Využití umělé inteligence v oblasti sociálního inženýrství

Bc. David Vítek

Diplomová práce
2024

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky
Ústav elektroniky a měření

Akademický rok: 2023/2024

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	**Bc. David Vítek**
Osobní číslo:	**A22372**
Studijní program:	**N1032A020003 Bezpečnostní technologie, systémy a management**
Specializace:	**Bezpečnostní management**
Forma studia:	**Prezenční**
Téma práce:	**Využití umělé inteligence v oblasti sociálního inženýrství**
Téma práce anglicky:	**Applications of artificial intelligence in the field of social engineering**

Zásady pro vypracování

1. Popište základní pojmy sociálního inženýrství a charakterizujte současně nejrozšířenější techniky sociálního inženýrství, které mohou využít umělou inteligenci.
2. Popište současně nejpoužívanější modely umělé inteligence.
3. Proveďte vlastní šetření se zaměřením na generování a úpravu hlasu za pomoci umělé inteligence a vyberte vhodné nástroje.
4. Navrhněte a realizujte sociální experiment za použití uměle generovaného hlasu.
5. Zpracujte a zdůvodněte opatření, které bude navazovat na předešlé testování.

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam doporučené literatury:

1. SALAHDINE, Fatima a KAABOUCH, Naima. Social Engineering Attacks: A Survey. Online. *Future Internet*. 2019, roč. 11, č. 4, s. 17. ISSN 1999-5903. Dostupné z: https://doi.org/10.3390/fi11040089. [cit. 2023-10-18]
2. MOUTON, Francois; MALAN, Mercia M.; LEENEN, Louise a VENTER, H.S. Social engineering attack framework. Online. In: *2014 Information Security for South Africa*. Johannesburg: IEEE, 2014, s. 1-9. ISBN 978-1-4799-3384-6. ISSN 2330--9881. Dostupné z: https://doi.org/10.1109/ISSA.2014.6950510. [cit. 2023-10-18]
3. UEBELACKER, Sven a QUIEL, Susanne. The Social Engineering Personality Framework. Online. In: *2014 Workshop on Socio-Technical Aspects in Security and Trust*. Vídeň: IEEE, 2014, s. 24-30. ISBN 978-1-4799-7901-1. ISSN 2325-1697. Dostupné z: https://doi.org/10.1109/STAST.2014.12. [cit. 2023-10-18]
4. HADNAGY, Christopher. *Social engineering: The Art of Human Hacking*. 2. Vyd. Wiley Publishing, 2018. ISBN 978--1119433385.
5. SARKER, Iqbal H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. Online. *SN Computer Science*. 2021, roč. 2, č. 6. ISSN 2662-995X. Dostupné z: https://doi.org/10.1007/s42979-021-00815-1. [cit. 2023-10-19]

Vedoucí diplomové práce: **Ing. Lukáš Králík, Ph.D.**
Ústav bezpečnostního inženýrství

Datum zadání diplomové práce: **20. listopadu 2023**
Termín odevzdání diplomové práce: **28. května 2024**

doc. Ing. Jiří Vojtěšek, Ph.D. v.r.
děkan

UNIVERZITA TOMÁŠE BATI VE ZLÍNĚ
Fakulta aplikované informatiky
L.S.

Ing. Milan Navrátil, Ph.D. v.r.
ředitel ústavu

Ve Zlíně dne 1. prosince 2023

Prohlašuji, že

- beru na vědomí, že odevzdáním diplomové práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně;
- byl/a jsem seznámen/a s tím, že na moji diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou práci nebo poskytnout licenci k jejímu využití jen připouští-li tak licenční smlouva uzavřená mezi mnou a Univerzitou Tomáše Bati ve Zlíně s tím, že vyrovnání případného přiměřeného příspěvku na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše) bude rovněž předmětem této licenční smlouvy;
- beru na vědomí, že pokud bylo k vypracování diplomové práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně, dne

David Vítek, v.r.
podpis studenta

ABSTRAKT

Diplomová práce se zabývá využitím umělé inteligence v oblasti sociálního inženýrství. Výzkum je zaměřen na podvodné telefonáty využívající generovaný syntetický hlas blízké osoby. Cílem diplomové práce je analyzovat nově vzniklou hrozbu v podobě podvodných telefonátů se syntetickým hlasem a na základě zjištění z analýzy vytvořit protiopatření. Teoretická část se zabývá základními pojmy, vlivem psychologie, rozšířeným modelem útoku a technikami, které již používají umělou inteligenci. Dále se teoretická část zabývá základní terminologií ohledně umělé inteligence a modely, které jsou hojně používané pro generování textu, obrázků a audio obsahu. Praktická část se zabývá analýzou dat získaných z ankety, sociálního experimentu a rozhovorů. Veškeré informace jsou následně interpretovány a na jejich základě vytvořeny opatření proti podvodným telefonátům se syntetickým hlasem.

Klíčová slova: sociální inženýrství, umělá inteligence, vishing, podvodné telefonáty, syntetický hlas

ABSTRACT

The thesis deals with the use of artificial intelligence in the field of social engineering. The research focuses on fraudulent phone calls using a generated synthetic voice of a close person. The aim of the thesis is to analyze the emerging threat of fraudulent phone calls with synthetic voice and to develop countermeasures based on the findings of the analysis. The theoretical part deals with basic concepts, the influence of psychology, an extended attack model and techniques that already use artificial intelligence. Furthermore, the theoretical part deals with the basic terminology regarding artificial intelligence and the models that are widely used to generate text, images and audio content. The practical part deals with the analysis of data collected from survey, social experiment and interviews. All the information is then interpreted and based on it, measures against fraudulent synthetic voice calls are developed.

Keywords: Social Engineering, Artificial Intelligence, Vishing, Fraudulent Phone Calls, Synthetic Voice

Rád bych poděkoval vedoucímu práce panu Ing. Lukáši Králíkovi, Ph.D. za veškeré rady a pomoc, které mi během zpracování mé práce poskytl.

Dále bych chtěl poděkovat všem účastníkům výzkumu za jejich ochotu a otevřenost. Poslední poděkování patří celé mé rodině a přátelům za jejich podporu při mém studiu.

Prohlašuji, že odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

ÚVOD

Umělá inteligence je jedním z nejrychleji rostoucích oborů v současnosti. Tento nástroj skrývá nepřeberné množství způsobů, jak pomáhat v různorodých odvětví jako doprava, energetika, průmysl, zdravotnictví a mnohé další. Umělá inteligence tedy může efektivně pomoci na mnoha místech a s mnoha činnostmi, ale co se stane, když začnou umělou inteligenci zneužívat jedinci nebo skupiny se špatnými úmysly? V oblastech sociálního inženýrství byl v posledních letech zaznamenán nárust nových způsobů a podvodů, které využívají ke své podpoře právě umělou inteligenci. Techniky poháněné umělou inteligencí skrývají podvody větší nebezpečí, jelikož jsou více personalizované i automatizované. Dokonce pomáhají usnadnit krádež identity člověka od jeho hlasu, až po jeho obličej. Phishing e-maily psané umělou inteligencí jsou k nerozeznání od běžných e-mailů psaných vedoucím nebo blízkou osobou. Hlasem v telefonních hovorech nemusí nutně mluvit přítel nebo rodinný příslušník, ale kdokoliv, kdo se rozhodl hlas syntetizovat a využívat jej k podvodům. Jak je patrné, umělá inteligence vnáší do světa sociálního inženýrství a sofistikovaných podvodů novou úroveň nebezpečí, na kterou je potřeba se řádně připravit. [1]

