# Security Evalution of Twitter

Adebukola Ademola 001184659, Tristan Read 001151378,
Tejesh Ramesh Bawa 00113097-8

## I. Introduction

Twitter has been at the forefront of the data breach movement. It has been a hotbed for security breaches, especially in recent years. In fact, many of these breaches can be traced back to Twitter. This is because Twitter has an open API that allows users to access and manipulate data on the site through third-party applications.
The reason this incident is interesting is because it shows how vulnerable online services can be to attacks. It also raises concerns about how companies should handle sensitive data in the future. This problem has implications for both the public and private sectors, as well as for individuals who consider whether they should use these services at all.
The data breach made it possible for the attackers to gain access to private messages and other sensitive information that users thought was secure. The breach exposed users' locations, there they were using their phones, which could help criminals identify where they live or work.
This incident highlights several key aspects of social media networks: first, users do not always have control over what happens with their information; second, there is no standard way of reporting problems; third, even if users do report problems, there is no guarantee those problems will be addresses; fourth, this lack of transparency makes it difficult for users to understand how their data is being used.
Measuring this as a performance metric, it would help to understand whether Twitter is doing enough to protect it users' privacy from their parties who could potentially access their data without their knowledge.

## II. Analysis and Discussion

This section will discuss various defensive measures that large co-operates like Twitter can take to improve their chances of survival in mitigating attacks before and after the attack. Furthermore, this section will analyse different attacks Twitter might face, and how attacker exploit vulnerabilities of these big co-operate giants.

### Defence

Researchers have undertaken substantial study on defence for information security. Typical information security systems (e.g., authentication, access control, information encryption, intrusion detection system, vulnerability scanning, and virus protection) had already offered a certain level of security, however with the evolution of diversification attacks, traditional defence has become insufficient. The current defence systems are insufficient to prevent many forms of assaults (Zheng et al., 2021).
Mitigating cyber security threats in an organisation can be difficult. This would be particularly true if the organisation has transitioned to remote working and now have less supervision over employee behaviour and networking devices. A comprehensive approach must include your whole IT infrastructure and be based on frequent risk assessments. Cyber-attacks may cost businesses billions of pounds and create significant harm. Organizations affected may lose sensitive data, pay penalties, and suffer reputational harm. Effective information security management must originate from the top down. A strong cyber security culture, backed up by frequent training, will guarantee that every employee understands cyber security is their duty. Good security and efficient working habits must coexist (IT Governance, 2016).
Cyber-attacks based on undocumented system vulnerabilities and backdoors remain the most serious danger to communication networks. The unpredictability of vulnerabilities, along with the limitations of recognised defence measures, forces administrators to adapt defence tactics and invent defence mechanisms in order to reverse the passive condition of being vulnerable to assaults and difficult to implement in cyber security (Zheng et al., 2021).

### SQL Injection

One possible attack that could be used to cause the data breach is SQL injection. SQL injection is when an attacker sends some arbitrary code as part of a payload that will be run on an SQL database, if the code that then parses this request doesn't serialize the data correctly, the attacker could have their arbitrary code be run on the server. This type of attack can be used to get almost any data from the database that it is being executed on. The reason that an SQL injection attack could be viable in this scenario is that because twitter uses a public API that allows users to access a lot of data freely. This means that an attacker could send an SQL query as part of an API request

that contains some arbitrary code that will be run on the database. This could be used to get sensitive or private information.

One of the most simple ways to perform an SQL injection attack is to prematurely end a query, this can be done by manually by adding a semi-colon to the end of a payload, which would invalidate the query that should be run and after this semi-colon the attacker can enter their own code.

To prevent this type of attack, the programmers can use various methods of protecting against SQL injection attacks.

Most programming languages have built in features that can escape SQL strings. String escaping will escape any special characters in the string that could be used to modify a query. Programmers could also manually filter out statements that look like or match known SQL injection attacks, this could in some cases cause legitimate data to be dropped from the user's request, however depending on the criticality of the table being written to, this may be a necessary safety measure.

Developers can also use "prepared statements" which is a way of sanitizing inputs to pre-defined fields. This way only specific data can be entered to set fields, this way only known data is inserted into the query and should prevent arbitrary attack data being executed on the SQL query.

