GEO-LOCATION AUTHENTICATION: IMPROVING SECURITY USING DATA LEAKS

Ivan Drankov, Rosita Nesheva

University of Mining and Geology “St. Ivan Rilski”, 1700 Sofia; E-mail: ivan.drankov@mgu.bg; rosita.nesheva@mgu.bg

ABSTRACT: Geo-location technology is becoming increasingly important for secure authentication purposes, allowing users to prove their identity based on their physical location. Geo-location algorithms are able to identify the exact location of a device or user, providing a unique identifier that can be used to verify their identity. This technology is particularly useful for verifying the location of mobile devices which are often used for online transactions and other sensitive activities. By using geolocation for authentication, organisations can ensure that only authorised users are able to access sensitive data or perform certain actions, improving security and reducing the risk of fraud or other unauthorised activities.

Key words: geo-location, Multifactor authentication, leaked data.

УДОСТОВЕРЯВАНЕ НА ГЕОГРАФСКО МЕСТОПОЛОЖЕНИЕ: ПОДОБРЯВАНЕ НА СИГУРНОСТТА ЧРЕЗ ИЗПОЛЗВАНЕ НА ИЗТЕКЛИ ДАННИ

Иван Дрънков, Росита Нешева

Минно-геоложки университет „Св. Иван Рилски“, 1700 София

РЕЗЮМЕ. Технологията за географско местоположение става все по-важна за целите на сигурното удостоверяване, като позволява на потребителите да доказват своята самоличност въз основа на физическото си местоположение. Алгоритмите за геолокация са в състояние да идентифицират точното местоположение на дадено устройство или потребител, предоставяйки уникален идентификатор, който може да се използва за удостоверяване на тяхната самоличност. Тази технология е особено полезна за проверка на местоположението на мобилни устройства, които често се използват за онлайн транзакции и други чувствителни дейности. Като използват геолокацията за удостоверяване, организациите могат да гарантират, че само оторизираните потребители имат достъп до чувствителни данни или извършват определени действия, като подобряват сигурността и намаляват риска от измами или други неоторизирани дейности.

Ключови думи: геолокация, удостоверяване, изтекли данни

1. Introduction to multifactor authentication and its connection to geo-location

As the digital world becomes increasingly interconnected, securing personal information and maintaining privacy has become a paramount concern. “Increase in cyber-crime now makes it pertinent to look beyond securing systems with passwords only. Even using the password-less approach may not suffice enough to provide the required security, although it is safer than using only a password” (Ivanov et al., 2022). Multi-Factor Authentication (MFA) has emerged as a crucial player in the effort to secure access to sensitive data and resources. Its adoption has been pivotal in thwarting a significant number of cyber attacks. MFA refers to the verification of a user's claimed identity by utilising a combination of two or more independent elements: something the user knows (knowledge), something the user has (possession), and something the user is (inherence). These elements, when employed together, provide a significantly higher level of security than when used independently. Geo-location represents an additional layer of security in MFA. It refers to the use of the geographical coordinates of the user to enhance the verification process. If a user's login activity is detected from a location inconsistent with their usual pattern, this can serve as an alert to potential security risks. The incorporation of geo-location algorithms adds robustness to MFA, enhancing the complexity of the authentication process and thereby increasing the difficulty for potential intruders.

2. Geo-location algorithms: principles and techniques of authentication

2.1. The Essence of Geo-Location

Before delving into the algorithms, it is important to understand geo-location as a concept. Geo-location refers to the identification or estimation of the real-world geographic location of an object, such as a mobile device or a computer terminal. In the context of MFA, geo-location can provide valuable data to verify whether a login attempt is legitimate or potentially malicious. Additional data can indeed improve the accuracy of many types of algorithms, including those used for geo-location. More data points often allow for a better understanding of patterns and behaviours, which can in turn make predictions or identifications more accurate, such as

leaked geo data that can be used to evaluate the legitimacy of a user.