Převážně telefonní podvody se syntetickým hlasem, které jsou i hlavním předmětem výzkumu, zaznamenávají nárust ve finančních škodách i v počtu obětí. Závažnost tohoto podvodu tkví v tom, že útočník může mluvit hlasem jakéhokoliv člověka, jen za pomoci několika sekundových nahrávek jeho hlasu. Podvod už se vyvinul do podoby, před kterou jsou lidé aktivně varováni vládními orgány USA. V květnu 2023 před podvodem se syntetickým hlasem varovala federální obchodní komise USA a v listopadu 2023 vydala FBI varování, ve kterém je nová hrozba zmíněna. V tomto varování bylo uvedeno, že FBI zaznamenala 195 případů podvodných telefonátů se syntetickým hlasem. Pro představu byla finanční ztráta v jenom z těchto případů okolo 1,9 miliónů dolarů. [2][3]

Téma jsem si vybral, protože mě oblast sociálního inženýrství zajímá již dlouhá léta. V této oblasti si rád zjišťuji nové trendy a opatření, které mi mohou pomoci se ubránit přílivu nových podvodů a zvýšit mou individuální bezpečnost. Dalším důvodem výběru tématu je pomoc běžným lidem seznámit se s pojmem sociální inženýrství a co vše může v kombinaci s umělou inteligencí napáchat.

Cílem diplomové práce je analyzovat nově vzniklou hrozbu v podobě podvodných telefonátů se syntetickým hlasem a na základě zjištění z analýzy vytvořit protiopatření. Dalším cílem je dozvědět se, jak je aktuálně hrozba vnímána ve společnosti a zda je informace o

existenci hrozby vůbec zachycena mezi lidmi. V neposlední řadě chci v rámci výzkumu zjistit, zda je vůbec možné využít hrozbu i v České republice. První kapitola v teoretické části se věnuje sociálnímu inženýrství včetně technik sociálního inženýrství, které mohou být poháněné umělou inteligencí. Také je zde zmíněn vliv psychologie v sociálním inženýrství a model útoku, který nastínil postup při útoku sociotechnika. Druhá kapitola se zabývá samotnou umělou inteligencí a modely pro generování textového, obrázkového a audio obsahu. Praktická část se zaměřuje na metodiku výzkumu a návrh sociálního experimentu od začátku po jeho konec. V rámci metodiky jsou popsány metody smíšeného výzkumu jako anketa a polostrukturovaný rozhovor. Anketa je cílena na širokou veřejnost pro zjištění míry informovanosti o daném tématu a zda jsou respondenti schopni rozpoznat syntetický hlas od reálného nebo zda aspoň ví, jak se bránit. Celý sociální experiment je koncipovaný jako podvod, při kterém jsou po oběti vyžadovány peníze. Tahle podoba experimentu se zvolila, aby se zjistilo, zda je reálné při aktuálních technologiích provést podvod se syntetickým hlasem i na území České republiky. Veškerá zjištěná data z ankety, sociálního experimentu a rozhovoru jsou analyzována a interpretována ve čtvrté a páté kapitole. Poslední kapitola vychází ze všech poznatků zjištěných z výzkumu, na kterých jsou postaveny opatření vhodná k obraně před syntetickým hlasem.

Práce je přínosná pro kohokoliv, kdo má touhu zvýšit svou bezpečnost v době, kdy jsou podvody s umělou inteligencí na vzestupu. Opatření, která v rámci práce vzniknou, budou moci aplikovat obyčejní lidé, aby se zabezpečili před podvodnými telefonáty se syntetickým hlasem. Dokonce i velké organizace si v nich mohou najít inspiraci a implementovat některé kroky do svých bezpečnostních politik.

I. TEORETICKÁ ČÁST

1 SOCIÁLNÍ INŽENÝRSTVÍ

První kapitola je zaměřena na charakteristiku sociálního inženýrství a další pojmy, které s ním souvisejí. Jeden z těchto pojmů je i psychologie, která je analyzována z pohledu šesti principů ovlivnění a jejich vlivu na osobnostní rysy oběti. Dále je zmíněn obecný postup neboli také model útoku, podle kterého může sociotechnik postupovat. Na závěr jsou popsány techniky, které mohou využívat umělou inteligenci k jejich vylepšení.

1.1 Charakteristika sociálního inženýrství

V oblasti bezpečnosti je sociální inženýrství (SI) chápáno jako umění manipulace a ovlivňování psychiky člověka tak, aby poskytl přístup k neautorizovaným údajům. [4] Sofistikovanější chápání SI poskytli Francois Mouton a kolektiv, kteří tvrdí, že SI je:

„Věda o využití sociální interakce jako prostředku k přesvědčení jednotlivce nebo organizace, aby vyhověli konkrétnímu požadavku útočníka, přičemž sociální interakce, přesvědčování nebo požadavek se týká subjektu, který může pomoci s přístupem k požadovaným informacím.“ [4]

Útoky SI jsou založeny na psychologické manipulaci, při které je oběť nucena stáhnout škodlivý software, sdělit neautorizovanou informaci, navštívit škodlivé stránky nebo provést činnost, která má za následek poškození organizace nebo samotného jedince. SI se zaměřuje především na hledání a využití chyb způsobených lidským faktorem. Využití lidských chyb je mnohem efektivnější způsob získání informace, než pracné a nákladné hledání slabin technických a digitálních systémů nebo vytváření hrozeb v podobě různých malwarů. [5]

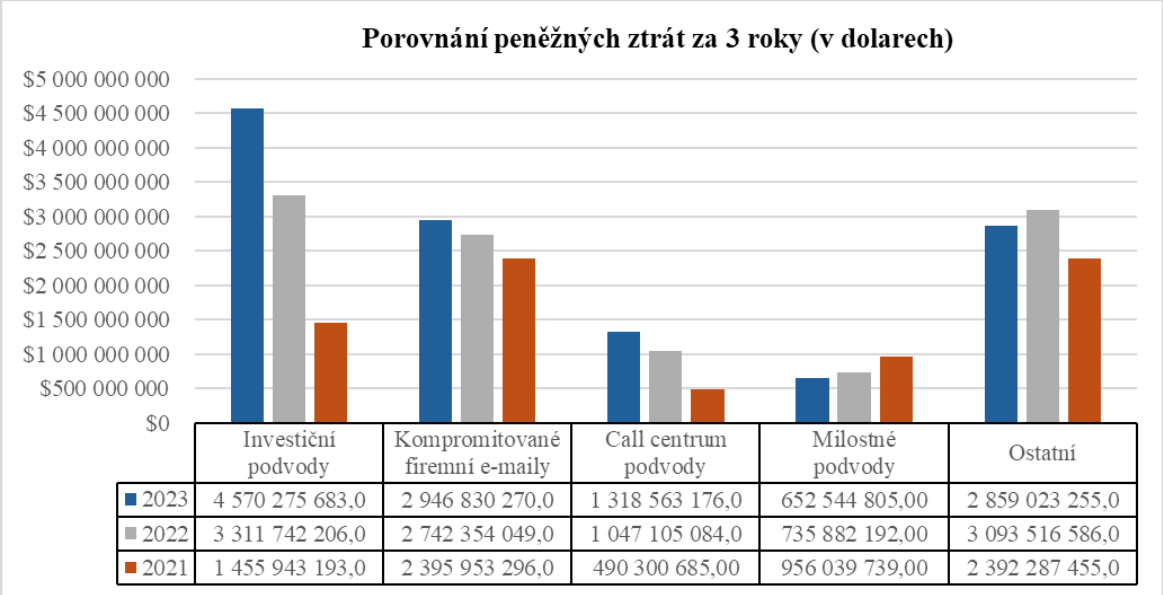
Se SI dále úzce souvisí pojem social engineer. Ten se do češtiny překládá dvěma způsoby jako sociální inženýr nebo sociotechnik. Dále je používán pouze pojem sociotechnik, který se v české literatuře i v jiných zdrojích objevuje častěji. [6][7][8] Sociotechnik je osoba, která provádí manipulaci nebo přesvědčování osob, aby prozradily utajované informace. Tyto informace mohou být dále využity k poškození dobrého jména jedince nebo organizace. Z popisu lze nesprávně usuzovat, že sociotechnika lze brát pouze jako podvodníka, ale nemusí tomu tak být pokaždé. Ozkaya tvrdí, že role sociotechnika může být prováděna negativním nebo pozitivním způsobem. U negativního působení, jak už bylo zmíněno na začátku odstavce, se sociotechnik zaměřuje na poškození jedince nebo organizace. Tuto činnost provádí převážně podvodníci, zloději identity a konkurenční firmy v oblasti průmyslové špionáže. Naopak pozitivní způsob vedení SI vede k legální prosperitě osob nebo organizací.