For dynamic SQL requests, where users may request customised data, the programmers can decide to use the "Enforce Least Privilege" principle. This principle states that the account used to access data on the database should only have the permissions that is required for the task that it is meant for, for example, not allowing the modification or reading of a user table.

The above methods wouldn't take too much time to modify and update the existing infrastructure for, meaning that potential downtime would be minimal. This is critical for a business as big as Twitter as downtime could negatively impact their customer base, and not implementing these changes would leave a bad view on the company if they were to suffer a larger breach.

**Social Engineering**

The advancement in cyber security has been tremendous in recent years as society has grown more aware of their online security and how their data is being utilized and safeguarded by large organizations such as Twitter, Meta, and others. As explained by (Mouton, Leenen and Venter, 2016) security mechanisms to protect sensitive information become more effective, humans remain vulnerable to manipulation, and so the human factor remains a weak link. A social engineering attack exploits this vulnerability by utilizing appropriate manipulation tactics to obtain sensitive information. One of the important ways to tackle this problem is by Improving Awareness of Social Engineering.

A strategy can be formed to train an internal staff person who would then teach other employees about security awareness through regular in-house training. External trainers, on the other hand, can be engaged for the same reason. The workshops mustn't be excessively long; rather, they must be presented in small chunks with periodic pauses. The lesson will be quickly received by the group, so they will not experience training strain as a result. Another crucial consideration is that the sessions should not include technical terminology. Employees who do not work in a technical role are not expected to grasp how a firewall works or how malware containment tools work. The training must be provided in basic, understandable language with strategic targets and an emphasis on controlling and mitigating Social Engineering (Saleem and Hammoudeh, 2018).

The intranet of the corporation may be quite useful in promoting security awareness campaigns. A corporation, for example, can implement a security course developed locally or externally by authorized experts and promote the program as a learning guide in a prominent part of the intranet. Managers must then push employees to study the material regularly to ensure that the knowledge is ingrained in their memory. The intranet is also a useful place to distribute security alerts to employees about current security threats, along with guidance on how to deal with the danger and who to report the event to (Saleem and Hammoudeh, 2018).

Screensavers will play a crucial role in raising staff security awareness. It can also be used to show brief reminders on issues such as password security, prohibiting tailgating, confronting anybody who does not have a business badge/pass, notifying any strange activities to departments concerned, and so on (Saleem and Hammoudeh, 2018).

Displaying colourful posters with large typography may be a useful eye-catcher. Posting concise and focused messaging on security risks affecting the organization may be a useful technique for raising staff awareness. General security reminders on posters must be cycled regularly, allowing employees to absorb different security messages with simplicity and convenience. Posters containing more significant and particular reminders, on the other hand, might be posted in a noticeable area of the office (Saleem and Hammoudeh, 2018).

Concise and straightforward reminders may also be sent to employees via printed handouts. In circumstances where there is no employee intranet or other tools that are available, this might be a cost-effective way for keeping employees aware of the hazards connected with Social Engineering. Managers might also set up a system in which these reminders are circulated across the office together with an employee name list and date. As a result, anyone who has gone through and comprehended the information may sign the form admitting that they have reviewed the security awareness reminders, while those who have not can be reminded again (Saleem and Hammoudeh, 2018).

**Attack**

Social media platform offers tools that improve the efficiency of virtual socialisation in the global world. It helps to facilitate seamless interaction and communication amongst people. Omolara et al, 2019.

Since users can freely share content on this dynamic platform (Twitter), Hacker' activities on the platform are on the increase and hackers are seriously utilising it for malicious intentions.

Attacks can be categorised into two:

a) Un-targeted attacks

In un-targeted attacks, Attackers aim for as many users, services, or devices as they can. Since there will be numerous vulnerable devices or services, they don't care who the victim is. They do this by employing strategies that profit from the Internet's accessibility, such as:

- Water holing is the practise of creating a fake website or compromising a real one to take advantage of users who are visiting it.
- Phishing is the practise of sending emails to large numbers of individuals asking for sensitive information (such as bank details) or enticing them to visit a false website.
- ransomware, which may spread malware that encrypts discs and demands money in exchange

b) Targeted attacks

An organisation is singled out in a targeted attack because the attacker has a particular interest in the company or has been paid to do so. Because it has been precisely designed to target the company's systems, procedures, or workers, both in the office and occasionally at home, a targeted attack frequently causes more harm than an untargeted one. The following are examples of targeted attacks:

- Malware Attack & spear phishing: which involves sending targeted recipients' emails that may include an attachment containing malicious software or a link that downloads malicious software.
- Deploying Botnet

**Malware Attack**

To spread malware on Twitter, as an attacker, I can encode the malware site as a short URL

A compromised trusted account is used to send out tweets because they contain short URL (Twitter is limited characters in a massage) so the user does not find it suspicious. However, I can disguise a shortened URL, so the user is not suspicious of the link. Then i can download malware to the user' endpoint.

