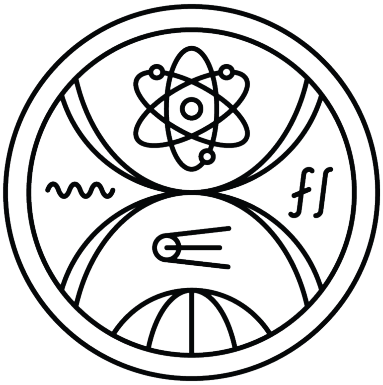
Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky



Analýza aktivít botov na sociálnych sieťach

Diplomová práca

2023
Richard Nagy

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Analýza aktivít botov na sociálnych sieťach

Diplomová práca

Študijný program:	Aplikovaná informatika
Študijný odbor:	Informatika
Školiace pracovisko:	Katedra aplikovanej informatiky
Školiteľ:	RNDr. Damas Gruska, PhD.

Bratislava, 2023
Richard Nagy

ZADANIE ZÁVEREČNEJ PRÁCE

[download this from AIS]

Poďakovanie: You can thank anyone who helped you with the thesis here (e.g. your supervisor).

Abstrakt

Táto diplomová práca sa zameriava na analýzu bezpečnosti sociálnych sietí voči botom vyvíjaným v Pythone a Seleniu. Boti sú rozdelení do troch generácií, pričom každá generácia predstavuje 5 najčastejších typov zlých botov na sociálnych sieťach a ich postupnú evolúciu. Cieľom je posúdiť efektívnosť obranných mechanizmov sociálnych sietí proti týmto botom. Práca poukazuje na riziká spojené s nevhodným využívaním botov, ako sú šírenie dezinformácií alebo neoprávnené zbieranie údajov, a obsahuje prehľad vývoja sociálnych sietí a botov, ich typov, metód detekcie a etických a právnych dilem.

Kľúčové slová: Boti, Sociálne siete, Python, Selenium, Kybernetická bezpečnosť, Detekcia botov, Analýza dát

Abstract

This thesis focuses on the security analysis of social networks against malicious bots developed in Python and Selenium. It categorizes bots into three generations, with each representing the five most common types of malicious bots on social networks and their evolutionary progress. The goal is to evaluate the effectiveness of social networks' defensive mechanisms against these bots. The study underscores the risks of improper bot usage, such as spreading disinformation or unauthorized data collection, and provides an overview of social networks and bots development, including their types, detection methods, and ethical and legal issues.

Keywords: Bots, Social Networks, Python, Selenium, Cybersecurity, Bot Detection, Data Analysis

Obsah

Úvod

V dnešnej digitálnej dobe sú sociálne siete neoddeliteľnou súčasťou nášho každodenného života. S tým sa však spájajú výzvy v oblasti zabezpečenia a integrity online platforiem. Jedným z kľúčových problémov, ktoré vyplynuli z rastúcej popularity sociálnych médií, je rastúci počet automatizovaných skriptov, známych ako 'boti'. Títo boti môžu vykonávať mnoho funkcií, od neškodných činností, ako je napr. pomoc pri navigácií na stránke alebo jednoduchá komunikácia s používateľom, až po sofistikované útoky smerujúce k narušeniu zabezpečenia a manipulácii verejného názoru.

V tejto diplomovej práci sa zameriavame na testovanie a analýzu bezpečnosti sociálnych sietí proti rôznym botom, s cieľom zistiť, ako efektívne sú existujúce obranné mechanizmy. Budeme vytvárať vlastných botov a pozorovať, ako na ne tieto mechanizmy reagujú. Význam tejto práce je zvýšený rastúcimi rizikami spojenými s nesprávnym alebo neetickým použitím botov, ktoré môžu byť využité napr. na šírenie dezinformácií, kyberšikanu, alebo na neautorizované zhromažďovanie osobných dát.

Tu bude rozbor jednotlivých kapitol.

Put your introduction here. By the way, you can cite articles [1, 2], books [3], websites [3], manuals [4, 3, 5], or other resources. You can reference other chapters, sections, images, etc. using the `cleveref` package like this: **??**.

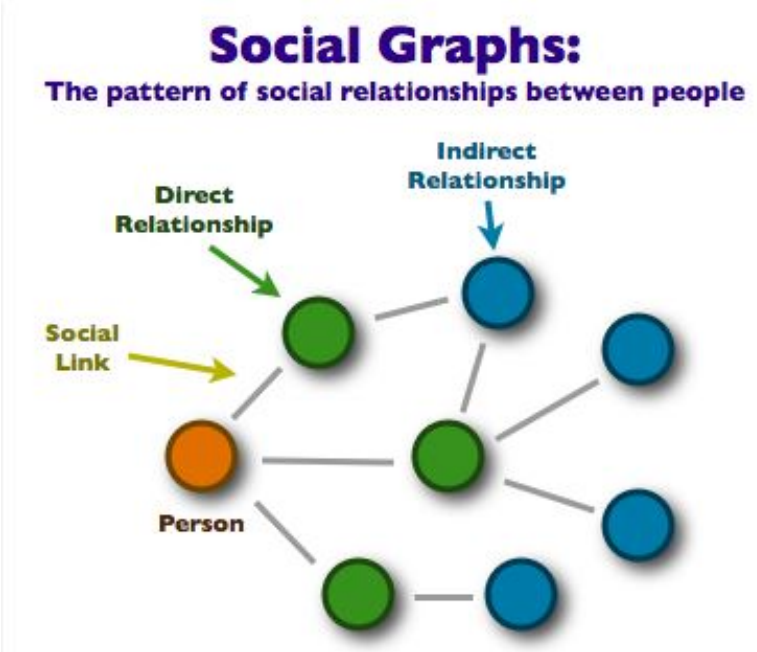
TODO: dokoncit vycuc kapitol

Kapitola 1

Prehľad problematiky

1.1 Úvod do teórie sociálnych sietí

Teória sociálnych sietí predstavuje komplexný pohľad na vzájomné vzťahy medzi jednotlivcami a organizáciami. Začala sa skúmať už v 60. rokoch 20. storočia [6]. Poskytla základy pre pochopenie vzájomných vzťahov a interakcií medzi jednotlivcami alebo skupinami. V sociálnych sieťach môžme analyzovať (obr1.1), ako sú jednotlivci a skupiny prepojené, ako informácie týmito sieťami cestujú a ako sa tým formujú naše názory a správanie. Ako základnú definíciu sociálnej siete môžeme považovať všeobecný koncept fungovania našej spoločnosti [7].



Obr. 1.1: Sociálna sieť, vzťahy medzi ľuďmi

1.1.1 Internetové sociálne siete

Internetové sociálne siete v dnešnej dobe predstavujú digitálne platformy, ktoré umožňujú jednotlivcom, skupinám alebo organizáciám vytvárať a zdieľať obsah, komunikovať, a udržiavať sociálne vzťahy. Tieto platformy sa stali súčasťou moderného života a majú významný vplyv na spoločenské interakcie alebo správanie jednotlivcov [5]. Medzi najznámejšie sociálne siete a zároveň aj tie, ktorým sa budeme venovať v tejto práci a rozoberieme ich podrobnejšie patria Facebook, Instagram, X (predtým známa ako Twitter) a LinkedIn.

