Advanced Research on Information Systems Security

ISSN: 2795-4609 | ISSN: 2795-4560
Print & Online

Phishing in Web 3.0: Opportunities for the Attackers, Challenges for the Defenders

Mohamed El Atoubi*

Georgia Institute of Technology, Atlanta, GA 30332, USA

Email: matoubi3@gatech.edu

Abstract

The emergence of Web 3.0 technologies is undoubtedly shifting the phishing threat landscape by empowering traditional attack vectors and enforcing attackers to adapt and leverage new techniques. Offering new opportunities for the attackers to exploit and adding new challenges for cybersecurity professionals. This paper aims to exhibit the evolution of phishing attacks in Web 3.0, the different techniques and novel technologies cybercriminals leverage for scaling phishing campaigns and explores the challenges and promising solutions defenders can adopt to narrow the strategic gap in this emerging component of the attack vector. The paper commences by summarizing the different wide web generations until it reaches Web 3.0 and starts exploring its architectural distinguishing features. Afterwards, it deductively treats the convoluting threats of the new web technology as an entrance to focus on phishing attacks by discussing the concept of phishing from various angles while shedding light on emerging threats from phishing 3.0 and potential solutions that can make a substantial impact.

Keywords: Web 3.0; Cybersecurity; Phishing; AI

Citation: M. Elatoubi, “Phishing in Web 3.0: Opportunities for the Attackers, Challenges for the Defenders”, ARIS2-Journal, vol. 3, no. 2, pp. 11–25, Dec. 2023.

DOI: https://doi.org/10.56394/aris2.v3i2.35

* Corresponding author. Email address: matoubi3@gatech.edu

1. Introduction

According to 2021 statistics from market.us, the number of internet users has continued its growth trend surpassing 5 billion users. The surprise is in the number of Web 3.0 developers which registered the highest number of users in one year around 34,391 [1]. This trend is likely to continue in the future. The wide spread of Web 3.0 technologies will bring many opportunities and create new job positions and revive existing ones. In a symmetrical manner, new threats are emerging with negative impacts on society and the economy itself. The focus here is particularly on cybercrimes. Attackers are leveraging the novel technologies of the Web revolution by shifting the traditional threat landscape by renovating their tactical arsenal to remain at a strategic vantage point. While all cybersecurity threats are of major importance for study. This paper focuses on phishing attacks in Web 3.0 as the literature seems to illustrate an incomprehensive image of these types of attacks in the third-generation wide web.

More than 90% of cyberattacks start with Phishing as a point of entry according to CISA [2]. Phishing attacks have always been the most efficient offensive tools in the hands of cybercriminals as they are able to bypass security checks manipulating and exploiting the human element. Today with the emergence of advanced technologies such the Web 3.0, the threat from these attacks is surely aggravating. The second section of this article studies the evolution of Web 3.0, its advantages, and proposes a polished technology architecture adapted to safeguard against phishing attacks. The third part discusses the convolution of inherent and novel threats that are gaining momentum with the upgrading transformation from Web 2.0 to 3.0. The fourth section drills down into how social engineering is applied to generate Phishing attacks in targeted campaigns. It also covers novel social engineering attacks in Web 3.0 and the use of AI to scale Phishing campaigns. The fifth section proposes an analytical method to detect Phishing 3.0 attacks by exposing the marking threat indicators that can be easily assessed by non-technical users and proposes actions to mitigate the risk. Finally, fundamental solutions gathered from trending literature including novel contributions are refined and proposed in efforts of tackling a new wave of advanced cyberattack campaigns.

2. The Web 3.0

2.1. Evolution of the Web 3.0

Web 3.0 is an advanced iteration that has emerged after the spread of Web 2.0. The Web 3.0 term emerged in the year 2014 by Ethereum cryptocurrency founders [3] but its components had already existed years ago. The necessity of a novel version of the web has precisely reached a non-turning point after the different weaknesses and issues became so pervasive with precedents versions of the world wide web. The monetization of personal data has launched a widespread race between a few tech entities to monopolize users' data in cyberspace. These "data-opolies" have shaped the Web 2.0 and are already moderately investing in the development of Web 3.0 [4]. Therefore, the motives behind developing a new version of the web are not only technical but often economical as well as with any invention in the modern world.

The most important concept that web 3.0 is built on and it is heavily discussed in the literature is the decentralization of the blockchain technology which is the network in Web 3.0. This feature ensures data privacy to a higher degree by preserving data ownership and eliminating dependence on third parties [5], a major feature not available in Web 2.0 and its predecessor. Enhanced privacy is a great perk that comes with Web 3.0; however, it is not the main goal of Web 3.0. Web 3.0 goal is to leverage the recent advancements made in Artificial Intelligence, Cloud Computing, Virtual/Augmented Reality, Social Computing and the Blockchain to open new horizons and provide unique experiences such as the Semantic Web allowing machines to be able to meaningfully understand human shared content using metadata embedded within the Web data [6]. One goal of Web 3.0 is achieving a hyperconnectivity experience to future users by connecting everyone and everything. The web evolution in general is to be taught of as an abstract wave full of new concepts, experiences and opportunities limited by the evolution of the infrastructure running the Web technology. The next section will discuss the architecture of Web 3.0 and its key strengths.

2.2. Web 3.0 architecture and strengths

In the quest of producing a foundational architecture for the Web 3.0, many researchers have attempted to mimic the OSI model layered architecture. Several papers brought different suggestions some were considered to be novel such as Jacksi who proposed the introduction of IPsec (IP Security) in the network layer [7], Zheng suggested the Ethereum blockchain as a separate layer [8], Yuqing [5] who expanded the model to six layers involving new telecommunications technologies such as 5G in the first layer called the infrastructure, the layer that follows is the network layer and the focus was based on open platforms like the Ethereum Virtual Machine, the third layer is the protocol layer and englobes different technologies like plasma protocol and the blockchain, the equipment layer is the layer just above that strives to immerse the users in hyped experiences using Augmented and virtual reality gadgets, the fifth layer is the software layer containing Artificial Intelligence centric software and operating systems, the final top layer is the interaction layer which is nothing but the decentralized applications layer. It appears that this architecture is a novel vision which combines a multidimensional stack of technologies and aligned with every Web 3.0 feature. The suggested architecture inspires from such models and expands to new horizons such as incorporating security and privacy in new englobing configurations. The following diagram illustrates how Web 3.0 architecture is visualized with added security and privacy preservation:

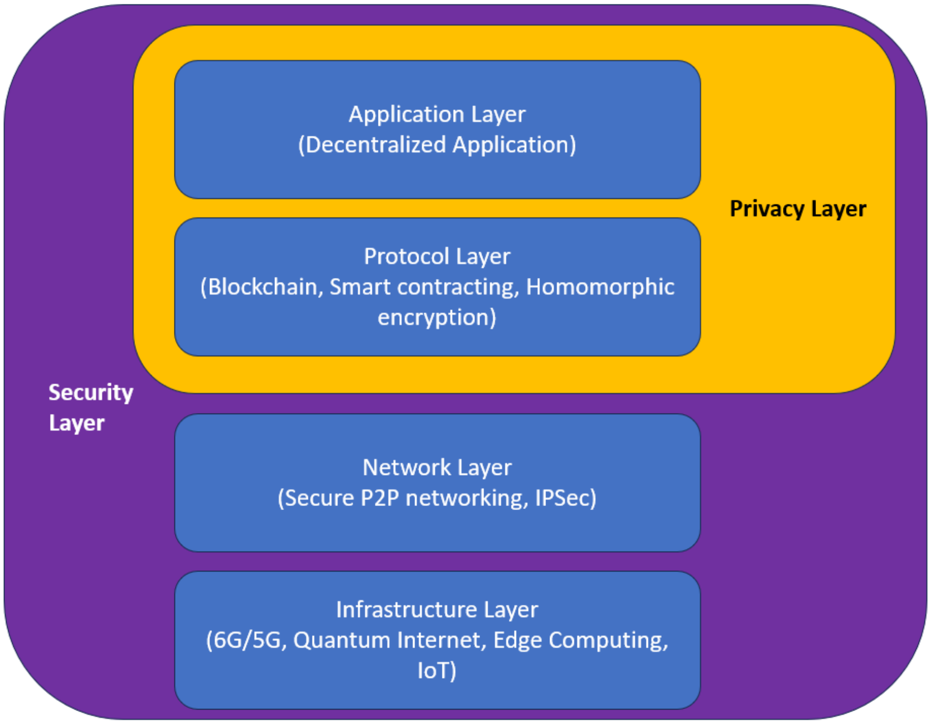

Application Layer
(Decentralized Application)
Privacy Layer
Protocol Layer
(Blockchain, Smart contracting, Homomorphic encryption)
Security Layer
Network Layer
(Secure P2P networking, IPSec)
Infrastructure Layer
(6G/5G, Quantum Internet, Edge Computing, IoT)

Figure 1: Web 3.0 Proposed architecture.

This proposed architecture overlays the strengths of Web 3.0 and is well aligned with its famously promoted capabilities. As shown in figure 1 above, the layers have stayed unchanged from what is portrayed in the literature. The englobing façade is the interaction layers representing the equipment used to fully experience the new Web. The layers are surrounded by the security layer in purple color emphasizing the necessity for holistic and total security through all the Web 3.0 technology stack. Security needs to be addressed not just at each layer of the architecture but within each transaction or interaction within the model. The privacy layer illustrated in orange color focuses only on the protocol and application layers and this is where information will start to make sense to humans and less sense to machines. Privacy enhancing methodologies and techniques are recommended at this stage where impact on performance is acceptable when weighed against identity and personal information risk exposures.

Web 3.0 has much potential than ever; the following list presents the advantageous core technologies reinforcing the Web of tomorrow:

- Blockchain: by definition, a distributed ledger containing growing encrypted records connected via cryptographic hashes [9]. This technology will play a critical role in securing users' data, reinforce their ownership, and achieve independence from "data-opolies" showcasing one the major strengths of Web 3.0.
- Edge computing: one of the paradigms of distributed computing technology and it is often confused with cloud computing. The difference is by large time-sensitivity of data being processed. Data is processed at the edges of the network. This is done to reduce network congestion and enhance latency for users. Edge computing is a necessity for the transformation towards the internet of things (IoT). This is enabled though use of advanced infrastructure such as edge servers, databases, and devices where 5G networks handle the carrying and routing of data [10].
- DApps: this is the popular term for decentralized applications. These applications do not communicate with a central datacenter, rather, they communicate through smart contracts on the blockchain. DApps communications can be simply taught of as apps using Peer-to-Peer communications like internet torrents, where data is shared and downloaded from multiple peers and not just from one source. Consequently, DApps also increase the level of privacy and security, although the latter is still at its early stages.
- Cryptography & privacy: Web 3.0 is here to stay and so must the exchanged data in its era, except that it must stay stored securely and safely. New advances in the field of cryptography such as post-quantum cryptography where data is assured to be secure and cannot be decrypted with a quantum computer. There is also homomorphic encryption that enables computers to assimilate the data and perform different computing operations without decrypting it.
- Next-gen telecom infrastructure: Web 3.0 will data compute intensive both from the rate data transfer and the computing performed on data as transit and when it comes to networking infrastructure, Web 3.0 will even challenge the end-to-end principle that founded the internet as it is known today. Fortunately, the progress made in mobile networks with the ongoing spread of 5G networks and the rapid catch up of 6G networks promises opens new horizons for accommodating Web 3.0 communications. Land internet such as fiber optics communications could handle the newly imposed data load, but the introduction of the quantum internet bears new opportunities such as ultra-high speed packets routing and additional security layers.
- Artificial intelligence: the advancements made in AI over the last two decades have boosted new economies and wrote new beginnings. AI is all about enhancing the digital experience of users and in this posture, it is well aligned with what Web 3.0 aims to achieve for the users. The field of Natural Language Processing (NLP) for example has been growing exponentially lately with revolutionizing applications such the generative pre-trained transformers (e.g., ChatGPT). Web 3.0 brings the notion of semantic web meaning that web pages will contain descriptors that will enable computers to understand content and leveraging NLP technology to autonomously assimilate, transform, and interact with human generated and shared content over the new Web.

