Şəbəkə trafikinin Naive Bayes və neyron şəbəkə əsaslı paralel klassifikasiya modeli

Yadigar İmamverdiyev[1], Babək Nəbiyev [2]
[1,2]AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
[1]yadigar@lan.ab.az, [2]babek@iit.ab.az

Xülasə – Trafikin monitorinqi sahəsində tədqiqatların icmalını apararkən, bu istiqamət üzrə bir çox tək mərhələli klassifikasiya sisteminə rast gəlmək mümkündür. Bu metodlar əsasən Naive Bayes və neyron şəbəkə əsaslı klassifikatorlardır. Onlardan ayrı-ayrılıqda sürətli və dəqiq klassifikasiya nəticələri əldə etmək üçün istifadə olunur. Bu məqalədə operativliyi və ya dəqiqliyi azaltmadan klassifikasiya xarakteristikalarını yüksəltmək məqsədilə iki mərhələli klassifikator təklif edilir.

Açar sözlər - şəbəkə trafiki, trafikin klassifikasiyası, naive bayes, feed-forward neyron şəbəkəsi

I. GİRİŞ

Trafikin dəqiq klassifikasiyası şəbəkə fəaliyyətinin təməlidir. İnformasiya və kommunikasiya texnologiyalarının sürətli inkişafı və geniş yayılması, rəqabətin kəskinləşməsi informasiya təhlükəsizliyinin təmin edilməsinin elmi-metodoloji prinsiplərinə əsaslanaraq və şəbəkə texnologiyalarının müasir inkişaf meyillərini nəzərə alaraq hüquqi, təşkilati, texniki və fiziki mühafizə tədbirlərini qarşılıqlı surətdə əlaqələndirməklə, korporativ şəbəkələrdə informasiya təhlükəsizliyinin təmin olunması üçün şəbəkə trafikinin tədqiqi vacibdir. Bu səbəbdən şəbəkələr, digər mürəkkəb sistemlər kimi fasiləsiz monitorinqə ehtiyac duyur. Monitorinq sistemi şəbəkənin vəziyyətinin proqnozlaşdırılması, kəsilmələrin-itkilərin idarə olunması və qabaqlayıcı tədbirlər görülməsi üçün hazırlanmalıdır.

Hücumların aşkarlanması şəbəkə texnologiyalarında ən aktual məsələlərdən biridir. DARPA təşkilatının verdiyi məlumatlara görə təhlükəsizliyi təmin olunmamış və İnternetə qoşulmuş kompüter 2-3 saatdan sonra artıq virusa yoluxmuş olur [1]. Bundan başqa, 2013-cü il ərzində dünyada baş verən informasiya təhlükəsizliyi hadisələrinin 32%-i DDoS hücumlarının payına düşür. Bu növ təhdidlərin aşkarlanması üçün avtomatlaşdırılmış alətlərdən (firewall, brandmauer, IDS və.s) istifadə olunur. Amma bu növ alətlər operativ əks-əlaqəni və şəbəkə trafikinin proqnozlaşdırılması funksiyalarını təmin edə bilmirlər.

Trafikin monitorinqi metodları şəbəkə trafikini daim nəzarətdə saxlayır və hər hansı bir hadisə baş verdiyi halda bu barədə məlumat verir. Bu bir növ antivirusun işləmə prinsipinə bənzəyir – sensorlardan keçən paketlər siqnaturlar bazası ilə müqayisə olunur və uyğunluq olduğu halda müəyyən tədbirlər görülür. Təəssüflər olsun ki, bu metod şəbəkə təhlükəsizliyi üçün etibarlı sayıla bilməz. Zərərli proqramların analizi və tədqiqi üzrə ixtisaslaşmış UNAM-CERT-in verdiyi məlumata görə, qurbanlardan hər həftə 2500-ə yaxın zərərli proqram aşkarlanır və bunların 15-i yeni növ təhdidlər olur. Bunları nəzərə aldıqda, siqnatur bazasının yenilənməsi prosesi çətin bir prosesə çevrilir. Bundan savayı, siqnatura bazasının böyüməsi ilə trafikin yoxlanılması prosesi də ləngiyərək sistemin effektivliyini aşağı salır. Bu problemin həlli, trafikin monitorinqi zamanı klassifikasiya prosesini apararaq anomal aktivliyin mənsubiyyətinin ilkin mərhələdə aşkarlanması nəzərdə tutulur.