1.1.2 Historický vývoj sociálnych sietí

Pôvod sociálnych sietí môžeme vystopovať späť až k vzniku webových stránok ako Classmates.com a SixDegrees.com. Tieto platformy položili základy pre súčasnú generáciu sociálnych sietí a predstavovali významný krok v ich vývoji [5].

Classmates.com a začiatky sociálnych sietí

Classmates.com, spustený v roku 199, sa považuje za jednu z prvých platforiem, ktorá sa zamerala na spájanie ľudí s ich bývalými spolužiakmi5 [5]. Táto stránka umožňovala užívateľom vytvárať osobné profily, uvádzať informácie o svojom štúdiu a vyhľadávať spolužiakov. Classmates.com predstavoval kľúčový krok v evolúcii sociálnych sietí, keďže predstavoval základné funkcie, ktoré sú bežné aj v súčasných sociálnych sieťach.

SixDegrees.com - priekopník sociálneho prepojenia

Neskôr, v roku 1997, Andrew Weinreich predstavil SixDegrees.com, čím sa otvoril nový rozmer v sociálnom prepojení. SixDegrees.com bol inovatívny v tom, že umožnil vytváranie užívateľských profilov, zoznamov priateľov a posielanie správ, čím položil základy pre moderné sociálne siete. Aj napriek tomu, že táto platforma ukončila svoju činnosť v roku 2001, je široko uznávaná ako prvá skutočná sociálna sieť a slúžila ako model pre budúce platformy5 [5].

Friendster a rozvoj sociálnych sietí

V roku 2002 Jonathan Abrams založil Friendster, ktorý priniesol nové funkcie, ako žiadosti o priateľstvo a možnosť vytvárať spojenia s ostatnými užívateľmi cez sieť priateľov5 [5]. Friendster predstavoval ďalší významný krok v evolúcii sociálnych sietí a mal významný vplyv na ich ďalší vývoj.

MySpace - fenomén personalizácie

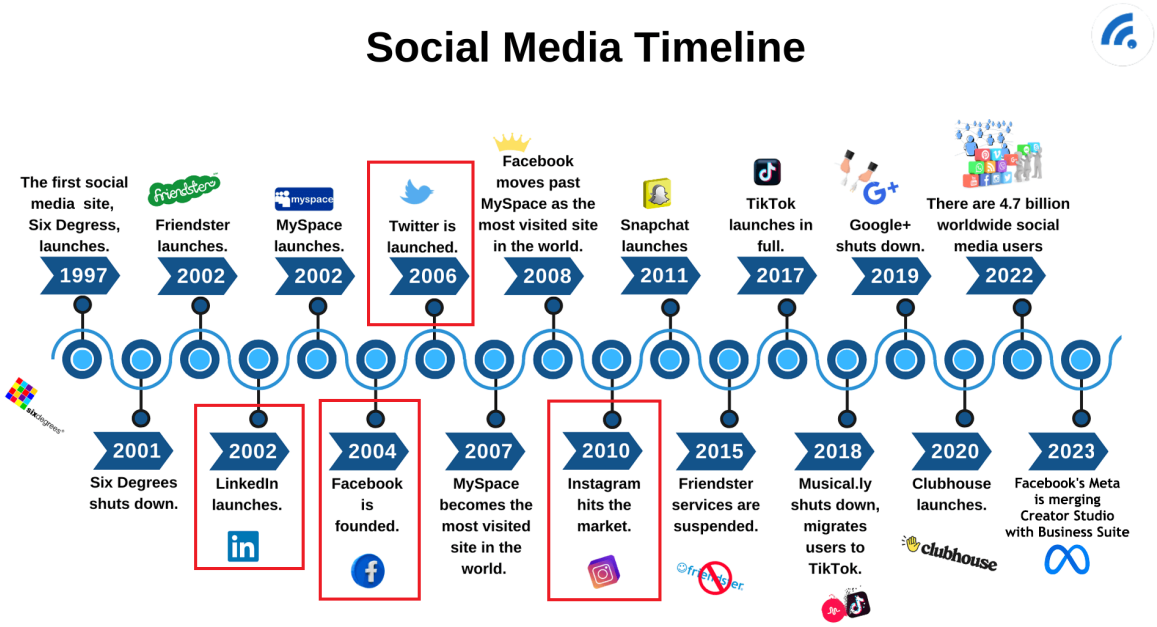
Následne v roku 2003 bol spustený MySpace, ktorý umožnil užívateľom vytvárať personalizované profily, zdieľať playlisty a spojiť sa s priateľmi5 [5]. MySpace rýchlo zís-

kal popularitu, najmä vďaka svojmu prepojeniu medzi hudobným a sociálnym odvetvím.

Facebook - globalizácia sociálnych sietí

V roku 2004 Mark Zuckerberg založil Facebook, ktorý sa rýchlo rozšíril z platformy pre študentov Harvardu na celosvetový fenomén5 [5]. Facebook ukázal globálny potenciál sociálnych sietí a pod vedením Zuckerberga neustále inovoval, rozširoval svoj dosah a vplyv, najmä prostredníctvom aplikácií ako Instagram a WhatsApp. Avšak, rast Facebooku bol sprevádzaný aj kontroverziami týkajúcimi sa ochrany osobných údajov a šírenia dezinformácií, čo zdôrazňuje etické komplikácie, ktoré sú spojené s nástupom sociálnych médií [8].

Vývoj sociálnych sietí predstavuje dynamickú a komplexnú oblasť, ktorá sa neustále vyvíja a mení. Od prvých krokov, ktoré urobili Classmates.com a SixDegrees.com, až po dnešné globálne platformy ako Facebook, sú sociálne siete neoddeliteľnou súčasťou našej digitálnej éry [5]. Ich vývoj zároveň poukazuje na potrebu neustáleho monitorovania tických a bezpečnostných otázok, ktoré sú s nimi spojené.



Obr. 1.2: Vývoj sociálnych sietí v čase

1.1.3 Prehľad a charakteristika najpopulárnejších sociálnych sietí

Facebook

Facebook je globálna sociálna sieť, ktorá pôsobí ako multifunkčná platforma umožňujúca užívateľom vytvárať a udržiavať sociálne kontakty, zdieľať obsah a komunikovať v reálnom čase. Platforma slúži ako prostriedok pre osobnú, ako aj profesijnú interakciu, poskytujúc nástroje na publikovanie príspevkov, fotografií, videí a na organizáciu udalostí [9]. Pre firmy a obchodníkov funguje Facebook ako marketingový nástroj, umožňujúci cielenú reklamu a komunikáciu so zákazníkmi. Vďaka množstvu užívateľov a pokročilých algoritmov na personalizáciu obsahu sa Facebook stal jednou z kľúčových platforiem pre tvorbu a šírenie digitálneho obsahu na internete.

