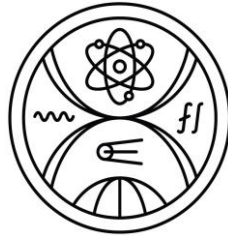


UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY
A INFORMATIKY



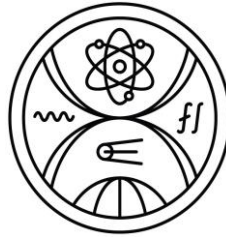
TELEGRAM - ANALÝZA SOCIÁLNEJ SIETE

Bakalárska práca

2023

Jozef Kolek

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY
A INFORMATIKY



TELEGRAM - ANALÝZA SOCIÁLNEJ SIETE

Bakalárska práca

Študijný program: Aplikovaná informatika

Študijný odbor: 2511 Aplikovaná informatika

Školiace pracovisko: Katedra aplikovanej informatiky

Školiteľ: RNDr. Damas Gruska, PhD.

Bratislava, 2023

Jozef Kolek



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Jozef Kolek
Študijný program: aplikovaná informatika (Jednoodborové štúdium,
bakalársky I. st., denná forma)
Študijný odbor: Informatika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický
Názov: Telegram - analýza sociálnej siete
Telegram - social network analysis
Anotácia: Práca sa sústreďí na analýzu správania sa používateľov sociálnej siete
Telegram. Skúma, ako z verejne dostupných informácií, možno získať údaje
o zvykoch, pohybe a pod. jej používateľov.
Vedúci: doc. RNDr. Damas Gruska, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.
Dátum zadania: 06.06.2022
Dátum schválenia: 14.12.2022 doc. RNDr. Damas Gruska, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Čestné vyhlásenie

Vyhlasujem, že som vypracoval bakalársku prácu samostatne na základe svojich vedomostí a skúseností a s použitím uvedenej odbornej literatúry.

V Bratislave dňa.....

.....

vlastnoručný podpis

Pod'akovanie

Chcem pod'akovať môjmu školiteľovi doc. RNDr. Damasovi Gruskovi, PhD. za jeho cenné rady a usmernenia, ktoré mi pomohli napísať túto bakalársku prácu.

Abstrakt

Bakalárska práca popisuje históriu sociálnych sietí, ich vývoj a súčasnosť. Poukazuje na bezpečnostné riziká vyplývajúce z publikácie informácií na sociálnych sieťach, typy útokov a odporúča metódy na elimináciu hrozieb. V praktickej časti popisuje metódy a nástroje použité pri získavaní informácií zo sociálnej siete Instagram a Telegram. V analytickej časti porovnáva sledovanosť Telegram skupín podľa jazyka, popularitu vybraných osobností publikujúcich na Telegrame a vplyv spoločenských udalostí na frekvenciu publikácie a sledovanosť príspevkov. Z dostupných telefónnych čísel prináša príslušnosť používateľov podľa krajín.

Kľúčové slová: sociálna sieť, hashtag, Telegram, bezpečnosť citlivých údajov, bot

Abstract

The bachelor's thesis describes the history of social networks, their development and the present. It points out security risks arising from the publication of information on social networks, types of attacks and recommends methods for eliminating threats. In the practical part, it describes the methods and tools used in obtaining information from the Instagram and Telegram social networks. In the analytical part, it compares the number of Telegram groups by language, the popularity of selected personalities publishing on Telegram and the impact of social events on the frequency of publication and the number of posts. From the available phone numbers, it brings the affiliation of users by country.

Keywords: social network, hashtag, Telegram, security of sensitive data, bot

OBSAH

1	Úvod.....	12
2	Sociálne siete.....	13
2.1	Sociálne siete dnešného sveta	14
2.1.1	SixDegrees.com	14
2.1.2	LiveJournal	15
2.1.3	Flickr.....	15
2.1.4	Facebook.....	15
2.1.5	LinkedIn.....	16
2.1.6	TikTok.....	16
2.1.7	Instagram	17
2.1.7.1	Hashtag.....	18
2.1.7.2	Instagramový profil.....	18
3	Telegram.....	19
3.1	Základné pojmy používané na Telegrame	19
3.1.1	Telegram Account – Profil.....	19
3.1.2	Verified account – Overený profil	20
3.1.3	Channels	20
3.1.4	Skupina	20
3.1.5	Bot.....	21
3.1.6	API.....	21
4	Teória grafov	22
4.1	Základné pojmy	22
4.2	Graf.....	22
4.3	Typy grafov.....	23
4.3.1	Orientovaný graf.....	23

4.3.2	Neorientovaný graf	23
4.3.3	Reprezentácia údajov - stĺpcový graf	24
4.4	Komplexné siete	24
4.4.1	Fenomén malého sveta	25
4.4.2	Bezškálové siete.....	26
4.4.3	Prirodzené siete.....	27
5	Bezpečnosť údajov	28
5.1	Nebezpečenstvá	28
5.2	Phishing	28
5.2.1	Klasický phishing	28
5.2.2	Spear phishing	29
5.3	Cookies	29
5.4	Bezpečnosť údajov.....	30
5.5	TikTok a bezpečnosť údajov.....	30
5.6	Bezpečnosť Európska únia a svet	30
5.7	Telegram a bezpečnosť	31
5.8	Ochrana pred nebezpečenstvami	32
6	Použité technológie a ich využitie.....	33
6.1	Python	33
6.2	Telethon.....	33
6.2.1	Získavanie členov skupín	33
6.2.2	Prehľadávanie konverzácií.....	34
6.2.3	Vyhľadávanie kanálov, skupín a užívateľov	35
6.2.4	Posielanie správ	35
6.2.5	Sťahovanie súborov z konverzácií.....	36
6.3	Pyvis.....	36
6.4	Phonenumbers.....	37

6.5	Combot.....	37
6.6	TGStat.....	37
6.7	TGScanRobot.....	37
6.8	Instagrapl	37
6.9	Selenium	38
7	Praktická časť.....	40
7.1	Analýza odberateľov jazyčných skupín v Telegrame	40
7.1.1	Ruské verzus Ukrajinské skupiny.....	40
7.1.2	Uzbekistan verzus Rusko a Ukrajina	44
7.1.3	Tadžikistan verzus Rusko a Ukrajina.....	44
7.1.3.1	Porovnanie s ruskými skupinami	44
7.1.3.2	Porovnanie s ukrajinskými skupinami	47
7.1.4	Kirgizsko verzus Rusko a Ukrajina	47
7.1.4.1	Porovnanie s ruskými skupinami	47
7.1.4.2	Porovnanie s ukrajinskými skupinami	47
7.1.5	Kazachstan verzus Rusko a Ukrajina	51
7.1.6	Vyhodnotenie	52
7.2	Analýza sledovanosti dvoch kanálov.....	52
7.2.1	Analýza sledovanosti príspevkov Dmitriho Medvedeva.....	52
7.2.2	Analýza sledovanosti príspevkov Volodymyra Zelenského	53
7.3	Frekvencia uverejňovania správ	53
7.3.1	Frekvencia uverejňovania Rian_ru	54
7.3.2	Frekvencia uverejňovania U_now	55
7.3.3	Zhodnotenie	56
7.4	Účasť botov v skupinách	56
7.5	Viditeľnosť telefónneho čísla.....	57
7.5.1	Národnosť užívateľov	57

7.6	Komunikácia ruských vojakov na Telegrame	58
7.7	Scraping dát na Instagrame	59
8	Záver.....	60

1 Úvod

Už sú za nami časy, keď sme používali internet len na posielanie mailov a vyhľadávanie informácií. V dnešnej dobe sa internet vďaka sociálnym sieťam stal komunikačnou platformou, kde ľudia nielen konzumujú obsah, ale ho aj aktívne vytvárajú a zdieľajú. Prispievajú svojimi myšlienkami, komentármi, fotografiami videami a dokonca podnikajú vo virtuálnom priestore. Vytvárajú si svoj okruh priateľov - komunikačné skupiny podľa profesie alebo záujmov. S intenzitou komunikácie narastá aj množstvo citlivých informácií publikovaných v sociálnych sieťach.

V druhej kapitole objasníme pojem sociálna sieť, jej vývin od začiatku po súčasnosť a taktiež v krátkosti popisujeme niektoré sociálne siete.

V tretej kapitole sa venujeme sociálnej sieti Telegram. Vysvetlíme si pojmy ako Telegram skupina, Telegram kanál a iné.

V štvrtej kapitole si vysvetlíme základy z teórie grafov. Zdefinujeme si pojmy orientovaný a neorientovaný graf. Skúmať budeme aj siete a ich vlastnosti.

V piatej kapitole sa zaoberáme bezpečnosťou údajov na internete. Riešime zneužitie dát vo všeobecnosti, súbory cookies alebo aj ochranu údajov v Európskej únii.

V šiestej kapitole uvedieme technológie, ktoré sme pri získavaní a vyhodnocovaní dát použili. Je to napríklad nástroj Pyvis na tvorbu grafov alebo Telethon na získavanie dát zo sociálnej siete Telegram.

V siedmej kapitole hľadáme spoločných odoberateľov Telegramových skupín, v ktorých sa komunikuje konkrétnymi jazykmi. Porovnáваме ukrajinské, ruské jazyčné skupiny s kazašskými, uzbeckými a inými jazyčnými skupinami. Ďalej analyzujeme sledovanosť dvoch Telegramových kanálov, ktoré patria dvom názorovo polárnym osobnostiam s približne rovnakým počtom odoberateľov. Sledujeme aj správanie ukrajinského a ruského média a ich reakciu na určité udalosti. Na záver uvádzame ďalšie analýzy ako účasť botov v skupinách alebo národnosť užívateľov na základe ich zisteného telefónneho čísla.

2 Sociálne siete

Sociálne siete predstavujú množstvo serverov umiestnených v dátových centrách po celom svete, vzájomne poprepájaných redundantnými vysokorychlostnými linkami, sú to ďalej webové stránky, počítačové a mobilné aplikácie a všetci používatelia týchto platforiem, ktoré umožňujú ľuďom vytvárať, zdieľať a šíriť informácie vo virtuálnych komunitách.

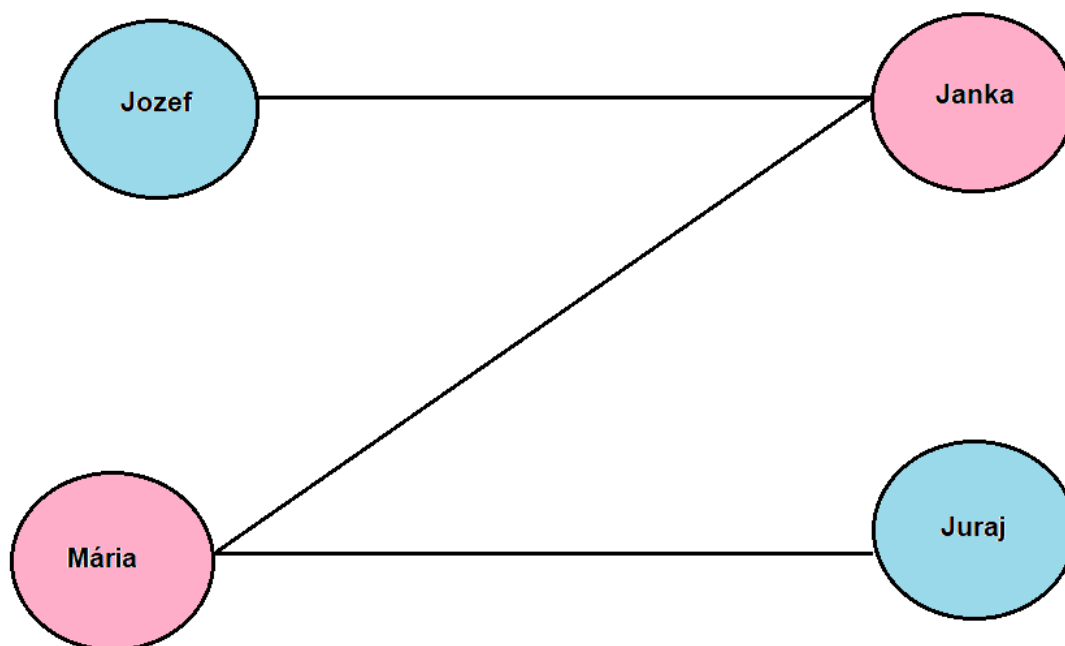
História sociálnych sietí siaha do 70. rokov 20. storočia, kedy vznikol počítačový systém BBS (Bulletin Board System). BBS slúži na posielanie textových správ a súborov cez internet. Ponúka predovšetkým textové rozhranie, ale niektoré webové verzie boli už vyvinuté s grafickým rozhraním. Vďaka posielaniu súborov a správ cez internet sa BBS stala súčasťou online komunity v 80-90. rokoch pred vznikom World Wide Web. V tomto období však nenastal výrazný rozvoj sociálnych sietí a ich masové používanie. Začiatkom 21. storočia došlo k vzniku sociálnych sietí menom SixDegrees, Myspace a LinkedIn. Išlo o historicky prvé siete, ktoré sa už podobali tým dnešným. Vznik Facebooku v roku 2004 odštartoval vznik ďalších sociálnych sietí a aj ich masové používanie. [1]

Rozvoj sociálnych sietí spôsobil revolúciu v spôsobe, akým ľudia komunikujú a spájajú sa navzájom. To umožnilo vytvárať si zoznam priateľov, zdieľať informácie len s určitými ľuďmi alebo skupinami. Je bežné, že keď sa nejaká sociálna sieť stala populárnou, tak inej sieti ubudli používatelia. V dnešnej dobe je to presne príklad Facebooku, kde ľudia ubúdajú a naopak TikTok a Instagram je čoraz populárnejší.

Z pohľadu informatiky došlo k algoritmickej zdieľaniu informácií. V rámci Facebooku došlo k vývinu edge-rank algoritmu, ktorý hodnotí interakciu používateľa s ponúknutým obsahom. Na základe toho sociálna sieť vie ponúknuť používateľovi relevantnejší obsah.[2]

Ak sa pozrieme na sociálnu sieť z hľadiska Teórie grafov, môžeme ju reprezentovať neorientovaným grafom. Vrcholy predstavujú používateľské účty a hrany predstavujú komunikáciu medzi dvoma účtami.

Na obrázku 1 vidíme jednoduchý príklad sociálnej siete, ktorá reprezentuje vzťahy - priateľstvá medzi Máriou, Jurajom, Jozefom a Jankou. Neorientované hrany spájajúce užívateľov, znamenajú priateľstvo medzi nimi.



Obrázok 1: Vzťahy medzi používateľmi sociálnej siete

2.1 Sociálne siete dnešného sveta

Sociálne siete prešli za posledných 20 rokov výraznou transformáciou. Zlepšili sa možnosti komunikácie, bezpečnosť aj poskytovanie obsahu. Sociálne siete si získali obrovské množstvo fanúšikov, ktorí siete často používajú. To logicky využili firmy ako priestor pre vlastnú reklamu. Zakladajú si profily na sociálnych sieťach, aby upútali pozornosť potenciálneho zákazníka. Siete im umožňujú aj organizovanie podujatí, rýchlejšiu a efektívnejšiu komunikáciu so zákazníkom. [3]

Nevýhodou je, že osobný kontakt medzi ľuďmi sa postupne vytráca. Častokrát sa ľudia stávajú sluhom sociálnych sietí a nie ich pánom. K súčasným najpoužívanejším sociálnym sieťam patria Facebook Instagram, Twitter, WhatsApp a Telegram. Spomenieme aj zoznamovacie služby Tinder a Badoo alebo aj Bumble.

2.1.1 SixDegrees.com

Sociálna sieť SixDegrees.com vznikla v roku 1997. Bola pomenovaná podľa teórie šiestich stupňov odlúčenia, ktorá hovorí o tom, že dvaja ľudia sa poznajú navzájom cez reťaz priateľov, ktorá je menšia rovná 6. Ľudia si môžu vytvoriť osobný profil, pridávať priateľov, zdieľať príspevky a vidieť vizuálnu reprezentáciu siete svojich priateľov. [4]

Táto sieť sa stala pomerne populárnou, ale nakoniec neobstála v konkurencii novších platforiem. V roku 2000 ju získala spoločnosť YouthStream Media Networks. [5]

2.1.2 LiveJournal

Jej vznik sa datuje do roku 1999. Bola populárnou najmä medzi mladými ľuďmi, ktorí ju používali na zdieľanie svojich myšlienok a skúseností s ostatnými priateľmi. Kľúčovou vlastnosťou siete bola možnosť vytvárať komunity na základe záujmov. Vznikali, tak skupiny na tému politiky, televíznych relácií alebo aktuálnych udalostí. V oblasti bezpečnosti došlo k zásadnejšiemu pokroku a inšpirovali sa tým viaceré sociálne siete. Užívatelia mohli nahlasovať nevhodné príspevky, mazať svoje príspevky alebo ich písať ako anonym. V dnešnej dobe už LiveJournal neexistuje.[6]

2.1.3 Flickr

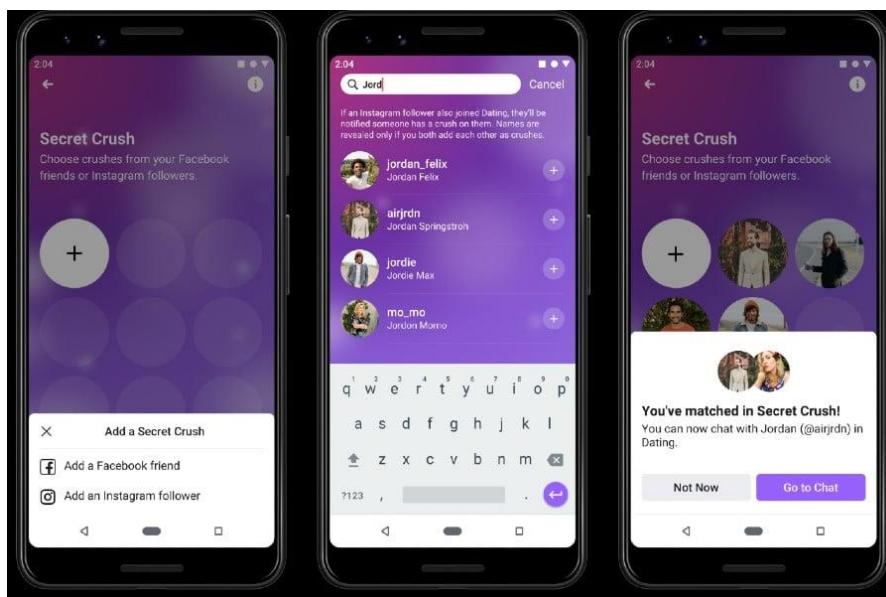
Ak by sme hľadali predchodcu Instagramu bol ním práve Flickr. Táto sieť slúži na ukladanie a zdieľanie fotografií medzi ľuďmi. Okrem zdieľania je k dispozícii široká ponuka funkcií na ochranu osobných údajov. Užívateľ si môže vybrať, s kým zdieľa fotografie, kto ju bude môcť komentovať alebo pridávať danú fotografiu do albumu. [7]

Celkovo Flickr umožnil budovanie komunít a bol obrovskou inšpiráciou pre vznik Instagramu.

