є захист стін користувача шляхом фільтрації небажаного контенту та соціальних мереж користувача медіа-зображення. Користувачі можуть змінювати правила фільтрації на свій розсуд. Користувач має контроль над тим, хто може надсилати повідомлення на його стіну. Для фільтрації тексту використовується метод класифікації коротких текстів.

1 ОСНОВНІ ВІДОМОСТІ ПРО МЕРЕЖУ ІНТЕРНЕТ

1.1 Загальне поняття про Інтернет

Інтерне́т (від англ. Internet) — всесвітня система взаємосполучених комп'ютерних мереж, що базуються на комплекті Інтернет-протоколів. Інтернет також називають мережею мереж. Інтернет складається з мільйонів локальних і глобальних приватних, публічних, академічних, ділових і урядових мереж, пов'язаних між собою з використанням різноманітних дротових, оптичних і бездротових технологій. Інтернет становить фізичну основу для розміщення величезної кількості інформаційних ресурсів і послуг, таких як взаємопов'язані гіпертекстові документи Всесвітньої павутини (World Wide Web — WWW) та електронна пошта.

В повсякденній мові слово Інтернет найчастіше вживається в значенні Всесвітньої павутини і доступної в ній інформації, а не у значенні самої фізичної мережі. Також вживаються терміни Всесвітня мережа, Глобальна мережа чи навіть одне слово Мережа, Іне́т, Тенета, Міжмережжя, Інтерне́трі або Не́трі. Все частіше Інтернет вживається і з малої літери, що можна пояснити паралелями з термінами «радіо», «телебачення», які пишуть з малої.

Історія Інтернету сягає досліджень 1960-х років, які проводилися на замовлення уряду США і мали на меті створення надійних розподілених комп'ютерних мереж, стійких до пошкоджень. Попередницею Інтернету стала мережа ARPANET (англ. Advanced Research Projects Agency Network), яка почавши функціонувати в кінці 1960-х, в кінці 1970-х об'єднувала близько 200 вузлів.

Урядове фінансування магістральної мережі Національного наукового фонду США в 1980-х, а також приватне фінансування для інших комерційних магістральних мереж в усьому світі призвело до участі в розробці нових мережевих технологій і злиття багатьох мереж. Комерціалізація в 1990-х міжнародної мережі

привела до її популяризації та впровадження в практично кожен аспект сучасного життя людини. З 2011 року більше 2,1 мільярда людей користуються послугами Інтернету.

Інтернет не має централізованого управління, правил використання чи доступу. Кожна складова мережа встановлює свої власні стандарти. Централізовано визначаються правила використання адресного простору Інтернет-протоколу та Системи доменних імен. Керує цим Інтернет корпорація з присвоєння імен та номерів (англ. Internet Corporation for Assigned Names and Numbers, або ICANN), міжнародна некомерційна організація з головним офісом у США. Технічне обґрунтування і стандартизацію основних протоколів (IPv4 та IPv6) проводить Internet Engineering Task Force (IETF), некомерційна організація, відкрите міжнародне співтовариство проектувальників, учених, мережевих операторів і постачальників послуг.

Мережа побудована на використанні протоколу IP і маршрутизації пакетів даних. В наш час Інтернет відіграє важливе значення у створенні інформаційного простору глобального суспільства, слугує фізичною основою доступу до веб-сайтів і багатьох систем (протоколів) передачі даних.

1.2 Структура мережі Інтернет

Образ Інтернету можливо представити як безкрайнього міста-магполіса, який складається з конгломерату окремих будинків, районів і кварталів, з'єднаних в одне ціле магістралями. Кожен такий окремий будинок - це самостійна автономна мережа, спільність яких за допомогою з'єднання каналами зв’язку і породжує Інтернет (рис. 1.1). А цементує основи Інтернету протокол TCP / IP – така мережева мова цього міста.

Становий хребет Інтернету складають його опорні мережі (CoreBackbone Network) провайдерів вищого рівня. Всі опорні мережі без обмежень обмінюються між собою Інтернет трафіком. Весь інший світ отримує доступ до хребта Інтернету вже через провайдерів першого (транснаціонального) рівня, що мають вихід в різні

країни. Слідом за провайдерами першого рівня розташовані мережеві провайдери другого рівня - національного і третього - регіонального, з'єднані між собою високошвидкісними каналами передачі даних, які надають доступ до Інтернету місцевим провайдерам - Internet Service Provider.

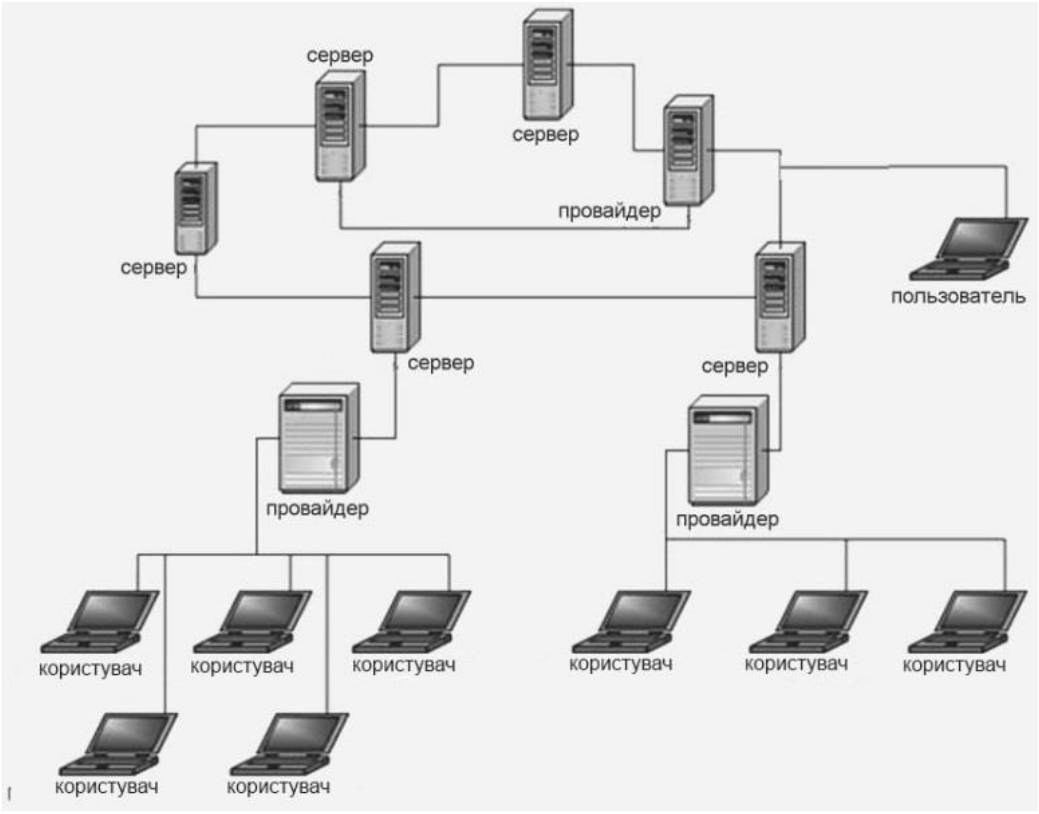

Рисунок 1.1 - Загальна схема глобальної мережі Інтернет

Провайдер - компанія, яка забезпечує вихід в Інтернет, тобто «постачальна» вас цією послугою. Саме провайдер на локальному рівні і забезпечує вихід в Інтернет індивідуальних користувачів. Кожен провайдер на своєму рівні вирішує всі організаційні, технічні та фінансові питання, представляючи в своїй особі перед вами всю всесвітню мережу.

Інтернет трафік - кількість переданої інформації,вимірюється в байтах.

Шлюз (gateway) - це комп'ютер або система комп'ютерів зі спеціальним програмним забезпеченням, що дозволяє зв’язатися двом мережам з різними протоколами. Найчастіше шлюзи зв’язують локальні обчислювальні мережі LAN (Local Area Network) з глобальною мережею WAN (Wide Area Network).

Маршрутизатор (router) - пристрій, який пов'язує мережі побудованих на основі одного протоколу, але різними типами мережевого обладнання.

Маршрутизатори зменшують трафік, пропускаючи в мережу тільки ті дані, які призначені саме для неї.

Протокол передачі даних - спеціальні набори правил, які забезпечують обмін інформацією як між окремими пристроями, так і між цілими мережами.

Комп'ютери, включені до світової мережі, мають абсолютно різну архітектуру і різне програмне забезпечення. Для забезпечення сумісності мереж були створені протоколи - спеціальні набори правил, забезпечуючи обмін інформацією як між окремими пристроями та процесами, так і між цілими мережами.

TCP (Transmission Control Protocol) відповідає за те, як буде проходити інформація з всесвітньої мережі. Він відповідає встановлення надійного з'єднання між комп'ютерами і пересилання даних, контролюючи оптимальний розмір пакетів даних, генеруючу повторну передачу пакету при збої.

IP (Internet Protocol) відповідає за те, куди буде надсилатися інформація, тобто завідує адресацією пакетів. Протоколом TCP виконується нарізка направлених файлів на пакети, кожен зі своїм точною адресою розміщення в структурі файлу. За місцем прибуття отримані фрагменти збираються в єдине ціле.

Головні особливості протоколів TCP / IP:

- відкритість стандартів, що розробляються незалежно від програмного і апаратного забезпечення мережі;

- незалежність від безпосередньої фізичної середовища передачі;

- унікальність адресації;

- стандартизованість протоколів високого рівня, які використовуються в сервісах.

Для того щоб інформація знайшла потрібну програму (адже на комп’ютері одночасно працює безліч різних програм різного призначення) існує система портів. Порт - спеціальний номер, який присвоюється кожному процесу на

комп'ютері і який виконує роль адреси відправника та адреси одержувача на транспортному рівні.

Розглянемо структурні складові Інтернету, апаратне та програмне забезпечення. Розглянемо рис. 1.2.

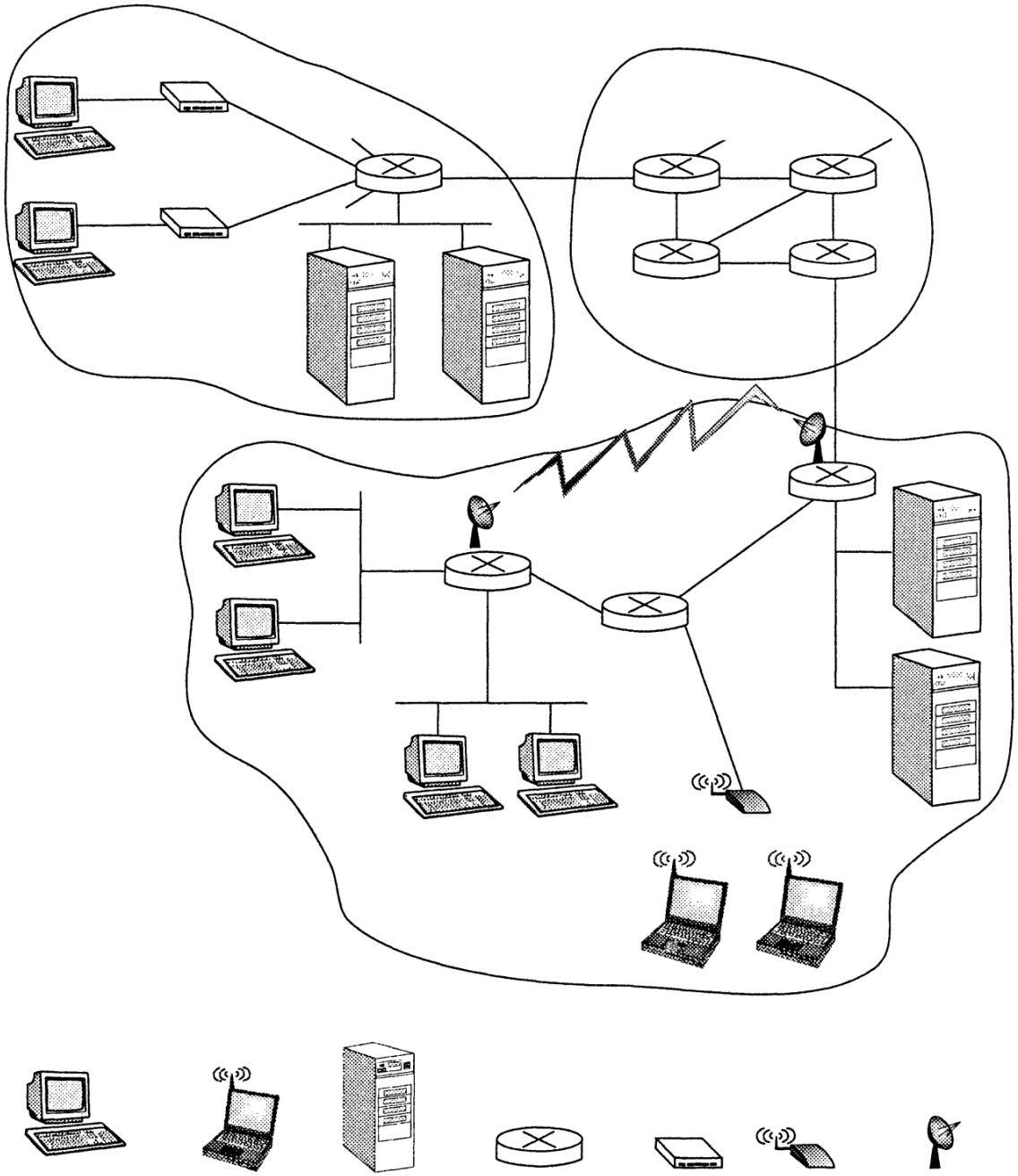

Рисунок 1.2 - Структурна схема Інтернету

Інтернет являє собою світову комп'ютерну мережу, яка зв'язує в єдине ціле мільйони обчислювальних пристроїв, розміщених в різних кутах земної кулі. Обчислювальні пристрої можуть бути настільними комп'ютерами, або так званими серверами, зберігаючи ми або передаючи інформацію, представлені в вигляді, наприклад, веб-сторінок або повідомлень електронної пошти. Останні роки до Інтернету під'єднуються нетрадиційні кінцеві системи, такі як PDA (Personal Digital Assistant – персональний цифровий помічник), телевізори, мобільні комп'ютери, автомобілі і навіть холодильники. По оцінкам спеціалістів в 2002 році в Інтернеті нараховували від 100 до 500 мільйонів кінцевих пристроїв, так званих хостів.

Кінцеві пристрої зв'язані між собою лініями зв'язку. Існує велика кількість ліній зв'язку, які використовують різні типи фізичних носіїв: коаксіальні, мідні, оптоволоконні кабелі, лінії радіозв'язку ті інші. Лінія зв'язку визначає швидкість передачі даних, а максимальну швидкість передачі даних називають пропускною здатністю лінії та виміряють в бітах на секунду.

Хости далеко не завжди напряму з'єднані між собою єдиною фізичною лінією зв'язку. Навпроти, типовою є ситуація, коли зв'язок здійснюється за допомогою множини послідовних ліній, з'єднаних спеціальними комутуючими пристроями – маршрутизаторами. Маршрутизатор приймає порцію пакетів, передану по одному з його вхідних каналів зв'язку, а потім перенаправляє її в один зі своїх вихідних каналів зв'язку. В термінології комп'ютерних мереж передані порції даних називають пакетами. Послідовність каналів зв'язку та маршрутизаторів, через які проходить пакет в процесі передачі, називається маршрутом, або шляхом, пакета в мережі. Шлях пакета не відомий завчасно і визначається безпосередньо в процесі передачі пакета. В Інтернеті кожній парі кінцевих систем не надають виділений маршрут, тому що використовується технологія комутації пакетів, при цьому різні пари кінцевих систем можуть одночасно користуватися одним й тим самим маршрутом або його частиною. Перші мережі з комутацією пакетів, які були створені в далекі 70-ті роки, є «далекими родичами» сучасного Інтернету.

Доступ кінцевих систем до Інтернету здійснюється за допомогою постачальників послуг Інтернету, або Інтернет провайдерів (Internet Service Provider, ISP). Інтернет-провайдери поділяються на резидентних, університетських та корпоративних. Інтернет-провайдер надає мережу маршрутизаторів і канали зв'язку. Як правило, Інтернет-провайдери пропонують декілька способів підключення хостів до мережі: комутоване модемне з'єднання, резидентне широко полосне підключення за допомогою кабельного модему або цифрової абонентської лінії (Digital Subscriber Line, DSL), високошвидкісний доступ через локальну мережу (Local Area Netwokr, LAN), а також безпровідний доступ. Крім того, Інтернет-провайдери здійснюють пряме підключення до мережі веб-сайтів. Для забезпечення зв'язку між віддаленими користувачами, а також для надання доступу до інформації, яка зберігається в Інтернеті, місцеві Інтернет-провайдери підключаються до Інтернет-провайдерів національної або інтернаціональної ланки. Остнанні використовують високошвидкісні маршрутизатори, з'єднані оптоволоконними кабелями. Кожний Інтернет- провайдер як нижньої, так верхньої ланки є адміністративною одиницею, яка передає дані по протоколу IP та дотримується угод про імена та адреси, прийнятих в Інтернеті.

Кінцеві системи, маршрутизатори та інші компоненти Інтернету використовують протоколи, які здійснюють управління прийомом та передачі інформації всередині Інтернету. Найбільш важливим протоколом в глобальній мережі є TCP (Transmission Control Protocol – протокол управління передачею) та IP (Internet Protocol – Інтернет-протокол). Протокол IP визначає формат пакетів, які передаються між хостами та маршрутизаторами. Стек основних протоколів, які використовуються в Інтернеті, відомий під назвою TCP/IP .