Trafikin monitorinqi sahəsində tədqiqatların icmalını apararkən, bu istiqamət üzrə bir çox bir mərhələli klassifikasiya sisteminə rast gəlmək mümkündür. Bu metodlar əsasən Naive Bayes və neyron şəbəkə əsaslı klassifikatorlardır. Bu metodlar ayrı-ayrılıqda sürətli və dəqiq klassifikasiya nəticələri əldə etmək üçün istifadə olunur. Bu məqalədə operativliyi və ya dəqiqliyi azaltmadan klassifikasiya xarakteristikalarını yüksəltmək məqsədilə iki mərhələli ardıcıl klassifikator təklif edilir.

II. ƏLAQƏDAR TƏDQİQATLARIN ANALİZİ

Trafikin klassifikasiyası İnternetin miqyasının genişlənməsi, əldə oluna bilən informasiyanın artması və şəbəkədə baş verən təhdidlərin sayının artması ilə son zamanlar çox diqqət çəkir. Trafikin klassifikasiyasının məqsədindən asılı olaraq, ayrı-ayrılıqda paketləri və ya bütov axını analiz etməklə bu trafikin generasiya olunma mənbəyini və xüsusiyyətini müəyyən etmək olar. Bu xüsusiyyətlər şəbəkənin təhlükəsiz idarə olunmasında marağı olanlar üçün mühümdür. Əslində, trafikin klassifikasiyası böyük bir sistemin əsas blokudur. Bu blok şəbəkədə anomal aktivliyi müəyyən edərək, onun risk səviyyəsini qiymətləndirir. Əvvəllər, İnternetlə bağlı olan proqram təminatları şəbəkənin nəqliyyat səviyyəsindən istifadə edirdilər, bu isə onları asanlıqla identifikasiya etməyə imkan verirdi. Son bir neçə ildə, qeyri-standart portlardan istifadə edən proqram təminatlarından (skype, bittorrent, VPN və s.) çox geniş istifadə olunmağa başlanılmışdır. Bundan əlavə, bir çox proqram təminatı öz varlığını gizlətmək üçün standart portlardan istifadə edir. Bu səbəbdən, dolğun klassifikasiyanın tətbiq olunması üçün paket yoxlama, statistika, maşın təlimi və davranış metodları standart vasitələrə çevrilmişdir.

Bu məqalənin əsas məqsədi şəbəkə trafikinin real vaxt rejimində klassifikasiyası prosesini aparmaqla anomal aktivliyin mənsubiyyətini ilkin mərhələdə aşkarlamaqdır.

Bu məqsədə nail olmaq üçün mövcud işlərdə aşağıdakı məsələlərə baxılmışdır:

1. Real vaxt rejimində klassifikasiya metodlarının tətbiqi.
2. Müasir şəbəkə arxitekturalarına uyğun olaraq trafikin klassifikasiyasının metod və modellərinin tədqiqi.
3. Normal trafik profilinin müəyyən olunması.
4. Anomal trafikin qiymətləndirilməsi metodlarının işlənməsi.
5. Bu prosesin kompleks formada realizasiyasına nəzarət.

Trafikin klassifikasiyası sahəsində istifadə olunan metodlar trafikin klassifikasiyasının qarşısında duran tələbləri ödəmədiyi kimi, məsələnin praktiki realizasiyası zamanı da problemlər yaradır. Bu sahədə müəyyən olunmuş metodların birbaşa müqayisəsi də dörd səbəbə görə çətin prosesə çevrilir: Birinci, hər bir korparativ şəbəkənin bu mühitə uyğun olaraq müxtəlif trafik axını olduğu üçün, ümumiləşdirilmiş normal trafik profilinin müəyyən olunması mümkün deyil. İkinci, yanaşmalardan müxtəlif funksiyaları əsas götürərək müxtəlif metodlardan istifadə edir və müxtəlif parametrlərin tənzimlənməsi üçün müxtəlif proqram əlavələrindən istifadə edir. Üçüncü, müəlliflər nəticələrini bəyan etsələr də, istifadə olunan proqram kodunu paylaşmırlar. Dördüncü, İnternetdən istifadə edən proqram təminatlarının sayı və növü gündən-günə artır və bunların bir çoxu yeni trafik generasiya edir. Yuxarıda göstərilən məsələləri həll etmək üçün [2]-də, üç trafikin klassifikasiya yanaşmasının hərtərəfli və ardıcıl qiymətləndirilməsi həyata keçirilmişdir: port-əsaslı, davranış-əsaslı və statistik.