Instagram

Instagram je sociálna sieť zameraná na zdieľanie fotografií a videí, ktorá umožňuje užívateľom vyjadriť sa a podeliť o svoje životné momenty prostredníctvom vizuálneho obsahu. S funkciou príbehov, IGTV a novších Reels, platforma poskytuje rôzne formáty na prezentáciu kreatívneho video obsahu. Instagram je tiež silný nástroj pre marketing, keďže značky spolupracujú s profilmi s vysokým počtom sledovateľov na propagáciu svojich produktov. Obsahuje aj funkcie ako priame správy (DMs) a možnosť pridávať hashtagy, ktoré podporujú spoločenskú interakciu a vyhľadávanie obsahu. Táto platforma sa tak stala nielen miestom pre osobné vyjadrenie, ale aj pre podnikanie, umenie a mediálne výmeny, vytvárajúc bohaté komunity okolo rôznych záujmov a aktivít [10].

X (predtým známa ako Twitter)

Twitter je sociálna sieť a mikroblogovacia platforma, kde užívatelia komunikujú prostredníctvom krátkych správ, známych ako tweety, ktoré sú limitované na určitý počet znakov. Táto platforma umožňuje rýchle šírenie informácií a slúži ako významný komunikačný kanál pre jednotlivcov, organizácie a verejných predstaviteľov na prenášanie myšlienok, názorov a aktualít v reálnom čase. S funkciami ako retweet, odpoveď a označenie „like“, Twitter podporuje interakciu a diskusie medzi užívateľmi a zároveň umožňuje sledovanie trendových tém prostredníctvom hashtagov [11].

LinkedIn

LinkedIn je profesionálna sociálna sieť, určená pre podnikateľské a profesionálne vzťahy, ktorá slúži ako centrum pre spojenie pracovníkov, náborárov a podnikateľov [12]. Umožňuje užívateľom vytvárať a udržiavať profesionálne kontakty, zdieľať svoj životopis a hľadať alebo ponúkať pracovné príležitosti. LinkedIn tiež poskytuje platformu na

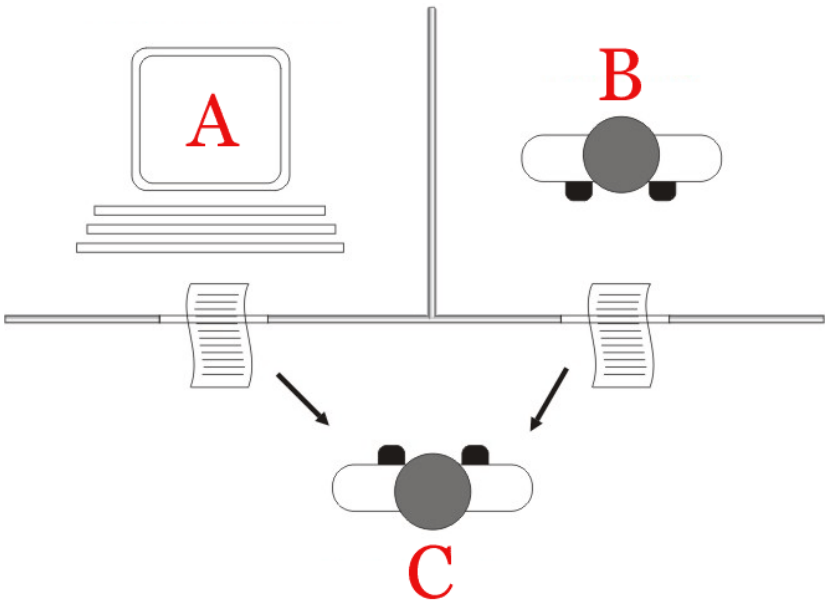
publikovanie a distribúciu odborného obsahu, čím podporuje odborný rast a vzdelávanie prostredníctvom článkov, štúdií a odborných diskusií. Ďalšou dôležitou funkcionalitou je schopnosť zobraziť a overiť profesionálne kvalifikácie a schopnosti, čo z neho robí cenný zdroj pre personalistov a manažérov pri hľadaní talentov.

1.1.4 Vznik a evolúcia botov

História vzniku botov sa začína v 50. rokoch 20. storočia, kedy počítačový vedec a matematik Alan Turing vyvinul Turingov test, známy aj ako Imitation Game(viď obr.1.3). Tento test bol navrhnutý ako jednoduchý experiment, ktorý mal slúžiť ako overenie, či môže stroj napodobniť ľudskú inteligentciu natoľko, že by človek nedokázal rozlíšiť, či odpovedá stroj alebo človek. Zahŕňal tieto tri postavy [13]:

- hráča A (stroj)
- hráča B (človeka)
- hráča C (tiež človek)

V tomto teste bol stroj (hráč A) a človek (hráč B) posudzovaní ďalším človekom (hráčom C), ktorý mal za úlohu zistiť, ktorý z nich je stroj a ktorý človek, na základe ich odpovedí na otázky. Tento test položil základ pre vývoj automatizácií a inštrukčných kódov.



Obr. 1.3: Diagram turingovho testu

V 60. rokoch vytvoril Joseph Weizenbaum program ELIZA (viď obr.1.4), ktorý bol schopný simulovať konverzáciu s človekom a bol považovaný za jeden z prvých algoritmov umelej inteligencie. ELIZA bola schopná viesť základnú konverzáciu a

napodobňovať niektoré formy ľudskej interakcie, čím položila základy pre dnešné dnešní botov používané v aplikácií [13].

Obr. 1.4: Používateľské rozhranie programu ELIZA

V 80. rokoch sa konalo prvé výročné zasadnutie American Association of Artificial Intelligence. V tomto desaťročí tiež debutoval AARON, bot ktorý bol schopný tvoriť originálne abstraktné umenie, ktoré bolo vystavovaný v galériách ako Tate Gallery, Stedelijk Museum a San Francisco Museum of Modern Art. Zaujímavosťou tohto obdobia sú aj prvé pokusy o vytvorenie samo-riaditeľných vozidiel, napr. projekt ALVINN z Carnegie Mellon University v roku 1989 [13].

V 90. rokoch sa boti začali viac orientovať na bežných spotrebiteľov. Jedným z príkladov je Tamagotchi z roku 1996, hoci nebolo vyslovene označené ako bot, interakcia s ním bola podobná. V roku 2000 bol vydaný SmarterChild, bot integrovaný do zoznamu priateľov na American Online Instant Messenger (AIM), ktorý bol priekopníkom v oblasti hlasového vyhľadávania. Začiatkom 20. storočia tiež zaznamenali pokrok v oblasti autonómnych vozidiel. [13].