2.1.4 Facebook

Sociálna sieť Facebook vznikla ako mini projekt Marka Zuckerberga v roku 2004 na pôde Harvardovej univerzity. Cieľom bolo, aby si návštevník vybral jednu z dvoch fotiek študentov univerzity, ktorá sa mu viac páči. Od tých čias sa Facebook pretransformoval na sieť ktorá má viac ako 2,96 miliardy používateľov [32]. V histórii Facebooku sa viackrát objavili kauzy s únikom dát a ponúkanie reklám, ktoré ohrozovalo súkromie užívateľov. [37]

V dnešnej dobe umožňuje Facebook používateľom komunikovať navzájom, vytvárať si zoznam priateľov, komunity s podobnými záujmami a takisto umožňuje kupovať a predávať produkty na službe menom Marketplace. Rok dozadu Facebook vytvoril „zoznamku“ na spôsob sociálnej siete Badoo. Na relevantnú ponuku používa zoznamy priateľov používateľov.



Obrázok 2: Facebook Dating zoznamka

2.1.5 LinkedIn

LinkedIn je profesionálna platforma, ktorá vznikla v roku 2003 a v súčasnosti je jej vlastníkom Microsoft. [9]

Táto sieť ma k dnešným dnom viac jak 900 miliónov užívateľských účtov.[8]. Primárne umožňuje užívateľom spojiť sa s inými ľuďmi, v tom istom pracovnom odvetví, prípadne predviesť svoje zručnosti, skúsenosti potenciálnym zamestnávateľom. Profilové údaje majú podobu životopisu. LinkedIn umožňuje podnikom propagovať svoje produkty a hľadať zamestnancov. [10]

Jedná sa o jednu z najväčších profesionálnych sieťových platforiem vôbec.

2.1.6 TikTok

Je v súčasnosti jednou z najpopulárnejších sietí. Jej predchodcom bola Music.ly, ktorá sa vyznačovala krátkymi 15-60 sekundovými videami. Vznikla v roku 2014. Jej videá častokrát obsahovali nejaký tanec, vtipnú reakciu, krátku scénu na nejakú známu pesničku. Za tri roky pôsobenia získala Music.ly zhruba 200 miliónov užívateľov. Jej formát videí prebral TikTok, ktorý vznikol neskôr v septembri roku 2017 a za 3 roky stihol nazbierať viac než miliardu užívateľov, čo sa pre porovnanie Instagramu podarilo za 8 rokov. Kľúčom tohto úspechu bol už spomínaný obsah a formát videí ale predovšetkým algoritmus na odporúčanie videí. [24]

Algoritmus je riadený podľa 3 kritérií:

1. Interakcia užívateľa - ktoré videá zdieľal, komentoval alebo označil ako obľúbené

2. Informácie o videu - popis, hashtagy, zvuky zaraďujúce video do nejakej kategórie, takisto aj dátum kedy video vyšlo
3. Nastavenia účtu, zariadenia – krajina z ktorej užívateľ pochádza, aké zariadenie využíva na prístup k TikToku

TikTok vďaka svojim krátkym videám a algoritmu umožňuje firmám veľmi rýchlo a vo veľkom robiť reklamu na ich produkty. Takisto vďaka algoritmu odporúčania môže video získať za krátky čas veľký počet zhládnutí. Spomeňme si napríklad na Black-lives Matter, ako rýchlo sa rozšírilo po svete. Za týmto úspechom stál vo veľkej miere práve TikTok a jeho používanie hashtagu. [24]



Obrázok 3: TikTok aplikácia

2.1.7 Instagram

Nasledovníkom siete Flickr je Instagram. Rovnako ako Flickr vznikol za účelom zdieľania pekných fotiek s ostatnými užívateľmi. K zdieľaniu fotiek sa vytvárajú a pridávajú tematické hashtagy, lokality fotiek, ktoré môžu užívatelia vo svojich príspevkoch spomenúť. Vzniknú tak kolekcie fotiek, ktoré sú viazané k danému hashtagu alebo lokalite. Neskôr Instagram pridal aj možnosť rýchleho posielania správ. [34] Instagram má v súčasnosti 2 miliardy aktívnych používateľov.[33] Stala sa z neho chatovacia sieť s niektorými obmedzeniami ale ku dnešným dňom je populárnejšia ako Messenger, WhatsApp alebo Telegram. V súčasnosti oblúba Instagramu rastie rýchlejšie ako u Facebooku, čo je spôsobené zmenou chovania používateľov. Používatelia Facebooku sú v priemere starší, a menej aktívni, naopak používatelia Instagramu sú výrazne mladší a omnoho aktívnejší, nebaví ich písať a čítať dlhé statusy na Facebooku, radšej si pozrú zaujímavú fotku s výstižným popisom, prípadne krátke video.

2.1.7.1 Hashtag

Hashtag predstavuje ľubovoľné slovné spojenie, ktoré užívatelia vytvoria a môžu používať vo svojich príspevkoch ako kľúčové slová. Vďaka hashtagom a lokalite vieme vytvárať komplexné siete grafov.

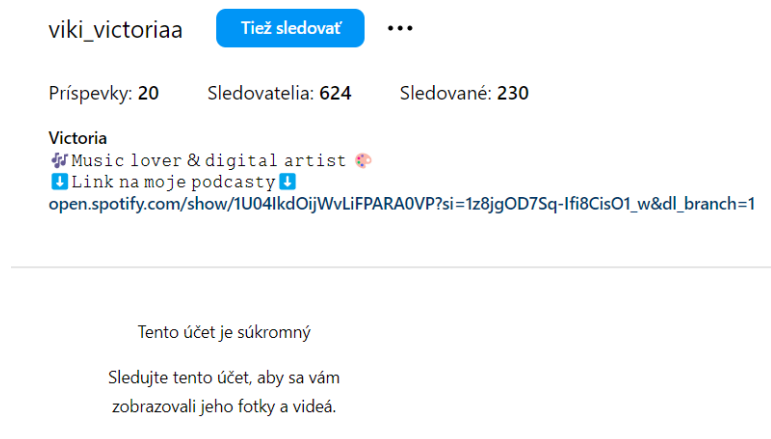
2.1.7.2 Instagramový profil

Každý instagramový účet obsahuje:

- počet osôb, ktoré sledujú užívateľov profil
- počet profilov, ktoré sleduje užívateľ
- počet príspevkov, ktoré užívateľ má uverejnených na svojom profile
- bio – užívateľov vlastný popis o ňom
- samotné príspevky a príbehy

Profil môže byť:

- verejný - po jeho načítaní máme prístup k všetkým jeho príspevkom, príbehom aj zoznamu jeho sledovateľov a ním sledovaných .
- súkromný - nevidíme jeho príspevky ani príbehy, ktoré uverejnil, takisto nevieme zobrazit' zoznam sledovateľov a ani ním sledovaných

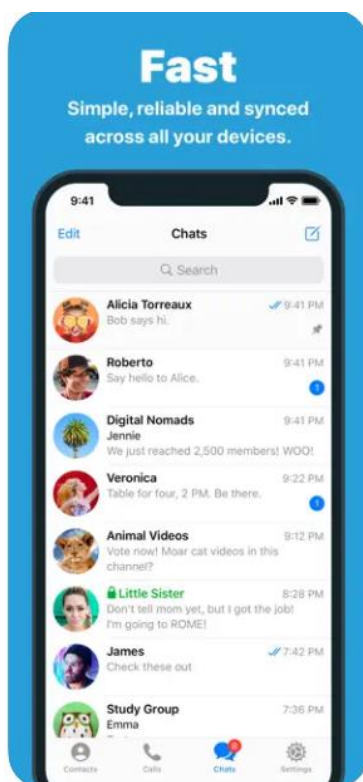


Obrázok 4: Súkromný účet

3 Telegram

Sociálna sieť Telegram bola vytvorená za účelom slobodnej komunikácie bez dohľadu vládnych agentúr. V roku 2011 ho založili bratia Durovci. Bola akýmsi upgradom WhatsAppu. Poskytuje ľuďom možnosť komunikácie pomocou okamžitých správ, vytvárania skupín a kanálov na zdieľanie obsahu, ako sú fotografie, videá a ľubovoľné súbory. Jej prioritou je bezpečnosť užívateľov. [38] Telegram má aktuálne 700 miliónov aktívnych používateľov. [33]

Na druhej strane však Telegram umožňuje vytváranie Botov a získavanie zaujímavých informácií.



Obrázok 5: Telegram aplikácia

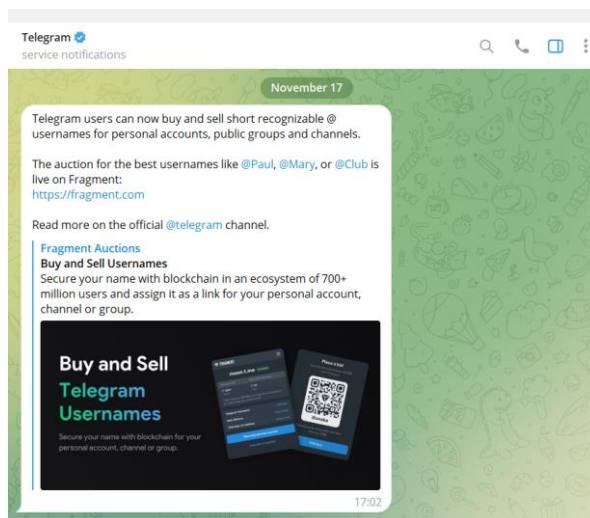
3.1 Základné pojmy používané na Telegrame

3.1.1 Telegram Account – Profil

Profil obsahuje telefónne číslo a profilovú fotku. Zároveň vieme pomocou neho komunikovať s inými ľuďmi, byť účastníkmi v skupinách, robiť si bot kontá alebo API.

3.1.2 Verified account – Overený profil

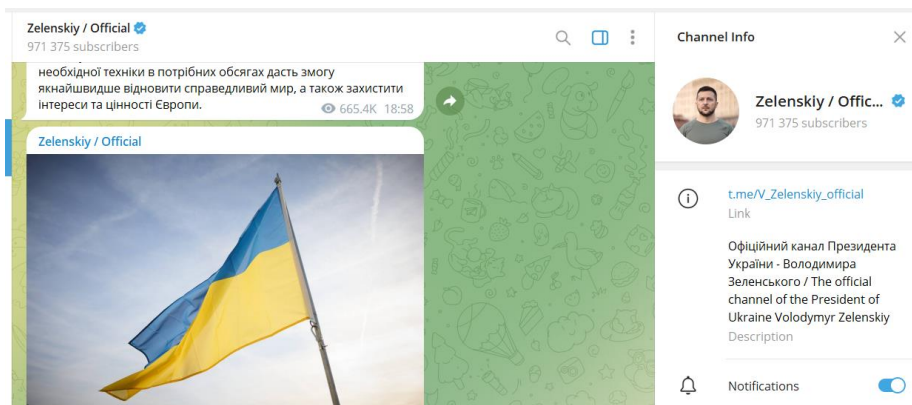
Profily, ktoré obsahujú symbol overenia. Sú to častokrát verejne známe osobnosti. Keďže dochádza k vytváraniu podvodných profilov, je tento symbol dôveryhodnosti veľmi nápomocný pre ľudí.



Obrázok 6: Overený profil na Telegrame

3.1.3 Channels

Každý profil si môže založiť svoj komunikačný kanál, pomocou ktorého vie jednostranne komunikovať s ľuďmi.[35] Každý kanál ma uvedený počet sledujúcich, užívateľské meno a fotografiu kanálu. Komunikácia má textovú podobu, ľudia môžu reagovať emotikonom, komentovať alebo zdieľať príspevok s inými ľuďmi.

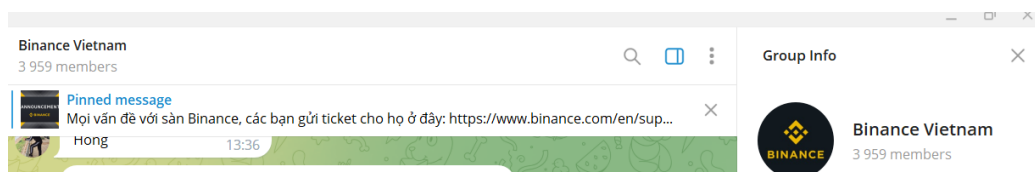


Obrázok 7: Komunikačný kanál

3.1.4 Skupina

Ľudia môžu vytvárať skupiny, v ktorej sa venujú určitej tematike.[35] Každá skupina obsahuje zoznam členov. Ľudia v skupine môžu komunikovať, vymieňať si informácie,

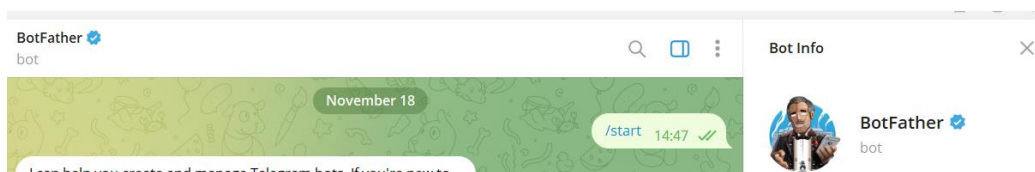
vyjadriť sa k téme. Konverzácia má podobu skupinového chatu na Messengeri alebo hocijakej inej chatovacej aplikácii. Zároveň sa dohliada na dodržiavanie pravidiel komunikácie.



Obrázok 8: Skupina

3.1.5 Bot

Každý užívateľ si môže vytvoriť svojho Bota – účet robota, ktorý vie komunikovať s inými užívateľmi. Najskôr ho ale musí druhá osoba kontaktovať, vie byť odoberateľom nejakých skupín a kanálov. Bot má v mene povinne názov Bot, aby sa odlišil od pravých užívateľských účtov. Správanie Botov je prísne monitorované komunitou. Zakladáme ho službou BotFather.



Obrázok 9: Bot

3.1.6 API

Telegram API je súbor nástrojov a protokolov, pomocou ktorých vieme zautomatizovať svojho Telegram robota tak, že dokáže posilať správy a zúčastňovať sa na živote v skupinách alebo kanáloch. Zároveň vieme vytvoriť vlastnú aplikáciu na Telegram, ktorá ponúka služby klientom. Každý užívateľ má k dispozícii práve jedno Telegram API. Každé Telegram API pozostáva z `api_id` – 7 miestneho čísla a `api_hash` 32-znakového hesla. Tak ako boty, aj činnosť Telegram API je prísne monitorovaná.[36]

4 Teória grafov

V praktickej časti sa budeme zaoberať grafmi, preto je nevyhnutné uviesť niekoľko pojmov.

4.1 Základné pojmy

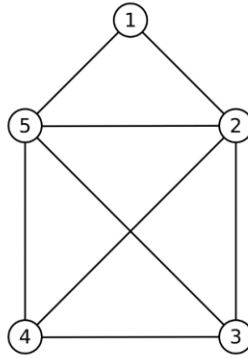
V tabuľke nižšie sú uvedené základné pojmy potrebné pre prácu s grafmi.

Názov, označenie	Popis
G	Reprezentácia konkrétneho grafu
Vrchol (uzol)	Základný prvok grafu, ktorý predstavuje nejakú entitu napríklad zviera, meno, mesto.
Hrana	Je čiara spájajúca dva vrcholy grafu. Predstavuje nejakú spoločnú vlastnosť, činnosť ktorá spája dva vrcholy
V	Reprezentácia množiny vrcholov grafu
E	Reprezentácia množiny hrán grafu.

