ISSN 2786-7226
DOI: 10.53920/ITS

Scientific journal

ITSYNERGY

2022 Issue 2 (3)

Науковий журнал

ITSYNERGY

2022 Випуск 2 (3)

Kyiv, 2022

SCIENTIFIC JOURNAL
IT SYNERGY

Published since 2021 year *Two time a year*

ISSN 2786-7226 Kyiv, 2022, Issue 2 (3)

Establishers: Academician Yuriy Bugay International Scientific and Technical University

CrossRef: http://doi.org/10.53920/ITS

The journal publishes the results of scientific research in the following specialties:
121 – Software engineering; 122 – Computer science; 125 – Cyber security;
172 – Telecommunications and radio engineering.

Recommended for publication by the decision of the Academician Yuriy Bugay International Scientific and Technical University (Ukraine), protocol № 04/2223 from 27.12.2022

Editorial board address: Scientific journal «IT SYNERGY», Academician Yuriy Bugay International Scientific and Technical University, provulok Magnitogorskyy, 3, Kyiv, 02094, Ukraine

☎: (095) 945-77-80 **e-mail:** journal@istu.edu.ua **web:** http://its.istu.edu.ua

Registered by the Ministry of Justice of Ukraine Certificate of state registration of the print media Series KB № 24967-14907P dated 20.09.2021

НАУКОВИЙ ЖУРНАЛ
IT SYNERGY

Засновано у червні 2021 року *Виходить 2 рази на рік*

ISSN 2786-7226 Київ, 2022, Випуск 2 (3)

Засновник: Заклад вищої освіти «Міжнародний науково-технічний університет імені академіка Юрія Бугая»

CrossRef: http://doi.org/10.53920/ITS

У журналі публікуються результати наукових пошуків зі спеціальностей:
121 – Інженерія програмного забезпечення; 122 – Комп'ютерні науки;
125 – Кібербезпека; 172 – Телекомунікації та радіотехніка.

Рекомендовано до друку рішенням Вченої ради ЗВО «МНТУ»,
протокол № 04/2223 from 27.12.2022

Адреса редакції: Науковий журнал «IT SYNERGY», ЗВО «МНТУ»,
провулок Магнітогорський, 3, м. Київ, 02094, Україна

☎: (095) 945-77-80 **e-mail:** journal@istu.edu.ua **web:** http://its.istu.edu.ua

Зареєстровано Міністерством юстиції України
Свідоцтво про державну реєстрацію друкованого засобу масової інформації
Серія КВ № 24967-14907Р від 20 вересня 2021 року

ЗМІСТ

Валерій Вікторович КОВАЛЬ, Олександр Всеволодович САМКОВ, Олександр Леонідович ОСІНСЬКИЙ, Богдан Олександрович САМКОВ

ПІДВИЩЕННЯ ДОСТОВІРНОСТІ СИНХРОІНФОРМАЦІЇ СИНЕРГЕТИЧНИХ МЕРЕЖ SMART ТЕХНОЛОГІЙ 6

Олександр Іванович ГОЛУБЕНКО, Олександр Олександрович ПІДМОГИЛЬНИЙ

GENERATIVE PRE-TRAINED TRANSFORMER 3 19

Андрій Вікторович ЛЕМЕШКО, Єлизавета Олександрівна НОВІЧЕНКО, Андрій Володимирович НЕДАВНІЙ

БЕЗПЕКА ДАНИХ В УКРАЇНІ ЗА ДОПОМОГОЮ ВИКОРИСТАННЯ ТЕХНОЛОГІЇ VPN 28

Ольга Миколаївна ТКАЧЕНКО, Владислав Олексійович СОСНОВИЙ

МОДЕЛЬ ПРОГНОЗУВАННЯ БЕЗПЕКИ МЕРЕЖІ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ 43

Сергій Станіславович КОРОТКОВ, Владислав Сергійович ЗАСАДЮК

АЛГОРИТМ ЗМЕНШЕННЯ ТРАНЗИТНИХ ПОТОКІВ ТРАНСПОРТНОЇ МЕРЕЖІ У ЗАДАНОМУ НАПРЯМКУ 55

CONTENTS

Valerii KOVAL, Oleksandr SAMKOV, Oleksandr OSINSKIY, Bogdan SAMKOV

INCREASING THE RELIABILITY OF SYNCHROINFORMATION OF SYNERGETIC NETWORKS OF SMART TECHNOLOGIES 7

Oleksandr GOLUBENKO, Oleksandr PIDMOGYLNYI

GENERATIVE PRE-TRAINED TRANSFORMER 3 20

Andrii LEMESHKO, Yelyzaveta NOVICHENKO, Andriy NEDAVNIY

DATA SECURITY IN UKRAINE USING VPN TECHNOLOGY 29

Olha TKACHENKO, Vladyslav SOSNOVYY

NETWORK SECURITY PREDICTION MODEL USED BY NEURAL NETWORKS 44

Serhii KOROTKOV, Vladyslav ZASADIUK

ALGORITHM FOR CHANGING TRANSIT FLOWS IN THE TRANSPORT METER AT A GIVEN DIRECTLY 55

УДК 621.391:621.396.688
DOI: https://doi.org/10.53920/ITS-2022-2-1

Валерій Вікторович КОВАЛЬ,
д-р техн. наук, проф.,
професор Національного університету біоресурсів
і природокористування України
ORCID ID: 0000-0003-0911-2538
Олександр Всеволодович САМКОВ,
д-р техн. наук, с.н.с.,
заступник директора з науково-технічної роботи
Інституту електродинаміки НАН України
ORCID ID: 0000-0003-2790-8564
Олександр Леонідович ОСІНСЬКИЙ,
начальник відділу Національної академії наук України
ORCID ID: 0000-0002-9921-699X
Богдан Олександрович САМКОВ,
аспірант Інституту електродинаміки НАН України
ORCID ID: 0000-0003-0080-1978

ПІДВИЩЕННЯ ДОСТОВІРНОСТІ СИНХРОІНФОРМАЦІЇ СИНЕРГЕТИЧНИХ МЕРЕЖ SMART ТЕХНОЛОГІЙ

Предметом дослідження є структура інтелектуальної системи та результати експериментальних досліджень її складових частин: супутникові навігаційні системи, обладнання для передачі синхросигналів з використанням PTP-протоколу, пристрій багатоканального моніторингу синхросигналів. Мета – розроблення структури інтелектуальної системи з пристроєм багатоканального моніторингу, яка забезпечить формування з підвищеною достовірністю синхроінформації, що використовується для проведення синхронних векторних вимірів на об'єктах синергетичних мереж SMART технологій.

В статті за результатами досліджень вітчизняних і закордонних фахівців представлено обґрунтування доцільності використання IP-мереж для передавання синхроінформації на основі PTP-протоколу. Експериментально перевірено і підтверджено використання обладнання українського виробництва для передачі по IP-мережам синхросигналів з точністю ± 1 мкс, застосування якого створить умови диверсифікації синхроінформаційного забезпечення. Запропоновано структуру інтелектуальної комп'ютерно-інтегрованої системи, яка за-

безпечить формування з підвищеною достовірністю синхроінформації, що використовується для проведення синхронних векторних вимірів на об'єктах електроенергетичних мереж.

Розроблено пристрій багатоканального моніторингу синхроінформації, який в автоматичному режимі забезпечує одночасний перегляд даних вимірів контрольованих сигналів, їх запис на запам'ятовуючі пристрої та формування інформації для підтримки прийняття рішень з метою підвищення достовірності синхроінформації електроенергетичних мереж SMART технологій. Пропонується подальше проведення наукових досліджень з метою створення інтелектуальної системи, яка забезпечить формування синхросигналів з покращеними показниками якості, а також матиме можливість виконувати безперервний багатоканальний моніторинг параметрів синхроінформації у реальному часі на об'єктах різних галузей економіки країни та може використовуватись в цілях підвищення обороноздатності і безпеки держави.

Ключові слова: *SMART технологія, синергія, векторні вимірювання, інтелектуальна система, синхроінформація, багатоканальний моніторинг.*

Valerii KOVAL
Doctor of Technical Sciences, Professor
Professor National University of Life and Environmental Sciences of Ukraine
ORCID ID: 0000-0003-0911-2538
Oleksandr SAMKOV
Doctor of Technical Sciences
Deputy Director for Scientific and Technical Work Institute of Electrodynamics of the NASU
ORCID ID: 0000-0003-2790-8564
Oleksandr OSINSKIY
Head of Department National Academy of Sciences of Ukraine
ORCID ID: 0000-0002-9921-699X
Bogdan SAMKOV
Postgraduate Institute of Electrodynamics of the NASU
ORCID ID: 0000-0003-0080-1978

INCREASING THE RELIABILITY OF SYNCHROINFORMATION OF SYNERGETIC NETWORKS OF SMART TECHNOLOGIES

The subject of the article is the structure of an intelligent system and the results of experimental studies of its constituent parts: satellite navigation systems, equipment for transmitting synchronization signals using the PTP

protocol, a device for multi-channel monitoring of synchronization signals. The main goal is to develop the structure of an intelligent system with the multi-channel monitoring device, which will ensure the formation of synchroinformation with increased reliability, which is used for conducting synchronous vector measurements at the facilities of synergistic networks of SMART technologies. Based on the research results conducted by domestic and foreign specialists, the article presents the justification of the feasibility of using IP networks for the transmission of synchroinformation based on the PTP protocol. It has been experimentally verified and confirmed that the use of Ukrainian-made equipment for the transmission of synchronization signals with an accuracy of ±1μs over IP networks would create conditions for the diversification of synchronoinformation support. The structure of an intelligent computer-integrated system is proposed. The system will ensure the formation of synchroinformation with increased reliability, which is used for conducting synchronous vector measurements at the facilities of electric power networks. A device for multi-channel monitoring of synchronous information has been developed, which provides the formation of information to support decision-making in order to increase the reliability of synchronoinformation of power grids of SMART technologies. It is proposed to carry out research in order to create an intelligent system that will ensure the formation of synchronization signals with improved indicators, as well as monitor the parameters of synchroinformation in real time at the facilities of various sectors of the country's economy and can be used for the purpose of increasing the state's defense capability.

Keywords: *SMART technology, synergy, vector measurements, intelligent system, synchroinformation, multi-channel monitoring.*

Постановка проблеми. Синергетична мережа є інтегрованою частиною SMART Grid системи, яка з допомогою функцій координації і контролю узгоджує синергії. SMART технології передбачають перетворення в цифрову форму основних вибіркових виміряних значень параметрів електроенергетичних мереж з можливістю їх дистанційного отримання і формування команд управління згідно концепції шини технологічного процесу визначеній стандартом МЕК 61850. Синхронізація процесів у часі, що здійснюється в мережах (електроенергетичній, інфокомунікаційній), є фундаментальною з точки зору їх ефективного та надійного функціонування. Зважаючи на зазначене, вирішення проблеми підвищення достовірності синхроінформації, з використанням

якої забезпечуються процеси часової синхронізації, є актуальним і важливим як в умовах штатного режиму роботи SMART Grid системи, так в умовах надзвичайних ситуацій.

Аналіз останніх досліджень і публікацій. В умовах сьогодення якість і надійність електропостачання споживачів різного призначення відзначається, як їх нормативно-правовим забезпеченням з врахуванням діючих вимог європейського законодавства, так і технічним станом об'єктів електроенергетики та інфокомунікацій. Процес адаптації вітчизняної електроенергетики по стандартам країн Євросоюзу є складною проблемою у зв'язку з її особливістю щодо дотримання всіх параметрів електроенергії на нормованих міжнародними стандартами рівнях. Одним із таких міжнародних нормативних документів є стандарт МЕК 61850, де наведена модель для систем передачі даних з метою вимірювання, моніторингу, автоматизації [1]. На території України технічна політика НЕК «Укренерго» у сфері розвитку та експлуатації магістральних та міждержавних електричних мереж має здійснюватись на нових підстанціях, а також на підстанціях із застарілим обладнанням, згідно стандарту СОУ НЕК 20.261:2021 [2].

Практичне вирішення сформульованої проблеми можливе за рахунок впровадження інтелектуальних засобів керування з використанням синхронізованих векторних вимірювань та застосуванням мульти-агентних комплексів управління, яке передбачає отримання синергетичного ефекту. Згідно визначення Г. Хакенена синергетичний ефект базується на узгодженості взаємодії елементів при утворенні структури як єдиного цілого [3]. У контексті синергетичної мережі SMART технологій це комплекси, системи, сукупність різних елементів, що працюють разом для отримання результатів, яких неможливо отримати жодним із елементів окремо. По суті, мережа являє собою набір взаємопов'язаних компонентів, які працюють разом із спільною метою: задоволення певної визначеної потреби [4].

Проблема якості і надійності електропостачання безпосередньо пов'язана з функціонуванням цифрових систем автоматизованого керування, які побудовані з використанням SMART технологій та виконують свої функції з прив'язкою до часу за рахунок використання синхроінформації [5-9]. В такому разі очевидним є те, що саме від характеристик синхроінформації залежить часова узгодженість взаємодій в системах автоматики та їх складових,

і як наслідок якість роботи синергетичних мереж SMART технологій в умовах штатного режиму і, що особливо важливо, в надзвичайних ситуаціях [8, 9].

Формування синхроінформації забезпечується пристроями синхронізації, а також засобами контролю і моніторингу параметрів синхросигналів, які використовуються для передавання синхроінформації [7, 9, 10]. Варто зауважити, що синхроінформація сформована на основі сигналів, які прийняті від супутникових навігаційних систем, не задовольняє вимогам щодо надійності і достовірності [7, 9, 10]. Разом з цим, формування і розповсюдження синхроінформації на основі засобів ДП «Укрметртестстандарт», які виконують функції по зберіганню національної шкали часу на рівні кращих національних шкал країн світу, не можуть задовольнити вимоги усіх споживачів [11]. Проблема покращення якісних показників синхроінформації в частині її формування, розповсюдження та моніторингу потребує подальшого дослідження і технічного рішення. В першу чергу це стосується споживачів критичної інфраструктури, якими є електроенергетичні мережі SMART технологій.

Мета статті – розроблення структури інтелектуальної комп'ютерно-інтегрованої системи з пристроєм багатоканального моніторингу синхросигналів, яка забезпечить формування з підвищеною достовірністю синхроінформації, що використовується для проведення синхронних векторних вимірів на об'єктах синергетичних мереж SMART технологій.

