MASHINALI O‘QITISHDAN FOYDALANIB STEGANOGRAFIK USULLAR ORQALI YASHIRINGAN MA’LUMOTLARNI ANIQLASH METODLARI

Zarif Xudoyqulov
Muhammad al-Alxorazmiy nomidagi
Toshkent axborot texnologiyalari universiteti, PhD, dotsent
zarif.khudoykulov@tuit.uz

Odiljonov Boburbek Abduvaxob o‘g‘li
Muhammad al-Alxorazmiy nomidagi
Toshkent axborot texnologiyalari universiteti, magistrant
boburbek068@gmail.com

ANNOTATSIYA

Ushbu maqola TCP va IP tarmoq protokollari sarlavhalari maydonlarida yashiringan maxfiy ma’lumotlarni aniqlash modeli taqdim etildi. Steganografiyani aniqlash uchun ishlatiladigan an’anaviy vositalar va usullar ste-ganografik paketlarning belgilarini aniqlash qiyin. Shu sababli, tadqiqotda steganografik paketlarning g‘ayritabiiy xatti-harakatlarini aniqlash uchun bir nechta mashinani o‘rganish usullari tajribadan o‘tkazildi. Modellarning imkoniyatlarini baholash uchun to‘g‘ri aniqlangan natijalar soni va ularning umumiy soni o‘rtasidagi bog‘liqlik formulasidan foydalaniladi.

Kalit so‘zlar: *Tarmoq steganografiyasi, TCP/IP sarlavha maydoni, stegopaketlar, steganaliz, mashinani o‘rganish, Tasodifiy o‘rmon.*

ABSTRACT

This paper presented a model for detecting confidential information hidden in TCP and IP network protocol header fields. Conventional tools and methods used for steganography detection are difficult to identify the signatures of steganographic packets. Therefore, several machine learning techniques were experimented in the study to detect abnormal behavior of steganographic packets. To evaluate the capabilities of the models, the relationship between the number of correctly identified results and their total number is used.

Keywords: *Network steganography, TCP/IP header field, stegopackets, ste-ganalysis, machine learning, Random Forest.*

KIRISH

Yashirin ma’lumotlar ataylab yoki qasddan boshqa ma’lumotlar ichida yashiringan ma’lumotlarni anglatadi va bu tarmoq xavfsizligiga jiddiy tahdid solishi mumkin. U fayllar, matn, tasvirlar yoki boshqa ma’lumotlar formatida bo‘lishi mumkin. Kiber-jinoyatchilar ko‘pincha tarmoqqa zarar yetkazishi yoki maxfiy ma’lumotlarni o‘g‘irlashi mumkin bo‘lgan zararli dasturlar, viruslar yoki boshqa zararli fayllarni yashirish uchun yashirin ma’lumotlardan foydalanadilar.

Yashirin ma’lumotlarni aniqlash uchun optik tarmoq xavfsizligi bo‘yicha mutaxassislar steganografiyani aniqlash usullaridan foydalanadilar.

Steganografiyani aniqlash - bu boshqa fayllardagi yashirin xabarlar yoki ma’lumotlarni aniqlash uchun ishlatiladigan usul. Kiber-jinoyatchilar ko‘pincha tasvirlar yoki boshqa ma’lumotlar fayllari ichidagi zararli fayllarni yashirish uchun steganografiyadan foydalanadilar. Steganografiyani aniqlash vositalari har qanday yashirin xabarlar yoki ma’lumotlarni aniqlash uchun faylni tahlil qiladi. Bu tarmoq ichida yashirin bo‘lishi mumkin bo‘lgan zararli kod yoki viruslarning o‘rnatilishini oldini olishga yordam beradi.

Maxfiy ma’lumotlarni yashirish uchun tarmoq protokollaridan foydalanish tobora muhim ahamiyat kasb etmoqda. Turli steganografik algoritmlar yashirin ma’lumotlar oqimi bilan paket ma’lumotlarining mazmunini modulyatsiya qiladi. So‘nggi yondashuvlar ko‘pincha tarmoq protokoli steganografiyasi yoki aniqlash uchun paket xususiyatlaridan foydalanadigan usullar orqali amalga oshiriladi. Biroq, bunday usullar tajovuzkor boshqa steganografiya usullaridan foydalanganda g‘ayritabiiy paketlarni aniqlay olish zarurati bilan bog‘liq murakkab muammoga duch keladi. Bunday hollarda mashinali o‘qitish samaraliroq bo‘ladi.