Tabuľka 1: Základné pojmy grafu

4.2 Graf

Graf predstavuje množinu vrcholov pospájaných hranami. Matematicky vieme graf vyjadriť ako $G = (V, E)$. Vieme, že E predstavuje množinu hrán grafu takisto vieme hranu vyjadriť zápisom (u, v) , ako dvojicu vrcholov u, v ktoré sú spojené čiarou.[11]

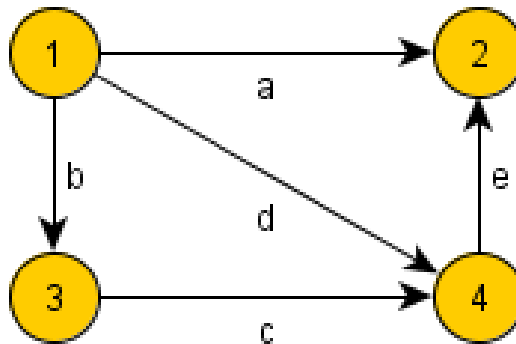


Obrázok 10: Neorientovaný graf

4.3 Typy grafov

4.3.1 Orientovaný graf

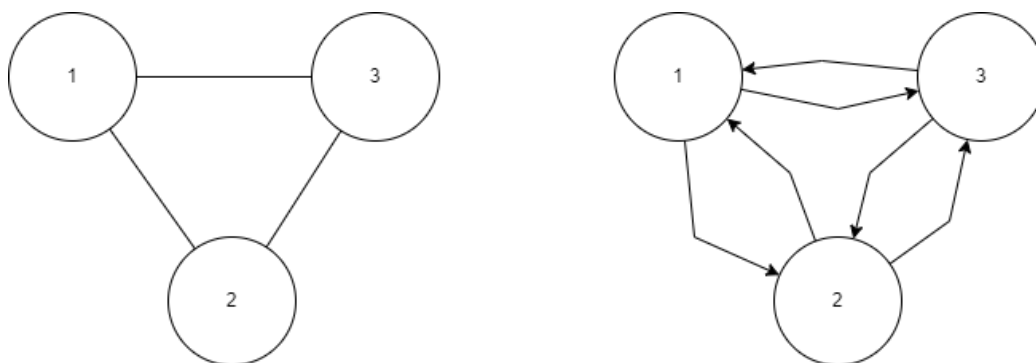
Graf $G = (V, E)$ predstavuje V množinu vrcholov a E množinu usporiadaných dvojíc vrcholov. Usporiadanú dvojicu môžeme nazvať aj ako orientovanú hranu, ktorá reprezentuje v grafickom znázornení šípku. Graficky (u, v) predstavuje šípku smerujúcu z vrcholu u do vrcholu v . Orientované grafy sa používajú v sociálnych sieťach a iných odvetviach. [12]



Obrázok 11: Orientovaný graf

4.3.2 Neorientovaný graf

Graf $G = (V, E)$ predstavuje V množinu vrcholov a E množinu neusporiadaných dvojíc vrcholov. Neorientovaná hrana predstavuje v tomto prípade neusporiadanú dvojicu vrcholov v grafickom znázornení čiaru. Príklad neorientovaného grafu môžeme vidieť na obrázku 12. Každý neorientovaný graf vieme pretransformovať na orientovaný graf, keď z každej neorientovanej hrany spravíme dve orientované hrany. Napríklad z neusporiadanej dvojice vrcholov (u, v) spravíme dve usporiadané dvojice (u, v) a (v, u) . [12]



Obrázok 12: Vyjadrenie neorientovaného grafu ako orientovaného.

4.3.3 Reprezentácia údajov - stĺpcový graf

Je graf, kde každý objekt má svoju hodnotu alebo frekvenciu. Graficky je to zaznačené obdĺžnikom, ktorý má pre každý objekt v grafe rovnakú šírku. Pri kreslení používame súradnicovú sústavu.



Obrázok 13: Stĺpcový diagram

Všimnime si, že na obrázku vodorovná os predstavuje merané objekty, kategórie a zvislá os ich hodnotu alebo frekvenciu. Graf skúma predaje vo Východnej Ázii za jednotlivé kvartály (štvrťrok) roku 2009.

4.4 Komplexné siete

Komplexné siete sú grafy so špeciálnymi vlastnosťami vyskytujúce sa v biológii, sémantike, strojárstve a iných oblastiach.

Vlastnosti siete [42]:

- klasterizácia vrcholov - skúma hustotu zhukovania vrcholov, ktoré sú medzi sebou spojené hranami a tým pádom tvoria susedné vrcholy
- mocninová distribúcia vrcholov - popisuje vrcholy a ich počet hrán
- prítomnosť hubov – prítomnosť vrcholov, ktoré majú hrany s väčšinou vrcholov grafu, graf vyzerá ako hviezda
- tranzitivita – ak je vrchol **a** spojený s vrcholom **b** a ten je spojený s **c**, je spojený aj vrchol **a** s **c**

Vlastnosti siete sú nevyhnutné pre správne určenie typu siete.

Typy sietí:

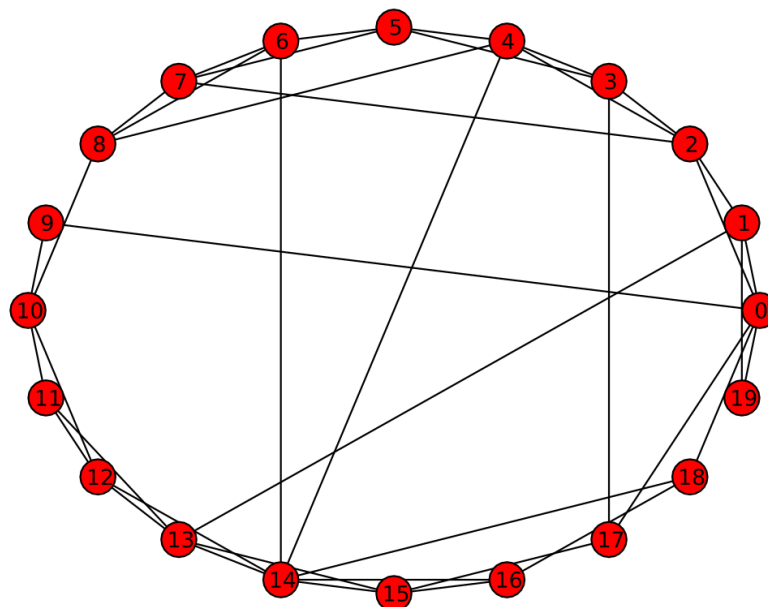
- prirodzená sieť
- sieť malého sveta
- bezškálová sieť

Typy sietí si podrobnejšie rozoberieme v nasledujúcich podkapitolách a neskôr ich znalosť využijeme v kapitole 7.

4.4.1 Fenomén malého sveta

Táto hypotéza nazývaná aj ako Šesť stupňov odlúčenia bola prvý krát sformulovaná v roku 1929 Frigyesom Karinthy. O pár desiatok rokov neskôr v roku 1967 bola otestovaná Stanleyom Milgramom. Jej podstatou je, že vzdialenosť medzi dvoma náhodne vybranými ľuďmi v grafe je maximálne 6. [43]

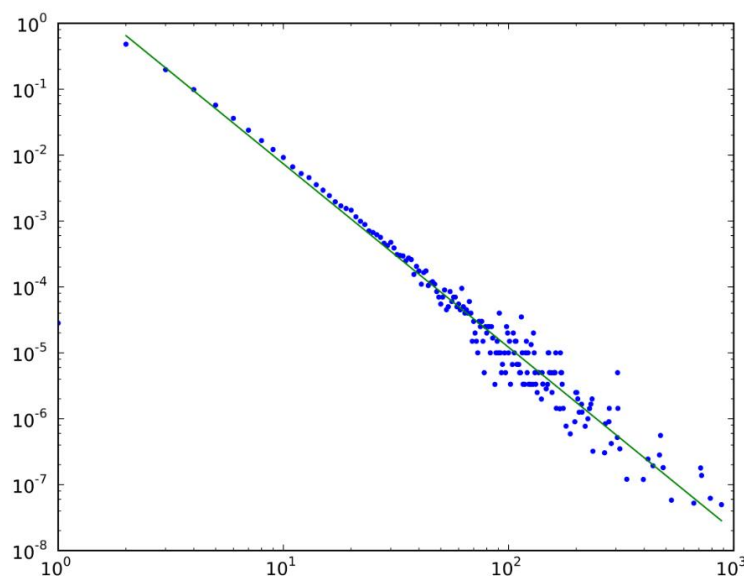
Na obrázku 14 môžeme vidieť graf, Watts-Storgatza z roku 1998, ktorý predstavuje model malého sveta. Klasterizácia je 0,75 a mocninová distribúcia vrcholov je tiež menšia ako 7. Na základe tohto modelu vznikol pojem sieť malého sveta, ktorá sa vyznačuje vysokou klasterizáciou a nízkou mocninovou distribúciou vrcholov. [43]



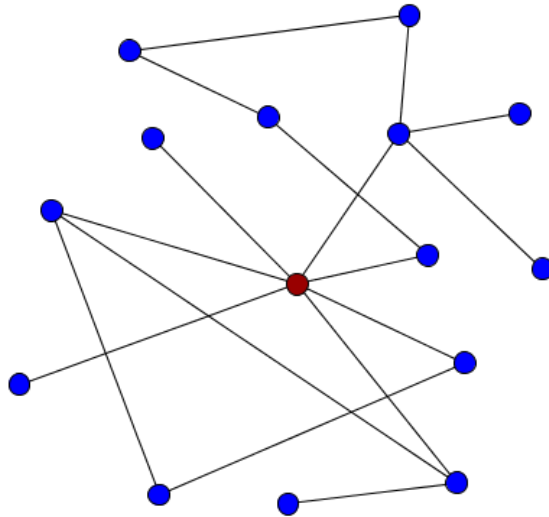
Obrázok 14: Watz-Storgatzov model malého sveta

4.4.2 Bezškálové siete

V roku 1999 vznikol vďaka Barabásimu a R. Albertovi algoritmus na generovanie bezškálových sietí. Algoritmus je založený na pribúdaní počtu vrcholov v čase ako aj prednostnom pripojení. To znamená, že ak pribudne nový vrchol, tak sa s veľkou pravdepodobnosťou spojí s vrcholom (hubom), ktorý má viacej spojení ako ostatné. Vďaka prítomnosti hubov sa bezškálové siete vyznačujú nízkou klasterizáciou a malou mocninovou distribúciou vrcholov. Graf distribúcie vrcholov pripomína tvarom mocninovú funkciu. [44]



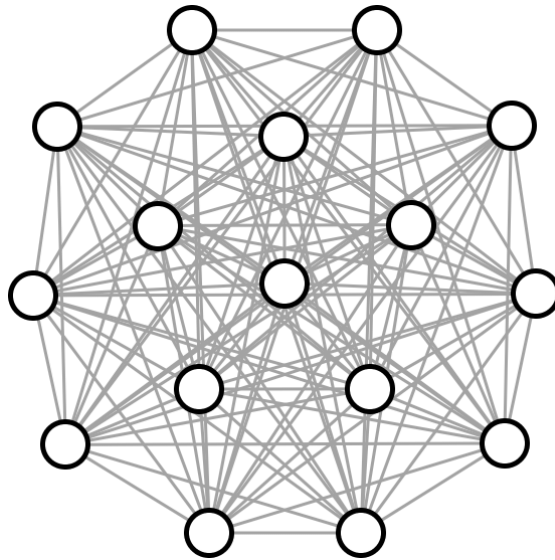
Obrázok 15: Graf mocninovej distribúcie vrcholov pri bezškálovej sieti



Obrázok 16: Bezškálová sieť

4.4.3 Prirodzené siete

Pojem prirodzených sietí vznikol v druhej polovici 20. storočia, kedy začali matematici a fyzici čoraz viac skúmať rôzne grafy, hustotu ich prepojení, prítomnosť hubov a iné vyššie spomínané vlastnosti. Prirodzená sieť obsahuje maximálne množstvo hrán, to znamená, že všetky jej vrcholy sú navzájom priamo prepojené. Klasterizácia tak nadobúda hodnotu 1. Vďaka tomu má mocninová distribúcia vrcholov tiež hodnotu 1. Prirodzená sieť sa tak podobá sieti malého sveta akurát v jej najextrémnejšom prevedení. [43]



Obrázok 17: Prirodzená sieť

5 Bezpečnosť údajov

Jednou z najdiskutovanejších tém dnešnej doby je aj ochrana osobných údajov. Čoraz častejšie dochádza k ich odcudzeniu, neoprávnenému nakladaniu s nimi alebo ich zneužitiu.

5.1 Nebezpečenstvá

Hrozba zneužitia citlivých údajov v dnešnej dobe číha na každom kroku. Dôvodom je často nerozvážnosť ľudí a zároveň vízia peňazí, kedy sociálne siete predávajú údaje o svojich používateľoch.

Medzi základné chyby, ktoré ľudia často robia, patrí zabezpečenie účtu slabým heslom, zdieľanie veľkého množstva osobných informácií s verejnosťou.[14]

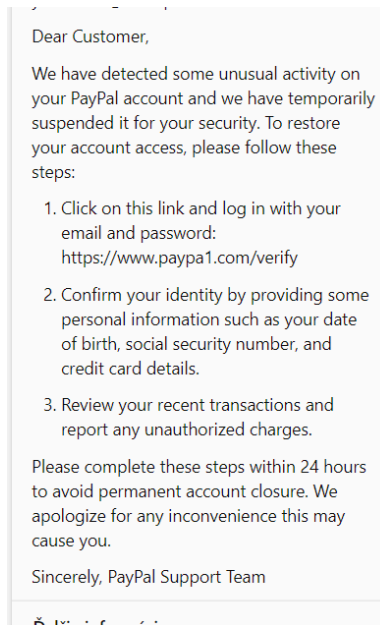
Na druhej strane sú tu firmy, častokrát priamo späté so sociálnymi sieťami, ktoré zbierajú a vyhodnocujú informácie o užívateľoch a neskôr ich predávajú záujemcom o ciele reklamu. Takisto máme aj útočníkov, hackerov, ktorí z tých istých dôvodov ako firmy vymýšľajú rôzne spôsoby, ako nenápadne získať citlivé dáta o užívateľoch.

5.2 Phishing

Medzi najznámejšie spôsoby kradnutia dát patrí phishing alebo v preklade lovenie dát. Útočník sa snaží tváriť čo naj dôveryhodnejšie, aby získal od prijímateľa jeho osobné informácie.

5.2.1 Klasický phishing

Úlohou útočníka je vytvoriť mail, ktorý bude obsahovať dôveryhodného odosielateľa. Mail sa snaží svojím obsahom presvedčiť adresáta, aby spätne zaslal nejaké dôležité údaje alebo klikol na odkaz. Po kliknutí na odkaz môže byť čitateľ hneď zavírený a okradnutý o citlivé dáta. V prefikanejšom prípade útočník vyrobil vierohodný web, kde človek zadá údaje a následne odíde nič netušiaci. Môže sa stať, že mail bude obsahovať nejaký spustiteľný exe súbor po ktorého spustení dostane čitateľ vírus. Cieľom klasického phishingu je oklamať čo najviac ľudí. Medzi klasický phishing patrí aj clone phishing. Ide o mailovú správu, ktorá obsahuje napodobeninu legitímneho mailu so škodlivou prílohou alebo odkazom.[15]



Obrázok 18: Falošná správa vyrobená Bing chatbotom

5.2.2 Spear phishing

Na rozdiel od klasického phishingu je cieľom oklamať špecifickú skupinu ľudí. Ide o najpopulárnejší spôsob získavania dát. Ako sa vyvíja bezpečnosť, tak sa zvyšuje aj šikovnosť útočníkov. Tí sa zameriavajú na získanie prístupu k bankovým účtom, sociálnym sieťam, dátam veľkých organizácií. V tomto prípade ide o komplexnejší a koordinovanejší spôsob útoku ako pri klasickom phishingu.[16]

5.3 Cookies

Medzi užitočné ale zároveň aj nebezpečné webové nástroje patria súbory cookies. Cookies majú podobu textového súboru a sú ukladané na disk používateľovho počítača. Výhodou ich textovej podoby je, že nie sú spustiteľné, tým pádom nepredstavujú priamo vírusovú hrozbu pre náš počítač. Delíme ich podľa dvoch kritérií podľa času uloženia a podľa využitia.[19]

Podľa času uloženia:

- dočasné - sú zapamätané len po dobu návštevy stránky
- permanentné - sú zapamätané, kým im nevyprší platnosť

Podľa účelu:

- základné: na zapamätanie prihlasovacích údajov, stavu prebiehajúcej úlohy
- funkčné: zapamätanie obsahu košíka pri nakupovaní, poskytovanie personalizovaných reklám

- reklamné: poskytovanie ponúk a rovnako reklám na základe osobného vyhľadávania

Pri návšteve webových stránok by sme si mali dávať pozor, aké typy cookies a na aké účely stránka využíva. Veľmi rýchlo môže prísť k úniku dát, ktoré by sme si neželali. [21]

5.4 Bezpečnosť údajov

Bezpečnosť údajov je ako sme už písali často diskutovanou témou. Nie všade vo svete platia rovnaké zákony na ochranu osobných údajov. Takisto nie je každý človek rovnako odolný voči hrozbám internetu.

5.5 TikTok a bezpečnosť údajov

Rovnako ako Telegram aj TikTok trpí nedostatočnou ochranou údajov o užívateľoch. Tento problém majú aj spoločnosti Meta (prevádzkovateľ sociálnej siete Facebook) a Google. TikTok vlastní materskú spoločnosť ByteDance, ktorá má svoje sídlo v Číne. Je dosť pravdepodobné, že vláda v Číne má prístup k jej údajom. V dôsledku týchto hrozieb sa používanie TikToku obmedzilo v niektorých krajinách Ameriky a Európy. Takýto problém bol historicky aj medzi USA a čínskou spoločnosťou Huawei.[25]

Poslednou hrozbou je poskytovanie dát firmám s ktorými TikTok spolupracuje. Firmy pomocou pixelu (sledovacieho softvéru) zbierajú obrovské množstvo dát o každom užívateľovi, ktorý ich službu alebo stránku využil. Následne ich poskytuje TikToku, ktorý dáta využíva údajne na reklamné účely. Vďaka tomu má TikTok informácie aj o ľuďoch, ktorí užívateľský účet na ňom nemajú založený.[26] Preto vlády odporúčajú správne zabezpečiť prehliadač a TikTok účet, aby sa poskytlo minimum údajov.

5.6 Bezpečnosť Európska únia a svet

Ak by sme hľadali vo svete krajiny, ktoré dbajú výrazným spôsobom na ochranu osobných údajov, boli by to krajiny Európskej únie. Európska únia si ustanovila ochranu osobných údajov ako prioritu. Bezpečnosť dát pri automatizovanom spracovaní vyústila v roku 2016 do formovania GDPR (General Data Protection Regulation), ktoré nadobudlo platnosť v máji 2018. [17]

GDPR vyžaduje súhlas používateľa na spracovanie jeho osobných údajov, pri akejkoľvek spoločnosti pôsobiacej v Európskej únii. Takisto má človek právo vedieť, ak došlo k úniku jeho údajov vplyvom hackerských útokov. V neposlednom rade má užívateľ právo žiadať spoločnosti o zmazanie jeho dát a nahlasovať prípadne aj žiadať pokuty pre spoločnosti,

ktoré bezpečnosť dát porušujú. Pokuty za nedodržanie nariadení GDPR sú až do výšky 4% z obratu firmy alebo 20 miliónov EUR, podľa toho čo je vyššie.[17] Prehľad najvyšších pokút je možné sledovať na stránke [41].

