Digital Identity Challenge: The Security and Convenience Dilemma

André Ferraz[1], **Carlos Ferraz**[2]

[1]Incognia
2479 East Bayshore Road, Suite 150 – Palo Alto – CA 94303 – USA

[2]Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Av. Jornalista Aníbal Fernandes – 50.740-560 – Recife – PE – Brazil

andre@incognia.com, cagf@cin.ufpe.br

Abstract. *This paper argues that the essential pieces of an enduring digital identity should be privacy, security, and convenience. Authentication should be frictionless. In this sense, the core of the digital identity of the future will be created around location sensing techniques. Incognia proposes a solution to secure and frictionless authentication for mobile apps that is composed of five steps. Its proprietary technology called environment fingerprinting can identify location spoofing and precisely determine the device's actual location. Incognia has found that most mobile logins, sensitive transactions, and purchases occur at trusted locations. To date, 90% of mobile logins and 89% of mobile banking sessions happen at a trusted location. Experimental results show false-negative rates below 0.004% and a decrease of over 85% of account takeover attacks.*

1. Introduction

Passwords were created in the '60s [Morris and Thompson 1979] when computing resources became shareable, and authentication became a requirement to guarantee the integrity of the system's data. Since then, authentication to internet-connected services has relied primarily on passwords. As more and more internet-connected services became available to help people perform day to day activities, i.e., the pervasiveness of today's Internet services, remembering passwords became increasingly challenging, leading users to start reusing passwords – in [Das et al. 2014], the authors estimated that 41-53% of users reuse the same password across multiple sites.

With frequent data breaches, passwords became publicly available, enabling bad actors to take over online accounts with ease [Raza et al. 2012]. Credential stuffing [Rees-Pullman 2020] is a technique in which an online fraudster uses leaked credentials to try login into as many services as possible, taking over accounts at scale. Multi-factor authentication (MFA) [Ometov et al. 2018] was then introduced to add a new layer of security to the process. Out-of-band authentication [Naor et al. 2020], including SMS and email-based one-time passwords (OTPs), biometrics, physical keys, and digital signatures, are also used as authentication factors.

In the US alone, identity fraud losses in 2020 reached a total of $56 billion (USD), growing 230% YoY (16.9B in 2019). 43B was a consequence of social engineering scams when bad actors usually act as customer support representatives to rip off people's money. The other 13B is due to data leaks that enable fraudsters to use information like stolen

credit cards and identities. As our mobile phones start to substitute cash and plastic cards, the problem will intensify. By 2025 there will be 5B mobile internet users[1]. Each of us will have a few personal devices that will interact with hundreds or thousands of other devices daily. Within such a wild environment with so much volume, it will be easy for bad actors to hide, making today's security technologies look like kindergarten. As the Internet of Things (IoT) becomes ubiquitous, the consequences of cyberattacks become physical and increasingly more critical. On the other hand, it will be impossible to live this life having to type passwords, taking selfies, and scanning our fingerprints, putting convenience and privacy at stake. So,

> *How to build a digital identity that endures and enables consumers to live a life that benefits from technology without feeling constant fear?*

The essential pieces of an enduring digital identity should be: i) privacy, as we know that no technology is impenetrable and flawless. So the ultimate protection is to avoid any information related to the user's real-world identity as much as possible; ii) security, as nowadays the attack surface is too broad. Our data are already spread across databases of several websites, mobile apps, and people with whom we interacted. The digital identity of the future must use dynamic and real-time information to validate the user's identity and context [Dey 2001]; iii) convenience, as the user experience to sign-up and login to most internet services has become worse, requiring increasingly complex passwords and additional actions for authentication, including OTPs, biometrics and other high-friction authentication factors. That should be the exception, applied only in high-risk situations. Authentication should be frictionless, invisible.

The one thing that will never change is that no matter how connected we become, we will always be physically present somewhere and interact with whatever internet service through a physical device in the same place we are. Therefore, the core of the digital identity of the future will be created around location sensing techniques.

2. Proposed Solution

With the growth of mobile devices, passwords and multi-factor authentication became a pain point for the consumer due to the friction it adds to the authentication process. Mobile devices popularized new sensors that enable deep understanding of user's behavior, including location sensors, such as the GPS, Wi-Fi, cellular and Bluetooth, motion sensors, such as accelerometer, gyroscope, compass, among others. People carry their mobile phones almost 100% of the time, enabling their sensors to identify behavioral patterns from the users. These behavioral patterns are unique and change continuously, enabling more secure authentication. Location is the most relevant behavior, given it can be linked to attributes of the user's real-world identity, including address information. In addition, location can be associated with contextual information, enabling a deeper understanding of the user's behavioral patterns. And yet, precise location information allows identifying online fraud hubs, where bad actors use to operate.

