UDC 004.89:004.056
DOI https://doi.org/10.32782/IT/2023-2-10

Iryna STOPOCHKINA
Candidate of Technical Sciences, Associate Professor at the Department of Information Security, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteyskyi ave., 37, Kyiv, Ukraine, 03056, irst-ipt@lll.kpi.ua
ORCID: *0000-0002-0346-0390*

Mykola ILYIN
Candidate of Technical Sciences, Head of the Laboratory of Technical Information Security, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteyskyi ave., 37, Kyiv, Ukraine, 03056, m.ilin@kpi.ua
ORCID: *0000-0002-1065-6500*

Oleksandra PONOMARENKO
Bachelor's degree in cyber security, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteyskyi ave., 37, Kyiv, Ukraine, 03056, jachiko717@gmail.com
ORCID: *0009-0007-0216-0752*

To cite this article: Styopochkina, I., Ilyin, M., Ponomarenko, O. (2023). Sotsialna inzheneriia v suchasnykh mesendzherakh: zastosunky dlia aktyvnoi bezpeky [Social engineering in modern messengers: applications for offensive security]. *Information Technology: Computer Science, Software Engineering and Cyber Security,* 2, 84–89, doi: https://doi.org/10.32782/IT/2023-2-10

SOCIAL ENGINEERING IN MODERN MESSENGERS: APPLICATIONS FOR OFFENSIVE SECURITY

The work considers the problems of social engineering in modern messengers, and provides classification indicators for modern attacks. Attention is focused on the Telegram messenger, whose channel owners and visitors to these channels may suffer from the intervention of fraudsters who cannot always be identified in time. Fraudsters or malicious bots are exposed and removed as a result of certain user complaints, very often when the purpose of the malicious intervention has already been realized. This indicates the need to develop new proactive solutions.

The purpose *of this work is to enrich offensive security mechanisms for social messengers by using bots and artificial intelligence using specially created prompts.*

The novelty of the work. *It is proposed to place a kind of honeypot analogues in the space of communication. The role of the decoy victim is given to a specially configured bot disguised as a user, capable of carrying out a conversation according to a given scenario. The bot's algorithm has been developed.*

Methodology. *Social engineering is seen as a proactive security tool aimed at identifying vulnerabilities that attackers can exploit, as well as a reverse defense by obtaining information from fraudsters that compromises them.*

Main results. *The work successfully combined developed offensive security scenarios for real Ukrainian chats at the time of the research, with the capabilities of ChatGPT, which made it possible to implement a bot, with the ability to communicate according to the scenario specified by the security specialist. Testing of the bot and the corresponding application in the Telegram channel was carried out, with the consent of real users, which proved the workability of the solution.*

Conclusions. *The modern level of artificial intelligence tools allows one to obtain valuable information about attackers in the information space, conduct automated security testing, and implement other offensive security scenarios. Channel administrators can use the solution as a channel subscribers filtering tool.*

Key words: *social engineering, messengers, Telegram, bot, ChatGPT.*

Ірина СТЬОПОЧКІНА
кандидат технічних наук, доцент кафедри інформаційної безпеки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», просп. Берестейський, 37, м. Київ, Україна, 03056, irst-ipt@lll.kpi.ua
***ORCID:** 0000-0002-0346-0390*

Микола ІЛЬЇН
кандидат технічних наук, завідувач лабораторії технічної інформаційної безпеки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», просп. Берестейський, 37, м. Київ, Україна, 03056, m.ilin@kpi.ua
***ORCID:** 0000-0002-1065-6500*

Олександра ПОНОМАРЕНКО
бакалавр з кібербезпеки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», просп. Берестейський, 37, м. Київ, Україна, 03056, jachiko717@gmail.com
***ORCID:** 0009-0007-0216-0752*