2.2. Common Techniques of Geo-Location MFA

2.2.1. IP-Based Geo-Location

“IP geo-location is the process of determining the real geographic location of an electronic device connected to the Internet, by its global network address” (Ivanov et al., 2022). The IP address serves as a key identifier. When a device establishes a connection with a network, the service detects the device's IP address that is unique to that specific connection at that particular time. Once the IP address is identified, it is matched with a specific geographical location through a lookup process. This procedure refers to an extensive database, often maintained by third-party entities, that correlates IP addresses to physical locations. These databases are compiled from various sources, such as Internet Service Providers (ISPs) and extensive data collection efforts, to map the likely geographic origin of an IP address. The information returned from this lookup might include the country, state, city, and sometimes even more granular details like postal codes. “IP geo-location error depends on the country, population density, and type of network device and ranges from several tens of metres to hundreds of kilometres. For the same input data, the results of different IP services can vary significantly” (ibid). This is the easiest way to MFA a user through a combination of their password and IP. This method is easy and cheap to implement; however, IP-based geo-location is not always precise. “Geo-location databases can claim country-level accuracy, but certainly not city-level” (Poese et al., 2011). Accuracy can vary greatly depending on a number of factors, such as:

IP addresses that are dynamic: ISPs often dynamically assign IP addresses to devices from a pool, and the same IP address can be assigned to different devices at different times.

The use of VPNs and proxies: VPNs and proxy servers can mask the true IP address of a device and make it appear as though the device is located in a different place. This can skew the results of IP-based geo-location.

Mobile and satellite connections: Devices connected via mobile networks or satellite internet may appear to be located at the position of the relevant ground station or data centre, which could be far from the device's actual location. These can lead to compromises, such as saving multiple IP addresses as a verified way to MFA a user. This can be exploited. For example, a malicious actor who owns leaked data from a user can exploit it, like hijacking a web browser session data and masking the IP of the new session to a known one previously leaked.

2.2.2. GPS-Based Geo-Location

GPS-based geo-location operates by receiving signals transmitted by GPS satellites orbiting the Earth. These signals are processed by a GPS receiver in the device, and the time taken for these signals to reach the receiver from the satellites is used to calculate the distance to each satellite. With signals from at least three satellites, the GPS receiver can then use this information to triangulate the device's precise location on Earth. It can be even used in real-time. Many systems use GPS to locate their assets and their moving. “Such a system has been tested, and it showed excellent results that can be used for positioning and navigating a vehicle with an accuracy of 10 m” (Abbas et al., 2019). However, as with any technology, GPS-based geo-location comes with its own set of challenges. Firstly, it requires a device to have a built-in GPS receiver. While this is common in many modern devices, especially smartphones, it is not universal, and therefore, the applicability of this method may be limited. GPS-based location can eliminate some of the drawbacks of IP geo-location, such as:

High Accuracy*:* GPS-based geo-location is renowned for its precision, often providing location data accurate to within a few meters.

Real-time Data*:* GPS can offer real-time positioning, which can be crucial in certain applications, such as navigation or emergency response.

Independent of Internet Connection*:* Unlike IP-based geo-location, GPS doesn't rely on an Internet connection. As long as a device has a GPS receiver and can access satellite signals, it can determine its location.

Universal Coverage*:* GPS satellites encircle the globe, ensuring that signals are available anywhere on Earth with a clear view of the sky, making the system globally applicable.

This method is more expensive and it has more challenges that come with its own drawbacks:

Requires Special Hardware*:* Not every device is equipped with a GPS receiver. To utilise GPS-based geo-location, a device needs the necessary hardware. This can typically add more cost to the device's cost of manufacturing increasing its price. Utilising GPS for location tracking can be energy-intensive. Frequent use can drain a device's battery more rapidly compared to other location-determining methods.

Signal Interference*:* Physical obstacles like tall buildings, mountains, or dense forests can interfere with GPS signals. Urban environments, in particular, can introduce the "urban canyon" effect, where GPS signals are disrupted by high-rise structures.

Dependence on Satellite Availability*:* For optimal accuracy, a GPS receiver needs to connect with at least four satellites. While there are many satellites in the GPS constellation, local conditions or technical issues can sometimes limit satellite availability.

This method seems to be superior for geo-location but it also has flows that can lead in malicious ways to manipulate a geo-based login challenge. The GPS data can be manipulated with mock data. The high accuracy is a good characteristic, but the service that is verifying needs to verify your last validated location every time a device is moved 10m or 20m. This is why GPS geo-location cannot be used alone to determine the validity of a user.

2.2.3. Hybrid Methods

Hybrid geo-location systems typically combine data from both IP-based and GPS-based methodologies, along with potentially integrating other location sources, like Wi-Fi positioning or cell tower triangulation. The overarching aim is to produce a more holistic and precise representation of a device's location, balancing the swift and universal accessibility of IP-based methods with the precision of GPS. “The hybrid algorithm gives better results compared to the two algorithms

GPS and Universal Mobile Telecommunications System (UMTS)” (Saadane et al., 2021).

