ELEKTRON XAT TIZIMIDA MA’LUMOTLARNI TASNIFLASH VA INTELLEKTUAL TAHLIL

Raximov Bobomurod Ramazonovich
TATU Mustaqil izlanuvchisi Tel:+998997601554

Annotatsiya: *Ushbu ilmiy maqolada elektron xat (e-mail) tizimlarida katta hajmdagi ma’lumotlarni avtomatik tasniflash va ularni intellektual tahlil qilishning zamonaviy yondashuvlari ko‘rib chiqiladi. Maqolada matnni oldindan qayta ishlash, xususiyatlarni ajratib olish, mashinaviy o‘rganish va chuqur o‘rganish modellarining samaradorligi tahlil qilinadi. Shuningdek, spamni aniqlash, mavzuga ko‘ra guruhlash, foydalanuvchi xatti-harakatlarini prognozlash kabi amaliy masalalar yoritiladi.*

Kalit so‘zlar: *elektron xat, ma’lumotlarni tasniflash, intellektual tahlil, mashinaviy o‘rganish, NLP.*

Аннотация: *В данной научной статье рассматриваются современные подходы к автоматической классификации и интеллектуальному анализу больших объемов данных в системах электронной почты. В статье анализируется эффективность предварительной обработки текста, извлечения признаков, моделей машинного обучения и глубокого обучения. Также рассматриваются практические вопросы, такие как обнаружение спама, группировка тем и прогнозирование поведения пользователей.*

Ключевые слова: *электронная почта, классификация данных, интеллектуальный анализ, машинное обучение, НЛП (обработка естественного языка).*

Abstract: *This scientific article reviews modern approaches to automatic classification and intelligent analysis of large amounts of data in electronic mail (e-mail) systems. The article analyzes the effectiveness of text preprocessing, feature extraction, machine learning and deep learning models. It also covers practical issues such as spam detection, topic grouping, and user behavior prediction.*

Keywords: *email, data classification, intelligent analysis, machine learning, NL*P.

KIRISH

Raqamli transformatsiya jarayonlari jadallashgan hozirgi davrda elektron xat (e-mail) tizimlari axborot almashishning eng muhim va keng tarqalgan vositalaridan biri bo‘lib qolmoqda. Davlat idoralari, ta’lim muassasalari, biznes subyektlari hamda jismoniy shaxslar kundalik faoliyatida elektron xatlardan faol foydalanmoqda. Natijada elektron pochta tizimlarida katta hajmdagi turli xil, asosan strukturalanmagan ma’lumotlar shakllanmoqda. Ushbu ma’lumotlarni samarali boshqarish, tezkor qayta ishlash va mazmunan tahlil qilish zamonaviy axborot jamiyati uchun dolzarb muammolardan biri hisoblanadi [1].

Elektron xat tizimlarida axborot oqimining ortib borishi bilan birga spam, fishing, zararli havolalar va ijtimoiy muhandislikka asoslangan xabarlar soni ham keskin ko‘paymoqda. Bu holat nafaqat foydalanuvchilarning vaqt resurslarini samarasiz sarflashiga, balki axborot xavfsizligiga jiddiy tahdid solmoqda. Shu sababli elektron xatlarni avtomatik tarzda tasniflash, muhim va ishonchli xabarlarni aniqlash, shuningdek, xavfli kontentni filtrlash zarurati yuzaga kelmoqda. An’anaviy qo‘lda saralash usullari katta hajmdagi ma’lumotlar bilan ishlashda yetarli samaradorlikni ta’minlay olmaydi.