Виклад основного матеріалу дослідження. Функціонування сучасних генеруючих потужностей, розподільчих мереж, споживачів електроенергетичних систем SMART технологій, в яких використовуються цифрові технології, залежить від якісних показників синхроінформації [1, 2, 5-11]. Синхроінформація використовуються для формування дискретних значень моментів часу в процесах цифрової обробки, передавання, збереження даних. В процесі синхронних векторних вимірів синхрофазорів (PMU - Phasor Measurement Unit), які встановлені на об'єктах електроенергетичних систем SMART технологій, синхроінформації забезпечує вирішення часової невизначеності. Результати вимірів пристроями PMU є вихідною цифровою інформацією для розрахунку стану електроенергетичних систем. Варто зауважити, що рішення сформовані автоматизованою системою керування електроенергетичних мереж SMART технологій залежать від достовірності та надійності отриманих даних векторних вимірів при-

строями PMU. Для отримання максимально достовірного виміру фазових кутів пристроями PMU в обов'язковому порядку потрібно на основі синхроінформаційних сигналів забезпечити «прив'язку» до часу дискретних моментів вимірів з точністю ± 1 мкс [1]. У разі погіршення показників якості синхроінформації, результуючі значення після вимірювання за допомогою синхрофазора можуть бути неточними та вважатимуться недійсними.

Згідно зі стандартом МЕК 61850, електротехнічні засоби електроенергетичних систем SMART-технологій повинні відповідати класам точності, починаючи від Т1 і закінчуючи Т5 (метрологічний сигнал системи інструментальної синхронізації) [1]. Синхроінформаційні сигнали високої точності можуть бути сформовані на основі радіоданих діючих супутникових навігаційних систем GPS, ГЛОНАСС, Galileo. Синхросигнали, що сформовані тільки за супутниковими сигналами, не можуть забезпечити високу інформаційну живучість зважаючи на те, що вони можуть бути спотвореними, як в умовах штатного режиму роботи, так і у надзвичайних ситуаціях [12]. Принциповим недоліком супутникових систем є залежність якості сигналу, що передається, від нестаціонарних характеристик середовища розповсюдження радіосигналів, яке є відкритим стосовно зовнішніх впливів, а також відсутність захисту сигналу від навмисного або випадкового спотворення різними діями [13]. Радіосигнали також можуть не забезпечувати задані показники якості синхроінформації зважаючи на невдале розташування приймальної антени, засмічення антени (листя, гіляки, пташині гнізда, тощо), підключення кабелів з порушеннями встановлених вимог. Навіть постійний вплив ультрафіолету може призвести до порушення роботи радіоприймачів.

Іншим способом формування синхроінформаційних сигналів може бути використання IP-мереж для відтворення національної шкали часу і частоти з прив'язкою до реального часу на основі механізму синхронізації PTP (PTP - Precision Time Protocol), який рекомендовано стандартом МЕК 61850-9 [1]. Варто відзначити, що енергетичні компанії все більше використовують Ethernet/IP і впроваджують мережі з багатопротокольною комутацією по мітках (MPLS) з метою гармонійного переходу до нових технічних рішень, які включають застосування синхрофазорів. Чисельні експериментальні дослідження PTP обладнання, яке використовується для передачі діючими IP-мережам синхроінформаційних сигналів, забезпечує точність ± 1 мкс [11, 14]. Результати випробувань

комплекту обладнання УС-1588 українського виробництва [15] наведено на рис. 1. З графіка залежності TIE (TIE - Time Interval Error) від часу можна зробити висновок про те, що максимальне відхилення часового інтервалу не перевищує однієї мікросекунди.

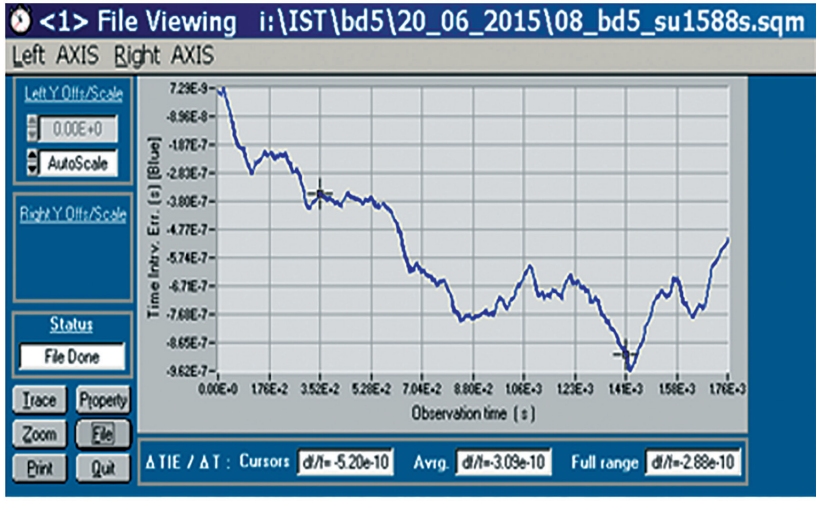

Рис. 1. Результати випробувань комплекту обладнання УС-1588 українського виробництва для випадку передавання синхроінформаційних сигналів діючою корпоративною IP-мережею

Джерело: [15]

Аналіз результатів досліджень [11, 14, 15] протоколу PTP, який отримав назву «енергетичний» профіль PTP, підтверджує можливість його практичного використання в електроенергетичних мережах SMART технологій. Можливість передавання синхросигналів діючими IP-мережам з використанням протоколу PTP створює умови для диверсифікації синхроінформаційного забезпечення в синергетичних мережах SMART технологій.

Вимоги щодо забезпечення заданої точності сигналів синхронізації за умови підвищення їх надійності і достовірності, обумовлюють актуальність створення автоматизованої комп'ютерно-інтегрованої системи формування синхросигналів (рис. 2) з розробкою програмно-апаратного забезпечення, включно

і з обчислювальним інтелектом [15]. Для виявлення і оцінки змін в синхроінформації пропонується ряд організаційно-технічних заходів, які забезпечують можливість проведення безперервного моніторингу (24x7) параметрів синхроінформаційних сигналів з обробкою результатів дистанційних вимірів показників якості у реальному часі (оптимізація, прогнозування, прийняття рішень).

Складовою інтелектуальної системи формування синхросигналів (ІСФС) є пристрій багатоканального моніторингу синхроінформації, який в автоматичному режимі формує цифрові дані про результати одночасних вимірів декількох синхросигналів у реальному часі та, використовуючи засоби IP-мереж, передає їх до/від сервера з метою забезпечення централізованого керування (рис. 2).

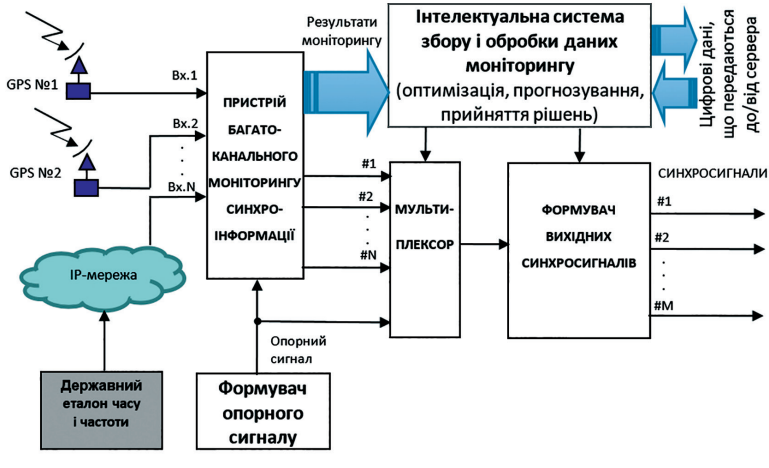

Рис. 2. Структура інтелектуальної системи формування синхросигналів

Процес вимірювань показників якості синхроінформаційних сигналів зводиться до порівняння з одиницею фізичної величини для отримання кількісної інформації, а при контролі фізичний параметр порівнюють з його нормою з метою визначення відхилень даного параметра [7]. Значна увага при зіставленні первинної інформації з заздалегідь встановленими вимогами приділяється формі подання, яка повинна бути легкою, як правило,

до візуального сприйняття, сприятливою для прийняття рішень. З цією метою інформація про стан синхроінформаційних сигналів надходить на екран у вигляді динамічних або статичних графіків, мнемонічних символів. Також формується сигнальна інформація та кількісні показники результатів моніторингу.

Створення інтелектуальної системи формування синхросигналів забезпечить можливість диверсифікації синхроінформаційного забезпечення та одночасного перегляду даних вимірів декількох контрольованих сигналів, їх запис на запам'ятовуючі пристрої та формування інформації для підтримки прийняття рішень з метою підвищення достовірності синхроінформації синергетичних мереж SMART технологій.

Запропонована інтелектуальна система може забезпечувати формування синхросигналів з покращеними показниками якості, а також виконувати безперервний моніторинг параметрів синхроінформації у реальному часі на об'єктах різних галузей економіки країни та використовуватись в цілях підвищення обороноздатності і безпеки держави.

Висновки та пропозиції. За результатами досліджень вітчизняних і закордонних фахівців обґрунтована доцільність використання IP-мереж для передавання синхроінформації на основі PTP-протоколу (IEEE-1588), рекомендованого міжнародним стандартом МЕК 61850.

Експериментально перевірено і підтверджено використання обладнання українського виробництва (УС-1588) для передачі по діючим IP-мережам синхросигналів з точністю ± 1 мкс, застосування якого створить умови диверсифікації синхроінформаційного забезпечення.

Запропонована структура інтелектуальної комп'ютерно-інтегрованої системи, яка забезпечить формування з підвищеною достовірністю синхроінформації, що використовується для проведення синхронних векторних вимірів на об'єктах синергетичних мереж SMART технологій.

Розроблено пристрій багатоканального моніторингу синхроінформації, який в автоматичному режимі забезпечує одночасний перегляд даних вимірів контрольованих сигналів, їх запис на запам'ятовуючі пристрої та формування інформації для підтримки прийняття рішень з метою підвищення достовірності синхроінформації синергетичних мереж SMART технологій.

Пропонується виконання наукових досліджень з метою створення інтелектуальної системи, яка забезпечить формування синхросигналів з покращеними показниками якості, а також матиме можливість виконувати безперервний багатоканальний моніторинг параметрів синхроінформації у реальному часі на об'єктах різних галузей економіки країни та може використовуватись в цілях підвищення обороноздатності і безпеки держави.

ЛІТЕРАТУРА

1. https://webstore.iec.ch/publication/6028.
2. Стандарт підприємства СОУ НЕК 20.261:2021 «Технічна політика НЕК «Укренерго» у сфері розвитку та експлуатації магістральних та міждержавних електричних мереж».
3. Dynamic of synergetic systems / Ed. by H. Haken. B. etc., 1980. 271 p.
4. Blanchard BS (2004). System engineering management (3rd ed.). Hoboken, NJ: John Wiley. p. 8. ISBN 978-0-471-29176-3.
5. Кириленко О.В., Басок Б.І., Базєєв Є.Т., Блінов І.В. Енергетика України та реалії голобального потепління // *Технічна електродинаміка*. 2020. № 3. С. 52-61. DOI: https://doi.org/10.15407/techned2020.03.052.
6. Інтелектуальні електричні мережі: елементи та режими. Під заг. Ред. Акад. НАН України О.В. Кириленко. К.: Ін-т електродинаміки НАН України, 2016. 400 с.
7. Автоматизований моніторинг сигналів синхронізації часу енергосистем: монографія / В.В. Коваль, О.В. Самков, І.В. Блінов, О.Л. Ламеко, І.В. Трач, С.Й. Поліщук, В.І. Вакась, В.В. Чопик, О.Л. Осінський, 2021. К.: Видавничий центр НУБіПУ, 2021. - 380 с.
8. Кириленко О.В., Блінов І.В., Парус Є.В., Трач І.В. Оцінка ефективності використання систем накопичення електроенергії в електричних мережах//*Технічна електродинаміка*. 2021. №4. С. 44–54. DOI: https://doi.org/10.15407/techned2021.04.044.
9. Valerii Koval, Vitaliy Lysenko, Oleksandr Osinskiy, Oleksandr Samkov, Mykola Khudyntsev. Infocommunication Technologies and Networks for Multichannel Monitoring of Synchronization Signals of SMART Grid and Microgrid Electrical Systems // Conference Procedings "International Scientific-Practical Conference «Problems of Infocommunications. Science and Technology»" (PICS&T-2019).

8 - 11 October 2019: Borys Grinchenko Kyiv University, Kyiv, Ukrain. – p.p.153-156. https://ieeexplore.ieee.org/document/8632078.

10. Koval V.V., Lysenko V.P., Kalian D.O., Osinskyi O.L., Samkov O.V. (2021) Improving Efficiency of the Phase-Locked Loop for Reference Oscillator of the Multichannel System for Time Synchronization Signals Telemonitoring. In: Vorobiyenko P., Ilchenko M., Strelkovska I. (eds) Current Trends in Communication and Information Technologies. IPF 2020. Lecture Notes in Networks and Systems, vol 212. Print ISBN 978-3-030-76342-8. Online ISBN 978-3-030-76343-5. Pages 60-79. Springer, Cham. https://doi.org/10.1007/978-3-030-76343-5_4.

11. Величко О.М., Коваль В.В., Самков О.В., Шкляревський І.Ю. Сучасні протоколи передачі шкали часу інтелектуальних електроенергетичних систем зі зниженою аварійністю // Науковий вісник Національного університету біоресурсів і природокористування України. Серія «Техніка та енергетика АПК». – К., 2016. – Вип.242. – С. 41-50.

12. Коваль В.В., Добровенко Д.О., Самков О.В., Осінський О.Л. Багатоканальний моніторинг синхронізуючих сигналів вимірювальних систем / Збірник XX Наук.-практ. конф. "Створення та модернізація озброєння і військової техніки в сучасних умовах''. м. Чернігів, 03 - 04 вересня 2020 р. / ДНДІ ВС ОВТ. – Чернігів: Видавець Брагинець О.В., 2020. – С.116.

13. Valerii Koval, Dmytro Kalian, Oleksandr Osinskiy, Oleksandr Samkov, Mykola Khudyntsev and Vitaliy Lysenko. Diagnostics of Time Synchronization Means of the Integrated Power Grid of SMART Technologies by Using an Optimal Performance System of Automatic Frequency Adjustment // 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2020; Lviv-Slavske; Ukraine; 25 February 2020 до 29 February 2020/ Conference Proceedings15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2020). 2020. - pp.269-276. https://ieeexplore.ieee.org/document/9088587.

14. Автоматизована система синхронізації цифрових сигналів: монографія / В.В. Коваль, О.В. Самков, М.М. Худинцев, Д.О. Кальян. – К.: ТОВ ЦП «Компринт», 2018. - 494 с.

15. Коваль В.В., Самков О.В., Піскун О.М., Медіна М.С., Головня М.В., Шкляревський І.Ю. Інформаційна система передавання еталонних значень шкали часу інтегрованих електроенергетичних мереж SMART-технологій // Вісник університету «Україна». Серія *«Інформатика, обчислювальна техніка, кібернетика»*. – К., 2019. – № 1(22), 2019. – С.231-239.