Příkladem mohou být psychologové, právníci, penetrační testeři nebo personalisté. V dalších kapitolách budou pojmy SI a sociotechnik brány pouze z negativního hlediska. [9][10]

SI je nedílnou součástí kyberkriminality a dle reportu IC3 a společnosti Verizon za rok 2023, je sociální inženýrství stále více se stupňujícím problémem. Report od společnosti Verizon za rok 2023 udává, že 74 % úniku dat bylo způsobeno chybou lidského faktoru a až 50 % útoků sociálního inženýrství jsou založeny na pretextingu. FBI zase předkládá, že za rok 2023 došlo: [11][12]

- V USA ke zcizení 12,5 miliardy dolarů v rámci kyberprostoru.
- V USA bylo napadeno 521 652 obětí (v rámci 20 dalších zkoumaných zemí bylo odhaleno dalších 315 880 obětí).
- K největším finančním ztrátám až 3,4 miliardy dolarů došlo u seniorů nad 60 let. Pro představu u populace pod 30 let došlo ke ztrátě pouze cca 401 milionů dolarů, což je skoro sedmínásobek v porovnání se skupinou seniorů.

Metody, které vykazovaly největší zisky byly založeny na sociálním inženýrství a jednalo se o podvodné investice, kompromitované firemní e-maily, call centrum podvody a milostné podvody (Graf 1). Všechny zmíněné podvody vychází z pretextingu a ztráta pouze u těchto útoků se vyšplhala na cca 9,5 miliard dolarů. Zbytek reportovaných hrozeb byl shrnut do položky ostatní, kde se jedná pouze o cca 2,8 miliardy dolarů (22 % z celkové částky). [11][12]



	Investiční podvody	Kompromitované firemní e-maily	Call centrum podvody	Milostné podvody	Ostatní
2023	4 570 275 683,0	2 946 830 270,0	1 318 563 176,0	652 544 805,00	2 859 023 255,0
2022	3 311 742 206,0	2 742 354 049,0	1 047 105 084,0	735 882 192,00	3 093 516 586,0
2021	1 455 943 193,0	2 395 953 296,0	490 300 685,00	956 039 739,00	2 392 287 455,0

Graf 1: Reportované ztráty od IC3 mezi lety 2021-2023. [11]

1.2 Psychologie v sociálním inženýrství

Psychologie nám poskytuje komplexní pochopení toho, jak lidé myslí, cítí a jednají. Toto pochopení je nezbytné pro efektivní realizaci sociálního inženýrství, protože nám umožňuje předvídat a ovlivňovat chování jednotlivců a skupin. Podle Kahnemana je pro ovlivnění člověka důležité, aby manipulace působila na oběť prostřednictvím jejího systému 1 (rychlé myšlení) nebo systému 2 (pomalé myšlení). Systém 1 lze chápat jako automatickou reakci na přijímané informace, která vychází z doposud naučených znalostí, ale také z evolučně zakořeněných pravidel. Systém 1 tedy reaguje a zpracovává informace za pomoci pouhé intuice. V oblasti SI tedy závisí úspěch ovlivnění oběti na útočníkově vizáži, stylu mluvy nebo chování. Druhý způsob, při kterém je nutné, aby osoba nad informacemi uvažovala, je označován za přesvědčování prostřednictvím systému 2. Vše závisí na důležitosti a srozumitelnosti informací, které jsou předávány oběti, a na jejím následném souladu nebo nesouladu s informací. Pokud komunikace nebude dostatečně přesvědčivá nebo bude složitá, kognitivní myšlení oběti přejde zpět na vnímání skrze systém 1. Sociotechnik využívá oba systémy k ovlivnění rozhodovacího procesu své oběti a k tomu mu dopomáhá celá řada faktorů. Mezi takové faktory patří nedostatek času, nedostatek znalostí o tématu, nedostatek osobní relevance k tématu atd. [13][14]

Cialdiny ve své knize *Zbraně vlivu* hovoří o šesti principech ovlivňování, které se běžně používají v rámci marketingu. [15] Tyto principy se dají účinně využít k přesvědčování a manipulaci lidí i v oblasti SI, což potvrzují autoři jako Francois Mounton a kolektiv, Kevin Mitnick a Christopher Hadnagy. [4][7][16]

Autorita – Hovoří-li se o autoritě, je možné rozlišovat dva významy autority. Autorita, která disponuje odbornými znalostmi jako doktor nebo vědec, u kterých je pro člověka typické nekriticky přejímat informace. Druhý význam spočívá v umístění člověka v rámci hierarchické struktury společnosti. Většina lidí se snáze podvolí policejnímu rozkazu nebo přání jejich nadřízeného, a to i ve věcech, které jsou v rozporu s jejich přesvědčením nebo etikou. Při ovlivnění autoritou postačí útočníkovi pouhé představení se jako autorita nebo doložení symbolu jako odznak, bílý plášť atd., který má stejný vliv. [14]

Reciprocita – Opravdu silná norma v rámci společnosti, která spoléhá na splacení laskavosti těm, od kterých byla přijata pomoc nebo dar. Tvoří základ pro navázání důvěry mezi jednotlivci a spoléhá na potřebu splatit závazky a tím naplnit lidskou potřebu po spravedlnosti. U osoby se silnou reciprocitou je možné očekávat, že splatí svůj dluh mnohem větší laskavostí,

než která byla spouštěčem. Dokonce i když o danou věc nemá oběť zájem, snaží se dárci odvděčit. [14]

Závazek a důslednost – Lidé mají nutkavou potřebu stát neohroženě za svými názory. Tohoto jevu zneužívá závazek a důslednost, jenž dbá na soulad mezi vyřčenými slovy a následným reálným chováním člověka. Mnohdy bývá důslednost tak silná, že osoby stále setrvávají ve svém výroku, i když jim byl poskytnut důkaz o jeho nepravdivosti. [16]

Sociální schválení – Lidé jsou již od počátku sociální bytosti, proto je mnohem jednodušší dojít ke změně názoru jedince, pokud se daný názor objevuje i v jeho okolí. Člověk jako takový se snaží přizpůsobit sociální sféře, v které se nachází. Je pro něj těžké dělat věci, které jdou proti přesvědčení dané sféry a radši pracuje na tom, aby se jí přizpůsobil. [14]

