HUJUMLARNI ANIQLASH TIZIMLARINI TAKOMILLASHTIRISH UCHUN BINAR TASNIFLAGICHLAR ALGORITMINI TAHLIL QILISH

Shirinov Bobir Botir o‘g‘li,
Muhammad al-Xorazmiy nomidagi TATU,
Axborot xavfsizligi kafedrasi assistenti,
sh.bobur1994@gmail.com
Shukurov Orziqul Pardayevich,
Muhammad al-Xorazmiy nomidagi TATU,
Axborot xavfsizligi kafedrasi assistenti,
orzubek90@bk.ru

Annotatsiya: Maqolada tarmoq hujumlarini aniqlash tizimlarining tuzilishi va ishlashining o’ziga xos xususiyatlari hamda tarmoqlarda paketlarni tahlillovchi algoritmlarni takomillashtirish imkonini beruvchi usullar ko’rib chiqilgan. Shuningdek, hujumlarni aniqlash tizimini takomillashtirishda binar klassifikatorlar usulidan foydalanish algoritmi orqali taklif etilgan model ko’rsatkichlarini baholash tahlil qilingan.

Kalit so’zlar: hujumlarni aniqlash tizimi, tarmoq paketlari, klassifikatorlar, detektorlar, kompyuter tarmoqlari, sensorlar, model va algoritmlar

Kirish. Hujumlarni aniqlash tizimini (IDS) ishlab chiqish axborot xavfsizligi sohasidagi ustuvor yo'nalishlardan biridir. Ushbu muammoni hal qilishning ahamiyati kompyuter tarmoqlari tahdidlarining doimiy o'sishi va xilma-xilligi bilan bog'liq bo'lib, ularning amalga oshirilishi turli tashkilotlarda jiddiy moliyaviy yo'qotishlarga olib kelishi mumkin. Kasperskiy laboratoriyasi statistik ma’lumotlariga ko‘ra, 2017-yilning birinchi choragida 479 milliondan ortiq kompyuter hujumlari aniqlangan va bartaraf etilgan bo‘lsa, 2018-yilning shu davrida bu ko‘rsatkich 796 million hujumdan oshgan [1]. Har yili hujumlarning bunday ko'payishi administratorlar va xavfsizlik bo'yicha tahlilchilar tomonidan sezilarli darajada ko'proq kuch va vaqt sarflashni talab qiladi. Korporativ tarmoq resurslarining xavfsizligini ta'minlash uchun muhim mahsulotlarni ishlab chiqarish bilan shug'ullanadigan kompaniyalar IDS komponentlari va unga xizmat ko'rsatadigan xodimlar ko'rinishidagi maxsus uskunalarni saqlashga qaratilgan katta moliyaviy va moddiy resurslarni sarflaydilar [9].

Shu sababli, anomal tarmoq ulanishlarini aniqlash vazifasi dolzarb bo'lib, ushbu maqolada taklif qilingan, hisoblash intellektining (HI) turli toifali usullari va signatura tahlilining kombinatsiyasidan (gibrid usuli) foydalanadigan yechim taklif etiladi.

Adabiyotlar sharhi: Ushbu muammolar haqida asosan chet el olimlari va professorlari ilmiy izlanish ishlarini olib borganlar va tezis, ilmiy maqolalar hamda bir qancha kitob nashrlarida o'z tajribalarini keltirib o'tishgan.

Bularga misol sifatida: Scarfone, Karen; Mell, Peter "Guide to Intrusion Detection and Prevention Systems (IDPS)" , Engin Kirda; Somesh Jha; Davide Balzarotti " Recent Advances in Intrusion Detection" , Denning, Dorothy E., "An Intrusion Detection Model" , Vaccaro, H.S., and Liepins, G.E., "Detection of Anomalous Computer Session Activity" , Singh, Abhishek. "Evasions In Intrusion Prevention Detection Systems" kabi nashrlarni keltirib o'tishimiz mumkin.