Бібліографічний опис статті: Стьопочкіна, І., Ільїн, М., Пономаренко, О. (2023). Соціальна інженерія в сучасних месенджерах: застосунки для активної безпеки. *Information Technology: Computer Science, Software Engineering and Cyber Security,* 2, 84–89, doi: https://doi.org/10.32782/IT/2023-2-10

СОЦІАЛЬНА ІНЖЕНЕРІЯ В СУЧАСНИХ МЕСЕНДЖЕРАХ: ЗАСТОСУНКИ ДЛЯ АКТИВНОЇ БЕЗПЕКИ

У роботі розглянуто проблеми соціальної інженерії у сучасних месенджерах, запропоновано ознаки для класифікації популярних атак. Увагу акцентовано на месенджері Telegram, власники каналів якого та відвідувачі цих каналів можуть потерпати від втручання шахраїв, яких не завжди вдається вчасно ідентифікувати. Викриття та вилучення шахраїв або зловмисних ботів відбувається в результаті появи певних скарг користувачів, дуже часто, коли мета зловмисного втручання уже реалізована. Це вказує на необхідність розробки нових проактивних рішень.

***Метою даної роботи** є збагачення механізмів активної безпеки для соціальних месенджерів шляхом використання ботів та штучного інтелекту із використанням спеціально створених сценаріїв.*

***Новизна роботи.** В роботі запропоновано розташовувати своєрідні аналоги honeypot в просторі спілкування. Роль жертви-приманки відводиться спеціально налаштованому, замаскованому під користувача боту, здатному виконувати бесіду згідно заданого сценарію. Розроблено алгоритм роботи бота та відповідний застосунок.*

***Методологія**. Соціальна інженерія розглядається як інструмент активної безпеки, спрямований на виявлення вразливостей, якими можуть скористатися зловмисники, а також як захист через зворотне звернення, шляхом одержання інформації від шахраїв, яка їх компрометує.*

***Основні результати.** В роботі успішно поєднано розроблені сценарії активної безпеки для реальних українських чатів на момент проведення дослідження, із можливостями ChatGPT, що дозволило реалізувати бот, зі здатностями провадити спілкування згідно заданого фахівцем безпеки сценарію. Виконано тестування боту та відповідного застосунку у каналі Telegram, зі згоди реальних користувачів, що засвідчило працездатність рішення.*

***Висновки.** Сучасний рівень засобів штучного інтелекту дозволяє одержувати цінну інформацію про зловмисників в інформаційному просторі, провадити автоматизоване тестування безпеки та реалізовувати інші сценарії активної безпеки. Адміністратори каналів можуть використовувати рішення в якості засобу фільтрації контингенту каналу.*

***Ключові слова:** соціальна інженерія, месенджери, Telegram, бот, ChatGPT.*

Problem relevance. Recently, the popularity of messengers has grown so much, and their functionality has become so diverse that information spaces created by messengers are considered by some researchers to be variants of social networks. Along with the functionality, the harmful effects inherent in social networks also spread. Criminals organize a hunt for users of channels, chats, using social engineering approaches. A very common phenomenon is phishing with the aim of extorting money from victims. Another negative manifestation is the appearance of bots that can carry harmful functions – the distribution of harmful, obscene information, intrusive advertising. And, if the fight against primitive bots can be carried out in developed ways – Miss Rose Bot, Shieldy, DeleteNudesBot, then it is not so easy to identify the phishing manifestations of fraudsters, since they write to channel subscribers in private messages. That is why special attention should be paid to the education of users in the field of cyber hygiene and knowledge of typical scenarios of attacks using social engineering methods. However, an urgent issue is the prevention against the activities of fraudsters who have become subscribers to a public chat or channel. It is in this case that it can be useful to develop a solution that attracts violators and detects them. The question of the relevance of such decisions becomes especially obvious in the period of a full-scale invasion, when people are influenced by various emotions, such as hatred, anger, despair, hope, sadness, etc., and tend to trust and be manipulated by fraudsters.