A solution to the problem stated in this paper should answer the question

How should location fit into our mobile authentication strategy?

[1]Global Message Services (GMS). What will global mobile market look like by 2025?. https://www.gms-worldwide.com/blog/what-will-global-mobile-market-look-like-by-2025/

Mobile authentication is challenging because it often forces app developers to choose between security and user experience. Today, most fraudulent transactions originate on mobile devices, but at the same time, mobile users are much more likely to abandon apps due to friction.

On mobile, we have seen that location intelligence can contribute directly to reducing friction during the authentication process and can detect account takeover fraud with ease. For example, more than 90% of logins to financial services apps happen from a *trusted location*, a place that is frequently visited by the user like home and workplace. Then,

> *How can we build secure and frictionless authentication solutions for mobile apps?*

The answer is composed of five steps:

Step 1 - Identify the user's context. To identify the user's context, leverage the data collected from the user's session, including device, network, and location data. These three categories of context data provide the strongest signals for mobile authentication.

- **Device data:** Identifying the user's mobile device is the first important step in reasoning about the user's context. A known device with good history is the first sign of trust. One challenge with device intelligence is that fraudsters are increasingly using tools and techniques, such as mobile emulators, to mimic devices. Any fraud detection system should be able to detect mobile emulation and other device manipulation techniques.
- **Network data:** Network data is another type of data that is relevant to a user's context. If the user is on a mobile network, the phone number can be frictionlessly verified using data provided by phone number intelligence companies, and if the user is on wifi, fraudulent activities can be mapped to specific networks.
- **Location data:** Finally, the mobile device's location data should be used to identify if the user's current location is part of their normal behavior pattern. IP data is the least precise form of location data but can be sufficient depending on the use case. GPS, when available, is more effective. The main issue with IP location and GPS is that fraudsters use tools like VPNs, proxies, GPS spoofing apps, and mobile emulators to mock their location easily.

Step 2 - Define levels of risk based on the context data. Device, network, and location data should be dynamically weighted as you verify each piece of data's efficiency at differentiating legitimate customers from fraudsters. Here are some examples of different weighting combinations:

- **Low risk:** known device, trusted location, known network;
- **Medium risk:** new device, trusted location, known network;
- **High risk:** known device, not trusted location, malicious network.

Step 3 - Weigh the level of contextual risk with the level of transaction risk. That should be based on the financial and reputational risk of each transaction. Examples of transactions that should be weighted differently are money transfer or payment, changing password, adding a credit card, adding payees, and withdrawal from savings account or investments.

Step 4 - Define the level of friction introduced for each risk level. Keeping friction low is essential. A bad review in the App Store or Reddit could be more damaging than fraud's direct financial costs.

- **Low risk:** frictionless authentication, or standard device authentication;
- **Medium risk:** full authentication – whatever you currently use: pin, passwords, 2FA, biometrics etc;
- **High risk:** block device.

Step 5 - Measure and optimize with context data. After implementing the first version of weighting for the context data, you should continuously monitor the results and adapt accordingly. Every business is different and attracts different types of users and fraudsters. Your judgment should be based on the level of risk and friction, as well as the precision of the solution and its cost.

Fraudsters are frequently retooling and changing their techniques. Any fraud prevention solution needs to evolve similarly to detect emulator attacks, location spoofing, SIM swaps, BOTs, and many others.

3. Implementation

By analyzing the continuous stream of sensor information from the mobile device, using a proprietary technology called *environment fingerprinting*, Incognia[2] can identify location spoofing and precisely determine the device's actual location. Also, Incognia learns the patterns unique to each device and starts anonymously attributing context to each device and location, such as the device owner's home location and work location.

With this new layer of location and device intelligence, mobile apps can now link user's devices to user-declared credentials. When a user starts an identity verification process, the user-supplied home address is compared to the device's location behavioral signals. If there is a good match, it means that there is a close link between the device and the user's identity. Therefore, the risk of accepting that customer is lower. Another exciting aspect is that this technique does not require the identity verification solution to store any direct personal identification, including the address, that is discarded after its verification.

