Kibertəhlükəsizlik problemlərinin həllində NLP effektiv vasitə kimi

Ramiz Şıxəliyev[1], Fərqanə Abdullayeva[2]
[1,2]İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
[1]*shikhramiz61@gmail.com,* [2]*a_farqana@mail.ru*

***Xülasə* — Məqalədə NLP-nin kibertəhlükəsizlikdə tətbiq sahələri araşdırılıb, əsas tədqiqat istiqamətləri müəyyən edilib. Sosial media mənbələrində hücumçular tərəfindən yazılmış mətn tipli kibertəhdid xarakterli informasiyasının klassifikasiyası üçün NLP-yə əsaslanan yanaşma təklif edilib. Təklif olunan metod əvvəlcə kontentdə kibertəhdidlərlə bağlı sentimental izləri analiz edir, sonra isə proqnozlaşdırır. Yanaşma təşkilatların təhdidlərin prioritetləşdirilməsi, təhdidlərin avtomatik modelləşdirilməsi kimi proaktiv qərarlar qəbul etməsinə imkan yarada bilər.**

Açar sözlər — Kiberhücum; CTI; NLP, Spam, Fişinq, Zərərli URL

I. GİRİŞ

Kibertəhlükəsizlik İnterneti və İnternetə qoşulmuş sistemləri (avadanlıq, proqram təminatı və verilənlər) kiberhücumlardan mühafizə etmək üçün resursların, texnologiyaların, proseslərin və idarəetmə elementlərinin təşkili təcrübəsidir [1]. Kibertəhlükəsizlik informasiyanın və sistemlərin kiberhücumlardan, verilənlərin sızmasından, şəxsi məlumatların oğurluğundan, zərərli proqramlardan, sosial mühəndislikdən, fişinqdən və digər kibertəhlükəsizlik təhdidlərindən mühafizəni əhatə edir [2].

Hər gün İnternetdə çox böyük həcmdə informasiya istehsal olunur və böyük əksəriyyəti strukturlaşdırılmamışdır. Yalnız strukturlaşdırılmış informasiyanın analizi əsasında kibertəhlükəsizlik problemlərini effektiv həll etmək mümkün deyil. Kibertəhlükəsizlik problemlərinin həlli üçün strukturlaşdırılmamış informasiyanın analiz edilməsinə ehtiyac vardır. Strukturlaşmamış informasiyanı analiz etməyə imkan verən bir sıra texnologiyalar mövcuddur, onlardan biri də Təbii Dilin Emalıdır (*ing.* Natural Language Processing, NLP). NLP kompüter elmləri, dilçilik və süni intellektin altsahəsidir. NLP mətn və ya nitq məlumatları şəklində təbii insan dilini emal etmək üçün hesablama linqvistikası, statistika, maşın təlimi və dərin təlim üsullarını birləşdirir. NLP qrammatika, semantika, praqmatika və morfologiyanın müxtəlif aspektlərini təhlil edərək təbii insan dilinin strukturunu və mənasını anlamağa imkan verir. Maşın tərcüməsi, məlumatların çıxarılması, ümumiləşdirmə, tibb, sorğu sistemləri və s. kimi müxtəlif sahələrdə istifadə olunur. NLP-nin kibertəhlükəsizlikdə tətbiq olunduğu ən maraqlı istiqamətlərdən biri kibertəhdidlərin intellektual analizidir (ing. Cyber Threat Intelligence, CTI). Bu istiqamət mətn tipli informasiya mənbələrindən təhdidlər

haqqında informasiyanın avtomatik çıxarılması və analizi ideyasına əsaslanır. Kibertəhdidlərin intellektual analizi sahəsində çox sayda üsullar işlənmişdir [3]. Kibertəhdidlər haqqında verilənlərin analizi mətn mənbələri üzərində həyata keçirilir. Belə mənbələrə kiberhücum strategiyalarının, prosedurlarının və alətlərinin şərh olunduğu təhdid hesabatlarını və onlayn bloq yazılarını, haker forumlarını, sosial media yazılarını, onlayn xəbər yazılarını misal göstərmək olar. Kibertəhdidlərin analizini aparmaq üçün supervizorlu və supervizorsuz maşın təlimi metodları təbii dilin emalı üsulları ilə birgə istifadə olunur. Təqdim olunan məqalədə NLP-nin kibertəhlükəsizlikdə tətbiqləri araşdırılıb, "Twitter" mənbələrində hücumçular tərəfindən yazılmış kibertəhdid xarakterli informasiyasının klassifikasiyası üçün yanaşma təklif edilib.