Pri spracovaní údajov vzniká problém s ich prenosom. V rámci dohôd o prenose dát došlo viackrát k porušeniu právnych predpisov, ktoré Európska únia ustanovila. Takisto sa dohoda o prenose dát medzi USA a Európskou úniou neustále formuje. Spoločnosť Google a Facebook mali už viackrát problémy s porušovaním ochrany dát v Európskej únii, za čo im Európska rada udelila pokuty v hodnote niekoľkých sto miliónov eur.[18] Okrem bezpečného prenosu dát užívateľa sa ustanovili účely, pri ktorých sa dáta o používateľoch môžu používať. Medzi USA a EÚ došlo k uzatvoreniu dohody o poskytnutí údajov z dôvodu teroristickej činnosti, to zahŕňa údaje, ktoré boli spracované leteckými alebo finančnými spoločnosťami.[17]

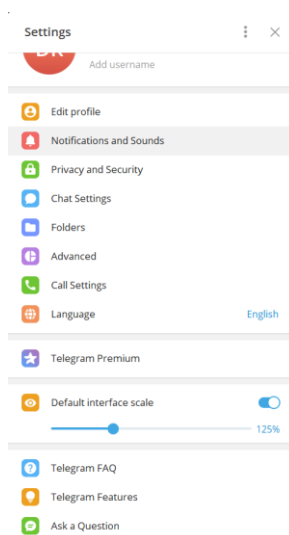
Ochrana údajov sa týka aj používania súborov cookies webovými stránkami na rôzne účely. Európska únia prikazuje, aby sa každý užívateľ pri navštívení stránky mohol slobodne vybrať, či povolí používanie cookies. alebo nie. Takisto prikazuje majiteľom stránkam neobmedzovať prístup tým ľuďom, ktorí ich použitie odmietli.[20]

Môžeme jasne vidieť, že nikde vo svete nie je tak dbané na bezpečnosť osobných údajov ako práve v Európskej únii. Robí to z nej vzor v bezpečnosti, ktorý môžu v budúcnosti nasledovať aj iné krajiny sveta.

5.7 Telegram a bezpečnosť

Je všeobecne známe, že Telegram je bezpečnou sieťou, kde vaše údaje nie sú v ohrození. Telegram využíva na prihlasovanie overenie telefónnym číslom. Telefónne číslo je viazané na SIM kartu a tá je štandardne zabezpečená minimálne PIN kódom. Po 3 neúspešných pokusoch je SIM karta blokována. Ďalej si Telegram ukladá naše kontakty, správy, skupiny v slabo zašifrovanej forme na svojich serveroch.[22]

Na druhej strane Telegram ponúka širokú ponuku nastavení na správu osobných údajov. Používateľ má vďaka tomu možnosť rozhodnúť sa, ktoré údaje o sebe skryje alebo nechá verejné.



Obrázok 19: Telegram nastavenia

5.8 Ochrana pred nebezpečenstvami

Dobrou prevenciou pred nebezpečenstvami je vzdelávanie používateľov sociálnych sietí a internetu. To je úloha pre každého jednotlivca, pre školy aj pre rodičov detí. Je priam nevyhnutnosťou mať v počítači nainštalovaný antivírusový program, ktorý rozozná škodlivé súbory alebo zabráni návšteve škodlivých stránok. Jedným z najznámejších takýchto produktov je Windows Defender od firmy Microsoftu, ktorý je súčasťou operačného systému Windows. Windows Defender zabezpečuje základnú ochranu počítača, platené antivírusové produkty však poskytnú vyššiu ochranu. Predstaviteľom plateného antivírusového riešenia je aj slovenský produkt ESET Internet Security. Mimo toho by človek mal dbať na bezpečné ukladanie prihlasovacích údajov a správne nastavenia používania súborov cookies.[23]

6 Použité technológie a ich využitie

V praktickej časti sme na získanie dát použili programovací jazyk Python s jeho knižnicami, ale aj rôzne Telegram aplikácie. Niektorým použitým technológiám sa budeme venovať podrobnejšie.

6.1 Python

Program Python je interpretovaný, interaktívny programovací jazyk, ktorého zakladateľom je Guido van Rossum. Stal sa akýmsi nastavbovým jazykom pre aplikácie, ktoré potrebovali programovateľné rozhranie. Výhodou jazyka bola možnosť vývoja modulov v programovacích jazykoch C alebo C++.[39]

Dnes Python ponúka širokú škálu programovacích knižníc na získavanie a analýzu dát zo sociálnych sietí.

6.2 Telethon

Je Python knižnica slúžiaca na prácu so sociálnou sieťou Telegram. Pre jej použitie je potrebné poznať svoje `api_id`, `api_hash` prípadne telefónne číslo.

Vďaka Telethonu vieme:

- sťahovať súbory z konverzácií,
- prezerat' text konverzácií,
- vyhľadávať skupiny alebo kanály,
- posielat' správy užívateľom.

My sme použili:

- prezeranie členov skupín,
- prezeranie pozretí jednotlivých príspevkov v kanáloch,
- časy uverejnenia príspevkov v kanáloch.

6.2.1 Získavanie členov skupín

Každý administrátor telegramovej skupiny spravuje zoznam členov. Zoznam môže byť viditeľný pre verejnosť alebo aj skrytý. [13] Ak je zoznam členov verejný, vieme pomocou rôznych nástrojov získať ďalšie informácie. My sme si zvolili metódu `get_participants()`.

Vstupným argumentom funkcie je názov skupiny, z ktorej chceme získať jej členov. Po jej zavolaní nám vráti zoznam maximálne 10-tisíc užívateľov. Pri správnej konfigurácii môže

získať až 70-80 percent z celkového počtu členov skupiny. V zozname nájdeme profily odoberateľov. Je to akoby sme si pozreli profil na Facebooku.

Takýto profil člena obsahuje zvyčajne užívateľské číslo, ktoré je jedinečné. Obsahuje aj Meno Priezvisko, telefónne číslo, poslednú aktivitu. Avšak poskytuje aj informáciu o tom, či daný profil mal už historicky nejaké problémy s dodržiavaním pravidiel na Telegramu alebo nie.

```
lang_code=None, emoji_status=None, usernames=[], User(id=203608975, is_self=False, contact=False, mutual_contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=8148742905157220040, first_name='sic.boi', last_name=None, username='sic.boi', phone=None, photo=UserProfilePhoto(photo_id=87449388925320635, dc_id=2, has_video=False, personal=False, stripped_thumb=b'\x01\x08\x08[0xaaQ0r06\x8a(\xa4\x84\x7f)'), status=UserStatusRecently(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), User(id=380983202, is_self=False, contact=False, mutual_contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=-6882920112200917462, first_name='Max', last_name=None, username='Mak7igorovych', phone=None, photo=None, status=UserStatusRecently(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), User(id=380709935, is_self=False, contact=False, mutual_contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=6582967637519504408, first_name='Artem', last_name='P', username='PnchKA', phone=None, photo=None, status=UserStatusLastWeek(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), User(id=756974209, is_self=False, contact=False, mutual_contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=8774001766679973935, first_name='5ynka xn16a', last_name=None, username='kisakiskiskis', phone=None, photo=None, status=UserStatusRecently(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), User(id=401055461, is_self=False, contact=False, mutual_contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=4452117447574577655, first_name='Dmytro', last_name='Bohdanov', username='Hansovich', phone=None, photo=UserProfilePhoto(photo_id=1722520089333442490, dc_id=4, has_video=False, personal=False, stripped_thumb=b'\x01\x08\x08\xad\x14\xcc\xac\x08c\xf9\xd1E\x15W\x11'), status=UserStatusRecently(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), User(id=892091189, is_self=False, contact=False, deleted=False, bot=False, bot_chat_history=False, bot_nochats=False, verified=False, restricted=False, min=False, bot_inline_geo=False, support=False, scam=False, apply_min_photo=True, fake=False, bot_attach_menu=False, premium=False, attach_menu_enabled=False, access_hash=-3396159859902238623, first_name='Яна Нагоніна', last_name=None, username='nadolinka', phone=None, photo=UserProfilePhoto(photo_id=3831502482260994967, dc_id=2, has_video=False, personal=False, stripped_thumb=b'\x01\x08\x08b\xce3\x0e\n\xber0L\x8f_\xe5E\x14R\xb9\xb2\x82g'), status=UserStatusRecently(), bot_info_version=None, restriction_reason=[], bot_inline_placeholder=None, lang_code=None, emoji_status=None, usernames=[]), total=2887)
```

Obrázok 20: Zoznam výpisu funkcie get_participants()

Ako môžeme vyčítať z obrázka, metódu get_participants() sme použili pri hľadaní prienikov medzi odoberateľmi rôznych jazykových skupín, botov v skupinách a aj pri získavaní telefónnych čísel. Celkovo bola táto funkcia veľmi užitočná a pomohla získať údaje pre viaceré štatistiky.

6.2.2 Prehľadávanie konverzácií

Telethon knižnica poskytuje metódu get_messages(), kde vstupným argumentom je názov skupiny alebo užívateľa. Ďalšími z voliteľných argumentami metódy sú:

- počet správ ktoré metóda vráti
- nastavenie či chceme získať od najstarších správ alebo najnovších
- dátum odkedy chceme správy získať.

Metóda vracia správy. Z každej správy vieme získať odosielateľa, čas poslania a ak ide o kanál či skupinu vieme získať koľko ľudí správu prečítalo. V praxi sme funkciu použili na porovnanie popularity telegramových kanálov, ich sledovanosť a pravidelné publikovanie príspevkov. Pri čítaní textov správ sa často stávalo, že program nevedel daný znak textu dekódovať. Išlo často o nejaký emotikon alebo písmeno z abjedy.

2023-05-28 15:26:29 12389 Оппозиционное турецкое агентство Анка утверждает, что Кылычдароглу лидирует с отрывом в 2% после обработки 44% бюллетеней во втором туре.

При этом по данным гостВ, Эрдоган набирает 56,36 % голосов, Кылычдароглу 43,64 % после обработки 49,39% бюллетеней.

Obrázok 21: Správa s počtom videní a časom uverejnenia

6.2.3 Vyhľadávanie kanálov, skupín a užívateľov

V prípade, že presne nepoznáme užívateľské meno, názov skupiny alebo kanálu, môžeme vytvoriť automatizovaný nástroj na poskytovanie alternatívnych používateľov, kanálov a skupín pomocou funkcie `Search_request()`.

Vstupným argumentom do tejto funkcie je text, ktorý chceme vyhľadávať. Okrem zadávaného textu môžeme ďalej nastaviť počet výsledkov, ktoré nám táto metóda má vrátiť. Výsledkom je slovník, ktorý obsahuje zoznamy kanálov, skupín alebo užívateľov. Nájdení používatelia obsahujú skoro identické informácie ako vracia metóda `get_participants()`.

```
Found(  
  my_results=[  
  ],  
  results=[  
    PeerUser(  
      user_id=1257189943  
    ),  
    PeerUser(  
      user_id=1117322539  
    ),  
    PeerUser(  
      user_id=115446950  
    ),  
    PeerUser(  
      user_id=470006949  
    ),  
    PeerUser(  
      user_id=2123703607  
    ),  
    PeerUser(  
      user_id=1006727013  
    ),  
  ],  
  chats=[  
  ],  
  users=[  
    User(  
      id=1257189943,  
      is_self=False,  

```

Obrázok 22: Výstup funkcie `Search_request ()`

6.2.4 Posielanie správ

Keď chceme s niekým komunikovať, otvoríme si Telegram a píšeme si s konkrétnym človekom. Pri konverzácii s bot účtom sú naše konverzácie častokrát jednotvárne. Väčšina botov obsahuje sadu naprogramovaných príkazov na ktoré dokáže odpovedať. Na posielanie správ vieme použiť metódu `send_message()`.

Jej prvým argumentom je adresát správy, a druhým argumentom je obsah správy, ktorú mu chceme poslať.

```

# user_id = []
# with open('dead_accounts_.txt','r') as f:
#     for i in f.readlines():
#         client.send_message('tgscanrobot',i.split(' ')[0])
#         for c,j in enumerate(client.get_messages('tgscanrobot',limit=2)):
#             if c == 1:
#                 print(str(j.message))
#                 try:
#                     user_id.append([i[:-1],str(j.message).split('\n')[6].split(
#                 except:

```

Obrázok 23: Funkcia send_message()

6.2.5 Sťahovanie súborov z konverzácií

Pri sťahovaní súborov z konverzácie môžeme využiť metódu `download_media()`, ktorá nám prílohu stiahne a uloží na zadané miesto.

Vstupnými argumentami funkcie sú objekt typu správa, ktorý dokážeme získať z metódy `iter_messages()`, ktorá funguje na podobnom princípe ako metóda `get_messages()`. Ďalším argumentom je miesto, kde chceme súbor uložiť. Každý objekt typu správa nesie v sebe aj booleovské premenné. Ak by sme prehľadávali konverzácie, vedeli by sme sťahovať len obrázky, a to vďaka premennej `message.photo`.

```

async def main():
    async for message in client.iter_messages('me'):
        print(message.id, message.text)
        if message.photo:
            print('File Name : ' + str(message.file.name))
            path = await client.download_media(message.media, "youranypathhere")
            print('File saved to', path) # printed after download is done

```

Obrázok 24: Funkcia download_media()

6.3 Pyvis

Je Python knižnica, ktorá je určená na vytváranie a vizualizáciu grafov. Grafu vieme nastaviť veľkosť, farbu, popisy vrcholov aj ich tvar. Ďalej vieme nastaviť váhu, popisy hrán ale aj veľkosť plochy, na ktorej sa bude graf zobrazovať. Po skompilovaní programu knižnica vytvorí html súbor s grafom, ktorý si vieme otvoriť v ľubovoľnom internetovom prehliadači. Graf môžeme podľa potreby priblížiť, alebo interaktívne kliknúť na vrcholy a kurzorom ich ťahať po ploche.

6.4 Phonenumbers

Na získanie národnosti užívateľov na základe predvoľby ich telefónneho čísla sme použili knižnicu `phonenumbers` a jej dve metódy `phonenumbers.parse("number")` a `geocoder.description_for_number(phonenumbers, "language")`. Prvá z metód zoberie telefónne číslo a overí, či má správny formát. To znamená že obsahuje len čísla a na začiatku má znamienko plus. Následne sa zavolá druhá metóda, ktorá nám nájde krajinu, ktorej telefónne číslo patrí. Výsledok vypíše v jazyku, aký si zvolíme. Ak nenašla funkcia krajinu vypíše prázdny reťazec.

```
# with open("nationality.txt", "w") as f:
#     for i in cisla:
#         phonenumbers = phonenumbers.parse("+" + i)
#         country = geocoder.description_for_number(phonenumbers, 'en')
```

Obrázok 25: Knižnica `phonenumbers`

6.5 Combot

Combot je telegramová aplikácia, ktorá poskytuje aktuálny prehľad verejných skupín na Telegramu. Skupiny sú na základe jazyka, v ktorom konverzácia prebieha, priradené ku danej krajine. Takisto Combot ponúka záujemcom plateného bota na spravovanie ich skupín (bezpečnosť užívateľov, dodržiavanie podmienok skupiny).

6.6 TGStat

Webová služba ktorá poskytuje katalóg telegramových kanálov. Každý kanál ma krajinu, a kategóriu do ktorej patrí. Takisto TGStat umožňuje vyhľadávať kanály podľa príspevkov. Vieme si prezerať zaujímavé štatistiky každého kanála, ich aktivitu, sledovanosť príspevky a rast jej členov.

6.7 TGScanRobot

Telegramová aplikácia, ktorá poskytuje o každom užívateľovi počet, názvy skupín a kanálov v ktorých je členom. Tato služba je platená ale počet skupín je zadarmo.

6.8 Instagrapi

Instagrapi je knižnica vyvinutá v jazyku Python na scrapovanie (extrahovanie) dát na Instagrame. Zároveň je jednou z najlepších knižníc, ktoré môžeme na scraping dát použiť,

vzhľadom na jej pravidelnú aktualizáciu kvôli stále meniacim sa bezpečnostným pravidlám Instagramu.

Vďaka metódam `photo` a `video download()` vieme sťahovať videá alebo fotky z ľubovoľného verejného profilu. Na základe `product_type` vieme rozlíšiť či je to príbeh, krátke video a iné.

Ďalšími užitočnými metódami sú `user_follow()` a `user_unfollow()`. Umožňujú začať a ukončiť sledovanie profilov. Takisto máme k dispozícii funkciu `user_followers()` a `user_following()`, ktoré umožňujú získať zoznam sledovateľov a profilov, ktoré sám daný užívateľ sleduje. Metódy fungujú iba na verejné profily.

Okrem týchto metód vieme získať aj príspevky na základe hashtagov a lokalít.

```
if m.media_type ==1:
    #Photo
    paths.append(cl.photo_download(m.pk)
elif m.media_type == 2 and m.product_type == 'feed':
    #Video
    paths.append(cl.video_download(m.pk)
elif m.media_type == 3 and m.product_type == 'igtv':
```

Obrázok 26: Funkcia video a photo download

6.9 Selenium

Selenium je bezplatný nástroj slúžiaci na automatizáciu testovania webových aplikácií v rôznych prehliadačoch. Podporuje viacero programovacích jazykov ako je Java, Python, C# a iné. [31] My sme Selenium využili pri telefónnych kontaktoch na získanie ich krajiny pôvodu.

