



МІНІСТЕРСТВО
ЕКОНОМІЧНОГО
РОЗВИТКУ І ТОРГІВЛІ
УКРАЇНИ

УКРАЇНА

(19) UA

(11) 119992

(13) U

(51) МПК

H04L 12/70 (2013.01)

(12) ОПИС ДО ПАТЕНТУ НА КОРИСНУ МОДЕЛЬ

(21) Номер заявки: **u 2016 12797**

(22) Дата подання заявки: **15.12.2016**

(24) Дата, з якої є чинними
права на корисну
модель: **25.10.2017**

(46) Публікація відомостей
про видачу патенту: **25.10.2017, Бюл.№ 20**

(72) Винахідник(и):

**Каптур Вадим Анатолійович (UA),
Князєв Олександр Андрійович (UA)**

(73) Власник(и):

**ОДЕСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ
ЗВ'ЯЗКУ ІМ. О.С. ПОПОВА,
вул. Кузнечна, 1, м. Одеса, 65029 (UA)**

(54) СПОСІБ АДАПТИВНОЇ ФІЛЬТРАЦІЇ УНІКАЛЬНИХ ІДЕНТИФІКАТОРІВ РЕСУРСІВ МЕРЕЖІ ІНТЕРНЕТ

(57) Реферат:

Спосіб адаптивної фільтрації унікальних ідентифікаторів ресурсів мережі Інтернет включає зменшення середнього часу обробки унікального ідентифікатора ресурсу мережі Інтернет всередині системи фільтрації контенту за рахунок адаптивної перебудови порядку слідування процедур оцінки відповідності унікальних ідентифікаторів ресурсів мережі Інтернет всередині системи фільтрації контенту. Також спосіб включає здійснення сортування послідовності виконання процедур оцінки відповідності унікальних ідентифікаторів ресурсів мережі Інтернет за принципом переміщення на початкові позиції тих процедур, що спрацьовували частіше за інші, за результатами моніторингу протягом певного періоду часу.

UA 119992 U

Запропонована корисна модель належить до техніки зв'язку, зокрема до фільтрації унікальних ідентифікаторів ресурсів (URI) мережі Інтернет.

Питання фільтрації нецільового контенту як на локальних (підприємства чи окремі користувачі), так і на глобальних рівнях досліджено у працях низки вітчизняних та закордонних вчених. Так, в роботах [1-3] висвітлено науково-методичні підходи щодо визначення найбільш ефективного способу організації системи фільтрації нецільового контенту в мережі організації. В працях [4, 5] досліджено світовий досвід управління інформаційною безпекою, її вплив на стабільність держави, сформовано методи захисту дитини в мережі Інтернет, а також наведено класифікацію систем виявлення небажаних втручань.

Основною технічною мірою захисту людини від негативної інформації в мережі Інтернет є технічна фільтрація інформації, що включає в себе різноманітні засоби фільтрації небажаного контенту. До основних засобів належать: фільтрація на базі HTTP проксі-серверу; фільтрація на базі DNS-серверу; фільтрація на базі брандмауера (firewall); фільтрація на базі веб-клієнта.

Існуючі способи на сьогоднішній день мають низку недоліків, які можуть бути повністю або частково вирішені за рахунок комбінування із іншими засобами, тому кожен з перерахованих вище способів фільтрації контенту може використовуватися як окремо, так і взаємодіяти з іншими, при цьому виникає поняття комплексної системи фільтрації контенту (КСФК), що включає в себе кілька систем фільтрації контенту (СФК), що містять певні методи фільтрації, будь-то використання проху-серверів або фільтрація на DNS-серверах [6].

В межах КСФК засоби фільтрації можуть працювати у послідовному (підсилюючи один одного) або в паралельному (доповнюючи один одного) режимі. В послідовному режимі адреса або контент, що пройшли процедуру фільтрації на одному з засобів - потрапляють на вхід іншого засобу з метою повторної обробки (наприклад за допомогою інших методів). В свою чергу паралельний режим припускає фільтрацію адреси через використання одного засобу фільтрації, а фільтрацію контенту з використанням іншого.