Several advantages underscore the increasing adoption of hybrid methods:

***Enhanced Accuracy**:* By drawing data from multiple sources, hybrid systems can often pinpoint a location with greater accuracy than any single method in isolation. For instance, in urban settings where the "urban canyon" effect may hinder GPS accuracy, Wi-Fi positioning can step in to refine the location estimate.

***Increased Reliability**:* Redundancy is a cornerstone of security, and by using a multi-pronged approach, hybrid systems ensure that if one method fails or provides ambiguous results, other methods can supplement or validate the data.

***Wider Applicability**:* While some devices might lack a GPS receiver, they might still be capable of Wi-Fi positioning or have access to IP-based location data. A hybrid system can tap into whatever sources are available, expanding its range of applicability.

However, integrating multiple methods does introduce new challenges:

***System Complexity**:* Merging data from several sources requires sophisticated algorithms to weigh, validate, and reconcile potentially conflicting data. This complexity can demand more advanced hardware and software solutions, potentially increasing costs.

***Higher Resource Usage**:* Tapping into multiple location data streams can lead to increased battery drain, especially if real-time, continuous tracking is needed.

***Privacy Concerns**:* With the system collecting data from various sources, there could be heightened concerns about user privacy. It becomes crucial to ensure data is anonymised and handled with the utmost care to maintain user trust.

In the context of multifactor authentication, hybrid geo-location methods present redundancy and versatility and can provide a robust layer of security, ensuring that even if one location method is compromised or inaccurate, the other can validate or correct the data. However, as with any complex system, careful design, regular maintenance, and a keen awareness of potential pitfalls are essential to harness their full potential.

3. CREATING MORE LAYERS FOR VERIFICATION OF GEOLOCATION MFA BY COMPARING PROBABLE MALICIOUS EXPLOITS FOUND IN LEAKED DATA

There are a lot of services, such as HaveIbeenPwned, that alert their users and clients for leaked information, but these services do not provide a way to check the raw data. It is up to the user or the client to recognise the leaked information.

Building such a service integrated into every system is impossible. Creating a global leaked database is the key to creating new methods to authenticate user geolocation, a global service that can provide cross-checking methods for validating potential exploits.

In this publication, we are going to use our own leaked data to bypass geo MFA and share the techniques used to achieve it. The service and system exploited will not be mentioned due to our own privacy.

3.1 Exploitation of IP-Based systems and preventing them from using the leaked databases

IP-based geo-location systems, while advantageous for their ease of implementation and ubiquity, are not without vulnerabilities. Exploiting these vulnerabilities can lead to a wide range of nefarious activities, ranging from misleading content delivery to serious security breaches. In a system where only geo-location is used to MFA a user, a lot of exploits can be used. In this case, the assumed system uses only passwords and geo-location for MFA. Such systems exist due to the inability or lack of hardware to use SMS authentication or another token system. This is common in IoT systems where hardware and software are limited. The IP geo-location method is cheap and easy to implement and commonly used in small IoT systems. Common ways to exploit IP based geolocation systems are:

IP Spoofing: One of the most common techniques used by attackers is IP spoofing, where they masquerade their device's IP address with another, often with the intent of deceiving systems about the device's true location or identity. In the case of the actor knowing the victim`s leaked IP addresses, this can lead to an easy bypass of the geo MFA.

In the case of our own leaked data, we used multiple IP addresses to log in to the service exploited in this case. Choosing potentially valid IP can be tricky, but by cross-referencing with other leaks from different sources a commonly used IP can be identified. Spoofing this IP from a Different location did not raise the chosen system alert.”The attack is generic in nature, i.e., all devices using vulnerable Wi-Fi-based Positioning Systems could potentially be exploited with the help of geo-tagged services that provide geo-location information publicly” (ibid).

VPN and Proxy Usage: Users or malicious entities might use VPNs or proxy servers to make their IP address appear as if it originated from a different geographic location. This can be done to bypass geo-restrictions, conduct location-based fraud, or obscure malicious activities.

In the case of our own leaked data, we used a VPN location based in our country and were able to log in from a different country without raising suspicion in the selected system.

There are more techniques to exploit IP-Based systems like IP Hopping or Man-in-the-Middle Attacks that can be used to achieve the same results. For this example, we only used the IP Spoofing and VPN/proxy method.

In this example, the systems that were exploited relied on databases with validated IP addresses. In order to prevent such behaviours, these systems can regularly check with a global leaked database for potential “harmful” IP addresses of the user and exclude them from the verified address list and promote new verification of this IP via another source, like email.