Na začiatok je nutné určiť, ktorý prehliadač budeme používať. My sme sa rozhodli pre Chrome, a preto sme si do premennej uložili `driver = webdriver.Chrome()`. Následne sme načítali stránku pomocou `driver.get("url")`. Po vybratí správneho elementu stránky sme manuálne zistili jeho id a pomocou funkcie `elem = driver.find_element(By.ID, "id")` sme ho získali. V našom prípade to bolo vyhľadávacie okno. Do neho sme pomocou funkcie `elem.send_keys("text")` zadali telefónne číslo, ktorého krajinu pôvodu sme chceli zistiť. Pomocou `Keys.RETURN` sme odsimulovali stlačenie tlačidla ENTER na klávesnici a nám sa zobrazili výsledky vyhľadávania. Na získanie html kódu novo načítanej stránky sme použili premennú `driver.page_source`. Na koniec celého vyhľadávania sme prehliadač zavreli pomocou `driver.close()`.

```

import time
from selenium import webdriver
from selenium.webdriver.common.keys import Keys
from selenium.webdriver.common.by import By
driver = webdriver.Chrome()
driver.get("https://phonenum.info/")
slovník = set()
for i in range(2):
    with open("russian_numbers" + str(i) + ".txt", "r") as f:
        for j in f.readlines():
            riadok = j.split("%")
            slovník.add(riadok[2][:-1])
pocet_cislo = 0,0
with open("nationality.txt", "w") as f:
    for c,i in enumerate(slovník):
        try:
            elem = driver.find_element(By.ID, "phoneInput")
            elem.clear()
            elem.send_keys(i)
            elem.send_keys(Keys.RETURN)
            output = driver.page_source
            start_index = output.find("<title>")
            end_index = output.find("0neparop:")
            f.write(i + "%" + output[start_index:end_index].split(":")[1].strip()[:-1] + "\n")
            print(c)
            break
        except:
            pass
# assert "No results found." not in driver.page_source
driver.close()

```

Obrázok 27: Selenium automatizovaný nástroj

7 Praktická časť

Telegram je sociálna sieť, ktorá obsahuje veľa údajov vhodných pre štatistiku. My sme hľadali spoločných odoberateľov rôznych jazyčných skupín. Takisto sme sa zaoberali popularitou kanálov a vplyvom udalostí na nich. V neposlednom rade sme zisťovali národnosť používateľov, ako aj účasť botov v skupinách. Ako poslednú sme popísali skúsenosť so súčasným scrapingom na Instagrame.

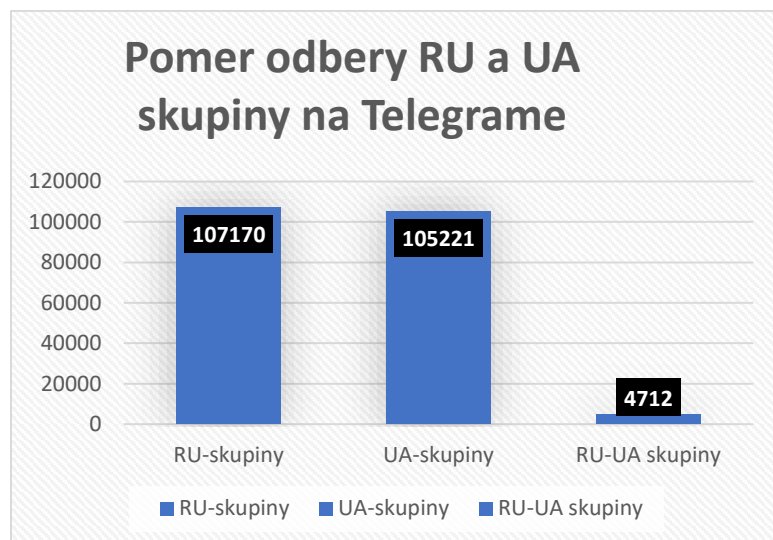
7.1 Analýza odoberateľov jazyčných skupín v Telegrame

V tejto časti sme vybrali vzorku zhruba 110000 ľudí ktorí odoberajú ruské a iných približne 110000 odoberateľov ukrajinských skupín a hľadali sme ich prienik. Neskôr sme zobrali vzorky ďalších krajín bývalého Sovietskeho zväzu, porovnali ich vzorky s ruskými a ukrajinskými a hľadali takisto ich prienik. Na základe jazyka, v ktorom v ktorom prebieha konverzácia v skupinách určíme či je ukrajinská, ruská alebo iná. Zoznam skupín sme čerpali z [40].

Pri získavaní dát sme narazili na zaujímavé skutočnosti. Pri získavaní užívateľov sme museli so správnym časovým odstupom žiadať jednotlivé skupiny o členov. V prípade veľkého množstva žiadostí v krátkom časovom intervale nás Telegram zablokoval a najbližších 24 hodín vrátil vždy nulové počty členov skupiny. Ani správne oneskorenie medzi požiadavkami nestačilo na získanie hľadaných informácií, nakoľko Telegram umožnil za deň stiahnuť maximálne vzorku 300 000 užívateľov a potom takisto vracal najbližších 24 hodín 0 užívateľov. Ďalším zaujímavým zistením bolo, že Telegram v rámci ochrany súkromia členov poskytuje maximálne 10000 členov jednej skupiny.

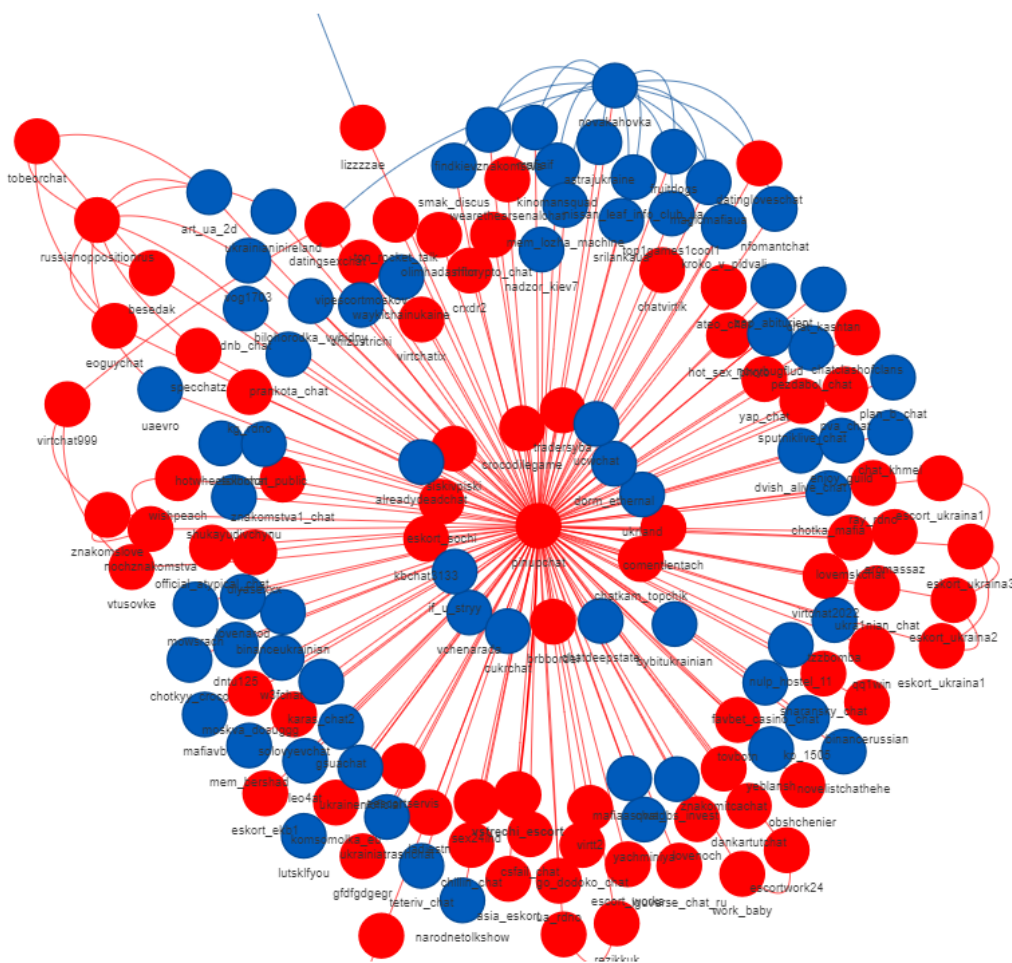
7.1.1 Ruské verzus Ukrajinské skupiny

Zo vzorky zhruba 110000 odoberateľov ruských skupín a 110000 odoberateľov ukrajinských skupín sme našli 4712 užívateľov, ktorí odoberajú aj ukrajinské aj ruské skupiny, čo tvorí približne 4% skúmaných odoberateľov.



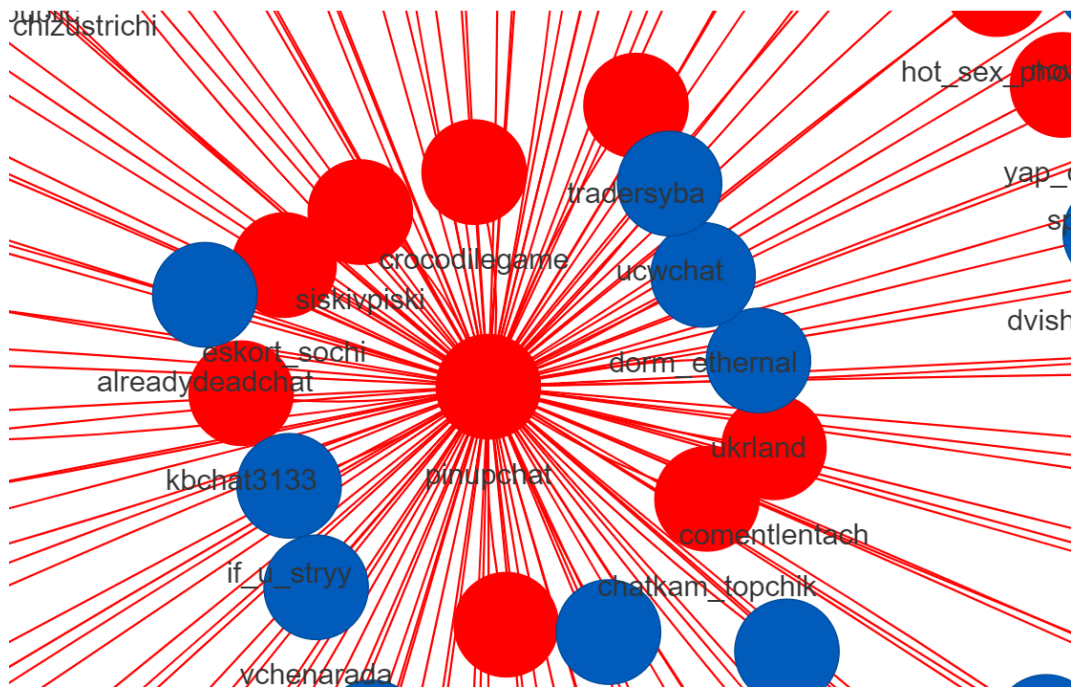
Graf 1: Odber skupín na Telegrame Rusko a Ukrajina

Pri zobrazení skupín ako vrcholov a hrán, ak majú viac ako 50 percent svojich odoberateľov spoločných, nám vznikol graf s 200 vrcholmi, 189 hranami a 49 izolovanými vrcholmi. Červený vrchol predstavuje ruskú skupinu a modrý ukrajinskú skupinu.

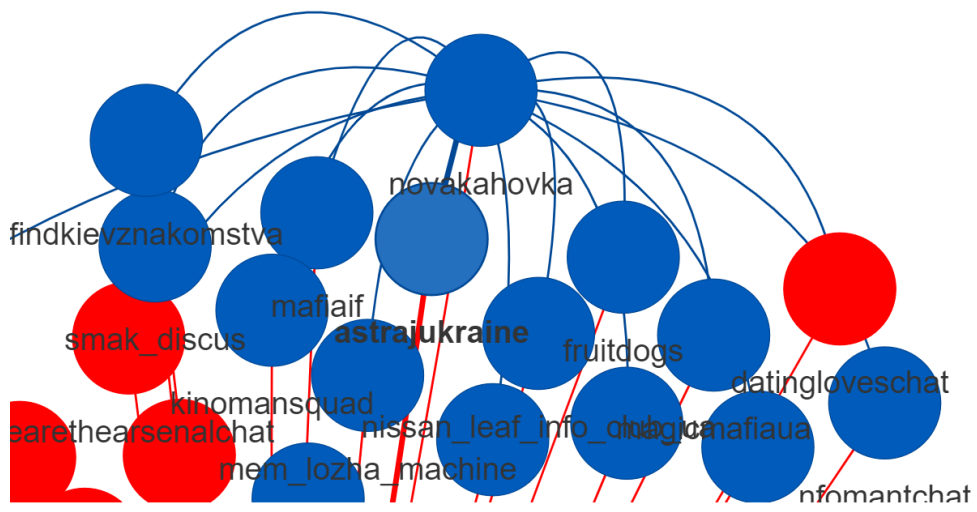


Obrázok 28: Sieť ruských a ukrajinských skupín

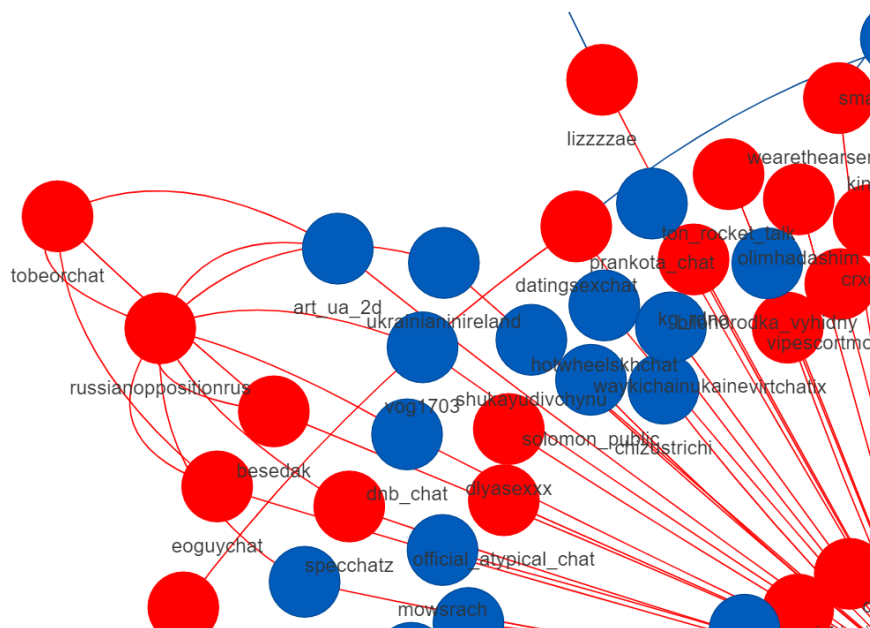
Z pohľadu vlastností graf obsahuje 4 huby. Najvýznamnejší hub tvorí ruská skupina pinupchat, ktorá je spojená so 145 ruskými a ukrajinskými skupinami. Druhým najvýznamnejším hubom je skupina novakahovka, ktorá spája ukrajinské skupiny. Ďalšími hubmi sú toberchat a russianoppositionrus, vrcholy ktorých spájajú ruské skupiny.



Obrázok 29: Hub pinupchat



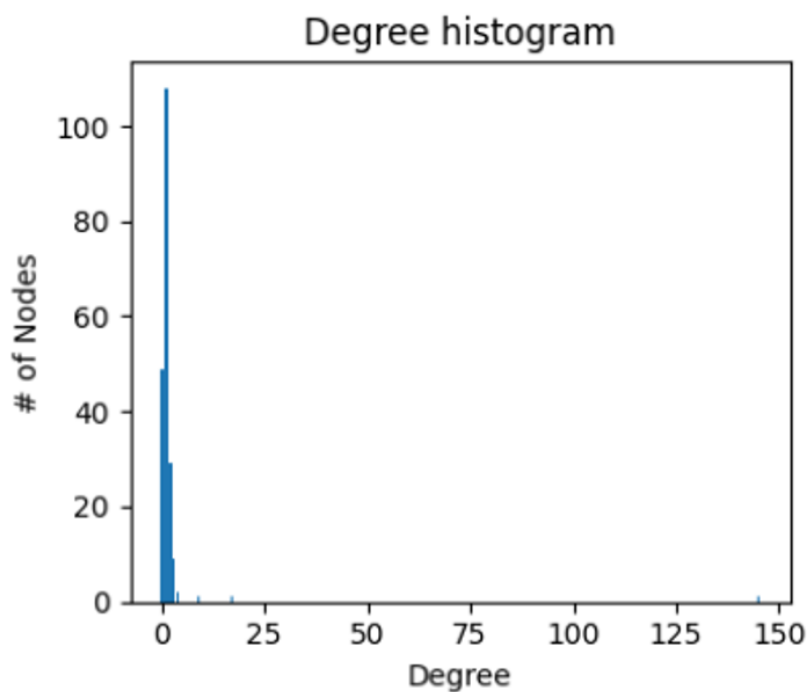
Obrázok 30: Hub novakahovka



Obrázok 31: Huby toberchat a russianoppositionrus

Ďalšou zo skúmaných vlastností je klasterizácia, ktorá nadobúda hodnoty od 0 po 1. V našom prípade je to 0,1335, čo je pomerne malá klasterizácia grafu. Tranzitivita grafu je 0.00786, čo je veľmi malá hodnota. Na histograme môžeme vidieť, že najkratšie cesty sú medzi väčšinou vrcholov menšie ako 6.

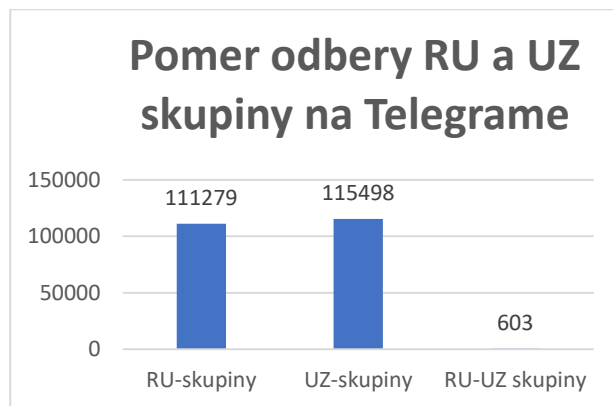
Konštatujeme, že sieť ruských a ukrajinských skupín je na základe nízkej klasterizácie a najkratších ciest medzi vrcholmi najviac podobná bezškálovej sieti. Tento faktor potvrdzuje aj prítomnosť hubu, ktorý spojený s oveľa viac vrcholmi ako ostatné vrcholy.