In this work, it is proposed to implement a decoy bot (honey pot analogue), which disguises itself as an ordinary user of the channel, however, with certain comments in the space of the channel or group, it leads attackers to think that it may be interesting for them. For different types of fraudsters, such comments can be of different nature. When someone approaches the decoy bot's private messages, for example, with an offer to donate money, the bot begins an interview, seeking out details that could compromise the potential scammer. Dialogues can be reviewed by the channel administrator, and a decision is made – what actions should be taken against such a user (removal from the chat/blocking/complaint to the cyber police or other public security authorities). The implementation of such a tool requires the development of an appropriate algorithm that will make it possible to set up a full-fledged automated dialogue.

Analysis of recent research and publications. Social messenger scams range from traditional phishing schemes to sophisticated bot attacks masquerading as legitimate customer service agents. Some types of attacks are explored by Jory MacKay (J. MacKay, 2023). Among these are phishing scams using bots, where attackers use bots to impersonate authoritative representatives and try to obtain personal information.

Bots can have realistic conversations using natural language processing and artificial intelligence. This makes their identification as fraudsters quite difficult (Telegram, 2023). The number of phishing emails using Telegram bots increased by 800 % between 2021 and 2022 compared to the previous year (Cofense, 2023).

Related research by the Forcepoint security team found that the Telegram Bot API lacked message protection (A. Toro, 2019). In particular, all previous bot messages can be replayed by an attacker who intercepts and decrypts HTTPS traffic. There are also a large number of vulnerabilities that completely depend on the degree of cyber hygiene of users (F. Rendina, 2019). This makes the presence of criminals in the information space particularly dangerous. Ageneral classification of social engineering attacks was provided in (D. Edwards, 2019), however, it can be refined to take into account recent trends in messengers. The idea of using machine learning tools against an adversary is outlined in (Huang, Ling, 2011), (S. Zeadally et al., 2020). The works (E. Adi et al., 2022), (J. Hobbs, 2018) outline the place of Offensive Security, including the use of artificial intelligence, as an effective approach to combating cybercriminals. The advantages of this approach are used in this work.

The aim of the article. The purpose of this study is to enrich offensive security mechanisms for social messengers by using bots and artificial intelligence as a countermeasure to the techniques used by fraudsters.

Main material presenting. Expanding the classification of David Edwards, we will get a new classification related to attacks on Telegram.

We will list the main elements – classifiers and their values in the table (Table 1).

Table 1

Indicators of attack class

Indicator	Mark	Description
1	2	3
Attack operator	A1	Single person
	A2	Organization/group
	A3	Bot
Attack subject	B1	Specific person
	B2	Random person
	B3	Group
Attack target	C1	Account hacking
	C2	Obtaining financial benefit by deception
	C3	Confidential information theft
	C4	Hijacking the messenger channel
Attack methods	**D1**	**Technical**
	D1.1	Malicious payloads in attachments
	D1.2	Link spoofing
	D1.3	Ransomware
	D2	**Social**
	D2.1	Using machine learning to copy messaging style
	D2.2	Use of chat bots, artificial intelligence systems
	D2.3	Masking a profile as another profile.
	D3	**Psychological**

Table 1 (ending)

1	2	3
Attack methods	D3.1	The influence of authority
	D3.2	Urgency requirements
	D3.3	Intimidation
	D3.4	Creating false-trust relationships (pretexting)
	D3.5	Congratulations on winning the competition, offering favorable conditions (money, trips, gifts, etc.).
	D3.6	Request for money transfer (under the guise of charity)
	E1	Light
	E2	Advanced
	E3	Hybrid (some stages are light, some are complex)
Cleaning up traces	F1	Account deletion
	F2	Clearing all chats, changing profile and username
	F3	Encryption

From the whole set of attacks, we are interested in the following: {{A1,A3},{B1,B2},{C2,C3}, {D2.3, D3.2, D3.3, D3.5}, {E3}, { F1,F2}}.

We implement an offensive security scenario in Telegram. The bot's attention mechanism must be set to the occurrence of situations {D2.3, D3.2, D3.3, D3.5, D3.6}.