REFERENCES

1. https://webstore.iec.ch/publication/6028.
2. Standard of SOU NEC 20.261:2021 «Technical policy of SE NEC «Ukrenergo» in the field of development and operation of trunk and interstate electric networks».
3. Dynamic of synergetic systems / Ed. by H. Haken. B. etc., 1980. 271 p.
4. Blanchard BS (2004). System engineering management (3rd ed.). Hoboken, NJ: John Wiley. p. 8. ISBN 978-0-471-29176-3.
5. Kyrylenko O.V., Basok B.I., Baseev Ye.T., Blinov I.V. Energy of Ukraine and the realities of global warming // Tekhnichna elektrodynamika. 2020. № 3. P. 52-61. DOI: https://doi.org/10.15407/techned2020.03.052.
6. Intelligent electrical networks: elements and modes. Under the general editorship of Acad. of the NAS of Ukraine, O.V. Kyrylenko. K.: Institute of electrodynamics of the NAS of Ukraine, 2016. 400 p.
7. Automated monitoring of time synchronization signals of power systems: monograph / V.V. Koval, O.V. Samkov, I.V. Blinov, O.L. Lameko, I.V. Trach, S.Y. Polishchuk, V.I. Vakas, V.V. Chopyk, O.L. Osinsky, 2021. K.: NUBiP of Ukraine, 2021. - 380 p.
8. Kyrylenko O.V., Blinov I.V., Parus Y.V., Trach I.V. Evaluation of efficiency of use of energy storage systems in electric networks // Tekhnichna elektrodynamika. 2021. №4. P. 44–54. DOI: https://doi.org/10.15407/techned2021.04.044.
9. Valerii Koval, Vitaliy Lysenko, Oleksandr Osinskiy, Oleksandr Samkov, Mykola Khudyntsev. Infocommunication Technologies and Networks for Multichannel Monitoring of Synchronization Signals of SMART Grid and Microgrid Electrical Systems // Conference Procedings "International Scientific-Practical Conference «Problems of Infocommunications. Science and Technology»" (PICS&T-2019). 8 - 11 October 2019: Borys Grinchenko Kyiv University, Kyiv, Ukrain. – p.p.153-156. https://ieeexplore.ieee.org/document/8632078.
10. Koval V.V., Lysenko V.P., Kalian D.O., Osinskyi O.L., Samkov O.V. (2021) Improving Efficiency of the Phase-Locked Loop for Reference Oscillator of the Multichannel System for Time Synchronization Signals Telemonitoring. In: Vorobiyenko P., Ilchenko M., Strelkovska I. (eds) Current Trends in Communication and Information Technologies. IPF 2020. Lecture Notes in Networks and Systems, vol 212. Print ISBN 978-3-030-76342-8. Online ISBN 978-3-030-76343-5. Pages 60-79. Springer, Cham. https://doi.org/10.1007/978-3-030-76343-5_4.

11. Velychko O.M., Koval V.V., Samkov O.V., Shklyarevsky I.Y. Modern protocols for the transmission of the time scale of intelligent electric power systems with reduced accident rate // Scientific Bulletin of the National University of Bioresources and Nature Management of Ukraine. Series «Agricultural machinery and energy». – K., 2016. – Issue 242. – P.41-50.

12. Koval V.V., Dobrovenko D.O., Samkov O.V., Osinsky O.L. Multi-channel monitoring of synchronization signals of measuring systems / Collection XX of Scientific and practical conf. on "Creation and modernization of weapons and military equipment in modern conditions". Chernihiv, 03 – 04 September, 2020. / DNDI VS OVT. – Chernihiv: Publisher Brahynets O.V., 2020. – P.116.

13. Valerii Koval, Dmytro Kalian, Oleksandr Osinskiy, Oleksandr Samkov, Mykola Khudyntsev and Vitaliy Lysenko. Diagnostics of Time Synchronization Means of the Integrated Power Grid of SMART Technologies by Using an Optimal Performance System of Automatic Frequency Adjustment // 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2020; Lviv-Slavske; Ukraine; 25 February 2020 до 29 February 2020/ Conference Proceedings15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET 2020). 2020. - pp.269-276. https://ieeexplore.ieee.org/document/9088587.

14. Automated system of synchronization of digital signals: monograph. / V.V. Koval, O.V. Samkov, M.M. Khudyntsev, D.O. Kalian. – K.: TOV TsP «Komprynt», 2018. - 494 p.

15. Koval V.V., Samkov O.V., Piskun O.M., Medina M.S., Golovnya M.V., Shklyarevsky I.Y. Information system for transmission of reference values of the time scale of integrated electric power networks of SMART-technologies // Bulletin of the University «Ukraine». Series «Computer Science, Computer Engineering, Cybernetics». – K., 2019. – № 1(22), 2019. – P.231-239.

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 23.11.2022

УДК 004.942
DOI: https://doi.org/10.53920/ITS-2022-2-2

Олександр Іванович ГОЛУБЕНКО,
канд. техн. наук, доцент
ЗВО «Міжнародний науково-технічний університет
імені академіка Юрія Бугая»
ORCID ID: 0000-0002-1776-5160
Олександр Олександрович ПІДМОГИЛЬНИЙ,
аспірант
Державний університет телекомунікацій
ORCID ID: 0000-0001-8689-2086

GENERATIVE PRE-TRAINED TRANSFORMER 3

GPT (Generative Pre-training Transformer) — це тип штучного інтелекту (AI), який використовує алгоритми машинного навчання для створення тексту природною мовою. Перша версія GPT, випущена в 2018 році, стала революційним досягненням у сфері ШІ та обробки природної мови (NLP). Однак він також мав деякі обмеження та проблеми, які були розглянуті в наступних версіях моделі.

Однією з головних проблем першої версії GPT була відсутність контролю над контентом, який вона генерувала. Модель було навчено на великому наборі даних тексту, створеного людиною, і вона змогла створити зв'язний і, здавалося б, людиноподібний текст на широкий спектр тем. Однак він часто створював текст, який був упередженим, образливим або іншим чином недоречним, оскільки він не міг повністю зрозуміти контекст або значення використаних слів.

Іншою проблемою першої версії GPT була її нездатність виконувати складніші завдання NLP, такі як переклад або конспектування. Хоча він міг створити зв'язний текст, він не міг зрозуміти значення чи структуру тексту так, як це може зробити людина.

Подальші версії GPT, такі як GPT-2 і GPT-3, вирішували ці проблеми та додавали нові можливості, такі як здатність виконувати складніші завдання NLP і генерувати більш зв'язний і відповідний контексту текст. Однак вони все ще мають обмеження і можуть давати необ'єктивні або невідповідні результати, якщо не використовувати їх відповідально.

Ключові слова: *штучний інтелект (AI), машинне навчання, обробка природної мови (NLP), генеративний передтренувальний трансформатор (GPT), генерація тексту, глибоке навчання, нейронна мережа.*

Oleksandr GOLUBENKO
Candidate of technical sciences, associate professor
IHE «Academician Yuri Bugay
international science and technical university»
ORCID ID: 0000-0002-1776-5160
Oleksandr PIDMOGYLNYI
Postgraduate
State University of Telecommunications
ORCID ID: 0000-0001-8689-2086

GENERATIVE PRE-TRAINED TRANSFORMER 3

GPT (Generative Pre-training Transformer) is a type of artificial intelligence (AI) that uses machine learning algorithms to generate text in natural language. The first version of GPT, released in 2018, was a revolutionary breakthrough in AI and natural language processing (NLP). However, it also had some limitations and issues that were addressed in subsequent versions of the model.

One of the main problems with the first version of GPT was the lack of control over the content it generated. The model was trained on a large dataset of human-generated text and was able to generate coherent and seemingly human-like text on a wide range of topics. However, he often produced text that was biased, offensive, or otherwise inappropriate because he could not fully understand the context or meaning of the words used.

Another problem with the first version of GPT was its inability to handle more complex NLP tasks such as translation or annotation. Although he could produce coherent text, he could not understand the meaning or structure of the text as a human could.

Later versions of GPT, such as GPT-2 and GPT-3, addressed these issues and added new capabilities, such as the ability to perform more complex NLP tasks and generate more coherent and context-appropriate text. However, they still have limitations and can produce biased or inconsistent results if not used responsibly.

Keywords: *artificial intelligence (AI), machine learning, natural language processing (NLP), generative pretraining transformer (GPT), text generation, deep learning, neural network.*

Постановка проблеми. Generative Pre-trained Transformer 3 (GPT-3) — це модель штучного інтелекту (AI) для генерації мови,

розроблена OpenAI, яка привернула значну увагу як ЗМІ, так і технічної спільноти. Маючи 175 мільярдів параметрів, GPT-3 наразі є найбільшою та найпотужнішою мовною моделлю з існуючих, а її можливості виходять далеко за рамки простого генерування тексту, що потребує детального аналізу та досліджень.

Аналіз останніх досліджень і публікацій. GPT-3 — це тип мовної моделі на основі Transformer це означає, що він використовує архітектуру transformer для обробки та генерації тексту. Архітектура трансформатора базується на ідеї самоуважності, що дозволяє моделі обробляти вхідні послідовності паралельно, а не послідовно. Це дає змогу моделі фіксувати довгострокові залежності та зв'язки між словами в реченні, що важливо для створення зв'язного тексту.

На додаток до самоуважності, мовні моделі на основі трансформаторів також використовують інші методи, такі як моделювання замаскованої мови та динамічне керування. Моделювання замаскованої мови передбачає маскування частини вхідного тексту та передбачення відсутніх слів на основі контексту, наданого незамаскованими словами. Це допомагає моделі навчитися розуміти зв'язки між словами та створювати текст, який є зв'язним і має сенс у контексті.

Динамічний контроль, з іншого боку, дозволяє моделі регулювати рівень деталізації та конкретності вихідних даних на основі вхідних даних. Це дозволяє моделі створювати текст, який підходить для даного контексту та завдання.

Мета статті – дослідження технічних особливостей GPT-3, аналіз потенційного застосування та впливу, а також розгляд деяких етичних проблем, пов'язаних з його використанням.

Виклад основного матеріалу. Архітектура трансформатора була представлена в оригінальній моделі GPT, але GPT-3 робить крок далі, використовуючи попередньо навчену версію моделі. Попереднє навчання передбачає навчання моделі на великому наборі даних і подальше її тонке налаштування для конкретного завдання, наприклад перекладу мови або відповідей на запитання.

Однією з ключових особливостей GPT-3 є його здатність генерувати зв'язний і схожий на людину текст. Це досягається за рахунок використання механізмів уваги, які дозволяють моделі враховувати контекст і зв'язки між словами в реченні. GPT-3 також має можливість виконувати широкий спектр мовних завдань,

включаючи переклад, узагальнення та відповіді на запитання, і навіть може виконувати завдання програмування [1].

Архітектура трансформатора, що використовується в GPT-3, базується на ідеї самоуважності, яка дозволяє моделі обробляти вхідні послідовності паралельно, а не послідовно. Це дозволяє моделі фіксувати довгострокові залежності та зв'язки між словами в реченні, що важливо для створення зв'язного тексту.

GPT-3 також використовує техніку під назвою масковане моделювання мови, яка передбачає маскування частини вхідного тексту та передбачення відсутніх слів на основі контексту, наданого незамаскованими словами. Щоб реалізувати масковане моделювання мови, частина вхідного тексту вибирається випадковим чином і замінюється спеціальним маркером, таким як «[MASK]». Потім модель навчається передбачати пропущене слово або слова на основі контексту, наданого незамаскованими словами. Наприклад, враховуючи вхідні дані «[MACKA] сиділа на килимку», модель буде навчена передбачати пропущене слово «кіт» на основі контексту, наданого іншими словами в реченні. Це допомагає моделі навчитися розуміти зв'язки між словами та створювати текст, який є зв'язним і має сенс у контексті [2].

На додаток до цих методів GPT-3 також використовує техніку, яка називається динамічним керуванням, яка дозволяє моделі регулювати рівень деталізації та конкретності вихідних даних на основі вхідних даних. Модуль динамічного керування може бути реалізований за допомогою різноманітних методів, таких як Transformer encoder або окрема нейронна мережа прямого зв'язку [5]. Transformer encoder в модулі керування обробляє вхідні дані та генерує керуючий сигнал, який використовується для налаштування рівня деталізації та конкретності на виході.

Динамічний контроль дозволяє GPT-3 генерувати текст, який відповідає заданому контексту та завданню. Наприклад, якщо введенням є запит на детальний опис особи, модуль керування генеруватиме керуючий сигнал, який повідомляє моделі створити більш детальний і конкретний опис. Якщо введенням є підказка із запитом про загальний огляд теми, модуль керування згенерує керуючий сигнал, який скаже моделі створити більш загальний і високорівневий опис. Це дозволяє GPT-3 генерувати текст, який підходить для заданого контексту та завдання.

GPT-3 також може включати зовнішні знання у свій вихід, що дозволяє генерувати текст, який є фактично точним та інформативним. Це досягається завдяки використанню трансформатора знань, який є окремим компонентом моделі, яка навчається на великому наборі даних із багатим на знання текстом.

GPT-3 був навчений на великому наборі текстових даних під назвою Internet Archive Books dataset. Цей набір даних складається з понад 8 мільйонів книг та інших текстів, які були оцифровані Інтернет-архівом, некомерційною організацією, яка працює над збереженням і наданням доступу до творів культури. Набір даних включає широкий спектр текстів, включаючи книги, статті та веб-сайти, і охоплює широкий діапазон мов і тем.

Окрім набору даних Internet Archive Books, GPT-3 також навчався на інших наборах даних, включаючи англійську Вікіпедію та набір даних WebText, який складається з тексту, взятого з Інтернету [4]. Комбінація цих наборів даних дозволяє GPT-3 навчатися на різноманітних і репрезентативних зразках тексту, що допомагає генерувати високоякісний і зв'язний результат.

Потенційне застосування та вплив

GPT-3 має потенціал для революції в обробці природної мови та штучному інтелекті, і він уже використовується для широкого спектру програм. Деякі з найбільш перспективних потенційних застосувань GPT-3 включають:

- Чат-боти: GPT-3 можна використовувати для створення чат-ботів, які можуть вести природні розмови з користувачами. Це може мати широкий спектр застосувань, наприклад, обслуговування клієнтів, освіта та розваги.
- Мовний переклад: GPT-3 може виконувати мовний переклад з високою точністю, що може мати значні наслідки для глобального спілкування та співпраці.
- Створення вмісту: GPT-3 має можливість генерувати текст, схожий на людину, який можна використовувати для автоматизації створення вмісту для веб-сайтів, соціальних мереж та інших платформ.
- Композиція музики: GPT-3 навіть використовувався для створення музики, демонструючи його здатність розуміти та генерувати складні моделі.

Одним із ключових застосувань GPT-3 є створення контенту [3]. Модель може генерувати високоякісні статті, публікації в блогах і со-

ціальних мережах, які можуть бути корисними для підприємств і організацій, які прагнуть швидко й ефективно створювати вміст. GPT-3 також можна використовувати для більш творчих додатків, таких як створення музики та написання історій [6].

Окрім можливостей створення мови, GPT-3 також має здатність виконувати такі завдання, як очищення та форматування даних, що може бути корисним для підприємств і організацій, які обробляють великі набори даних.