I can disseminate the information using two approaches

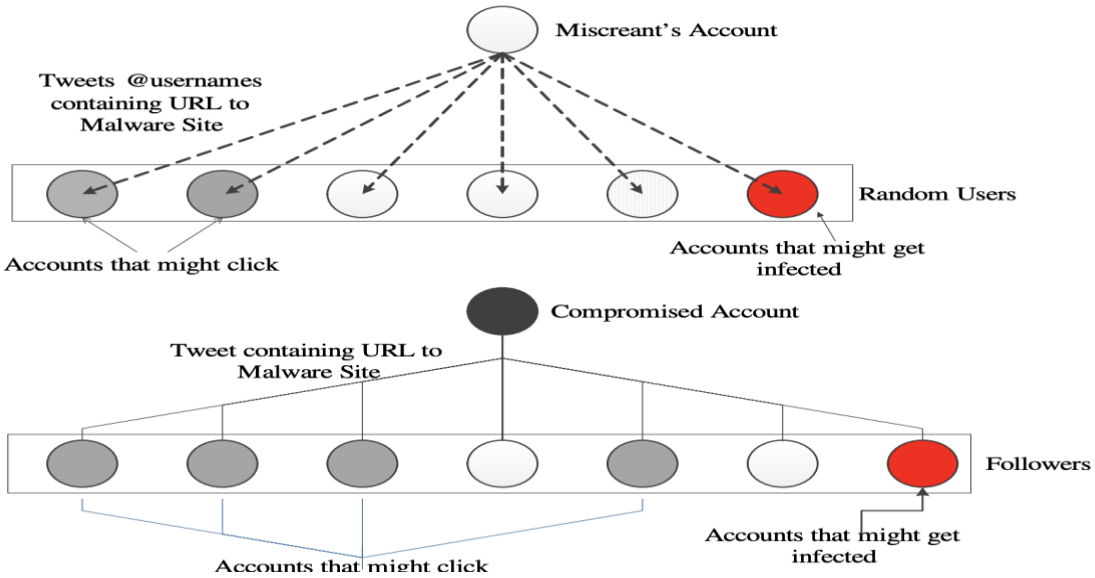
a) Using lots of @username in tweet and assume some users click on the link
b) Control the user' account through compromise, post tweet to the followers.

As an attacker, I inject Javascript code into a user's twitter account page. The malware collects the twitter user' authentication token that give authorisation for call to Twitter' API then i can post malicious tweet on behalf of the user.

Sample of tweet posted by malware using a user' account

a) The new king (king Charles 111) will earn more than his mother the late Queen, check his salary by clicking the link.
b) Watch Lizzo' BET award performs of her Hit song.

The strategic attack is demonstrated in the diagram below:




The first picture demonstrates the method of sending targeted messages to users while the second demonstrate the procedures of using a hijacked account to send tweets to followers. This attack is like Koobface Botnet where the attacker tries to trick the users into visiting a malware site and coarse them into downloading the malware under the guise of updating their software or flash drive. This type of attack can used on many users at the same time.

**An Advanced Attack on Twitter Using Clickjacking Technology**

Robert Hansen and Jeremy Grossman were the first to introduce clickjacking attack in 2008. A malicious web page is created to deceive the user into making inadvertent click that is useful to the attacker. It spreads worm, steal password and cookies. It can delete personal email, send spam. (Clickjacking for Shells, 2011). This use of sophisticated attack makes use of twitter' user follower model. There is need for the tweet to be retweeted by the user clicking a link.

As an attacker, I can insert embedded script on a twitter link that can be executed without the user' knowing about it. It can be a button on a link attached to a tweet (shown to user to perform a different purpose). When the user clicks on the button, this will run a malicious script from another sire.