З погляду системи прийняття рішення про дозвіл або блокування запиту, всі КСФК можна умовно розділити на два типи:

Системи миттєвого прийняття рішення. КСФК цього типу базуються на послідовному виконанні процедур перевірки відповідності URI заданими правилами з можливістю передчасного переривання процесу оцінки в разі позитивного (або негативного) спрацьовування тієї чи іншої процедури;

Системи накопиченого аналізу. КСФК цього типу також базуються на послідовному виконанні процедур перевірки відповідності URI заданим правилам, приймаючи рішення про дозвіл або заборону доступу до того чи іншого інформаційного ресурсу на підставі відомостей, що надходять від певної сукупності виконаних процедур (найчастіше - всіх процедур відповідності, що входять до складу КСФК).

Зміна послідовності виконання процедур перевірки відповідності URI для КСФК другого типу, як правило, не викликає зміни підсумкового часу обробки запиту і, таким чином, КСФК цього типу не можуть бути оптимізовані за цим параметром. В свою чергу КСФК миттєвого прийняття рішення можуть бути оптимізовані за рахунок запровадження адаптивної фільтрації унікальних ідентифікаторів ресурсів мережі Інтернет.

Структурно процес аналізу URI в КСФК, який досліджено у низці наукових праць, зокрема [6], складається з таких етапів:

1. Запити (які в нашому випадку представлені у вигляді URI) надходять до КСФК та обробляються вбудованою системою прийняття рішень, що базується на відомостях про профіль користувача (або групи користувачів) та прийняту для нього (або групи) політику фільтрації.

2. Система прийняття рішень, відповідно до результатів проведеного аналізу, а також згідно зі способом отримання запиту (запит надійшов до DNS-серверу, запит надійшов до проху-серверу тощо), формує послідовність засобів фільтрації (один або більше), що мають бути задіяні для оцінки доцільності блокування (або надання доступу) до ресурсу, що представлено відповідним URI.

3. В свою чергу, кожен з засобів фільтрації (проху-сервер, DNS-сервер, міжмережевий екран тощо) після отримання для аналізу URI формує послідовність процедур оцінювання відповідності URI заданим правилам з врахуванням доступних конкретному засобу можливостей (аналіз за доменом або доменною зоною, використання регулярних виразів, аналіз за номером порту тощо), а також із використанням відповідних "білих" або "чорних" списків (відповідно до політик фільтрації, що мають застосовуватись до конкретних користувачів).

4. У разі, якщо хоча б одна з процедур підтвердить необхідність блокування ресурсу, що представлено відповідним URI, КСФК має прийняти рішення про його блокування (наприклад, через надсилання у відповідь сторінки із повідомленням про блокування). У разі, якщо жодна з процедур не підтвердить необхідності блокування ресурсу, або хоча б одна з них підтвердить його входження до "білого" списку КСФК має забезпечити доставку запитаного контенту користувачеві.

У процесі аналізу, результат обробки URI буде безпосередньо залежати від того, яка саме за порядком процедура оцінювання прийме те чи інше рішення. Так, наприклад, у разі позитивного спрацьовування першої ж процедури (наявність ресурсу в "білому" або "чорному" списку) час обробки унікального ідентифікатора ресурсу буде найменшим, а у разі невідповідності ІЖІ жодному з правил жодної з процедур - найбільшим.

Технічний результат, що досягається при здійсненні запропонованого способу полягає в скороченні часу обробки унікальних ідентифікаторів ресурсів мережі Інтернет за рахунок перестановки процедур СФК всередині КСФК.