Незважаючи на свої вражаючі можливості, GPT-3 не позбавлений обмежень. Одним із потенційних обмежень GPT-3 є його залежність від даних, на яких він навчався [7]. Якщо навчальні дані містять упереджений або невідповідний вміст, модель може створити упереджений або невідповідний вихід. Важливо ретельно розглянути різноманітність і якість даних, які використовуються для навчання GPT-3, і переконатися, що вони є репрезентативними для населення або завдання, для якого вони використовуються.

Іншим потенційним обмеженням GPT-3 є його вартість і вимоги до ресурсів. Модель вимагає великої кількості обчислювальних ресурсів і даних для ефективного навчання, що може зробити її дорогим і ресурсомістким у використанні [8]. Це може обмежити його доступність для деяких організацій.

Висновки та пропозиції. Загалом GPT-3 — це потужна та універсальна мовна модель, яка здатна виконувати широкий спектр мовних завдань і генерувати високоякісний текст. Однак важливо ретельно розглянути потенційні обмеження та етичні міркування моделі, використовуючи її для конкретних цілей.

Є кілька ключових особливостей і характеристик GPT-3, які роблять його унікальним і потужним інструментом для досліджень НЛП і ШІ. Ось деякі з основних ключів до розуміння GPT-3:

- Великий масштаб: GPT-3 є однією з найбільших мовних моделей, коли-небудь розроблених, із 175 мільярдами параметрів. Це дозволяє генерувати високоякісний текст і виконувати складні завдання NLP з рівнем продуктивності, якого важко досягти з меншими моделями.
- Трансформаторна архітектура: GPT-3 використовує трансформаторну архітектуру, яка є типом нейронної мережі, яка широко використовується в завданнях НЛП. Архітектура трансформатора дозволяє GPT-3 обробляти введе-

ний текст більш ефективним і результативним способом, дозволяючи генерувати текст, який є більш зв'язним і відповідним контексту.

- Моделювання замаскованої мови (MLM): GPT-3 використовує техніку під назвою моделювання замаскованої мови (MLM), яка передбачає маскування частини вхідного тексту та прогнозування відсутніх слів на основі навколишнього контексту. Це допомагає GPT-3 дізнатися про зв'язки між словами та структурою мови.
- Динамічне керування: GPT-3 представляє нову техніку, яка називається динамічним керуванням, яка дозволяє моделі адаптувати свої результати на основі контексту чи поточного завдання. Це дозволяє GPT-3 генерувати текст, який є більш актуальним і підходить для різних ситуацій.

GPT-3 є потужним і універсальним інструментом, який має потенціал зробити революцію в дослідженнях НЛП та ШІ. Його великий масштаб і розширені можливості роблять його цінним ресурсом для дослідників і розробників, які працюють над широким спектром програм і проектів.

ЛІТЕРАТУРА

1. Language models are better than humans at next-token prediction Authors: Buck Shlegeris, Fabien Roger, Lawrence Chan, Euan McLean. arXiv:2212.10560.

2. Self-Instruct: Aligning Language Model with Self Generated Instructions. Authors: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi. arXiv:2212.10509.

3. Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions Authors: Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, Ashish Sabharwal. arXiv:2212.08072.

4. Foresight -- Deep Generative Modelling of Patient Timelines using Electronic Health Records Authors: Zeljko Kraljevic, Dan Bean, Anthony Shek, Rebecca Bendayan, Joshua Au Yeung, Alexander Deng, Alfie Baston, Jack Ross, Esther Idowu, James T Teo, Richard J Dobson. arXiv:2212.04037.

5. Demystifying Prompts in Language Models via Perplexity Estimation Authors: Hila Gonen, Srini Iyer, Terra Blevins, Noah A. Smith, Luke Zettlemoyer. arXiv:2211.09267.

6. Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality.

Authors: Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, Xiang Ren. arXiv:2211.07615.

7. UGIF: UI Grounded Instruction Following. Authors: Sagar Gubbi Venkatesh, Partha Talukdar, Srini Narayanan. arXiv:2210.17497.

8. Leveraging Pre-trained Models for Failure Analysis Triplets Generation. Authors: Kenneth Ezukwoke, Anis Hoayek, Mireille Batton-Hubert, Xavier Boucher, Pascal Gounet, Jerome Adrian. arXiv:2210.17238.

REFERENCES

1. Language models are better than humans at next-token prediction Authors: Buck Shlegeris, Fabien Roger, Lawrence Chan, Euan McLean. arXiv:2212.10560

2. Self-Instruct: Aligning Language Model with Self Generated Instructions. Authors: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi. arXiv:2212.10509

3. Interleaving Retrieval with Chain-of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions Authors: Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, Ashish Sabharwal. arXiv:2212.08072

4. Foresight -- Deep Generative Modelling of Patient Timelines using Electronic Health Records Authors: Zeljko Kraljevic, Dan Bean, Anthony Shek, Rebecca Bendayan, Joshua Au Yeung, Alexander Deng, Alfie Baston, Jack Ross, Esther Idowu, James T Teo, Richard J Dobson. arXiv:2212.04037

5. Demystifying Prompts in Language Models via Perplexity Estimation Authors: Hila Gonen, Srini Iyer, Terra Blevins, Noah A. Smith, Luke Zettlemoyer. arXiv:2211.09267

6. Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality Authors: Pei Zhou, Hyundong Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, Xiang Ren. arXiv:2211.07615

7. UGIF: UI Grounded Instruction Following. Authors: Sagar Gubbi Venkatesh, Partha Talukdar, Srini Narayanan. arXiv:2210.17497

8. Leveraging Pre-trained Models for Failure Analysis Triplets Generation. Authors: Kenneth Ezukwoke, Anis Hoayek, Mireille Batton-Hubert, Xavier Boucher, Pascal Gounet, Jerome Adrian. arXiv:2210.17238

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 23.11.2022

УДК 004.021
DOI: https://doi.org/10.53920/ITS-2022-2-3

Андрій Вікторович ЛЕМЕШКО,
д-р філософії, доц.,
Державний університет телекомунікацій
ORCID ID: 0000-0001-8003-3168
Єлизавета Олександрівна НОВІЧЕНКО,
студентка,
Державний університет телекомунікацій
ORCID ID: 0000-0002-4354-4513
Андрій Володимирович НЕДАВНІЙ,
студент
Державний університет телекомунікацій
ORCID ID: 0000-0002-4383-8933

БЕЗПЕКА ДАНИХ В УКРАЇНІ ЗА ДОПОМОГОЮ ВИКОРИСТАННЯ ТЕХНОЛОГІЇ VPN

Початок XXI століття анонсував собою глобалізацію в усьому світі та швидкий розвиток інформаційних технологій, а саме мережі Інтернет, що в свою чергу допомогло зробити великий стрибок у розвитку передачі даних та доступності інформації серед великого кола людей.

Сьогодні Інтернет знаменує собою глобальний віртуальний світ з безліччю безкоштовної інформації та даних. Попит на нього зростає щодня, що спонукає постачальників інтернет послуг на постійний розвиток ринку.

На сьогоднішній день важко уявити світ без вільного доступу до Інтернету. На жаль, уряди деяких країн на законодавчому рівні обмежують доступ до тих чи інших ресурсів, що, в свою чергу, збільшує попит на розвиток та використання VPN технологій та сервісів. Деякі користувачі використовують VPN для анонімності в мережі Інтернет та отримання доступу до заблокованих ресурсів. Інші - користуються данною технологією для захисту особистої інформації. Під час вибору VPN-сервісу деякі користувачі керуються якістю послуг, які будуть надаватись, а інші - їх вартістю. Власники VPN-сервісів постійно вдосконалюють якість свої послуг та впроваджують нові технології.

Попит на VPN-сервіси, після початку повномасштабного російського вторгнення, в Україні виріс в рази - не тільки за рахунок блокування українських медіа ресурсів, а й за рахунок появи IT-армії України. Завдяки чому деякі VPN-сервіси почали безкоштовно надавати доступ українцям до своїх серверів.

VPN має декілька рівнів захисту такі як: шифрування даних, аутентифікація джерела даних, перевірка хешу, що в свою чергу забезпечує конфіденційність передаваємих даних в Інтернеті. В сумісності, це все допомагає підвищити рівень захисту особистих даних користувачів.

Громадяни України, котрі залишаються на тимчасово окупованих територіях, в більшій мірі, мають доступ тільки до російського медіа простору, за рахунок того, що український ресурси блокуються, а операторів зв'язку «глушать» та знищується їх інфраструктура. За допомогою VPN-сервісів вони можуть отримати доступ до українського медіапростору.

Ключові слова: *інформаційні технології, VPN, попит, сервіси, Інтернет, користувачі, доступ, інформація.*

Andrii LEMESHKO
Doctor of Philosophy, Associate Professor
State University of Telecommunications
ORCID ID: 0000-0001-8003-3168
Yelyzaveta NOVICHENKO
Student
State University of Telecommunications
ORCID ID: 0000-0002-4354-4513
Andriy NEDAVNIY
Student
State University of Telecommunications
ORCID ID: 0000-0002-4383-8933

DATA SECURITY IN UKRAINE USING VPN TECHNOLOGY

The beginning of the 21st century heralded globalization worldwide and the rapid development of information technologies, namely the Internet, which in turn helped to make a big leap in the development of data transmission and the availability of information among a large number of people.

Today, the Internet represents a global virtual world with a wealth of free information and data. The demand for it is growing every day, which encourages Internet service providers to constantly develop the market.

Today, it is difficult to imagine a world without free access to the Internet. Unfortunately, the governments of some countries at the legislative level limit access to certain resources, which, in turn, increases the demand for the development and use of VPN technologies and services. Some users use VPN for anonymity on the Internet and access to blocked resources. Others use this technology to protect personal information. When choosing a VPN service, some users are guided by the quality of the services that will be provided, while others are guided by their cost. Owners of VPN services constantly improve the quality of their services and introduce new technologies.

The demand for VPN services, after the beginning of the full-scale Russian invasion, in Ukraine has grown many times - not only due to the blocking of Ukrainian media resources, but also due to the appearance of the IT army of Ukraine. As a result, some VPN services began to provide Ukrainians with free access to their servers.

VPN has several levels of protection, such as: data encryption, data source authentication, hash verification, which in turn ensures the confidentiality of transmitted data on the Internet. In compatibility, all this helps to increase the level of protection of personal data of users.

Citizens of Ukraine, who remain in the temporarily occupied territories, to a greater extent, have access only to the Russian media space, due to the fact that Ukrainian resources are blocked, and communication operators are «jammed» and their infrastructure is destroyed. With the help of VPN services, they can get access to the Ukrainian media space.

Key words: *information technologies, VPN, demand, services, Internet, users, access, information.*

Постановка проблеми. Загарбники намагаються усіляко відгородити людей на окупованих територіях від цивілізованого світу. Один зі способів – блокування зв'язку та інтернету.

Окупанти на тимчасово захоплених територіях намагаються приєднати українців до російських інтернет-мереж, де стоїть обладнання для фільтрації інтернет-трафіку, що дозволило заблокувати безліч українських та міжнародних веб ресурсів. Щоб уникнути слідкування та обійти обмеження, українським корис-

тувачам, яких під'єднали до російських мереж, необхідно використовувати VPN-сервіси.

VPN (Virtual Private Network) – це віртуальна приватна мережа, яка забезпечує шифрування трафіку між клієнтом та VPN-сервером і зміну IP-адреси. При підключенні до VPN створюється захищений канал між комп'ютером користувача і VPN-сервером. Дані в ньому надійно зашифровані: інтернет-провайдер не дізнається локації користувача та веб ресурсів, які він відвідував. Оновлена IP-адреса зазвичай створюється з іншого міста або країни.

VPN-сервіси дозволяють користуватися ресурсами, доступ до яких заборонено за географічним принципом або на підставі рішень органів влади. Завдяки VPN можна вільно відвідувати заблоковані сайти, достатньо лише вибрати та завантажити додаток на свій комп'ютер або мобільний пристрій

Мета статті – дослідження роботи технології VPN, а також основні її принципи. Представити причини використання даної технології. Розглянути типи та протоколи VPN.

Завданнями даної роботи:

- створити поради, які допоможуть обійти блокування VPN та безпечно користуватися месенджерами на окупованих територіях України;
- визначити основні структурні елементи зовнішньоекономічної стратегії держави, зосереджуючи увагу на інноваційній діяльності, для визначення основних принципів розвитку інноваційних процесів та проблем інвестиційної привабливості України.

Виклад основного матеріалу. VPN – віртуальна приватна мережа, що являє собою тип веб-служби, яка дозволяє користувачам приховувати активність у мережі, особистість і розташування під час роботи в мережі.

Зазвичай для доступу в Інтернет комп'ютер створює загальнодоступне з'єднання з постачальником інтернет-послуг. VPN створює приватне з'єднання між комп'ютером і віддаленим сервером, що належить постачальнику VPN. Це цифрове підключення або тунель шифрує дані користувача, щоб ніхто інший не міг їх побачити. Він також маскує IP-адресу користувача, щоб ніхто інший не міг його відстежити. В результаті робота в Інтернеті стає безпечнішою, надійнішою та анонімною [1].

По самому визначенню VPN-підключення:

1. віртуальне тому, що в процесі підключення не задіяні фізичні кабелі;
2. приватний, тому що через це підключення ніхто інший не може бачити дані користувача чи дії в Інтернеті;
3. мережеве, тому що кілька пристроїв — комп'ютер користувача і VPN-служба — працюють разом, щоб підтримувати встановлений зв'язок.

Головні причини використання VPN

1. Захист даних.

Конфіденційні дані, такі як робочі листи, платіжна інформація та позначки розташування, постійно передаються в Інтернеті. Цю інформацію можна відстежити та легко використовувати, особливо в загальнодоступній мережі, де будь-хто, хто має доступ до мережі, потенційно може отримати доступ до особистих даних. Підключення VPN перетворює дані на код і робить їх нечитаними для всіх, хто не має ключа шифрування. Він приховує активність у мережі, щоб ніхто інший не міг її побачити.

2. Робота з дому.

Сьогодні віддалена робота поширена як ніколи. За допомогою VPN віддалені працівники можуть отримати доступ до ресурсів компанії через приватне з'єднання з будь-якого місця, якщо вони можуть виходити в Інтернет. Це дає співробітникам велику гнучкість, а також гарантує, що дані компанії залишаться захищеними та безпечними навіть у загальнодоступній мережі Wi-Fi.

3. Отримання доступ до регіонального вмісту або транслювання його з будь-якого місця.

Деякі сайти та служби обмежують свій медіаконтент залежно від географічного положення, що означає, що може не бути доступу до певних видів контенту. VPN маскує або підробляє розташування локального сервера, щоб він виглядав так, ніби він знаходиться в іншому місці, наприклад, в іншій країні.

4. Обхід блокування та спостереження.