IP va TCP sarlavha maydonlari tarmoqdagi eng keng tarqalgan steganografik usullardan biridir [1]. Tarmoq steganografiyasi usullari bitta oqimdagi bir qator paketlar orasidagi vaqt mavjudligidan foydalanadi. U ketma-ket paketlarni modulyatsiya qilish uchun ba’zi paketlarni ataylab kechiktirishga asoslangan. TCP va IP sarlavha maydonlari maxfiy ma’lumotlarni yuklash uchun o‘zgartiriladi. Ko‘pgina steganografik usullar maxfiy ma’lumotlarni aniq yashirishi mumkin, ammo ba’zi steganaliz usullari maxfiy ma’lumotlarni ochib berishi mumkin. Ko‘pgina steganografik usullar xabar paketlarining foydali yuklarini o‘zgartiradi. Chunki bu usullarda tarmoq steganografiyasini fosh qilish qiyinroq hisoblanadi, ammo bu usullarda uzatilishi mumkin bo‘lgan maxfiy ma’lumotlarning miqdori nisbatan kamroq. [2] foydali yuklarni o‘zgartirish orqali tarmoq paketi sarlavhasida maxfiy ma’lumotlarni yashirish usulini muhokama qiladi. Bundan tashqari, tarmoqning

ketma-ket paketli oqimlarini o‘zgartirish orqali ma’lumotlarning maxfiyligini ta’minlash mumkin.

Tarmoq steganografiyasining uchta mashhur usuli mavjud [3]. Bular quyidagilar:

- Tarmoq paketining sarlavhasini yoki foydali yukini o‘zgartirish usullari.
- Paketli oqimlar strukturasini o‘zgartirish usullari.
- Gibrid usullar.

Yuqoridagi tarmoq steganografiya usullarini aniqlash usullari 1-rasmda keltirilgan.

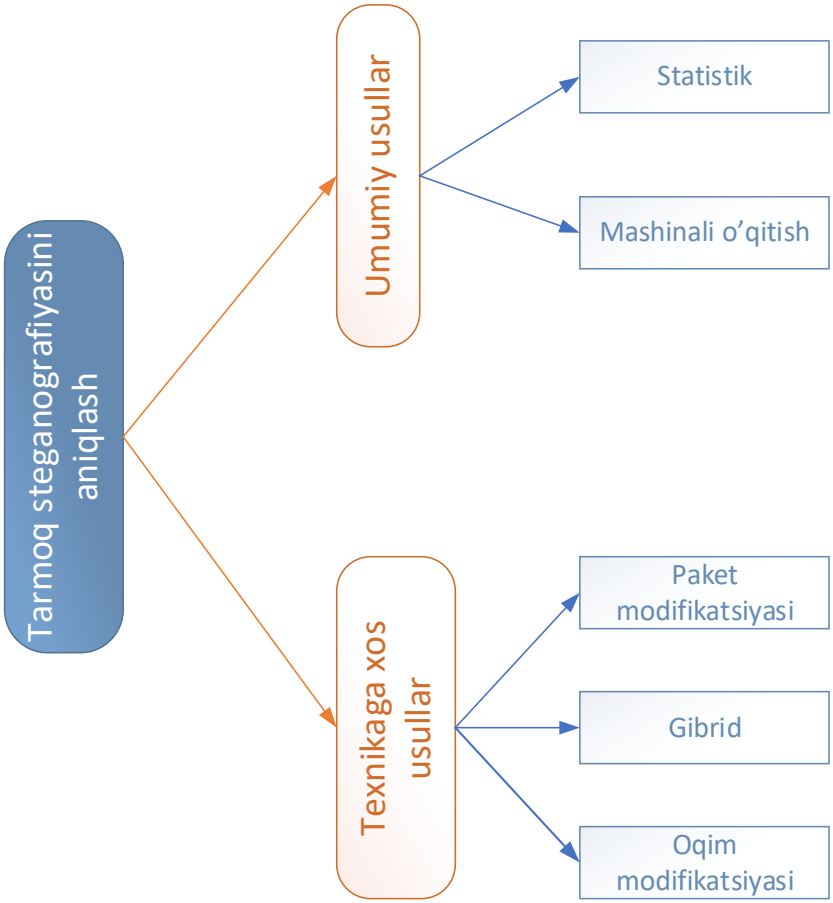
Paketning sarlavha qismiga maxfiy ma’lumotlarni kiritish tarmoq paketining sarlavha maydonlarini o‘zgartirish usulining asosi hisoblanadi [4]. Shu tarzda IP va TCP protokollarining sarlavha maydonlari ko‘pincha ishlatiladi.