Oblíbenost – Princip spoléhá na to, že lidé raději vyhoví těm osobám, které mají rády a cítí k nim sympatie nebo v nich vidí podobnosti vůči vlastní osobě. Důležitý faktor ovlivňující oblíbenost je fyzická přitažlivost. I když se zdá, že odsouzení lidí na základě jejich vzhledu je v dnešní době pasé, opak je zde pravdou. Fyzická přitažlivost má tak silný vliv, že oběť má často tendenci přiřazovat přitažlivé osobě i vlastnosti, kterými vůbec nedisponuje. Dalším faktorem je podobnost, kdy oběť důvěřuje lidem, ve kterých se zrcadlí její chování, názory nebo jiné prvky. Tato podobnost bývá často tak povrchní, že osoba dokáže více souznít s člověkem, který se stejně jmenuje nebo má narozeniny ve stejný den. [15]

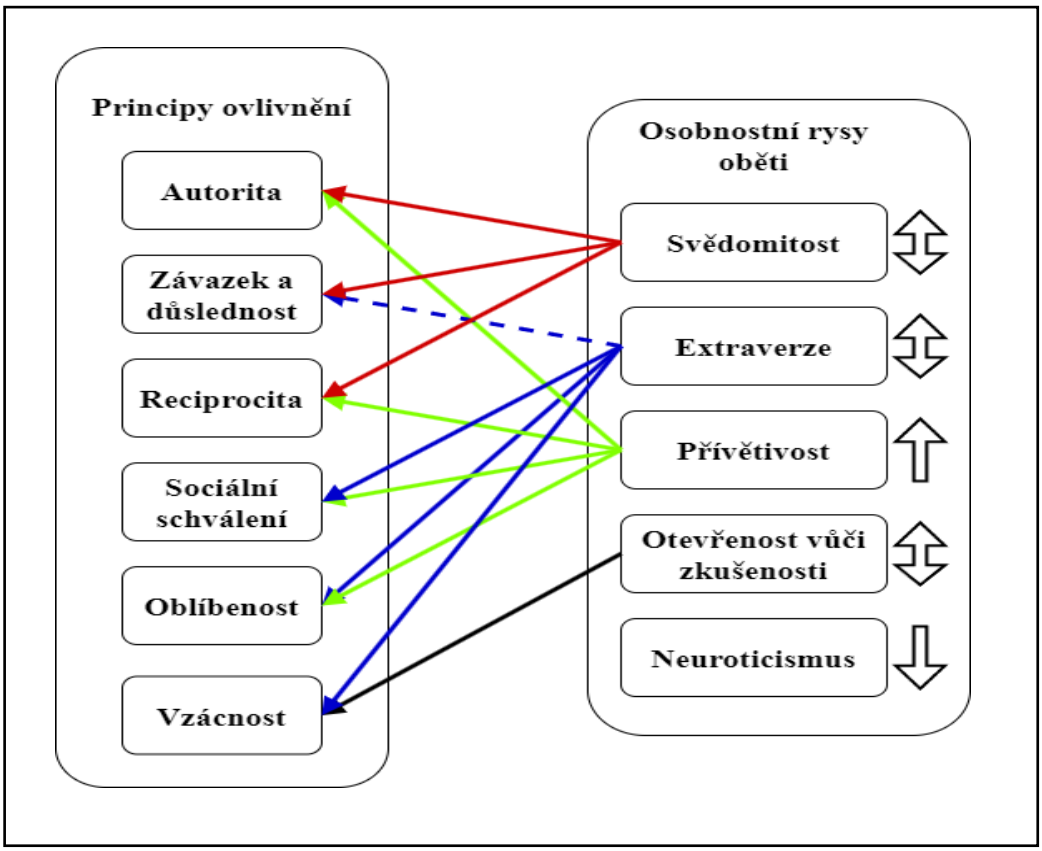
Vzácnost – Pokud je něčeho nedostatek, je v lidské povaze považovat danou věc za vzácnou a upřednostnit ji před ostatními. Jakmile se produkt nebo služba stává nedostatkovým nebo je omezena na určitý časový interval, je v oběti vyvolána silná touha po tom, co se stalo vzácným. To samé platí pro informace. Jakmile se k informacím nedá jednoduše dostat nebo jsou podávány jako exkluzivní, je pro oběť těžší jim odolat. [15]

1.2.1 Vliv osobnostních rysů

Osobnost je možné definovat jako soubor vlastností tvořený emocemi, myšlenkami a chováním jedince. Osobnost každého člověka je jedinečná a její podoba závisí na mnoha proměnných jako genetika a sociální prostředí, ve kterém se osobnost buduje. V oblasti sociálního inženýrství je možné k popisu osobnostních rysů využívat zjednodušený pětifaktorový model. Tento model se skládá z pěti osobnostních rysů, které lze ještě dělit na menší podskupiny. Používá se v mnoha oblastech výzkumů a veškeré úrovně jsou kontinuálně

zkoumány již od 50. let 20. století. Pro sociální techniky je důležité, že mohou využít jak pozitiv, tak negativ jednotlivých úrovní. [14]

Na obrázku (Obrázek 1) níže jsou viditelné vazby, které stanovují náchylnost osobnostních rysů vůči principům ovlivnění. Plné čáry uvádí zvýšenou pravděpodobnost úspěchu útoku a čárkované sníženou. Šipky vedle osobnostních rysů přesněji popisují, zda zvyšují, snižují nebo nijak nepůsobí na proveditelnost útoku. [14]



Obrázek 1: Asociace mezi osobnostními rysy a principy ovlivnění. [14]

Tento model je pouze obecný a zkoumá převážně osoby, které se charakterizují daným rysem. Nebyly brány v potaz všechny negativně se projevující osobnostní rysy, ale i tak tento model poskytuje důležité informace o vlivu psychologie v SI. Dále jsou již popsány jednotlivé rysy pětifaktorového modelu a jejich zneužití v sociálním inženýrství. [14]

Svědomitost – Tento faktor popisuje osobu jako spolehlivou, pořádkumilovnou, disciplinovanou, pracovitou a precizní. Opakem osobnostního rysu jsou osoby nedbalé, líné, lhostejné a nespolehlivé. [17]

Svědomitá osoba je odolnější proti hrozbám, pokud dodržuje pravidla a normy, které jsou stanoveny a této osobě sděleny. Vážné ohrožení může hrozit od útočníků využívající

autority, kdy jsou schopni svědomitým lidem nastavit jiná pravidla. Další možnosti ovlivnění jsou reciprocita, závazek a důslednost. Co se týče závazku a důslednosti, má tento princip vliv pouze tehdy, když je závazek veřejný nebo souvisí s pravidly. Naopak u principu oblíbenosti, sociálního důkazu a vzácnosti se náchylnost neprojevila, protože se zde neapeluje na žádná pravidla, které by musela oběť splnit. Pokud se v organizaci nachází osoba v kontrastu se svědomitým člověkem, je možné osobu označit za nespolehlivou, nezodpovědnou a náchylnější na nebezpečí. [14]

Extraverze – Silná potřeba založená na kontaktu s ostatními lidmi. Extrovertní lidé jsou hovorní, aktivní, sociabilní, optimističtí a mají silnou orientaci na druhé. V opačném případě jsou lidé tiší, nekomunikativní a uzavření do sebe. [17]