3.1. Architecture

The system is provided in the form of a Mobile SDK to be integrated with the customer's app, allowing it to continuously stream sensor information from the mobile device. The main modules of the system's architecture are:

- the **Mobile SDK**;
- **Visit detection**: sensor data are analyzed to identify if the user has exited or entered a new environment. A change in the Wifi scans and GPS signals can indicate this displacement;
- **Frequent locations**: if the environment is part of the user's behavioral graph or closely related to it, it indicates a higher probability of the transaction being legitimate;

[2]https://www.incognia.com/

- **Visit classification**: the signals are analyzed to understand if the current environment is known based on the GPS and Wifi readings;
- **Evidence list**: additional evidence is provided to support the risk assessment, regarding
    1. device integrity: location spoofing, vpn/proxy, root, emulator detection;
    2. device reputation;
    3. location reputation;
    4. travel feasibility;
    5. address verification and home/work inference.

3.2. Privacy

To ensure user privacy, Incognia does not link the location data to any direct personal identifier and retains the least possible amount of location data to perform its digital identity services. Additionally, Incognia invests in proprietary and open-source data anonymization technologies with a particular focus on location privacy.

4. Experimental Results

Results shown here have been accomplished through the analysis of data gathered from over 60 million mobile devices.

4.1. Identifying account takeover by using location behavior

With the rapid increase in the use of mobile apps, more people than ever are managing their finances, paying bills, and shopping from their mobile devices. The reason is simple: convenience. However, as more transactions originate on mobile, the channel has become a bigger target for account takeover (ATO) attacks. While many existing technologies help reduce the threat of account takeovers, fraudsters continue to employ new techniques, and ATO remains a growing problem.

Location behavior gathered from sensors on mobile phones is now being leveraged to significantly mitigate ATO attacks on mobile. Incognia's location technology works in the background, passively building a user's *location behavioral profile*. As it collects data, Incognia's platform learns which locations are trusted for each user and assigns context to them, like home or work. This location profile also enables Incognia to differentiate legitimate travel or device changes from an account takeover.

Location offers a level of resolution and continuity that enables anomalies to be easily detected. These anomalies include inconsistent velocity patterns or radical changes of location behavior. This approach not only protects mobile accounts but it reduces unnecessary friction for legitimate users.

4.2. Location information is one piece of data that fraudsters always conceal or spoof

Incognia has found that most mobile logins, sensitive transactions, and purchases occur at trusted locations, and there is a nearly perfect correlation between mobile transactions initiated at a trusted location and low risk. For example, to date, Incognia has found that 90% of mobile logins and 89% of mobile banking sessions happen at a trusted location. That shows that evaluating whether a user is at a trusted location is a highly precise way

of measuring risk. We routinely see false-negative rates below 0.004% and a decrease of over 85% of account takeover attacks.

Fraudsters are continually retooling and evolving their techniques, and location intelligence is helping companies stay ahead of the game. Additionally, Incognia's fast-growing mobile network, including tens of millions of devices, creates a consortium of fraud data, enabling companies across verticals to work together to defeat fraud.

5. Conclusion

In this paper, we have seen that the core of the digital identity of the future will be created around location sensing techniques for three reasons: 1) Privacy: our location behaviors are unique as our fingerprints, so it is strong enough to authenticate us without requiring our real-world identity or biometric data. Location data can be very sensitive, so it requires state-of-the-art anonymization technology and continuous improvement. Authenticating with data that are not directly linkable to our real-world identity is an essential step towards the future's digital identity; 2) Security: location behaviors are inherently dynamic but also have some interesting additional properties. The first is its uniqueness: there can't be two or more of us at a different location simultaneously. Location behaviors are also associated with places that are not publicly accessible such as your home and workplace. Hence, a fraudster has more difficulty entering it to impersonate your digital identity – mobile sensors provide behavioral signals that are hard to mimic; 3) Convenience: location behaviors are just our standard behaviors. We already carry our digital tokens with us at all times, so no action is required to authenticate using location. That is the ultimate level of convenience.

The ideal solution would be to transform our anonymized physical behavior into dynamic authentication tokens for the Internet of Things, enabling friction-free, private, and secure communication with every device.

References

Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wang, X. (2014). The tangled web of password reuse. In *NDSS*, volume 14, pages 23–26.

Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, 5(1):4–7.

Morris, R. and Thompson, K. (1979). Password security: A case history. *Communications of the ACM*, 22(11):594–597.

Naor, M., Rotem, L., and Segev, G. (2020). The security of lazy users in out-of-band authentication. *ACM Trans. Priv. Secur.*, 23(2).

Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., and Koucheryavy, Y. (2018). Multi-factor authentication: A survey. *Cryptography*, 2(1).

Raza, M., Iqbal, M., Sharif, M., and Haider, W. (2012). A survey of password attacks and comparative analysis on methods for secure authentication. *World Applied Sciences Journal*, 19(4):439–444.

Rees-Pullman, S. (2020). Is credential stuffing the new phishing? *Computer Fraud & Security*, 2020(7):16–19.