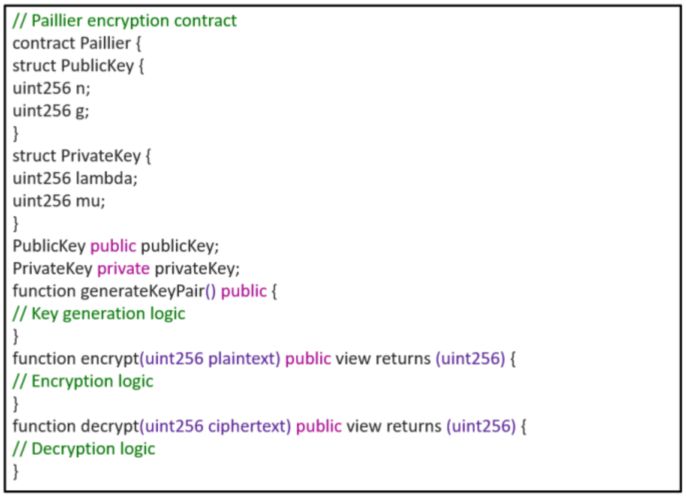
3. Convolution of threats in the Web 3.0 cyberspace

3.1. Web 3.0 challenges for cybersecurity

Convolution in mathematics is an operation where the shape of the first operand is modified by the second operand, hence, the reason this term was used in the section's title. To describe how the threat landscape is changing with the evolution of the Web 3.0 technology, it would be wise to first glance through the affected areas and what was familiar with scholars previously as form of threats. Web 3.0 can be thought of as human centric contrary to Web 2.0 which is data centric and that is due to the decentralization and privacy notches. Web 3.0 will be running human lives by a large difference than Web 2.0. Web 3.0 will affect social computing, financial technologies, innovation, governance and law-making [11]. New challenges and threats in all these areas are rising and this trend will continue for quite some time before any action is taken and this gap is generally well known whenever a new technology becomes available for application. The power of Semantic web which is one of the promised Web 3.0 features will drastically change experiences on the web such as the way social media platforms introduce their services, or the way ecommerce companies target their customers as the value of a user's metadata and traceability will be more detailed and available at a cheap price. This same feature will make the task of intelligence gathering about victims much easier than it already is. The rate of successful spamming and phishing campaigns will see an increase, but the most successful and accurate attacks will be Spear-phishing attacks. Cybercriminals will be able not just to gather superficial information about their targets but also track their behaviors, know their habits history, and anticipate or predict their future actions making the malicious phishing emails bypass most security measures and obstacles. The second area of immense concern is just governance and regulations. Policy making and regulations follow a lengthy process, and it causes delay which deepens the gap between of void leaving attackers enough room to exploit the vulnerability in the judicial system well enough before moving on to new tactics avoiding punishment. This well know phenomenon in political science as strategic adaptation gap and the gap is formed due to the delay for policy makers to catch up to an emerging technology and regulate it [12]. The legislative scenery for the last 5 years at least has been crippled unable to react to the growing spread of technologies such as cryptocurrencies, NFTs and financial fraud related to money laundering. Many national regimes have been hesitant to regulate this new financial industry. For example, few countries have been bouncing between prohibiting the use of cryptocurrencies and allowing them a second time. When cybercriminals witness such confusion from law makers and enforcement entities, they only grow confident and utilize the time window to its fullest. The most feature Web 3.0 is cherished for by technology enthusiasts is the privacy enhancement and data ownership in the new cyberspace. Web 3.0 offers privacy without bias which grants cybercriminals immunity and comfort to continue their operations without getting caught. This will pose new challenges to the current intrusion analysis tactics and new frameworks need to be elaborated. The level of anonymity that decentralization offers will make tracking Web 3.0 users quite challenging, hence current intrusion analysis frameworks like "The Diamond Model" or "The MITRE Att&ck" might turn out to be inefficient. The next paragraph will discuss in detail the challenges facing the financial industry and the technological issues that arise from Blockchain.

3.2. Threats from Blockchain smart contracts

One of the major innovative leaps that reshaped the Blockchain technology mechanisms is the appearance of smart contracts. The decentralization theory principle of Web 3.0 stipulates that no third-party entity should be involved in transactions between the seller and the buyer, however this idea would not be practical in the real world unless a trusted third party exists which can guarantee the smooth transition of funds from each party to the other and ensure their rights, thus, the invention of smart contracts transaction mechanism. Just as ordinary paper contracts in the real world, smart contracts are virtual contracts functioning on top of the Blockchain [13]. Smart contracts serve the same benefits. Smart contracts are self-executing codes categorized as transaction protocols, they control and execute agreements and transactions actions and events in an automated manner [14]. A smart contract acts as an intermediary guardian of value. From a technical perspective, a smart contract is viewed as the abstraction of a class containing different functions or methods that handle the construction of the data field which corresponds to the creation of the contract, another method handles the writing of contract terms depending on the agreements conditions, a security layer that contains encryption methods to reverse the immutability feature of these contracts, other methods responsible for event logging and storage in the Blockchain, and finally a self-destructing method to free the occupied space taken by the contract [13]. Shafaq & Faiza [15] have technically categorized two state types of these contracts. The first state type is a smart contract in a *constant state* which can be interpreted as a contract which cannot be modified. The second state type is the *writeable state*" and this means that the states can be saved in the Blockchain. The figure below shows a few selected methods that can be found in a smart contract class under Solidity programming language [16].



```
// Paillier encryption contract
contract Paillier {
struct PublicKey {
uint256 n;
uint256 g;
}
struct PrivateKey {
uint256 lambda;
uint256 mu;
}
PublicKey public publicKey;
PrivateKey private privateKey;
function generateKeyPair() public {
// Key generation logic
}
function encrypt(uint256 plaintext) public view returns (uint256) {
// Encryption logic
}
function decrypt(uint256 ciphertext) public view returns (uint256) {
// Decryption logic
}
```

Figure 2: Smart Contract code classes.

The operational process of initiating and resolving smart contracts is as follows: First, there are two essential parties in each transaction, a buyer, and a seller. The successful conclusion of a transaction would lead to the transfer of assets ownership from the seller to the buyer in exchange of funds from the buyer to the seller. Since neither party do not want to involve a third party to handle the transaction operation and want to trade in a trusted way. They invoke a smart contract where they specify terms, clauses, and conditions. The contract terms cannot be changed. Once both parties involved meet the conditions, the contract is executed automatically and handles the transaction operation. All smart contract operations are running via the smart contract platform (e.g., Ethereum VM) in a transparent and confidential way offering both parties identity protection feature and access to the chronological events. If for any reason the transaction is incomplete, the whole operation will be canceled, and parties refunded [17]. As a result, the smart contract would be terminated but its traceability is stored in the Blockchain. The following figure illustrates the smart contract operational process.