Lidé s extrovertními vlastnostmi často porušují bezpečnostní zásady, aby vyhověli požadavkům útočníka a zalíbili se mu. Tohoto faktu mohou využít principy oblíbenosti a sociálního důkazu. Extrovertní lidé často vyhledávají vzrušení. Když tedy získají něco vzácného, mohou podlehnout vzrušení a být náchylnější na princip vzácnosti. Kladem pro extravertní lidi může být menší citlivost vůči závazku a důslednosti. Takový jedinec má menší tendenci splnit závazek, ke kterému se zavázal. Větší odolnost vůči útokům vykazovali lidé introvertní, kteří jsou méně společenští a jsou pravým opakem extrovertních osob. [14]

Přívětivost – U tohoto faktoru se předpokládá, že osoby jsou konzervativnější, dobrosrdečnější, laskavější a ochotni pomáhat druhým. Osoby vykazující opačné vlastnosti jsou vnímány jako pomstychtivé, podezíravé a bezcitné. [17]

Přívětiví jedinci jsou mnohem náchylnější na rizika, jelikož jsou důvěřivější než jiné osobnostní rysy a jsou méně skeptičtí vůči narušení vlastního soukromí. Tento typ jedince může být jednoduše ovlivněn, pokud je navázán vztah založený na důvěře. Ve vztahu mezi osobou a sociálním technikem mohou být efektivně využity techniky ovlivnění jako apel na autoritu, reciprocita a další techniky založené na zneužívání důvěřivosti. [14]

Otevřenost vůči zkušenosti – Tito jedinci mají velkou touhu k vyhledávání nových zážitků, objevování a tolerování neznámého. Jedinci s vysokou mírou otevřenosti tíhnou ke zvědavosti, všestrannosti, tvořivosti a originalitě. Lidé s opačnými rysy se vyznačují konvenčností, neuměleckostí a přízemností. [17]

Otevření lidé vyhledávají nové zkušenosti s menšími obavami o jejich bezpečí a tím pádem se mohou stát jednoduchým cílem pro útočníky. Tito jedinci špatně hodnotí rizika, která jim mohou hrozit a značně je podceňují. Nastává tedy situace, při které otevření jedinci častěji

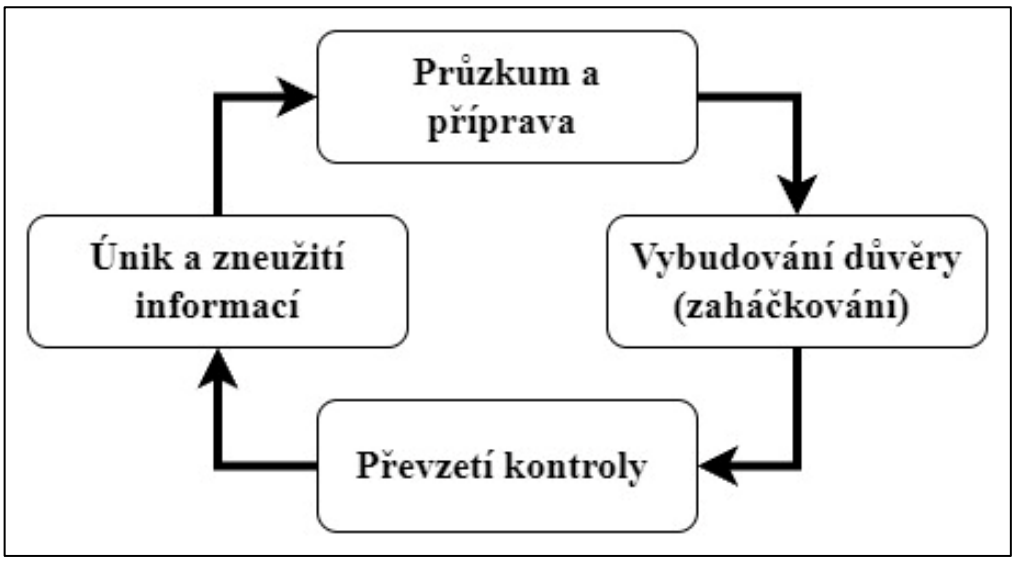
porušují pravidla kybernetické bezpečnosti tím, že otevřou upravený odkaz nebo stáhnou infikovaný soubor. Výše uvedený model počítá s tím, že osoba otevřená vzdělávání je odolnější vůči hrozbám a má technologické a jiné znalosti, které ji činí mnohem odolnější vůči hrozbám. Jedinou náchylností je zde vzácnost nebo také nedostatek, jež může v otevřeném jedinci iniciovat pocit ohrožení svobody, což je pro něj nemyslitelné. [14]

Neuroticismus – Může vést ke stavům nejistoty, paranoii, pesimismu, lability a neklidu. Lidé s nízkým skóre neuroticismu jsou mnohem stabilnější, vyrovnanější, spokojenější a klidnější. [17]

Lidé projevující se neuroticismem jsou méně náchylní na kybernetické útoky, jelikož nedůvěřují ostatním a nezveřejňují osobní informace. To má za následek větší odolnost vůči phishingovým emailům, nevyžádaným telefonátům a dalším typům útoku. Pesimismus je u neurotických jedinců základním prvkem jejich obrany, jelikož je vždy očekáváno nejhorší možné vyvrcholení situace. Proto se v rámci modelu počítá i s menší náchylností na manipulaci. [14]