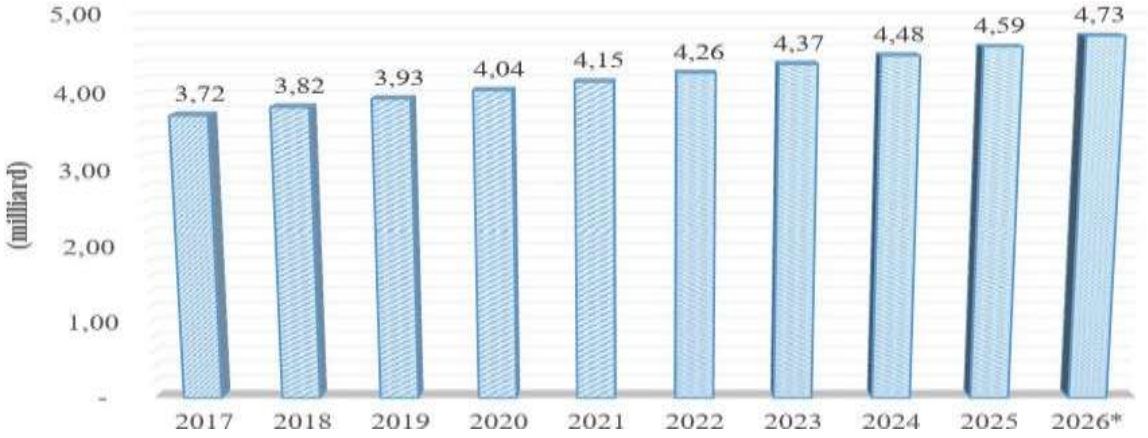
Mazkur muammoni hal etishda sun’iy intellekt va mashinaviy o‘rganish texnologiyalariga asoslangan intellektual tahlil usullari muhim ahamiyat kasb etadi. Xususan, tabiiy tilni qayta ishlash (Natural Language Processing – NLP) metodlari yordamida elektron xatlarning matnli tarkibini tahlil qilish, semantik mazmunini aniqlash va ularni oldindan belgilangan toifalarga ajratish imkoniyati mavjud. Bunday yondashuvlar xatlarni “muhim”, “oddiy”, “spam”, “reklama” yoki “xavfli” kabi sinflarga ajratishda yuqori aniqlikni ta’minlaydi [2].

Elektron xatlarni intellektual tahlil qilish nafaqat xavfsizlikni oshirish, balki foydalanuvchi tajribasini yaxshilashga ham xizmat qiladi. Masalan, foydalanuvchining oldingi xatti-harakatlari asosida xatlarni ustuvorlik bo‘yicha tartiblash, avtomatik javob tavsiyalari berish yoki ma’lum mavzudagi xabarlarni guruhlash imkoniyatlari yaratiladi. Bu esa ish samaradorligini oshirib, axborot bilan ishlash jarayonini optimallashtiradi.

Shu nuqtai nazardan, elektron xat tizimlarida ma’lumotlarni tasniflash va intellektual tahlil qilish masalalarini ilmiy va amaliy jihatdan o‘rganish muhim ahamiyatga ega. Ushbu maqolada elektron xatlarni avtomatik tasniflashning nazariy asoslari, zamonaviy intellektual algoritmlar va ularning amaliy qo‘llanilishi tahlil qilinadi.

TAHLIL VA NATIJALAR

Bugungi kunda xabar almashish platformalari va ilovalarining yaratilishi hamda rivojlanishiga qaramay, elektron pochta tizimi aloqa infratuzilmasi uchun ma’lumotni uzatish va konfidensial axborotni almashishning muhim vositalaridan biri bo‘lib qolmoqda. Internet servislari ko‘lamini ortishi bilan birgalikda, dunyo bo‘ylab elektron pochta foydalanuvchilarining soni oshib bormoqda [1] (1.1- rasm).



1.1 – rasm. Butun dunyo bo‘ylab elektron pochta foydalanuvchilari soni (2017-2026 yillar)

Biroq, uning keng tarqalganligi bilan bir qatorda, xavfsizlik tahdidlari yanada keskinlashib tobora murakkablashib bormoqda va undan foydalanish tashkilotlar uchun eng katta risklardan biri bo‘lib qolmoqda.

Elektron pochta bugungi kunda faqat muloqot qilish uchun emas, balki onlayn servislarda ro‘yxatdan o‘tish va turli tizimlardan foydalanish huquqini olish uchun qo‘llanilmoqda.

Nomaqbul xabarlar - foydalanuvchilarning ma’lum soniga ularning roziligisiz tijoriy, firibgarlik yoki zararli mazmunda jo‘natiladigan elektron pochta xatlari. Elektron pochta spam xabarlari zamonaviy axborot kommunikatsiya infrastrukturasidagi eng dolzarb muammolardan biri. Nomaqbul xabarlar tarmoqning o‘tkazish qobiliyatini sarflab foydalanuvchilarning pochta qutilarida joy egallaydi va haqiqiy elektron xatlarning kirib kelishiga to‘sqinlik qiladi, shu bilan birgalikda butun tarmoq tizimi resurslarining haddan tashqari yuklanishiga olib keladi, bu esa elektron pochta xizmatlari va aloqa kanallarining uzluksiz ishlashiga salbiy ta’sir ko‘rsatadi. Spam xabarlar elektron pochta tizimiga ta’sir etuvchi quyidagi risklarni yuzaga keltiradi:

– ishchi stansiyalarni zararkunanda dasturlar yordamida zararlanishi (viruslar, qurtlar va troyan dasturlar);

– kiruvchi ko‘p sonli nomaqbul xabarlar oqibatida tarmoqning o‘tkazish qobiliyatining yuklanishi;