3.2 Exploitation of GPS-Based Geo-Location and preventing it from using the leaked databases

GPS-based geo-location, often used for its precision, is increasingly becoming popular for mobile applications. Most smartphones have built-in GPS and rely on them for geo MFA.

However, the system is not impermeable to threats and other ways to exploit it. In the case of phone geolocation, there is a threshold for the valid location. The most common ways to exploit GPS location are GPS Spoofing, GPS Jamming, Replay Attacks, and Data Interception.

The method exploited using our leaked GPS data was GPS Spoofing. The exploited service was only verifying the GPS data as MFA. From our own leaked GPS data, we were able to determine the most commonly used GPS location by cross-referencing it with other leaked databases. The exploited service did not raise any suspicion.

To prevent such exploitation, the valid location database can be cross-referenced for leaked data on a global database.

3.3. Exploitation of Hybrid Geo-Location and preventing it from using the leaked databases

Hybrid geo-location systems, combining the strengths of multiple location-determining techniques, such as IP-based, GPS-based, Wi-Fi positioning, and cell tower triangulation, aim to offer enhanced accuracy and reliability. The main weaknesses revolve around the complexity of such a system. The most common are:

Compound Vulnerabilities: One of the most evident challenges is that a hybrid system inherits the vulnerabilities of each individual technique it integrates. For instance, if it combines GPS and IP-based methods, it can be susceptible to both GPS spoofing and IP spoofing.

Increased Attack Surface: With multiple channels of data input, attackers have more entry points to exploit. They might jam or spoof one signal, while simultaneously launching a cyber-attack on another, complicating defence mechanisms.

Data Conflict and Confusion: In situations where different location methods return conflicting data, determining the true location can become challenging. Malicious actors can exploit this confusion, intentionally feeding one of the systems false data to obscure their actual location.

Chain Reaction Failures: A compromise in one geo-location method can sometimes lead to cascading failures in other systems, especially if they rely on each other for validation or calibration.

In the case of Hybrid Geo MFA, it is the same process of exploitation as exploiting its components one by one. This is enough for leaked data to both exploit IP-based and GPS-based MFA at the same time. In the case of my leaked data, I wasn't able to find an application suitable for testing.

4. Conclusions

Using leaked data to Bypass MFA is a niche that is not explored widely. Our test had great results with systems with low-security bypassing geo MFA. All the experiments were made on our own data and no laws were violated. To achieve this method on a big scale, malicious actors have to own a lot of leaked data. The use of leaked data is a treading and ever-evolving problem. To counter this, new improvement to the already existing leaked database monitoring service must be created. At the moment, this kind of cooperation is between security researchers working with leaked data and their clients in private. An open-source alternative, like the suggested global leaked database system, can be beneficial. To conclude the need for such a global system, it is necessary to use a big amount of data and a big enough service, like a social network, to experiment on.

Used literature:

Bao, James , Yen Tsui. 2005. *Fundamentals of Global Positioning System Receivers. A Software Approach*. John Wiley & Sons Inc, 352 p.

Célestin Matte, Jagdish Prasad Achara, Mathieu Cunche. 2015. *Device-to-Identity Linking Attack Using Targeted Wi-Fi Geolocation Spoofing*, Researchgate.

Kingsley-Hughes, Kathie. 2005. *Hacking GPS*. John Wiley & Sons Inc, 337 p.

Yinusa Olatunde Bamimore, Ayorinde Oduroye. 2023. *Moving Beyond Password-less Approach -Multi-Factor Authentication (MFA) in Cyber Space*. researchgate

Singh, Satinder Pal. 2011. *IP Geolocation in Metropolitan Areas*. University of Maryland, 97 p.

References:

Abbas, A. H., Habelalmateen, Mohammed I., Syukran Jurdi, L. Audah, N. A. M. Alduais. 2019. GPS-based location monitoring system with geo-fencing capabilities. - *Advances in Electrical and Electronic Engineering*, November 2019.

Ivanov, Mihail, Alexander Polunin. 2022. Improving the Accuracy of IP Geolocation Based on Public IP Geoservices Data. - *Informatics and Automation 21* (4), July 2022, 758-785.

Poese, Ingmar, Steve Uhlig, Dali Kaafar, Benoit Donnet, Bamba Gueye. 2011. IP Geo-location databases: Unreliable?. - *ACM SIGCOMM Computer Communication*, April 2011.

Saadane, Rachid, Mostafa Belkasmi, Mohammed El Koutbi, Mohamed Rachid Dadda. 2012. Evaluation of hybrid geo-location based on UMTS and GPS technologies. *Multimedia Computing and Systems (ICMCS), International Conference*, May 2012.