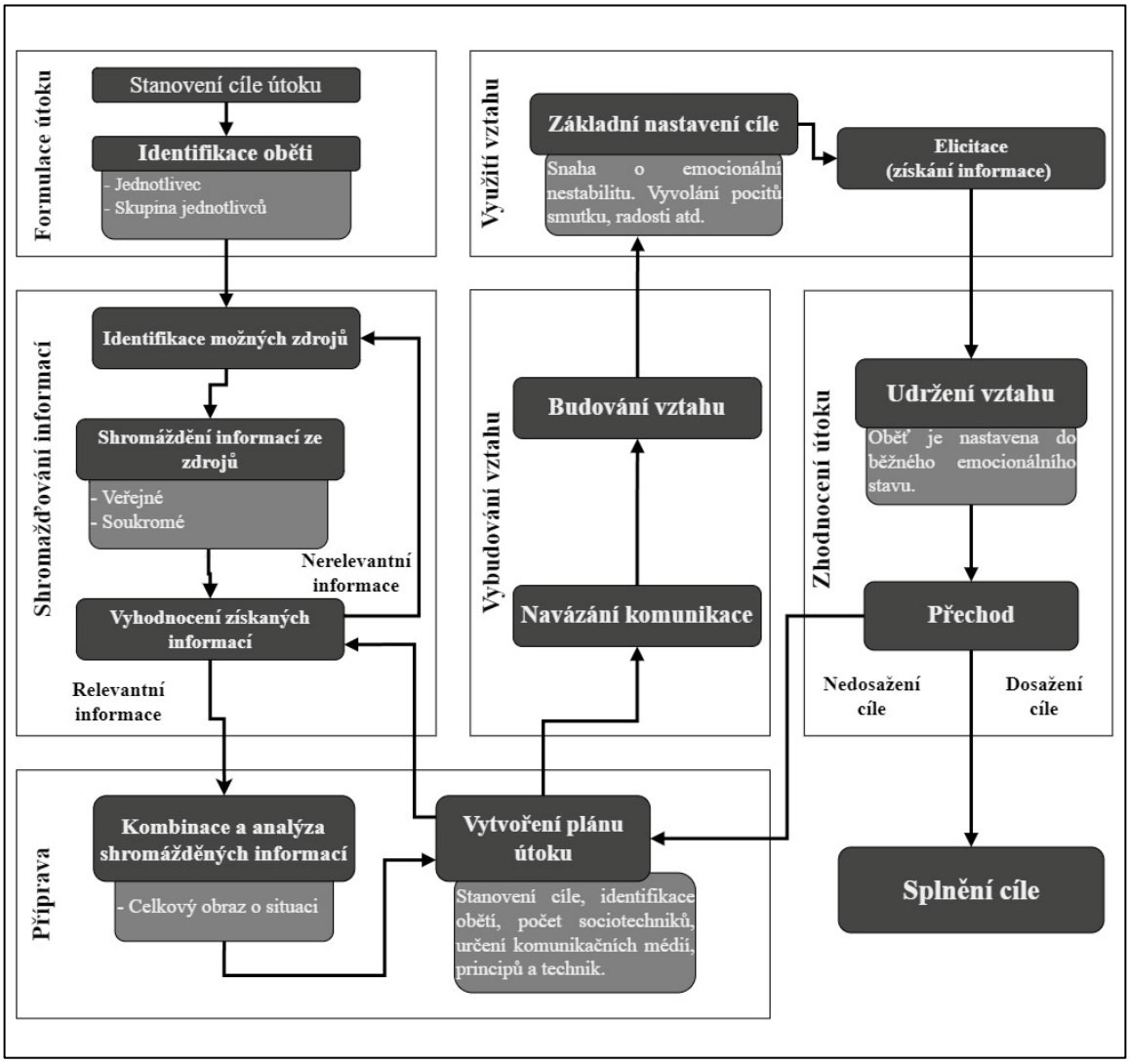
1.3 Model útoku sociálního inženýrství

Ačkoliv je každý útok SI unikátní, je možné konstatotvat, že se každý útok řídí podle určitých skutečností. Zjednodušený a velmi spolehlivý model útoku stanovil Kevin Mitnick ve své knize *Umění klamu*. V této knize stanovil životní cyklus útoku (Obrázek 2), který se skládá z průzkumu a přípravy, vybudování důvěry (zaháčkování), převzetí kontroly a v neposlední řadě úniku a zneužití informací. [7]



Obrázek 2: Základní model dle Mitnicka. [7]

Model Kevina Mitnicka je sice základním popisem, jak postupovat během útoku, ale jedná se pouze o stručný popis. Mitnickův cyklus je totiž mnohem složitější, než se může na první pohled zdát a obsahuje mnohé skutečnosti, které autor nezmínil. Podrobnější náhled byl poskytnut Francoisem Mouton a kolektivem (Obrázek 3), kteří základní cyklus rozpracovali do podrobného modelu a poukázali na potřebu zpětnovazebních smyček. Základní model tedy představoval útok jako neustále se opakující smyčku. Zatímco model dle Moutona říká, že je potřeba na konci vybraných fází zhodnotit úspěšnost a podle vyhodnocení se buď vrátit na začátek aktuální nebo jiné fáze nebo pokračovat dál. [18]



Obrázek 3: Model útoku dle F. Moutonu a kolektivu. [18]

Oproti základnímu modelu dle Mitnicka je tento model rozšířen o fázi **formulace útoku**, ve které se specifikuje cíl útoku a identifkuje oběť. Stanovení cíle musí být zcela jednoznačné, aby se dle něj dala vybrat ideální oběť. Co se týče oběti, zde je možné vybrat jednotlivce nebo skupinu jednotlivců, která má přístup k požadovaným informacím a jejich ovlivnění pomůže k dosažení cíle. [18]

Aby bylo možné vybudovat silnou důvěru mezi sociotechnikem a obětí, je nutné shromažďovat kvalitní informace. Ve druhé fázi tedy dochází ke **shromažďování informací.** V prvním kroku se nejprve identifikují možné zdroje, ze kterých se ve druhém kroku získávají informace. Tyto zdroje mohou být veřejné, jako např. webové stránky firmy, sociální sítě, blogy a fóra nebo soukromé. Pro získání soukromých zdrojů je navíc potřeba proniknout na určitá místa, dostat se do osobních účtů nebo použít tzv. dumpster diving, při kterém se prohledávají popelnice kvůli zisku důvěrných dokumentů. Celý proces se opakuje, dokud sociotechnik ve třetím kroku neusoudí, že má dostatek relevantních zdrojů. [18]

Ve fázi **příprava** dochází k finální kontrole všech informací před započetím konfrontace oběti. Prvním krokem je kombinace všech získaných informací na jejichž základě se vytvořil celkový obraz útoku. Následující krok už jen vychází ze všech informací a vytváří se plán útoku, ve kterém jsou specifikovány prvky jako stanovení cíle, identifikace oběti, výběr komunikačního média (přímo nebo nepřímo), počet sociotechniků, princip ovlivnění nebo využitá technika SI. [18]

Vybudování důvěryhodného vztahu s obětí je zásadní fází pro dosažení cíle. Není-li vybudován dostatečně silný vztah, možnost úspěchu rapidně klesá. S obětí je na začátku fáze navázán kontakt prostřednictvím určeného média. Komunikace se odehrává buď osobně, e-mailem, na sociálních sítích nebo jiným způsobem. Časově velice náročným krokem je finální budování vztahu s obětí, ve kterém se aktivně používají principy ovlivnění a techniky SI. [18]

Pokud se podaří s obětí navázat plnohodnotný vztah, přechází se k **využití vztahu**. Oběť je nejprve nastavena do požadovaného emočního rozpoložení, při kterém je pro ni velmi těžké kontrolovat své reakce. Pro oslabení oběti mohou pomoci stavy smutku, štěstí, deprese, vzteku atd. Když je cíl vyveden z emocionální stability, může sociotechnik provést elicitaci (získání informace) a dostat se k požadované informaci. Takovou informací mohou být hesla, citlivé údaje, firemní tajemství a mnohé další. [18]