MUHOKAMA VA NATIJALAR

Mashinani o‘rganishdan foydalangan holda paket xususiyatlariga asoslangan tarmoq steganografiyasini aniqlash

Steganaliz - maxfiy xabarga ega bo‘lmagan yopiq obyektlarda yashirin ma’lumotlar mavjudligini aniqlash sohasi. Steganalizning ikki turi mavjud: ko‘r steganaliz va maqsadli steganaliz. Ko‘r steganaliz algoritmlari umumiy steganografik algoritmlarni aniqlash uchun mo‘ljallangan, maqsadli steganaliz algoritmlari esa bitta steganografik algoritm uchun mo‘ljallangan. Tarmoq steganaliz paradigmasida bir nechta aniqlash algoritmlari mavjud. Masalan, TCP da ISN ketma-ketlik raqami yoki IP sarlavhasi asosida yashirish [5].

Mashinali o‘qitish asosida tarmoq steganografiyasini aniqlash birinchi navbatda real vaqtda tarmoq trafigida paketlarni yig‘ishni talab qiladi. Wireshark, tcpdump va tshark kabi vositalar tarmoq trafigining qismlarini to‘playdi va tahlil qiladi. Har bir vosita o‘zining afzalliklari va kamchiliklariga ega. Ushbu maqola tarmoq paketlarini olish uchun Wireshark-dan foydalanadi.



1-rasm. **Turli usullar yordamida tarmoq steganografiyasini aniqlash**

Mashinali o‘qitish asosida tarmoq steganografiyasini aniqlash birinchi navbatda real vaqtda tarmoq trafigida paketlarni yig‘ishni talab qiladi. Wireshark, tcpdump va tshark kabi vositalar tarmoq trafigining qismlarini to‘playdi va tahlil qiladi. Har bir vosita o‘zining afzalliklari va kamchiliklariga ega. Ushbu maqola tarmoq paketlarini olish uchun Wireshark-dan foydalanadi.

Yuqorida aytib o‘tilganidek, IP va TCP sarlavha maydonlarini o‘zgartirish usullari va ularni qolgan usullardan ajratib turadigan asosiy xususiyatlar quyidagilardir:

- Transmitter sifatida mashhur va standart protokollardan foydalanish.
- Har bir paket uchun jami 49 bit tarmoqli kengligi oling.
- Agar paket qismlarga bo‘linmasa, undagi o‘zgarishlar uning tarmoqdagi harakatlariga ta’sir qilmaydi [6].

Paket maydonlarini o‘zgartirish uchun hozirgacha taklif qilingan steganaliz usullari quyidagilarni o‘z ichiga oladi:

- Sarlavha va foydali yuk tahlili, shu jumladan Verification Tags qiymatlari tahlili; "oddiy" foydalanuvchilar (steganografiyadan foydalanmaydigan foydalanuvchilar) va

shubhali foydalanuvchilar tomonidan uzatiladigan oqimlarning qiymatlarini taqqoslash; "oddiy" foydalanuvchilar (steganografiyadan foydalanmaydigan foydalanuvchilar) va shubhali foydalanuvchilar tomonidan uzatiladigan Stream Sequence Number qiymatlari o‘rtasidagi taqqoslash; Payload Stream identifikatorining qiymatini tekshiring; a_rwnd qiymatlari va olingan qiymatlarni tahlil qilish; takroriy qismlarning o‘rtacha sonini tahlil qilish; Umumiy kalit identifikatori qiymatlarini tahlil qilish; To‘ldirish ma’lumotlarini tahlil qilish; yuborilgan IP-manzillar mavjudligini tekshirish; oddiy foydalanuvchi (steganografiyadan foydalanmaydigan foydalanuvchi) va shubhali foydalanuvchilar tomonidan uzatiladigan Heartbeat Info Parametr qiymatlarini solishtirish; Tasodifiy sonlarni tahlil qilish; odatiy foydalanuvchi va shubhali foydalanuvchi tomonidan uzatiladigan ASCONF-so‘rov korrelyatsiyasi identifikatori qiymatlari o‘rtasidagi taqqoslash. Yuqoridagi usullar paketni o‘zgartirishning barcha usullarini, jumladan foydali yukni o‘zgartirish, sarlavhani o‘zgartirish va gibrid texnikani o‘z ichiga oladi.

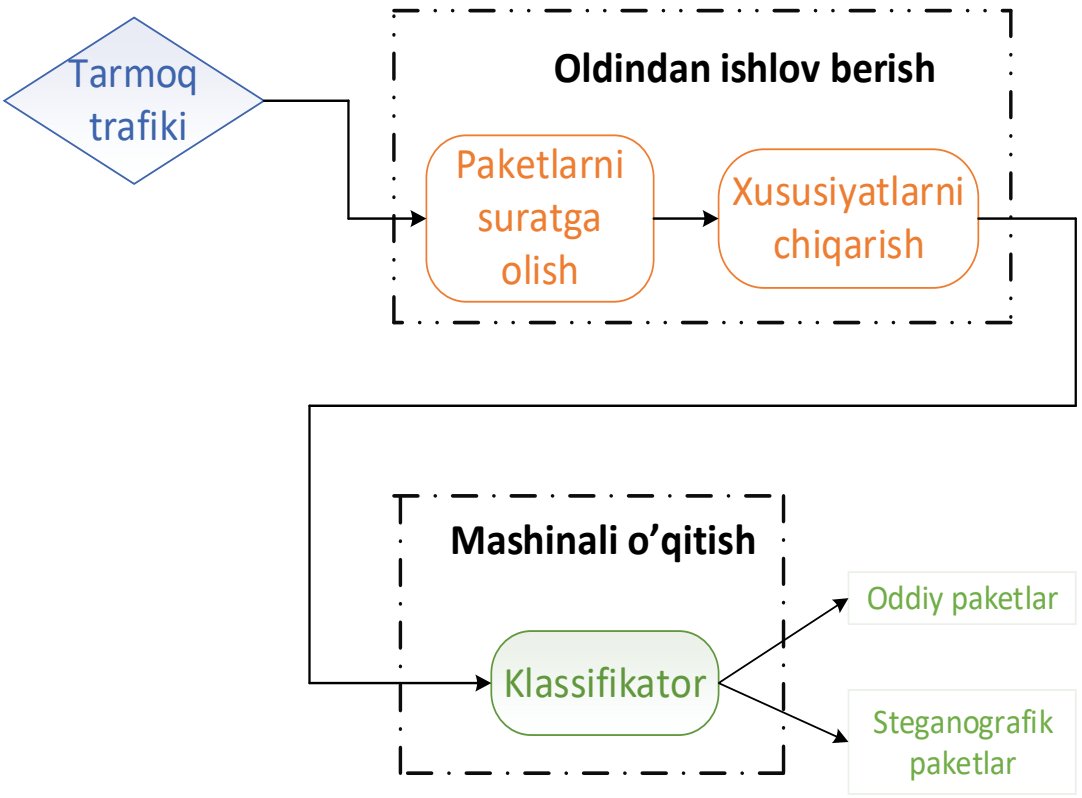
- Sarlavha nazorat summasi monitoringi qayta uzatilgan IEEE 802.11 freymlari uchun nazorat summalarini solishtiradi. Agar nazorat summasi foydali yuk va sarlavha uchun bir xil bo‘lsa va bunday jarayonlar tez-tez sodir bo‘lsa, HICCUPS kabi steganografik texnikadan foydalanilgan bo‘lishi mumkin. Usul sarlavhani o‘zgartirishning steganografik texnikasiga asoslangan.
- TCP tartib raqamining eng muhim bitini kuzatishni o‘z ichiga olgan sarlavha ma’lumotlarining tanlangan asosiy yoki lotin xususiyatlarini kuzatish. Usul sarlavhani o‘zgartirishning steganografik texnikasiga asoslangan.