Те що ми зазвичай називаємо словом «Інтернет», - це так званий «відкритий Інтернет». Крім загальнодоступного Інтернету існує багато закритих (приватних) комп'ютерних мереж, побудованих по тому ж принципу що і глобальна мережа. Як правило, приватні мережі призначені для використання всередині різних фірм та організацій. Вони не можуть обмінюватися інформацією з зовнішньою мережею, за виключенням повідомлень, які проходять через так звані брандмауери,

контролюючі потік повідомлень, які входять чи виходять з мережі. Подібні мережі об’єднують терміном інтранет. Ця назва співзвучна назві «Інтернет» і відображає той факт що в зачинених мережах використовують такі ж самі хости, маршрутизатори, канали зв’язку та протоколи, канали зв’язку, що і в відкритому Інтернеті. З точки зору технологій і розвитку існування Інтернету забезпечується створенням, перевіркою та впровадженням Інтернет-стандартів. Стандарти виробляє «проблемна група розробок для інтернету» (Internet Engineering Task Force, IETF). Документи які створює група мають назву RFC (Requests For Comments)

1.3 Мережа інтернет з погляду обслуговування

З точки зору обслуговування Інтернет можливо поділити на декілька пунктів:

- Інтернет дозволяє розподіленим програмам, працюючих на кінцевих системах, здійснювати обмін даними між собою. В число таких програм входять віддалений робочий стіл, електронна пошта, засоби навігації в веб, засоби передачі відео та аудіо даних, Інтернет телефонія, мережеві комп’ютерні ігри, засоби обміну даними та інше. Маємо підкреслити що веб це не окрема комп’ютерна мережа, а одне з багатьох розподілених програмних засобів;

- Інтернет надає своїм розподіленим програмам два типи служб: надійну службу з встановленням логічного з’єднання та ненадійну службу без встановлення логічного з’єднання. Ці поняття означають наступне: надійна служба з встановленням логічного з’єднання гарантує, що дані які передаються відправником будуть доставлені отримувачу повністю (без втрат та пошкоджень) та в початковому порядку; ненадійна служба без встановлення логічного з’єднання не надає ніяких гарантій відносно доставки. Як правило розподілене програмне забезпечення підтримує один з двох типів передачі;

- на даний час Інтернет не дає гарантій відносно того, скільки часу знадобиться для передачі даних від відправника до отримувача. І, якщо не врахувати можливість підвищення пропускної здатності каналу доступу до вашого

Інтернет-провайдера, ми можемо затребувати в Інтернеті більш високої якості обслуговування, якщо готові заплатити за це.

1.4 Сучасні мережеві протоколи

Будь-який рух інформації в Інтернеті між двома або більше пристроями підпорядковується протоколу. Так протоколи маршрутизаторів визначають шлях пакета від відправника до отримувача; реалізовані апаратно протоколи мережевих карт двох фізично з'єднаних комп'ютерів контролюють потік бітів, які передаються по мережевому кабелю; протоколи контролю перезавантаження, які використовуються в кінцевих системах, потрібні для контролю частоти передачі пакеті. Інтернет повністю заснованих на протоколах.

В якості прикладу, одночасно простого і який наглядно ілюструє суть мережевого протоколу, розглянемо що відбувається в момент, коли ми відтворюємо запит до веб серверу. Графічно дану ситуацію ілюструє рис. 1.3. Спочатку комп'ютер відсилає серверу повідомлення із запитом на встановлення з'єднання і чекає на відповідь. Сервер приймає запит та відправляє повідомлення у відповідь, яке підтверджує встановлене з'єднання. Таким чином, знаючи, що тепер можливо запросити веб-документ, комп'ютер відсилає серверу ім'я ресурсу, а сервер повертає заданий ресурс користувачу.

Розглянувши приклади протоколів які демонструють дві найбільш значимими складових протоколу – повідомлення і дія, можу сформулювати наступне визначення: Протокол визначає формат та чергу повідомлень, якими обмінюються два або більше пристроїв, а також дії, які виконуються при передачі або прийомі повідомлень. Протоколи дуже широко використовуються взагалі в мережах, так і в мережі Інтернет.

1.5 Ядро комп’ютерних мереж

Предметом розгляду стане взаємодія між маршрутизаторами, точніше механізми передачі даних від одного хоста до іншого. Елементи структури мережі, які відносяться до ядра показані на рис. 1.3.

1.5.1 Комутації каналів та пакетів в мережі

Існує два фундаментальних підходу до організації ядра мережі: комутація каналів та комутація пакеті. При комутації каналів відбувається

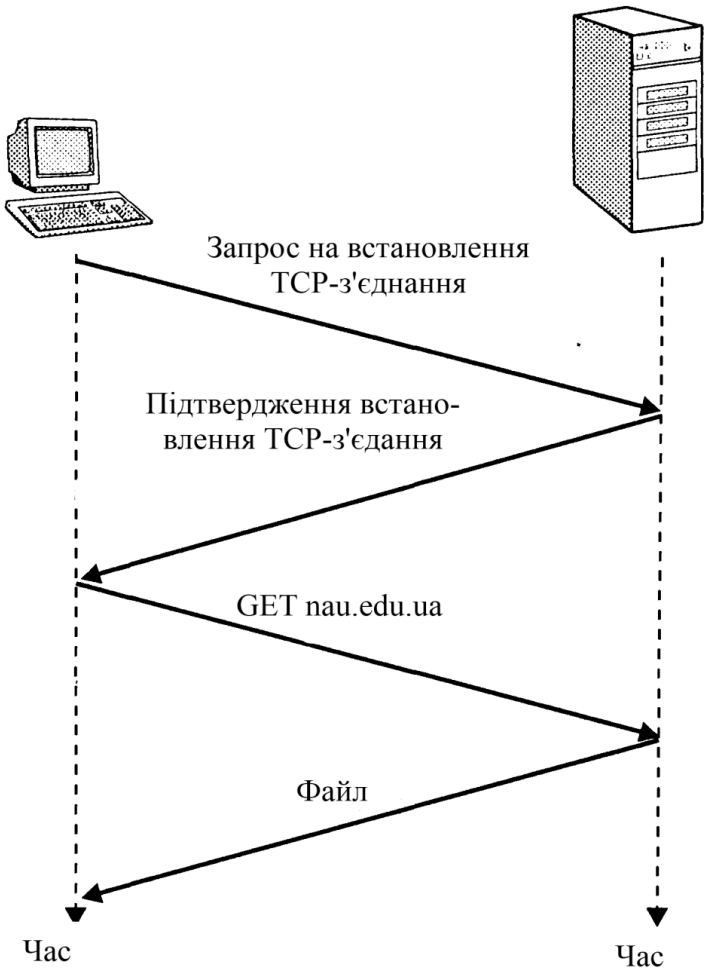

Рисунок 1.3 - Елементи структури мережі, які відносяться до ядра

Протокол спілкування між комп’ютерами резервування на час сеансу зв’язку необхідних ресурсів на всьому мережевому шляху. При комутації пакетів ресурси запитуються при необхідності і виділяються за потрібністю. Іноді декілька повідомлень можуть намагатися використати лінію зв’язку одночасно, тому існує необхідність в організації черги повідомлень.

Сучасний Інтернет є типовою системою з комутацією пакетів. Як правило, пакет проходить через множину каналів, однак ніяких резервувань частотних полос при цьому не відбувається. У випадку перенавантаження каналу, пакет буде змушений чекати в черзі звільнення. Таким чином з точки зору швидкодії

Інтернет намагається доставити пакет з максимальними зусиллями, час доставки не гарантовано.

В мережах з комутацією каналів комутатори з’єднані меж собою лініями зв’язку. Кожна з ліній може підтримувати одночасно n каналів зв’язку. Хости напряму з’єднані з одним із комутаторів. Між парами хостів встановлюється виділене сковзне з’єднання. Таким чином, щоб хост А мав можливість передавати пакети хосту Б, потрібно зарезервувати одну полосу частот на кожній з ліній зв’язку, з’єднуючих хост А і Б.

Прихильники технології комутації пакетів завжди звертали увагу на серйозний недолік мереж з комутацією каналів, який заключається в тому, що виділені канали неможливо звільнити під час простою. Іншою причиною за яку комутація каналів викликає обґрунтовану критику, це необхідність в складному сигнальному обладнанні для управління комутаціями і виділенням частотних полос каналам зв’язку.

В сучасних комп’ютерних мережах відбувається автоматичне розбиття великих за об’ємом повідомлень на більш малі фрагменти, пакети. Пакет є одиницею передачі даних. При передачі пакет проходить через послідовність ліній зв’язку та комутаторів, звичайно названих маршрутизаторами. Передача пакета по лінії зв’язку здійснюється монопольно, з максимальною швидкістю, яку може забезпечити лінія зв’язку. Більшість маршрутів використовують механізм передачі з проміжним накопиченням. Це означає, що перед тим як почати передачу в вихідну

лінію зв’язку, маршрутизатору потрібно завершити процес прийому пакета в буфер. Таким чином в маршрутизаторах виникає затримка накопичення, зумовлена необхідністю очікування закінчення прийому пакета.

Кожний маршрутизатор має множину вхідних і вихідних ліній зв’язку. Кожна вихідна лінія має буфер, який називають вихідним буфером. В буфері зберігаються пакети, призначені для передачі по лінії зв’язку. Буфери грають ключову роль в механізмі комутації пакетів. Якщо при закінченні прийому пакета виявляється що лінія зв’язку зайнята, то пакет ставиться в чергу в вихідному буфері. Таким чином крім затримки накопичення в буфері в маршрутизаторах присутня затримка очікування. Затримки очікування є змінними показниками і залежать від завантаженості каналу. Оскільки розміри буферів обмежені, можлива ситуація коли місця в буфері буде недостатньо для розташування нового пакету. В такому випадку виникне втрата пакету – буде втрачений або новий пакет, або один за пакетів які знаходяться в черзі.

На рис. 1.4 приведена структура простої мережі з комутацією пакетів. Тут пакети представлені у вигляді тривимірних брусків. Ширина бруска відповідає довжині пакету. В даному прикладі всі пакети мають однакому довжину. Припустимо що хости А та В посилають пакету хосту Е, при цьому зв'язок хостів А та В з першим маршрутизатором здійснюється за допомогою ліній зв’язку Ethernet. Маршрутизатор направляє пакет на зовнішню лінію зв’язку. Якщо лінія перенавантажена, пакети очікують її звільнення в черзі. Подивимося, що відбувається при одночасній передачі пакетів хостами А і В. Очевидно, що ніякої синхронізації між хостами немає, і, відповідно, неможливо завчасно передбачити порядок передачі пакетів. Цю особливість називають стичним мультиплексором. Противники комутації пакетів часто висувають тезис про те, що комутація пакетів не дозволяє організувати мережеве обслуговування в реальному часі (наприклад забезпечити передачу звука та відео), пояснюючи це непередбачуваними затримками в при передачі пакетів всередині мережі. Прихильники комутації пакетів помічають що дана технологія дає можливість більш ефективно

організувати розподілення пропускної можливості лінії зв’язку, а також є більш простою, ефективною та дешевшою.

1.5.2 Сегментування повідомлень в мережі

В більшості сучасних мережах з комутацією пакетів передаючий хост розбиває довгі повідомлення, які генерує програма, на менші пакети. Ці пакети отримує адресат, з яких збирає вихідне повідомлення. Значною перевагою розбиття на пакети заключається в тому, що час передачі повідомлення, як правило, значно скорочується порівняно з передачею повідомлення цілим.

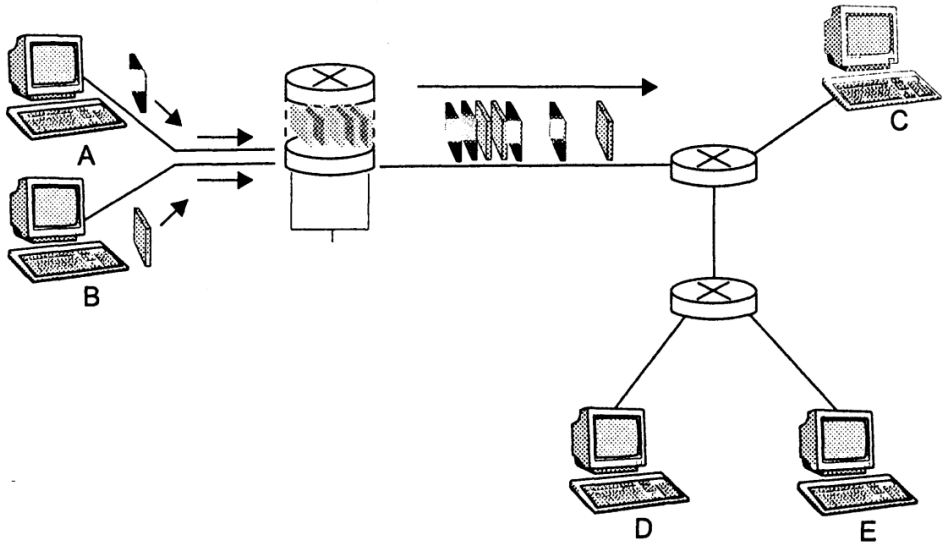

Рисунок 1.4 - Структура мережі з комутацією пакетів

1.5.3 Передача повідомлень в мережі

Існує два основних класи комп’ютерних мереж з комутацією пакетів: дейтаграмні мережі та мережі з віртуальним каналом. Ці два класи відрізняються між собою механізмом передачі пакетів в мережі. Мережі, в я яких передача здійснюється на основі аналізу адреси отримувача, називають дейтаграмними. Дейтаграмний спосіб передачі характерний для Інтернету. Якщо все ж таки в

мережі використовується механізм передачі з віртуальним каналом, то кажуть мережі з віртуальними каналами. До останніх відносять мережі які підтримують протокол X.25, ретрансляцію кадрів, асинхронний режим передачі.

Віртуальний канал (Virtual Channel, VC) характеризується трьома складовими:

- маршрутом, по якому передаються всі пакети від відправника до отримувача;

- номерами віртуального каналу, по одному номеру на кожну лінію зв'язку, які створюють маршрут;

- записами в таблицях трансляції номерів віртуального каналу, які є в кожному з комутаторів на маршруті.

Після того як з'єднання між отримувачем та відправником встановлено (створено віртуальний канал), відправник може почати пересилку пакетів з відповідними номерами віртуального каналу. Оскілки кожна лінія зв'язку має свій номер віртуального каналу, кожний раз при проходженні пакету через комутатор, останній повинен автоматично змінювати для пакета значення номеру віртуального каналу. Новий номер віртуального каналу пакет отримує за допомогою таблиці трансляції номерів віртуального каналу.

Концепцію віртуального каналу ілюструє рис. 1.5. Припустимо що хост А запросив віртуальний канал з хостом В,і мережа встановила віртуальний канал з маршрутом A-PS1-PS2-B, назначивши лініям зв'язку номери 12, 22 та 32 відповідно. Таким чином, кожний пакет який відправляється через хост А, має номер 12, а пакети які відправляються за маршрутизаторів PS1 та PS2, номери 22 та 32 відповідно.

В дейтаграмній мережі кожний пакет що передається включає в себе інформацію про адресу отримувача, який має ієрархічну структуру. Кожний раз при отриманні пакета комутатор аналізує фрагмент адреси пакета і направляє його на відповідну лінію зв'язку. Комутатор забезпечений таблицею маршрутизації, який зв'язує кінцеві адреси або її частини з лініями зв'язку. Після зчитування

заголовка відбувається виділення адреси, який використовується в якості індексу таблиці маршрутизації.

Дейтаграмні мережі, на відміну від мереж з віртуальними каналами, не використовують інформацію про стан з'єднань в своїх комутаторах. Фактично люба мережа, побудована на дейтаграмній передачі, не контролює інформаційні потоки всередині себе, оскілки рішення про шлях слідування любого пакета приймається виключно на основі його адреси призначення і не залежить від з'єднання між хостами. Простота дейтаграмного механізму дає привід для критичних нарікань на адресу віртуальних каналів, відносно складності останніх. Прихильники віртуальних каналів відповідають на ці нарікання тим, що технологія забезпечує краще мережеве обслуговування програм.

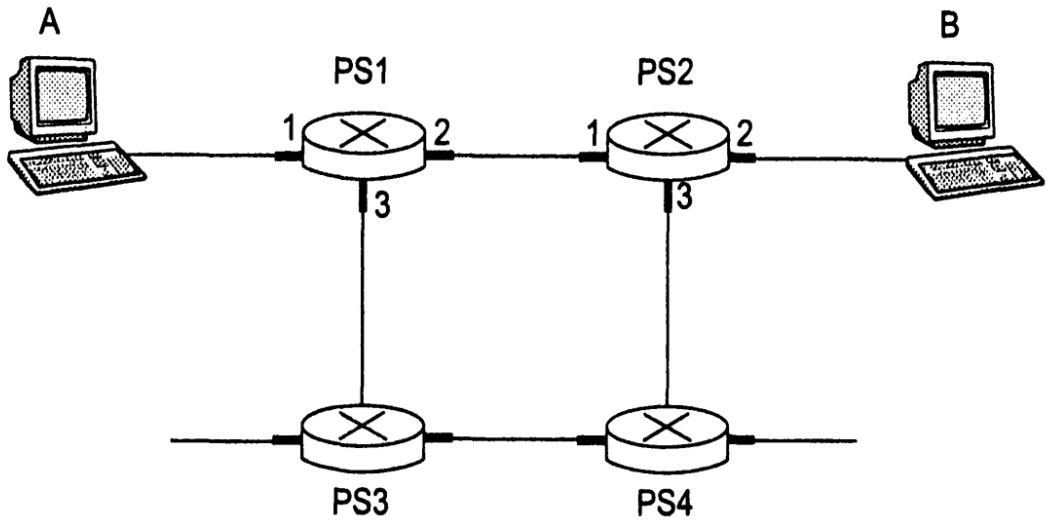

Рисунок 1.5 - Проста мережа з віртуальним каналом

1.6 Доступ до комп'ютерної мережі та її фізична середа

Доступ до мережі можливо умовно класифікувати наступним чином:

- резидентний доступ використовується для підключення до мережі домашніх кінцевих систем;

- корпоративний доступ використовується для підключення до мережі кінцевих систем, що належать приватним або державних організаціях;

- мобільний доступ, що використовується для підключення різних портативних пристроїв.

Але, часто на практиці зустрічаються порушення приведеної класифікації, тому ці поняття відповідають тільки як найбільш типові випадки доступу.

1.6.1 Резидентний доступ

Спочатку, як правило, резидентний доступ здійснювався за допомогою модему та комутованої телефонної лінії через підключення до місцевого Інтернет-провайдера. Модем перетворював цифрові сигнали домашньої системи в аналогові сигнали, які передаються через телефонний кабель, який являє собою мідну виту пару. Сигнал приймається стороною Інтернет провайдера, де модем здійснює зворотне перетворення аналогових сигналів в цифрові та передає їх на вхід маршрутизатора. Таким чином типовий резидентний доступ до мережі забезпечується парою модемів, з’єднаних за допомогою комутованої лінії. На сьогодні модеми дозволяють забезпечити швидкість передавання даних за швидкістю 56 Кбіт/с, що, в наші часи, вважається дуже малою.

Ця проблема отримала своє рішення у вигляді нових широкополосних засобів передачі інформації. Існують два основні засоби широкополосного доступу: цифрові абонентські лінії (Digital Subscriber Line, DSL) та оптоволоконно-коаксиалні кабелі (Hybrid Fiber Coaxial Cable, HFC).

Як правило DSL-доступ забезпечується телефонними компаніями, іноді сумісно з Інтернет-провайдерами. DSL-технологія нагадує звичайних модемний доступ по телефонному кабелю, але дозволяє, за рахунок скорочення відстані від користувача до модему Інтернет-провайдера, значно підвищити швидкість передавання інформації. Звичайно швидкість між обміну між сторонам асиметричні, при цьому швидкість передачі у напрямку користувача значно вище, чим швидкість передачі в сторону Інтернет-провайдера.