Для вибору найбільш ефективного порядку проходження процедур оцінювання, необхідно враховувати часові характеристики кожної задіяної процедури, тобто при надходженні запиту, відбувається перевірка на збіг в кожному списку певної процедури. Враховуючи той факт, що процедур в конкретному способі фільтрації може бути досить багато, позначимо як S_{ij} кількість записів у списку, що знаходиться на j -м місці в межах i -го засобу фільтрації (згідно чинного порядку застосування в межах КСФК). Кожна перевірка списку здійснюється певною процедурою $R=\{R_1, R_2, R_3 \dots R_k\}$, де k - кількість видів процедур.

У процесі перевірки на збіг у списку певної процедурою R , витрачається певний час t . Відповідно, кожна перевірка певними процедурами $\{R_1, R_2, R_3 \dots R_k\}$, буде супроводжуватися конкретним часом $\{T_{R1}, T_{R2}, T_{R3} \dots T_{Rk}\}$.

Очевидно, що для більшості процедур оцінки відповідності, час обробки конкретною процедурою безпосередньо залежить від розміру списку, що перевіряється тобто $T_{Rk}=\{f(S, \Delta t)\}$, де Δt - час обробки одного запису в списку.

Враховуючи той факт, що в процесах перевірки на збіг URI в різних списках буде витрачатися різний час, що залежить від місцезнаходження URI в самому списку, можна зробити висновок, що чим ближче URI буде перебувати до початку списку, тим менше буде витрачатися час на обробку запиту (табл.). Виходячи з цього, проведені в роботі дослідження дозволяють констатувати, що при відсутності URI в списку буде витрачатися максимальний час обробки процедури ($S_{ij} \cdot \Delta t$), зважаючи на те, що має бути проаналізований весь список на наявність збігу. Відсутність URI в списку є аналогічною місцезнаходженню URI останнім записом в списку, тобто на його пошук буде витрачатися максимальний час. Відповідно, мінімальний час обробки процедури (Δt) буде в тому випадку, якщо URI буде знаходитися першим записом в списку, що є часом обробки одного запису, тобто Δt . Якщо ж URI знаходиться у середині списку,

то час обробки біде дорівнювати $\frac{S_{ij}}{2} \cdot \Delta t$.

Далі, враховуючи той факт, що час обробки буде максимальним, якщо у списках не буде виявлено URI або ж загальний час обробки запиту перевищить прийнятне (з точки зору вимог до якості обслуговування) значення, необхідно ввести поняття про систему подій, де в ролі події виступатиме вихід з засобу фільтрації, з подальшим прийняттям рішення.

Процес прийняття рішення представимо у вигляді системи подій $\Omega_{\text{проц}}=\{A_1, A_2, A_3 \dots A_n\}$, де A_n - подія виходу з засобу фільтрації після n -ої процедури (наприклад прийнято рішення фільтрувати вхідний запит).

Таблиця

Залежність часу обробки t від місця розташування URI в списку

Місцезнаходження запису URI	Час обробки t , с
URI відсутній у списку	$S_{ij} \cdot \Delta t$
URI знаходиться першим записом у списку	Δt
URI знаходиться останнім записом у списку	$S_{ij} \cdot \Delta t$

URI знаходиться у середині списку	$\frac{S_{ij}}{2} \cdot \Delta\tau$
-----------------------------------	-------------------------------------

Зважаючи на те, що алгоритм не може наперед визначити яка саме подія настане, ці події можна вважати випадковими з відповідними ймовірностями $P(A_1, A_2, A_3 \dots A_n)$ [9]. Відповідно

$$\sum_{i=1}^{S_{ij}} P(A_i) = 1.$$

5 Таким чином, при настанні події A_r час обробки запиту буде дорівнювати $\Delta\tau$. В свою чергу при настанні події A_2 , він буде дорівнювати $2 \cdot \Delta\tau$, а при настанні події A_n - $S_{ij} \cdot \Delta\tau$.