[3]-də IP trafikin analizi sahəsində əsas problemlər izah edilir və trafik generasiya edən vasitələrin aşkarlanması yolları tədqiq olunur. Trafik paket səviyyəsində ayrı-ayrılıqda tədqiq olunur və axın kateqoriyaları əsasında detallaşdırılaraq müxtəlif yanaşmalar vasitəsilə yoxlanılır. Bu prosesin sonunda isə hansı yanaşmanın konkret hansı problemin həlli olması uyğunluğu müəyyən olunur.

Tədqiqatlar göstərir ki, şəbəkə trafikinin klassifikasiyası üçün ən əlverişli yanaşmalardan biri maşın təlimi metodlarından istifadə etməkdir. [4]-də real vaxt rejimində trafikin klassifikasiyası üçün kvazi real vaxtda statistik klassifikasiya sxemi təqdim edilmişdir. Bu proses şəbəkə axınını əhəmiyyətli dərəcədə demultipleksləşdirir və sonra sadələşdirilmiş Bayes alqoritmindən istifadə edərək bütün şəbəkə səviyyələri üzrə trafik axınını klassifikasiya etməyə imkan verir.

İnternet trafikin identifikasiya edilməsi şəbəkənin idarə edilməsi üçün ən vacib vasitədir. Bu operatorlara trafikin matrisini və tələblərini proqnozlaşdırmağa, təhlükəsizlik əməkdaşlarını trafikin anomal davranışını müəyyən etməyə və tədqiqatçılara isə təhlükələrin vaxtında müəyyən olunaraq qarşısının alınması üçün reallığa yaxın model işləmələrinə kömək edir. [5]-də informasiyanın mənbəyi, istiqaməti və port nömrəsi məlum olmadan trafikin yüksək dəqiqliklə klassifikasiya edilməsi qeyd olunmuşdur. Bunun üçün Bayes təlimli neyron şəbəkəyə əsaslanan maşın təlimindən istifadə olunur. Trafik paketlərindən alınmış, bir və ya bir neçə başlıqdan ibarət olan informasiya təlim və yoxlama prosesində istifadə olunur. Klassifikasiya üçün paketin məzmunun oxunmaması digər sistemlərlə müqayisədə emal prosesini daha da sürətləndirir.

Hazırda trafikin dəqiq klassifikasiyası üçün paketlərin dərin analizindən istifadə olunur ki, bu da öz növbəsində normal trafik profilinin çıxarılması üçündür. Bu prosesin realizasiyası istifadəçilərin şəxsi məlumatlarına təhlükə yaradır, güclü prosessor, böyük əməli yaddaş tələb edir. [6]-də paketin birinci dörd baytının oxunaraq daha asan və sadə klassifikasiya metodu təklif olunur.

Proqram təminatlarına uyğun olaraq IP trafikin klassifikasiyası müasir şəbəkə idarəetmə mərkəzinin tərkib hissəsidir. Buna baxmayaraq, OSI modelinin nəqliyyat və ya tətbiqi səviyyələri üzrə trafikin analizi sürətlə effektivliyini itirir. [7]-də təklif olunan axının klassifikasiya mexanizmi, nəzarətdə olan IP trafikin üç əsas xüsusiyyətinə əsaslanır: paketlərin həcmi, intervalı və çatması. Bu xüsusiyyətlərdən istifadə edərək trafikin klassifikasiyası aparılıb, amma bu metodda protokolların işarələnərək izlənməsi prosesi aparılır. Bu metod inkişaf mərhələsində olmasına baxmayaraq alınan nəticələr ümüdvericidir.

Şəbəkədə informasiya axınını öyrənmək üçün məlumat az olduqda, trafikin klassifikasiyasının səmərəliliyini artırmaq üçün [8]-də yeni klassifikasiya sxemi təklif olunmuşdur. Təqdim olunmuş sxemdə, nəqliyyat axınları diskretləşdirilmiş statistik funksiyalar vasitəsilə təsvir olunur və axın korrelyasiya haqqında informasiya isə bag-of-flow (BoF) kimi modelləşdirilir. BoF əsaslı trafikin klassifikasiyasına kombinə olunmuş klassifikator və səmərəliliyin nəzəri təhlili çərçivəsində baxılır. Bundan başqa, yeni BoF trafikin klassifikasiya metodu əsasında Naive Bayes-in korrelyasiya olunmuş axınların proqnozlarını ümumiləşdirmək təklif olunur. Eksprimentlərin nəticələrinə görə, təklif olunan metod mövcud olan ən yaxşı metodlardan daha yüksək klassifikasiya səmərəlilik göstəricilərinə malikdir.