II. NLP-NİN KİBERTƏHLÜKƏSİZLİKDƏ TƏTBİQLƏRİ

Müdaxilələrin aşkarlanmasının ənənəvi üsulları nümunə uyğunluğundan, ya da qara siyahıdan istifadə edir. Hər iki halda naməlum və ya yeni zərərli hücumlar müəyyən edilə bilinmir. Davranışa əsaslanan üsul daha yüksək dəqiqliklə hücumları aşkar edir. Bu üsul görünməz zərərli trafik kimi bəzi tanınmış hücumları aşkar edə bilmir. NLP-nin istifadəsi aşkarlanma dəqiqliyini artıra bilər, çünki NLP əsaslı aşkarlama mexanizmləri hücum metodlarından asılı olmur [4].

NLP-ni zərərli proqramların aşkarlaması üsulu kimi istifadə etmək mümkündür. Məsələn, NLP ilə icra edilə bilən faylların ASCII (American Standard Code for Information Interchange) sətirlərindən istifadə edərək zərərli proqramların aşkarlanması üsulu təklif edilmişdir [5]. ASCII sətirləri sözlərə bölür və sözlər təhlükəsiz və zərərli kimi təsnif edilir. NLP-dən istifadə edərək tez-tez istifadə olunan sözlərin korpusu xüsusiyyət vektoruna çevrilir. Bu zaman, söz tezliyi ilə müqayisədə NLP qeyri-adi sözlərə ehtiyac duymur.

Zərərli URL-lər (Uniform Resource Locator) bugünkü kiberhücumların və kiberfırıldaqların əksəriyyətində istifadə edilir və e-poçt, mətn mesajları, pop-up-lar və ya reklamlar vasitəsilə istifadəçilərə göndərilir. Bu URL-lərin kliklənməsi və ya skanlanması oğurlanmış e-poçt hesablarına, fişinqə başlanmasına, zərərli proqramların və casus proqramların yüklənməsinə, həmçinin ciddi maliyyə itkilərinə səbəb ola bilər. E-poçtlarda zərərli URL-lərin aşkarlanması üçün maşın təliminə əsaslanan ansambl təsnifatı yanaşması təklif edilmişdir [6]. Bu yanaşmada struturlaşdırılmamış URL sətirlərindən çıxarılan statik leksik xüsusiyyətlərdən istifadə edilir. Bu xüsusiyyətlərdən istifadə aşkarlama dəqiqliyini və sürətini artırır, çünki o, URL-lərin axtarışını və ya qara siyahıya salınmasını tələb etmir.

Spam filtrasiyası texnologiyaları e-poçt mesajlarında spam və ya fişinqi skan etmək üçün NLP mətn təsnifatından istifadə edir [7]. NLP həm spamerlərin şablonlarını və göndərdikləri mesajların növlərini müəyyən

etmək üçün, həm də e-poçtun daxili strukturunu başa düşmək üçün istifadə edilə bilər.

Fişinq kontekstində NLP, özünü insan kimi təqdim edən maşın tərəfindən göndərilən e-poçt mətnindəki botların və ya spamların davranışını anlamaq üçün istifadə edilə bilər [8]. Sosial mediadan əsas resurs kimi istifadə etdikdə fişinq daha da təsirli olur. NLP-dən istifadə sosial şəbəkələr və digər mənbələr vasitəsilə şəxsi məlumatların toplanması prosesini avtomatlaşdırır və bu, hücumun müəyyən edilməsini çətinləşdirir.

Adətən hücumun aşkarlanması və qarşısının alınması üzrə tədqiqatlar şəbəkənin özünün xüsusiyyətlərinə əsaslanan birbaşa sübutlara, məsələn, anormal trafikin monitorinqinə yönəlir. Lakin dolayı sübutlarla hücumların mənbəyi müəyyən edilə bilər: məsələn, sosial media mətni vasitəsilə. Sübut kimi yalnız sosial şəbəkələrdən istifadə edərək DDoS (Distributed denial-of-service) hücumlarının aşkar edilməsi üçün NLP modellərinin istifadəsi təklif edilmişdir [9]. NLP modelləri hücumları aşkar etmək üçün istifadəçi mesajlarından kifayət qədər məlumat çıxarmağa imkan verir.