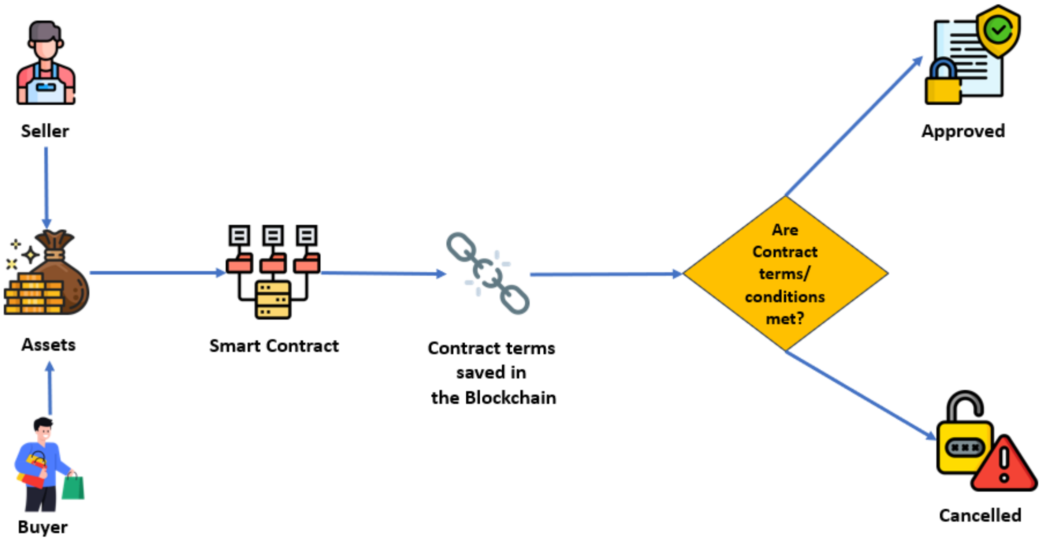

Seller
Assets
Buyer
Smart Contract
Contract terms saved in the Blockchain
Are Contract terms/ conditions met?
Approved
Cancelled

Figure 3: Smart Contract operational process.

Despite researchers trying to demonstrate over the years the security and independence of smart contracts as an additional trusted operational mechanism for the Blockchain, there have been few events that effectively perturbated all those convictions. Taking the May 2016 attack incident as an example where few known figures from the Ethereum society launched the Decentralized Autonomous organization (DAO) smart contract. 50 million USD worth of Ethereum cryptocurrency was stolen due to a vulnerability that came from low quality code [18]. The vulnerability known as reentrancy allowed the attacker to create a loophole requesting funds numerous times before the balance is updated. This incident demonstrated how a negligible bug in smart contracts code can have a tremendous and devastating impact in the Blockchain ecosystem. A second example incident with comparable destructive effect would be the Parity Wallet hacking incident. Parity Wallet is a secure Ethereum wallet offering robust features and user-friendly interface developed by Parity technologies [19]. The attack incident occurred one year after the DAO smart contract attack and precisely in July 2017 [19]. The attacker exploited a vulnerability in the multi-signature - responsible for securing transactions- wallet smart contracts. The source of the vulnerability came from coding mistake in the multi-signature process. The attack can be divided into two phases which are represented by transmitting two transactions aiming at taking ownership of the victim's wallet and draining the funds [20]. The attack resulted in a loss of 150,000 Ethereum (ETH) or 30M in USD.

The only valid statement that can describe the security situation of smart contracts is to say that it is unacceptable, however, anticipation to judgement would be unjust as the technology is still in its early phases of development and the research activities are yet to gain full momentum and be able to catch up to the on growing and fast fluctuating threat landscape.

4. Phishing attacks in Web 3.0

4.1. Phishing attacks

The National Institute of Standards and Technology (NIST) defines phishing as a technique for attempting to acquire sensitive data, such as bank account numbers, through a fraudulent solicitation in email or on a web site, in which the perpetrator masquerades as a legitimate business or reputable person [21]. Phishing attacks are the point of entry for most cyber-attacks. In fact, 91% of intrusive attacks happen through emails in the delivery phase of the cyber kill chain [22]. Throughout the years, phishing attacks have seen a significant development incorporating novel social engineering techniques and exploiting every means of communication technologies. Most phishing attacks are trapped in a few types that differ on few levels such as delivery technology used, the efficiency of the message and the vulnerability the embedded payload aims to exploit. The following list is briefly detailing the common types of phishing that circulate in the cyberspace:

- Traditional Phishing: also known as email phishing, is the first type of phishing attacks that gained popularity among cyber criminals and is still effective until the present day. The malicious email would simulate an ordinary trusted email that could be expected from the victim. The attacked might spoof or mimic an email address and pretend to be a trusted entity known to the target. The attacker would solicitate valuable information - usually the victim's access credentials – using emotional incentives such as fear, greed, or a sense of urgency. The attacker then tricks the victim to download a malicious file to be later executed to establish a command-and-control connection or transfer the victim to a fabricated landing web page that mimics an original sign in page to retrieve credentials.
- Spear Phishing: as the name entails, this type of attack is highly targeted and requires intensive intelligence gathering and reconnaissance techniques involving the victim's public and even private life. These attacks are also conducted through emails. The body and subject of the email usually involves a topic of interest about the victim's life such as a scholarship announcement if the victim is a student, an urgent email from the bank the victim is dealing with, or an email calling for an urgent and limited heavy discount for a good or a service only the victim knows about but has showed interest for it via social media. Cybercriminals can also perform intelligence gathering operations through tailgating the victim and studying closely their habits and behaviors.
- Ice Phishing: This type of attacks propelled due to the rise of cryptocurrencies where attackers can impersonate wallet software, tamper with transaction addresses, or impersonate legitimate dealers during a transaction to drain the victim's account from funds by initiating multiple transactions at several periods. The DAO Badger attack that occurred

between the months of November and December of 2021 resulted in theft of millions of US dollars' worth of cryptos [23].

