Жовноватюк Р. М., к.т.н., с.н.с.,
Варшавський П. В., курсант
Житомирський військовий інституту імені С. П. Корольова

СТВОРЕННЯ БАЗИ ВЕБ-ОРІЄНТИРІВ У СИСТЕМІ IP-ГЕОЛОКАЦІЇ

IP-геолокація може бути використана для виявлення аномальної активності, аутентифікації користувачів, фільтрації контенту, захисту від DDoS-атак, а також для інших заходів кібербезпеки та контролю в мережі Інтернет. Наразі відомо декілька методів IP-геолокації. Зокрема це GeoPing, CBG, SOI, TBG та Octant методи [1, 2]. Найбільш перспективним методом IP-геолокації вважається дещо модифікований CBG метод. При його реалізації є необхідність знаходження веб-орієнтирів, їх координати використовуються для визначення координат цілі за IP-адресою. У якості веб-орієнтирів обираються організації, компанії, наукові установи та урядові установи, які розміщують свої веб-служби локально. Такі орієнтири є стабільними, надійними та довготривалими.

У доповіді зазначено, що для розробки програмного забезпечення по створенню бази веб-орієнтирів використовувалась технологія WindowsForms з використанням мови програмування C# та фреймворка NET. Структура проекту складається з сукупності окремих компонентів, таких як ядро парсерів, ядро роботи з API, ядро роботи з базою даних, моделей, що описують структуру JSON-файлів та даних, що описують сутність «орієнтир». При написанні програмного коду застосовувався об'єктно-орієнтований принцип програмування.

Створення бази веб-орієнтирів розпочинається зі складання списку потенційних веб-орієнтирів, для чого використовуються каталоги з навчальними закладами і компаніями України.

Особливо важливою є інформація про фізичне розташування установи та посилання на її веб-сайт. Ця інформація може бути використана для перевірки, чи існує локальний веб-сервер конкретного закладу за вказаною адресою. Важливо враховувати, що не всі компанії мають власний локальний веб-сервер, оскільки деякі можуть орендувати віддалений веб-сервер або використовувати віртуальний хостинг. Для перевірки цього використовуються веб-додатки з власним API.В результаті обробки даних з геосервісів отримується інформація про географічне розташування IP-адреси, а саме про країну розміщення, область та місто. Цей етап отримав назву «грубої фільтрації».

Далі порівнюється фізична адреса та дані розташування серверу, на якому хоститься веб-сайт компанії. Оскільки на сайті адреса може бути у різному вигляді, то для обробки такої інформації реалізовано спеціальні модулі. Модулі здійснюють перевірку таких елементів: назва країни, населеного пункту, вулиці, номера будинку та поштового коду міста тощо. Якщо елемент – непотрібна інформація (номер квартири, поверху тощо), то ці дані ігноруються. В результаті роботи отримується клас даних: місто, вулиця, номер будинку, поштовий код установи.

Далі за допомогою сервісу LocationQ, для якого реалізовано окремий клас даних, визначаються координати. З метою підвищення точності ці координати перевіряються на сервісі GeoApiFy, який повертає зворотні дані про місто, вулицю та номер будинку за координатами: якщо дані співпадають, то це означає, що координати визначені правильно.

Заключним етапом є обробка всієї наявної інформації та записування в базу даних веб-орієнтирів. Дані, отримані шляхом парсингу веб-ресурсів, зберігаються в окремих таблицях. Кожен запис у таблиці обробляється за таким алгоритмом:

1. Перевірка доступності сайту.
2. Отримання даних з геосервісів (виклик окремого методу).
3. Верифікація отриманої інформації.
4. Запис у базу даних веб-орієнтирів такої інформації: IP-адреса веб-сайту, домене ім'я, повна фізична адреса, адреса з геосервісів та координати.

У результаті розроблено програмне забезпечення, що створює базу даних веб-орієнтирів системи IP-геолокації, що має функцію оновлення відомостей. Подальша робота спрямована на розробку та програмну реалізацію алгоритму визначення місцеположення об'єктів за їх IP-адресою.

Список використаних джерел

1. Wang Y., Burgener D., Flores M., Kuzmanovic A. and Huang C. (2011) Towards Street-Level Client Independent IP Geolocation. In NSDI'11. Proceedings of the 8th USENIX conference on networked systems design and implementation, pp: 27-36.

2. Eriksson B., Barford P., Sommers J. andNowak R. (2010) A Learning-based Approach for IP Geolocation. IN PAM'10 Proceedings of the 11th international conference on Passive and active measurement, pp: 171 – 180.