Порядок проходження можна вважати оптимальним у тому випадку, якщо середній час обробки запиту до прийняття того чи іншого рішення буде мінімальним. Приймаючи, що середній час порівняння URI запиту із записом в списку блокування і середній час тестування
10 URI регулярним виразом є величини постійні і приблизно еквівалентні, а середній час обробки запиту залежить від кількості записів у списках фільтрації та кількості регулярних виразів, отримаємо формулу для визначення "загального часу обробки вхідного запиту":

$$T_{\text{обробки}}^{ij} = \sum_{v=1}^{S_{ij}} (P(A_v) \cdot v \cdot \Delta\tau), \quad (1)$$

15 де v - кількість записів у списку; $\Delta\tau$ - час обробки одного запису в списку (із застосуванням відповідної процедури), $P(A_v)$ - імовірність позитивного спрацьовування процедури на позиції v чинного списку.

Час обробки вхідного запиту всіма процедурами відповідності в межах однієї СФК можна визначити за формулою:

$$20 \quad T_{\text{обробки}}^i = \sum_{j=1}^{M_i} T_{ij}, \quad (2)$$

де M_i - кількість процедур відповідності в межах однієї СФК.

Процес прийняття рішення в процесі роботи всього алгоритму також супроводжується системою подій $\Omega_{\text{зас}} = \{A_1^{\text{вих}}, A_2^{\text{вих}}, A_3^{\text{вих}} \dots A_{M_i}^{\text{вих}}\}$. При цьому процес виходу з алгоритму

25 характеризується ймовірностями $P(A_1^{\text{вих}}, A_2^{\text{вих}}, A_3^{\text{вих}} \dots A_{M_i}^{\text{вих}})$.

Зважаючи на це, формула для обчислення часу середньої затримки буде наступною:

$$T_{\text{обробки}}^i = \sum_{j=1}^{M_i} (P(A_j^{\text{вих}}) \cdot \left[\sum_{v=1}^{S_{ij}} (P(A_v) \cdot v \cdot \Delta\tau) \right]). \quad (3)$$

Технічно задача вирішується в такий спосіб (креслення):

30 1. В адаптивну комплексну систему фільтрації контенту (АКСФК) надходить потік запитів від користувачів (груп користувачів), в процесі запиту від користувачів, система визначає тип запиту (URI, IP-адресу тощо). Після цього відбувається перевірка на наявність користувача в системі. Якщо ж у базі даних відсутній профіль конкретного користувача, то відбувається його додавання в систему, а також застосування для даного профілю послідовності засобів і процедур
35 фільтрації небажаного контенту за умовчанням.

2. Якщо ж профіль користувача присутній в базі даних, то для нього буде задіяний механізм, що підбирає найкращу послідовність використання способів і процедур фільтрації таким чином, щоб методи і процедури, які найчастіше спрацьовували, були в пріоритеті і спрацьовували першими. Відбувається це в процесі фіксування частоти застосування засобів і процедур для
40 кожного конкретного користувача (або групи користувачів) і процесів блокування, які будуть зберігатися в базі даних.

3. Далі, на основі вхідної рекомендованої послідовності відбувається перевірка на наявність даної послідовності, якщо ж даної послідовності рекомендованих параметрів не виявлено або ж недостатньо (в основному для нових профілів), то в даному випадку буде застосовуватися
45 раніше згадана послідовність засобів і процедур фільтрації небажаного контенту за умовчанням. У разі виявлення такої послідовності, буде задіяний механізм підбору найкращої послідовності. Після цього відбувається перевірка з подальшим прийняттям рішення про блокування або ж доступ на запитуваний ресурс.

4. У процесі роботи АКCFK відбувається оновлення даних для кожного користувача про рішення блокування або ж доступу, на основі якої і буде працювати перебудова послідовності засобів (і процедур усередині методу) з погляду пріоритетності, про що було сказано раніше.

5. Також даний алгоритм підтримує зворотний зв'язок з користувачем, оповіщаючи його про те, що ресурс заблокований, у разі, якщо система визначила запитуваний ресурс шкідливим (на основі збігу в "чорному списку").