- Smishing and Vishing: Following the pattern and purpose of phishing attacks, the "SM" letters stand for the Short Messaging Service (SMS) and the "V" in Vishing is the first letter in the word "Voice". Smishing attack is executed with a short phone message while Vishing is done over the phone. Popular type of Smishing attacks are attacks that create a sense of urgency where the attacker impersonated the victim's bank account system manager and prompts the user to reset their bank account password for the bank's mobile application through a link that drive the victim to a fake landing page meant to retrieve the victim's real credentials. An example of Vishing is where attackers try to highjack SIM cards by calling telecom operators and pretending to be the SIM card owners and ask for transfer of the phone number to a new SIM card.
- Business Email Compromise (BEC): Also known as Whaling in the field, these kinds of phishing emails tend to mobilize social engineering principles such as the principle of authority. The attackers pretend to be a figure of authority in the organization with higher organizational hierarchy than the victim. The final purpose is to convince the victim to act in the favor of the attacker such as handing credentials of senior management personnel or escalating account privileges in the company's internal system.

4.2. Novelties in social engineering techniques

In traditional cybersecurity strategic thinking, it was thought that the important requisite to achieving a system's security is the constant implementation of new security technologies on both the hardware and software but primitively the software was precedent. History has shown on many occasions that such a vision was narrow and lacked consistency for dealing with different emerging threats. The focus was then shifted to policy thinking, standardization, processes implementation, and security measures patching. This shift in the global cybersecurity community has reinforced the worldwide security posture but hasn't broken the will and perseverance of cybercriminals. Since the beginning, a hacker's first target has always been the weaker ring in the cyber chain which is the human element. There are always patches for equipment but none for humans. Potent attackers study their opponents with much focus more than they study the target's IT architecture. They leverage different social engineering techniques to achieve their goals efficiently and with the least number of resources. Social engineering enabled attackers to test and prove their limitless creative intellectual capabilities and even contribute to this field in a quasi-academic way. Acquiring sensitive information by manipulating a victim's behavior is no longer a mystery. As it has been mentioned in previous section that social engineering techniques involve pressing on the right human psychological points like fear, intimidation, greed, social reputation, and status in order for the attack to fulfill its objectives. The combination of these points with the exact right rhythm of play forms what is known to be social engineering techniques. Recent technical advancements in this area of study are discussed below in more detail.

- Typo squatting: usually internet users do not pay too much attention on their typos while online due to the number of distractions modern web surfing experiences have. Users might type "g00gle.com" instead of "google.com" which is the legitimate URL for the Google search engine [24]. The first could be a malicious website that imitates Google's user interfaces and is made to steal credentials, download certain malware, or simply sell the user competing products with the original website. Attackers also use this technique in phishing emails to impersonate legitimate domains especially with the spread of sub domains from trusted companies. The technique is also used when attackers try to obfuscate a URL or an email address exploiting the weakness of scripting system in computers.
- Reverse social engineering: if social engineering techniques are about transitioning the victim from a neutral psychological state to a state controllable by the attacker leading to sensible reaction from the victim that aligns with the attacker's objectives, then reverse social engineering is the exact opposite. This is where the attacker pretends to be in a tough state and would require immediate help from the victim tricking them to indirectly operate under his command. An example would be an attacker pretending they lost their personal business phone and important contact business numbers and would need to transfer the phone number to a new SIM card, therefor, attempting a SIM card hijacking operation using a reverse social engineering technique.
- Outsmarting smart contracts: this is one of the novel techniques of social engineering techniques applied to exploit cognitive weaknesses in the Ethereum Blockchain platform. The work Nikolay Ivanov & Qiben Yan [25] explores different social engineering attacks aimed at smart contracts. One of the demonstrated attacks is the pre-calculation of a future contract address and replacing Externally Owned Accounts (EOA) with a non-payable contract. The attack

begins when the user initiates a smart contract and deposits funds. The attacker intercepts the deposits event, then deploys a future non-payable contract after extracting the first real contract address with a special function. When the user initiates the funds withdrawing event. The attacker replaces the addresses, and the user receives a payment failure event in exchange for the deposited funds. This is an example of how naïve smart contracts are designed.

4.3. Large Language Models (LLMs) generating phishing emails

In recent years and with the major advancements made in Artificial Intelligence, Large Language Models have captured the attention of the Web community worldwide. In June 2020, ChatGPT 3 was launched a Large Language Models (LLM) trained on 175 billion parameters [26]. Then it was followed by a newer version named ChatGPT 3.5 offered for free public usage recently with version 4 as a premium upgrade. LLMs offer human-like cognitive services such as text processing, limited creative writing, style mimicking, content generation, explanation of material and summarizing. While these platforms can be used to increase human productivity and shorten the duration of tasks. Attackers have found novel ways to bypass embedded security locks aimed at preventing malicious use of these models. Cybercriminals are leveraging features of creative content generations to generate unsigned malicious aiding software that can be unknown to traditional antivirus software. They can leverage LLMs to conduct open-source intelligence in cheaper, more efficient, and faster ways than humans. They also craft brilliant prompts and generate targeted phishing emails. In fact, Julian Hazell found that hackers leveraging Anthropic's advanced model can generate one thousand spear phishing emails under just $10 U.S Dollars and for a 2 hour' time length [27]. The process of creating phishing emails using LLMs is divided into two phases. The first phase is the intelligence gathering phase, where information is gathered about the target to draft a relevant message to the target's interests. While LLMs cannot scrap the Web looking for such information [28], however, if they are giving the right input data embedded in the prompt, they are able to generate genuine phishing messages undistinguishable from human produced content. The second phase is construction phase after the right prompt has been engineered it can easily be inputted to generate the required content. Engineering prompts with such intents must take into consideration design adaptations to bypass security measures in LLMs to block malicious use of the model. To conclude, AI has offered cybercriminals an immense opportunity to strategize their phishing attacks with high multidimensional scalability at fraction of the cost to none.

5. Phishing 3.0 is phishing at scale

Phishing 3.0 is characterized by automation, scalability, and precision. These features are all offered by modern AI systems. This section will briefly characterize phishing 3.0 attacks by analyzing existing features to enable potential victims in detecting such emails and prevent breaches from cybercriminals. A phishing 3.0 detection analysis methodology is proposed.