У деяких регіонах доступ до певних сайтів або служб може бути недоступним через урядові обмеження, цензуру або стеження. Спуфінг розташування дає користувачам можливість обходити брандмауери, переглядати заблоковані веб-сайти і вільно переміщатися в мережі.

5. Заборонити відстеження з боку Інтернет-провайдера та сторонніх осіб.

Постачальники послуг Інтернету (ISP) реєструють та відстежують історію відвідувань через унікальну IP-адресу вашого пристрою. Ця інформація потенційно може бути продана стороннім рекламодавцям, передана уряду або залишена вразливою перед порушенням безпеки. Шляхом маршрутизації на віддалений VPN-сервер замість серверів інтернет-провайдера VPN маскує IP-адресу користувача, запобігає відстеженню постачальником інтернет-послуг та зберігає конфіденційність особистих даних [2].

Типи VPN

Існує чотири основні типи VPN:

1. VPN-брандмауер оснащений як брандмауером, так і можливостями VPN. Цей тип використовує захист, що надається брандмауерами, для обмеження доступу до внутрішньої мережі та забезпечує переведення адрес, автентифікацію користувача, аварійні сигнали та протоколювання.

2. Апаратна VPN забезпечує високу пропускну здатність мережі, а також покращує продуктивність та надійність, але є дорогою.

3. Програмний VPN забезпечує гнучкість з погляду управління трафіком. Це найкраще, коли кінцеві точки не контролюються однією стороною і при використанні різних брандмауерів та маршрутизаторів.

4. Безпечний рівень сокету (SSL) VPN дозволяє користувачам підключатися до VPN-пристроїв за допомогою веб-браузера. SSL використовується для шифрування трафіку між веб-браузером та пристроєм VPN.

Протоколи VPN

Протоколи тунелювання VPN пропонують різні функції та рівні безпеки, і для кожного з них є переваги та недоліки. Існує п'ять основних протоколів тунелювання VPN:

1. SSTP використовує протокол HTTPS для передачі трафіку через брандмауери та веб-проксі, які можуть блокувати інші протоколи. SSTP надає механізм перенесення трафіку протоколу «точка-точка» (PPP) каналом SSL. Використання PPP дозволяє підтримувати надійні методи аутентифікації, а SSL забезпечує безпеку на рівні транспорту з розширеним узгодженням ключів, перевіркою шифрування та цілісності.

2. PPTP дозволяє зашифрувати багатопротокольний трафік і обернути його в заголовок, який буде надіслано через мережу інтернет-протоколу (IP). PPTP можна використовувати для віддаленого доступу та VPN-з'єднань «точка-точка». При використанні інтернету PPTP-сервер є VPN-сервером з підтримкою PPTP з одним інтерфейсом в інтернеті та другим інтерфейсом корпоративної інтрамережі. PPTP використовує з'єднання протоколу керування передачею для керування тунелями та інкапсуляції загальної маршрутизації для перенесення кадрів PPP для тунельованих даних.

3. L2TP дозволяє зашифрувати багато-протокольний трафік, а потім використовувати будь-який носій, що підтримує доставку даних PPP, наприклад IP або асинхронний режим передачі. L2TP – це комбінація PPTP та Layer 2 Forwarding (L2F). L2TP представляє найкращі функції PPTP та L2F. На відміну від PPTP, L2TP покладається на IP безпеку (IPsec) у транспортному режимі для служб шифрування. Комбінація L2TP та IPsec відома як L2TP/IPsec. Обидва L2TP та IPsec повинні підтримуватись як клієнтом VPN, так і VPN-сервером. L2TP/IPsec - ідеальна передова таємність.

4. OpenVPN – це програмна програма з відкритим вихідним кодом, яка реалізує методи VPN для створення безпечних з'єднань «точка-точка» або «сайт-сайт» у маршрутизованих або мостових конфігураціях та засобах віддаленого доступу. Він використовує власний протокол безпеки, який використовує SSL/TLS для обміну ключами. OpenVPN дозволяє одноранговим вузлам аутентифікувати один одного за допомогою секретного ключа, сертифіката або імені користувача та пароля. Більшість провайдерів VPN, які використовують OpenVPN, використовують пряму таємність.

5. IKEv2 – це протокол на основі протоколу IPSec, який використовується у Windows 7 та вище. IKEv2 – це стандарт наступного покоління для безпечного обміну ключами між однорангові VPN-пристроями. IKEv2 особливо корисний в автоматичному відновленні VPN-з'єднання, коли користувачі тимчасово втрачають свої з'єднання.

Отже, розглянувши VPN-протоколи, можна порівняти їх за наявністю таких критеріїв, як: багатопротокольне тунелювання, підтримка аутентифікації та шифрування, управління потоком

даних у тунелі, управління правами користувачів, сфера використання, перспективи розвитку. Дане порівняння представлено в таблиці 1 [3].

Таблиця 1. Порівняльні характеристики протоколів VPN

Критерії	Протоколи			
	L2F	L2TP	IPSec	SSL/TSL
Багатопротокольне тунелювання	Так	Так	Так	Ні
Підтримка аутентифікації та шифрування	Відсутнє	Слабке	Присутнє	Дуже надійне
Управління потоком даних у тунелі	Ні	Ні	Так	Так
Управління правами користувачів	Ні	Ні	Ні	Так
Сфера використання	Віддалений доступ через провайдера	Віддалений доступ через провайдера	Для реалізації особистого рішення	Для реалізації особистого рішення
Перспективи розвитку	Слабкі	Існують	Сильні	Сильні

VPN шифрування військового рівня

Шифрування відігравало велику роль протягом тисячоліть. З розвитком сучасних технологій він залишався в авангарді інновацій, постійно розвиваючись. Хоча шифрування змінювалося формою і розвивалося протягом історії, воно завжди працювало однаково. Щоб працювати, шифрування потребує безпечного ключа, який розшифровує та шифрує. Без ключа, здатного розшифрувати зашифрований текст чи алгоритм, шифрування залишиться.

В онлайн-спілкуванні користувачі можуть застосовувати програмне забезпечення для шифрування, щоб захистити свої особисті дані від будь-кого, хто може захотіти заволодіти ними. Використовуючи математичний алгоритм, шифрування шифрує дані

в такий спосіб, що можна розшифрувати лише з допомогою певного ключа. Існує безліч типів шифрування, кожен із яких забезпечує певний рівень безпеки. В даний час найнадійнішим і найчастіше використовуваним шифруванням є AES, також відомий як Advanced Encryption Standard. Він пропонує ключ шифрування розміром до 256 біт, шифрування, яке може протистояти практично будь-якій хакерській атаці.

Як випливає з назви, VPN-шифрування військового рівня – це стандартне шифрування, яке використовується військовими установами. Це найвище доступне шифрування AES. Оскільки військові установи часто діють непомітно, вони зазвичай використовують найкращі доступні протоколи безпеки, щоб гарантувати, що кожна частина інформації залишиться прихованою та зашифрованою.

У зв'язку з цим, коли VPN-сервіс, такий як ZoogVPN пропонує своїм клієнтам стандарт шифрування VPN військового рівня, це означає, що вони надають їм найбезпечніший і надійніший протокол шифрування, доступний в даний час. Користувачі можуть бути повністю впевнені у надійності та безпеці сервісу.

Через те, що це дуже просунута і складна технологія шифрування, не багато VPN-сервісів можуть собі це дозволити. Саме тому більшість VPN-сервісів використовують звичайне чи, у деяких випадках, розширене шифрування. Це все ще досить безпечно та надійно, але набагато слабше, ніж шифрування VPN військового рівня. В ідеалі при виборі VPN-сервісу користувачі повинні шукати той, який пропонує максимально можливі стандарти шифрування.

VPN-шифрування військового рівня — це найкраще, що можна купити за гроші прямо зараз. Він може протистояти практично будь-якій спробі вторгнення. Це пов'язано з тим, що у нього багато рівнів безпеки, тому щоразу, коли хакер порушує один із рівнів, система автоматично закриває його точку входу, і йому доводиться шукати інший вхід. У той самий час, коли система розпізнає порушення, вона починає відстеження самого хакера як контрзаходи. Цей тип шифрування – найкращий спосіб запобігти та зупинити будь-які потенційні порушення безпеки. Щоб відповісти на питання, наскільки безпечно шифрування VPN військового рівня, це надзвичайно безпечно, і цього є більш ніж достатньо для захисту кожного онлайн-користувача. Хоча більшість онлайн-користувачів навіть не замислюються про цей аспект свого VPN-сервісу [6].

Рекомендовані VPN-сервіси для жителів тимчасово окупованої території

1. ExpressVPN;

Цей VPN-сервіс займає перше місце у більшості рейтингів. Його рекомендують різні технологічні видання, наприклад, The Verge. Він зареєстрований на Британських Віргінських островах, тому не підпадає під юрисдикцію урядів США та Європи й не передає дані користувачів до поліції за запитом.

ExpressVPN використовує алгоритм AES (Advanced Encryption Standard) з 256-бітним ключем, також відомий як AES-256. Ця технологія схвалена урядом США і використовується фахівцями з безпеки у всьому світі для шифрування секретної інформації.

Недоліком є ціна: пакет на місяць коштує $9,9, але ви можете зекономити, придбавши пакет одразу на пів року або рік.

2. VPN Unlimited;

Цей сервіс рекомендує Нацполіція України й такі великі ресурси як Engadget і vpnMentor. Країна реєстрації – США.

Компанія обіцяє захист конфіденційності своїх користувачів, тому не зберігає інформацію про використані сервери, зміст отриманих даних, історію браузера та загальний час з'єднання.

Сервіс VPN Unlimited переказує на допомогу українській армії 30% від своїх доходів і спочатку запропонував українцям безоплатне використання VPN-сервісу на рік. Проте зараз роздачу промокодів на безплатне користування тимчасово припинили.

3. TunnelBear;

Ще один VPN-сервіс, який радить Нацполіція та консультанти з цифрових технологій. Місце його реєстрації – Канада. За замовчуванням використовує надійний алгоритм AES 256. Також має захист від блокування VPN за допомогою технології GhostBear. Таким чином використання VPN буде складніше виявити державним органам, організаціям й інтернет-провайдерам.

Цей сервіс підтримує українців під час війни. Після російського вторгнення він запропонував українцям 100 ГБ інтернет-трафіку безплатно.

4. NordVPN;

VPN-сервіс і флагманський продукт компанії Nord Security, що працює в галузі кібербезпеки й підтримується Литовським стартап акселератором. Отримав хороші відгуки провідних технологічних видань світу. Країна реєстрації – Панама.

Протокол шифрування та безпека – NordLynx, OpenVPN, IKEv2. Шифрування AES-256. Приватний DNS, подвійне шифрування даних і підтримка Onion, блокування реклами та шкідливих програм, Kill-Switch.

Коштує цей VPN від $4,13 на місяць.

5. ClearVPN.

Це VPN від української компанії MacPaw, який став цьогорічним переможцем Global Infosec Award у сфері кібербезпеки.

Попри легкий у використанні функціонал, над проєктом працювали професіонали кібербезпеки. Всі дані користувачів захищені сучасними протоколами, AES-256 шифруванням і додатковою функцією Kill Switch. Розробники гарантують швидке з'єднання та безпеку. Також застосунок має no-logs політику, яка забороняє збирати будь-яку інформацію щодо дій користувача: від IP-адреси до запитів.

Під час війни ClearVPN став безплатним для українців, тож кожен може завантажити його собі, щоб безпечно користуватися мережею.

У Держспецзв'язку зазначають, що краще обрати VPN, в якому є шифрування військового класу (наприклад, AES-256), асортимент протоколів безпеки (OpenVPN, L2TP, IKEv2, WireGuard та інші), захист від витоків через DNS, технологія TrustedServer, яка видаляє всі ваші дані при кожному перезапуску, та функцію автоматичного аварійного відключення [4].

Підтримка зв'язку на окупованих територіях України

На тимчасово окупованих територіях росіяни перевіряють телефони, вимикають зв'язок та обмежують можливість читати українські новини.

На сьогодні до переліку тимчасово окупованих територій частково входять: Донецька, Харківська, Луганська, Запорізька, Херсонська області та АР Крим.

На них окупаційна влада направила інтернет-трафік через Росію – так вони хочуть запровадити цензуру, обмежити мешканцям регіону доступ до українських і світових джерел інформації та тримати людей у тотальній невідомості.

При затриманні чи перевірці документів перевіряються фотографії, читається листування, інтернет переходить під тотальний контроль.

Поради як налагодити зв'язок і убезпечити користування смартфоном та інтернетом:

1. Очищайте телефон. Щоб не спровокувати ворога, видаліть всі листування, фотографії, історію браузера та історію дзвінків. Застосовуйте вбудований функціонал додатків, що забезпечують конфіденційність (приватні вкладки браузера, що не зберігають історію, секретні чати із таймером видалення повідомлень).

2. Використовуйте блокування. Можна встановити графічний пароль або спеціальний код. Відключіть функції розблокування обличчям та відбитком пальця – так стороннім буде складніше увійти до важливих програм.

3. Зберігайте копії всіх важливих документів на хмарі або у прихованих папках та альбомах.

4. Використовуйте секретні чати та надсилайте повідомлення, що зникають через деякий час. Фотографуйте через месенджери, не зберігайте ці фотографії на телефоні.

5. Налаштуйте свій смартфон у режим віддаленого управління. У сучасних моделях смартфонів є можливість віддаленого стирання всіх даних на вашому телефоні (для цього необхідно знати пароль до хмарного облікового запису телефону (Google Account, Apple ID, Samsung Account).

6. Використовуйте VPN.

Сенс роботи VPN – перенаправляти інтернет-з'єднання через сервери в різних точках земної кулі з метою приховати вашу реальну IP-адресу і тим самим обійти блокування.

Обхід VPN-блокування на окупованих територіях

Поради для обходу VPN-блокування:

1. Використовуйте різні VPN-сервіси. Зазвичай блокують IP-адреси найпопулярніших VPN. Але всі доступні VPN заблокувати неможливо.

2. Встановіть свій приватний VPN-сервер. Це означає, що у вас буде своя IP-адреса VPN, яку не буде заблоковано. Щоб його створити, не потрібно бути IT-експертом. В інтернеті є багато статей та посібників зі створення користувацьких VPN-серверів. Але в такому випадку ви не отримаєте стандартних гарантій безпеки, як під час роботи з платними VPN.

3. Використовуйте інший протокол VPN. Сучасні VPN роблять з урахуванням набору з різних протоколів. Деякі з них ставлять в основу конфіденційність і безпеку користувача. Інші призначені для забезпечення швидкої роботи. Залежно від того, яку програму

VPN встановлює користувач, він може змінити протокол, перейшовши на панель налаштувань і вибравши необхідний VPN-протокол. Найкращі VPN-протоколи: WireGuard, OpenVPN, L2tp/IPsec, iKev2.

4. Налаштуйте свій браузер для обходу блокування. Використовуйте спеціальні VPN-розширення для браузера.