Materiallar va usullar. Anormal tarmoq ulanishlarini aniqlash uchun taklif etilayotgan metodologiyaning umumiy strukturasi quyidagi besh bosqichdan iborat:

1. klassifikatorlar daraxtini qurish;
2. tarmoq ulanishlari parametrlarini shakllantirish;
3. tarmoqqa ulanish parametrlarini oldindan qayta ishlash;
4. klassifikatorlar daraxtining kenglikdagi ierarxik o'tishi;
5. anomal tarmoq ulanishlarini aniqlash.

IDS komponentlarini operator va tizim tomonidan bajariladigan harakatlarni ko'rsatadigan ushbu usulning har bir bosqichiga batafsilroq to'xtalib o'tamiz.

Usbu usulning birinchi bosqichi tayyorgarlik sifatida tavsiflanishi mumkin, u individual ikkilik tasniflagichlar (detektorlar) tuzilishini tanlashni o'z ichiga oladi: qatlamlarning o'lchamlari va soni, o'qitish parametrlari va algoritmlari, faollashtirish funksiyalari turlari va yadro funksiyalari. Har bir detektor uchun o'rganish qoidalari to'plamini tuzish mumkin [2]. Har bir bunday guruhdagi detektorlar bittaga-hamma (one-vs-all), birga-bir (one-vs-all) yondashuvlar yoki ularning turli xil kelib chiqishiga asoslangan tasniflagichga birlashtiriladi.

Birinchi yondashuvda har bir $F_{jk}^{(i)}: R^n \to \{0,1\}$ $(k = 1, \dots, m)$ detektor $\{(\boldsymbol{x}_l, [c_l = k])\}_{l=1}^{M}$ ma'lumotlarga o'rgatiladi va $F_j^{(i)}$ detektorlar guruhining ishlashi istisno prinsipi yordamida tasvirlanadi:

$$F_j^{(i)}(z) = \begin{cases} \{0\} \\ \left\{k \mid F_{jk}^{(i)}(z) = 1\right\}_{k=1}^{m}, \end{cases} \quad \text{agar } \forall k \in \{1, \dots, m\} \; F_{jk}^{(i)}(z) = 0$$

Ikkinchi yondashuvda har bir $C_{m+1}^2 = \frac{(m+1)*m}{2}$ detektorlar $F_{jk}^{(i)}$ yorlig'i bilan faqat ikkita $k0$ va $k1$ sinfga mansub ob'ektlar to'plamiga o'rgatiladi, bu yerda $0 \le k_0 < k_1 \le m$ va $F_j^{(i)}$ detektorlari guruhining ishlashi max-wins formulasi bilan beriladi:

$$F_j^{(i)}(z) = \left\{ arg \max_{c \in \{0,\dots m\}} \sum_{k=c+1}^{m} \left[F_{jck}^{(i)}(z) = 0\right] + \sum_{k=0}^{c-1} \left[F_{jkc}^{(i)}(z) = 1\right] \right\}.$$

1-jadvalda kirish ob'ektini bir yoki bir nechta (m + 1) sinf belgilari bilan bog'lash uchun mo'ljallangan ko'p sinfli modelga detektorlarni birlashtirish uchun ko'rib chiqilgan sxemalarning xususiyatlari ko'rsatilgan.