5.1 Background

Next-gen anti-phishing email security systems implement AI and machine learning techniques to detect phishing emails, however, as with any statistical learning model used for detection there will always be a margin for error. The wide use of generative AI tools can really shape the email writing style and content to bypass these detection models using feature of contextual understanding and efficient prompt engineering such prompting the AI tool to generate emails with similar styling to emails from a popular newsletter the victim is following and so on.

5.2 Methodology

Mailboxes today are bombarded with emails that mostly remain unfiltered as they couldn't be recognized by email security solutions. It becomes significantly impossible for users to intuitively detect malicious emails from legitimate ones. The proposed methodology starts by extracting and analyzing phishing 3.0 emails and especially ones generated by generative AI. The user can analyze suspicious emails by observation and match extracted writing characteristics with the ones described on the method to build an informed decision. Phishing 3.0 usually tend to have the following characteristics:

- Unweighted writing tone: Writing content generated by AI usually doesn't distinguish between the weight of the topic and the formal tone of writing. For example, the written content can be simple as a fake email asking for a reset of account credentials but will be written in a highly formal tone that would make the human reader feel something is unusual and there is no necessity for using high level grammar and sophisticated vocabulary. The well-engineered prompt can deviate from this weakness.
- Inconsistency: There will always be a margin of error with any AI system. When the email content seems to be almost irrelevant or off main topic at certain positions this can be an indicative of generative AI content.
- Personalization: Lengthy conversations with LLM bots can show signs of generic content which lacks personalization in the conversation and can exclude prior given information about the interlocutor after multiple iterations of replies.
- Speed: If engaged in an email conversation and the reply emails are too fast for human speeds to generate this is a strong sign of generative AI especially if the replies are dynamically reacting to your questions and comments.

5.3 Recommended actions

- Verification: users are required to verify and check the authenticity of email senders and their mail service domains. If the email seems to come from an unknown destination or has a poorly formatted domain. It is recommended to flag the email and forward it to security teams.
- Protecting sensitive information: refraining from sharing personal or sensitive information such as access credentials of privilege escalation passwords is not recommended even in utmost situations unless proper verification is validated by competent entities.
- Maintaining privacy: Web 3.0 transfers the responsibility of data privacy and security from third parties directly to the user in exchange for complete anonymity. Web 3.0 users must handle their own data protection and preservation operations. Therefore, users need to become knowledgeable about data privacy best practices and what tools to use. Users must choose a comprehensive and accredited source of cybersecurity knowledge to avoid nuances and confusion.

6. Proposed solutions

While many papers have suggested and discussed potential solutions for security issues in Web 3.0, none have attempted to address the issue of phishing cyberattacks, despite being a pivotal determinant in most attack campaigns at the present. This section aims to propose a set of solutions and recommendations in hopes of strengthening active research and ongoing efforts to secure this layer of cybersecurity. The proposed solutions are strategized to prioritize limiting access to personal information which is the valuable commodity that feeds the process of generating phishing content. It is important for now to formulate all anti-phishing solutions to this idea as it is the only controllable parameter at the moment.

6.1. Privacy enhancing

Data privacy is at the core of Web 3.0 architecture principles. Web 3.0 benefits from privacy enhancing protocols, cryptography schemes, and decentralized applications. Danyal Namakshenas [29] pushed for decentralized identity preserving data pools on the Blockchain, where personal data can be stored encrypted and immutable. Users can share their personal information with third parties at their own will. While this proposition seems to be perfectly compatible with the decentralization principle, it raises questions about its practicalities as to how these identity hubs can sustain their operations and how can their security auditing be transparent without threatening their security posture. A more isolated approach is for the user to take complete responsibility for their own personal information storage locally such as in their e-wallet and provide it through secure peer-to-peer channels whenever they need to.

6.2. Homomorphic encryption

Impactful phishing campaigns often rely on the target's personal information gathered from different sources whether via open sources or through buying datasets from data breaches sold in the dark web. Homomorphic encryption protects personal data in an unprecedented way. Homomorphic encryption allows for machine operations on the encrypted data without any encryption and thus, data can be utilized freely with no risk of exposure. Homomorphic encryption holds great promises for the Web 3.0 data privacy and security technology stack. Once fully implemented, cybercriminals will be disincentivized to sustain data theft attack operations.

6.3. Novel detection methods for phishing content

Phishing content is divided into two categories, phishing emails serving as first communication contact with potential victim and phishing web landing pages where the victim is tricked into providing their confidential information. With the rise of the AI revolution, researchers have explored ideas into implementing these technologies for cybersecurity as they usually handle repetitive tasks well enough to achieve full automation. The interest in this section is focused on Large Language Models (LLMs) applications to detect phishing emails and websites more rapidly and accurately than traditional reputation and scanning techniques still in use by cybersecurity solutions. Current anti-phishing solutions do apply machine learning and AI techniques to analyze phishing content, however they lack the contextual understanding that LLMs have. They usually analyze the provided content with parameters gained from the training set used to build the AI model and they find difficulty to adapt and evolve with the threat landscape as cybercriminals are always innovating maneuverable counterattacks to bypass these intelligent systems. For detecting phishing websites, Takashi Koide et al. [30] proposed a new approach where ChatGPT LLM is used to scan web pages and classify them as phishing websites or not. The researchers used a process where websites are fetched then dismantled in separate elements containing the URL, reduced content free from HTML tags, images are run through an OCR algorithm to extract graphic text information, and then using ChatGPT4 with predesigned prompt, the classification results are produced in matter of seconds. A comparative study was also conducted between ChatGPT precedent model version 3.5 and the premium version 4. Performance results were quite promising with both LLMs versions achieving 98.3% precision with ChatGPT 4 scoring more in recall and accuracy with 98.4% [30].

LLMs have also proven to be worthy of detecting human and AI crafted phishing emails. The experiment run by Fredrik Heiding [31] where 20 phishing emails generated by LLMs and humans were fed into different LLMs models such as ChatGPT, Bard, LLaMA, and Claude for malicious intent detection and phishing classifications. Claude LLM outperformed all its counterparts. Claude was able to detect 75% of the human generated emails, 100% of the GPT-generated emails, 100% of the V-Triad (an advanced methodology for writing phishing emails) emails, and 100% of the emails generated by mix of V-Triad and GPT [31].