Obrázok 32: Distribúcia stupňov vrcholov ruské a ukrajinské skupiny

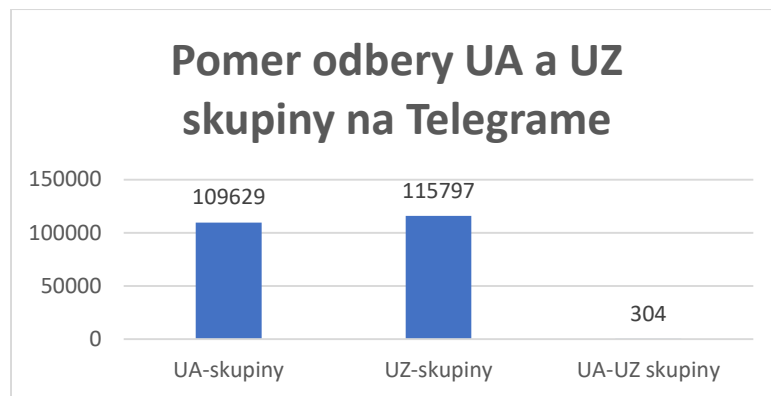
7.1.2 Uzbekistan verzus Rusko a Ukrajina

Zo vzorky zhruba 110000 odoberateľov ruských skupín a 115000 odoberateľov uzbeckých skupín sme identifikovali 603 užívateľov, ktorí odoberajú aj uzbecké aj ruské skupiny, čo tvorí približne 0,5% skúmaných odoberateľov.



Graf 2: Odber skupín na Telegrame Rusko a Uzbekistan

Zo vzorky skoro 110000 odoberateľov ukrajinských skupín a 115000 odoberateľov uzbeckých skupín sme zistili 304 užívateľov, ktorí odoberajú aj uzbecké aj ukrajinských skupiny, čo tvorí približne 0,25% skúmaných odoberateľov.



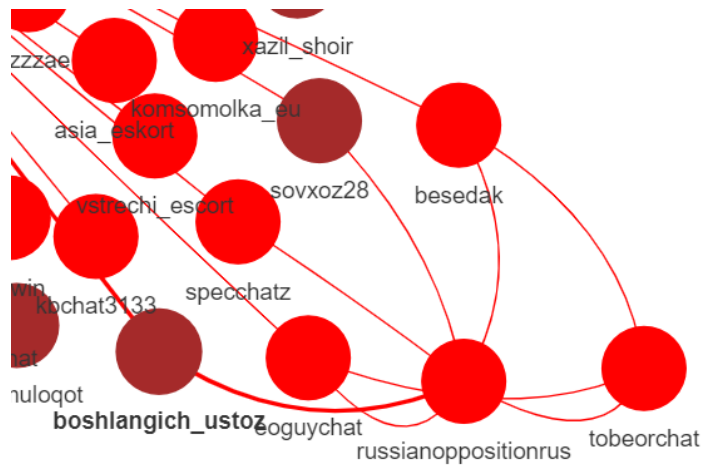
Graf 3: Odber skupín na Telegrame Ukrajina a Uzbekistan

7.1.3 Tadžikistan verzus Rusko a Ukrajina

7.1.3.1 Porovnanie s ruskými skupinami

Zistili sme, že okolo 110000 odoberateľov ruských skupín a 110000 odoberateľov tadžických skupín iba 236 užívateľov odoberá súčasne aj tadžické aj ruské skupiny, čo tvorí približne 0,21% skúmaných odoberateľov.

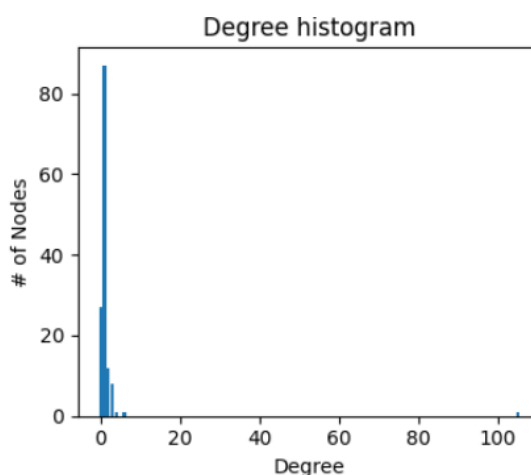
Ďalšími hubmi sú toberchat a russianoppositionrus vrcholy, ktoré spájajú ruské skupiny aj jednu tadžickú skupinu.



Obrázok 34: Hub russianoppositionrus ruské a tadžické skupiny

Ďalšou zo skúmaných vlastností je klasterizácia. V našom prípade je to 0,0812, čo je ešte menšia klasterizácia grafu ako pri ukrajinsko-ruských skupinách. Tranzitivita grafu je 0.00598, čo je opäť veľmi malá hodnota. Na histograme môžeme vidieť, že najkratšie cesty sú opäť medzi väčšinou vrcholov menšie ako 6.

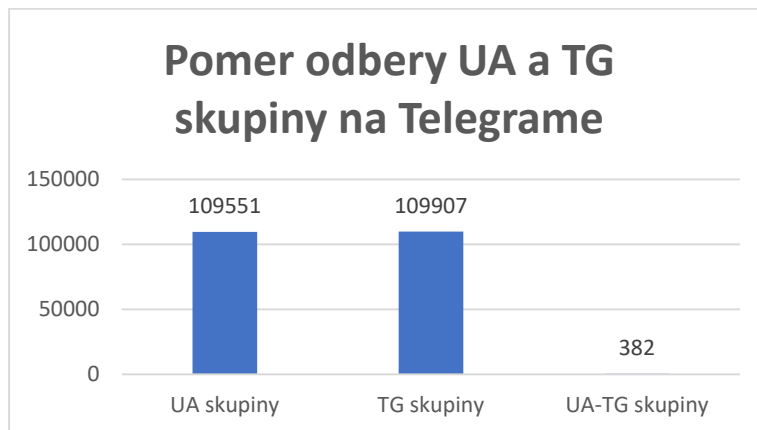
Konštatujeme, že sieť ruských a tadžických skupín je opäť na základe nízkej klasterizácie a najkratších ciest medzi vrcholmi najviac podobná bezškálovej sieti. Tento faktor potvrdzuje aj prítomnosť hubu, ktorý je spojený s oveľa viac vrcholmi ako ostatné vrcholy.



Obrázok 35: Distribúcia vrcholov tadžických a ruských skupín

7.1.3.2 Porovnanie s ukrajinskými skupinami

Zo vzorky okolo 10000 odoberateľov ukrajinských skupín a 110000 odoberateľov tadžických skupín sa našlo 382 užívateľov, ktorí odoberajú aj tadžické aj ukrajinských skupiny, čo tvorí približne 0,34% skúmaných odoberateľov.

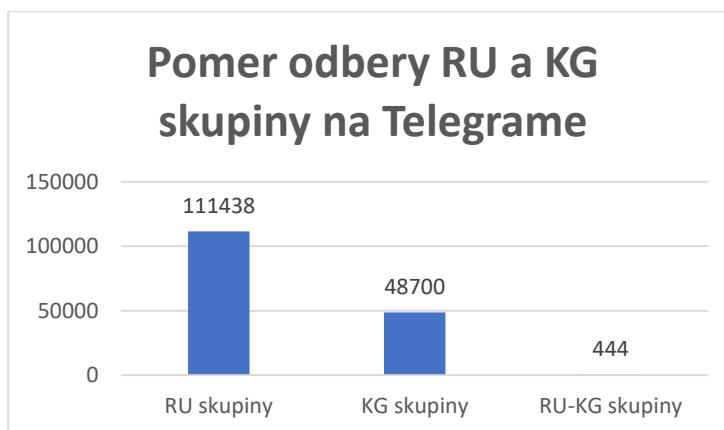


Graf 5: Odber skupín na Telegrame Ukrajina a Tadžikistan

7.1.4 Kirgizsko verzus Rusko a Ukrajina

7.1.4.1 Porovnanie s ruskými skupinami

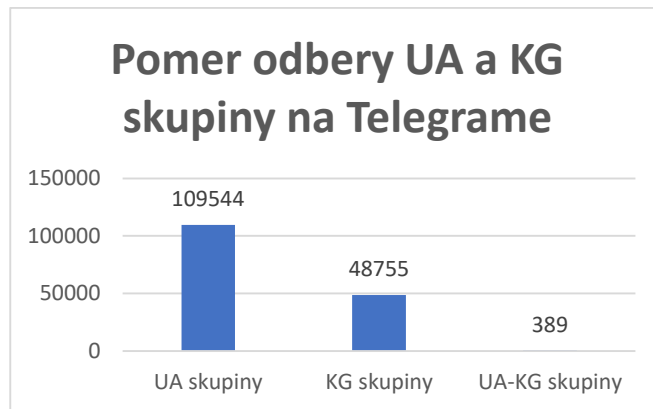
Zo vzorky zhruba 110000 odoberateľov ruských skupín a 50000 odoberateľov kirgizských skupín sa našlo 444 užívateľov ktorí odoberajú aj kirgizské aj ruské skupiny, čo tvorí približne 0,55% skúmaných odoberateľov.



Graf 6: Odber skupín na Telegrame Rusko a Kirgizsko

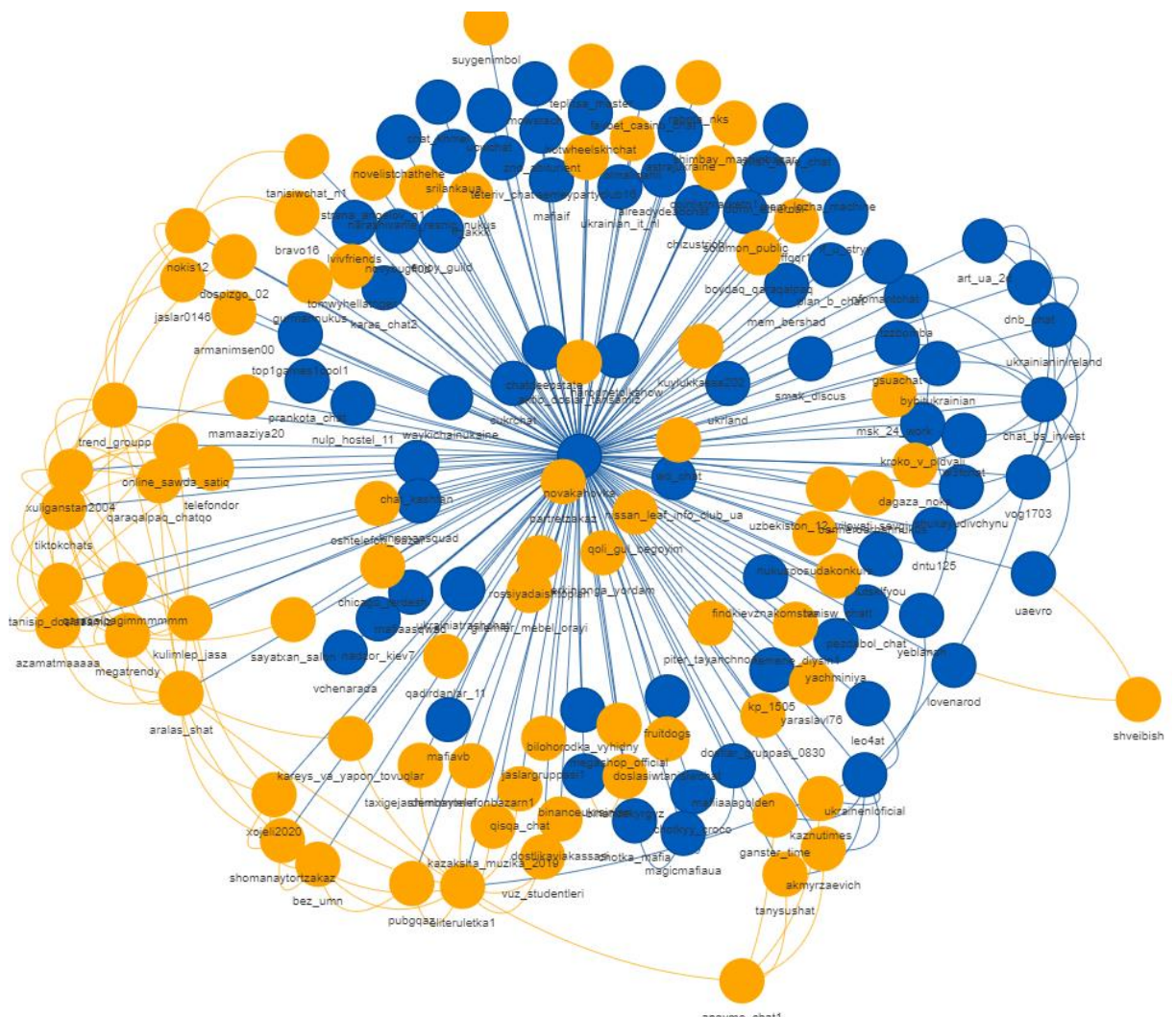
7.1.4.2 Porovnanie s ukrajinskými skupinami

Zo vzorky približne 110000 odoberateľov ukrajinských skupín a 50000 odoberateľov kirgizských skupín sa našlo 389 užívateľov, ktorí odoberajú aj kirgizské aj ukrajinských skupiny, čo tvorí približne 0,48% skúmaných odoberateľov.

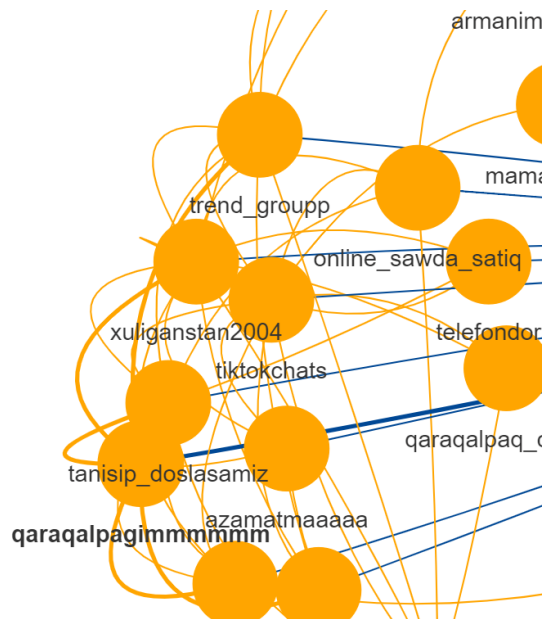


Graf 7: Odber skupín na Telegrame Ukrajina a Kirgizsko

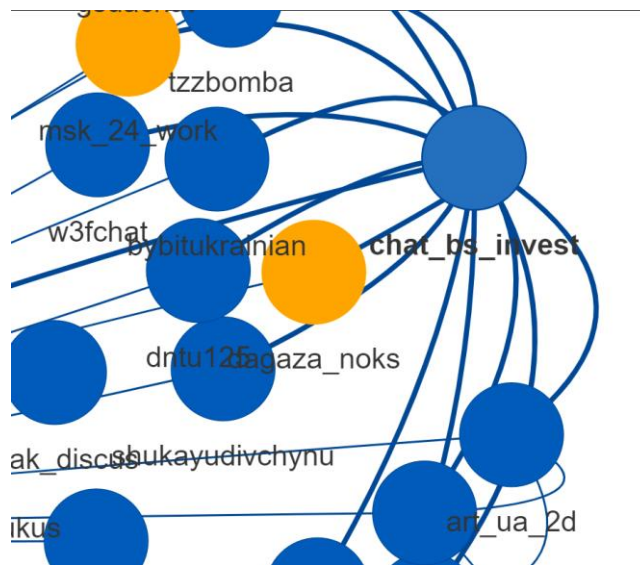
Pri zobrazení skupín ako vrcholov a hrán ak majú viac ako 30 percent svojich odoberateľov spoločných, nám vznikol graf s 191 vrcholmi, 266 hranami, 40 izolovanými vrcholmi. Modrý vrchol predstavuje ukrajinskú skupinu a oranžový zas kirgizskú skupinu.



Obrazok 36: Sieť spoločných odoberateľov kirgizských a ukrajinských skupín



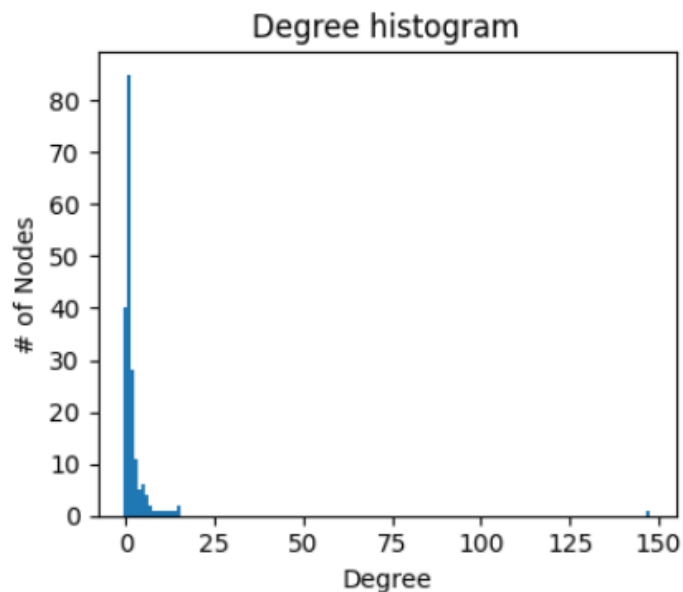
Obrázok 39: Huby trend_group a qaraqalpagimmmmmmm



Obrázok 40: Hub chat_bs_invest

Ďalšou zo skúmaných vlastností je klasterizácia. V našom prípade je to 0,2546, čo je 2 krát väčšia klasterizácia grafu ako pri ukrajinsko-ruských skupinách. Tranzitivita grafu je 0.046, čo je opäť veľmi malá hodnota. Na histograme môžeme vidieť, že najkratšie cesty sú opäť medzi väčšinou vrcholov menšie ako 6.

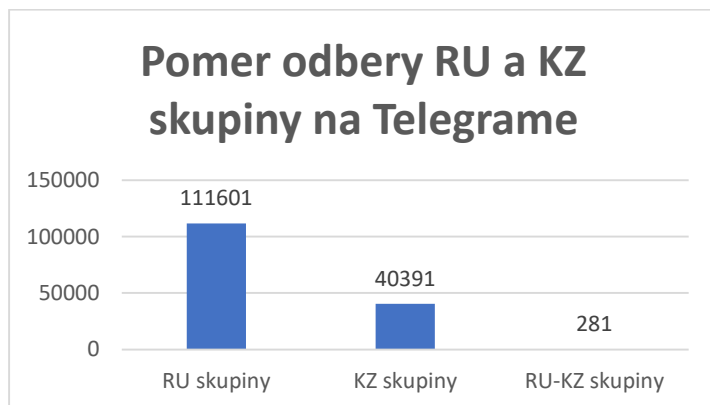
Konštatujeme, že sieť kirgizských a ukrajinských skupín je na základe nižšej klasterizácie a najkratších ciest medzi vrcholmi najviac podobná bezškálovej sieti. Tento faktor potvrdzuje aj prítomnosť niekoľkých hubov, ktoré majú hrany s oveľa viac vrcholmi ako iné.