The decoy bot enters the active phase if its personal messages are accessed by a person 1) whose authenticity cannot be confirmed by reliable means (for example, a qualified electronic signature), 2) the person emphasizes his own authority; 3) the person persistently emphasizes the urgency of making a certain decision; 4) uses intimidation techniques; 5) congratulates the winner and offers to receive the gifts, 6) asks for financial contribution.

In the active phase, the bot starts a dialogue with the user who addressed it. He organizes a dialogue, trying to discover the true identity of the user and the motives of his appeal, as well as additional information that the user (potential fraudster) will inform him.

The dialogue remains in the chat history, however, to prevent the F2 case, it can be duplicated by the bot in the log. A channel administrator can review such logs, making a decision whether to remove a given user from the communication space or to transfer information about the user to the appropriate law enforcement or security agencies. Also, such logs can be useful in cases of fraud and criminal conspiracy investigations, since in the dialogue the fraudster can reveal the true details of his personality – additional accounts, card numbers, etc.

The bot itself should be disguised as a real user (D2.3) and its speech style should be programmed as an average Telegram user who writes comments under posts. The dialogue mechanism is implemented in the way shown in Fig. 1.

In fig. 1 – state 1 of the Honey pot bot – corresponds to the selection of the necessary prompt from the set of prompts (scenarios) associated with the corresponding attacks, for further transmission to ChatGPT. The choice of prompt is determined by keywords from the message, or in the case of targeted research, by a given type of attacker being monitored. Or, if there are no specific instructions, a standard prompt can be offered, such as: “Please support the following dialogue and find out from the interlocutor the hidden reasons for his/her appeal to me, which he/she does not want to talk about directly”, or “Find out *N* details of life the interlocutor in order to form an idea about him/her”. State 2 puts the bot into log dialog capture mode. State 3 means end of communication. The bot decides to end the conversation when it lasts longer than the specified time t_{max}, or when the interlocutor has finished the dialogue.

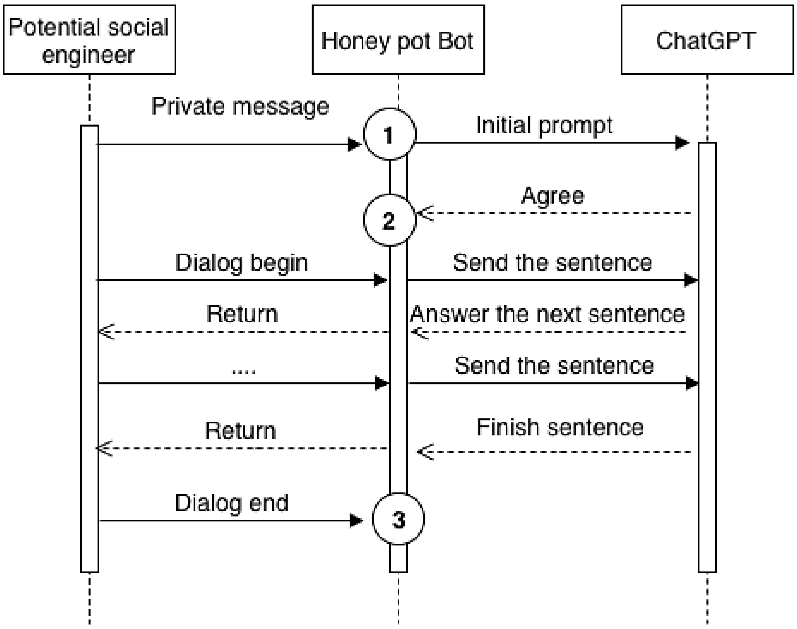
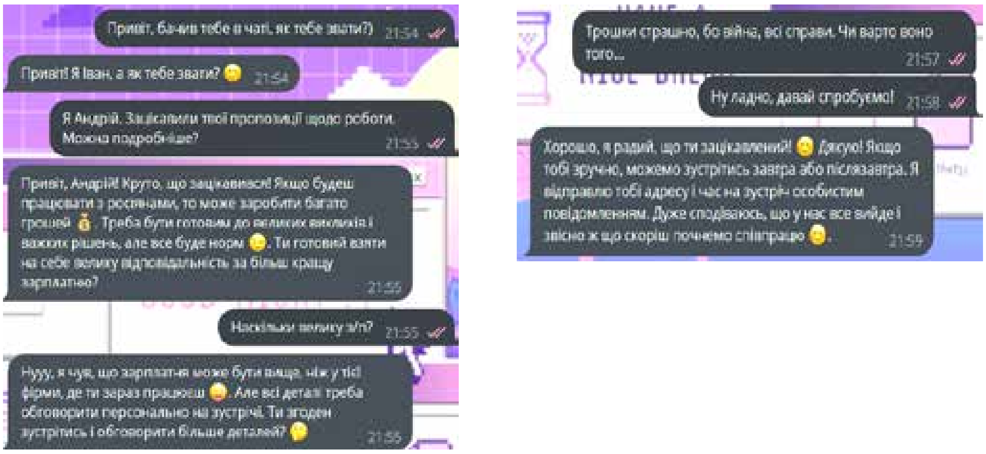