TAKLIF ETILAYOTGAN USUL

Taklif etilayotgan steganografik paketlarni aniqlash usuli

Maqolada steganografiyani samarali aniqlash uchun tarmoq steganaliz sxemasi keltirilgan. Taklif etilayotgan usulda tarmoq paketi oqimidan funksiyalar to‘plami aniqlanadi va yashirin xabar mavjudligini aniqlash uchun ushbu funksiyalar yordamida nazorat qilinadigan klassifikator o‘rgatiladi. Shuningdek, tavsiya etilgan steganalizning samaradorligini tekshirish uchun keng ko‘lamli tajribalar to‘plami o‘tkazildi.

Mashinani o‘rganish asosida tarmoq steganografiyasini aniqlash uchun taklif qilingan model 2-rasmda keltirilgan.



2-rasm. **Mashinani o‘rganish yordamida steganografik paketlarni aniqlash modeli**

Paketga xos xususiyatlar har bir paketdan alohida yig‘iladi. Paketlarning asosiy anomal xatti-harakati bu xususiyatlardir. Bu standart yoki steganografik paketmi, bu xatti-harakatlardan aniqlanishi mumkin [7]. Maqolada tasvirlangan xususiyatlarni qayta ishlatishdan tashqari, paketlarda yashiringan maxfiy ma’lumotlarni aniqlashni yaxshilash uchun bir nechta yangi xususiyatlar taklif etiladi.

Paketlarga xos xususiyatlarning ma’lumotlar to‘plamini olgandan so‘ng, mashinani o‘rganish algoritmi ushbu yozuvlarni oddiy yoki steganografik paketlar ekanligini ko‘rsatadi. Tarmoq paketlarining anomal xatti-harakatlarini aniqlash uchun mashinani o‘rganishni qo‘llash tanish, ammo tarmoq steganografiyasini aniqlashda keng qo‘llanilmagan. Ushbu dissertatsiya ishimizda Random Forest algoritmlaridan foydalanishni taklif qilamiz.

Quyidagi 1-jadvalda TCP/IP-da tarmoq steganografiyasini belgilaydigan asosiy xususiyatlar ro‘yxati keltirilgan.

1-jadval.

TCP/IP da tarmoq steganografiyasini aniqlash uchun foydalaniladigan xususiyatlar.