Obrázok 41: Distribúcia vrcholov kirgizských a ukrajinských skupín

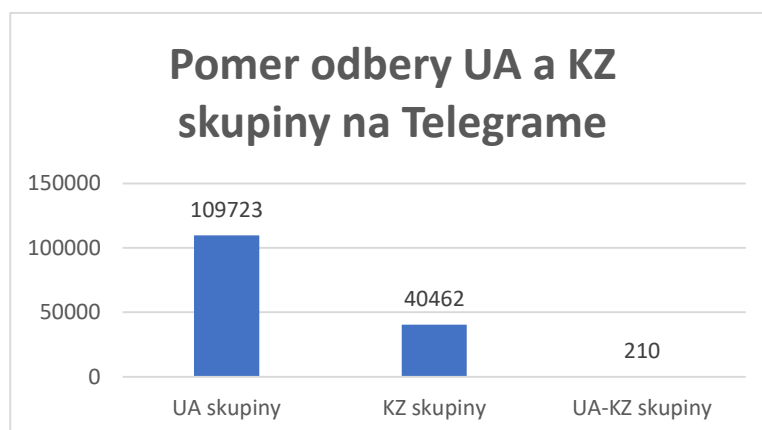
7.1.5 Kazachstan verzus Rusko a Ukrajina

Zo vzorky takmer 110000 odoberateľov ruských skupín a 40000 odoberateľov kazašských skupín sa našlo 281 užívateľov ktorí odoberajú aj kazašské aj ruské skupiny, čo tvorí približne 0,37% skúmaných odoberateľov.



Graf 8: Odber skupín na Telegramme Rusko a Kazachstan

Zo vzorky zhruba 110000 odoberateľov ukrajinských skupín a 40000 odoberateľov kazašských skupín sa našlo 210 užívateľov ktorí odoberajú aj kazašské aj ukrajinských skupiny, čo tvorí približne 0,28% skúmaných odoberateľov.



Graf 9: Odber skupín na Telegrame Ukrajina a Kazachstan

7.1.6 Vyhodnotenie

Z výsledkov môžeme usúdiť, že najväčšími odoberateľmi ruských skupín sú ľudia hovoriaci po ukrajinsky, potom kirgizsky, uzbecky, kazašsky a tadžicky.

Najväčšími odoberateľmi ukrajinských skupín sú ľudia hovoriaci po rusky, a potom kirgizsky, tadžicky, kazašsky, uzbecky.

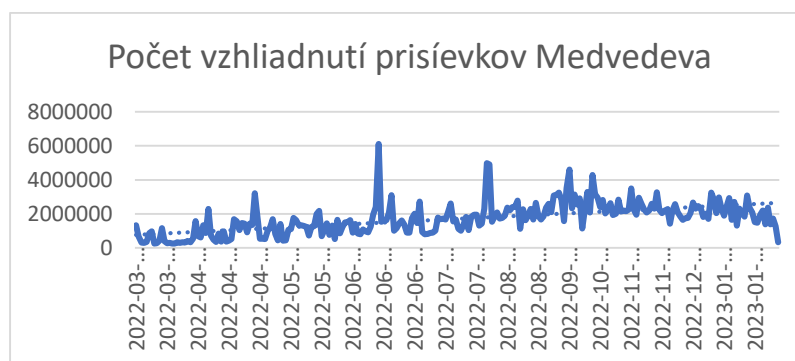
Najväčším prekvapením je prienik medzi Ukrajinskými a Ruskými skupinami je 8 krát väčší než majú ostatné prieniky, ktoré sme vyhodnocovali.

7.2 Analýza sledovanosti dvoch kanálov

Na analýzu sledovanosti kanálov sme si vzhľadom na aktuálnu tému vojny vybrali dve osobnosti na Telegrame s približne miliónom sledujúcich. Prezidenta Ukrajiny Volodymyra Zelenského a bývalého prezidenta Ruskej Federácie Dmitriho Medvedeva.

7.2.1 Analýza sledovanosti príspevkov Dmitriho Medvedeva

Z grafu 11 môžeme jasne vidieť, že jeho sledovanosť a aj popularita rýchlo zvýšila 7. júna 2022, kedy jeho príspevok videlo viac ako 6 miliónov ľudí. Nejednalo sa pritom o žiadnu významnú udalosť na vojnovej fronte s Ukrajinou. Pán Medvedev sa príspevkom snažil prebudiť ruské vlastenectvo. Dňa 27. júla 2022 jeho príspevok sledovalo 5 miliónov ľudí, zámer bol ten istý. Záujem o jeho príspevky od tej doby klesol a k dnešnému dňu predstavuje približne 2 milióny pozretí na príspevok. Ako si môžeme všimnúť uverejňoval s frekvenciou 1 príspevok za deň.



Graf 10: Príspevky Dmitriho Medvedeva

7.2.2 Analýza sledovanosti príspevkov Volodymyra Zelenského

Na rozdiel od p. Medvedeva uverejňuje p. Zelenský svoje príspevky niekoľkokrát za deň. Jeho sledovanosť príspevkov rýchlo stúpila už po druhom týždni začatia vojnového konfliktu medzi Ukrajinou a Ruskom. Jeho najsledovanejšie príspevky si pozrelo skoro 7 miliónov užívateľov. Niektoré sú z obdobia 6.marca 2022 a posledné u respondentov veľmi obľúbené sú z 31.decembra 2022. Priemerne frekvencia sledovanosti jeho príspevkov predstavuje okolo 3 miliónov pozretí na príspevok, čo zobrazuje graf 12.



Graf 11: Príspevky Volodymyra Zelenského

7.3 Frekvencia uverejňovania správ

Ruské aj ukrajinské médiá sú známe tým, že ak dôjde k nejakej zaujímavej udalosti na domácej scéne alebo priamo vojnovej udalosti, začnú publikovať v aplikácii Telegram množstvo filozofických príspevkov. Našou úlohou bolo túto skutočnosť overiť.

Zobrali sme si ruské médium s 2,5 milióna sledovateľov a ukrajinské médium s 1,7 milióna sledovateľov. Skúmali sme počet sekúnd medzi časmi uverejnenia daných príspevkov.

7.3.1 Frekvencia uverejňovania Rian_ru

Zaznamenali sme, že správy boli prvý mesiac vojnového konfliktu medzi Ukrajinou a Ruskom uverejňované každých 5 minút. Neskôr sa frekvencia uverejňovania znížila na každých 10 minút. Najdlhšie sa nezverejňovalo 6 hodín. V priemere sa uverejňovalo každých 11 minút a 3 sekundy.

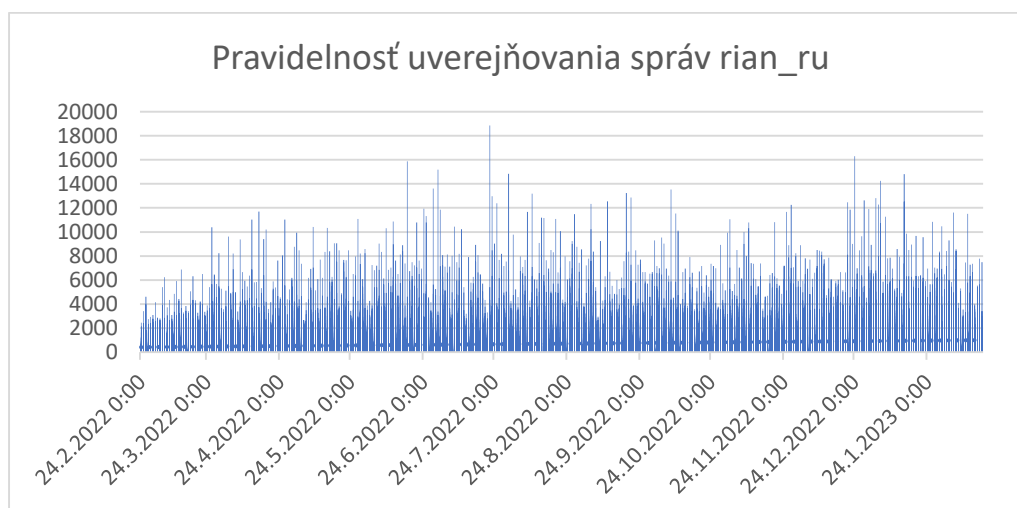
Jednou z najvýznamnejších udalostí bolo potopenie krížnika Moskva 14. apríla 2022. V týchto dňoch bola frekvencia publikovania správ každých 11 minút a 40 sekúnd. Pohybovalo sa slabo nad priemerom. Počet publikovaných správ bol 138, čo predstavoval slabý nadpriemer.

Ďalšou udalosťou bolo ukrajinská protiofenzíva na Donbase 10. septembra 2022. V tomto prípade rian_ru uverejňovalo správy každých 21 minút, čo je menej, než priemer. Aj počet správ bol iba 40, čo bolo 3 krát menej než priemer. Ďalšie dni bol priemerný ruch.

Anektovaním ukrajinského územia Ruskom dňa 30. septembra 2023 bolo publikovaných 196 správ, čo predstavuje o 70 viac než priemer. Pravidelnosť uverejňovania bola každých 10 minút 55 sekúnd, čo je málo nad priemerom. Táto udalosť o niečo viac zaujala média.

Poslednou významnejšou udalosťou bol výbuch Kerčského mosta dňa 8. októbra 2022.

V týchto dňoch sa zverejnilo okolo 173 správ, čo je nadpriemer. Pravidelnosť uverejňovania správ bola každých 11 minút 54 sekúnd, teda skoro priemerne uverejňovanie.



Graf 12: Frekvencia uverejňovania správ na Rian_ru

Po zhodnotení, konštatujeme, že média viac reagovali iba prvý mesiac vojnového konfliktu Ukrajina a Rusko. Potom záujem o vojnový konflikt u médií opadol. Ďalšie

významné udalosti záujem médií neovplyvnili. Naopak pri ukrajinskej protiofenzíve záujem opadol a anexia s výbuchom Kerčského mosta len jemne zvýšila produkciu správ.

7.3.2 Frekvencia uverejňovania U_now

Zaznamenali sme z grafu 13, že za prvé tri mesiace vojnového konfliktu Ukrajiny a Ruska sa uverejňovalo každé 3 minúty. Neskôr sa interval zväčšil na 10,5 minúty. U_now uverejňoval správy v priemere každých 7 minút a 36 sekúnd. Počet uverejňovaných správ za jeden deň bol priemerne 187.

V prípade udalosti potopenie krížnika Moskva 14. apríla 2022, bola perióda publikovania správ každé 4 minúty 25 sekúnd. Čo bolo dvakrát častejšie ako obyčajne. Počet textových správ bolo publikovaných 396, čo bolo viac ako dvojnásobok denného priemeru zverejňovania.

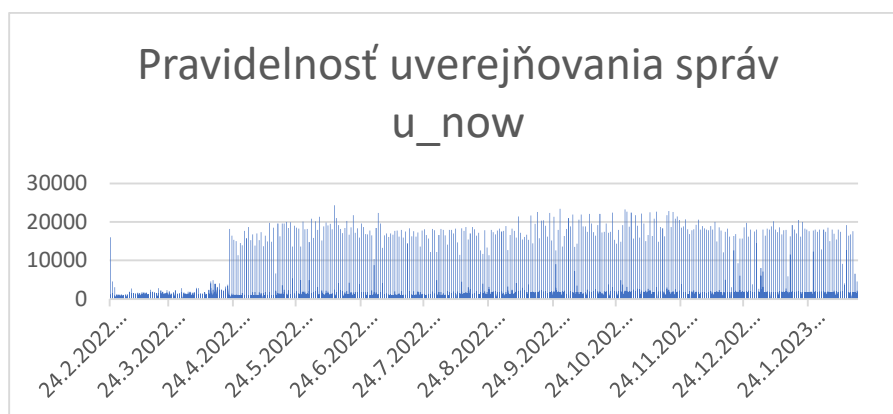
Pri ukrajinskej protiofenzíve na Donbase 10. septembra 2022 respondenti uverejňovali správy každých 13 minút 53 sekúnd, čo je dvakrát menej, než priemer. Počet postov bol iba 108, čo bolo veľmi málo oproti inými udalosťami vojnového konfliktu. Tu mohlo nastať mediálne ticho, aj najbližšie dni tak bolo, aby sa zamýšľaná operácia podarila.

Anektovaním ukrajinského územia Ruskom dňa 30.septembra 2023 bolo publikovaných 226 správ, o 40 viac než bol priemer. Pravidelnosť uverejňovania bola každých 7 minút 22 sekúnd, čo je priemer. Táto udalosť médiá málo zaujala.

Poslednou významnejšou udalosťou bol výbuch Kerčského mosta dňa 8.októbra 2022.

V týchto dňoch sa zverejnilo na U_now okolo 253 správ, teda nad priemer. Pravidelnosť uverejňovania správ bola každých 7 minút, čo je slabo nad priemerom.

Konštatujeme teda, že médiá citlivejšie reagovali prvé dva mesiace vojnového konfliktu Ukrajiny a Ruska. Neskôr záujem o vojnový konflikt u médií upadol. Potopenie krížnika médiá zaujalo oveľa viac. Naopak ukrajinská protiofenzíva z taktického hľadiska vyvolala mediálne ticho, anexia neprekročila priemer a výbuch Kerčského mosta spôsobil viac publikovaných správ.



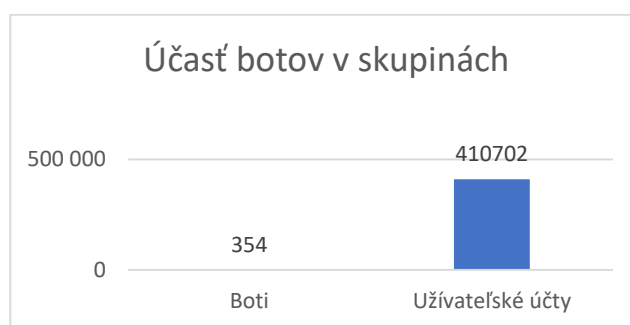
Graf 13: Frekvencia uverejňovania správ na U_now

7.3.3 Zhodnotenie

Z grafu 13 a 14 vidíme, že ruské médiá reagovali s ukrajinskými médiami rovnako citlivo prvý mesiac konfliktu a následne pri výbuchu Kerčského mosta. Neskôr si našli svoj priemer a reagovali na udalosti takticky, podľa záujmu verejnosti o túto tematiku. Rozdiel bol jediný: ukrajinské médiá na rozdiel od ruských reagovali intenzívnejšie na potopenie krížnika v Čiernom mori .

7.4 Účasť botov v skupinách

Sieť robotizovaných účtov, ktoré organizovane zdieľajú, komentujú príspevky, reagujú na komentáre navzájom, považujeme za veľmi populárny trend v používaní sociálnych sietí. Aplikácia Telegram na túto činnosť vymyslela bot účty, ktoré môže každý užívateľ za určitých pravidiel používať. My sme sa preto snažili zistiť, koľko bot účtov sa v skupinách nachádza.

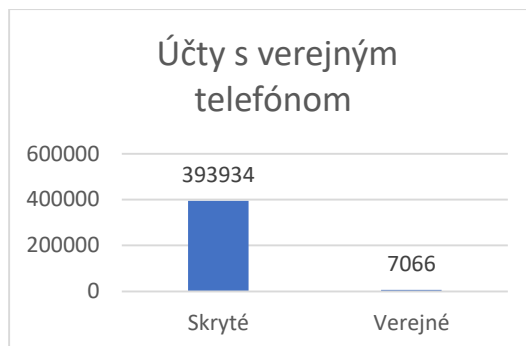


Graf 14: Účasť botov v skupinách

Na grafe 14 pozorujeme, že zo vzorky 411006 členov asi 200 rôznych skupín bolo iba 354 bot účtov. Čo predstavuje zanedbateľnú časť, ani nie desatinu percenta. Samozrejme existujú aj bežné užívateľské účty, ktoré sú spravované automatizovane a teda sa chovajú ako neoficiálni boti. Takéto účty sme nevyhodnocovali.

7.5 Viditeľnosť telefónneho čísla

Pri založení účtu Telegram, je telefónne číslo verejné, ak ho chce používateľ skryť, musí zmeniť nastavenia svojho profilu. Pri sťahovaní informácií o členoch skupín sme zistili, že len malá časť používateľov má telefónne číslo v profile zverejnené.



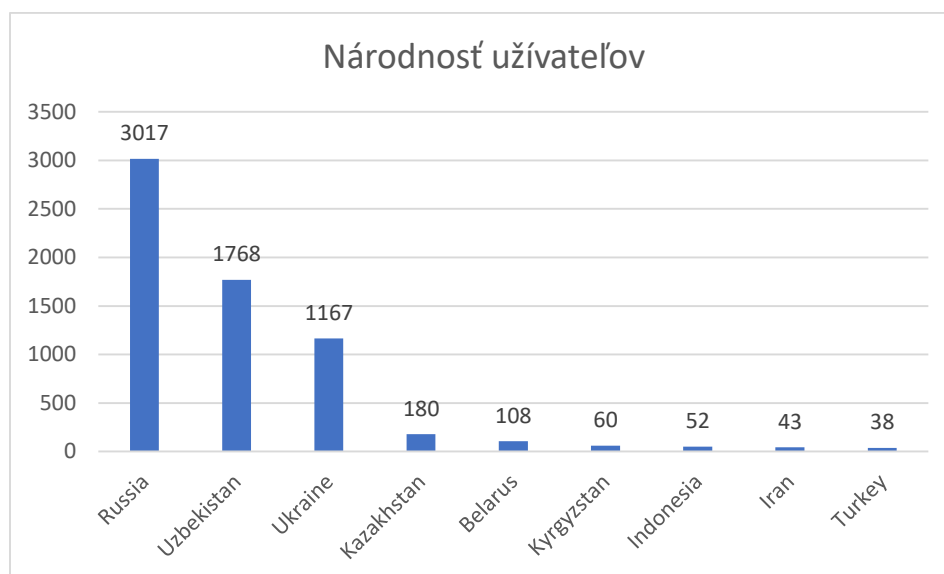
Graf 15: Nastavenia telefónneho čísla na účte

Zo vzorky 401000 účtov sme zistili, že 393934 užívateľov má skryté telefónne číslo a 7066 má zverejnené telefónne číslo, čo predstavuje 1,7%.