– ma’lumotlar va hisoblash resurslari saqlanuvchi makondan noto‘g‘ri foydalanish;

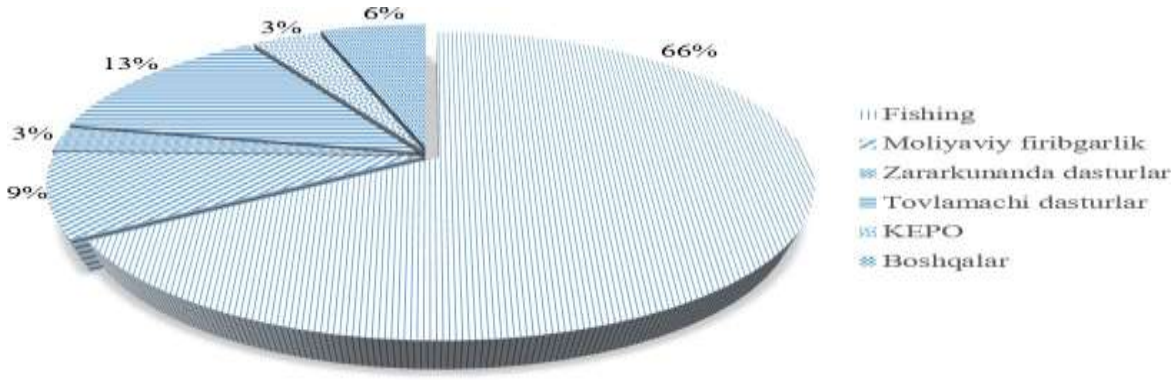
– moliyaviy firibgarliklar (fishing va u bilan bog‘liq boshqa xujumlar natijasida moliyaviy yo‘qotishlar) [2].

Dunyo pochta trafigida nomaqbul elektron xabarlarning o‘rtacha ulushi 1.2-rasmda keltirilgan.

1.2-rasm. Dunyo pochta trafigida nomaqbul elektron xabarlarning o‘rtacha ulushi (2011-2023 yillar)

Butun dunyona elektron pochta tizimlari xavfsizligidagi asosiy muommolardan biri hisoblanib, zararli dasturlar tarqalishining asosiy yo‘nalishi biri bo‘lib qolmoqda. Spam-xabarlarning har xil turlari mavjud, ular foydalanuvchi qurilmalarini zararlash va

ishdan chiqarishga qaratiladi. Ushbu turdagi xabarlardan ko‘plab hujumlarni amalga oshirishda foydalaniladi. Ular orasida, fishing xujumi niyati buzuqlar tomonidan qo‘llaniladigan xavfli usul bo‘lib, barcha xujumlarning uchdan ikki qismini tashkil qilmoqda. Keng tarqalgan nomaqbul xabarlarning turlari bo‘yicha statistikasi 1.3-rasmda keltirilgan.



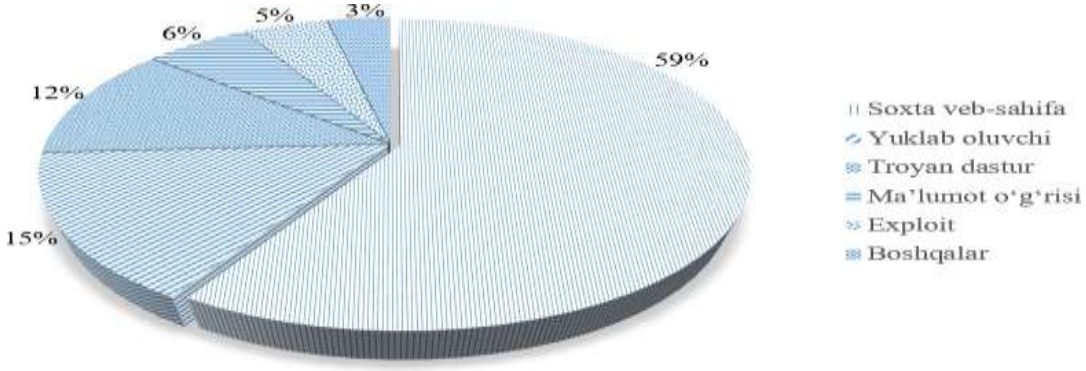
1.3.-rasm. Nomaqbul xabarlarning turlari bo‘yicha statistikasi