[9]-da skype trafikin digərlərindən seçilməsində iki maşın təlimi klassifikasiya metodları müqayisə olunur ki, bunlar Naive Bayes və neyron şəbəkələridir. Bunun nəticəsində hansı metodun daha effektiv olunması müəyyən olunur. Modellərin yoxlanılması üçün şəbəkə alətlərindən (NETSTAT və Tcpdump) istifadə olunur, şəbəkə axınında bütün paketlər ələ keçirilib işarələnir. Bu yolla trafikin axınının statistik xarakteristikaları çıxarılır. Bu funksiyalar klassifikatorların öyrədilməsi və yoxlanılması ilə yanaşı yuxarıda qeyd etdiyimiz işarələmələr isə qiymətləndirmələrdə "əsas həqiqət" kimi qəbul olunur. Eksprimentlərdən sonra müəyyən olunub ki, Naive Bayes dəqiqlikdə neyron şəbəkələrdən aşağı olsa da, hesablama sürəti nisbətən yüksəkdir.

Maşın təliminin bir neçə metodu arasında trafikin klassifikasiya dəqiqliyinin yoxlanılması üçün [10]-da müqayisəli təhlil aparılıb. Bu təhlildə IP-trafikin klassifikasiyası üçün C4.5, Naive Bayes, ən yaxın qonşular, radial bazis funksiyaları (RBF) metodlarından istifadə edilib. Aparılan təhlil onu göstərir ki, C4.5 metodundan istifadə edərkən, digər metodlarla müqayisədə qərarların ağacı 93,33% dəqiqlik verir. Bu metod klassifikasiya xətalarını azaldaraq daha yaxşı nəticələr əldə etməyə imkan verir.

III. TƏDQIQAT OBYEKLƏRİ

Klassifikasiyanın tətbiqi üçün obyeklərin parametrlərinin müəyyənləşdirilməsi tələb olunur. Klassifikator bu parametrlərə uyğun olaraq hər bir obyektə uyğun sinif seçir.

Bizim yanaşmada klassifikasiya obyekti trafik axınıdır. Bu trafik axını müəyyən qovşaqlar arasında ötürülən və qəbul olunan bir və ya bir neçə paketdən ibarətdir. Paket konteyner formasında müəyyən olunur və IP-ünvanlar, qovşaqlar haqqında məlumat, protokollar (məsələn - ICMP, TCP və ya UDP), UDP və TCP olduğu halda portların nömrələrindən ibarətdir. TCP qoşulmaları zamanı axın müəyyən uzunluğa və xüsusi semantikaya malikdir. Bu məqalədə klassifikasiya prosesini yüngülləşdirmək üçün yalnız TCP qoşulmaları nəticəsində yaranan trafik axını tədqiq olunur. Şəkil 1-də təsvir olunan hər bir obyek klassifikasiya üçün verilənlər kimi istifadə olunur. Bu obyektlər dupleks trafikin hər iki istiqaməti üçün nəzərdə tutulub

Obyektlər
Axın müddəti
TCP Port
Paketlərin çatma vaxtı
Faydalı trafikin həcmi
Entropiya əsasında Bandwidth səmərəliliyi
Furye çevirməsilə paketlərin çatma vaxtı

Şək 1. Klassifikasiya obyektləri

Klassifikasiya	Nümunələr
HƏCM	ftp
VERİLƏNLƏR BAZASI	postgres, sqlnet oracle, ingres
İNTERAKTİV	ssh, klogin, rlogin, telnet
MAIL	imap, pop2/3, smtp
SERVIS	X11, dns, ident, ldap, ntp
WWW	www
P2P	KaZaA, BitTorrent, GnuTella
HÜCUMLAR	Internet soxulcanlar və virus hücumları
OYUNLAR	Half-Life
MEDİA	Windows Media Player, Real

Şək 2. Şəbəkə trafikinin sinifləri.