1-jadval

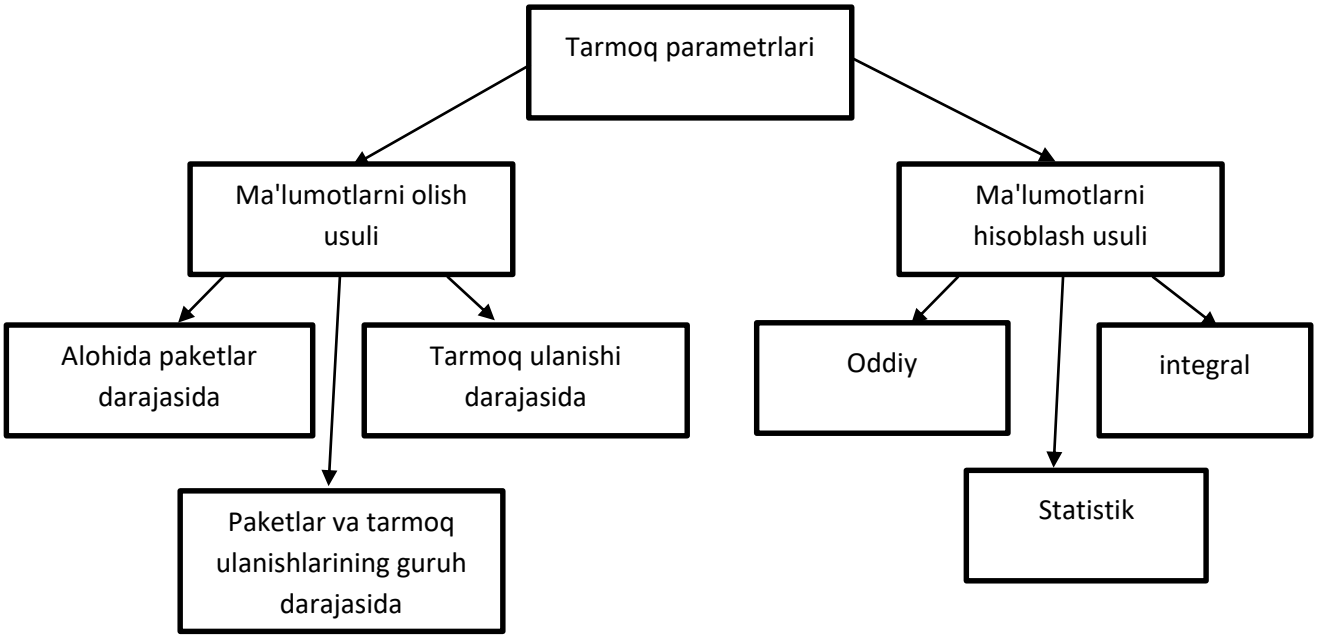
Detektorlar birlashma sxemalarining xarakteristikalari

Birlashma sxemasi	O'rganiladigan detektorlar soni	Ob'ektlarni tasniflashda ishtirok etadigan detektorlarning minimal soni	Tasniflashda ishtirok etadigan detektorlarning maksimal soni
Birga-hamma	m	m	m
Birga-bir	$\frac{(m+1)*m}{2}$	$\frac{(m+1)*m}{2}$	$\frac{(m+1)*m}{2}$
Birga-bir	m	1	m
Yo'naltirilgan asiklik grafik	$\frac{(m+1)*m}{2}$	m	m

Ushbu metodika uni amalga oshiradigan tizimlarning taqsimlangan arxitekturasini nazarda tutadi, bunda ma'lumotlar ikkilamchi tugunlar - sensorlar tomonidan to'planadi va yig'ilgan ma'lumotlar oqimining barcha qayta ishlanishi markazlashtirilgan server - kollektorda amalga oshiriladi [7].

Detektorlar tomonida bajariladigan metodikaning ikkinchi bosqichi ishlov berilmagan paketlarni tarmoq ulanishlariga yig'ish uchun ishlab chiqilgan algoritmni qo'llash, ularning parametrlarini ajratib olish va pastki qator namunalarini qidirish algoritmlarining bir nechta ishlab chiqilgan parallel modifikatsiyalari yordamida signatura tahlilini o'tkazishdan iborat. Shu maqsadda tanlangan Snort signaturaga asoslangan Aho-Korasik va Boyer-Mur algoritmlarining ishlashi o'rganildi va ularning takomillashtirilgan analoglari OpenMP va CUDA texnologiyalari yordamida amalga oshirildi [4].