Приховування VPN-програми

Велика ймовірність, що при контакті з окупаційними військами, наприклад у разі перетину лінії розмежування, вони вимагатимуть ваш смартфон на перевірку. Завжди вчасно очищуйте історію браузера та месенджерів. Видаляйте соціальні мережі, які можуть видати вашу проукраїнську позицію та створити вам проблеми під час спілкування з окупантами. Не забувайте також видаляти фотографії та файли (наприклад, завантаження програм VPN). Але намагайтеся, щоб телефон не виглядав підозріло чистим, залиште мем, фотографії сім'ї, нейтральну соцмережу, один очищений месенджер.

Якщо у вас є можливість, видаліть програму VPN, оскільки вона може привернути увагу окупанта. Якщо вам потрібно залишити його на телефоні, є кілька способів приховати його.

На Android-смартфонах ви можете користуватися функцією «приховати програму», якщо ваша версія операційної системи це підтримує. Ви також можете встановити сторонній лаунчер, наприклад Nova Launcher і змінити дизайн іконки будь-якої програми. Так ви зможете замаскувати VPN та будь-яку іншу програму, наприклад, під калькулятор [5].

Висновки та пропозиції. В умовах війни російські війська можуть блокувати доступ до українських державних сайтів та ЗМІ на тимчасово окупованих територіях України. Громадянам України важливо отримувати інформацію про евакуацію, рекомендації державних органів та роботу волонтерів. Якщо користувач знаходитьсяна території, де сили окупантів блокують доступ до життєво важливих джерел інформації, йому знадобляться засоби обходу блокування сайтів – технологія VPN.

Розроблені поради необхідно використовувати для жителів тимчасово окупованої території України для забезпечення безпеки свої даних, а також для доступу до правдивої інформації.

ЛІТЕРАТУРА

1. Чи є безпечний VPN [Електронний ресурс]. Режим доступу: https://azure.microsoft.com/ru-ru/resources/cloud-computing-dictionary/what-is-a-vpn/#what-is-a-vpn.

2. Використання технології vpn для забезпечення інформаційної безпеки [Електронний ресурс]. Режим доступу: https://cyberleninka.ru/article/n/ispolzovanie-tehnologii-vpn-dlya-obespecheniya-informatsionnoy-bezopasnosti.

3. Для тих, хто в окупації: підбірка безпечних VPN-сервісів для користування інтернетом [Електронний ресурс]. Режим доступу: https://henichesk.city/articles/219181/dlya-tih-hto-v-okupacii-pidbirka-bezpechnih-vpn-servisiv-dlya-koristuvannya-internetom.

4. Держспецзв'язку: Що таке VPN, і як ним безпечно користуватись [Електронний ресурс]. Режим доступу: https://www.kmu.gov.ua/news/derzhspeczvyazku-shcho-take-vpn-i-yak-nim-bezpechno-koristuvatis.

5. VPN, Tor та спілкування у месенджерах. Як підтримувати зв'язок на окупованих територіях. Детальна інструкція [Електронний ресурс]. Режим доступу: https://forbes.ua/innovations/vpn-tor-ta-spilkuvannya-u-messendzherakh-yak-pidtrimuvati-zvyazok-na-okupovanikh-teritoriyakh-detalna-instruktsiya-10082022-7616.

6. Що таке VPN-шифрування військового рівня? [Електронний ресурс]. Режим доступу: https://zoogvpn.com/military-grade-vpn-encryption/.

REFERENCES

1. Is VPN secure [Electronic resource]. Access mode: https://azure.microsoft.com/ru-ru/resources/cloud-computing-dictionary/what-is-a-vpn/#what-is-a-vpn.

2. Using vpn technology to ensure information security [Electronic resource]. Access mode: https://cyberleninka.ru/article/n/ispolzovanie-tehnologii-vpn-dlya-obespecheniya-informatsionnoy-bezopasnosti.

3. For those who are under occupation: a selection of safe VPN services for using the Internet [Electronic resource]. Access mode: https://henichesk.city/articles/219181/dlya-tih-hto-v-okupacii-pidbirka-bezpechnih-vpn-servisiv-dlya-koristuvannya-internetom.

4. State Communications: What is a VPN and how to use it safely [Electronic resource]. Access mode: https://www.kmu.gov.ua/news/derzhspeczvyazku-shcho-take-vpn-i-yak-nim-bezpechno-koristuvatis.

5. VPN, Tor and communication in messengers. How to maintain communication in the occupied territories. Detailed instructions [Electronic resource]. Access mode: https://forbes.ua/innovations/vpn-tor-ta-spilkuvannya-u-messendzherakh-yak-pidtrimuvati-zvyazok-na-okupovanikh-teritoriyakh-detalna-instruktsiya-10082022-7616.

6. What is military-grade VPN encryption? [Electronic resource]. Access mode: https://zoogvpn.com/military-grade-vpn-encryption/.

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 02.12.2022

УДК 004.942
DOI: https://doi.org/10.53920/ITS-2022-2-4

Ольга Миколаївна ТКАЧЕНКО,
д-р техн., проф.,
Київський національний університет імені Тараса Шевченка
ORCID ID: 0000-0001-7983-9033
Владислав Олексійович СОСНОВИЙ,
асистент кафедри комп'ютерної інженерії
Державний університет телекомунікацій
ORCID ID: 0000-0002-3217-4537

МОДЕЛЬ ПРОГНОЗУВАННЯ БЕЗПЕКИ МЕРЕЖІ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

В статті розглянуто чотири алгоритми, а саме алгоритм SVM, алгоритм нечіткої кластеризації, алгоритм кластеризації K-Means і алгоритм Apriori. Деталізуємо 4 різних кроки безпеки користувачів мережі та їх контролю доступу статті є розробка надійної моделі прогнозування безпеки мережі. Розроблена модель виявлення вторгнень, побудована з використанням нейронних мереж. Модель виявлення вторгнень виявляє аномалії та атаки на основі зловживання. Модель виявлення вторгнень також виконує три типи завдань класифікації. Завдання включають класифікацію між появою атаки чи звичайним випадком, класифікацією між різними типами атаки чи звичайним випадком . Модель виявлення вторгнень також показує точність класифікації, час виконання та обсяг використання пам'яті. Цілями моделі виявлення вторгнень є висока точність, малий час виконання та мінімальний обсяг використання пам'яті. Модель виявлення вторгнень, побудована за допомогою нейронних мереж, відповідає цілям високої точності, малого часу виконання та мінімального використання пам'яті.

Ключові слова: *алгоритм SVM, алгоритм кластеризації K-Means, алгоритм Apriori, Алгоритм нечіткої кластеризації, Нейронна мережа.*

Olha TKACHENKO
Doctor of technical sciences, professor
Taras Shevchenko National University of Kyiv
ORCID ID: 0000-0001-7983-9033
Vladyslav SOSNOVYY
assistant of the department of computer engineering
State University of Telecommunications
ORCID ID: 0000-0002-3217-4537

NETWORK SECURITY PREDICTION MODEL USED BY NEURAL NETWORKS

The article discusses four algorithms, namely the SVM algorithm, the fuzzy clustering algorithm, the K-Means clustering algorithm, and the Apriori algorithm. We detail the 4 different steps of network user security and their access control. The article is the development of a reliable network security prediction model. An intrusion detection model built using neural networks has been developed. The intrusion detection model detects anomalies and abuse-based attacks. The intrusion detection model also performs three types of classification tasks. Tasks include classification between the occurrence of an attack or a normal case, classification between different types of attack or a normal case. The intrusion detection model also shows classification accuracy, execution time, and memory usage. The goals of the intrusion detection model are high accuracy, low execution time, and minimal memory usage. An intrusion detection model built using neural networks meets the goals of high accuracy, low execution time, and minimal memory usage.

In today's world, networks are becoming increasingly complex, interconnected, and widely used. Today, network traffic is growing almost exponentially. Networks also become more vulnerable to attack by hackers or anyone with malicious intent to disrupt network systems. Vulnerable networks are at risk of a blow to the economy and the destruction of confidential information. Thus, there is a need to improve network vulnerability detection mechanisms and improve network security prediction. The network security prediction model also aims to reduce memory consumption and improve the detection of different types of attacks in terms of timing and accuracy. In the network security prediction model, the memory consumption was low, and the time spent to detect attacks was also low. Attack detection accuracy is also high.

The above methods used to design the model are also easy to design. The above methods are also much more cost-effective since the use of neural networks is free. In addition, calculations are simplified by using this model. Therefore, using a neural network is also an effective way to develop a network security prediction model. Thus, the use of neural networks is recommended for the development of any type of network security prediction model. Future tasks are to develop models that will detect any intrusions even more accurately and quickly.

Keywords: *SVM algorithm, K-Means clustering algorithm, Apriori algorithm, Fuzzy clustering algorithm. Neural network.*

Постановка проблеми. У сучасному світі мережі стають все складнішими, взаємопов'язаними та широко використовуються. Сьогодні мережевий трафік зростає майже в геометричній прогресії. Мережі також стають більш вразливими до атак з боку хакерів або будь-кого зі злими намірами зруйнувати мережеві системи. Вразливі мережі знаходяться під загрозою удару по економіці та знищенні конфіденційної інформації. Таким чином, існує потреба у вдосконаленні механізмів виявлення вразливості мережі та покращенні прогнозування безпеки мереж. Модель прогнозування безпеки мережі також має на меті зменшити споживання пам'яті, а також покращити виявлення різних типів атак з точки зору часу та точності.

Аналіз останніх досліджень і публікацій. Було розглянуто моделі виявлення атак. Усі моделі мають спільну мету — точніше, ефективніше та швидше виявляти вразливості в мережі. Для досягнення мети були запропоновані різні алгоритми [1]. Запропонований метод виявлення вторгнень, повністю заснований на алгоритмі K-методі [2]. Був запропонований гібридний алгоритм виявлення вторгнень, заснований на K-методі та дереві вибору [3]. Для попередньої обробки 41- функції у статистичному наборі використовували моніторинг функцій даних [4].

Мета статті — опис створення файлів, які використовуються для виявлення аномальних атак. Описати деталі кількості атак або звичайних випадків для процесу виявлення аномалій або атак на основі неправильного використання. Описати спосіб видалення непотрібних або малокорисних непотрібних даних, щоб отримати оптимальну кількість даних для класифікацій. Використати нейронну мережу для виявлення різноманітних атак. Розрізнити різні

процеси класифікації, які відбуваються в атаках виявлення аномалій, атаках виявлення неправильного використання та окремих типів атак. Виявити аномалії за допомогою класифікації виникнення атаки або звичайного випадку.

Виклад основного матеріалу. Розглянемо чотири алгоритми, а саме алгоритм SVM, алгоритм нечіткої кластеризації, алгоритм кластеризації K- Means і алгоритм Apriori. Далі деталізуємо 4 різних кроки безпеки користувачів мережі та їх контролю доступу. Алгоритм SVM використовується для вирішення проблем класифікації. Базуючись на базовій конструкції статистичного принципу, до процесу обчислення додається функція ядра для відображення проблеми низької розмірності в просторі високої розмірності і отримує простір рішень високої розмірності. Це означає, що використання алгоритму SVM розблокує приховані шаблони у великій кількості даних, щоб виявити інформацію. Після завантаження інформації ,система може ідентифікувати часові ряди або тенденцію розвитку даних і робити точні висновки [5]. Було також використано метод автоматичного отримання найбільш оптимальних параметрів Гауса для отримання найкращої гіпосфери [6]. Була вдосконалена версія, у якій алгоритм K-методу змінено на комбінацію з Apriori для досягнення правильного значення виявлення Root to Learn і User to Root за інформацією БД KDDCUP99, встановленою на 98% і 79% [7]. Була запропонована ідея використання набору правил розмірності, змішаного з одно-класовою SVM [8].

Алгоритм нечіткої кластеризації виглядає наступним чином:

- Визначення функції подібності.
- Встановлення відповідної нечіткої матриці подібності відповідно до функції подібності.
- Обчислення нечіткого відношення та використання плоского методу. Також включає випередження при знаходженні транзитивного закриття.
- Класифікація відповідно до надзвичайних порогів і отримання специфічного динамічного ефекту кластеризації.
- Ступені групуються разом у набір серій.
- Алгоритм аналізу зразків використовується для виявлення нападу з можливої послідовністю нападу.
- Встановлення відповідної нечіткої матриці подібності, відповідно до функції подібності.

- Обчислення нечіткого відношення та використання плоского методу. Також включає випередження при знаходженні транзитивного закриття.
- Класифікація відповідно до надзвичайних порогів і отримання специфічного динамічного ефекту кластеризації.
- Ступені групуються разом у набір кандидатських серій.
- Алгоритм аналізу зразків використовується для виявлення режиму нападу з можливою послідовністю.

Новий підхід до генерації нечітких правил став пропозицією, в якій кластери в шаблоні навчання встановлюються відповідно до техніки кластеризації нечіткого С-методу, відповідно до характеристик кожного шаблону та кластера [9].

Алгоритм кластеризації K-means припускає, що необхідні значення кластеризації відомі, однак фактично в аналізі безпеки значення k зазвичай невідомі. І вибір початкового центру кластеризації К-ознак набору правил є важливим. Міні-серія К-ознак використовується для розділення звичайного набору даних і набору даних атаки на кластери однакового розміру окремо, а центр кожного кластера використовується як індекс кластера. Існує метод вибору репрезентативних екземплярів з кожного кластера. Репрезентативність елемента пов'язана як з щільністю, так і з відстанню. Вища репрезентативність збільшує ймовірність того, що елемент буде обраний як репрезентативний . Після відбору кожному репрезентативному елементу присвоюється вага. Цей крок не тільки зменшує розмір вихідних даних, але й зберігає максимальну кількість інформації [1].

Алгоритм апріорі використовується для аналізу внутрішніх асоціацій правил безпеки фактів, оскільки він має значущість високої якості. Проблемою цього алгоритму є часте сканування бази даних транзакцій і непомірний набір додаткових параметрів очікування [10, 11]. Значення мінімальної підтримки та мінімальної довіри мають величезний вплив на результати виявлення. Алгоритм Апріорі був вперше запропонований [10].