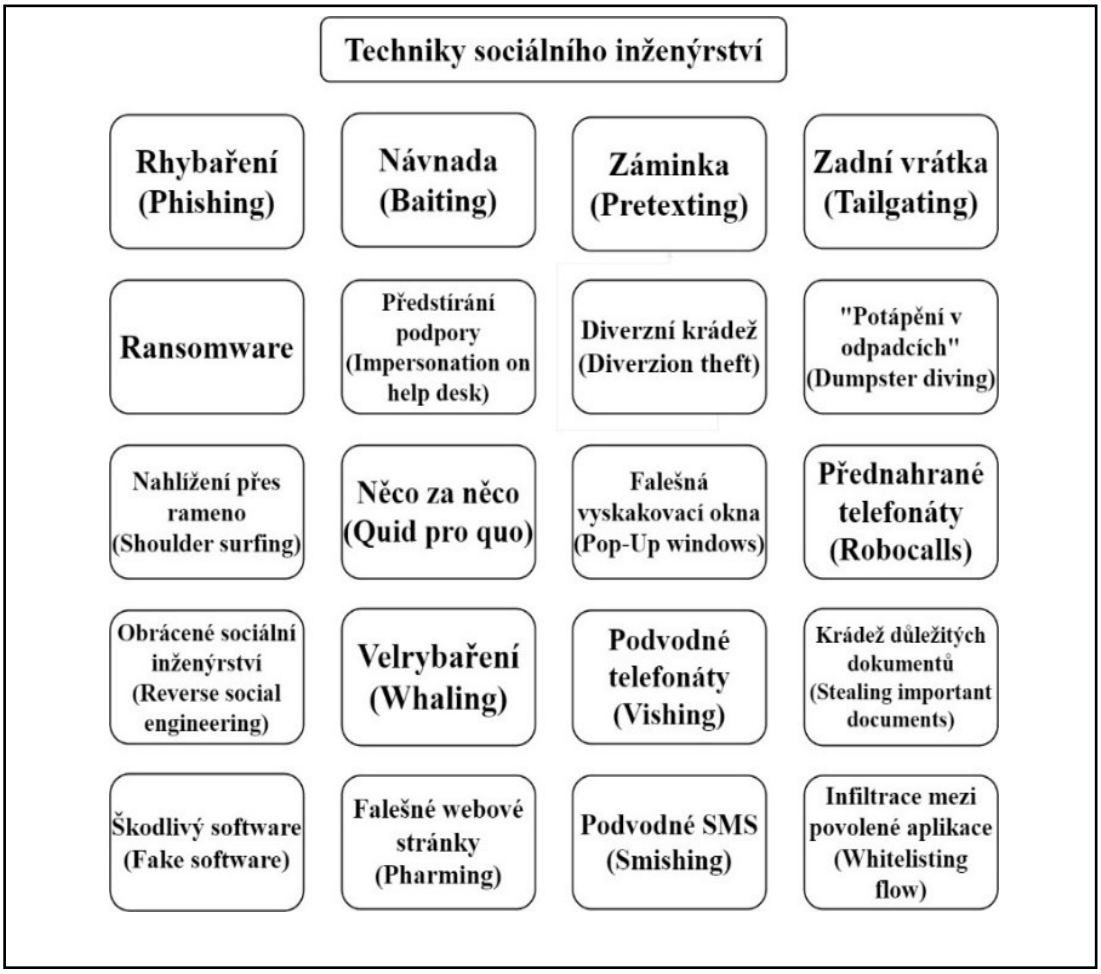
Na počátku **zhodnocení útoku** je ještě potřeba navrátit oběť do emocionálně stabilního stavu. Tím se docílí toho, že oběť nebude příliš přemýšlet o situaci, která nastala, a sociotechnik nadále zůstane v anonymitě. V úplném závěru se sociotechnik rozhoduje o úspěchu či neúspěchu útoku. Splní-li se cíl útoku, je útok ukončen a informace zneužita. Pokud je tomu naopak, přechází se zpět do přípravné fáze a dle nových skutečností se upravuje plán nebo přechází zpět ke sběru informací. [18]

Na závěr této podkapitoly je potřeba dodat, že je spousta modelů popisujících průběh útoku. Tento model byl zmíněn proto, že mi přišel jako nejkomplexnější a poskytoval ucelený pohled na útok od jeho začátku po jeho konec. Samozřejmě, že jsou zde další podobně kvalitní modely, kdy určitě za zmínku stojí pyramidový model dle Hadnagy, který je tvořen z pěti úrovní a každá úroveň slouží jako pevný základ pro další. Pyramida je tvořena sběrem informací, vytvořením záminky, naplánováním útoku, započetím útoku a zhodnocením útoku. Postup připomíná model dle Mountona, ale chybí mu určité fáze a zpětnovazební smyčky, které by jej činili mnohem informativnějším. [16]

1.4 Techniky sociálního inženýrství

Techniky SI jsou souborem metod a strategií, které se používají k manipulaci s lidským chováním, emocemi a myšlením za účelem dosažení stanovených cílů. Techniky SI jsou rozmanité a je těžké vytvořit systemizaci, která by každé technice našla své místo. Přesto zde existuje hojně používané rozdělení na útok vedený skrze techniku a útok založený na osobním kontaktu. Bohužel i toto rozdělení není všemocné a techniky z jedné kategorie mohou spadat i do druhé. Pro základní popis je ale dané rozdělení postačující. [19]

Útoky technického charakteru využívají média jako e-mail, telefon nebo jiná média, která umožňují vzdálenou manipulaci. Sociotechnik je skryt za počítačem nebo jiným zařízením, které mu umožňuje útočit na více obětí, a přitom zůstat v anonymitě. U útoků založených na osobním kontaktu sociotechnik osobně komunikuje s obětí a snaží se od ní získat informace. Mezi tyto útoky patří i fyzické techniky založeny na prohledávání počítačů, odpadkových košů atd. přímo v organizaci za fyzické přítomnosti sociotechnika. Fyzická přítomnost sociotechnika skrývá značná rizika, mezi něž patří malý počet lidí, které lze ovlivnit, zvýšená pravděpodobnost odhalení apod. Na obrázku níže je možné vidět některé techniky SI (Obrázek 4). [19]



Obrázek 4: Některé techniky sociálního inženýrství. [19]

V dalších podkapitolách budou rozebírány jednotlivé techniky, při kterých je možné využít umělou inteligenci (dále jen UI). Dopodrobna je UI rozebrána v kapitole 2. Techniky SI poháněné UI jsou více nebezpečné kvůli jejich sofistikovanosti, automatizaci a hlavně personalizaci. Nové funkce, které UI skrývá, mohou značně zefektivnit přípravu a průběh SI útoku, a to těmito způsoby: [20]

- Analyzování velkých souborů strukturovaných i nestrukturovaných dat, ze kterých může extrahovat důležité informace zneužitelné pro útok. Tímto způsobem lze ve velmi nízkém čase identifikovat zranitelnosti jedince nebo slabá místa systémů.
- Na základě získaných informací je možné generovat různé možnosti provedení útoku a poté na základě získaných dat vytvořit optimalizovaný plán.
- UI může na základě vstupních dat generovat texty, které svým stylem přesně odpovídají stylu psaní známých, vedoucích atd.
- Personalizované texty je možné vytvářet i v jiných jazycích s mnohem menším počtem chyb, než tomu bylo do teď.

UI může generovat textový, obrázkový, zvukový, ale taky video obsah. V rámci generování textu se sociotechnik zaměřuje na vytváření personalizovaných e-mailů nebo falešných dokumentů, které napodobují styl textu blízké osoby nebo autority. Generování obrazového a video obsahu je vhodným základem pro tvorbu deeepfake videí nebo upravených fotek. Poslední ze zmíněných je generování zvukového nebo také hlasového obsahu pro napodobení hlasu osob v blízkém vztahu k oběti. Zbytek informací ohledně UI bude uveden v kapitole 2. Dále jsou řešeny pouze techniky, které jsou v současnosti již vystaveny vlivu UI [20] [21]

1.4.1 Pretexting

Pretexting (záminka) funguje na principu použití falešného a smyšleného příběhu, který je určen k zmanipulování oběti za účelem získání citlivých informací, stáhnutí malwaru nebo k jiné činnosti, která by měla za následek poškození vytěžované osoby nebo organizace. Metodu pretextingu je možné považovat za jádro dalších technik jako jsou phishing, vishing a mnohé další. Zvládá však efektivně fungovat i sama o sobě. [19]

Dle reportu FBI za rok 2023, který se týkal internetových zločinů, je technika pretextingu jedna z nejnebezpečnějších metod. V rámci následujících příkladů z reportu je možné vidět, jak efektivní může metoda pretextingu být: [11]