According to Agam & Joab 2011, An attacker can also use opaque layer which trick a user into clicking a link that somewhat leads a user into somewhat a harmless page. This is another form of attack that can used on twitter. I can then hijack when the user clicks to the other website.

A link can be sent to the user' email which is like twitter' (a mirror site) making them assume they are still on their account, but I can gain insight to their personal information and login credentials.

The clickjacking attack can also exploit the weakness of short-URL providers which encodes a new short-URL for different users. This attack has the benefits of propagating down the Twitter tree, with the additional benefits of making it difficult for Twitter to analyse the different short-URLs due to the amount of information generated and traversing through the network.

**Social Engineering Attack with SQL Injection**

To launch a successful cyberattack, the attacker must first identify and exploit system vulnerabilities. As social engineering is the most vulnerable weakness in every organisation, large or small. This vulnerability may be exploited using attacks such as baiting, spear phishing, and so on. However, social engineering may be a tough strategy since substantial study on the target is required, as a lack of knowledge on the target might lead to failure and shut down the attack entirely.

This attack will mainly be focused on Social Engineering as its vulnerability, and this vulnerability will be exploited by the use of phishing attacks on the employees who are new to the system and bypass any existing MFA (Multi-factor authentication).

Following the Covid-19 epidemic, many individuals found that working from home is a viable option. However, as noted by (Irwin, 2021), a remote workforce is significantly more vulnerable to threats in the absence of security safeguards provided by office systems, such as firewalls and blacklisted IP addresses, among other things. Employees' cloud files, emails, and third-party services are all vulnerable, and the fact that so much data is exchanged digitally simply amplifies the attack surface.

A remote workforce poses several risks since employees rely on personal home networks and their personal equipment to execute tasks. They may experience technical difficulties and might have to contact the IT department. According to (Irwin, 2021), 70% of remote employees reported IT issues during the epidemic, with 54% being required to wait for up to three hours. This demonstrates how many individuals suffer with technical issues, as well as the reason why any attacker would desire to target this element of the system.

An attacker can target this by posing as a member of the IT staff and acting as though they are responding to a reported problem with the technical problems the employees are experiencing. As an attacker can trick the employee into believing that there is a technical problem with the organization's VPN and that it must be rebooted via their log-in credentials.

The attackers may attempt to deceive employees into visiting a phishing website that looks similar to the official organisation VPN website and has a similar domain name. The attacker inserts information into the actual website while the employee enters their credentials into the phishing website. Because this is a fraudulent log-in, an MFA notification will be triggered, and a few workers will likely authenticate themselves.

This is where social engineering plays a role because attackers must use confidential information about workers to persuade them that perhaps the attacker is authentic and hence trustworthy. An effective social engineering attack allows the attacker to penetrate into the internal network of these large multinational corporations such as Twitter.

The initial compromise enables the attackers to explore internal websites and gain a better understanding of the organization's information network. This flaw may also allow the attacker to access the intranet websites and read the organization's classified data. To gain more understanding of how the internal system work the attacker can use SQL Injection to understand how their database is being stored. This will allow the attacker to gain access of unlimited accounts and personal data of people around the world who use the organisations services.

To overcome the organization's protective coding, the attacker can utilise a form of SQL Injection known as Alternate Encodings. Through its injected strings, the attacker may well have organised other types of encoding, such as hexadecimal, ASCII, or Unicode character encoding. As a result, scanning and detecting procedures are ineffective against it (w3resource, August 19 2022).

Illegal/Logically Incorrect Queries are another sort of SQL Injection. This allows the attacker to learn about the kind and design of the Web Application's back-end database. The attack could be viewed as the initial phases of subsequent attacks. When an invalid query is made to a database, certain software applications provide the default error code, which the attacker exploits. They insert code into susceptible or injectable parameters, causing syntax, type conversion, or logical errors. Type error may be used to determine the data kinds of certain columns (w3resource, August 19 2022).

The attacker can also utilise the Inference SQL Injection technique, which is applied to well-secured database that do not produce any useable feedback or informative error messages. Typically, the assault is constructed in the manner of a true false assertion. Following the discovery of the vulnerable parameter, the attacker injects numerous conditions (to determine whether they are true or not) through query and carefully observes the situation. If the assertion is true, the page will continue to function normally. If false, the page operates drastically differently than it typically would. Blind Injection is the name given to this form of injection. Another sort of inference attack is known as a Time Attack. In this approach, an attacker creates a conditional statement, injects it through the susceptible parameter, and collects information based on time delays in the database's response (w3resource, August 19 2022).