1.6.2 Корпоративний доступ

Як правило в державних або приватних організаціях доступ до Інтернету здійснюється за допомогою локальних мереж (LAN), які з’єднують кінцеві системи з периферійним маршрутизатором. Серед усіх технологій локальних мереж, найбільш поширеною є технологія Ethernet. В ній для з’єднання кінцевих систем між собою та периферійним маршрутизатором використовується мідна вита пара. На сто годній цей спосіб підключення до Інтернету широко використовується для підключення домашніх систем.

1.6.3 Мобільний доступ

Прорив в області бездротових технологій, як і виникнення глобальної Мережі, привів до значних змін у сфері телекомунікацій. В 2000 році в Європі було більше власників мобільних телефонів, ніж власників машин або персональних комп'ютерів. В наш час існує два основні засоби безпровідного підключення до глобальної мережі. Бездротові локальні мережі дозволяють користувачам обмінюватися даними через базову станцію, часто звану точкою безпровідного доступу, перебуваючи на відстані десятків метрів від неї. Як правило, базова станція має підключення до Інтернету за допомогою кабелю і здатна поєднати користувачів з глобальною мережею. У бездротових мережах з віддаленим доступом базова станція управляється постачальником телекомунікаційних послуг і забезпечує доступ користувачів на відстані до десятків кілометрів.

Бездротові локальні мережі, засновані на технології IEEE 802.11b, відомого як бездротова Ethernet-мережа і Wi-Fi, в даний час отримують масове поширення в різних навчальних, комерційних, розважальних організацій і навіть при домашньому користуванні. Подібні технології, реалізовані в будівлі, дозволяють користувачам задіяти електронну пошту або займатися подорожам по веб-сторінкам в будь-якій точці цієї будівлі.

Зараз спостерігається розширення кола домашніх користувачів, які разом з широкосмуговим доступом застосовують недорогі бездротові локальні мережі для створення потужних домашніх комп'ютерних мереж. Мережа включає в себе портативний та персональний комп'ютери, базову станцію, з якою пов'язаний портативний комп'ютер, кабельний модем, який здійснює підключення персонального комп'ютера до Інтернету, і, нарешті, маршрутизатор, що з'єднує базову станцію і персональний комп'ютер з кабельним модемом. Описана мережа дозволяє двом членам сім'ї мати широкосмуговий доступ в Інтернет, причому один з них може при цьому ходити з однієї кімнати в іншу.

Отримуючи доступ до Інтернету через бездротову локальну мережу, ви пов'язані необхідністю знаходитися на відстані не більше декількох десятків метрів від базової станції. Це допустимо для домашнього або корпоративного користування. Якщо ви знаходитеся в машині або за містом на відпочинку, на допомогу приходять технології мобільного телефонного зв'язку, які забезпечують доступ до глобальної Мережі на відстані десятків кілометрів від базової станції.

Технології WAP (Wireless Access Protocol - протокол бездротового доступу). WAP-телефони нагадують звичайні мобільні телефони, проте мають дещо збільшений екран і забезпечують низькошвидкісний доступ в Інтернет.

Замість мови HTML в WAP-телефонах використовується мова розмітки WML (WAP Markup Language), оптимізований під низькошвидкісний доступ і невеликий екран. Протокол WAP в Європі підтримується стандартом мобільного зв'язку GSM. У зв'язку з поширенням технології GPRS (General Packet Radio Service - основна служба радіотрансляції пакетів) очікується зростання популярності WAP.

Зараз телекомунікаційні компанії роблять великі інвестиції в безпровідні технології третього покоління (Third Generation, 3G), які дозволять здійснювати віддалений бездротовий доступ в Інтернет з комутацією пакетів на швидкостях не нижче 384 Кбіт/с. 3G -системи забезпечать високошвидкісний доступ до веб-ресурсів і інтерактивного відео, а також телефонний зв'язок з більш високою якістю звуку, ніж у звичайних телефонних мережах.

1.7 Фізична середа передачі даних в комп’ютерній мережі

Передача між пристроями відбувається шляхом поширення електромагнітних хвиль або оптичних сигналів у фізичному середовищі. Фізична середа може приймати вельми різноманітні форми, причому на шляху прямування пакета ці форми можуть змінюватися. Прикладами фізичних середовищ є мідна вита пара, коаксіальний кабель, багатомодовий оптоволоконний кабель, територіальні та супутникові радіоканали.

Фізичні середовища можна розділити на два типи: провідні і безпровідні. Провідні середовища передачі припускають присутність твердотільного провідника і включають оптоволоконний кабель, мідну виту пару і коаксіальний кабель. У бездротовому середовищі передача здійснюється без участі твердих провідників; цей тип середовища використовується в безпровідних локальних мережах і при супутниковому зв'язку.

Мідна вита пара

Мідна вита пара є найдешевшим і найбільш популярним видом кабелів.

Протягом більш ніж 100 років вита пара активно використовується в телефонних мережах. Можна сміливо стверджувати, що більше 99% всіх кабелів, з'єднуючих абонентів з телефонними комутаторами, є мідними витими парами. Вита пара складається з двох ізольованих мідних дротів товщиною 1 мм, укладених в спіральну оболонку. Усередині оболонки дроти переплетені один з одним, щоб знизити рівень електричних перешкод, що виникають між парою провідників. Зазвичай перед поміщенням пар всередину кабелю їх забезпечують додатковими захисними екранами.

Неекранована вита пара (Unshielded Twisted Pair, UTP), як правило, використовується в офісних локальних мережах, розташованих в одній будівлі.

З появою в 80-ті роки оптоволоконних ліній зв'язку багато фахівців прогнозували, що вони з часом повністю витіснять низькошвидкісну виту пару. Однак вита пара виявилася не настільки безперспективною. Неекрановану виту пара категорії 5 дозволяє отримати швидкість передачі даних 100 Мбіт / с на

відстанях до сотні метрів. На менших відстанях можна добитися ще більшої швидкості. Цей тип кабелю ще довго може займати домінуюче положення в сфері локальних офісних мереж. Вита пара активно використовується для резидентного доступу в Інтернет.

Коаксіальний кабель

Коаксіальний кабель, як і вита пара, складається з двох мідних провідників, але ці провідники, на відміну від витої пари, розташовані не паралельно, а концентрично (коаксіально). Із застосуванням особливих видів ізоляції і екранування коаксіальний кабель дозволяє домогтися більш високих швидкостей передачі даних, ніж кручена пара. Коаксіальні кабелі поділяються на два види: з не модульованою передачею і з модульованою передачею. Коаксіальний кабель з не модульованою передачею має опір 50 Ом і товщину близько 1 см; до його фізичних переваг можна віднести легкість і гнучкість. Коаксіальний кабель з модульованою передачею має опір 75 Омі і має велику товщину, вагу і меншу гнучкість в порівнянні з кабелем з не модульованою передачею. Часто кабель з модульованим передачею використовують в системах кабельного телебачення. Лінії кабельного телебачення в поєднанні зі спеціальними кабельними модемами здатні забезпечити зв'язок користувачів з Інтернетом на швидкості до 20 Мбіт/с і вище. При передачі по кабелю з модульованим передачею відбувається попередня модуляція (перенесення) аналогових сигналів в потрібну смугу частот.

Оптоволоконна середовище передачі являє собою тонкий і гнучкий кабель, всередині якого поширюються світлові імпульси, що несуть інформацію про передані біти. Навіть простий оптоволоконний кабель здатний передавати дані на величезних швидкостях в десятки і навіть сотні гігабіт на секунду. Оптоволоконні лінії не схильні електричним наведенням, мають дуже низький рівень ослаблення сигналу на одиницю довжини і мають значну стійкість до механічних впливів. Перераховані переваги зробили оптоволоконні лінії зв'язку вельми привабливою технологією для передачі інформації на великі відстані, особливо для міжнародних і міжконтинентальних комунікацій.

Оптоволоконне середовище

Активно використовується для передачі даних в Інтернеті. Однак висока вартість оптичних пристроїв (маршрутизаторів, приймачів і передатчиків) робить недоцільним (з економічних причин) застосування оптоволоконних ліній зв'язку для передачі на короткі відстані, наприклад, в локальних офісних мережах або для резидентного домашнього доступу.

Територіальні радіоканали

Радіоканали передають сигнали за допомогою електромагнітних хвиль радіодіапазону. Їх перевага полягає в тому, що для зв'язку не потрібно твердотільної провідника сигналів (отже, немає необхідності в його прокладці), тобто користувач може бути мобільним, є потенціал у збільшенні відстані передачі. Характеристики радіоканалу залежать від середовища передачі радіохвиль і відстані між кінцевими системами. До факторів середовища передачі відносять загасання сигналу внаслідок поширення в середовищі, проходження через поглинаючі предмети, взаємодії з відбитими електромагнітними хвилями, а також хвилями, що виходять від інших джерел випромінювання.

Супутникові радіоканали

Супутник зв'язку організовує взаємодію між двома або більше наземними прийомопередавачами. Він приймає сигнали одного частотного діапазону, проводить їх регенерацію за допомогою повторювача, а потім передає сигнали в іншому частотному діапазоні. Існують два типи супутників: геостаціонарні і низькоорбітальні.

Відмінною рисою геостаціонарних супутників є те, що вони не міняють свого положення щодо заданої точки земної поверхні. Це досягнуто шляхом поміщення супутника на орбіту, віддалену приблизно на 36000 км від земної поверхні. Значну відстань, яку потрібно долати сигналу, обумовлює велику затримку його поширення, 250 мс. Проте супутникові канали, за допомогою яких не складає ніяких труднощів досягти швидкостей передачі в сотні мегабіт в секунду, активно використовуються в телефонії та Інтернеті.

Низькоорбітальні супутники знаходяться значно ближче до земної поверхні, ніж геостаціонарні, і обертаються навколо неї, подібно місяцю. Для постійного

покриття певних ділянок земної поверхні доводиться розміщувати на орбіті кілька супутників. В даний час є чимале число проектів низькоорбітальних систем зв'язку. Зокрема, в майбутньому передбачається використання подібних систем для передачі даних в Інтернеті.

1.8 Інтернет провайдери і магістралі Інтернету

У загальнодоступному Інтернеті мережі доступу з'єднуються з іншими мережами за допомогою ієрархії мереж Інтернет-провайдерів, зображеної на рис. 1.6. Знизу ієрархії перебувають мережі резидентних Інтернет-провайдерів, до яких звичайно підключаються кінцеві системи. Верхня частина ієрархії представлена мережами так званих Інтернет-провайдерів першої ланки. З одного боку, мережі цих Інтернет-провайдерів володіють типовими рисами комп'ютерних мереж - наявністю маршрутизаторів і зв'язків з іншими мережами. З іншого боку, мережі Інтернет-провайдерів першої ланки мають свою специфіку. По-перше, їх лінії зв'язку зазвичай забезпечують швидкість передачі не нижче 622 Мбіт/с, а іноді 2,5-10 Гбіт/с. По-друге, маршрутизатори мереж Інтернет-провайдерів першої ланки повинні функціонувати з гранично високою швидкістю для того, щоб не викликати затримок пакетів. По-третє, всі мережі Інтернет-провайдерів першої ланки безпосередньо сполучені між собою. По-четверте, до кожної мережі Інтернет-провайдера першої ланки підключено велику кількість мереж Інтернет-провайдерів другої ланки та інших комп'ютерних мереж. На-кінець, по-п'яте, район мережевого охоплення Інтернет-провайдера першої ланки є міжнародним.

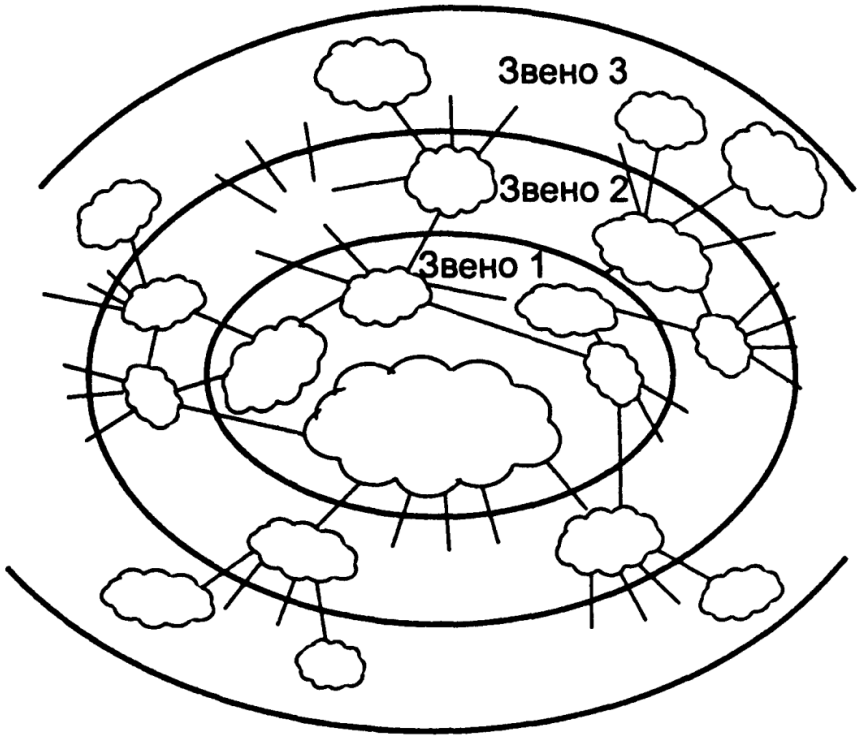

Рисунок 1.6 - Зв’язок між мережами Інтернет-провайдерів

Мережі Інтернет-провайдерів першої ланки часто називають магістралями Інтернету. Мережі Інтернет-провайдерів другої ланки, як правило, мають регіональну територію охоплення і підключаються до декількох мереж Інтернет-провайдерів першої ланки. Кажуть, що Інтернет-провайдери другої ланки є споживачами послуг Інтернет-провайдерів першої ланки. Великі компанії та установи підключають свої корпоративні мережі безпосередньо до мереж Інтернет-провайдерів другої і навіть першої ланок і вважаються споживачами їх послуг. Споживачі оплачують послуги своїх Інтернет-провайдерів.

Мережі Інтернет-провайдерів другої ланки також можуть з'єднуватися між собою і обмінюватися інформацією без участі Інтернет-провайдерів першої ланки. Нижче в ієрархії розташовані мережі Інтернет-провайдерів, які підключаються до мереж Інтернет-провайдерів другої ланки (до однієї або декількох).

На останньому щаблі ієрархії перебувають мережі доступу. Заплутують попередню концепцію, полягає в тому, що деякі Інтернет-провайдери першої ланки одночасно можуть бути Інтернет-провайдерами другої ланки, до мереж яких приєднані мережі крупних організацій та Інтернет-провайдерів нижчих ланок.

Точки, в яких мережа Інтернет-провайдера зв'язується з мережами інших Інтернет-провайдерів (розташованих вище, нижче або на одному ієрархічному рівнів), називаються точками присутності (Points of Presence, POP). Як правило, точка присутності являє собою одну або декілька груп маршрутизаторів мережі, до яких підключені маршрутизатори іншої мережі. У Інтернет-провайдерів першої ланки зазвичай є безліч точок присутності в різних географічних регіонах. Для підключення до Інтернет-провайдеру більш високої ланки споживач зазвичай орендує високошвидкісні лінії зв'язку у будь-якої телекомунікаційної компанії (що є «третьою стороною» в угоді) і з'єднує свої маршрутизатори з точкою присутності Інтернет-провайдера. Можливі з'єднання двох мереж одночасно в декількох точках присутності.

Крім сполучення через точки присутності, використовується також механізм з’єднання через точки доступу в мережу (Network Access Points, NAP), кожна з яких може належати сторонній телекомунікаційній компанії або магістральному Інтернет-провайдеру і управлятися ними. Зазвичай подібна схема використовується при підключенні мереж Інтернет-провайдерів другої ланки один до одного і до мереж Інтернет-провайдерів першої ланки, а Інтернет- провайдери першої ланки частіше воліють сполучати свої мережі між собою через точки присутності. Оскільки в точках доступу необхідно забезпечувати комутацію і передачу величезних обсягів інформації в одиницю часу, їх реалізація вельми непроста технологічно.

Топологія Інтернету є складною і складається з декількох десятків мереж Інтернет-провайдерів першого та другого ланок і сотень мереж менш великих Інтернет-провайдерів регіонального і локального масштабу. Останні підключаються до перших, які, в свою чергу, сполучені між собою за допомогою точок присутності або точок доступу.

1.9 Стек протоколів мережі Інтернету

Комунікаційна модель Інтернету складається з п'яти рівнів: фізичного, канального, мережевого, транспортного і прикладного. Замість термінів «одиниця обміну мережевого рівня »,« одиниця обміну канального рівня »і т. д. ми будемо використовувати спеціальні імена. Одиниці обміну канального рівня ми назвемо кадрами, одиниці обміну мережевого рівня - дейтаграммами, одиниці обміну транспортного рівня - сегментами, а одиниці обміну прикладного рівня - повідомленнями. Для одиниць обміну фізичного рівня зазвичай не передбачається спеціального імені. Комунікаційна модель Інтернету та одиниці обміну її рівнів зображені на рис. 1.7.

Підтримка протоколів може бути апаратною, програмною або змішаною. Протоколи прикладного рівня, такі як HTTP і SMTP, а також протоколи транспортного рівня практично завжди підтримуються програмно. Навпроти, протоколи фізичного і канального рівнів, тісно пов'язані з середовищем передачі даних, підтримується апаратно мережевою картою. Мережевий рівень, що знаходиться в центрі комунікаційної моделі, може підтримуватися як апаратно, так і програмно.

Далі дано характеристики кожного з п'яти рівнів комунікаційної моделі Інтернету.

Прикладний рівень

Прикладний рівень, як випливає з його назви, призначений для підтримки

мережевих програм. Є безліч протоколів прикладного рівня, з яких найбільш важливими є HTTP (для подорожей по веб-сторінкам), SMTP (для електронної пошти) і FTP (для обміну файлами). Розробка власного протоколу прикладного рівня не становить особливих труднощів.