Необхідно переконатися, що конфіденційні дані не витікають до тих, хто має зловмисні наміри. Через це існують конкретні цілі контролю доступу для різних користувачів. Існує 4 таких основних ролі користувача. Вони є постачальниками даних, збирачами даних, майнерами даних і особами, які приймають рішення [12]. Крім того, вкрай важливо переконатися, що конфіденційні

факти не потрапили до тих, хто має злі наміри. Завдяки цьому існують цілі контролю доступу для надзвичайних ролей споживачів безпеки мережі. Існує 4 таких основних ролі користувача. Вони є постачальниками інформації, збирачами інформації, добувачами інформації та особами, які приймають рішення [12]. Для постачальників даних мета контролю доступу полягає в тому, щоб ефективно контролювати кількість конфіденційних даних, які розкриваються іншим. Для досягнення цієї мети можна використовувати інструменти захисту, щоб обмежити доступ інших до їхньої інформації, просувати дані на аукціоні, щоб отримати достатню компенсацію за втрату конфіденційності, або фальсифікувати інформацію, щоб приховати свою справжню особу Для збирачів даних мета контролю доступу полягає в тому, щоб запустити корисні факти для майнерів фактів, не розкриваючи особи постачальників записів і конфіденційну статистику щодо них. Для досягнення цієї мети необхідно розробити правильні моделі конфіденційності для кількісної оцінки можливої втрати контролю доступу під час виняткових атак, а також застосувати стратегії анонімізації статистики [12]. Для майнерів даних мета збереження конфіденційності полягає в тому, щоб отримати правильні результати видобутку записів, зберігаючи при цьому конфіденційну статистику нерозкритою ні в рамках методу видобутку записів, ні в результатах видобутку. Для досягнення цієї мети може бути обраний правильний підхід для регулювання інформації до виконання алгоритмів позитивного видобутку. Крім того, протоколи стабільного обчислення можуть бути використані для забезпечення безпеки приватних даних і конфіденційної статистики, що міститься в навченій моделі. Для осіб, які приймають рішення, мета контролю доступу полягає в тому, щоб зробити правильний висновок, наближаючи достовірність результатів аналізу фактів, які вони отримують. Щоб досягти цієї мети, можна використовувати методи походження, щоб підказати повернуті записи отриманих фактів або створити класифікатор [12].

Було створено дві системи одна для виявлення нападів на основі аномалій, а інша для виявлення нападів на основі зловживання. Ці системи мали приблизно 4500 записів. Вхідні дані поділялися на набір даних (75%) для навчання нейронної мережі та тестовий набір даних (25%) для навченої нейронної мережі. Першою метою методу було спрощення фактів для обробки. Спрощення

передбачає відмову від познак, які є менш корисними. Перевага полягає в тому, що позбавлення від ознак зменшує розміри даних, що обробляються для підвищення продуктивності нейронної мережі. Недоліком може бути той факт, що якщо ключові атрибути будуть випадково видалені, точність виявлення вторгнень знизиться. Усі постійні ознаки було видалено, а ознаки, які передають найбільший відсоток дисперсії, було видалено додатково. Сценарії «R» також перевіряли на задовільні характеристики на основі дисперсії всередині зразків. Не чіпали атрибути , які не вносять навіть 1 відсотка сукупної варіації в набір фактів. Для зношування виявлення вторгнень для атак, які переважно базуються на аномаліях, і атак, які переважно базуються на неправильному використанні, було створено два файли: набір даних аномалії та набір даних неправильного використання. У наборі інформації про виявлення аномалії клас або змінна передбачення були або нормальними, що представляло повсякденний випадок, або атаку. Набір фактів виявлення неправильного використання мав змінну категорії «Звичайна» або «Назва нападу», яка представляє певний вид нападу, наприклад Smurf, NMap, Rootkit тощо.

Очищення даних було досягнуто для файлів, що складаються з аномалії набору даних і неправильного використання набору даних. Використання Weka для отримання вибору атрибутів аномалії набору даних і вибору атрибутів неправильного використання набору даних означало набагато менші атрибути, які сприяли прискоренню NN. Пакет «neuralnet» доступний у R і має відкритий код. Його почали використовувати для IDS та аналізу на основі нейронної мережі. Пакетна угода надала можливості як для створення нейронної мережі, так і для проведення класифікації.

В цій роботі було розглянуто понад 4500 випадків атак. Було обрано 10 типів атак, включаючи атаки Neptune, NMap, PortSweep, Satan, Smurf, BufferOverflow, FTPWrite, GuessPassword, Back і Rootkit. В процесі виявлення атак було реалізовано виявлення вторгнень на основі аномалій. Для виявлення атак реалізовано виявлення вторгнень на основі зловживання, щоб запропонувати матрицю плутанини, точність класифікації, час, витрачений на впровадження, і споживання ресурсів. Було створено класифікацію серед 10 нападів і звичайного випадку. Для деяких атак система виконувала виявлення зловживань.

Атака з виявленням аномалій У наборі результатів, наведених у таблиці 1, є деталі точності під час виявлення атаки з аномаліями, час виконання та обсяг споживання пам'яті [13]. Існує класифікація виникнення атаки або звичайного випадку в цьому процесі. Значення, зазначені в таблиці 1, вказують на кількість записів. Значення координат (Attack, Attack) і значення (Attack, Normal) координати дорівнюють 389 і 6 відповідно. Значення координати (Атака, Атака) вище, ніж координата (Атака, Норм), як показано в таблиці 1. Це означає наявність атаки. Точність класифікації висока – 99,57 відсотка. Мінімальний час виконання становить 3,9979 секунди. Використання пам'яті є мінімальним і становить 2191,311 Kbit, як показано в таблиці 2.

Таблиця 1. Виявлення аномальних атак

	Axis2	
Axis1	Anomaly attack	Normal
Anomaly Attack	389	6
Normal	3	763

Таблиця 2. Результати виявлених атак

Точність	99,57%
Час виконання	3.9979 с
Використання пам'яті	2189.311 Kbs

Атака виявлення неправильного використання В наборі результатів, показаних у таблиці 3, є деталі точності виявлення атаки неправильного використання, час виконання та обсяг споживання пам'яті [13]. Існує класифікація між 10 типами атак і нормальними випадками. У таблиці 3, як у випадку виявлення аномалії, є 2 осі, тобто вісь 1 і вісь 2. Значення, згадані в таблиці 3, вказують на кількість записів. Значення (Назад, Назад), (Buffer Overflow, Buffer Overflow), (Guess Password, Guess Password), (Neptune, Neptune), (Nap, NMap), (Port Sweep, Port Sweep), (Satan, Satan) і (Smurf, Smurf) координати є найвищими серед усіх значень рядка (значення 67, 4, 12, 57, 75, 748, 63,

59 і 58 відповідно). Це вказує на легшу класифікацію та більшу ймовірність нападу. Координати (FTP Write, FTP Write), (Rootkit, Rootkit) значно низькі (значення 1 і 0 відповідно). Значення координат (FTP Write, Normal) і (Rootkit, Normal) дорівнюють 0 і 1 відповідно, що не є найвищим значенням рядка. Таким чином, нормальний випадок, як показано в таблиці 3, відсутній. Як показано в таблиці 4, точність класифікації висока і становить 98,1%. Час виконання 48,9282 секунди вищий, ніж у попередньому випадку, через більший обсяг інформації, однак все ще низький для обробленої інформації. Використання пам'яті 2988,14 Кб, збільшене завдяки більшій інформації, але не дуже високе.

Таблиця 3. Атака з виявленням неправомірного використання — відомості про записи

Axis1 / Axis2	Повернення	Bufer Overflow	FTP	Guess Pass.	Neptune	NMap	Normal	Port Sweep	Rootkit	Satan	Smurf
Повернення	67	0	0	0	0	0	0	0	0	0	0
Bufer Overflow	0	4	0	0	0	1	1	0	0	0	0
FTP	0	0	1	1	0	0	0	0	0	0	0
Guess Password	1	0	0	12	0	1	0	0	0	0	0
Neptune	0	0	0	0	57	0	0	1	0	0	3
NMap	0	0	0	0	0	75	0	0	0	0	0
Normal	0	0	0	0	0	0	748	0	1	0	0
Port Sweep	0	0	0	0	0	0	0	63	0	0	1
Rootkit	0	0	1	0	3	0	1	3	0	0	1
Satan	0	0	0	0	0	3	0	1	1	59	1
Smurf	0	0	0	0	0	0	0	0	0	0	58

Таблиця 4. Результати атаки виявлення неправильного використання

Точність	98.10%
Час використання	48.9282 с
Використання пам`яті	2988.14 Kbs

Висновки та пропозиції. У моделі прогнозування безпеки мережі споживання пам'яті було низьким, час, витрачений на виявлення атак, також низький. Також висока точність виявлення атак. Вищевказані методи, які використовуються для проектування моделі, також прості в проектуванні. Крім того, обчислення спрощуються завдяки використанню цієї моделі. Отже, використання нейронної мережі також є ефективним способом розробки моделі прогнозування безпеки мережі. Таким чином, використання нейронних мереж рекомендовано для розробки будь-якого типу моделі прогнозування безпеки мережі. Майбутні завдання полягають в розробці моделей, які виявлятимуть будь-які вторгнення ще точніше та швидше.

ЛІТЕРАТУРА

1. Qiuhua W Xiaoqin O Jiacheng Z 2019 A Classification algorithm based on data clustering and data reduction for intrusion detection system over big data KSII Trans. on internet and information systems 13 pp 3714-3732.

2. Jianliang M Haikun S Ling B 2009 The application on intrusion detection based on K-means cluster algorithm Int. Forum on Information Technology and Applications (Chengdu) pp. 150-152.

3. Aung Yi Y Myat M M 2018 Hybrid intrusion detection system using K-Means and classification and regression trees algorithms IEEE 16th Int. C. on Software Engineering Research, Management and Applications (SERA) (Kunming) pp 195-199.

4. Ravale U Marathe N Padiya P 2015 Feature selection based hybrid anomaly intrusion detection system using K-Means and RBF kernel function Procedia Computer Science pp 428-435.

5. Xiaoyi H 2020 Network security situation prediction based on grey relational analysis and support vector machine Algorithm Int. J. of network security 22 pp 177-182.

6. Xiao Y Wang H Xu W 2015 Parameter selection of gaussian kernel for one-class SVM IEEE Trans. on Cybernetics 45 pp 927-939.

7. Song C Ma K 2009 Design of intrusion detection system based on data mining algorithm Int. Conf. on signal processing systems (Singapore) pp 370-373.

8. Rochim A F Aziz M A Fauzi A 2019 Design log management system of computer network devices infrastructures based on ELK stack Int. C. on Electrical Engineering and Computer Science (ICECOS) (Batam Island) pp 338-342.

9. Jin C Ye Z Wang C, Yan L Wang R 2018 A network intrusion detection method based on hybrid rice optimization algorithm improved fuzzy C-means IEEE 4th Int. Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS) (Lviv) pp 47-52.

10. Han J W Kamber M P 2011 Data mining: concepts and techniques (3rd Edition. Elsevier Science).

11. Huang Z 2019 Research and implementation of intrusion detection based on host log". Proc. of the Int. Conf. on Big Data Engineering, pp 98–106.

12. Lei X Chunxiao J Jian W Jian Y, Yong R 2014 Information security in big data: privacy and data mining IEEE Access 2 pp 1149-1176.

13. Pratik M, Saunil D, Ravina D 2015 Intelligent Security Systems-Intrusion Detection System https://github.com/jgera/Network-Intrusion-Detection-System/blob/master/report.pdf.

REFERENCES

1. Qiuhua W Xiaoqin O Jiacheng Z 2019 A Classification algorithm based on data clustering and data reduction for intrusion detection system over big data KSII Trans. on internet and information systems 13 pp 3714-3732.

2. Jianliang M Haikun S Ling B 2009 The application on intrusion detection based on K-means cluster algorithm Int. Forum on Information Technology and Applications (Chengdu) pp. 150-152.

3. Aung Yi Y Myat M M 2018 Hybrid intrusion detection system using K-Means and classification and regression trees algorithms

IEEE 16th Int. C. on Software Engineering Research, Management and Applications (SERA) (Kunming) pp 195-199.

4. Ravale U Marathe N Padiya P 2015 Feature selection based hybrid anomaly intrusion detection system using K-Means and RBF kernel function Procedia Computer Science pp 428-435.

5. Xiaoyi H 2020 Network security situation prediction based on grey relational analysis and support vector machine Algorithm Int. J. of network security 22 pp 177-182.

6. Xiao Y Wang H Xu W 2015 Parameter selection of gaussian kernel for one-class SVM IEEE Trans. on Cybernetics 45 pp 927-939.

7. Song C Ma K 2009 Design of intrusion detection system based on data mining algorithm Int. Conf. on signal processing systems (Singapore) pp 370-373.

8. Rochim A F Aziz M A Fauzi A 2019 Design log management system of computer network devices infrastructures based on ELK stack Int. C. on Electrical Engineering and Computer Science (ICECOS) (Batam Island) pp 338-342.

9. Jin C Ye Z Wang C, Yan L Wang R 2018 A network intrusion detection method based on hybrid rice optimization algorithm improved fuzzy C-means IEEE 4th Int. Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS) (Lviv) pp 47-52.

10. Han J W Kamber M P 2011 Data mining: concepts and techniques (3rd Edition. Elsevier Science).

11. Huang Z 2019 Research and implementation of intrusion detection based on host log". Proc. of the Int. Conf. on Big Data Engineering, pp 98–106.

12. Lei X Chunxiao J Jian W Jian Y, Yong R 2014 Information security in big data: privacy and data mining IEEE Access 2 pp 1149-1176.

13. Pratik M, Saunil D, Ravina D 2015 Intelligent Security Systems-Intrusion Detection System https://github.com/jgera/Network-Intrusion-Detection-System/blob/master/report.pdf.

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 07.12.2022

УДК 004.021
DOI: https://doi.org/10.53920/ITS-2022-2-5

Сергій Станіславович КОРОТКОВ,
асистент кафедри комп'ютерної інженерії
Державний університет телекомунікацій
ORCID ID: 0000-0002-4090-5934
Владислав Сергійович ЗАСАДЮК,
студент
Державний університет телекомунікацій
ORCID ID: 0000-0002-8088-4277

АЛГОРИТМ ЗМЕНШЕННЯ ТРАНЗИТНИХ ПОТОКІВ ТРАНСПОРТНОЇ МЕРЕЖІ У ЗАДАНОМУ НАПРЯМКУ

В статті розроблено алгоритм зменшення транзитних потоків. В результаті виконання алгоритму величина потоку на будь-якому розрізі мережі буде максимальною, а сумарний потік буде складатися зі зменшених транзитних потоків, що дозволить підвищити ефективність використання транспортних засобів. Потоки розподілені таким чином, що після застосування до них принципу суперпозиції величини сумарного результуючого потоку на дугах не перевищують їх пропускних здібностей. Розглядається знаходження інтегрального максимального потоку транспортної мережі у заданому напрямку.

Ключові слова: *потоки машин, теорема Форда-Фалкерсона, максимальний потік, алгоритм зменшення транзитних потоків.*

Serhii KOROTKOV
Assistant of the department of computer engineering
State University of Telecommunications
ORCID ID: 0000-0002-4090-5934
Vladyslav ZASADIUK
Student
State University of Telecommunications
ORCID ID: 0000-0002-8088-4277

ALGORITHM FOR CHANGING TRANSIT FLOWS IN THE TRANSPORT METER AT A GIVEN DIRECTLY

An algorithm for reducing transit flows has been developed. As a result of the execution of the algorithm, the amount of flow on any section of the network will be maximum, and the total flow will consist

of reduced transit flows, which will increase the efficiency of the use of vehicles. The flows are distributed in such a way that after applying the principle of superposition to them, the values of the total resulting flow on the arcs do not exceed their throughput capacities. Finding the integral maximum flow of the transport network in a given direction is considered. The algorithm consists of:

1. Finding the integral maximum flow of the transport network in a given direction (ZV) for each time interval, the matrices of the values of the maximum flows and the efficiencies of these flows are composed, the elements of which are the values of the maximum flows and the efficiencies of the flows for each combination of input and output.