Fishing xabarlar

Ijtimoiy muhandislik usullaridan foydalangan holda amalga oshiriladigan xujumlarning eng keng tarqalgan shakllaridan biri bo‘lib, kiberxavfsizlik sohasidagi ko‘plab insidentlar fishing xabarlaridan boshlanadi. Fishing xujumidan asosiy maqsad foydalanuvchining muhim (tizim qayd yozuvlari va shaxsiy parollar, moliyaviy hisob raqamlar) ma‘lumotlarini yig‘adigan soxta va zararali veb- sahifaga yo‘naltirish, kiritilgan ma‘lumotlarni niyati buzuq tomonidan belgilangan manzilga yuborishdan iborat. Elektron pochta fishing xujumning ko‘plab turlari mavjud, ular quyidagi qadamlar orqali amalga oshiriladi:

- tarkibida havola yoki yuklab oluvchi bo‘lgan firibgarlik elektron xati yaratiladi;
- tegishli vosita yordamida jabrlanuvchiga xat jo‘natiladi;
- havola foydalanuvchini firibgarlik veb- sahifasiga yoki zaruriy dasturni yuklashga yo‘naltiradi;
- veb-sahifa foydalanuvchidan maxfiy va qayd yozuvi ma’lumotlarini kiritishni yoki dasturni yuklab o‘rnatishni so‘raydi. Ushbu ma’lumotlar niyati buzuq tomonidan yig‘iladi va shu bilan keyingi xujumlarda foydalaniladi [3].

Fishing xabarlarining turlari bo‘yicha statistikasi 1.4-rasmda keltirilgan.

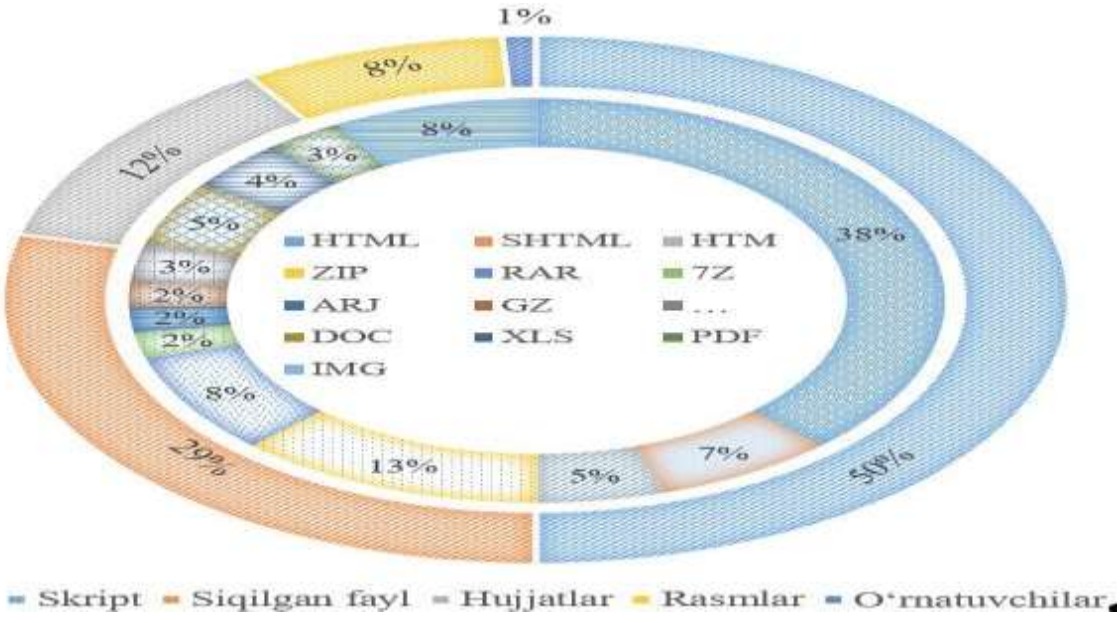


1.4 – rasm. Fishing xabarlarni turlari bo‘yicha statistikasi

Zararli qo‘yilmalar

Bugungi kunda elektron pochtaning zararli qo‘yilmalari zararkunanda dasturlarni tarqatishning eng keng tarqalgan usullaridan biri bo‘lib foydalanuvchi kompyuteriga hujum qilish uchun mo‘ljallangan. Ushbu elektron xatlar, tarkibidagi zararli qo‘yilmalar, xujjatlar elektron fayllar tusida niqoblangan bo‘lishi mumkin. Deyarli har qanday turdagi fayl elektron pochtaga biriktirilishi mumkin, shuning uchun kiberjinoyatchilar ma’lumotlarni yo‘qotish yoki o‘g‘irlash imkoniga ega zararli dasturlarni o‘rnata oluvchi fayllarni ushbu elektron xatlarga biriktiradilar.