Рисунок 1.7 - Стек протоколів Інтернету та одиниці обміну різних рівнів

Транспортний рівень

Головна функція транспортного рівня полягає у передачі повідомлень прикладного рівня між клієнтом і сервером. В Інтернеті існують два транспортні протоколи: TCP і UDP. Протокол TCP забезпечує передачу з встановленням логічного з'єднання, тобто надійну передачу з контролем переповнення. Крім того, TCP виробляє розбиття довгих повідомлень на більш короткі і контролює перевантаження. Контроль перевантаження зводиться до примусового зниження швидкості передачі кінцевої системи при високому завантаженні мережі. Протокол UDP забезпечує передачу повідомлень без встановлення логічного з'єднання, тобто не надійний вид зв'язку, де допускаються спотворення і втрати даних

Мережевий рівень

Мережевий рівень забезпечує передачу дейтаграм між двома хостами і базується на двох основних протоколах. Перший протокол визначає поля дейтаграми та інтерпретацію їх вмісту маршрутизаторами і кінцевими системами. Цей протокол є єдиним протоколом мережевого рівня в Інтернеті і називається IP. Другим протоколом є один з численних протоколів маршрутизації, призначених для визначення шляхів дейтаграм від відправника до адресата. Число протоколів маршрутизації величезне. Інтернет є мережею мереж, а кожна мережа підтримує

власний протокол маршрутизації, і визначається адміністратором мережі. Незважаючи на функціональні відмінності між протоколом IP і протоколами маршрутизації, а також на широке розмаїття останніх, їх зазвичай об'єднують під загальним іменем IP, підкреслюючи цим їх сполучну роль в організації глобальної мережі.

Протокол транспортного рівня (TCP або UDP) передає сегмент і адреса призначення протоколу IP мережного рівня, а протокол IP мережного рівня доставляє сегмент кінцевому хосту і передає його назад транспортному рівню.

Канальний рівень

Мережевий рівень забезпечує передачу пакета через серію маршрутизаторів між кінцевими системами. Для переміщення пакета (дейтаграми) від одного вузла до іншого мережевий рівень вдається до службам канального рівня. Таким чином, основна функція канального рівня полягає у передачі дейтаграм між вузлами на маршруті.

Канальний рівень використовує спеціальний протокол, орієнтований на використану лінію зв'язку. Іноді протоколи канального рівня забезпечують надійну передачу між вузлами. Відмінність надійної передачі на транспортному і канальному рівнях: протокол TCP забезпечує надійність на всьому шляху проходження повідомлення, а протокол канального рівня – лише між парою вузлів. До протоколів канального рівня відносяться Ethernet і PPP; іноді аналогічні функції несуть технології асинхронної передачі даних (ATM) і ретрансляції кадрів. Оскільки шлях від відправника до адресата звичайно складаються з ланцюжка різнорідних ліній зв'язку, передача дейтаграми може здійснюватися різними канальними протоколами.

Фізичний рівень

Якщо призначенням канального рівня є передача кадрів між сусідніми

вузлами мережі, то фізичний рівень забезпечує передачу між вузлами окремих бітів інформації. Протоколи фізичного рівня також безпосередньо залежать від використаної лінії зв'язку (мідної витої пари, одномодового оптоволокна і т. п.). Технологія Ethernet підтримує безліч протоколів фізичного

рівня, призначених для підтримки кручений пари, коаксіального кабелю, оптоволоконного кабелю і деяких інших видів ліній. У кожній з ліній зв'язку механізми передачі біта різні.

2 МЕТОДИ УПРАВЛІННЯ ІНТЕРНЕТ КАНАЛОМ

2.1 Класичні елементи системи управління трафіком в комп'ютерній мережі

Обмеження вихідного трафіку (shaping) - це механізм, за допомогою якого пакети затримуються перед передачею з тим, щоб швидкість передачі відповідала бажаної. Це один з найбільш часто використовуваних механізмів управління трафіком. Як побічний ефект, даний механізм дозволяє згладжувати вибухонебезпечне трафік.

Планування (scheduling) - це механізм, який дозволяє упорядковувати або змінювати порядок об'єкти між входом і виходом конкретної черги. Прикладами планувальника можуть послужити алгоритми FIFO, SQF, WRR та інші.

Класифікація (classifying) - механізм, що розділяє пакети для різної обробки, можливо в різні черги. У процесі прийому, маршрутизації і передачі пакетів, мережеве пристрій може по-різному їх класифікувати. Це може бути маркування, яке зазвичай відбувається на кордоні мережі з єдиним адмініструванням, або ж може виконуватися індивідуально на кожному проміжному вузлі.

Обмеження вхідного трафіку (policing) - черговий елемент системи якості обслуговування, що обмежує трафік. Цей механізм приймає пакети до певної швидкості, а над частиною трафіку перевищила заданий поріг виконується певна дія. Наприклад, можна знищувати трафік, або перекласифікувати його. Не дивлячись на те, що в даному випадку теж використовується концепція буфера токенів, він не підтримує можливість затримки пакетів на відміну від механізму обмеження вихідного трафіку

Знищення (dropping) - механізм, що знищує дані. Наприклад, він використовується при переповненні буфера даних обмежувача вихідного трафіку

Знищення (dropping) - механізм, що знищує дані. Наприклад, він використовується при переповненні буфера даних обмежувача вихідного трафіку.

Маркування (marking) - механізм зміни пакета. Зверніть увагу, що це не fwmark. Цілі MARK і - mark утиліт iptables і ipchains відповідно, модифікують метадані пакета, а не сам пакет.

2.2 Вирішення проблеми безпечного використання ресурсів в мережі Інтернет

Проблему безпечного і продуктивного використання Інтернет ресурсів можна вирішити двома способами.

Перший - радикальне заборона використання Інтернету без необхідності. Якщо прийняти принцип «заборонено все, що явно не дозволено», користувачам дозволяється доступ тільки до строго певних сайтів. Другий спосіб - більш гнучкий, він дозволяє користувачам діяти за принципом «дозволено все, що не заборонено». У цьому випадку співробітник може вільно користуватися ресурсами Інтернету, проте його дії перебувають під контролем. Це означає, що якщо користувач виконає дії, що суперечать політиці безпеки, це буде виявлено і попереджено.

В даний час «радикальний» спосіб досі знаходить застосування. Він використовується, в першу чергу, організаціями, в котрих циркулює інформація з грифом «секретної ». До таких організацій належать різні науково-дослідні інститути,військові організації, державні органи і спеціальні служби. У таких

«секретних» організаціях існують інструкції та документи, які суворо регламентують поведінку користувачів, пов'язані з отриманням інформації і її передачею за межі організації. А це значно полегшує діяльність контролюючих служб щодо забезпечення потрібного рівня захисту.

Інший приклад «радикального» способу - застосування в компаніях так званих Інтернет-кіосків, коли користувачам надається доступ до Інтернет ресурсів через виділені термінали. Як правило, в цьому випадку дії користувачів суворо регламентується, а трафік, що проходить через цей термінал, контролюється спеціальними засобами.

Більшість же комерційних організацій і компаній віддають перевагу більш гнучкий спосіб регламентації спілкування із зовнішнім світом. Далі розглядається саме цей спосіб, оскільки саме при його застосуванні виникають суттєві проблеми. І полягають вони тому, що практично неможливо однозначно визначити, до якої інформації слід забороняти доступ.

Щоб забезпечити гнучкий контроль користування Інтернет ресурсів, необхідно ввести в компанії відповідну політику використання ресурсів. Ця політика може реалізовуватися як вручну, так і автоматично. Ручна реалізація означає, що в організації є спеціальний штат співробітників, які в режимі реального часу або по журналах маршрутизаторів, проксі-серверів або міжмережевих екранів ведуть моніторинг активності користувачів. Такий моніторинг є проблематичним, оскільки вимагає великих трудовитрат. Крім того, він вимагає не тільки відстеження активності користувача, але і категоризації сайтів, які відвідують користувачі, а провести роботу по категоризації сайтів силами співробітників IT - підрозділів окремої компанії практично неможливо.

Щоб уникнути описаних вище проблем і забезпечити гнучкий контроль використанням Інтернет ресурсів, компанія повинна дати адміністратору інструмент для реалізації політики використання ресурсів компанії. Цій меті служить так звана контент фільтрація. Її суть полягає в декомпозиції об'єктів інформаційного обміну на компоненти, аналізу вмісту цих компонентів, визначення відповідності їх параметрів прийнятій в компанії політиці використання Інтернет ресурсів та здійсненні визначених дій за результатами такого аналізу. У разі фільтрації веб-трафіку під об'єктами інформаційного обміну мають на увазі веб-запити, вміст веб-сторінок, файли які передаються за запитом користувача і т.д.

Об'єктивна потреба викликала безліч програмних продуктів, передбачених для контролю вмісту інформаційного обміну. Всі вони в тій чи іншій ступені виконують покладену на них завдання. Однак необхідно мати на увазі, що ніяка автоматична система не дає 100% гарантії безпеки без діяльної участі людини в процесі фільтрації. Будь-яка технологія - тільки додатковий інструмент у руках

пекло адміністратора. І від того, наскільки система адекватна завданням, які ставить перед собою адміністратор, залежатиме, чи вдасться знизити рівень ризиків, пов'язаних з використанням Інтернету.

2.3 Засоби контролю використання ресурсів в мережі Інтернет

У цьому розділі коротко описані наявні на ринку засоби фільтрації веб-трафіку. Ці засоби можна умовно розділити за типами і методам фільтрації. В даний час відомо три типи засобів, здатних в тій чи іншій мірі забезпечити контроль використання Інтернет ресурсів на корпоративному рівні.

До першого типу відносяться міжмережеві екрани, проксі-сервери, маршрутизатори і подібні їм засоби фільтрації. Другий тип - це сучасні антивірусні програми, які володіють базовими можливостями тематичної фільтрації. До третього типу відносяться спеціалізовані засоби, розроблені безпосередньо для контролю використання Інтернет ресурсів.

Кожен тип засобів контролю використання Інтернет ресурсів призначений для фільтрації на різних рівнях мережевої ієрархії. Засоби першого і другого типів напряму не призначені для контролю вмісту інформаційного обміну по каналах Інтернету. Так, міжмережеві екрани, проксі-сервери і маршрутизатори здійснюють фільтрацію трафіку на мережевому і транспортному рівнях. Спеціалізовані засоби контентної фільтрації здійснюють її на прикладному рівні. Якщо говорити про антивірусні програми, то з засобами третього типу їх об'єднує саме здатність здійснювати деякі базові функції тематичної фільтрації. Загалом, засоби першого і другого типів можна вважати достатньо ефективними для компаній, в яких контроль вмісту інформаційного обміну не є першочерговим завданням (наприклад, там, де конфіденційна інформація не циркулює в корпоративної мережі). В організаціях, де активно використовується Інтернет і при цьому актуальне завдання контролю доступу користувачів до

Інтернет ресурсів і захисту від витоку конфіденційної інформації, застосування таких засобів недостатньо. Це обумовлено наступними недоліками засобів першого і другого типів:

- обмежена кількість параметрів, за яким можлива фільтрація трафіку;

- обмежені можливості по фільтрації тексту в запитах користувачів і завантажених сторінках;

- неможливість декомпозиції об'єктів інформаційного обміну, отже, відсутність більш глибокої фільтрації (наприклад, тип файлів визначається за декларативним, а значить, не обов'язково реальному розширення, при цьому відсутні засоби для визначення реального типу файлів за сигнатурою);

- відсутність необхідної гнучкості при реалізації політики використання Інтернет ресурсів;

- відсутність функціональності, яка забезпечує контентну фільтрацію. Наприклад, не всі міжмережеві екрани і проксі-сервери підтримують контроль переданих даних на рівні команд протоколів.

Спеціалізовані програмні засоби контролю використання Інтернет ресурсів призначені для перевірки даних що передаються на відповідність тих чи інших умов інформаційного обміну і виконання відповідних дій за підсумками перевірки. Програмне забезпечення цього типу можна розділити за місцем використання на клієнтські і серверні. Клієнтське програмне забезпечення працює на кожній робочій станції, де потрібно забезпечити контроль використання Інтернет ресурсів.

Застосування клієнтського програмного забезпечення в організаціях неефективно, виключаючи окремі специфічні випадки. Клієнтське програмне забезпечення постійно працює на комп'ютері користувача і ненадійно з точки зору безпеки, оскільки потенційно його конфігурація і налаштування можуть бути спотворені (підроблені) досвідченим користувачем або спеціалізованим шкідливим кодом. Крім того, таке програмне забезпечення вимагає великих трудовитрат на налагодження і супровід, оскільки ці операції здійснюються для кожної робочої станції окремо.

В корпоративному секторі найбільш ефективне застосування програмного забезпечення, яке функціонує на виділеному сервері, оскільки воно розроблено спеціально для використання в корпоративних мережах і найбільш повноцінно відповідає вимогам по функціональності і безпеки. До таких вимог відносяться:

- централізоване розміщення (забезпечує зручність встановлення, налаштування та впровадження);

- розміщення на сервері і обмеження доступу до налаштувань ПО (забезпечує безпеку, оскільки виключена можливість зміни налаштувань користувачами).

2.4 Класифікація корпоративних засобів контролю використання ресурсів мережі Інтернет

Далі розглядаються серверні (а значить, корпоративні) засоби фільтрації веб-трафіка. Говорячи про їхню класифікацію необхідно виділити основні ознаки, за яким вони розрізняються:

- використовувані методи фільтрації;

- місце розміщення в інфраструктурі мережі та спосіб впливу на трафік (технології pass by і pass through);

- можливість вибору режиму функціонування - standalone (автономні) і integrated (вбудовуються);

- глибина політики фільтрації;

- підхід до оповіщення про дії користучів;

- можливість генерації різних видів звітів за даними фільтрації.

Коли говорять про класифікацію по методам фільтрації, класи визначаються набором параметрів, за якими проводиться перевірка вмісту інформаційного обміну. Системи використовують різні набори перевірок в залежності від покладених на засоби фільтрації завдань.

Найбільш типові й поширені системи, в яких основним способом фільтрації веб-трафіку є перевірка адрес Інтернет-ресурсів (URL). Звичайно, ці системи застосовують і інші способи (фільтрація по командам протоколів, іменами

користувачів, IP? Адресами робочих станцій і типами файлів), проте саме результат перевірки адреси сайту служить підставою для рішення про блокування сайту. Такі системи, як правило, фільтрують тільки запити користувачів, не перевіряючи файли сайтів на наявність забороненого політикою безпеки вмісту.

Існують системи, в яких реалізований комплексний підхід до фільтрації трафіку. В них перевірки рівноцінні за значимістю, а склад набору перевірок визначається завданнями, покладеними на систему контролю. Відмінністю таких систем є те, що вони мають найбільш широкий набір перевірок, серед яких однією з найважливіших є перевірка вмісту тексту запитів і сайтів.

Існуючі системи розрізняються за місцем розміщення в корпоративній мережі. Їх можна розділити на дві великі групи - працюючих як проксі-сервери (технологія pass through) і працюють як аналізатори пакетів (packet sniffer), що перехоплюють пакети потрібних протоколів і перевіряють їх на наявність забороненого вмісту (технологія pass by).

Системи, що використовують технологію pass through, працюють як звичайні HTTP-проксі. Як правило, їх встановлюють між клієнтськими місцями і зовнішнім проксі-сервером або брандмауером. Клієнтські місця повинні бути налаштовані так, щоб вони використовували потрібний проксі-сервер.

До переваг систем, що використовують технологію pass by, можна віднести їх прозорості для користувача і можливість встановлення без переналаштування клієнтських місць. Недоліком є те, що вони не завжди можуть справлятися з великим потоком даних, які передаються по контрольованих протоколах.

Говорячи про режими функціонування системи, розрізняють автономні (standalone) і вбудовані рішення (integrated). Автономні системи, що встановлюються в розрив і працюють незалежно від інших підсистем, більш надійні з точки зору якості фільтрації. В них можливість помилок при фільтрації знижена до мінімуму за рахунок повного контролю над переданими даними. Зазвичай такі рішення використовуються спільно із зовнішніми проксі- серверами з метою збільшення швидкості доступу за рахунок кешування данних.

Деякі компанії поставляють програмне забезпечення для контролю використання Інтернет ресурсів у вигляді вбудованих модулів до існуючих проксі-серверів або міжмережевих екранів. Досить часто модулі фільтрації створюються для широко використовуваних проксі-сервера Microsoft ISA.

Крім того, деякі системи фільтрації поставляються у вигляді окремих серверів, а доступ до них здійснюється по протоколу ICAP, який є стандартом для контролю та модифікації запитів і відповідей, що проходять через проксі- сервер, що підтримує даний протокол. Досить часто у вигляді ICAP-серверів реалізуються антивірусні комплекси (Symantec Antivirus, Dr.Web, TrendMicro). До переваг продуктів, реалізованих у вигляді таких серверів, можна віднести те, що вони можуть взаємодіяти з різними типами проксі-серверів і міжмережевих екранів, які реалізують функцію клієнта ICAP, а також те, що сервери можуть виконувати перевірку тільки певних типів даних. Недоліком цього підходу слід вважати те, що при використанні протоколу ICAP сервер фільтрації може не мати повного контролю над переданими даними.

2.5 Перевірка адрес Інтернет ресурсів

Один з основних способів фільтрації веб-трафіку - перевірка адрес Інтернет ресурсів. Мова йде про фільтрування по URL, який передбачає собою повний шлях до конкретного документу або розділу на сервері або комп’ютері, підключеному до Інтернету. Використовуючи даний спосіб фільтрації, можна заборонити доступ як до всього сайту або домену, так і до його частини або окремій сторінці. Рекомендується вибирати засоби фільтрації, які можуть забезпечити блокування доступу і в тому випадку, коли користувач замість повного URL вказує тільки IP адресу сервера або комп’ютера в мережі.

Системи контролю використання Інтернет ресурсів, які застосовують даний спосіб фільтрації, використовують спеціальний сервіс по категоризації сайтів.

2.6 Керування каналом мережі Інтернет на основі ОС Linux

Що може запропонувати нам Linux, ось далеко неповний список з того, що може запропонувати нам операційна система Linux:

- Управляти пропускною здатністю на окремих комп'ютерах.
- Управляти пропускною здатністю до окремих комп'ютерів.
- Допоможе "роздати" пропускну здатність по-справедливості.
- Захистити вашу мережу від DDoS-атак.
- Запобігти нападу з вашої мережі на сервери в Інтернет.
- Розпаралелити кілька серверів, з метою рівномірного розподілу навантаження.
- Обмежити доступ до ваших комп'ютерів.
- Обмежити доступ ваших користувачів до інших вузлів мережі.
- Виконувати маршрутизацію на основі UID, MAC-адрес, що виходять IP-адрес, номерів портів, типу обслуговування, часу доби і вмісту.

На сьогоднішній день ці додаткові можливості не отримали широкого розповсюдження. На те є ряд причин: хоча наявна документація досить докладна, вона майже не містить практичних рекомендацій. А питання управління трафіком взагалі не освітлені.