Významným momentom pre botov bolo zavedenie Siri od Apple v roku 2011, ktorá umožnila hlasové vyhľadávanie a osobnú asistenciu priamo z telefónu. To odštartovalo rýchly rast botov medzi používateľmi a otvorilo dvere pre možnosti hlasového vyhľadávania v kombinácií s Internetom. Amazon vydal Echo len tri roky po uvedení Siri, ktorý využíva bota menom Alexa na zodpovedanie otázok a riadenie domácej automatizácie. Google vstúpil do tohto priestoru v roku 2015 s uvedením Google Home. V roku 2016 oznámil Facebook, že jeho platforma Messenger môže byť integrovaná s botmi na poskytovanie automatizovaného obsahu, ako sú aktualizácie počasia, dopravy a prispôsobené komunikácie [13].

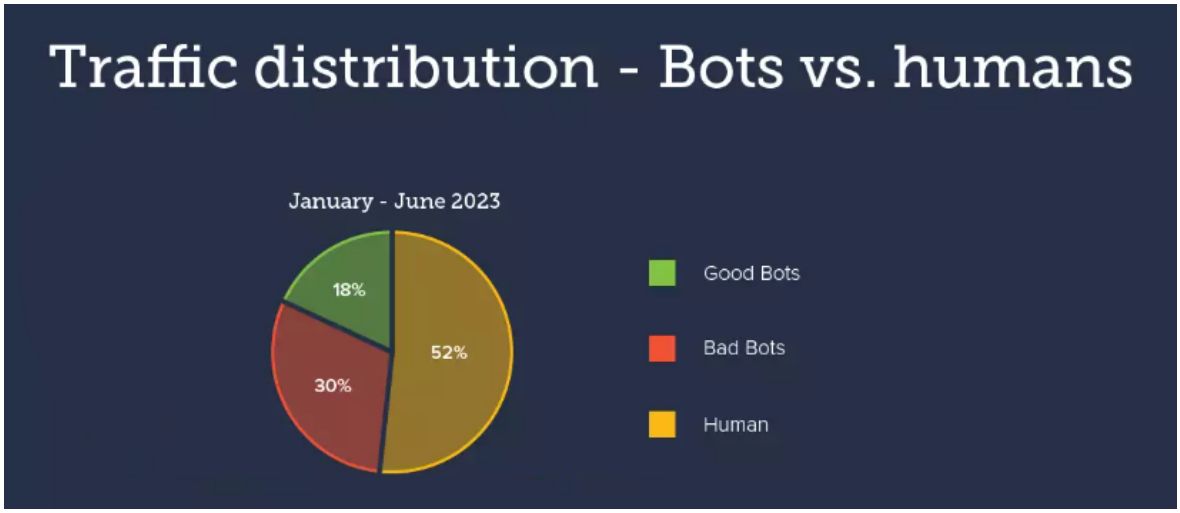
Táto história nám ukazuje, ako sa boti vyvíjali od počiatočných teoretických konceptov a experimentov v 50. rokoch cez prve praktické aplikácie v 60. a 70. rokoch, až po dnešné sofistikované aplikácie v oblasti hlasového vyhľadávania, osobnej asistencie a automatizácie. Každé desaťročie nám prinieslo významné technologické inovácie, ktoré formovali súčasnú podobu a využitie botov.

1.2 Automatizácia na internete

Väčšina webových návštevníkov nie sú ľudia, ale boti. Boti (z angličtiny "bots"), známi aj ako internetové alebo webové roboty, sú programy navrhnuté na vykonávanie automatizovaných úloh na internete. Podľa najnovšej správy bezpečnostnej firmy Baracuda [14], ktorá každoročne vydáva hodnotenie online aktivity botov, ktorý je založený na analýze takmer 17 miliárd návštev webových stránok z celkovo 100 000 domén, celkovo tvoria boti až 52% (viď obr.1.5) všetkých procesov na internete. Niektorí nám môžu pomáhať udržiavať naše profily na sociálnych sieťach, iný sa môžu vydávať za ľudí a vykonávať škodlivé DDoS útoky. Vo všeobecnosti môžeme týchto botov rozdeliť na dve hlavné kategórie (viď. aj tab.1.1):

- Dobrí boti
- Zlí boti

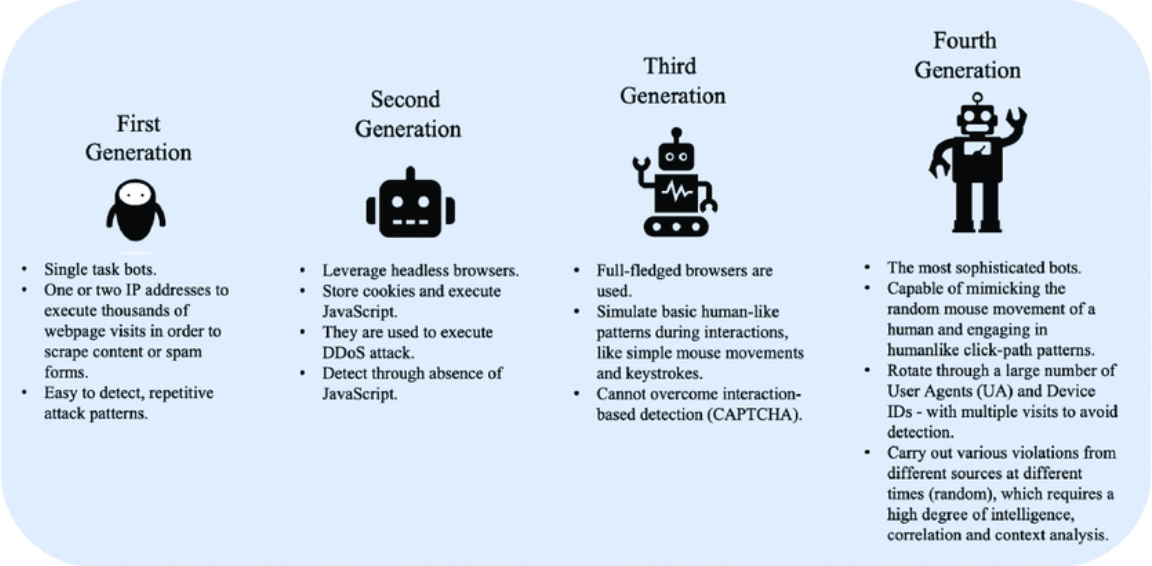
Kým dobrí boti zvyčajne prispievajú k pozitívnym používateľským skúsenostiam tým, že automatizujú rutinné úlohy, interagujú s používateľmi, alebo zbierajú dáta na analýzu, zlí boti môžu byť zneužití na šírenie dezinformácií, spamu alebo vykonávanie neetických aktivít.



Obr. 1.5: Aktivita botov na internete (jún 2023) [14]

1.2.1 Generácie internetových botov

Evolúcia botov je príkladom neustáleho technologického rozvoja a schopnosti adaptácie na neustále sa meniace prostredie. Počas evolúcie botov môžeme pozorovať štyri generácie, pričom každá generácia je charakteristická značím zlepšením v napodobňovaní ľudskej aktivity a interakcie, čím sa stávajú čoraz ťažšie odlíšiteľnými od skutočných užívateľov.