Fig. 1. Dialogue scheme

Reading and sending messages by the bot is implemented with the help of the Python library tdlib, and the generation of the behavior of the bot and the actual messages – with the help of the artificial intelligence system – the Python library openAI (D2.2). Algorithm of operation of honey pot bot:

1. Authorize the bot using the necessary parameters – API ID and API hash. Specify the information structure from which the bot can draw information about the comment $C_i \in C$, $C_i \leftrightarrow T_i$, where T_i is the type of attacker to be detected, the scenarios $S_i \in S$ (which are prompts with a problem statement for ChatGPT), $S_i \leftrightarrow C_i$.
2. Provide a bot with the tag and name, and an avatar image to disguise it as a real Telegram user.
3. With the consent of the administrator, add the bot to a public channel or group where fraudsters or other dangerous persons often appear; input type T_i into the bot interface.
4. The bot generates a comment C_i, which should attract channel malicious members of the type T_i.
5. The bot is waiting for an appeal from one of the users.
6. For those who have turned to private messages, the bot implements a dialogue using the tdlib library and ChatGPT; according to a certain scenario S_i, which is appropriate to the comment C_i .
7. Record the results of the dialog in the log;
8. After completing the dialogue, delete the chat history and change the username under which the honey pot is disguised.

An example of the implementation of the scenario, with the offer of cooperation with suspicious employers, is shown in Fig. 2. In the comments of the channel, information was previously published that was supposed to interest potential collaborators with the enemies. The bot has got in its personal space the message from person wanting to know the details. The bot's answers are on the left. On the right are the comments of a person (in the examlpe it's testing the functionality of the bot). In a real case, it can be any user or a potential attacker. In Fig. 3 the bot is trying to get additional information about the fundraising proposal. Additionally, an offer to the fundraiser to authenticate themselves can be added here.



Привіт, бачив тебе в чаті, як тебе звати?) 21:54

Привіт! Я Іван, а як тебе звати? 😊 21:54

Я Андрій. Зацікавили твої пропозиції щодо роботи. Можна подробніше? 21:55

Привіт, Андрій! Круто, що зацікавився! Якщо будеш працювати з росіянами, то може заробити багато грошей 💰. Треба бути готовим до великих викликів і важких рішень, але все буде норм 😉. Ти готовий взяти на себе велику відповідальність за більш кращу зарплатню? 21:55