Elektron pochta tarkibida ko‘rish yoki chop etish uchun yuklab olishni taklif qiluvchi fayl mavjud bo‘ladi. Foydalanuvchi faylni ochganida biriktirilgan skriptlar, makroslar yoki tizim zaifliklari avtomatik ravishda ishga tushadi hamda kompyuterga troyan yoki viruslarni masofadan yuklab oladi va ishga tushiradi. Zararkunanda dasturlar turli xil va murakkab ko‘rinishda bo‘lib, ular ma’lumotlarga zarar yetkazishi (o‘chirib yuborishi, shifrlashi), masofadan ulanishga ruxsat berishi, ishchi stansiyani to‘liq yoki qisman ishdan chiqarishi hamda blokirovkalashi mumkin [4]. Zararli qo‘yilmalarning keng tarqalgan turlari 1.5- rasmda ko‘rsatilgan.



1.3 – rasm. Zararli qo‘yilmalarning turlari b‘yicha statistikasi

Korporativ elektron pochtani obro‘sizlantirish (BEC- business email compromise) - maqsadli kiberhujum turi bo‘lib, kiberjinoyatchi o‘zini ishonchli shaxs (kompaniyaning vakolatli vakili) sifatida ko‘rsatadi va maxfiy ma‘lumotlarni taqdim etish uchun elektron pochta xabarlarini muayyan shaxslar (kompaniya hodimlari yoki hamkorlari)ga yuboradi. Mazkur turdagi hujum quyidagi qadamlar orqali amalga oshiriladi:

– kompaniyaning faoliyati va boshqaruv strukturasi xususida ma’lumot to‘plash;

– kompaniyaning vakolatli vakilining elektron pochta qayd yozuvi

ma’lumotlarini olish uchun maqsadli hujum qilish;

– elektron pochta qayd yozuvidan vakolatli hodimlarga ishonchli shaxs nomidan xabar yuborish [5].

Xulosa

Xulosa qilib aytganda, elektron xat tizimlarida ma’lumotlarni tasniflash va intellektual tahlil qilish zamonaviy axborot jamiyati sharoitida muhim ilmiy-amaliy ahamiyat kasb etadi. Elektron pochta orqali uzatiladigan axborot hajmining ortib borishi, ularning mazmunan xilma-xilligi va xavfsizlikka bo‘lgan talablarning kuchayishi avtomatlashtirilgan va intellektual yondashuvlarni joriy etishni taqozo etmoqda. An’anaviy qo‘lda saralash usullari katta hajmdagi elektron xatlar bilan ishlashda yetarli samaradorlikni ta’minlay olmaydi, bu esa zamonaviy algoritmlarga asoslangan tizimlarga ehtiyojni yanada oshiradi.

Tadqiqot davomida elektron xatlarni tasniflashda sun’iy intellekt, mashinaviy o‘rganish va tabiiy tilni qayta ishlash texnologiyalarining imkoniyatlari tahlil qilindi. Ushbu yondashuvlar yordamida xatlarning semantik mazmunini aniqlash, spam va fishing xabarlarni yuqori aniqlik bilan aniqlash hamda foydalanuvchi ehtiyojlariga mos ravishda axborotni guruhlash mumkinligi aniqlandi. Ayniqsa, matnli ma’lumotlar bilan ishlashda statistik va neyron tarmoqlarga asoslangan modellarning samaradorligi yuqori natijalar berayotgani kuzatildi.

FOYDALANILGAN ADABIYOTLAR:

1. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. — 4th ed. — Pearson Education, 2021.

2. Aggarwal C. C. Machine Learning for Text. — Springer, 2018.

3. Ганиев С. К., Хамидов Ш. Ж. Безопасность систем электронной почты: проблемы и решения. Электронный сборник научных статей по материалам XXI Международной научно- методической конференции “Информатика: проблемы, методы, технологии”, Воронеж, 11-12 февраля 2021. - 690-699 с.

4. Muralidharan T., Nissim N. Improving malicious email detection through novel designated deep-learning architectures utilizing entire email. Neural Networks Vol. 157, January 2023, pp. 257-279.

5. Papathanasiou, A., Liontos, G., Liagkou, V., Glavas, E. Business Email Compromise (BEC) Attacks: Threats, Vulnerabilities and Countermeasures. A Perspective on the Greek Landscape. J. Cybersecur. Priv., Vol. 3, Iss. 3. September 2023, 3, pp 610–637.

6. Raja S.H., Afrooz F. New method for assessing risks of email. International Conference on Graphic and Image Processing (ICGIP 2012), Proceedings of the SPIE. Vol. 8768, pp 6.