Obr. 1.6: 4 generácie botov [ZDROJ]

História

Prvá Generácia: Jednoúčelové boty

Prvá generácia botov sú jednoduché automatizované programy, ktoré môžu vykonávať veľké množstvo návštev na webových stránkach s cieľom scrapovať obsah alebo spamovať formuláre [15]. Sú relatívne jednoduchý na detekciu a následne zablokovanie vzhľadom na ich opakujúce sa vzorce útokov a obmedzenému počtu IP adries, z ktorých operujú.

Títo jednoduchí boti obvykle pochádzajú z dátových centier a využívajú proxy IP adresy a nekonzistentných užívateľských agentov (UAs - User Agents - reťazce identifikujúce prehliadač, ktorý používateľ používa pri prístupe na webové stránky). Často generovali tisíce requestov iba z jednej alebo dvoch IP adries. Operujú prostredníctvom nástrojov na scrapovanie, ako sú ScreamingFrog alebo DeepCrawl [15]. Sú najľahšie detegovateľné, keďže nemôžu udržiavať cookies, ktoré väčšina webových stránok využíva. Okrem toho zlyhávajú pri JavaScriptových kontrolách ako napr. Recaptcha alebo CloudFlare, pretože ich nevedia spustiť. Botov prvej generácie je možné blokovať blacklistom ich IP adries a User Agentov, ale taktiež aj ich spoločnou kombináciou.

Druhá Generácia: Bezhlavé prehliadače

Boti druhej generácie sa vyznačujú používaním vývojových a testovacích nástrojov webových aplikácií, známych ako "bezhlavé"(headless) prehliadače (napríklad PhantomJS alebo SimpleBrowser), a tiež neskorších verzií Chrome a Firefox, ktoré umožňujú fungovanie v bezhlavom režime [15]. Na rozdiel od botov prvej generácie títo boti môžu udržiavať cookies a spúšťať JavaScript. Experti na botov začali používať bezhlavé prehliadače vzhľadom na rastúce používanie JavaScriptových výziev na webových stránkach a aplikáciách.

Títo boti sa používajú na útoky DDoS na aplikácie, scrapovanie, spamovanie formulárov, skreslenie analytických dát a podvody s reklamou.

Je možné identifikovať prostredníctvom ich prehliadačových charakteristík a nastavení, vrátane prítomnosti špecifických JavaScriptových premenných, manipulácie s iframom, relácií a cookies. Po identifikácii bota je možné ho zablokovať na základe jeho digitálnych odtlačkov. Takisto jedna z metód detekcie týchto botov je aj analýza metrík a typických používateľských ciest, pričom sa hľadajú veľké rozdiely v premávke (traffic) medzi rôznymi sekciami webovej stránky. Tieto rozdiely môžu odhaľovať prítomnosť botov, ktorí majú v úmysle vykonávať rôzne typy útokov, ako je prevzatie účtu alebo scrapovanie.

Tretia Generácia: Plnohodnotné prehliadače

Tretia generácia botov využíva pri svojej činnosti plnohodnotné prehliadače. Títo boti sú schopný simulovať základné ľudské interakcie, ako sú napr. jednoduché pohyby myšou a stláčania kláves. Avšak ich pohyby sú stále príliš strojové a môžu mať problém preukázať ľudskú náhodnosť v ich správaní [15].

Boti tretej generácie sa používajú na hackovanie účtov, DDoS útoky na aplikácie, zneužívanie API, podvody s reklamou a mnoho ďalších.

Títo boti sú ťažšie detegovateľní na základe charakteristík zariadenia a prehliadača. Na ich zistenie je potrebná analýza používateľského správania založená na interakcii, keďže tieto boty zvyčajne nasledujú programatickú sekvenciu prechádzania URL adries.

Avšak, aj napriek tejto zvýšenej sofistikovanosti, tieto boty stále nedokázali prekonať interakčne založenú detekciu, ako sú napr. CAPTCHA testy.

Štvrtá Generácia: Najsofistikovanejšie boty

Štvrtá generácia botov predstavuje vrchol v evolúcii botov s najmodernejšími technológiami a s pokročilými charakteristikami ľudských interakcií, vrátane náhodného pohybu kurzora myši namiesto pohybu po priamkach. Tieto boty tiež môžu meniť svojich užívateľských agentov (UAs) pri rotácii tisícov IP adries. Rastúce dôkazy poukazujú na to, že vývojári botov vykonávajú "hijacking správania", čo znamená zaznamenávanie

spôsobu, akým sa reálny používatelia dotýkajú a posúvajú na mobilných aplikáciách, aby presnejšie napodobňovali ľudské správanie na webovej stránke alebo v aplikácii. Toto zneužitie správania ich robí oveľa ťažšie detegovateľnými, keďže ich aktivity nie je ľahké rozlíšiť od aktivít skutočných používateľov a navyše ich široké rozšírenie je pripisované k veľkému počtu používateľov, ktorých prehliadače a zariadenia boli zneužité.

Boti štvrtej generácie sa používajú na preberanie účtov, DDoS útoky na aplikácie, zneužívanie API, podvody s reklamou a mnoho ďalších.

TTíto boti sú masovo distribuovaní naprieč desiatkou tisíc IP adries, pričom často vykonávajú "low and slow"útoky, aby prešli bezpečnostnými opatreniami. Detekcia týchto botov na základe charakteristík interakcie, ako sú vzory pohybov myši, by viedla k vysokému počtu falošných pozitívnych výsledkov. Súčasné techniky sú preto nedostatočné na odhalenie takýchto botov. Sú potrebné technológie založené na strojovom učení, ako je analýza hlbokého správania, ktoré modely strojového učenia na identifikáciu úmyslu botov s najvyššou presnosťou teda na presnú detekciu botov štvrtej generácie bez falošných pozitívnych výsledkov.

V kontexte sociálnych sietí môžu mať boti všetkých štyroch generácií významný vplyv. Od jednoduchého spamovania a scrapovania obsahu až po sofistikované šírenie dezinformácií alebo vykonávanie koordinovaných útokov, boti predstavujú výzvu pre bezpečnosť online priestoru. Ich neustály vývoj si vyžaduje rovnako dynamické metódy detekcie a ochrany, aby sa udržala dôveryhodnosť a bezpečnosť digitálneho prostredia.

1.2.2 Úlohy a vplyv botov na internete

Boti v digitálnom prostredí môžu zastupovať rôzne úlohy, od užitočných až po škodlivé, a majú významný dopad na celkový online ekosystém:

- Pomocní boti: Mnohí boti sú programovaní na vykonávanie užitočných úloh, ako je poskytovanie zákazníckej podpory, automatizácia opakovaných úloh, zber a analýza dát alebo poskytovanie informácií v reálnom čase. Napríklad chatboti na webových stránkach môžu efektívne pomáhať zákazníkom pri riešení ich otázok a problémov [16] [17].