10 Запропонований спосіб адаптивної комплексної фільтрації контенту, зменшує середній час обробки унікальних ідентифікаторів ресурсів мережі Інтернет у КCFK за рахунок сортування процедур всередині CFK. Також, при своєму застосуванні запропонований спосіб фільтрації контенту на обчислювальній техніці нового покоління, де час обробки одного запиту є відносно незначним, сумарний виграш, при обробці великої кількості вхідних запитів, буде постійно зростати, що буде позитивно відображатись на загальному часу обробки всіх вхідних запитів.

На кресленні наведено алгоритм забезпечення заданого класу обслуговування.

Умовні позначення (креслення):

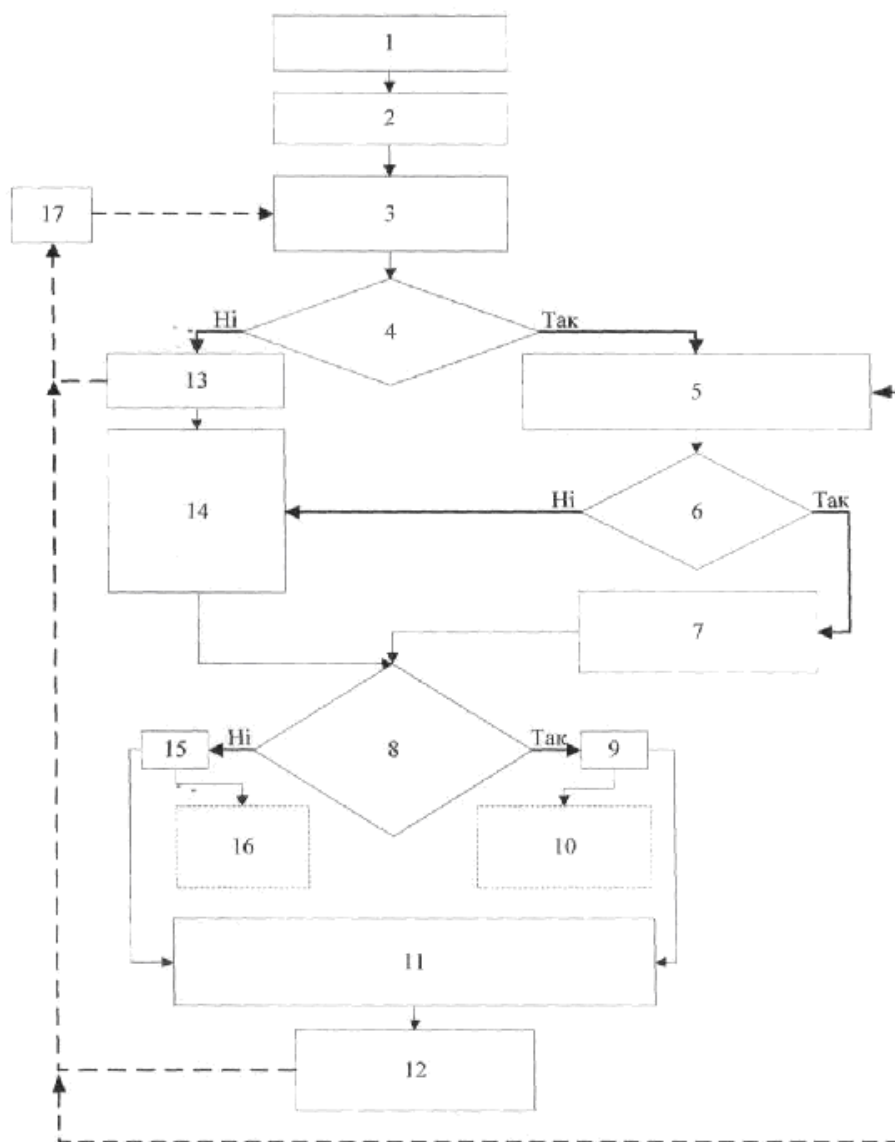
- 1 Отримання вхідного запиту системою фільтрації.
- 2 Визначення відправника запиту.
- 3 Пошук профілю відправника в базі даних.
- 4 У відправника є профіль?
- 5 Отримання рекомендованої для профілю послідовності засобів і процедур фільтрації.
- 6 Послідовність знайдено?
- 7 Формування послідовності засобів і процедур фільтрації.
- 8 Ресурс повинен бути заблокований?
- 9 Перенаправлення вхідного запиту в чорний список.
- 10 Повідомлення про блокування ресурсу.
- 11 Збереження відомостей (статистичної інформації) про результат блокування/надання доступу до профілю відправника.
- 12 Оновлення профілю відправника.
- 13 Додавання нового профілю у базу даних.
- 14 Використання послідовності засобів і процедур фільтрації за замовченням.
- 15 Перенаправлення вхідного запиту в білий список.
- 16 Надання доступу до ресурсу.
- 17 База даних.