Hər axının bir sıra unikal xüsusiyyətləri və davranış xarakteristikasının parametrləri var. Bu məlumatlar klassifikasiya üçün giriş diskriminatorunu təşkil edir.

Klassifikasiya prosesində əsas yanaşma trafikin siniflərə bölünməsi ideyasıdır (şəkil 2.). Qeyd etmək lazımdır ki, hər bir axın bir sinifə aid edilsə də, siniflərin xüsusiyyətləri unikal deyildir. Misal üçün, HƏCM sinifinə aid olan FTP trafiki idarə olunma və informasiya mübadiləsi axınlarından ibarət olsa da bir sinifə aid olunub.

IV. KLASSİFİKATORLAR

Hər bir obyekt $A=\{a_1, a_2, \dots, a_n\}$ atribut qiyməti ilə təsvir olunur. Təlim toplusunda obyektlərin hansı siniflərə mənsub olması məlumdur: (A_1, C_1), (A_2, C_2), ..., (A_m, C_m), burada C_1, C_2, ..., $C_m \epsilon C$ siniflərin nişanlarıdır. Klassifikasiya məsələsi – atributları ilə verilirmiş obyektin (A) hansı sinfə aid olmasını (C nişanını) müəyyən etməkdən ibarətdir.

Naive Bayes klassifikatoru sadə klassifikasiya sxemidir. Bayes klassifikatoru aposterior ehtimalın maksimumluğu prinsipinə əsaslanır. Klassifikasiya olunan obyekt üçün hər bir sinifin həqiqətə oxşarlıq funksiyası hesablanır və onların əsasında siniflərin aposterior ehtimalları tapılır. Obyekt aposterior ehtimalı maksimal olan sinfə aid edilir:

$$h(A)= \arg \max_{c \epsilon C} p(c|A) \quad (1)$$

Bayes teoremindən istifadə etməklə yuxarıdakı qərar funksiyasını çevirmək olar

$$H(A)= \arg \max_{c \epsilon C} p(c|A) = arg \max_{c \epsilon C} \frac{p(A|c)p(c)}{p(A)} = arg \max_{c \epsilon C} p(A|c)p(c). \quad (2)$$

burada P(A) sabitdir.

Naive Bayes metodunda atributların statistik asılı olmaması fərz olunur. Ona görə p(A|C)-ni hesablamaq sadələşir.

$$p(A|c) = \prod_{i=1}^{n} p(a_i|c)p(c) \quad (3)$$

Naive Bayes klassifikatorunun qərar funksiyası

$$h(A) = arg \max_{c \epsilon C} \prod_{i=1}^{n} p(a_i|c)p(c) \quad (4)$$

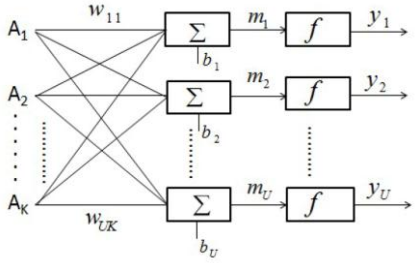
şəklinə düşür. Naive Bayes klassifikasiyasını aparmaq üçün təlim toplusundan istifadə etməklə p(ai|c) və p(c) ehtimal paylanmalarını qiymətləndirmək lazımdır.

Naïve Bayes klassifikatoru digər klassifikasiya metod-larından fərqli olaraq çox sadədir, çünki təlim məlumatlarını yalnız bir dəfə keçmək yetərlidir və sadə əlaqələr olan yerlərdə yüksək nəticələr verir.

Neyron şəbəkə insanın sinir sistemini təqlid edən yüksək qeyri-xətti mürəkkəb sistemdir. Əslində, neyron şəbəkə çoxlu sayda neyronların müəyyən strukturda birləşməsidir və real fiziki dünyanın müxtəlif hadisələrini modelləşdirməyə xidmət edir. Hazırda, neyron şəbəkələr tanıma, klassifikasiya, identifikasiya aləti kimi geniş istifadə olunur. Neyron şəbəkəsinin yaradılması və istifadəsi üçün lazım olan tipik proseduralar aşağıdakılardır:

1. Məlumatların və çıxış məlumatlarının toplanması, onların təlim verilənlərinə bölünməsi və bütün bu məlumatların dolğunluğunun yoxlanması;
2. Neyron şəbəkənin arxitekturunun qurulması;
3. Neyron şəbəkənin təlim verilənlərindən istifadə edərək öyrədilməsi;

Bu məqalədə trafikin klassifikasiyası üçün feed-forward neyron şəbəkəsi istifadə olunmuşdur. Feed-forward neyron şəbəkə strukturu giriş, gizli və çıxış laylarından ibarətdir. Şəkil 3-də neyron şəbəkənin strukturu göstərilib. Bir qayda olaraq, f funksiyası istifadəçi tərəfindən, digər parametrlər isə təlim tərəfindən müəyyən olunur.