- Marketing a propagácia: Boti sú často používané na marketingové účely, ako je zvyšovanie návštevnosti webstránok, šírenie obsahu, zvyšovanie interakcie na sociálnych sieťach a zlepšovanie online viditeľnosti značky alebo produktu. Tieto aktivity však môžu mať aj negatívne dôsledky, ako je manipulácia s verejným vnímaním alebo skresľovanie metrík sledovanosti [16].

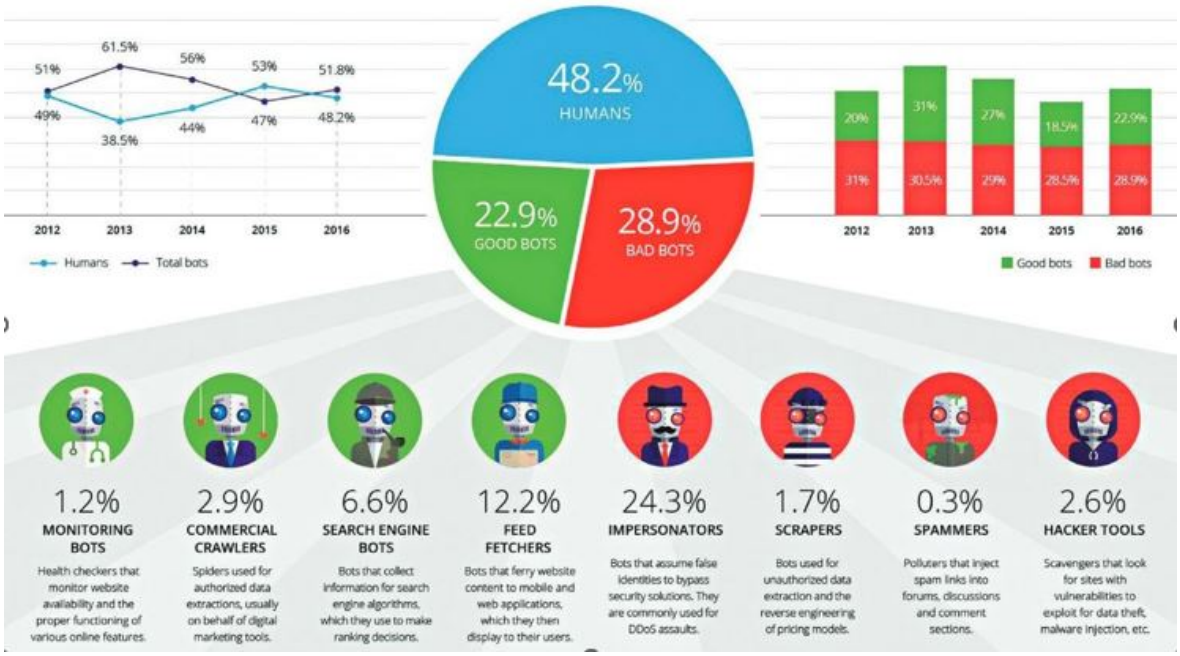
- Škodlivé aktivity: Škodliví boti sú používaní na nelegálne alebo neetické účely, vrátane šírenia falošných správ, vykonávania DDoS útokov, phishingu, spamovania a

krádeže identity. Títo boti môžu spôsobiť významné škody, ako je narušenie služieb, strata dôvery používateľov a finančné straty pre postihnuté organizácie [17] [16].

- Vplyv na verejnú mienku a voľby: Boti majú významný vplyv na verejnú mienku a politické diskusie. Výskumy ukázali, že boti boli aktívne používaní na ovplyvňovanie verejnej mienky počas významných politických udalostí, ako boli Brexit, prezidentské voľby v USA a parlamentné voľby vo Francúzsku a Nemecku [17].
- Výzvy v Detekcii: S narastajúcou sofistikovanosťou botov sa stáva ich detekcia náročnejšou. Vyspelí boti môžu imitovať ľudské správanie a unikať detekčným metódam, čo vyžaduje pokročilé technológie ako strojové učenie a hlboká analýza správania na presnú identifikáciu a zablokovanie týchto botov [16].

1.2.3 Typy botov na sociálnych sieťach

Na sociálnych sieťach môžme nájsť rôzne typy botov(viď tab. 1.1), niektoré poskytujú užitočné služby, ako sú aktuálne informácie o počasí, športové výsledky, správa profilu a pridávanie príspevkov, iné sú maskovaní ako bežní používatelia a sú využívaný na rôzne škodlivé účely. V tejto práci sa orientujeme na škodlivejších botov, ktorý majú najpočetnejšie zastúpenia na sociálnych sieťach:



Obr. 1.7: Zastúpenie rôznych typov botov na internete (2016)

Troll Boti

Troll boti sú navrhnutý tak, aby provokovali a zapájali sa do online diskusií. Často vytvárajú rozruch a konflikty pomocou kontroverzných alebo provokatívnych komentárov,

čím narušujú komunikáciu a vytvárajú konflikty.

Spamoví Boti

Títo boti sú známi pre šírenie nevyžiadanej pošty, reklám a odkazov na rôzne malvéry. Spamoví boti môžu zaspamovať sociálne siete a e-mailové schránky nežiaducimi správami, čo môže viesť k zníženiu dôveryhodnosti a celkového užívateľského dojmu.

Scrapovací Boti

Scrapovací boti sú používaní na automatické zbieranie obsahu a dát z webových stránok alebo sociálnych sietí. Títo boti môžu zbierať citlivé údaje alebo kopírovať obsah bez povolenia, čo predstavuje problém autorských práv a ochrany súkromia.

Hoax Boti

Títo boti šíria falošné informácie alebo hoax správy na ovplyvnenie verejnej mienky alebo na podporu určitých názorov. Môžu byť využívané na politické kampane, šírenie falošných správ alebo na manipuláciu verejných diskusií.

Impostor Boti

Impostor boti napodobňujú skutočné používateľské profily s cieľom klamať alebo získavať informácie. Môžu byť používaný na krádež identity, phishing alebo na získavanie dôvernosti iných používateľov.

Promotion Boti

Boti, ktorí falšujú popularitu, vytvárajú ilúziu vysokej popularity určitých príspevkov, účtov alebo tém prostredníctvom fake lajkov, komentárov a zdieľaní, čím ovplyvňujú vnímanie týchto príspevkov a rozširujú dosah obsahov.

Scam Boti

Scam boti sú programovaný na sledovanie a zbieranie osobných údajov používateľov bez ich vedomia alebo súhlasu, čím predstavujú vážne riziko pre ochranu súkromia.