Більшість дистрибутивів Linux, як і більшість ОС UNIX, нині використовують досить древні утиліти arp, ifconfig і route. Поки що ці інструменти працюють досить адекватно, але іноді, на ядрах Linux версії 2.2 і вище, вони можуть поводитися досить несподівано. Мережева підсистема, в ядрах 2.2 і вище, була повністю переписана. Новий мережевий код дав збільшення продуктивності і більш високі експлуатаційні характеристики, що робить Linux ще привабливішим на ринку операційних систем.

Фактично, реалізація мережевої підсистеми в Linux, що виконує класифікацію, маршрутизацію і фільтрацію, виявилася навіть повнішою, ніж в спеціалізованих маршрутизаторах, міжмережевих екранах і інших облаштуваннях управління трафіком.

У міру появи нових розробок, вони «нашаровувалися» поверх існуючих реалізацій в існуючих операційних системах. Це постійне нашарування привело до того, що код, вирішальний завдання управління мережевим трафіком, часом проявляв дуже дивну поведінку.

Наново переписана, реалізація мережевої підсистеми дозволила досягти таких характеристик, які раніше були просто недоступні.

Linux має досить складну систему управління пропускною спроможністю, названою Traffic Control (Управління Трафіком). Вона підтримує різні методи класифікації, ділення по пріоритетах, спільного використання, і обмеження трафіку, що як входить, так і витікаю чого.

Якщо наш маршрутизатор обслуговує складну мережу, то треба задовольняти потреби різних людей, обслуговування яких, ймовірно, повинне відрізнятися. База політик маршрутизації дозволяє реалізувати це за допомогою набору таблиць маршрутизації.

2.6.1 Політики маршрутизації в мережі Інтернет

Якщо ви хочете використати цю можливість, переконаєтеся що ядро зібране з підтримкою «IP : advanced router» і «IP: policy routing».

Коли ядру необхідно вибрати маршрут, воно визначає відповідно до якої таблиці це треба робити. По-умовчанню, визначені три таблиці. Стара утиліта route змінює таблиці main і local, як і утиліта ip.

Якщо ми хочемо зробити щось цікаве, то треба задати правила, що використовують різні таблиці маршрутизації. Це дозволить нам перевизначити загальносистемну таблицю маршрутизації.

2.6.2 Маршрутизація через декілька каналів/провайдерів

Перше питання полягає в тому, як організувати маршрутизацію так, щоб відповіді на запити, що приходять через певного провайдера, скажемо провайдера 1, йшли через того ж провайдера.

Друге питання полягає у балансуванні навантаження між двома провайдерами. Це не складно, якщо у нас вже налагоджений роздільний доступ, описаний в попередньому розділі.

Замість вибору одного з провайдерів в якості маршруту за замовчуванням, ви настроюєте т.з. multipath маршрут. У стандартному ядрі це забезпечить балансування навантаження між двома провайдерами. Результатом команди буде попеременний вибір маршруту за замовчуванням. Ви можете змінити параметр weight, так щоб один з провайдерів отримував велике навантаження.

Зверніть увагу, що балансування не буде ідеальним, оскільки вона ґрунтується на маршрутах, а маршрути кешуются. Це означає, що маршрути до часто відвідуваних сайтів не проходитимуть через різних провайдерів.

2.6.3 Дисципліни обробки черг для управління пропускною здатністю комп’ютерної мережі

Організація черги визначає спосіб відсилання даних. Важливо розуміти, що ми можемо керувати лише швидкістю передачі даних, що відправляються. У тому вигляді, в якому зараз існує Internet, ми не можемо контролювати обсяг вхідного трафіку. Це щось на зразок поштової скриньки. Немає ніякого способу впливати на те, який обсяг пошти приходить до вас, хіба що спілкуючись з кожним респондентом.

Продовжуючи нашу аналогію, це можна порівняти з викиданням половини вашої пошти в надії на те, що люди перестануть вам писати. Різниця лише в тому, що у випадку з Internet цей прийом спрацьовує

Якщо у вас є маршрутизатор і ви хочете обмежити швидкість завантаження у внутрішній мережі, вам потрібно це робити на внутрішньому інтерфейсі маршрутизатора, з якого дані передаються ваших комп'ютерів.

Крім того, ви повинні бути впевнені, що контролюєте вузьке місце з'єднання. Так, якщо у вас є 100-мегабітний мережева карта і маршрутизатор із з'єднанням в 256 Кбіт / сек, ви повинні переконатися, що не посилаєте даних більше, ніж ваш маршрутизатор може передати. Інакше канал буде контролювати маршрутизатор і саме він буде обмежувати доступну пропускну здатність. Нам потрібно, так би мовити, «володіти чергою» і бути самим повільним ланкою. На щастя це легко реалізується.

Прості безкласові дисципліни обробки черги

Як вже говорилося, дисципліни обробки черги визначають спосіб передачі даних. Безкласові дисципліни, загалом, отримують дані, змінювють порядок, вносять затримку або знищують їх.

Вони можуть використовуватися для обмеження пропускної здатності інтерфейсу цілком, без будь-якого поділу за класами. Вкрай важливо, щоб ви зрозуміли призначення цього типу черг перед тим, як ми перейдемо до класових дисциплін!

Найбільш поширеною дисципліною є pfifo_fast - вона використовується за замовчуванням. Кожна з дисциплін має свої переваги і недоліки. Не всі з них досконально протестовані.

pfifo_fast – ця дисципліна працює, як видно з назви, за принципом «першим прийшов, першим пішов» (First In, First Out). Це означає, що жоден пакет не отримує спеціальної обробки. Однак це не зовсім так. Дана чергу має три, так званих, «смуги». У кожній «смузі» пакети обробляються за принципом FIFO. Але смуга 1 не буде обслуговуватися до тих пір, поки є пакети в смузі 0. Аналогічно, поки є пакети в смузі 1, не обробляється смуга 2.

Ядро враховує значення поля пакета Type of Service, і направляє пакети з встановленим прапором мінімальна затримка в смугу 0.

Ми не можемо конфігурувати pfifo_fast, оскільки її параметри жорстко «зашиті».

Token Bucket Filter (TBF) проста дисципліна черги, яка передає надходять пакети зі швидкістю не перевищує адміністративно заданий поріг, але з можливістю перевищують його коротких сплесків.

TBF дуже точна дисципліна, при цьому вона не створює серйозних навантажень на мережу і процесор. Якщо вам потрібно просто обмежити швидкість на інтерфейсі, то це перший кандидат на користування. Реалізована TBF у вигляді буфера, який постійно заповнюють токенами з заданою швидкістю. Найбільш важливим параметром буфера є його розмір, що визначає кількість збережених токенів.

Кожен прибуваючий токен співставляється з одним пакетом даних з черги після чого видаляється. Зв'язавши цей алгоритм з двома потоками - токенів і даних, одержимо три можливих ситуації:

- Дані прибувають зі швидкістю рівною швидкості входять токенів. У цьому випадку кожен пакет має відповідний токен і проходить чергу без затримки.

- Дані прибувають зі швидкістю меншою швидкості надходження токенів. У цьому випадку лише частина існуючих токенів буде знищуватися, тому вони стануть накопичуватися до розміру буфера. Далі, накопичені токени можуть використовуватися при сплесках, для передачі даних зі швидкістю яка перевищує швидкість надходження токенів.

- Дані прибувають швидше, ніж токени. Це означає, що в буфері скоро не останеться токенів, що змусить дисципліну призупинити передачу даних. Ця ситуація називається «перевищенням». Якщо пакети продовжують надходити, вони починають знищуватися.

Остання ситуація дуже важлива, оскільки дозволяє адміністративно обмежувати доступну смугу пропускання.

Накопичені токени дозволяють пропускати короткі сплески, але при тривалому перевищенні пакети будуть затримуватися, а в крайньому випадку -

знищуватися. Врахуйте, що в реальній реалізації дисципліни, токени відповідають байтам, а не пакетам.

Не дивлячись на те, що нам ймовірно нічого не доведеться змінювати, дисципліна TBF має певні параметри. У першу чергу це:

- Limit - це кількість байт, які можуть бути поміщені в чергу очікування токенів. Цю ж величину можна задати параметром latency, який визначає максимальний «вік» пакета в черзі TBF. В останньому випадку, до уваги береться розмір буфера, швидкість і, якщо задана, пікова швидкість (peakrate).

- Розмір буфера в байтах. Максимальна кількість байт, для яких токени можуть бути доступні миттєво. В цілому, чим більше гранична швидкість, тим більше повинен бути розмір буфера. Наприклад, для обмеження на швидкості 10 Мбіт / с на платформі Intel, вам потрібен буфер розміром як мінімум 10 Кбайт, щоб досягти заявленої швидкості!

- Якщо буфер занадто малий, пакети можуть знищуватися. Це пов'язано з тим, що кожен тік таймера буде генеруватися більше токенів, ніж може поміститися у вашому буфері.

- Пакет нульового розміру все одно використовує смугу пропускання. У мережах ethernet, будь-який пакет має розмір не менше 64 байт. MPU задає мінімальну кількість токенів для пакета.

- Обмеження швидкості. Якщо буфер заповнений токенами, то вхідні пакети будуть проходити чергу без всяких затримок.

- Якщо на момент надходження пакету є вільні токени, пакет пройде чергу без будь-яких затримок. Так би мовити, зі швидкістю світла. Можливо, це не зовсім те, чого ви хочете, особливо якщо ви використовуєте великий буфер.

- Параметр peakrate задає швидкість, з якою елемент може проходити чергу. Згідно теорії, це досягається організацією достатньої затримки між проходять пакетами.

- Очевидно, що максимальне значення peakrate, рівне 1 Мбіт / сек, накладало б сильне обмеження на область застосування цієї дисципліни. Однак, завдання

великих значень peakrate можливо. Досягається це за рахунок проходження за один інтервал часу більше одного пакета даних.

- За замовчуванням, значення mtu одно одному пакету, тобто за раз проходить тільки один пакет.

Stochastic Fairness Queueing (SFQ) - проста реалізація сімейства алгоритмів справедливою очередизації. Вона не так точна, як інші дисципліни, але вимагає менше розрахунків, і при цьому порівну розподіляє доступну смугу пропускання між сеансами.

Ключовим поняттям в SFQ є потік, який приблизно відповідає сеансу або потоку TCP/UDP. Трафік поділяється на достатню кількість черг типу FIFO, по одній на кожен потік. Після цього, всі черги обробляються в циклічному порядку, тим самим забезпечуючи кожному сеансу рівні шанси на передачу даних.

Завдяки цьому досягається дуже рівна поведінка, яка не дозволяє будь-якому діалогу пригнічувати інші. SFQ називається «стохастичною» т.к. насправді для кожного сеансу чергу не формується, а трафік ділиться на обмежену кількість черг на основі хеш-алгоритму.

Через використання хешу, кілька сесій можуть потрапити в одну і ту ж чергу, що зменшує шанси на передачу кожного сеансу. Для того, щоб ця проблема не відчувалася, SFQ часто змінює алгоритм хешування, тому, якщо сесії і потраплять в одну чергу, тривати це буде лише кілька секунд.

Варто зауважити, що SFQ ефективний тільки якщо вихідний інтерфейс повністю завантажений. В іншому випадку ніякого позитивного ефекту спостерігатися не буде.

Зокрема, застосування SFQ на Ethernet-інтерфейсі до якого підключений кабельний модем або DSL-маршрутизатор абсолютно безглуздий без обмеження смуги пропускання!

SFQ в значній мірі самоконфігурірующаяся:

- Perturb Інтервал зміни алгоритму хешування. Хорошим значенням є 10 секунд.

- Quantum кількість байт виведених з черги за один раз. За-замовчуванням дорівнює 1 пакету максимально можливого розміру (MTU).

- limit загальна кількість пакетів, які можуть бути поміщені в чергу SFQ (наступні пакети будуть знищуватися).

Класові дисципліни обробки черг

Класові дисципліни широко використовуються у випадках, коли той або інший вид трафіку необхідно обробляти по різному. Прикладом класової дисципліни може служити CBQ - Class Based Queueing (дисципліна обробки черг на основі класів). Вона настільки широко відома, що багато хто ідентифікує поняття «Дисципліна Обробки Черг» з назвою CBQ, проте це далеко не так.

CBQ - один із старих алгоритмів і крім того - один з найскладніших. На жаль він може далеко не все. Це може виявитися несподіванкою для тих, хто свято вірить в те, що якщо яка-небудь досить складна технологія поширюється без документації, то це краща технологія з наявних варіантів.

Коли трафік передається на обробку класовій дисципліні, він має бути віднесений до одного з класів(класифікований). Визначення приналежності пакету до того або іншого класу виконується фільтрами. Дуже важливо розуміти, що саме фільтри викликаються з дисципліни, а не навпаки!

Фільтри, приєднані до дисципліни, повертають результат класифікації (грубо кажучи клас пакету), після чого пакет передається в чергу, що відповідає заданому класу. Кожен з класів, у свою чергу, може складатися з підкласів і мати свій набір фільтрів, для виконання точнішої класифікації своєї долі трафіку.

Крім того, у більшості випадків класові дисципліни виконують шейпінг (формування) трафіку, з метою переупорядковування пакетів(наприклад, за допомогою SFQ) і управління швидкістю їх передачі. Це безперечно необхідно у разі перенаправлення трафіку з високошвидкісного інтерфейсу (наприклад, ethernet) на повільний(наприклад, модем).

Кожен з інтерфейсів має одну витікаючу кореневу дисципліну. За-замовчуванням це нагадує раніше дисципліну - pfifo_fast. Кожній дисципліні і кожному класу призначається унікальний дескриптор, який який може

використовуватися подальшими інструкціями для посилання на ці дисципліни і класи. Окрім витікаючої дисципліни, інтерфейс так само може мати і дисципліну, яка проводить управління трафіком.

Дескриптори дисциплін складаються з двох частин - старшого і молодшого номерів, у виді: <старший>:<молодший>. Кореневій дисципліні загальноприйнято привласнювати дескриптор '1:', що еквівалентно запису '1: 0'. Молодший номер в дескрипторі будь-якої дисципліни завжди '0'.

Старші номери дескрипторів класів завжди дублюють старший номер дескриптора свого «батька».

Дисципліна PRIO фактично ніяк не обмежує трафік, вона лише виконує його класифікацію на основі приєднаних до неї фільтрів. Ви можете розглядати дисципліну PRIO як потужнішу версію pfifo_fast, в якій кожна із смуг є окремим класом, а не простою чергою FIFO. Постановка пакету в чергу виконується дисципліною PRIO на основі фільтрів, заданих вами. За замовчуванню створюються три класи. Ці класи містять звичайні дисципліни FIFO, але вони можуть бути замінені дисциплінами будь-якого типу, які вам тільки доступні. Коли необхідно витягнути пакет з черги, то першим перевіряється клас :1. Кожен подальший клас перевіряється тільки у тому випадку, якщо в попередньому немає жодного пакету.

Ця «розкидати» трафік по пріоритетах, ґрунтуючись не лише на прапорах TOS. Ви можете так само додати інші дисципліни до зумовлених класів, що підвищить можливості управління трафіком, в порівнянні з pfifo_fast. Формально, дисципліна PRIO відноситься до розряду планувальників типу Work- Conserving.

Дисципліна CBQ. Одна з найскладніших дисциплін. Це найоб'ємніша, найнезрозуміліша і найзаплутаніша дисципліна організації черг. Це не тому, що автори алгоритму некомпетентні, а тому, що ідеологія цього алгоритму абсолютно не співпадає з ідеологією Linux.

Крім того, що ця дисципліна є класовою, вона так само може виконувати і шейпінг трафіку, правда саме ця її сторона є найслабкішим місцем. Якщо ви спробуєте обмежити 10 мегабітовий канал величиною в 1 мегабіт, то виявиться, що

з'єднання просто простоюватиме 90% усього часу. Замість визначення об'єму трафіку, CBQ вимірює час в мікросекундах між запитами і на основі отриманого часу розраховується середня завантаженість каналу. Такий алгоритм роботи не завжди дає потрібні результати. Наприклад, що якщо мережевий інтерфейс не може забезпечити повне завантаження каналу на усю його можливу ширину, із- за неякісного драйвера? Як тоді правильно визначити час простою?

Проблема стає ще гостріше, якщо вам доводиться мати справу з такими речами, як PPP через Ethernet або PPTP через TCP/IP. Ефективна пропускна спроможність в даному випадку може бути визначена як пропускна.

Обмеження пропускної спроможності в CBQ виконується за рахунок визначення проміжку часу між проходженням сусідніх пакетів середнього розміру. В процесі роботи вимірюється ефективний час простою, як експоненціальне зважене середнє по ковзаючому вікну. До речі, UNIX розраховує величину loadaverage(середня величина навантаження) аналогічним чином.

Розрахунковий час простою віднімається із зваженого середнього, в результаті виходить величина avgidle. Повністю завантажений канал має величину avgidle рівну нулю -- проміжок часу між пакетами точно співпадає з розрахунковим. У разі перевищення заданого обмеження, величина avgidle стає негативною. Якщо перевищення досягає деякого порогу, CBQ призупиняє передачу. З іншого боку, після декількох годин простою, величина avgidle може вийти занадто великою і це приведе до того, що канал «відкриється» на усю ширину. Щоб цього не відбувалося, величина avgidle обмежується числом maxidle.

Hierarchical Token Bucket CBQ занадто складна і слабо оптимізована для більшості типових ситуацій. Його підхід точніше відповідає конфігураціям, коли необхідно розподілити задану смугу пропускання між різними видами трафіку на смуги гарантованої ширини, з можливістю запозичення.

HTB працює точно так, як і CBQ, але, на відміну від останньої, принцип роботи заснований не на обчисленні часу простою, а на визначенні об'єму трафіку, що повністю відповідає назві Token Bucket Filter.

Хоча конфігурація HTB -- завдання досить складне, проте конфігурації добре масштабуються. У випадку ж з CBQ процес конфігурації стає занадто складним навіть в найпростіших випадках.

Intermediate queueing device. Облаштування IMQ не є дисципліною обробки черги, але тісно з ними пов'язано. У Linux, дисципліни обробки черг приєднуються до мережевих пристроїв і все, що поміщається в чергу пристрою, потрапляє спочатку в чергу дисципліни обробки черги. Із-за цього підходу існують два обмеження:

- Обмеження пропускної спроможності повноцінно працює тільки для витікаючого трафіку(дисципліна обробки трафіку, що входить, існує, але її можливості мізерні в порівнянні з повнокласовими дисциплінами).

- Дисципліна обробки черги обслуговує трафік тільки для одного інтерфейсу, немає ніякої можливості задати глобальні обмеження.

Облаштування IMQ намагається вирішити ці проблеми. За допомогою підсистеми фільтрації ОС Linux можна певні пакети направляти через цей псевдо-интерфейс, до якого підключаються різні дисципліни обробки черг. Таким чином, можна управляти смугою пропускання, трафіку, що як входить, так і загального.