Şəxsi və konfidensial məlumatlar tez-tez sosial media mesajlarında yer alır. Bu, istifadəçilərin müəyyən sosial qruplarda qarşılıqlı əlaqələrinin sayını artırmaq istəyi, eləcə də məxfilik riskləri barədə zəif məlumatı ilə bağlıdır. Məlumatın məxfiliyini qiymətləndirə bilən texnologiyalar istifadəçini müəyyən məlumatların açıqlanması ilə bağlı risklər barədə xəbərdar etməklə məxfiliyin qorunmasını təmin edə bilər. İstifadəçilər tərəfindən şəxsi məlumatları aşkar etmək üçün istifadə olunan təkrarlanan təbii dil nümunələri vasitəsilə məxfi məlumatların avtomatik tutulması üçün üsul təklif edilmişdir [10].

NLP-nin tətbiq olunduğu sahələrdən biri də Kiber Təhdidlərin Analizidir (Cyber Threat Analysis). Kiber Təhdidlərin Analizi terminina dünyanın aparıcı təşkilatları müxtəlif təriflər vermişdir. SANS İnstitutunun verdiyi tərif aşağıdakı kimidir:

CTI – qarşı tərəf, onun imkanları, tətbiq etdiyi texnikalar və taktikalar haqqında məlumatın toplanması və onların analizidir (SANS Institute).

CTI verilənləri – kibertəhdidləri aşkarlamağa, qiymətləndirməyə, izləməyə, cavab verməyə imkan verən istənilən informasiyadır. Bu informasiya aşağıdakıları əhatə edir:

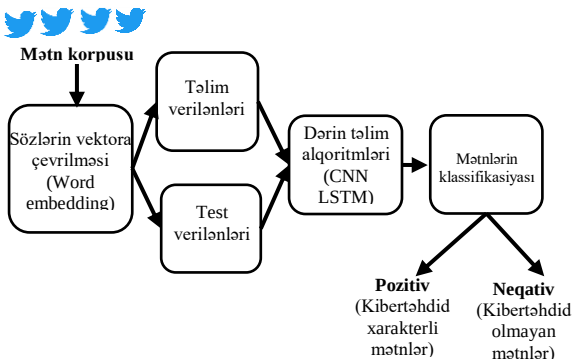
- kompromentasiya indikatorları (file hash, İP address)
- təhdid aktorunun istifadə etdiyi taktikalar, texnikalar və proseduralar
- təhdidləri aşkarlamaq üçün təklif edilən əməliyyatlar
- hücumu həyata keçirmək üçün alətlərin adları
- hücum vektoru.

CTI məlumatlarının çıxarılması üçün aşağıdakı üsullar istifadə edilə bilər:

- CTI əsaslı mətnlərin klassifikasiyası;
- Hücum taktikalarının, texnikalarının, prosedurlarının çıxarılması;

- Kibertəhlükəsizliklə bağlı açar sözlərin çıxarılması;
- Kibertəhlükəsizliklə bağlı hadisələrin identifikasiyası;
- Kibertəhdid xəbərdarlıqlarının generasiyası;
- Haker resurslarının analizi;
- Proqram təminatı boşluqları haqqında informasiyanın çıxarılması;
- Təhdid hesabatlarının klassifikasiyası.

İşdə CTI mətnlərinin klassifikasiyası məsələsinə baxılmışdır. CTI mətnlərinin klassifikasiyası üçün təklif edilmiş *Word2Vec+CNN* modelinin arxitekturası Şəkil 1-də verilmişdir.



Şəkil 1. CTI mətnlərinin klassifikasiyası üçün təklif edilmiş Word2Vec+CNN modelinin arxitekturası

Haker forumları, təhdid hesabatları, sosial media yazıları, onlayn məqalələr CTI mətnlərinin analizi üçün mənbələrdir. Təqdim olunan işdə sosial media saytı olan "Twitter" verilənləri kibertəhlükəsizlik hadisələrinin analizi məqsədi ilə istifadə edilmişdir. Təklif edilmiş yanaşmada sosial media saytlarından mətnlər götürülür, mətnlərin alqoritmlərə tanıdılması üçün onlar ilk növbədə ədədlərə çevrilməlidirlər. Bu prosesi həyata keçirmək üçün NLP-də vektorizasiya əməliyyatından istifadə edilir. İngilis ədəbiyyatlarında bunu "Word Embedding" adlandırırlar. Mətnlərin vektorizasiyasını həyata keçirmək üçün müxtəlif üsullar vardır. Bunlardan biri "Word2Vec" üsuludur. "Word2Vec" modelinin iki növü vardır. *Skip-gram* və *CBOW* (Continuous bag-of-words). Təqdim olunan işdə vektorizasiya skip-gram əsasında aparılmışdır. DDoS hücumunun proqnozlaşdırılması üçün kiberhücumla bağlı terminlər lüğətindən istifadə edilmişdir. İstifadə edilmiş lüğət *UFONet, Low Orbit Ion Cannon (LOIC), Wannacry, Wannacrypt, Petya, Wcry, Petrwrap, Hakerlər, DDoS, DoS, Wikileaks, lulzsec* və s. kimi təhlükəsizliklə bağlı sözlərin adlarını ehtiva edir.