Good Bots	Bad Bots
Chat Bot	Spam Bot
Information Bot	Promotion Bot
Service Bot	Phishing Bot
Monitoring Bot	Disinformation Bot
Research Bot	Tracking Bot
Helper Bot	Scam Bot
Indexing Bot	Scraper Bot
Authentication Bot	Hacker Bot

Tabuľka 1.1: Niektorý z najčastejšie sa vyskytujúcich botov na sociálnych sieťach

1.2.4 Metódy detekcie a boja proti botom na sociálnych sieťach

V súčasnom prostredí sociálnych sietí je detekcia a boj proti botom kľúčovou výzvou. Existuje niekoľko prístupov a techník, ktoré sociálne siete využívajú na identifikáciu botov:

Analýza správania

Táto metóda sa zameriava na rozpoznanie neobvyklých alebo automatizovaných vzorcov správania, ako sú opakujúce sa aktivitý, neštandardné časy príspevkov alebo rýchlosť interakcií, ktoré nesúhlasia s typickým ľudským správaním. [18] [19]

Strojové učenie a umelá inteligencia

Algoritmy strojového učenia a umelá inteligencia sa využívajú na identifikáciu botov na základe rozsiahlych dátových sád. Tieto systémy sa môžu učiť z minulých dát a neustále sa adaptovať na nové stratégie, ktoré boti používajú. [19]

Analýza Sociálnych Sietí

Tento prístup zahŕňa analýzu štruktúry a dynamiky sociálnych sietí, ako sú vzťahy a interakcie medzi účtami, aby sa identifikovali skupiny botov alebo neobvyklé vzorce sieťovania. [18]

Využitie CAPTCHA a JavaScriptových výziev

Na odlišenie botov od ľudí sa často používajú CAPTCHA testy a JavaScriptové výzvy. Tieto metódy skúmajú, či návštevníci stránky dokážu spracovať určité prvky, ako sú cookies alebo CAPTCHA, čo je úloha, s ktorou majú boti problémy. [19]

Využitie štatických a výzvových nástrojov

Štatické nástroje na analýzu identifikujú požiadavky na web a hlavičkové informácie spojené so škodlivými botmi, zatiaľ čo výzvové metódy aktívne testujú, či návštevník je človek alebo bot. Tieto nástroje môžu efektívne identifikovať a blokovať nežiaduce boty. [19]

Behaviorálne prístupy

Tieto prístupy zahŕňajú analýzu behaviorálneho podpisu každého návštevníka, aby sa zistilo, či sa správa v súlade s normálnym správaním používateľa. Porovnávajú sa behaviorálne vzorce s predchádzajúcimi známymi podpismi škodlivých botov. [19]

1.2.5 Etické a právne dilemy spojené s používaním botov na sociálnych sieťach

V súčasnej dobe je používanie botov na sociálnych sieťach sprevádzané rôznymi etickými a právnymi dilemami.

- **Ovplyvňovanie verejnej mienky**: Boti na sociálnych sieťach môžu výrazne ovplyvňovať verejnú diskusiu, najmä v kontexte dezinformačných kampaní. Regulácie a politiky platforiem týkajúce sa používania sociálnych botov sú preto kľúčové. Avšak, ako sa uvádza v štúdii CNA, regulácie v tejto oblasti sú často ťažko zoskupiteľné a rozptýlené. [20]

- **Dilemy sociálnych sietí** Spoločnosti sociálnych sietí čelia vlastným dilemám pri vytváraní účinných predpisov pre botov. Po amerických voľbách v roku 2016 vzniká pod tlakom verejnosti väčší tlak na zabezpečenie týchto hrozieb. Platformy sa ale tiež obávajú prílišného regulovania, čo by mohlo viesť k odvolaniu ich obmedzenej imunity podľa CDA 230. [20]

- **Dôsledky politiky platfomy** Platformy často riešia správanie botov prostredníctvom iných politík týkajúcich sa zakázaného správania. Dôsledky porušenia pravidiel jednotlivých platforiem sa líšia, pričom stránky často posudzujú konkrétne porušenie, závažnosť priestupku a históriu používateľa na platforme. [20]

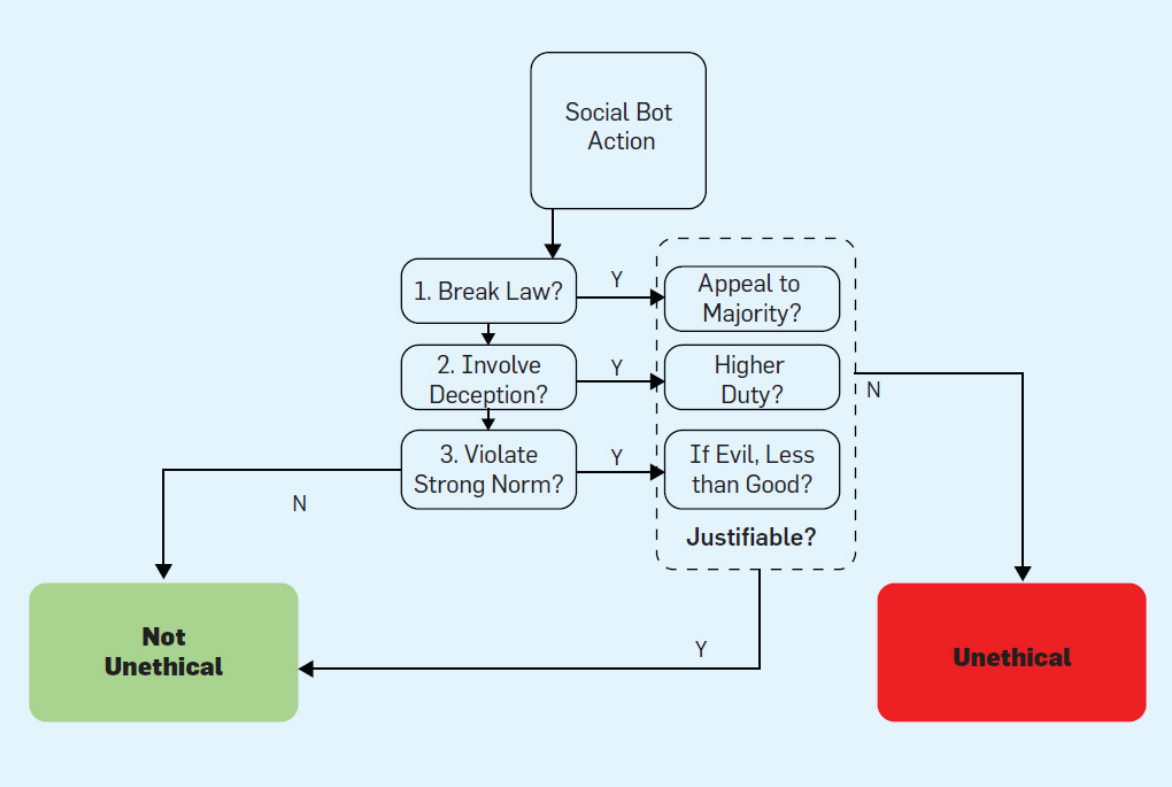
- **Budúce výzvy** Hrozby spojené s dezinformáciami a zlomyseľnými botmi budú pravdepodobne naďalej spôsobovať problémy americkému vládnemu aparátu. Očakáva sa, že sociálne mediálne platformy budú niesť veľkú časť bremena pri riešení týchto problémov, avšak úloha a rozsah ich samoregulácie zostáva otvorenou otázkou. [20]

- **Medzinárodné úsilie v EÚ**: Medzinárodné úsilie v EÚ v súvislosti s botmi na sociálnych sieťach je súčasťou Kódexu správania o dezinformáciách, ktorý bol prvýkrát zavedený v roku 2018 a posilnený v roku 2022. Tento Kódex obsahuje opatrenia zamerané na:

  Znižovanie manipulatívneho správania: Kódex posilňuje opatrenia na zníženie manipulatívneho správania používaného na šírenie dezinformácií, ako sú falošné účty, amplifikácia riadená botmi, napodobňovanie a škodlivé deepfakes. Tým sa priamo zameriava na techniky používané botmi na sociálnych sieťach.