2.6.4 GRE та інші тунелі

У ОС Linux підтримуються 3 типи тунелів. Це тунелювання IP в IP, GRE тунелювання і тунелі не-ядерного рівня(як, наприклад, PPTP).

Тунелі можуть використовуватися для дуже незвичайних і цікавих речей. Також вони можуть посилити ситуацію, якщо вони конфігуровані неправильно. Не задавайте маршрут за умовчанням через тунель, якщо тільки ви точно не упевнені в тому, що робите. Ще, тунелювання збільшує навантаження на систему і мережу, тому що додаються додаткові IP- заголовки. Зазвичай, це 20 байт на пакет. Таким чином, якщо звичайний розмір пакету (MTU) в мережі дорівнює 1500 байтам, то при пересилці по тунелю, пакет може містити тільки 1480 байт. Це не обов'язково

стає проблемою, але пам'ятаєте про необхідність правильного налаштування фрагментації пакетів, якщо ви сполучаєте великі мережі.

IPSEC: безпечна передача даних протоколами IP через Інтернет

IPSEC є безпечною версією протоколу IP. Поняття «безпеки», в даному випадку, означає можливість шифрування і аутентифікації. У чистому вигляді, з технічної точки зору, «безпека» означає тільки шифрування, проте, досить легко показати, що цього недостатньо - ви можете обмінюватися шифрованими даними, але не мати при цьому гарантій, що видалена сторона саме та, яку ви чекаєте.

Шифрування, в IPSEC, виконується протоколом ESP(Encapsulating Security Payload - Інкапсульовані Захищені Дані), аутентифікація -- протоколом AH(Authentication Header -- Заголовок Аутентифікації). Ви можете конфігурувати їх обох, або один з них. І ESP, і AH спираються на Security Association(захищений віртуальний канал, або контекст безпеки). Security Association - однонапрямлене логічне з'єднання (від відправника до одержувача) між двома системами, що підтримують протокол IPSec, яке однозначно ідентифікується наступними трьома параметрами:

- індексом захищеного з'єднання (Security Parameter Index, SPI - 32-бітова константа, використовувана для ідентифікації різних SA с однаковими IP- адресою одержувача і протоколом безпеки);

- IP- адресою одержувача IP- пакетів (IP Destination Address);

- протоколом безпеки (Security Protocol - AH або ESP).

Щоб забезпечити необхідний рівень безпеки, ми повинні б передавати відомості про конфігурацію по надійних каналах. Якби нам довелося настроювати видалений хост через telnet, то будь-яка третя особа запросто могла б отримати секретні відомості, і така конфігурація буде далеко не безпечна. Крім того, як тільки секретна інформація стає відомою кому-небудь, вона перестає бути секретною. Знання секретних відомостей дасть не так багато видаленому користувачеві, але ми маємо бути абсолютно упевнені в тому, що канали зв'язку з нашими партнерами дійсно надійно захищені. Ця упевненість вимагає великої

кількості ключів, якщо у нас є 10 партнерів, то необхідно мати не менше 50 різних ключів.

Окрім проблеми, пов'язаної з необхідністю узгодження ключів, існує також необхідність в періодичній їх зміні. Якщо третя сторона зможе перехопити наш трафік, то рано чи пізно вона буде в змозі «розколоти» ключ. Це може бути відвернене за рахунок періодичної зміни ключів, але цей процес вже вимагає автоматизації.

Інша проблема полягає в тому, що при роботі з ключовою інформацією «вручну», як це описано вище, ми заздалегідь точно визначаємо алгоритми і використовувану довжину ключа, що у свою чергу вимагає тісної координації з видаленою стороною. Бажано було б мати можливість визначення ширшої політики призначення ключів, наприклад так: «Ми можемо використати алгоритми 3DES і Blowfish.

Рішення цих проблем бере на себе Протокол Обміну Ключами – IKE (Internet Key Exchange), що дозволяє обмінюватися згенерованими, автоматично і випадковим чином, ключами. Передача ключів здійснюється за допомогою асиметричної технології кодування, відповідно до зумовлених алгоритмів.

У Linux IPSEC 2.5, реалізація цих можливостей виконана у вигляді демона KAME 'racoon' IKE.

Однак, Internet, в більшості своєму, заснований на протоколі TCP / IP, а в нього є кілька властивостей, які можуть нам допомогти. TCP / IP не може дізнатися пропускної здатності мережі між двома хостами, тому він починає передавати дані все швидше і швидше (це називається «повільний старт»). Коли пакети починають губитися через перевантаження передавальної середовища, передача гальмується. Насправді все трохи складніше і розумніше, але про це пізніше.

2.7 Інші системи керування для ОС Windows

2.7.1 Lan2net Traffic Shaper

У Lan2net Traffic Shaper реалізоване унікальне для Windows систем динамічне управління завантаженням каналу, побудоване на принципах диференційованого обслуговування (differentiated services). Подібний механізм застосовується в рішеннях на базі Linux. Принцип диференційованого обслуговування полягає в застосуванні різної якості обслуговування (quality of services) для задоволення різних потреб користувачів каналу. При диференційованому обслуговуванні мережевий трафік розділяється на класи, і до кожного класу застосовується індивідуальне нормування параметрів трафіку. Це дозволяє найгнучкіше управляти завантаженням мережевого каналу. Для різних груп протоколів, IP адрес і тому подібне можна встановлювати пріоритети, задавати максимальну і мінімальну ширину каналу. В результаті, це дозволяє добитися того, що найбільш важливі сервіси або користувачі отримають максимальну ширину каналу і швидкість їх роботи з Інтернет не впаде, коли трафік з нижчим пріоритетом займатиме канал. Меню програми показано на рис. 2.1.

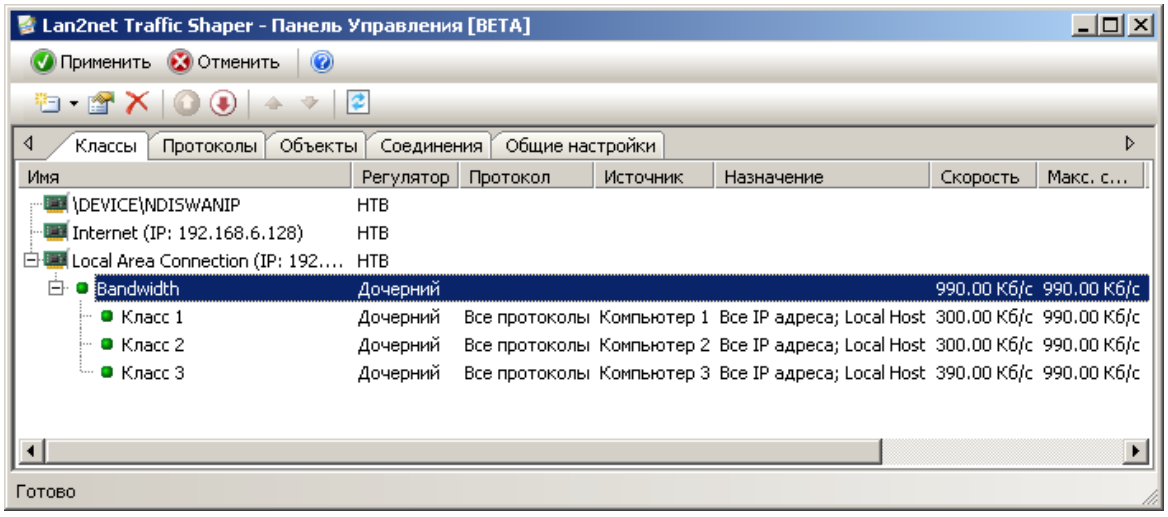

Рисунок 2.1 - Меню програми Lan2net Traffic Shaper

При диференційованому обслуговуванні мережевий трафік розділяється на класи, і до кожного класу застосовується індивідуальне нормування параметрів трафіку. Як показують дослідження, застосування подібної технології дозволяє поліпшити якість обслуговування користувачів навіть більшою мірою, чим придбання додаткової смуги пропускання у провайдера.

У загальному вигляді, якість обслуговування характеризується декількома, частенько взаємозв'язаними параметрами. Найбільш очевидний ключовий параметр тут, це - завантаження каналу, тобто, з якою швидкістю дані передаються по каналу. Для управління завантаженням каналу застосовуються різні схеми (queueing disciplines) або, більше простими словами, регулятори (throttles). Регулятори діляться на два типи: с розділенням трафіку на класи (classful) і без розділення (classless). Регулятори без розділення трафіку на класи обмежують швидкість передачі усього трафіку того, що проходить через регулятор. Регулятори з розділенням трафіку на класи розділяють трафік, що проходить через них, на класи, і обмежують смугу пропускання для кожного класу індивідуально.

2.7.2 UserGate Proxy & Firewall

UserGate Proxy & Firewall — це комплексне рішення для організації загального доступу в Інтернет з локальної мережі, обліку трафіку і захисту корпоративної мережі від зовнішніх загроз. UserGate є ефективною альтернативою дорогому програмному і апаратному забезпеченню і призначений для використання в компаніях малого і середнього бізнесу, вікно програми показано на рис. 2.2.

UserGate забезпечує комплексний захист локальної мережі, завдяки наявності двох вбудованих антивірусних модулів від провідних розробників антивірусних програм — Лабораторії Касперского і Panda Security. Антивірусні модулі проводять сканування усіх типів мережевого трафіку, включаючи поштовий, HTTP і FTP -трафік. На додаток до антивірусної перевірки в UserGate

вбудований міжмережевий екран, що забезпечує надійний захист мережі від зовнішніх атак.

UserGate використовує комплексний підхід до забезпечення безпеки локальної мережі і сучасні методи боротьби з Інтернет-загрозами, такими, як віруси, шкідливі програми і хакерські атаки.

Функції інформаційної безпеки включають:

- Захист від вірусів
- Міжмережевий екран
- Розширений драйвер NAT
- Підтримка VPN- з'єднань

Захист від вірусів

Необхідність захисту локальної мережі від різних мережевих загроз, зокрема вірусів, Інтернет-черв'яків або троянов не можна недооцінювати, оскільки простий недогляд може мати непоправні наслідки для будь-якого бізнесу. Питання «Яке антивірусне застосування вибрати»? — що найчастіше задається на форумах, присвячених безпеці в Інтернет. Компанія Entensys співпрацює з двома світовими лідерами в області розробки антивірусного ПО — Лабораторією Касперского і Panda Security — з метою надати своїм користувачам вибір антивірусного рішення для використання у складі UserGate.

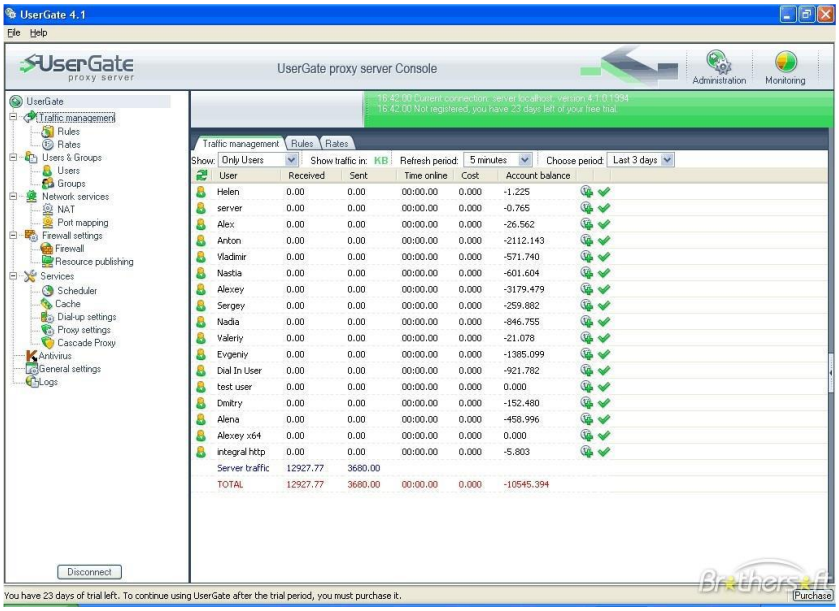

Рисунок 2.2 - Вікно програми UserGate Proxy & Firewall

Користувачі можуть за бажанням використати той або інший антивірусний модуль, або активувати обидва модулі для максимального захисту. При цьому можна комбінувати антивірусний захист UserGate із захистом файлової системи на локальних машинах за допомогою третього антивірусного рішення.

Міжмережевий екран

Міжмережевий екран (брандмауер) в UserGate дозволяє захистити локальну мережу від несанкціонованого доступу ззовні, одночасно надаючи можливість відкрити доступ до внутрішніх ресурсів, таким як поштовий, веб- або VPN- сервер в локальній мережі.

Розширений драйвер NAT

Версія UserGate 5 містить новий, розширений варіант драйвера NAT. Функція маршрутизації тепер дозволяє адміністраторові створювати локальні підмережі і настроювати обмін пакетів між ними. Наявність підтримки протоколів IP-телефонії і публікація ресурсів дозволяють використати сучасні способи комунікації і спільної роботи.

Підтримка VPN- з'єднань

UserGate підтримує передачу трафіку через протоколи PPTP і L2TP для з'єднання VPN- сервера з VPN- клієнтами локальної мережі. Крім того, можна використати публікацію мережевих ресурсів, щоб зробити VPN- сервер локальної мережі доступним видалено.

Контроль і статистика

За допомогою UserGate можна контролювати доступ в Інтернет окремих співробітників компанії і їх груп. Вбудований модуль Entensys URL Filtering дозволяє блокувати доступ до небажаних ресурсів як окремо, так і по категоріях сайтів. UserGate також дозволяє контролювати застосування, встановлені на клієнтських машинах, дозволяючи або забороняючи тому або іншому застосуванню вихід в Інтернет. Детальні статистичні звіти доступні як безпосередньо з програми, так і видалено за допомогою веб-браузеру.

UserGate дозволяє здійснювати повний контроль над використанням Інтернет-трафіку в компанії і надає адміністраторові детальну статистику. На

основі даних статистики керівництво може визначати політику доступу в Інтернет в цій компанії, яка потім реалізується за допомогою гнучкої системи правил управління трафіком в UserGate.

Контроль доступу в Інтернет включає наступні функції:

- Користувачі і групи
- URL- фільтрація трафіку
- Контроль застосувань
- Швидкість і квотування
- Контроль трафіку і гнучка система звітів
- Модуль веб-статистики
- Білінгова система

Користувачі і групи

У основі роботи UserGate лежить поняття «користувач», яке визначається як комп'ютер або група комп'ютерів, об'єднаних загальною ознакою. В якості такої ознаки може виступати IP - або MAC- адреса, пара «логін/пароль», обліковий запис в Active Directory або обліковий запис Windows. До усіх користувачів застосовуються правила розподілу трафіку, а також ведеться статистика і облік відвідування Інтернет-ресурсів.

Щоб спростити управління трафіком, адміністратор може об'єднати користувачів в групи за допомогою функції «Додати в групу». Інший спосіб угрупування користувачів полягає у використанні одного з декількох методів авторизації - наприклад, пари «логін/пароль». Адміністратор може вибрати будь-який спосіб, або обидва способи разом для ефективного управління безліччю користувачів, комп'ютерів або підмереж.

URL- фільтрація трафіку по категоріях сайтів

Нецільове використання Інтернету на робочому місці є серйозною проблемою для працедавця і робить негативний вплив на продуктивність праці співробітників, безпеку локальної мережі, і збереження конфіденційних даних. Щоб уникнути потенційних загроз такого використання, невід'ємною частиною

системи безпеки корпоративної мережі повинні стати механізми фільтрації відвідуваних веб-ресурсів.

Робота модуля фільтрації сайтів заснована на технології Entensys URL Filtering, яка також вилкористовується в GateWall DNS Filter. При цьому використовується категотризиційна база, в якій міститься 500 млн. сайтів в 82 категоріях. Адміністратор може забороняти доступ до окремих сайтів, до категорій сайтів, або до тих сайтів, адреси яких містять задані фрагменти слів. База спеціально адаптована для використання російськомовними користувачами і містить до 10 мільйонів російськомовних сайтів.

Контроль програм

Кількість програм, які так чи інакше використовують Інтернет-з'єднання для своєї роботи неухильно збільшується з кожним роком. Згідно з недавніми дослідженнями, застосування для миттєвого обміну повідомленнями(інтернет-пейджери) використовуються у більш ніж 80% організацій, і ця цифра постійно росте. Виникає необхідність контролю за діяльністю таких застосувань з метою захистити локальну мережу від зовнішніх загроз.

Контроль (фільтрація) програм — це технологія, що дозволяє обмежувати або блокувати трафік конкретних Інтернет- програм. Ця технологія має подвійне призначення: по-перше, вона дає можливість адміністраторові блокувати роботу певних Інтернет-застосувань (наприклад, ICQ або MSN), а по-друге — захищає локальну мережу від шкідливого ПО, яке може проникати в локальну мережу через такі застосування.

Обмеження трафіку і швидкості доступу

UserGate надає адміністраторові широкі можливості по контролю швидкості передачі даних між локальною мережею і Інтернетом. Встановити обмеження за швидкістю можна в модулях «Управління трафіком» і

«Управління шириною каналу». Перший модуль призначений для налаштування обмежень швидкості по окремих користувачах і групах, тоді як другою дозволяє встановлювати обмеження швидкості для конкретного

мережевого адаптера, протоколу (TCP або UDP), IP- адреси джерела або одержувача і порту.

Окрім швидкості, UserGate дозволяє також обмежувати об'єм трафіку і час перебування в мережі для користувачів і груп. При цьому, у розпорядженні адміністратора знаходиться широкий набір функцій, що дозволяє йому створювати правила, які задовольнятимуть будь-яким заданим вимогам. Наприклад, можна створити правило, яке стає активним при виконанні певних умов, таких як настання вказаного часу доби або використання певного протоколу.

Статистика відвідування Internet і система звітів

UserGate надає адміністраторові повну статистику про використання Інтернет в компанії по окремих користувачах і групах. Детальна статистика є основою для ухвалення рішень керівництвом про необхідність обмеження доступу до тих або інших ресурсів, або блокування роботи деяких Інтернет- застосувань.

Веб-статистика UserGate

Доступ до статистики UserGate можна отримати через Інтернет з будь-якої точки світу, використовуючи звичайний браузер. Інформація відображається не лише в табличному виді, але і у вигляді діаграм і графіків, що істотно полегшує сприйняття звітів. Об'єм доступної статистичної інформації залежить від рівня доступу користувача.

Білінгова система

Вбудована в UserGate білінгова система автоматично проводить розрахунок вартості роботи користувача в мережі Інтернет виходячи із заданої ціни, часу і/або об'єму трафіку. Можна встановлювати тарифи як для окремого користувача, так і для групи користувачів. Існує можливість перемикання тарифів залежно від часу доби, дня тижня або адреси сайту.