Джерела інформації:

1. ПП. Воробієнко // Єдина система обмеження доступу до нецільових ресурсів мережі Інтернет в освітніх закладах України. / ПП. Воробієнко, В.А. Каптур, В.А. Коляденко, В.О. Самодід //Комп'ютер у школі та сім'ї. - 2009. - № 8. - С. 30-34.
- 20 2. Kaptur V. // Current status and prospects of the content filtering methods in the telecommunication networks. // Ukrainian Scientific Journal of Information Security, 2014.- Vol. 20, issue 2. - P. 113-119.
3. Каптур В.А. Система фільтрації SMS-повідомлень в мережі оператора мобільного зв'язку. /В.А. Каптур, А.Г. Ложковський, М.В. Фурмур, Р.В. Чумаков // Наукові праці ОНАЗ ім. О.С. Попова. - 2011. - № 2. - С. 19-24.
- 25 4. Баранов А.А. / Региональная инициатива "Создание центра по защите детей в сети Интернет для региона СНГ". / А.А. Баранов, В.А. Каптур //Региональное подготовительное собрание для стран СНГ к ВКРЭ-14., Кишинёв, Молдова, 19-21 февраля 2013, Документ RPM-CIS13/08.
- 30 5. Корченко А.А., Ахметова С.Т. / Базовые признаки классификации систем обнаружения вторжений. / Сучасні інформаційно-телекомунікаційні технології: матеріали науково-технічної конференції (м. Київ, 17-20 листопада 2015 р.). У 5 томах. - Том 4. Сучасні технології інформаційної безпеки. - Київ, ДУТ, 2015. - С. 25-27.
- 35 6. Каптур В.А., Поднебесний І.А. / Формування профілів ефективної оцінки URI в комплексних системах фільтрації контенту. Вимірювальна та обчислювальна техніка в технологічних процесах: Матеріали XIV Міжнар. наук.-техн. конф. (5-10 Червня 2015 р., м. Одеса); ОНАЗ ім. О.С. Попова. - Одеса-Хмельницький: ХНУ, 2015. - С. 26-29.

ФОРМУЛА КОРИСНОЇ МОДЕЛІ

- Спосіб адаптивної фільтрації унікальних ідентифікаторів ресурсів мережі Інтернет, що включає зменшення середнього часу обробки унікального ідентифікатора ресурсу мережі Інтернет всередині системи фільтрації контенту за рахунок адаптивної перебудови порядку слідування процедур оцінки відповідності унікальних ідентифікаторів ресурсів мережі Інтернет всередині системи фільтрації контенту, який **відрізняється** тим, що включає здійснення сортування послідовності виконання процедур оцінки відповідності унікальних ідентифікаторів ресурсів мережі Інтернет за принципом переміщення на початкові позиції тих процедур, що спрацьовували частіше за інші, за результатами моніторингу протягом певного періоду часу.



Комп'ютерна верстка А. Крижанівський

Міністерство економічного розвитку і торгівлі України, вул. М. Грушевського, 12/2, м. Київ, 01008, Україна

ДП "Український інститут інтелектуальної власності", вул. Глазунова, 1, м. Київ – 42, 01601