2. Compilation of a list of flows, in which all elements are ranked from «worst» in terms of efficiency to «best» according to the data of the unit square and the reachability matrix.

3. Rejecting flows and obtaining a set of the most effective flows, forming a matrix of integral transit flow in the selected direction.

4. Compilation of the total set of arcs, which shows how much the integral flow of this arc exceeds its carrying capacity.

5. From the set of saturated arcs, the element with the maximum value of n1 is selected.

6. In each of the transit flows of the set where this selected arc meets, it is necessary to reduce the flow by n1 times so that after performing the superposition of the arc flows.

7. After that, the sets are rearranged.

The described actions are performed as long as there is at least one element in the set with n greater than zero.

Key words: *machine flows, Ford-Falkerson theorem, maximum flow, transit flow reduction algorithm.*

Постановка проблеми. Для того щоб раціонально організувати рух транспортних потоків необхідно оцінити максимальний потік у мережі, знайти найефективніше розподіл потоку, виявити вузькі місця та своєчасно їх ліквідувати. Одночасно з цими завданнями слід оцінити сумарні витрати транспортних засобів під час їх руху з початкового пункту в кінцевий.

Аналіз останніх досліджень і публікацій. Алгоритми: побудову, аналіз, застосування імітаційного моделювання та дослідження динаміки транспортних потоків досліджували: Жогаль С.І., Максим І.В., Поляков К.Ю., Соболь І.М., Новіков Ф.А., Сукач О.І.

Мета статті – висвітлити принцип роботи розробленого алгоритму зменшення транзитних потоків.

Виклад основного матеріалу. Для знаходження інтегрального максимального потоку транспортної мережі заданому напрямку (ZV) для кожного часового інтервалу складаються матриці величин максимальних потоків та ефективностей цих потоків, елементами яких є значення максимальних потоків та ефективностей потоків по кожному поєднання входу та виходу. Матрицю величин потоків позначимо $\varphi = \|\varphi_{ij}^{max}\|$, i=$\overline{1,m}$, j=$\overline{1,n}$. Матрицю ефективностей потоків позначимо $\varphi=\|\varphi_{ij}\|$, i=$\overline{1,m}$, j=$\overline{1,n}$. Отримані матриці нормуються за максимальним елементом:

$$\varphi^* = \|\varphi_{ij}^*\| = \left\| \frac{\varphi_{ij}}{\max ij\ \varphi_{ij}} \right\|, \quad \Phi^* = \|\Phi_{ij}^*\| = \left\| \frac{\Phi_{ij}}{\max ij\ \Phi_{ij}} \right\|, \ i=\overline{1,m},\ j=\overline{1,n}.$$

В результаті всі елементи матриць φ^*, Φ^* задовольняють нерівності $0 < \varphi_{ij}^* < 1$, $0 < \Phi_{ij}^* < 1$. У прямокутній системі координат $\varphi^* 0 \Phi^*$ відмічаємо точки $(\varphi_{ij}^*, \Phi_{ij}^*)$. Причому, через те, що елементи матриць нормовані за максимальним елементом, всі точки будуть у межах одиничного квадрату, лівий нижній кут якого суміщений з початком координат.

Усі точки одиничного квадрату порівнюються за допомогою деякої метрики. Складається список потоків, в якому всі елементи ранжуються від «найгіршого» за ефективністю до «найкращого» відповідно до даних одиничного квадрата [1].

Одночасно складається матриця досяжності $D = \|d_{ij}\|$ i=$\overline{1,m}$, j=$\overline{1,n}$, де $d_{ij} = 1$, якщо є потік з i-того входу в j-тий вихід, і $d_{ij} = 0$, якщо немає потоку з i-того входу в j-тий вихід, і $d_{ij} = 0$, якщо немає потоку з i-того входу в j-тий вихід.

Список проглядається, починаючи з потоку, найгіршого за ефективністю. Поточний потік виключається зі списку, якщо його виняток не залишає жодну початкову вершину без вихідного потоку і залишає жодну кінцеву вершину без вхідного потоку. При виключенні зі списку потоку, який йде з i-того входу в і - вихід, модифікується матриця досяжності D, в якій на перетині цього рядка і і-того стовпця одиниця замінюється нулем.

Таким чином, поточний потік з i-того входу в j-тий вихід виключається зі списку в тому випадку, якщо після його виключення та модифікації матриці досяжності D у її i-тому рядку буде хоча б одна одиниця і в j-му стовпці також буде хоч би одна одиниця.

Якщо ж виняток потоку веде до того що, що у матриці досяжності i-тий стовпець чи j-та рядок складатиметься з нулів, то потік зі списку потоків не виключається, й у списку переходимо наступного потоку.

В результаті відбраковування потоків отримуємо безліч найбільш ефективних потоків $ZV^e = \{Z_i,V_j\}$, таких, що кожен із входів пов'язаний транзитним потоком, принаймні, з одним виходом. Тобто кожен із входів транспортної мережі має, принаймні, один потік, що виходить з нього, а кожен з виходів мережі має хоча б один вхідний до нього потік [2]. Для безлічі таких потоків ZV^e застосовується принцип суперпозиції, коли відповідні їм матриці розподілу потоків $X^{ij} = \|\varphi_{kl}^{ij}\|$ підсумовуються, утворюючи матрицю інтегрального транзитного потоку за вибраним напрямом $\|\sum X\|$. Причому, завдання суперпозиції потоків вирішується таким чином, щоб в мережі могли одночасно існувати всі потоки, що залишилися із множини ZV^e.

З цією метою складається загальна множина дуг мережі $DN = \{<d_{ij},n>\}$. Елементами цієї множини є пари $<d_{ij},n>$, що складаються з вказівника насиченої дуги d_{ij} з i-того вузла в j-тий , і числа n , що показує, наскільки інтегральний потік цієї дуги перевищує її пропускну здатність. З безлічі насичених дуг DN вибирається елемент, у якого величина n1 максимальна. Цей елемент списку $<d_{ij},n1>$ описує дугу, для якої величина n1 сумарного потоку, побудованого з транзитних потоків, що залишилися, найбільше перевищує пропускну здатність дуги. Тому в кожному з транзитних потоків множини DN, де зустрічається ця обрана дуга, необхідно зменшити потік в n1 раз для того, щоб після виконання суперпозиції потоків дуги, що залишилися d_{ij} виявилася насиченою. Зменшення проводиться для кожного потоку з множини $\{Z_iV_j\}$ по всіх шляхах, які насичували дугу, що розглядається, в ході виконання алгоритму Форда-Фалкерсона, пропорційно їх вкладу в насичення дуги. Після цього множини DN перебудовуються через те, що було зроблено зменшення кожного з транзитних потоків, що спричинило зміну величин n для гілок мережі, які були задіяні при зменшенні потоку [3].

Описані дії виконуються до тих пір, поки в множині DN є хоча б один елемент, у якого n більше нуля. Як тільки у всіх елементів множини DN величини n стануть негативними або рівними нулю, процес зменшення транзитних потоків закінчується і як рішення

задачі знаходиться результуючий інтегральний потік. При цьому потоки розподілені таким чином, що після застосування до них принципу суперпозиції величини сумарного результуючого потоку на дугах не перевищують їх пропускних здібностей.

В результаті виконання алгоритму суперпозиції потоків відповідно до теореми Форда-Фалкерсона величина потоку на будь-якому розрізі мережі буде максимальною, а сумарний потік буде складатися зі зменшених транзитних потоків [4].

Для ілюстрації роботи алгоритму знаходження інтегрального потоку розглянемо ділянку транспортної мережі.

Множина входів у мережі представлено вершинами 1 і 2. Множина виходів задається вершинами 9 і 11. У мережі розглядаються такі транзитні потоки: (1,...,9), (1,...,11), (2,...,9) і (2,...,11).

Для кожного із зазначених транзитів вирішується завдання про максимальний потік окремо, в результаті чого отримуємо чотири матриці максимальних потоків для цих транзитів (φ max1_9, φ max1_11, φ max2_9, φ max2_11), відповідно чотири матриці ефективностей (Ф1_9, Ф1_11, Ф2_9, Ф2_11), а також відповідні розподіли величин потоків по гілках мережі (X1_9, X1_11, X2_9, X2_11).

Для формування списку найбільш ефективних потоків відкидаємо найменш ефективні потоки таким чином, щоб не залишити жоден вхід хоча б без одного потоку, що виходить, і жоден вихід хоча б без одного вхідного потоку. Для прикладу, що розглядається, найбільш ефективними потоками виявилися потоки 1→9 з величиною потоку: 41, 1→11 з величиною потоку: 43 і 2→9 з величиною потоку: 34.

Підсумовуючи матриці максимальних потоків |X1_9|, |X1_11| та |X2_9|, отримуємо матрицю інтегрального транзитного потоку |∑X|. Шляхом поелементного віднімання матриці |∑X| з матриці пропускних здібностей мережі |C| отримуємо матрицю |C – ∑X|.

Негативні значення елементів матриці свідчать про недостатню пропускну здатність відповідних гілок мережі в разі руху через мережу одночасно всіх залишених потоків [5]. Вибираємо гілку, сумарний потік на якій найбільше перевищує її пропускну здатність. Для прикладу ця гілка (3,6) - найменший елемент матриці |C – ∑X| .

Вибираємо всі шляхи, які насичували гілку (3,6) у ході розв›язання задачі про максимальний потік для кожного з транзитних напрямків та величини Δ, на які збільшувався потік цими шляхами.

Для транзитного напрямку 1→9 це будуть шляхи:
(1,3) → (3,6) → (6,9), Δ =16;
(1,3) → (3,6) → (6,10) → (10,9), Δ = 4.
Для транзитного направлення 1 → 11:
(1,3) → (3,6) → (6,10) → (10,11), Δ = 20.

У транзитному напрямку 2→9 жоден із шляхів у ході вирішення задачі про максимальний потік не проходив через вершину (3,6). Для того, щоб сумарний потік зміг пройти через гілку (3,6), зменшуємо потоки транзитних напрямів обраними шляхами так, щоб весь сумарний потік зменшився на 20 одиниць, причому по кожному шляху потік зменшується пропорційно величині Δ. Для транзитного напрямку 1→9 потік шляхом 1 зменшуємо на 8 одиниць, а потік шляхом 2 зменшуємо на 2 одиниці. Для транзитного напрямку 1→11 потік шляхом 1 зменшуємо на 10 одиниць. З цією метою віднімаємо від елементів матриць |X1_9|, |X1_11| відповідних вузлам шляхів величини, на які зменшується потік по дорозі [6].

Далі перераховуються значення елементів матриці |C–∑X|. В результаті одержуємо нову матрицю, у якої елемент (3,6) дорівнює 0, так як потік через гілку мережі (3,6) було зменшено.

Процес зменшення аналізованих потоків повторюється до того часу, поки у матриці |C–∑X| залишатимуться негативні елементи. Коли всі елементи матриці |C–∑X| виявляться позитивними, це значить, що пропускних здібностей гілок мережі достатньо для того, щоб усі зменшені транзитні потоки змогли існувати в мережі одночасно [7]. У цьому випадку алгоритм зменшення транзитних потоків закінчується і розв'язанням задачі про максимальний потік буде сумарна матриця |∑X|.

Величина потоку після зменшення транзиту 1→9 склала 24, транзиту 1→11 – 25, транзиту 2→9 – 15. Сумарна величина трьох потоків становила 64.

Висновоки та пропозиції. В ході написання статті був виконаний аналіз робіт по знаходженню максимального потоку, а також огляд аналітичних моделей дослідження операцій і теорій автоматичного керування. В результаті був розроблений алгоритм суперпозиції потоків у відповідності до теореми Форда-Фалкерсона, який дозволяє підвищити ефективність використання транспортних засобів.

ЛІТЕРАТУРА

1. Томас Х. Кормен. Алгоритми: побудова та аналіз. - 2-ге вид. - М.: «Вільямс», 2006. - 1296с.

2. Жогаль С.І., Максим І.В. Завдання та моделі дослідження операцій. Навчальний посібник. – Гомель: БелДУТ, 1999. – 4.1: Аналітичні моделі дослідження операцій. – С 109.

3. Поляков К.Ю. Теорія автоматичного керування. Санкт-Петербург, 2008. С. 4-20.

4. Соболь І.М. Метод Монте-Карло. Москва: Наука, 1968. 64 з.

5. Новіков Ф.А. Дискретна математика програмістів. - 3-тє. - СПб.: Пітер, 2008. - С. 277-279. - 384с.

6. Сукач О.І. Застосування імітаційного моделювання на дослідження динаміки транспортних потоків регіону. Вісті Гомельського державного університету імені Ф. Скорини. – 2006. – № 4 (37). – С. 96-99.

7. Зайченко Ю.П. Дослідження операцій: Навчальний посібник. Київ: Видавничий дім «Слово», 2002. – 320 с.

REFERENCES

1. Thomas H. Corman. Algorithms: construction and analysis. - 2nd edition. - M.: «Williams», 2006. - 1296p.

2. Zhogal S.I., Maksym I.V. Tasks and models of operations research. Tutorial. - Gomel: BelDUT, 1999. - 4.1: Analytical models of operations research. - C 109.

3. Polyakov K.Yu. Theory of automatic control. St. Petersburg, 2008. P. 4-20.

4. I.M. Sobol Monte Carlo method. Moscow: Nauka, 1968. 64 p.

5. Novikov F.A. Discrete mathematics of programmers. - the 3rd. - St. Petersburg: Peter, 2008. - P. 277-279. - 384 p.

6. Sukach O.I. The application of simulation modeling to the study of the dynamics of transport flows in the region. News of Gomel State University named after F. Skoryna. – 2006. – No. 4 (37). - pp. 96-99.

7. Zaichenko Yu.P. Operations Research: A Study Guide. Kyiv: «Slovo» Publishing House, 2002. - 320 p.

СТАТТЯ НАДІЙШЛА ДО РЕДАКЦІЇ 07.12.2022

Науковий журнал

«IT SYNERGY»

Заклад вищої освіти
«Міжнародний науково-технічний університет
імені академіка Юрія Бугая»
Вип. 2 (3), Київ, 2022. – 62 с.

Відповідальний за випуск О. І. Бражнікова

Підписано до друку 27.12.2022.
Формат 60x90/16. Ум. друк. арк. – 3.56. Тираж 100. Зам. № 170.

Друк: Видавництво ТОВ «А-ЦЕНТР».
Свідоцтво про реєстрацію Серія ДК № 599 від 14.01.2001 р.
04112, м. Київ, вул. Івана Гонти, 3А.