Організація доступу в Інтернет

За допомогою UserGate можете організувати доступ в Інтернет для співробітників вашої компанії через NAT або проксі-сервер, одночасно працювати через декілька Інтернет-провайдерів, а також оптимізувати споживання Інтернет-трафіку з тим, щоб уникнути навантаження на мережу і понизити витрати на

трафік. Підтримка протоколів IP- телефонії дозволяє скористатися перевагами VoIP- рішень, щоб на їх основі створити сучасну комунікаційну інфраструктуру компанії.

Доступ в Інтернет

Забезпечення доступу в Інтернет і контроль трафіку є основним завданням придбання і установки проксі-сервера в компанії. За допомогою UserGate можна організувати доступ користувачів локальної мережі в мережу Інтернет як через NAT, так і через HTTP-, FTP - і інші типи проксі-серверів. Гнучкість і різноманіття функцій UserGate дозволяють адміністраторові мережі настроїти сервер так, щоб він відповідав найсерйознішим вимогам безпеки і продуктивності.

- Забезпечення доступу в Інтернет
- Проксі-сервери для різних протоколів
- Робота з декількома провайдерами
- Управління шириною каналу
- Кешування
- Підтримка IP- телефонії

Забезпечення доступу в Інтернет

UserGate дозволяє організувати доступ в Інтернет комп'ютерів у вашій локальній мережі, використовуючи будь-який тип Інтернет-підключення, такий як DSL, ISDN, кабельне підключення, комутований доступ або WiFi. UserGate служить проміжною ланкою між Інтернет і локальною мережею, і має як зовнішню IP- адресу для роботи в Інтернет, так і один або декілька локальних IP- адрес для роботи з комп'ютерами в локальній мережі.

Оскільки при використанні UserGate увесь Інтернет-трафік проходить тільки через сервер UserGate, такі завдання як управління трафіком, перегляд статистики закачувань і захист локальної мережі від зовнішніх загроз здійснюються централізований.

Проксі-сервери для різних протоколів

UserGate може служити проксі-сервером між локальною мережею і Інтернет. Функція проксі доступна таких протоколів, як HTTP (с підтримкою HTTPs і «FTP через HTTP»), FTP, SOCKS, POP3, SMTP, SIP, і H.323. При цьому,

підтримується функція «прозорий проксі» при використанні якої немає необхідності в ручному налаштуванні браузерів користувачів. Крім того, ви можете вказати конкретний мережевий інтерфейс, на якому працюватиме проксі-сервер.

Використання HTTP- проксі може служити різним цілям, таким як прискорення відповідей на запити користувачів (кешування), і пов'язана з цим економія трафіку. Метою може бути також необхідність фільтрації Інтернет-трафіку і заборона доступу до деяких ресурсів.

Робота з декількома провайдерами

У UserGate підтримується одночасна робота з декількома Інтернет-провайдерами. Ця функція дозволяє надати доступ в Інтернет різним користувачам через різні провайдери, а також автоматично перемикати користувачів на резервне з'єднання у разі, якщо з'єднання з основним Інтернет- провайдером не працює.

Управління шириною каналу

Із зростанням числа Інтернет-з програм, і, отже, об'єму трафіку виникає необхідність оптимізувати споживання трафіку. Управління шириною каналу в UserGate є функцією, покликаною вирішити це завдання.

Кешировання

Кешировання є однією з функцій HTTP - і FTP--проксі в UserGate, яка прискорює відкриття веб-сторінок або завантаження файлів, а також дозволяє істотно економити витрати на трафік, що входить. При використанні кешировання результати веб-запитів користувачів зберігаються на локальний диск комп'ютера, на якому встановлений UserGate.

Підтримка IP- телефонії

Підтримка протоколів SIP і H.323 дозволяє використати проксі-сервер UserGate в якості VoIP- шлюзу як для програмних, так і для апаратних IP-телефонів. При використанні SIP проксі-сервера в моніторингу UserGate буде відображена уся інформація про поточний стан з'єднання(реєстрація, дзвінок,

очікування та ін.), а також ім'я користувача, телефонний номер того, що дзвонить, час розмови і кількість переданих і отриманих байт. Ця ж інформація буде записана і у базу статистики UserGate.

Адміністрування мережі

UserGate включає DHCP- сервер для динамічного призначення IP- адрес в локальній мережі і функцію публікації ресурсів, яка дає можливість отримати доступ ззовні до ресурсів компанії усередині локальної мережі. Функція маршрутизації дозволяє передавати дані між двома локальними підмережами. І нарешті, сервер UserGate можна адмініструвати видалено, підключаючись до нього з будь-якої точки світу.

Мережеве адміністрування

За допомогою UserGate можна виконувати деякі рутинні операції, що дозволяє спростити мережеве адміністрування. Наприклад, вбудований DHCP-сервер автоматизує процес видачі IP- адрес комп'ютерам і іншим пристроям в локальній мережі. Якщо комп'ютер з UserGate підключений до декількох локальних мереж, сервер UserGate можна настроїти як маршрутизатор (router), забезпечивши прозорий, двонаправлений зв'язок між локальними мережами. Публікація ресурсів дозволяє надати доступ до внутрішніх ресурсів компанії, наприклад до Web, FTP, VPN або до поштового сервера. І, нарешті, видалене адміністрування дає можливість видалено підключатися по локальній мережі або через Інтернет з будь-якого комп'ютера, на якому встановлена Консоль Адміністрування UserGate.

- DHCP- сервер
- Маршрутизація
- Публікація ресурсів
- Видалене адміністрування

DHCP- сервер

DHCP (англ. Dynamic Host Configuration Protocol — протокол динамічної конфігурації вузла) — це мережевий протокол, що дозволяє автоматично

отримувати IP— адреси пристроям в локальній мережі у момент їх підключення, тим самим позбавляючи системного адміністратора від ручного налаштування.

Маршрутизація

Якщо комп'ютер з UserGate підключений до декількох локальних мереж, сервер UserGate можна настроїти як маршрутизатор (router), забезпечивши прозорий, двонаправлений зв'язок між локальними мережами.

Публікація ресурсів

Часто виникає необхідність надання доступу до внутрішніх ресурсів компанії ззовні. Як правило, до таких ресурсів відносяться Web-, FTP-, VPN - або поштовий сервер. Для того, щоб надати такий доступ за допомогою UserGate необхідно створити правило перенаправлення запитів на комп'ютер в локальній мережі, на якому запущена відповідна служба.

Видалене адміністрування

До сервера UserGate можна підключатися по локальній мережі або видалено через Інтернет з будь-якої точки світу. Для цього досить встановити на комп'ютер Консоль Адміністрування UserGate, і вказати в налаштуваннях з'єднання IP- адресу і порт сервера UserGate.

Можливість видаленого адміністрування сервера UserGate особливо корисна у разі, коли необхідно адмініструвати декілька серверів UserGate в різних місцях(наприклад, декількох Інтернет-кафе). При цьому, адміністрування здійснюється з однієї і тієї ж консолі — усі доступні сервера відображаються в списку «З'єднання», і можна видалено підключитися до будь-якого з них.

2.7.3 Traffic Inspector

Сертифіковане комплексне рішення для організації і контролю доступу в Інтернет. Не вимагає дорогого мережевого устаткування, забезпечує гнучку тарифікацію, надійний мережевий захист, розподіл завантаження, точний облік і статистику, економію трафіку і робочого часу. Вигляд програми показано на рис. 2.3.

Traffic Inspector спеціально створений для того, щоб об'єднати і доповнити усе різноманіття мережевих можливостей операційних систем Microsoft Windows, тому вам не доведеться робити спеціальних налаштувань. Все, що працювало раніше, працюватиме і після установки програми - знадобиться тільки авторизація користувачів.

Організація доступу в Інтернет. Контроль трафіку

Користувачі(у програмі вони називаються клієнти) можуть працювати як безпосередньо, через NAT, так і через проксі-сервер. Для кожного користувача створюється окремий обліковий запис(чи підвантажується з Active Directory), і усі його дії в мережі відображаються у вигляді простих і зрозумілих звітів.

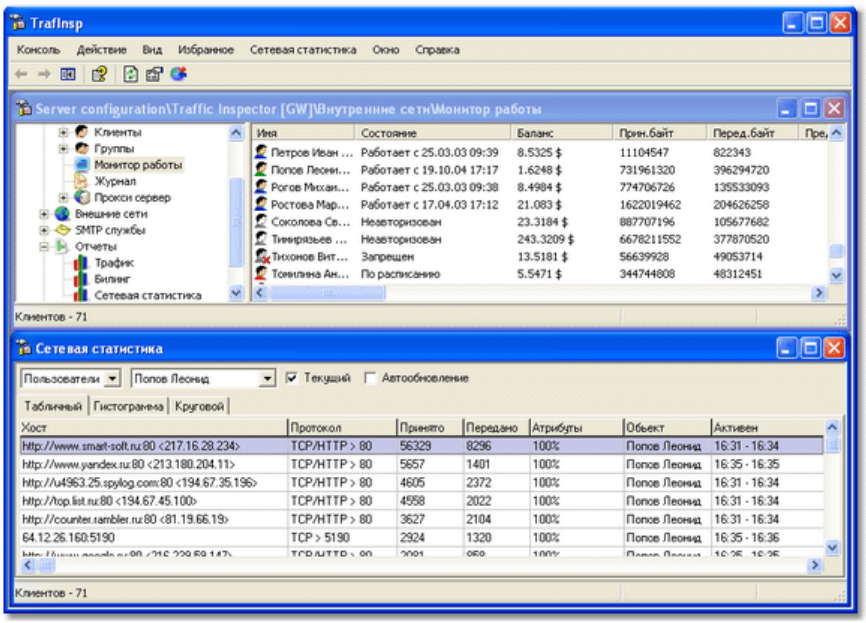

Рисунок 2.3 - Вигляд вікна програми Traffic Inspector

Облік трафіку. Система білінга (billing)

Підрахунок трафіку в програмі відбувається по кожному користувачеві з точністю до байта, причому ви самі визначаєте одиницю обліку, ліміти, кредити, блокування, фільтри і розклади. Система биллинга Traffic Inspector має сертифікат відповідності зв'язку, що гарантує виняткову точність розрахунків.

Проксі-сервер і економія

Використання проксі-сервера Traffic Inspector дозволяє кешувати часто використовувані інтернет-ресурси, а також блокувати банери, рекламні вставки, графіку, музику або відео і забороняти небажані сайти або їх розділи.

Безпека і захист мережі. Firewall

Захист мережі організований двома рівнями: мережевий екран забезпечує захист від зовнішніх атак, а система блокування і сповіщення при надмірній мережевій активності служить для внутрішнього контролю безпеки.

Фільтрація спаму

При використанні поштового шлюзу Traffic Inspector є можливість застосувати систему блокування спаму на внутрішньому поштовому сервері.

Антивірусний захист

Окрім функції своєчасного виявлення зараження мережевими вірусами, для перевірки трафіку на проксі-сервері і поштовому шлюзі Traffic Inspector передбачені додаткові модулі антивірусного захисту.

Управління швидкістю і маршрутизацією

Traffic Inspector дозволяє задавати обмеження швидкості для користувачів або груп з динамічним розподілом навантаження, а система управління маршрутизацією Advanced Routing дає можливість направити трафік на різні канали доступу, у тому числі на супутник.

Видалений контроль і статистика

У програмі є ряд засобів для видаленого управління і моніторингу системи. Використовуючи Traffic Inspector, ви завжди будете в курсі справи про стан мережі, де б ви не знаходилися.

Traffic Inspector може бути використаний як в організаціях для безпечного і ефективного використання інтернет-підключення, так і в невеликих підприємствах, що роблять послуги з передачі даних : провайдерів, інтернет- кафе, готелях і хот-спотах. Щоб переконатися, що Traffic Inspector вам підходить, ми рекомендуємо безкоштовно перевірити в роботі його повнофункціональну версію.

3 ФІЛЬТРАЦІЯ КОНТЕНТУ САЙТІВ СОЦІАЛЬНИХ МЕРЕЖ ЗА ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Для побудови соціальної мережі або соціальних зв'язків між людьми ми використовуємо соціальні мережі, такі як Facebook, Twitter, додатки тощо. Використовуючи ці медіа, користувачі можуть ділитися своїми поглядами та думками про певні речі. Багато людей використовують свої медіа в особистих інтересах, розвагах, на ринку акцій або в ділових цілях. На сьогоднішній день безпека користувачів є основною проблемою для соціальних мереж. Соціальні мережі в Інтернеті надають невелику підтримку щодо фільтрації контенту. У цій статті ми запропонували систему, яка забезпечує безпеку щодо шкідливого контенту, який розміщується на сайтах соціальних мереж. Для фільтрації контенту, який може бути небажаними повідомленнями, маркованими зображеннями або вульгарними зображеннями, ми запропонували трирівневу архітектуру. Користувач також може використовувати функцію автоматичного блокування.

У сучасному житті соціальні мережі відіграють дуже важливу роль. Люди проводять більшу частину свого часу в соціальних мережах, спілкуючись і ділячись своїми ідеями. Використовуючи ці медіа, люди можуть ділитися своєю інформацією або обмінюватися різними типами контенту, такими як зображення, відео, текстові або аудіо повідомлення. Багато людей коментують цей спільний контент. Люди отримують зворотній зв'язок на будь-який матеріал, яким вони поділилися на стіні. Іноді це може бути позитивна відповідь, негативна відповідь або пропозиції, які є дуже корисними для покращення. За даними Facebook, користувачі створюють 90 біт контенту щомісяця; щомісяця вони обмінюються більш ніж 30 мільярдами біт контенту (веб-посилання, новини, записи в блогах, нотатки, фотоальбоми тощо).

Користувачі можуть розміщувати в соціальних мережах будь-який тип контенту. Прикладами можуть бути небажані текстові повідомлення, брендовані фотографії, непристойні, порнографічні зображення, особисті пустотливі коментарі тощо. Інші користувачі можуть бачити ці дописи та коментувати їх.

Соціальний імідж користувача може постраждати в результаті такого повідомлення. Як наслідок, захист стіни такого користувача є критично важливим. До певного моменту Facebook забезпечує захист. Лише обрана група людей на Facebook має доступ до стін інших людей, таких як друзі, друзі друзів або створені групи друзів. Користувач має можливість заблокувати зображення свого профілю. Однак, оскільки фільтрація на основі контенту не підтримується, такі небажані повідомлення не можуть бути попереджені. Метою даної системи є захист стін користувача шляхом фільтрації небажаного контенту та соціальних мереж користувача медіа-зображення. Користувачі можуть змінювати правила фільтрації на свій розсуд. Користувач має контроль над тим, хто може надсилати повідомлення на його стіну. Для фільтрації тексту використовується метод класифікації коротких текстів.

Для навчання використовується словниковий список. Список словника містить слова та їх клас. Для фільтрації зображень з мітками ми використали алгоритм розпізнавання тексту (OCR). OCR може витягувати текст із зображення і зберігати його в текстовому файлі. Потім ми порівнюємо витягнуті дані з набором даних. Якщо збіг знайдено, то зображення буде відфільтровано системою. Для фільтрації зображень ми використовували техніку скін-шерифа, яка складається з алгоритму виявлення шкіри та алгоритму виявлення порнографії. Алгоритм виявлення шкіри може виділити всі ділянки шкіри із зображення і позначити піксель, який класифікується як шкіра, сірим кольором, а не шкіра - білим. Алгоритм виявлення порнографії використовується для обчислення максимальної площі нешкіряного пікселя. Якщо вона перевищує максимальне значення, зображення вважається вульгарним.

3.1 Фільтрація на основі вмісту

У системі фільтрації на основі вмісту документ затверджується шляхом порівняння профілю документа з профілем користувача за допомогою методів пошуку інформації, таких як частота термінів і зворотна частота документів (TF-

IDF). Відповідно до попередніх відгуків та вибору користувача, характеристики користувача були зібрані та профілізовані. Система потребує зв'язку між елементами, який допомагає рекомендувати документ користувачеві. Система починає зі збору інформації про предмет, наприклад, про поведінку, показання і т.д. для предмета, пов'язаного з хворобою. Потім система пропонує користувачеві пройти тест, щоб оцінити матеріал. Потім система порівнює не оцінений пункт з пунктом профілю користувача і присвоює оцінку не оціненому пункту, в результаті чого користувачеві будуть представлені пункти, які оцінюються відповідно до присвоєної оцінки

3.2 Спільна фільтрація

Спільна фільтрація - це метод, який фільтрує інформацію, що підтримує інтерес користувача (тобто історію) і, відповідно, рейтинг інших користувачів з еквівалентним інтересом. Ця техніка фільтрації працює з великою групою людей, щоб знайти меншу групу користувачів зі схожими інтересами. Він створює ранжований список пропозицій. Він відстежує речі, які подобаються користувачам, і об'єднує їх, щоб створити ранжований список пропозицій. Він широко використовується в багатьох системах фільтрації або рекомендаційних системах, особливо в додатках електронної комерції. Прикладами таких програм є Amazon.com, YouTube, eBay тощо, де рекомендації щодо нових продуктів надаються користувачеві на основі його минулої історії покупок, вподобань та антипатій подібних користувачів.

3.3 Персоналізація вмісту OSN на основі політик

В OSN нещодавно з'явилися деякі системи класифікації операторів для персоналізації входу. Великій кількості користувачів сервісів мікроблогів було запропоновано кілька методів класифікації коротких текстових повідомлень. Twitter, пристрій фокусується на спільноті категорій і забезпечує оновлення

контенту. Споживач бачитиме ті типи твітів, які підтримують його інтереси. Кутер і Гольбек, з іншого боку, створили додаток довіри до фільмів, який використовує довірчі відносини OSN та інформацію про походження для налаштування доступу до місця розташування. Однак цим системам бракує політичного рівня фільтрації, який дозволяє користувачеві бачити результати процесу класифікації і вибирати, як відфільтрувати небажані дані.

3.4 Система для фільтрації небажаних дописів

У цьому розділі ми представляємо запропоновану архітектуру системи для фільтрації небажаних дописів на стіні користувача. Поточна система складається з трьох рівнів, як показано на рисунку 3.1. Інтерфейс користувача забезпечується рівнем менеджера соціальної мережі. Зовнішній підтримується шаром "Додатки для соціальних мереж". Третій рівень - це графічні інтерфейси користувача (GUI), які використовуються для відображення результатів. Користувачі спілкуються з системою через графічний інтерфейс користувача (GUI) для налаштування та управління своїми ФР/БЛ (чорними списками).