Hodisalarga asoslangan tarmoq trafigi analizatori amalga oshirildi, uning yordamida 106 ta tarmoq parametrlari orasida ulanish davomiyligi, foydalanilgan tarmoq xizmati, xost tomonidan maxsus paketlarni jo'natish intensivligi, faol ulanishlar soni, ma'lum bir IP-manzillar juftligi o'rtasida (DoS-hujumlaridan biri), TCP ulanishining hozirgi holati, mavjudligining turli belgilari. TCP, UDP, ICMP va IP darajalarida paketlarni skanerlash (15 xil skanerlash kodlari) tahlil qilindi. Ushbu parametrlarning tasnifi 1-rasmda ko'rsatilgan.



1-rasm. Tarmoq parametrlarining tasnifi

Metodikaning uchinchi bosqichi RPC / SSL protokoli orqali uzatiladigan va ulanish parametrlarini o'z ichiga olgan sensorlardan keladigan paketlarni aniqlashdan boshlanadi. Kollektor va sensorlarning o'zaro ta'sirini ta'minlash uchun tanlov protokollarning ushbu to'plamiga to'g'ri keladi, chunki ular tezkor va xavfsiz ma'lumotlarni jo'natishni kafolatlaydi. RPC muvaffaqiyatli vaqt sinovidan o'tgan texnologiya bo'lib, u ikkilik ma'lumotlar oqimlarining ixcham uzatilishini osongina tashkil etish imkonini beradi. SSL, o'z navbatida, ma'lumotlarni uzatuvchilar o'rtasida shifrlangan kanal yaratish uchun keng qo'llaniladi [3].

Detektorlarni to'g'ridan-to'g'ri o'rganishdan oldin, bu parametrlar ularning kuchli o'zgaruvchanligi ta'sirini kamaytirish uchun oldindan qayta ishlanadi. Hisoblash resurslari bo'yicha metodologiyaning to'rtinchi bosqichi eng ko'p vaqt talab qiladi va quyidagi rekursiv takrorlanadigan ketma-ketliklardan iborat: joriy klassifikatorning bog'liqliklarini hisoblash, joriy klassifikator uchun kirish signallarini yaratish, joriy klassifikatorni o'rganish [6].

Klassifikatorlarni saqlash uchun maxsus daraxt strukturasi ishlab chiqilgan bo'lib, u barcha bog'liqlik zanjirlari bo'ylab yuqori darajadagi klassifikatordan detektorlar bilan ifodalangan terminal tugunlarigacha samarali pasayish imkonini beradi. Har bir klassifikatorni o'rganish uning bog'liqliklari ro'yxatida ko'rsatilgan quyi klassifikatorlarni o'rganish uchun so'rovni hosil qiladi va yuqori klassifikatorning kirish ma'lumotlarini shakllantirish uchun ularning chiqish ma'lumotlarini yaratadi [7].

Metodikaning beshinchi bosqichi ikkita rejimni o'z ichiga oladi: samaradorlikni baholash rejimi va ish rejimi. Birinchi rejimda tasniflash modellarining sifatini baholash ko'rsatkichlari hisoblab chiqiladi, ikkinchi rejimda tizim diagnostikasi

aniqlangan tarmoq ulanishining haqiqiy sinfini ehtimoliy qiymat orqali amalga oshiradi.

Tahlil va natijalar. G tasniflash modeli sifatini baholash uchun quyidagi ko'rsatkichlar aniqlandi:

1. *TPR* - tarmoq hujumini aniqlashning aniqlik darajasi:

$$TPR = \frac{TP}{TP+FN} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \neq 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \neq 0\}_{i=1}^{M^*}},$$

bu yerda $TP = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \neq 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}$ - to'g'ri hisoblangan anomal birikmalar soni, $FN = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \neq 0\}_{i=1}^{M^*}$ - *TP* - II turdagi xatolar soni (o'tkazib yuborilgan hujumlar soni).