All of these SQL Injection techniques can help the attacker obtain a better grasp of the system database and internal networks. This gives the attacker the ability to go farther and hijack certain important accounts and publish whatever they want, such as hacking accounts on July 15, 2020, and tweeting out a double bitcoin plan. In addition, the attacker can also request a ransomware from Twitter.

**III. Conclusions**

Big and internationally prominent social media platforms, such as Twitter, largely self-regulate. There is no specific state or federal regulator with the authority to oversee sufficient cybersecurity policies to avoid fraud,

misinformation, and other systemic dangers to social media behemoths. A regulator with relevant knowledge should be charged with monitoring and overseeing these firms' cybersecurity (Department of Financial Services, 2020).
The Twitter Hack highlights the dangers that world faces when fundamentally powerful organisations are permitted to govern themselves. It is critical for all of us—consumers, voters, government, and industry—to protect vital social media against misuse (Department of Financial Services, 2020). It is also important to recognize that newer and improved defensive methods will be developed to protect against threats; yet attackers will develop newer attacks to exploit flaws in these new defensive techniques.
It is essential to consider that in the cyber security perspective, defenders have no notion where the next assault will come from and must anticipate and predict which vulnerabilities the attackers will exploit. The attackers use a direct approach in attacking the system, whereas the defenders must develop a hypothesized secure system to counter the attackers.

**References**


Department of Financial Services. (2020). *Twitter Investigation Report*. [online] Available at: https://www.dfs.ny.gov/Twitter_Report [Accessed 15 Nov. 2022].
Irwin, L. (2021). *The Cyber Security Risks of Working From Home.* [online] IT Governance UK Blog. Available at: https://www.itgovernance.co.uk/blog/the-cyber-security-risks-of-working-from-home [Accessed 12 Nov. 2022].
Mouton, F., Leenen, L. and Venter, H.S. (2016). Social engineering attack examples, templates and scenarios. *Computers & Security*, [online] 59, pp.186–209. doi:10.1016/j.cose.2016.03.004.
OWASP (2013). SQL Injection | OWASP. [online] Owasp. Available at: https://owasp.org/www-community/attacks/SQL_Injection.
Saleem, J. and Hammoudeh, M. (2018). (PDF) *Defense Methods Against Social Engineering Attacks*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/319097404_*Defense_Methods_Against_Social_Engineering_Attacks* [Accessed 15 Nov. 2022].
w3resource. (2019). *SQL Injection Tutorial - w3resource*. [online] Available at: https://www.w3resource.com/sql/sql-injection/sql-injection.php [Accessed 13 Nov. 2022].
What Is SQL Injection? | Cloudflare UK. (n.d.). *Cloudflare*. [online] Available at: https://www.cloudflare.com/en-gb/learning/security/threats/sql-injection/.
Zheng, Y., Li, Z., Xu, X. and Zhao, Q. (2021). Dynamic defenses in cyber security: Techniques, methods and challenges. *Digital Communications and Networks*, [online] 8(4), pp.422–435. doi:10.1016/j.dcan.2021.07.006.
Agam Shah, Joab Jackson, 2011. Doj Charges Seven in Massive Clickjacking Scheme. *Network World IDG News Service*
A. E. Omolara, A. Jantan, O. I. Abiodun, V. Dada, H. Arshad, and E. Emmanuel, "A Deception, 2019. *Model Robust to Eavesdropping over Communication for Social Network Systems*," no. Im, pp. 1–21, 2019
Clickjacking for Shells, 2011. OWASP Wellington, New Zealand Chapter Meeting.
Clickjacking, 2012. *The Open Web Application Security Project*.
Sandler, R. (2020, July 22). *Twitter Says Hackers Accessed Direct Messages From 36 Users, Including One Dutch Elected Official*. *Forbes.*
IT Governance (2016). *What is Cyber Security? | IT Governance UK.* [online] Itgovernance.co.uk. Available at: https://www.itgovernance.co.uk/what-is-cybersecurity.