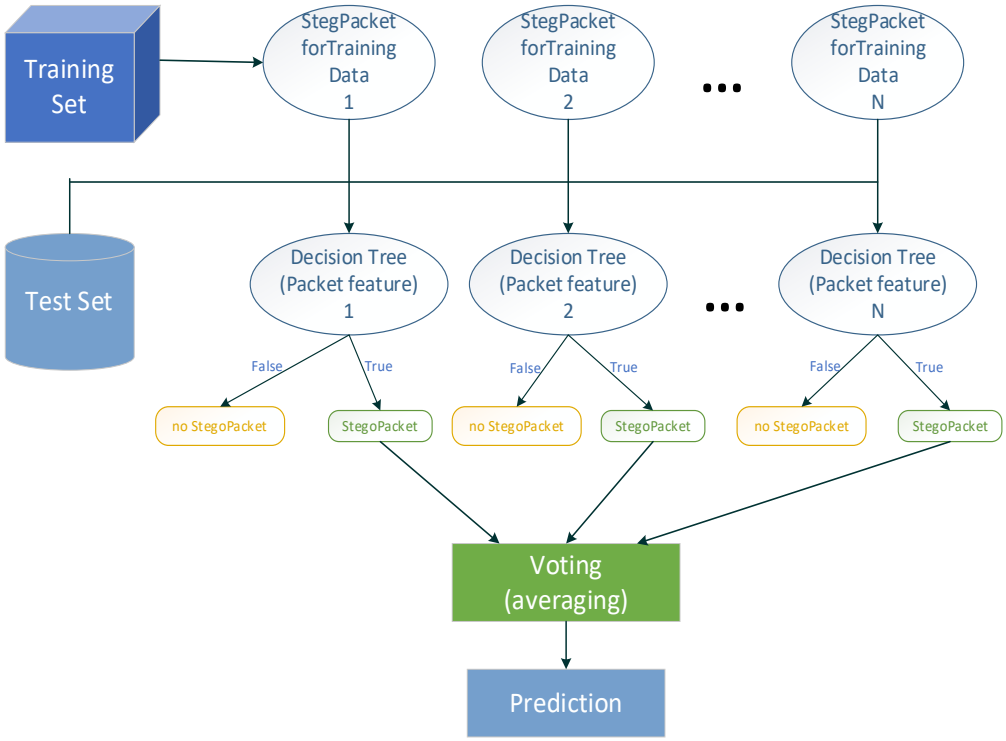
IP_		
Xususiyat	**Tavsif**	**Turi**
_id	Identifikatsiya	Int
_bayroqlar	Bayroqlar	Int
_frag_offset	Fragment ofset	Int
_ nazorat summasi	Sarlavhani tekshirish summasi	Int
_ttl	Yashash vaqti	Int
_tos	Xizmat turi	Int
_src	Manba manzili	Str
_dst	Belgilangan manzil	Str
_son_variant	Variantlar soni	Int
TCP_		
Xususiyat	**Tavsif**	**Turi**
_bayroqlar	Bayroqlar	Int
_ nazorat summasi	Tekshirish summasi	Int
_seq	Tartib raqami	Int
_bayroqlar_res	Zaxiralangan	Bool
_shoshilinch_ko‘rsatkich	Shoshilinch ko‘rsatkich	Int
_ack	Tasdiqlash raqami	Int
_srcport	Manba porti	Int
_dstport	Belgilangan port	Int

Ba’zi TCP/IP sarlavha atributlari xususiyatlari mashinani o‘rganishga asoslangan tarmoq steganografiyasi usullaridan foydalangan holda steganografik paketlarni aniqlash uchun ishlatiladi.

1-jadvalda keltirilgan xususiyatlar faqat TCP sarlavhasi va IP sarlavhasida yashiringan maxfiy ma’lumotlarni aniqlashga qaratilgan. Agar tarmoq steganografiyasining yangi usullari paydo bo‘lsa, bu atributlar yuqori aniqlikdagi aniqlash natijalarini bermaydi, shuning uchun mashinani o‘rganish bu muammoni engib o‘tadi.

Steganografiyani aniqlash usuli samaradorligini oshirish uchun mashinani o‘rganish algoritmlari va modellaridan foydalanish tavsiya etiladi. Bu algoritm va modellarga quyidagilar kiradi: K-Yaqin qo‘shnilar (KNN), Naive Bayes (NB) - Bayes teoremasiga asoslangan shartli ehtimollik klassifikatorlari sinfi. Ular xususiyatlar

orasidagi farazlarning mustaqilligini qabul qiladilar; Vektorli mashinalarni qo‘llab-quvvatlash (SVM) nazorat ostidagi mashinani o‘rganish algoritmi bo‘lib, regressiya va tasnifni modellashtirish orqali ma’lumotlarni tahlil qilishni osonlashtiradi.



3-rasm. **Taklif etilgan steganografik paketni aniqlash uchun tasodifiy o‘rmon algoritmi**

Tasodifiy o‘rmonlar yoki tasodifiy qaror o‘rmonlari tasniflash, regressiya va boshqa vazifalarni o‘rganish uchun texnikadir. Ushbu algoritm ko‘plab qarorlar daraxtlarini qurish orqali ishlaydi. Har bir daraxt alohida ishlaydi va mustaqil natija beradi [10]. Ushbu ishda ushbu usuldan foydalanish tavsiya etiladi (3-rasm).