6.4. Web 3.0 reputation systems

Trust and reputation systems are more than important for the growth of Web 3.0 especially due to the private nature of the technology. In Web 2.0, Web security reputation systems depended only on application and network vulnerabilities scanning for calculating a risk index for the assessed Web domain. The complexities and anonymities of Web 3.0 will drag the process of threat hunting and disclosure longer than usual, at least initially. The Web 3.0 reputation systems must include users feedback in risk calculations index. Data collected from users' feedback systems must be verifiable and confirmed by cybersecurity expert entities before inclusion. Intelligence from untraditional sources such as Dark Web should also be analyzed and assessed for inclusion if found valuable. Threat information sharing between public and private sectors need to standardize for the greater good. Adding these two components to Web 3.0 reputation systems will outperform current systems and establish trust on the internet of the future.

6.5. Zero Trust Access

The distributed nature of Web 3.0 amplifies inherited cybersecurity challenges and simultaneously encourages the rise of new ones. The rise of remote access after the last pandemic and the convenience of cloud applications introduced many weaknesses revolving around authentication and data privacy in the network. Zero trust access (ZTA) by default concept enforces multi-dimensional authentication for each requested access to the IT infrastructure. It uniformly treats users and devices accessing the network whether they connect internally or externally. ZTA verifies users & devices identities. The verification playbook rules ensure that the user is the owner of the device, the identity of the user in the organization's matches the identity of the current access requester. ZTA has many advanced features in terms of analyzing users' behaviors and patterns when signed into the system. These features can detect anomalies such as the impossible travel case (when a user's location changes in few minutes between two continents), it can also flag access request to certain services as suspicious during holidays or outside ordinary working hours. Implementing ZTA is a gradual journey that starts by integrating the solution at authentication points in the IT architecture alongside existing networking services such as VPNs. The choice of ZTA needs to consider many requirements like privacy, coexistence, scalability, redundance, and impact on stakeholders. The journey to ZTA is achieved through continuous assessments overtime and must finish by covering all data access points.

6.6. Phishing awareness

The human element is the indispensable catalyzer for every successful phishing attack. Cybercriminals will exploit human behaviors, routine habits, ideologies, and emotions to the advantage of their goals and targets. Intensified security measures will remain ineffective in stages where humans become the sole barrier. Many contributions from cybersecurity public and private sectors enlightened the public about the importance of cybersecurity as a collective responsibility to help fill the awareness gap between the average internet user and the expert cybercriminal. The traditional approach is creating generic cybersecurity awareness content and broadcasting through all means of communication. Although this approach is neither meaningless nor useless, it hasn't generated a satisfactory impact. Steve Sheng et al. found that phishing awareness educational materials reduced users' likelihood to input information into phishing landing pages by 40% [32]. The suggested approach is to diversify content and align it with different criteria such as the educational level of the receiver, their professional area of expertise, and their online interests. Awareness against phishing in Web 3.0 should include prioritized short security advice that is easy to assimilate. Defensive countermeasures should be included in education material with online practical scenarios. The material must comprehend the educational background of the receiver. Phishing awareness in Web 3.0 must focus on educating the public on how to leverage Web 3.0 privacy capabilities and the potential threats and social engineering attacks associated with its technology stack.

7. Conclusion

The title of this article suggests that new opportunities as interesting as they may appear always hold challenges and carry risks as soon as they are implemented. The complexities and security challenges Web 3.0 uncovers are beyond the responsibility of technology leaders to hold users accountable for shielding themselves from the revolving threats. Web 3.0 is not only a technological transformation; it is a novel social and economic experience that soon humanity will begin to adapt to either willingly or by force of circumstances. Thus, the transformation must also be accompanied by a philosophical critique of future situations that Web 3.0 will impose over cyberspace. Cybersecurity can no longer be illustrated as neither private nor public good.

Phishing attacks have always been a major concern in cybersecurity studies. The emergence of Web 3.0 is providing new legitimate concepts, tools, and techniques exploitable by cybercriminals for malicious and disruptive intents. Most cybercriminals will be focusing on the human element's lack of cybersecurity knowledge and best practices involving Web 3.0 technologies. Raising awareness by unifying and accrediting sources of cybersecurity and data privacy awareness is the future approach to win this challenge and achieve a safer cyberspace.

This promising approach must be universal and must not exclude any honest initiative or contribution as insignificant it might be assessed by today's measure it will make an impact by measures of the future.

References

[1] "Web 3.0 Statistics and Facts" Internet: https://market.us/statistics/information-and-communication/web-3-0-statistics/, Oct. 27, 2022 [Dec. 20, 2023].

[2] "General Information" Internet: https://www.cisa.gov/stopransomware/general-information, [Dec. 24, 2023].

[3] Gilad, Edelman. "What Is Web3, Anyway?" Internet: https://www.wired.com/story/web3-gavin-wood-interview/, Nov. 29, 2021 [Oct. 3, 2023].

[4] Maurice, E. Stucke. "Here Are All the Reasons It's a Bad Idea to Let a Few Tech Companies Monopolize Our Data" Internet: https://hbr.org/2018/03/here-are-all-the-reasons-its-a-bad-idea-to-let-a-few-tech-companies-monopolize-our-data, Mar. 27, 2018 [Nov. 01, 2023].

[5] Fan, Yuqing, et al. "The current opportunities and challenges of Web 3.0". arXiv:2306.03351 [cs], Jun. 6, 2023 [Oct. 3, 2023].

[6] Sónia, Silva. "Web 3.0 and Cybersecurity – Short Paper" *Advanced Research on Information Systems Security, an International Journal (ARIS2) (2022)* [On-line]. Volume 2, No 2, pp 39-49. Available: https://doi.org/10.56394/aris2.v2i2.21 [Sep. 27, 2023].

[7] Jacksi, Karwan & Abass, Shakir. "Development History Of The World Wide Web" *International Journal of Scientific & Technology Research* [On-line]. Volume 8, (2019), pp 75-79. Available: https://www.researchgate.net/publication/336073851_Development_History_Of_The_World_Wide_Web, [Oct. 5, 2023].

[8] G. Zheng, L. Gao, L. Huang and J. Guan. (2020, August 31). *Ethereum Smart Contract Development in Solidity.* (1st edition). [On-line]. Available: https://doi.org/10.1007/978-981-15-6218-1 [Oct. 21, 2023].