Podľa aktuálnych dostupných informácií aplikácia Telegram naďalej vyžaduje prihlásenie pomocou telefónneho čísla alebo identity zakúpenej za kryptomeny.[30]

7.5.1 Národnosť užívateľov

Na základe zoznamu asi 7-tisíc čísel a telefónnej predvoľby 180 krajín sa nám podarilo zistiť krajinu, v ktorej majú používatelia registrované telefónne číslo. Najskôr sme sa snažili získať údaje pomocou Selenium knižnice, ktorá automatizovane za nás získavala údaje z verejne dostupného web servera v internete. Tento spôsob sa ukázal nakoniec ako časovo neefektívny, nakoľko nás web server po niekoľkých stovkách požiadaviek zablokoval. Preto sme neskôr použili Python knižnicu phonenumbers, ktorá nám národnosť tisícok telefónnych čísel prezradila behom pár sekúnd.



Graf 16: Nastavenia telefónneho čísla na účte

Zo vzorky 7063 telefónnych čísel sa nám na základe predvoľby podarilo zistiť krajinu registrácie až 7027 užívateľov čo predstavuje 99,4 %. Najviac zistených užívateľov 3017 má telefónne číslo registrované v Rusku. Následne sú to Uzbeki s 1768 účtami a poslednou krajinou sú Ukrajinci s 1167 užívateľmi. Aj telefónne čísla potvrdzujú, že rusko-jazyčných skupín je na Telegramе najviac zo všetkých krajín sveta.

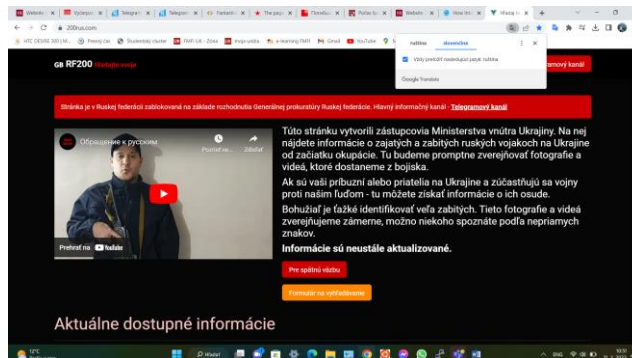
7.6 Komunikácia ruských vojakov na Telegramе

Vzhľadom na situáciu vo svete sa prevažná komunikácia Rusov na sociálnych sieťach sústreďuje práve na Telegramе. Média častokrát prinášajú informácie na základe príspevkov ruských blogerov, politických a vojenských osobností. Naším cieľom bolo zistiť, či počas vojenského ťaženia, kedy prišlo o život viacero vojakov, je možné zachytiť túto udalosť aj na Telegramе.

Prvým krokom bolo získať z nezávislých zdrojov, mená vojakov. Ruská vláda však zakázala ruským médiám zverejňovať mená mŕtvych vojakov.[27] Ani miestne úrady nezverejňovali mŕtvych. Vplyvom týchto podnetov bol vytvorený projekt Look for your own Ukrajinským ministerstvom vnútra. Cieľom bolo poskytnúť fotografie, videá a akékoľvek informácie rodinám, ktorých najbližších do vojnu povolali.[28] Ruská vláda však opätovne zakázala prístup svojim ľuďom k tejto stránke.[29] Dnešným dňom má projekt skôr symbolický charakter. Je na ňom uverejnených pár videí a fotiek. Na získanie nejakých špecifických informácií musí človek vyplniť formulár na základe, ktorého mu informácie budú poskytnuté.

Hľadali sme aj na iných zdrojoch, podarilo sa nám získať malú vzorku zhruba asi 500 mien mŕtvych ruských vojakov, pre cieľ nášho výskumu to však nebolo postačujúce. Treba

podotknúť, že jeden zo zdrojov časom tieto mená vymazal. Presné informácie o počte padlých majú v rukách iba štátne orgány. Bežní ľudia sa k nim normálnym spôsobom nedostanú.



Obrázok 42: Webová stránka Look for your own

7.7 Scraping dát na Instagrame

Scraping údajov na Instagrame vzhľadom na meniace sa bezpečnostné podmienky spoločnosti Meta sa stáva čoraz náročnejším. Naším cieľom bolo vytvoriť profil zameraný na domácich miláčikov (zvieratká) a pomocou rôznych stratégií profil popularizovať a analyzovať nových sledovateľov.

Pri zisťovaní sledovateľov a sledujúcich sme používali knižnice Instagrami. Pri každom spustení programu sme sa museli prihlásiť instagramovým účtom. Získanie zoznamu sledovateľov bolo veľmi pomalé z dôvodov hrozby zablokovania účtu.

Výsledkom celého úsilia nakoniec bolo zablokovanie účtu a IP adresy na dobu 48 hodín. Zároveň ak sme sa pokúsili prihlásiť na hociktorý účet cez zablokovanú IP adresu, prihlásenie nám bolo následne znepriístupnené. Instagram je veľmi citlivý na časté prihlásenia a zároveň má nastavené počty údajov, ktoré môžeme stiahnuť. V porovnaní s Telegramom je Instagram oveľa lepšie chránený voči scrapovaniu dát.

8 Záver

Naším cieľom bolo analyzovať sociálne siete. Zistili sme, že vzhľadom na neustále zlepšovanie ich zabezpečenia to nie je jednoduché. Sociálne siete automaticky blokujú účty, ktoré sa snažia získať väčšie množstvo informácií, dokážu tiež blokovat' IP adresy z ktorých sa podozrivé účty prihlasujú. Preto sme oproti pôvodnému zámeru analyzovať viacero sociálnych sietí väčšinu analýz realizovali na sieti Telegram, ktorá je k týmto činnostiam nateraz tolerantnejšia. Pre hlbšiu analýzu by boli potrebné sofistikovanejšie nástroje, napríklad také, ktoré sú schopné distribuovať zber dát v čase a cez viaceré servery s rôznymi IP adresami. Napriek obmedzeniam sme dokázali zistiť zaujímavé štatistiky ohľadom sledovanosti kanálov, dostupných telefónnych čísel používateľov a spoločných sledovateľov viacjazyčných skupín. Na druhej strane máme sociálne siete ako Facebook a Instagram, ktoré sa zabezpečením posunuli výrazným smerom dopredu za posledné roky. Môžeme to vidieť na diplomovej práci od Kristíny Dvorskej, ktorá dokázala rok dozadu získať zaujímavé údaje z Instagramu. Dnes už také informácie nie sú dostupné.

Zlepšenie zabezpečenia sociálnych sietí je pozitívny trend, zostáva nám dúfať, že aj my ako používatelia sa budeme chovať a publikovať zodpovedne, pretože čokoľvek necháme na sociálnych sieťach môže byť zneužitú proti nám.

ZOZNAM POUŽITEJ LITERATÚRY

[1]

Saqib Shah - The history of social networking

(30.1.2023)

Dostupné na: <https://www.digitaltrends.com/computing/the-history-of-social-networking/>

[2]

EdgeRank: Give your Facebook page the edge

(2.2.2023)

Dostupné na: <https://www.wordtracker.com/academy/social/facebook/edgerank-facebook-ranking>

[3]

C.C. Aggarwal. Social Network Data Analytics. Springer US, 2011.

(2.2.2023)

[4]

Wesley Szuway Shu, Yu-Hao Chuang – Percieved Benefits of 6-Degree-Separation Social Networks, 2011

(2.2.2023)

[5]

Chenda Ngak – Then and now: a history of social networking sites 2011

(2.2.2023)

Dostupné na: <https://www.cbsnews.com/pictures/then-and-now-a-history-of-social-networking-sites/>

[6]

Christine Moellenberndt- Livejournal loyalty and melodrama: Stakeholder relations in web 2.0, May 2013

(2.2.2023)

Dostupné na:

https://www2.sjsu.edu/anthropology/docs/projectfolder/Moellenberndt_Christine_thesis.pdf

[7]

Andrew Smock Leveraging social media for learning communities of practice on flickr, Proquest, 2012

(2.2.2023)

[8]

(2.2.2023)

Dostupné na: <https://about.linkedin.com/>

[9]

Lisa Eadicicco - LinkedIn CEO:Here's Why I Sold the Company to Microsoft, 2016

(2.2.2023)

Dostupné na: <https://time.com/4366106/linkedin-microsoft-why/>

[10]

W.David Albrecht – LinkedIn for Accounting And Business Students, The clute Institute,

2011

(2.2.2023)

[11]

Balakrishnan, V. K. (1997). Graph Theory (1st ed.). McGraw-Hill. ISBN 978007-005489-9

(14.2.2023)

[12]

Rina More, Archana Patil, Depak Dandwate, Umesh Tupe - Theory in Various Fields,

International Journal for Research Trends and Innovation, 2021

(14.2.2023)

Dostupné na: <https://ijrti.org/papers/IJRTI2101001.pdf>

[13]

How to hide the list of group members from your Telegram groups

(14.2.2023)

Dostupné na: <https://www.techmesto.com/hide-list-of-telegram-group-members/>

[14]

Lenka Bazalová, Pohľad na sociálne siete a využívanie

(20.2.2023)

Dostupné na: https://is.muni.cz/th/fhfak/DS_Bazalova.pdf

[15]

Jana Chvátalová, Phishing

(20.2.2023)

Dostupné na: https://is.ambis.cz/th/dd18i/BP_Jana_Chvatalova_Archive.pdf

[16]

Sociálne inžinierstvo

(20.2.2023)

Dostupné na: <https://www.csirt.gov.sk/socialne-inzinierstvo.html>

[17]

(20.2.2023)

Dostupné na: <https://www.europarl.europa.eu/factsheets/sk/sheet/157/ochrana-osobnych-udajov>

[18]

(28.2.2023)

Dostupné na: <https://refresher.sk/128556-Zuckerbergova-Meta-porusila-nariadenie-EU-o-osobnych-udajoch-Musi-zaplatit-pokutu-390-milionov-eur>

Dostupné na: <https://tvnoviny.sk/zahranicne/clanok/279810-google-zaplaci-pokutu-viac-ako-4-miliardy-eur-potvrdil-to-uz-aj-vseobecny-sud-eu>

[19]

David Lancefield - Department for Culture, Media & Sport, 2011

(28.2.2023)

Dostupné na:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/77641/PwC_Internet_Cookies_final.pdf

[20]

Ochrana osobných údajov

(28.2.2023)

Dostupné na: https://european-union.europa.eu/cookies_sk

[21]

Tomáš Mahrík - Cookie lišty, 2022

(28.2.2023)

Dostupné na: <https://www.webhostingcentrum.sk/cookie-listy-2022/>

[22]

Michal Reiter – Telegram predstavuje riziko. Tu je pravda o Telegram , 2022

(5.3.2023)

Dostupné na: <https://touchit.sk/telegram-predstavuje-riziko-tu-je-pravda-o-telegram/408095>

[23]

Ján Hainc - Bezpečnosť detí na internete, 2018

(5.3.2023)

Dostupné na: <https://is.ambis.cz/th/z4dmk/BP-Hainc-Jan.pdf>

[24]

Viktoría Kudyn - Fenomén sociální sítě TikTok a její vliv a vnímání v Česku

(5.3.2023)

Dostupné na:

<https://dspace.cuni.cz/bitstream/handle/20.500.11956/124399/130299878.pdf?sequence=1&isAllowed=y>

[25]

Katarína Šimková – Bezpečnosť siete TikTok nie je jediným problémom: Prečo je téma blokovania sociálnej siete na pretrase?, 2023

(5.3.2023)

Dostupné na: <https://vosveteit.zoznam.sk/bezpecnost-siete-tiktok-nie-je-jediny-problemom-preco-je-tema-blokovania-socialnej-siete-na-pretrase/>

[26]

Richard Zliechovsky – TikTok vás môže sledovať aj v prípade, ak jeho aplikáciu nepoužívate

(20.3.2023)

Dostupné na: <https://vosveteit.zoznam.sk/tiktok-vas-moze-sledovat-aj-v-pripade-ak-jeho-aplikaciju-nepouzivate/>

[27]

(20.2.2023)

Dostupné na: <https://msk1.ru/text/incidents/2022/06/01/71377118/>

[28]

Kateryna Tyschenko – Russia blocks website about captured and killed RF soldiers

(20.3.2023)

Dostupné na: <https://www.pravda.com.ua/eng/news/2022/02/27/7326424/>

[29]

Roskomsvoboda - How Internet censorship changed in Russia during the 1st year of military conflict in Ukraine

(20.3.2023)

Dostupné na: <https://ooni.org/post/2023-russia-a-year-after-the-conflict/>

[30]

Umar Shakir - Now Telegram users don't need a phone number — they can buy a fake one with crypto

(6.4.2023)

Dostupné na: <https://www.theverge.com/2022/12/7/23498236/telegram-fragment-phone-number-crypto-verification-monetization>

[31]

Selenium test je nástroj na automatizáciu testovania webových aplikácií

(10.4.2023)

Dostupné na: https://msg-life.sk/clanky/digitalizacia/selenium-it-tester/?gclid=CjwKCAjw586hBhBrEiwAQYEnHeIAmP71nCmY4GBbgP6xr7y-PmKase5B_8mnBRga505Ub6uYWluJixoCtUcQAvD_BwE

[32]

(16.4.2023)

Dostupné na: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

[33]

(16.4.2023)

Dostupné na: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

[34]

Agung Purnomo – Instagram Literature: Insights from Scientometric Application, ResearchGate, 2020

(29.4.2023)

Dostupné na:

https://www.researchgate.net/publication/344461448_Instagram_Literature_Insights_from_Scientometric_Application/link/5f789c74a6fdcc0086556392/download

[35]

Telegram: Know the difference between channels and groups

(29.4.2023)

Dostupné na: <https://www.indiatoday.in/information/story/what-is-the-difference-between-telegram-channel-and-telegram-group-1850196-2021-09-07>

[36]

(29.4.2023)

Dostupné na: <https://core.telegram.org/>

[37]

Martin Hodás - Z univerzít do celého sveta: Facebook má 19 rokov. Ako začínal?

(30.1.2023)

Dostupné na: <https://zive.aktuality.sk/clanok/138029/facebook-ma-19-rokov-z-univerzit-do-celeho-sveta/>

[38]

Ako a kým bol vytvorený "telegram"

(30.1.2023)

Dostupné na: <https://sk.ateasyday.com/articles/programs/kak-i-kem-sozdavalsya-telegram.html>

[39]

BLAKE, Author Troy. TIOBE Index for January 2021, 2021-01-18

(30.1.2023)

[40]

(4.3.2023)

Dostupné na: www.combot.org .

[41]

(13.3.2023)

Dostupné na: <https://www.privacyaffairs.com/gdpr-fines/>

[42]

Bjarke Monsted, Sune Lehman, Morten Morup - Social Spreading in Complex Networks, 2019

(5.5.2023)

[43]

(18.5.2023)

Angel Garrido – A Survey on Complex Networks, 2011

[44]

(22.5.2023)

R. Albert and A.L. Barabasi – Statistical mechanics of complex networks, The American Physical Society, 2002

Zdroje použitých obrázkov

Obrázok 2

(5.3.2023)

Zdroj: <https://vosveteit.zoznam.sk/nova-internetova-zoznamka-facebook-dating-je-oficialne-dostupnou-aj-v-europe/>

Obrázok 3

(5.3.2023)

Zdroj: https://techcrunch.com/wp-content/uploads/2020/10/Apple_clips-update-iphone12-hero_10282020.jpg?w=1390&crop=1

Obrázok 5

(16.3.2023)

Zdroj: <https://play->

lh.googleusercontent.com/wlwY1vowGxTrvQMRDHJD21iYBG7S_E09QbKPY-L6dqEc4UxCi0fDQiNuCBonFYED1yU=w2560-h1440-rw

Obrázok 11

(18.3.2023)

Zdroj:

https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.algorithmmy.net%2Farticle%2F1369%2FGraf&psig=AOvVaw1QSszlOdm_BZM50G4xH9AF&ust=1680347511428000&source=images&cd=vfe&ved=0CBAQjRxqFwoTClkyraEhv4CFQAAAAAdAAAAABAJ

Obrázok 13

(9.4.2023)

Zdroj:

https://www.google.com/url?sa=i&url=https%3A%2F%2Fsupport.microsoft.com%2Fsk-sk%2Foffice%2Fprezentovanie-%25C3%25BAadajov-v-st%25C4%25BApcovom-grafed89050ba-e6b6-47de-b090-e9ab353c4c00&psig=AOvVaw0Y229V_SO1bWrJBVi0ePGn&ust=1680347587097000&source=images&cd=vfe&ved=0CBAQjRxqFwoTCKii5tqEhv4CFQAAAAAdAAAAABAE

Obrázok 14

(20.4.2023)

Zdroj:

https://upload.wikimedia.org/wikipedia/commons/thumb/8/8f/Watts_strogatz.svg/1280px-Watts_strogatz.svg.png

Obrázok 15

(25.4.2023)

Zdroj:

https://upload.wikimedia.org/wikipedia/commons/thumb/a/a8/Barabasi-albert_model_degree_distribution.svg/1280px-Barabasi-albert_model_degree_distribution.svg.png

Obrázok 16

(2.5.2023)

Zdroj:

https://mathinsight.org/media/image/image/small_network_hub.png

Obrázok 17

(15.5.2023)

Zdroj:

<https://www.researchgate.net/publication/330552912/figure/fig1/AS:718078228840448@1548214599145/Fully-connected-network-where-each-country-is-represented-as-a-node-and-edges-indicate.ppm>