  Posilnenie používateľov a výskumníkov: Kódex zabezpečuje, aby používatelia boli lepšie chránení pred dezinformáciami prostredníctvom nástrojov na ich rozpoznanie a označovanie. Zároveň poskytuje výskumníkom lepší a širší prístup k údajom platformy, čo umožňuje efektívnejšie štúdium vplyvu a rozšírenia botov.

  Podpora overovateľov faktov: Nový Kódex rozširuje pokrytie overovateľov faktov vo všetkých členských štátoch EÚ a jazykoch, čo zvyšuje konzistentnosť používania overovateľov faktov na platformách. Toto je kľúčové pre identifikáciu a riešenie dezinformácií šírených botmi. [21]



Obr. 1.8: Rozdelenie botov na etických a neetických, podľa akcií, ktoré vykonávajú

Kapitola 2

Teória

An example chapter with sections.

2.1 Algoritmy na detekciu botov

2.1.1 Prehľad a klasifikácia algoritmov

2.1.2 Algoritmy založené na strojovom učení

2.1.3 Behaviorálne algoritmy

2.1.4 Výzvy a obmedzenia existujúcich algoritmov

2.2 Zabezpečenie sociálnych sietí

2.2.1 Autentifikácia a autorizácia užívateľov

2.2.2 Mechanizmy prevencie

2.2.3 Prípady útokov a obranné stratégie

2.3 Generácie internetových botov

2.3.1 Prvá generácia: Základné skripty

2.3.2 Druhá generácia: Pokročilé interakcie

2.3.3 Tretia generácia: Samoučiace sa systémy

2.4 Potrebné technológie a nástroje

2.4.1 Python a jeho knižnice

2.4.2 Selenium: Automatizácia webového prehliadača

2.4.3 PySQLite3: Práca s databázam

Kapitola 3

Praktická časť

An example chapter.

3.1 Návrh a implementácia

3.1.1 Vývojové prostredie a nástroje

3.1.2 Návrhový proces databázy

3.1.3 Návrhový proces botov

3.1.4 Programovanie potrebných modulov

3.1.5 Programovanie botov

3.1.6 Testovacie scenáre

3.2 Testovanie zabezpečení sociálnych sietí

3.2.1 Definícia testovacích kritérií

3.2.2 Automatizácia testov

3.3 Úprava a optimalizácia botov

3.3.1 Optimalizácia kódu a výkonu

3.3.2 Zlepšenia zabezpečenia

Kapitola 4

Výsledky

Kapitola "Výsledky"sa zameriavame na analýzu a hodnotenie bezpečnostných mechanizmov sociálnych sietí voči botom vyvíjaným v Pythone a Seleniu. Podrobne opisujeme postupy vyvinuté na tvorbu botov, ktorí sú klasifikovaní do troch generácií podľa ich schopností a komplexnosti. Metodológia zahŕňa proces testovania týchto botov na rôznych platformách sociálnych sietí, zameriavajúc sa na ich schopnosť obchádzať bezpečnostné protokoly a získavanie údajov. V tejto časti sa tiež skúmame ich vplyv na šírenie dezinformácií a kyberšikanu. Záverečná časť kapitoly sumarizuje výsledky a diskutuje o ich dôležitosti pre pochopenie a zlepšenie bezpečnostných opatrení na ochranu užívateľov sociálnych sietí.

4.1 Zhromaždené dáta a analýzy

4.1.1 Metodika zhromažďovania dát

4.1.2 Štatistické spracovanie

4.1.3 Interpretácia výsledkov

4.2 Diskusia a porovnanie s existujúcimi riešeniami

4.2.1 Porovnanie účinnosti

4.2.2 Diskusia o nálezoch

4.2.3 Návrhy na zlepšenie

Záver

Put your conclusion here.

Literatúra

[1] Zineb Ellaky, Faouzia Benabbou, and Sara Ouahabi. Systematic literature review of social media bots detection systems. *Journal of King Saud University - Computer and Information Sciences*, 35(5), 2023.

[2] M. Aljabri, R. Zagrouba, and A. Shaahid. Machine learning-based social media bot detection: a comprehensive literature review. *Social Network Analysis and Mining*, (20), 2023.

[3] Maksim Kalameyets. *Algorithms and techniques for bot detection in social networks*. PhD thesis, Université Paul Sabatier - Toulouse III, 2021.

[4] Adam Zoltan Kenyeres. Social media bot detection. Master's thesis, University of Technology Lulea, Sweden, 2021.

[5] Danah M. Boyd and Nicole B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13:210–230, 2008.

[6] J. A. Barnes. Class and committees in a norwegian island parish. *Human Relations*, 7(1), 1954.

[7] Jan H. Kietzmann, Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre. Social media? get serious! understanding the functional building blocks of social media. *Business Horizons*, 54(3), 2011.

[8] Patrick Mutabazi. The evolution of social media: How did it begin, and where could it go next? LinkedIn, 2023.

[9] Ralf Caers, Tim De Feyter, Marijke De Couck, Talia Stough, Claudia Vigna, and Cind Du Bois. Facebook: A literature review. *New Media & Society*, 15(6), 2013.

[10] Elise Moreau. What is instagram, anyway. Here's what Instagram is all about and how people are using it [online], 2018.

[11] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600, 2010.

[12] O LinkedIn. About linkedin. LinkedIn Corporation, 2022.

[13] Amanda Zantal-Wiener. Where do bots come from? a brief history. HubSpot, 2021.

[14] Amber Jackson. Us face api attacks as bad bots account for 72 Cybermagazine, 2023.

[15] Meet the four generations of bots. Radware, 2023.

[16] Ralf Rodriguez. Understanding social media bots: A comprehensive guide. Softlist, 2023.

[17] Social bots – the technology behind fake news. Ionos, 2022.

[18] Bots. Imperva, 2023.

[19] What types of bots are there? Fastly, 2023.

[20] Kasey Stricklin and Megan McBride. Social media bots: Laws, regulations, and platform policies. CNA, 2020.

[21] Brooke Tanner. Eu code of practice on disinformation. Brookings, 2022.