2. *FPR* - noto'g'ri ijobiy ko'rsatkich:

$$FPR = \frac{FP}{FP+TN} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} = 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} = 0\}_{i=1}^{M^*}},$$

bu yerda $FP = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} = 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}$ – I turdagi xatolar soni (noto'g'ri ijobiylar soni), $TN = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} = 0\}_{i=1}^{M^*}$ - *FP* - to'g'ri hisob-langan oddiy ulanishlar soni.

3. *CCR* - birikmalar tasnifining to'g'rilik ko'rsatkichi:

$$CCR = \frac{CC}{TP+FN+FP+TN} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \overline{\{c\}} = G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)}\}_{i=1}^{M^*}},$$

bu yerda $CC = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \overline{\{c\}} = G(z_i)\}_{i=1}^{M^*}$ - normal va anomal birikmalardan tashkil topgan birlashtirilgan ma'lumotlar to'plamida sinfi to'g'ri aniqlangan elementlarning umumiy soni.

4. *ICR* - noto'g'ri tasniflash darajasi:

$$ICR = \frac{IC}{TP+FN+FP+TN} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \notin G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)}\}_{i=1}^{M^*}},$$

bu yerda $IC = \#\{z_i \mid z_i \in X_C^{(TS)} \wedge \bar{c} \notin G(z_i)\}_{i=1}^{M^*}$ - noto'g'ri tasniflash holatlari soni. Bu holatda 1 - *CCR* - *ICR* qiymati 0 ga teng bo'lishi shart emas va uni to'g'ri tasniflashning qarama-qarshi ko'rsatkichi sifatida talqin qilish mumkin.

5. *GPR* – aniqlashda umumlashtiruvchi qobiliyat ko'rsatkichi:

$$GPR = \frac{\overline{TP^*}}{\overline{TP^*}+\overline{FN^*}} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \setminus X_C^{(LS)} \wedge \bar{c} \neq 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)} \setminus X_C^{(LS)} \wedge \bar{c} \neq 0\}_{i=1}^{M^*}},$$

bu yerda $\overline{TP^*}, \overline{FN^*}$ ko'rsatkichlari mos ravishda to'g'ri hisoblangan anomal birikmalar soni va II turdagi xatolar soni, M^* kardinalligining $X_C^{(TS)}$ boshqaruv majmuasining yagona ma'lumotlari bo'yicha hisoblanadi, $X_C^{(LS)}$ - ta'lim to'plamining har qanday ma'lumotlarini qat'iyan istisno qilish holati.

6. *OPR* – aniqlashda haddan tashqari moslashish darajasi:

$$OPR = \overline{TPR} - GPR = \frac{\#\{z_i \mid z_i \in X_C^{(LS)} \wedge \bar{c} \neq 0 \wedge 0 \notin G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(LS)} \wedge \bar{c} \neq 0\}_{i=1}^{M^*}} - GPR,$$

bu yerda $\overline{TPR}$ - $\bar{M}$ kardinallikning $X_C^{(LS)}$ boshqaruv majmuasining yagona ma'lumotlari bo'yicha aniqlashning to'g'riligi ko'rsatkichi.

7. *GCR* – tasniflashda umumlashtirish qobiliyati ko'rsatkichi:

$$GCR = \frac{\overline{CC^*}}{\overline{TP^*}+\overline{FN^*}+\overline{FP^*}+\overline{TN^*}} = \frac{\#\{z_i \mid z_i \in X_C^{(TS)} \setminus X_C^{(LS)} \wedge \overline{\{c\}} = G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(TS)} \setminus X_C^{(LS)}\}_{i=1}^{M^*}},$$

bu yerda $\overline{CC^*}$, $\overline{FP^*}$, $\overline{TN^*}$ mos ravishda to'g'ri tasniflangan birikmalar soni, I turdagi xatolar soni va to'g'ri tan olingan normal birikmalar soni, M^* kardinalligining $X_C^{(TS)}$ boshqaruv majmuasining yagona ma'lumotlari bo'yicha hisoblanadi, $X_C^{(LS)}$ - ta'lim to'plamining har qanday ma'lumotlarini qat'iyan istisno qilish holati.