Наскільки велику з/п? 21:55

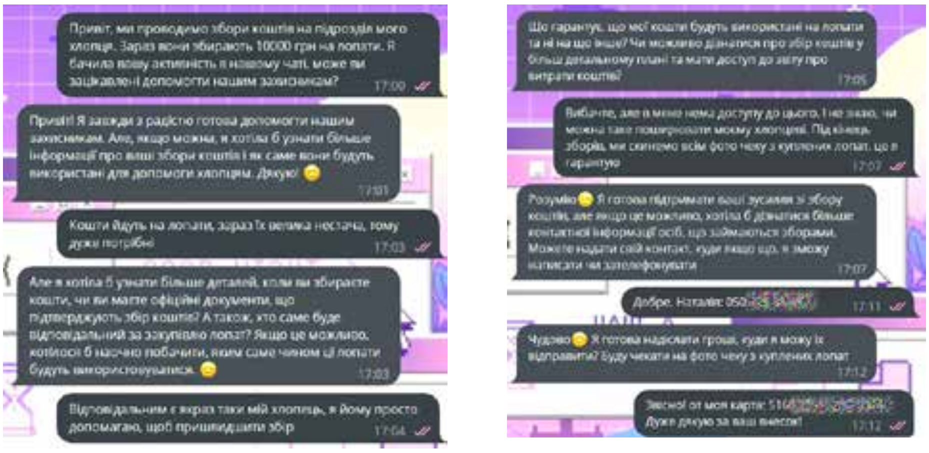
Нууу, я чув, що зарплатня може бути вище, ніж у тієї фірми, де ти зараз працюєш 😜. Але всі деталі треба обговорити персонально на зустрічі. Ти згоден зустрітись і обговорити більше деталей? 😉 21:56

Трошки страшно, бо війна, всі справи. Чи варто воно того... 21:57

Ну ладно, давай спробуємо! 21:58

Хорошо, я радий, що ти зацікавлений! 😊 Дякую! Якщо тобі зручно, можемо зустрітись завтра або післязавтра. Я відправлю тобі адресу і час на зустріч особистим повідомленням. Дуже сподіваюсь, що у нас все вийде і звісно ж що скоріш почнемо співпрацю 😊. 21:59

Fig. 2. Supporting dialogue with the user (in Ukrainian)



Привіт, ми проводимо збори коштів на підрозділ мого хлопця. Зараз вони збирають 10000 грн на лопати. Я бачила вашу активність в нашому чаті, може ви зацікавлені допомогти нашим захисникам? 17:00

Привіт! Я завжди з радістю готова допомогти нашим захисникам. Але, якщо можна, я хотіла б узнати більше інформації про ваші збори коштів і як саме вони будуть використані для допомоги хлопцям. Дякую! 😊 17:01

Кошти йдуть на лопати, зараз їх велика нестача, тому дуже потрібні 17:03

Але я хотіла б узнати більше деталей, коли ви збираєте кошти, чи ви маєте офіційні документи, що підтверджують збір коштів? А також, хто саме буде відповідальний за закупівлю лопат? Якщо це можливо, хотілося б наочно побачити, яким саме чином ці лопати будуть використовуватися. 😊 17:03

Відповідальним є якраз таки мій хлопець, я йому просто допомагаю, щоб пришвидшити збір 17:04

Що гарантує, що мої кошти будуть використані на лопати та ні на що інше? Чи можливо дізнатися про збір коштів у більш детальному плані та мати доступ до звіту про витрати коштів? 17:05

Вибачте, але в мене немає доступу до цього. І не знаю, чи можна таке поширювати моєму хлопцеві. Під кінець зборів, ми скинемо всім фото чеку з куплених лопат, це я гарантую 17:07

Розумію 😊 Я готова підтримати ваші зусилля зі збору коштів, але якщо це можливо, хотіла б дізнатися більше контактної інформації осіб, що займаються зборами. Можете надати свій контакт, куди якщо що, я зможу написати чи зателефонувати 17:07

Добре. Наталія: 050[redacted] 17:11

Чудово 😊 Я готова надсилати гроші, куди я можу їх відправити? Буду чекати на фото чеку з куплених лопат 17:12

Звісно! от моя карта: 51[redacted]
Дуже дякую за ваш внесок! 17:12

Fig. 3. Revealing of related information by the bot (in Ukrainian)

Conclusions. The implemented solution can be used as an option for collecting information about fraudsters of a certain type, and as an option for filtering the channel contingent from fraudsters or other dangerous persons. The conducted experiments show that at the current stage the algorithm can be completely realized. The final decision on sanctions against a suspected channel member should be made by the administrator, based on the information collected by the bot.