Şək 3. Neyron şəbəkənin strukturu

Şəkil 3-də $A_i (1 \leq j \leq K)$ bu layın girişidir;

$w_{ij}(1 \le i \le U; 1 \le j \le K)$ müvafiq neyronlar arasında çəki; $b_i(1 \le i \le U)$ bu yerdəyişmədə i neyronuna əlavə olunur ; f bu layın aktivləşmə funksiyasıdır; $y_i(1 \le i \le U)$ müvafiq neyronun çıxışıdır; $m_i = w_{i1}A_1 + w_{i2}A_2 + \cdots + w_{ik}A_k + b_i; (1 \le i \le U)$; $y_i = f(m_i)$. Şəkil 3-də göstərildiyi kimi, müxtəlif layları birlikdə feed-forward neyron şəbəkəsi ilə birləşdirdikdə əvvəlki layın çıxışı növbəti layın girişi olur.

İlkin konfiqurasiyadan sonra, neyron şəbəkə modeli w çəki matrisinin qiyməti və b yerdəyişməsi dəyərlərinin tənzimlənməsi yolu ilə yenilənir.

V. NƏTİCƏ

Bu məqalədə iki mərhələli klassifikatorun kompüter şəbəkələrinin informasiya təhlükəsizliyinin monitorinqi sistemlərində təhdidlərin operativ aşkarlanması üçün tətbiqinin mümkünlüyü analiz edilmişdir. Naive Bayes və feed-forward neyron şəbəkəsi əsasında tətbiq olunan multi klassifikatorun şəbəkə təhlükəsizliyi insidentlərinin müəyyən olunması üçün əsas mərhələləri nəzərdən keçirilmişdir. İnformasiya təhlükəsizliyi hadisələrinin operativ aşkarlanması üçün yanaşma təklif edilmişdir.

ƏDƏBİYYAT

[1] Amanda A. Surviving Security: How to Integrate People, Process, and Technology, Auerbach Publications, 2003.

[2] Kim H., Claffy K., Fomenkova M., Browlee N., Barman D., Faloutsos M. Comparison of Internet Traffic Classification Tools / Internet Measurement Research Group Workshop on Application Classification and Identification, 2007, pp. 11.

[3] Callado, A. Kamienski, C. Szabo, G. Gero, B. Kelner, J. Fernandes, S. Sadok, D. A Survey on Internet Traffic Identification // IEEE Communications Surveys & Tutorials, 2009, vol.11, no.3, pp. 37 – 52.

[4] Wei Li, Kaysar A., Robert D., Andrew M., Approaching Real-time Network Traffic Classification, Technical Report, 2006.

[5] Auld T., Andrew M., Gull S.F., Bayesian Neural Networks for Internet Traffic Classification // IEEE Transactions on Neural Networks, 2007, vol.18, no.1, pp. 223 – 239.

[6] Shane A., Richard N., Libprotoident: Traffic Classification Using Lightweight Packet Inspection, Technical Report, 2012.

[7] Manuel C., Maurizio D., Francesco G., Luca S., Traffic Classification through Simple Statistical Fingerprinting // Association for Computing Machinery's Special Interest Group on Data Communications Computer Communication Review, 2007, vol.37, no.1, pp. 5-16.

[8] Jun Z., Chao C., Yang X., Wanlei Z., Yong X., Internet Traffic Classification by Aggregating Correlated Naive Bayes Predictions// , IEEE Transactions on Information Forensics and Security, 2012, vol.8, no.1, pp. 5-15.

[9] Mohammad J., Skype Traffic Classification: Naive Bayes or Neural Networks, Report, University of Toronto, 2010

[10] Jamuna A, Vinodh Ewards S.E., Efficient Flow based Network Traffic Classification using Machine Learning // International Journal of Engineering Research and Applications, 2013, vol.3, no.2, pp.1324-1328.