8. *OCR* – tasniflashda haddan tashqari moslashish darajasi:

$$OCR = \overline{CCR} - GCR = \frac{\#\{z_i \mid z_i \in X_C^{(LS)} \wedge \{\bar{c}\} = G(z_i)\}_{i=1}^{M^*}}{\#\{z_i \mid z_i \in X_C^{(LS)}\}_{i=1}^{M^*}} - GCR,$$

bu yerda $\overline{CCR}$ - M^* kardinalligining $X_C^{(TS)}$ boshqaruv majmuasining yagona ma'lumotlari bo'yicha tasniflash to'g'rilik ko'rsatkichi.

Xulosa va takliflar

1. Detektorlarni birlashtirish uchun beshta blokli crossvalidatsiya va past darajadagi one-vs-all sxemasi qo'llanilganda, $GCR - ICR$ ko'rsatkichining 1.275% ga oshishi eksperimental tarzda olingan. Ushbu metodikani ob'ektni tasniflashning umumiy muammolarini hal qilishda ob'ektlarning alohida intellektual yadrosi sifatida IDS dasturiy ta'minotni amalga oshirishda foydalanish mumkin.

2. IDS qurish uchun ishlab chiqilgan model foydalanish bo'yicha tavsiyalar tarmoq hujumlarini aniqlash uchun parallelizatsiya mexanizmlaridan foydalanish, IDS uchun asos sifatida mashina kodidan foydalanish va xotiraga kirishni optimallashtirishga qaratilgan yondashuvlarni ishlab chiqishni o'z ichiga oladi. Ushbu izlanishda olingan tajriba natijalari ushbu tavsiyalardan foydalanishning maqsadga muvofiqligini tasdiqlaydi.

3. Izlanishni yanada rivojlantirish istiqbollari yangi turdagi anomaliyalarga xos xususiyatlarni hisobga olgan holda hisoblangan tarmoq parametrlari ro'yxatini kengaytirish va tarmoq hujumlarining zamonaviy sinflariga moslashtirish uchun taklif etilayotgan IDS arxitekturasini takomillashtirishdan iborat.

Foydalanilgan adabiyotlar:

1. Usmanbayev Daniyorbek Shukhratovich. (2022). ANALYSIS OF EXISTING THREATS AND VULNERABILITIES IN COMPUTER NETWORKS. Academicia

Globe: Inderscience Research, 3(10), 53–56. https://doi.org/10.17605/OSF.IO/XVGRH .

2. Shukhratovich, U. D. (2022). Specific Features Of The Structure And Operation Of Network Attack Detection Systems. Open Access Repository, 8(04), 224-228.

3. D. Usmanbayev, "Improving and Evaluating Methods Network Attack Anomaly Detection," 2021 International Conference on Information Science and Communications Technologies (ICISCT), 2021, pp. 1-5, https://doi.org/10.1109/ICISCT52966.2021.9670073

4. Chandrasekhar, A. M. Intrusion detection technique by using k-means, fuzzy neural network and SVM classifiers / A. M. Chandrasekhar, K. Raghuveer // In Proceedings of International Conference on Computer Communication and Informatics (ICCCI). - IEEE. 2013. - Pp. 1-7.

5. Ennert, M. Testing of IDS model using several intrusion detection tools / M. Ennert, E. Chovancova, Z. Dudlakova // Journal of Applied Mathematics and Computational Mechanics. - 2015. - Vol. 14, no. 1. - Pp. 55-62.

6. Chunayev, N., & Shirinov, B. (2023). EXPERIMENTAL CHARACTERIZATION OF FILTERING MODEL DISPLAY PROCEDURE NUMBER. INTERNATIONAL CONFERENCES, 1(21), 17–22. Retrieved from http://erus.uz/index.php/cf/article/view/943

7. Pardayevich, S. O. (2022). AXBOROT XAVFSIZLIGI RISKLARI TASNIFI VA BAHOLASH USULLARI. Komputer texnologiyalari, 1(10).