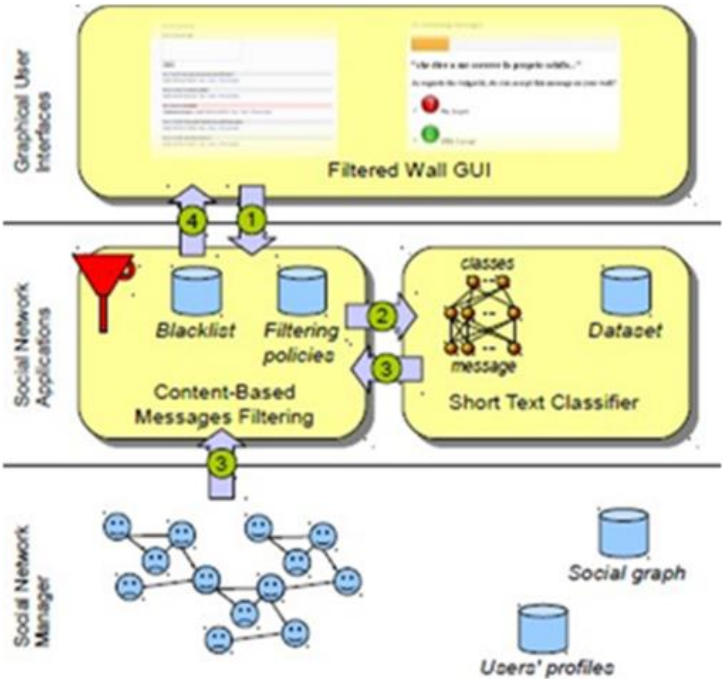

Рисунок 3.1 - Запропонована система для фільтрації небажаних дописів на стіні користувача

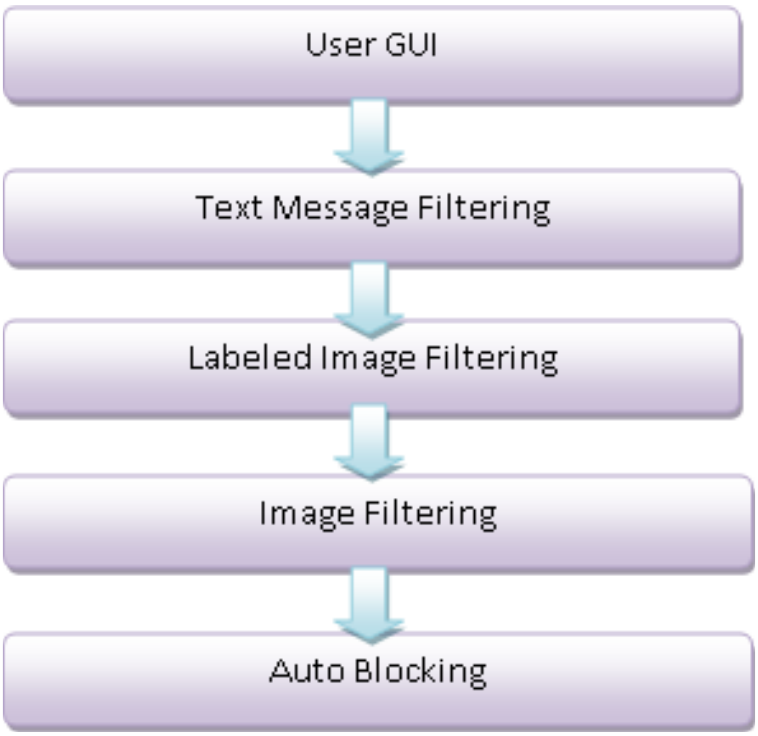

Рисунок 3.2 - Рівні запропонованої архітектури системи

Як показано на рисунку 3.2, він описує рівні запропонованої системи. Графічний інтерфейс користувача використовується для розміщення повідомлень, які можуть бути у будь-якій формі, наприклад, текстове повідомлення, зображення з міткою або зображення. Для автентифікованого користувача передбачена функція автоматичного блокування.

Щоб розмістити повідомлення на стіні, користувач повинен спочатку перейти на стіну, яка розділена на три розділи: фільтрація тексту, фільтрація зображень з мітками та фільтрація зображень.

Дані контенту отримуються за допомогою текстового класифікатора для фільтрації тексту. Він відповідає збору даних. Якщо знайдено збіг, матеріал буде приховано від перегляду на стіні.

Фільтрація зображень: За допомогою алгоритму виявлення шкіри та алгоритму виявлення порнографії ми відфільтрували небажані зображення зі стіни. Якщо зображення знайдено в наборі даних, то воно не відображається на стіні.

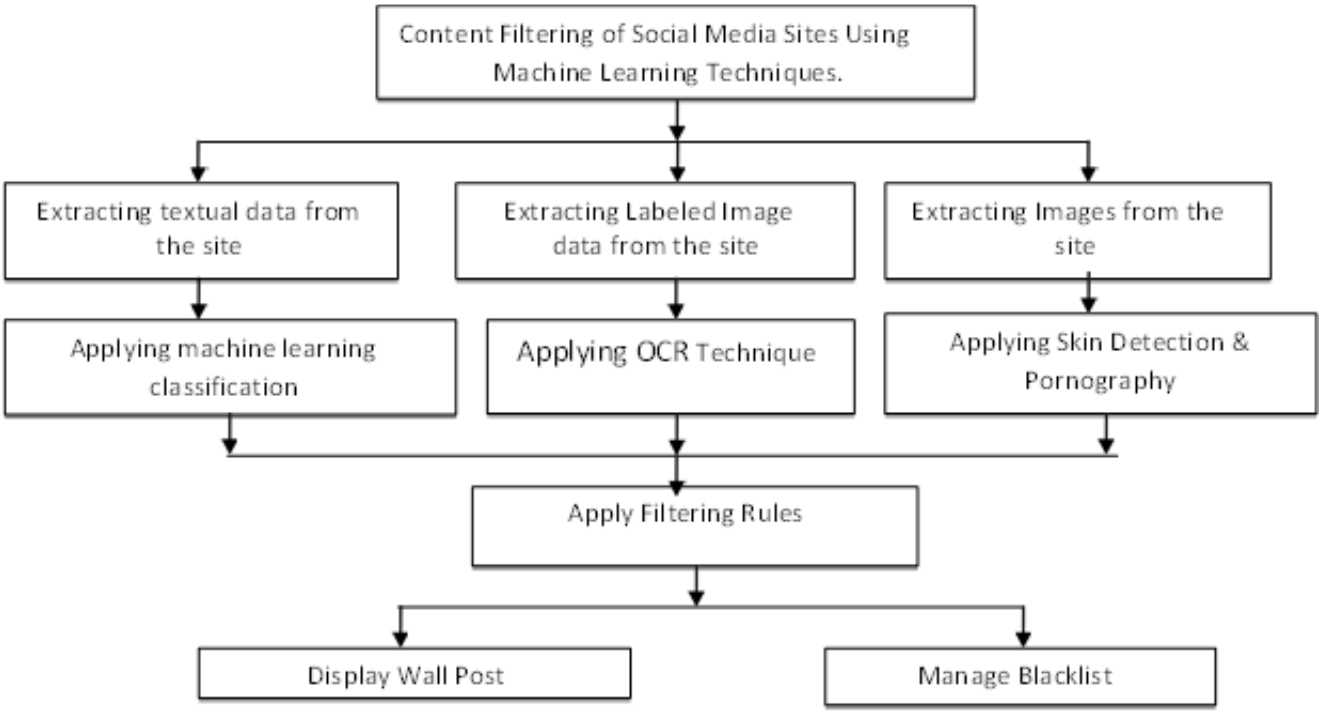

Рисунок 3.3 - Потік запропонованої системи

Для фільтрації зображень з мітками ми використали обидва попередні методи. Ми можемо виділити текст із зображення за допомогою методів розпізнавання тексту і застосувати правила фільтрації повідомлень.

Згідно з цією логікою, вихідні дані будуть відфільтровані та опубліковані на стіні.

3.5 Керування чорним списком та правила фільтрації

3.5.1 Правила фільтрації

Користувачі можуть вказати, який контент має бути заблокований або відображений на відфільтрованій стіні, використовуючи правила фільтрації. Правила фільтрації визначаються відповідно до профілю користувача та використання соціальних мереж. Автор - це особа, яка визначає правила.

Користувачі OSN позначаються через Creator Spec, a Content Spec - це булевий вираз

На рисунку 3.4 показано сторінку "Залишити коментар". Користувач може вибрати стіну друга і опублікувати вміст, наприклад, текстові повідомлення, позначені зображення або зображення.

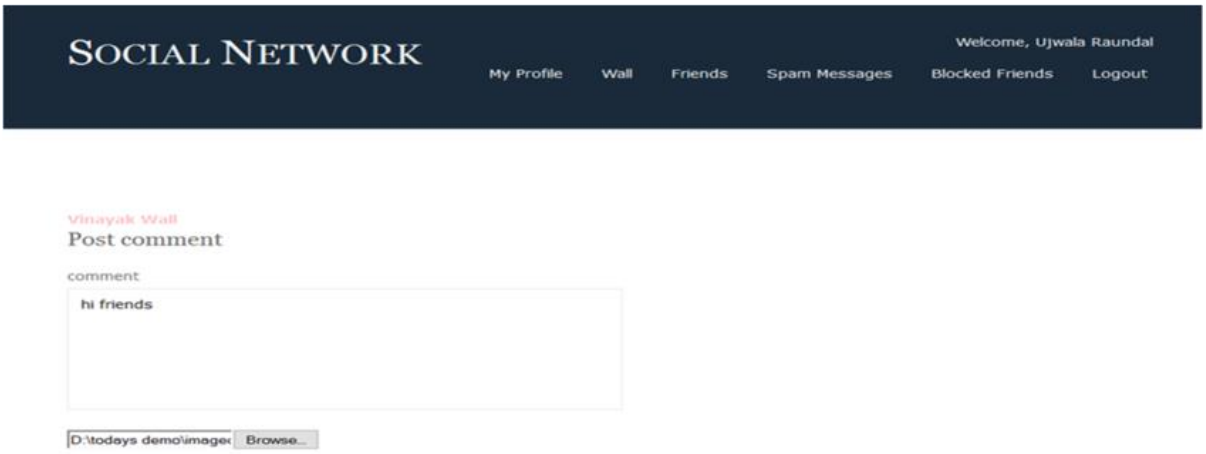

Рисунок 3.4 - Сторінка коментарів до публікації

Рисунок 3.5 показує список небажаних повідомлень та користувачів з чорного списку.

Рисунок 3.5 - Список небажаних повідомлень і користувачів із чорного списку

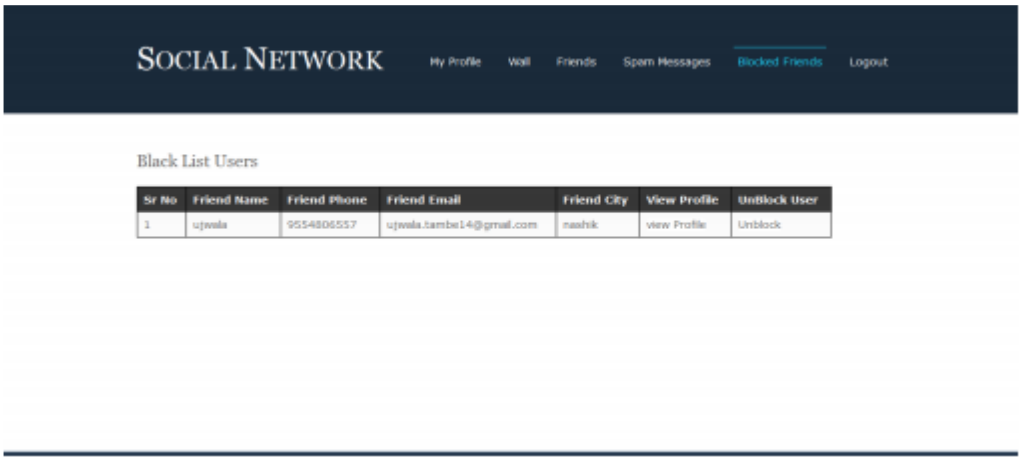

Рисунок 3.6 - Спам-повідомлення та користувачі з чорного списку

На рисунку 3.7 показано вульгарне зображення, яке було відфільтроване системою. Його не можна розміщувати на стіні.

Рисунок 3.7 - Фільтрація зображення

ВИСНОВКИ

В роботі розглянуто програмні та апаратні компоненти, з яких складаються як комп'ютерні мережі взагалі, так і Інтернет зокрема, починаючи з «периферії» комп’ютерних мереж, кінцевих систем та програм, а також транспортних послуг, що надаються програмам, запущеним на кінцевих системах.

Потім увагу було зосереджено на самий нижній з точки зору комунікаційної моделі фізичний рівень передачі даних: було розглянуто основні середовища передачі і технології доступу до глобальної мережі, структуру Інтернету, представивши її як мережу мереж і як ієрархію мереж Інтернет-провайдерів, що дозволила глобальній мережі з легкістю включати в себе нові сегменти.

Інша частина була присвячена основним аспектам функціонування комп'ютерних мереж. Спочатку було розгялуто основні причини затримок і втрат пакетів в процесі передачі, концепції багаторівневої комунікаційної моделі та протоколами кожного з рівнів, складових архітектурну основу комп'ютерних мереж.

В роботі було ознайомлено з технологіями і видами керування Інтернет каналом і зроблено висновки, що для кожної платформи існує немало засобів для управління Інтернет каналом.

У цій дипломній роботі представлено систему для фільтрації небажаних повідомлень, маркованих зображень та небажаних зображень зі стіни користувача. Для фільтрації текстових повідомлень система використовує текстові класифікатори та методи машинного навчання. Для фільтрації зображень з мітками ми використали методи розпізнавання тексту. За допомогою алгоритму виявлення порнографії ми можемо виявляти вульгарні зображення. Наша система також передбачає систему автоматичного блокування. Таким чином, тільки авторизовані користувачі можуть розміщувати повідомлення на стіні користувача.

ПЕРЕЛІК ПОСИЛАНЬ

1. Ванетті, Елізабетта Бінагі, Олена Феррарі, Барбара Кармінаті, Морено Карулло Факультет комп'ютерних наук та комунікацій, Університет Інсубрії "Система фільтрації небажаних повідомлень на стіні користувача OSN" IEEE Transactions on Knowledge And Engineering Flight Data: 25 the Year 2013

2. Ф. Себастьяні, "Машинне навчання в автоматизованій категоризації текстів", ACM Computing Surveys, т. 34, № 1, с. 1-47, 2012. J. Leskovec, D. P. Huttenlocher, and J. M. Kleinberg, "Predicting Positive and negative links in online social networks," inProc. 19th Int. Conf. World Wide Web, 2010, с. 641-650.

3. Vinaitheerthan Renganathan1*, Ajit N Babu2 та SuptendraNath Sarbadhikari3 "Підручник з концепцій та методів фільтрації інформації для біомедичного пошуку".

4. B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in Twitter to improve information filtering," in Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Searching, SIGIR 2020, 2020, pp. 841-842.

5. J. Golbeck, "Combining provenance with trust in social networks for semantic web content filtering," in Provenance and Annotation of Data, ser. Конспект лекцій з інформатики, Л. Моро та І. Фостер, ред.. Springer Berlin / Heidelberg, 2006, vol. 4145, pp. 101-108.

6. Pennock DM, Horvitz E, Lawrence S, Giles CL (2021) Спільна фільтрація за допомогою діагностики особистості: Гібридний підхід на основі пам'яті та моделей. Матеріали Шістнадцятої конференції з невизначеності в штучному інтелекті (UAI-2021), Morgan Kaufmann Publishers Inc 473-480.

7. Богер З., Куфлик Т., Шапіра Б., Шовал П. (2000) Фільтрація інформації та автоматична ідентифікація ключових слів за допомогою штучних нейронних мереж. Матеріали 8-ї Європейської конференції з інформаційних систем.

8. Дженнінгс А, Хігучі Х (1993) Нейронна мережа з моделлю користувача для персональної служби новин. Моделювання користувача та адаптована до користувача взаємодія 3: 1-25.

9. P. J. Hayes, P. M. Andersen, I. B. Nirenburg, and L. M. Schmandt, "Tcs: a shell for content-based text categorization," in Proceedings of 6th IEEE Conference on Artificial Intelligence Applications (CAIA-90) IEEE Computer Society Press, Los Alamitos, US, 2018, pp. 320-326.

10. Н. Я. Бєлкін та В. Б. Крофт, "Фільтрація інформації та інформаційний пошук: Дві сторони однієї медалі?" Communications of the ACM, vol. 35, no. 12, pp. 29-38, 2022.

11. P. W. Foltz та S. T. Dumais, "Персоналізована доставка інформації: Аналіз методів фільтрації інформації", Communications of the ACM, vol. 35, no. 12, pp. 51-60, 2013.

12. S. Zelikovitz and H. Hirsh, "Improving short text classification using unmabeled background knowledge," in Proceedings of 17th International Conference on Machine Learning (ICML-00), P. Langley, Ed. Стенфорд, США: Morgan Kaufmann Publishers, Сан-Франциско, США, 2017, с. 1183- 1190.

13. В. Бобічев та М. Соколова, "Ефективний та надійний метод класифікації коротких текстів", в AAAI, Д. Фокс та К. П. Гомес, Ред. AAAI Press, 2018, с. 1444-1445.

14. http://www. myvocabulary. com/word-list/violencevocabulary

15. https://archive. ics. uci. edu/ml/datasets/SMS+Spam+Collection

16. http://www. bannedwordlist. com/lists/swearWords. Txt

17. Wang, J., De Vries, A. P. and Reinders, M. J. (2006). Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 501–508.

18. Billsus, D. and M.J. Pazzani, 2018. Learning collaborative information filters. Proceeding of the 15th International Conference on Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp: 46-54.

19. Billsus, D. and M.J. Pazzani, 2020. User modeling for adaptive new access. User Mod. User-adapted Interac., 10(2-3): 147-180.

20. Liu, H., Hu, Z., Mian, A., Tian, H. and Zhu, X. (2014). A new user similarity model to improve the accuracy of collaborative filtering. KnowledgeBased Systems, Vol. 56, pp. 156–166.

21. Sun, D., Luo, Z. and Zhang, F. (2021). A novel approach for collaborative filtering to alleviate the new item cold-start problem. In Communications and Information Technologies (ISCIT), 2021 11th International Symposium on, IEEE, pp. 402–406.

22. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. And Sartin, M. (2019). Combining content-based and collaborative filters in an online newspaper. In Proceedings of ACM SIGIR workshop on recommender systems, Vol. 60.

23. Basu, C., Hirsh, H. and Cohen, W. (2018). Recommendation as classification: Using social and content-based information in recommendation. In Proceedings of the national conference on artificial intelligence, pp. 714–720.

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