NƏTİCƏ

NLP üsulları strukturlaşdırılmamış mənbələrdən informasiyanın çıxarılmasını avtomatlaşdıra bilər. Texnologiya inkişaf etdikcə, NLP üsulları kibertəhlükəsizlikdə getdikcə daha çox istifadə olunur. NLP kibertəhlükəsizliyin çoxlu sayda müxtəlif məsələlərinin həllinə tətbiq edilə bilər. Nəzərdən keçirilən tətbiqlər kibertəhlükəsizlikdə NLP üsullarının əhəmiyyətini nümayiş etdirir, həmçinin kibertəhlükəsizlik sahəsində NLP-yə ehtiyacı təsdiqləyir.

Kiberhücumların sosial media verilənləri əsasında proqnozlaşdırılması istiqamətində geniş tədqiqatlar aparılmışdır. Mövcud işlər aşkarlama sisteminin siniflәndirilmiş nümunələr və qeyd olunmuş sayda əlamətlər əsasında öyrədilməsi ideyasına əsaslanır. Bu tip yanaşmalar dinamik təbiətli kiberhücumların aşkarlanmasını həyata keçirə bilmir. İşdə sosial media kontentində kibertəhdidlərin identifikasiyası üçün metod təklif edilmişdir. Bu məqsədlə təklif olunan metod əvvəlcə kontentdə kibertəhdidlərlə bağlı sentimental izləri analiz edir, sonra isə proqnozlaşdırır.

İSTİNADLAR

[1] D. Craigen, N. Diakun-Thibault, and R. Purse, "Defining Cybersecurity," Technol. Innov. Manag. Rev., vol. 4, no. 10, pp. 13–21, 2014.

[2] P.S. Seemma, N. Sundaresan, M. Sowmiya, "Overview of Cyber Security," IJARCCE, vol. 7, pp. 125–128.

[3] R.M. Alguliyev, R.M. Aliguliyev, F.J. Abdullayeva, "The improved LSTM and CNN models for DDoS attacks prediction in social media," International Journal of Cyber Warfare and Terrorism (IJCWT), vol. 9, no. 1, pp. 1-18, 2019.

[4] S. Das, M. Ashrafuzzaman, F. T. Sheldon, and S. Shiva, "Network Intrusion Detection using Natural Language Processing and Ensemble Machine Learning," in 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Dec. 2020, pp. 829–835

[5] Ryo Ito and Mamoru Mimura, Detecting Unknown Malware from ASCII Strings with Natural Language Processing Techniques, 14th Asia Joint Conference on Information Security (AsiaJCIS), Kobe, Japan, 09 September 2019.

[6] A. Joshi, L. Lloyd, P. Westin, and S. Seethapathy, "Using Lexical Features for Malicious URL Detection - A Machine Learning Approach," 2019, p. 6.

[7] S. Srinivasan, V. Ravi, M. Alazab, S. Ketha, A. M. Al-Zoubi, and S. Kotti Padannayil, "Spam Emails Detection Based on Distributed Word Embedding with Deep Learning," in Machine Intelligence and Big Data Analytics for Cybersecurity Applications, Y. Maleh, M. Shojafar, M. Alazab, and Y. Baddi, Eds. Cham: Springer International Publishing, 2021, pp. 161–189.

[8] R. Verma, N. Shashidhar, and N. Hossain, "Detecting Phishing Emails the Natural Language Way," in Computer Security – ESORICS 2012, Berlin, Heidelberg, 2012, pp. 824–841.

[9] N. Chambers, B. Fry, and J. McMasters, "Detecting Denial-of-Service Attacks from Social Media Text: Applying NLP to Computer Security," in Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), New Orleans, Louisiana, Jun. 2018, pp. 1626–1635.

[10] G. Canfora, A. Di Sorbo, E. Emanuele, S. Forootani, and C. A. Visaggio, "A Nlp-based Solution to Prevent from Privacy Leaks in Social Network Posts," in Proceedings of the 13th International Conference on Availability, Reliability and Security, New York, NY, USA, Aug. 2018, pp. 1–6.