[9] A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder. (2016, July 19). *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction.* (1st edition). [On-line]. Available: https://press.princeton.edu/books/hardcover/9780691171692/bitcoin-and-cryptocurrency-technologies [Oct. 22, 2023].

[10] Hua, Haochen & Li, Yutong & Wang, Tonghe & Dong, Nanqing & Li, Wei & Cao, Junwei. (2023). "Edge Computing with Artificial Intelligence: A Machine Learning Perspective" *ACM Computing Surveys* [On-line]. Volume 55, pp 1-35. Available: https://dl.acm.org/doi/10.1145/3555802 [Dec. 17, 2023].

[11] Wensheng Gan, Zhenqiang Ye, Shicheng Wan, and Philip S. Yu. "Web 3.0: The Future of Internet" In Companion Proceedings of the ACM Web Conference 2023 (WWW '23 Companion). Association for Computing Machinery, New York, NY, USA, pp. 1266–1275. Available: https://doi.org/10.1145/3543873.3587583 [Oct. 9, 2023].

[12] Mika Kuikka. "Threat Perceptions and Strategic Adaptation in 21st Century Finland", Master thesis in military science (strategy), Swedish Defense University, Sweden, 2023 [Oct. 10, 2023].

[13] Vitalik, Buterin. "A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM." Available: https://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf , 2015 [Oct. 10, 2023].

[14] Alexander, Savelyev. "Contract Law 2.0: «Smart» Contracts As the Beginning of the End of Classic Contract Law" *Higher School of Economics Research Paper* [On-line]. Available: http://dx.doi.org/10.2139/ssrn.2885241, Dec. 14, 2016 [Nov. 1, 2023].

[15] Shafaq Naheed Khan, Faiza Loukil, Chirine Ghedira-Guegan, et al. "Blockchain smart contracts: Applications, challenges, and future trends". *Peer-to-Peer Netw. Appl.* [On-line]. Volume 14, (2021), pp 2901–2925. Available: https://doi.org/10.1007/s12083-021-01127-0, [Nov. 13, 2023].

[16] Solidity Academy. "Data Privacy and Homomorphic Encryption in Solidity: Safeguarding Confidentiality in Smart Contracts" Internet: https://medium.com/@solidity101/data-privacy-and-homomorphic-encryption-in-solidity-safeguarding-confidentiality-in-smart-a6f2b313b33f, Jun. 21, 2023 [Nov. 10, 2023].

[17] Susmita Pathak. "Smart Contracts" Internet: https://www.wallstreetmojo.com/smart-contracts/, [Nov. 15, 2023].

[18] Klint Finley. "WIRED: A 50 million hack just showed that the DAO was all too human" Internet: https://www.wired.com/2016/06/50-million-hack-just-showed-dao-human/, Jun. 18, 2016 [Nov. 20, 2023].

[19] Web3author. "PARITY Wallet Hack: What, When and How?" Internet: https://medium.com/@web3author/parity-wallet-hack-demystified-all-you-need-to-know-91b8dcb5b81, Jun. 30, 2023 [Nov. 20, 2023].

[20] S. Sayeed, H. Marco-Gisbert and T. Caira. "Smart Contract: Attacks and Protections". in *IEEE Access*, Volume 8, (2020), pp. 24416-24427, Available: DOI: 10.1109/ACCESS.2020.2970495.

[21] Michael Nieles, Kelley Dempsey, Victoria Yan Pillitteri. "An Introduction to Information Security". NIST Special Publication 800-12, Available: https://doi.org/10.6028/NIST.SP.800-12r1, June 2017 [Nov. 18, 2023].

[22] "91% of all cyber attacks begin with a phishing email to an unexpected victim". Internet: https://www2.deloitte.com/my/en/pages/risk/articles/91-percent-of-all-cyber-attacks-begin-with-a-phishing-email-to-an-unexpected-victim.html, Jan. 9, 2020 [Nov. 19, 2023].

[23] Microsoft Threat Intelligence. "'Ice phishing' on the blockchain". Internet: https://www.microsoft.com/en-us/security/blog/2022/02/16/ice-phishing-on-the-blockchain/, Feb. 16, 2022 [Nov. 20, 2023].

[24] Amrut Kajave, Shazny Ahmed Hussain Nismy. "How Cyber Criminal Use Social Engineering To Target Organizations". arXiv:2212.12309 [cs.CY], Dec. 7, 2022 [Nov 25. 2023].

[25] Nikolay Ivanov, Qiben Yan. "Et tu, Blockchain? Outsmarting Smart Contracts via Social Engineering". arXiv:2209.08356v1 [cs.CY], Sep. 17, 2022 [Nov 26. 2023].

[26] Bernard Marr. "A Short History Of ChatGPT: How We Got To Where We Are Today". Internet: https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=15891e9b674f, May 19, 2023 [Dec. 7, 2023].

[27] Julian Hazell. "Spear Phishing With Large Language Models". arXiv:2305.06972 [cs.CY], May. 11, 2023 [Dec. 15, 2023].

[28] Maryia Stsiopkina. "How to Use ChatGPT for Web Scraping in 2023". Internet: https://oxylabs.io/blog/chatgpt-web-scraping, Jul. 7, 2023 [Dec. 16, 2023].

[29] Danyal Namakshenas. "Web3.0 Security: Privacy Enhancing and Anonym Auditing in Blockchain-based Structures". Master of Cybersecurity and Threat Intelligence Thesis, University of Guelph, ON, Canada, 2023.

[30] Takashi Koide, Naoki Fukushi, Hiroki Nakano, Daiki Chiba. "Detecting Phishing Sites Using ChatGPT". arXiv:2306.05816 [cs.CR], Jun 9, 2023 [Dec. 16, 2023].

[31] Fredrik Heiding, Bruce Schneier, Arun Vishwanath, Jeremy Bernstein, Peter S. Park. "Devising and Detecting Phishing: Large Language Models vs. Smaller Human Models". arXiv:2308.12287 [cs.CR], Aug 23, 2023 [Dec. 18, 2023].

[32] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. "Who falls for phish? a demographic analysis of phishing susceptibility and effectiveness of interventions". *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2010, pp. 373–382. Available: https://doi.org/10.1145/1753326.1753383 [Dec. 20, 2023].