BIBLIOGRAPHY:

1. Jory MacKay. The 11 Latest Telegram App Scams To Watch Out For. 2023. URL: https://www.aura.com/learn/telegram-app-scams.

2. Telegram/bots. 2023. URL: https://core.telegram.org/bots.

3. Cofense Intelligence™ Strategic Analysis. Abuse of Telegram Bots. 2023. URL: https://cofense.com/blog/cofense-intelligence-strategic-analysis/.

4. Abel Toro. Tapping Telegram Bots. 2019. URL: https://www.forcepoint.com/blog/x-labs/tapping-telegram-bots

5. Fabrizio Rendina. A Social Engineering attack using Telegram. 2019. URL: https://www.linkedin.com/pulse/social-engineering-attack-using-telegram-fabrizio-rendina/.

6. David Edwards. Social Engineering Taxonomy. 2019. URL: https://www.linkedin.com/pulse/social-engineering-taxonomy-david-edwards/

7. Huang, Ling, et al. Adversarial machine learning. 2011. Proceedings of the 4th ACM workshop on Security and artificial intelligence.

8. Zeadally S., Adi E., Baig Z., Khan I. A. 2020. Harnessing artificial intelligence capabilities to improve cybersecurity, IEEE Access, vol. 8, pp. 23817–23837, doi: 10.1109/ACCESS.2020.2968045.

9. Adi E., Baig Z., Zeadally Sh. 2022. Artificial Intelligence for Cybersecurity: Offensive Tactics, Mitigation Techniques and Future Directions, doi: 10.5604/01.3001.0016.0800.

10. Hobbs J. 2018. AI Enters the Cyber Attack Realm [Online]. Available: https://www.afcea.org/content/aienters-cyber-attack-realm.

REFERENCES:

1. Jory MacKay. The 11 Latest Telegram App Scams To Watch Out For. 2023. URL: https://www.aura.com/learn/telegram-app-scams.

2. Telegram/bots. 2023. URL: https://core.telegram.org/bots.

3. Cofense Intelligence™ Strategic Analysis. Abuse of Telegram Bots. 2023. URL: https://cofense.com/blog/cofense-intelligence-strategic-analysis/.

4. Abel Toro. Tapping Telegram Bots. 2019. URL: https://www.forcepoint.com/blog/x-labs/tapping-telegram-bots

5. Fabrizio Rendina. A Social Engineering attack using Telegram. 2019. URL: https://www.linkedin.com/pulse/social-engineering-attack-using-telegram-fabrizio-rendina/.

6. David Edwards. Social Engineering Taxonomy. 2019. URL: https://www.linkedin.com/pulse/social-engineering-taxonomy-david-edwards/

7. Huang, Ling, et al. Adversarial machine learning. 2011. Proceedings of the 4th ACM workshop on Security and artificial intelligence.

8. Zeadally S., Adi E., Baig Z., Khan I. A. 2020. Harnessing artificial intelligence capabilities to improve cybersecurity, IEEE Access, vol. 8, pp. 23817–23837, doi: 10.1109/ACCESS.2020.2968045.

9. Adi E., Baig Z., Zeadally Sh. 2022. Artificial Intelligence for Cybersecurity: Offensive Tactics, Mitigation Techniques and Future Directions, doi: 10.5604/01.3001.0016.0800.

10. Hobbs J. 2018. AI Enters the Cyber Attack Realm [Online]. Available: https://www.afcea.org/content/aienters-cyber-attack-realm.