Trening bosqichida ushbu algoritm birinchi navbatda stegopacket xususiyatlariga ko‘ra olingan namunalarni kiritish bilan boshlanadi va keyin har bir namuna uchun qaror daraxtini yaratadi. Daraxt barcha namunaviy ma’lumotlar va sinov namunalari yordamida yaratiladi. Oxirgi bosqichda daraxtlarni o‘stirish texnikasidan foydalangan holda har bir yuklab olish uchun natijalar bashorat qilinadi.

Testerning ma’lumotlarini bashorat qilish uchun taqsimlash sinov bosqichida amalga oshiriladi. Yakuniy qismda bashorat qilish barcha daraxtlar bo‘ylab bashorat qiluvchilarni o‘rtacha hisoblash orqali amalga oshiriladi.

XULOSA

Xulosa qilib aytganda, maxfiy ma’lumotlarni aniqlash optik tarmoq xavfsizligining muhim jihati hisoblanadi. Kiber jinoyatchilar optik tarmoqqa jiddiy

zarar yetkazishi yoki maxfiy ma’lumotlarni o‘g‘irlashi mumkin bo‘lgan zararli dasturlar yoki boshqa zararli fayllarni yashirish uchun yashirin ma’lumotlardan foydalanishi mumkin. Oxir oqibat, yashirin ma’lumotlarni aniqlash va tarmoq xavfsizligini oshirishning kaliti hushyor va faol bo‘lishdir. Muntazam ravishda fayllar va hujjatlarni skanerlash, tarmoq trafigini kuzatish va dasturiy ta’minotni yangilab turish yashirin ma’lumotlarning aniqlanishi va tarmoq xavfsizligini ta’minlashga yordam beradi. Ushbu choralarni ko‘rish orqali korxonalar va jismoniy shaxslar o‘zlarini kiberjinoyat va ma’lumotlar buzilishi tahdididan yaxshiroq himoya qilishlari mumkin.

ADABIYOTLAR RO‘YXATI

[1] B. Xu, J. Z. Wang, and D. Y. Peng, “Practical Protocol Steganography: Hiding Data in IP Header,” *Proceedings - 1st Asia International Conference on Modelling and Simulation: Asia Modelling Symposium 2007, AMS 2007*, pp. 584–588, 2007, doi: 10.1109/AMS.2007.80.

[2] K. Szczypiorski, “A performance analysis of HICCUPS - A steganographic system for WLAN,” *1st International Conference on Multimedia Information Networking and Security, MINES 2009*, vol. 1, pp. 569–572, 2009, doi: 10.1109/MINES.2009.248.

[3] A. S. Nair, A. Sur, and S. Nandi, “Detection of packet length-based network steganography,” *Proceedings - 2010 2nd International Conference on Multimedia Information Networking and Security, MINES 2010*, pp. 574–578, 2010, doi: 10.1109/MINES.2010.126.

[4] C. H. Rowland, “Covert Channels in the TCP/IP Protocol Suite,” *undefined*, vol. 2, no. 5, 1997, doi: 10.5210/FM. V2I5.528.

[5] S. Cabuk, C. E. Brodley, and C. Shields, “IP Covert Channel Detection,” *undefined*, vol. 12, no. 4, Apr. 2009, doi: 10.1145/1513601.1513604.

[6] A. Noskov and F. Jakab, “Analysis of network protocols the ability of concealing the information,” *ICETA 2016 - 14th IEEE International Conference on Emerging eLearning Technologies and Applications, Proceedings*, pp. 245–249, Dec. 2016, doi: 10.1109/ICETA.2016.7802098.

[7] D. X. Cho, D. T. H. Thuong, and N. K. Dung, “A Method of Detecting Storage Based Network Steganography Using Machine Learning,” *Procedia Comput Sci*, vol. 154, pp. 543–548, 2018, doi: 10.1016/J.PROCS.2019.06.086.