- Investiční podvody – V podvodu je útočník představen jako finanční poradce, který nabízí rychlé zhodnocení peněz až o několik desítek procent. Nejčastější investiční podvody se týkaly kryptoměn, u kterých je jednoduché argumentovat dobrým zhodnocením. Tento podvod je uváděn pro rok 2023 jako nejzávažnější a byl zaznamenán obrovský nárust oproti roku 2022, kdy ztráta vzrostla z 3,31 na 4,57 miliard dolarů. [11]
- Kompromitace firemních e-mailů – Útočník se vydává za vedoucího nebo důležitého obchodního partnera firmy, který má ve firmě určitou autoritu. Cílí na zaměstnance firmy, které žádá o pomoc s přihlášením do firemního bankovního účtu nebo požaduje zaplacení přiložené faktury. Tímto podvodem bylo za rok 2023 získáno přes 2,94 miliardy dolarů. [11]
- Call centrum podvod – Podvod začíná telefonátem, kdy se podvodník nejčastěji představuje jako technická podpora nějaké služby nebo státní zaměstnanec (policista, finanční úředník atd.). Snaží se cílovou osobu přesvědčit, že má potíže jako např. nezaplacené pokuty, nezaplacené daně, napadení účtu atd. Pod výhružkou pokuty a

časového nátlaku je oběť nucena k zaslání peněz útočníkovi. Celková ztráta se vyšplhala na 1,31 miliard dolarů. [11]

Úspěch pretextingu je možné značně ovlivnit za pomoci UI, a to využitím generativních modelů. Především modely pro generování textu dokážou zpracovávat a napodobovat schopnost přirozeného jazyka jedinců a záhy na to generovat rozsáhlé texty přizpůsobené na míru oběti. Textový model je dokonce možné naučit i chyby a různé typy slov, které oběť běžně využívá a tím vytvořit text, který je těžce rozeznatelný od originálu. Další výhodou v používání UI je v schopnosti generovat text v jiných jazycích. Tím je eliminována stávající slabina pretextingu, ale i phishing e-mailů, které obsahovaly spoustu pravopisných chyb. Všechny zmíněné výhody jsou posunuty ještě o level výše, jelikož UI může celý proces výběru oběti a generování textu zautomatizovat. To vkládá útočníkovi do rukou možnost pracovat individuálně a být stejně nebezpečný jako organizované skupiny podvodníků. [22]

1.4.2 Phishing

Technika založena na bezkontaktním způsobu, během kterého bývá oběti zaslána věrohodně vypadající zpráva, jenž nabádá oběť ke stažení škodlivého softwaru nebo poskytnutí citlivých údajů. Ve většině případů se jedná o podvodné e-maily nebo podvržené webové stránky, ale používají se i telefonáty, SMS nebo Pop-Up reklamní okna. Podle způsobu vykonání útoku můžeme phishing dělit na: [23]

- Email phishing – Nejrozšířenější typ phishingu, kdy útočníci rozesílají co nejvíce podvodných e-mailů od legitimní organizace. E-mail se snaží vzbudit pocit ohrožení a donutit oběť k úkonu jako je kliknutí na škodlivý odkaz nebo přihlášení na falešné stránce. [23]
- Spear phishing – Oproti předešlé metodě se spear phising zaměřuje na jednotlivce v organizaci, na které cílí s více personalizovanými e-maily. K této metodě je potřeba provést dostatečně velký průzkum oběti, aby bylo možné e-mail co nejlépe zacílit. [23]
- Whale phishing – Útočník se nezaměřuje na podřadné pracovníky, ale útok vede vůči vrcholovému managementu, jinak řečeno jde po "velkých rybách". [23]
- Social media phishing – Útočník si vytvoří falešný profil na sociálních sítích jako je Facebook, TikTok a Instagram, aby z obětí získal citlivé údaje. [23]

- Smishing – funguje na podobném principu jako e-mailový phishing, ale místo e-mailu jsou používány zprávy SMS. [23]
- Další podkategorie mohou být třeba clone phishing, pharming nebo v další podkapitole více rozpracovaný vishing. [23]

I zde v rámci phishingu má UI své uplatnění. V e-mailové konverzaci mohou modely pro generování textu pomoci s personalizací e-mailů na základě vstupních dat. Tato data jsou tvořena informacemi o oběti, které lze získat ze sociálních sítí, e-mailové komunikace atd. To má za následek rapidní zvýšení výkonnosti personalizovaných phishing e-mailů jako spear phishing a whale phishing. Rozpoznatelným znakem phishing e-mailů byla často špatná gramatika a velké množství chyb, jelikož útočník nebyl rodilým mluvčím. Proto útočníci začali využívat modely UI, které umožňují přeložit zprávy a zachovat správnost sdělení i v jiném jazyce. [24]

Nově vzniklou podkategorií phishingových útoků je Indirect Prompt Injection. Tento útok vznikl na základě implementace chatbotů (Copilot v Microsoft Edge) s UI přímo do prohlížečů. Tito chatboti využívají velké jazykové modely, které jsou schopny analyzovat prohlíženou webovou stránku a podat o ní uživateli souhrn nebo jiné informace. Metoda Indirect Prompt Injection upravuje vstupní data do jazykových modelů, aby ovlivnila jeho chování a chatbot místo souhrnu stránky nebo rad vylákal z oběti platební údaje a odkázal ji na falešnou stránku. Samotný obsah na falešné stránce nemusí vypadat podezřele, ale mezi pixely jsou skryty řádky škodlivého textu, které nejsou na první pohled patrné. Chatbot takový text bez problému rozezná a bere jej jako normální obsah stránky. Takto provedená injektáž promění chatbota na sociotechnika, který se může vydávat za podporu daného prohlížeče a získávat z oběti osobní údaje. [25]

1.4.3 Vishing

Tato technika je podkategorií phishingu, někdy také známá jako hlasový phishing. Útočník se pokouší prostřednictvím telefonního hovoru získat od své oběti finanční prostředky, osobní nebo citlivé údaje. Během telefonátu se útočník představí jako zaměstnanec vládní organizace, bankéř nebo technická podpora. Snaží se vystavit oběť situaci, kterou je potřeba v krátkém časovém intervalu řešit, jinak by byla oběť sankcionována nebo by přišla o lukrativní příležitost. [26]

Útočníkům hraje do karet i tzv. spoofing, který zvládá napodobovat telefonní čísla. Oběti se mohou zobrazovat na telefonu jakákoliv čísla, a to i čísla organizací jako Policie ČR nebo bankovní instituce. [26]

Další velmi oblíbený nástroj pro útočníky jsou telefonáty za pomoci Voice over Internet Protocol (VoIP). Telefonát zprostředkovaný prostřednictvím internetu je pro bezpečnostní složky hůře dohledatelný, což útočníkům poskytuje další vrstvu ochrany. [26]

Vishing je čím dál více spojován z uměle generovaným hlasem. Jak bylo zmíněno v úvodu i federální obchodní komise USA varovala před vishingovými útoky, při kterých byl využíván uměle generovaný hlas rodinných příslušníků k podvodům. Útočník nemusí kopírovat pouze hlas členů rodiny, ale stejně efektivně napodobí hlas šéfa, spolupracovníka nebo kohokoliv. [2][20]

V rámci zasedání rady v USA, která řešila hrozbu podvodů s využitím UI, prohlásil Tahir Ekin, profesor a ředitel Centra pro analýzu a datovou vědu, že: [27]

"Umělá inteligence zesiluje dopad podvodů, zvyšuje jejich věrohodnost a emocionální působivost díky personalizaci." [27]

Poukázal tím tak na jeden z bodů zasedání, který se týkal pana Garyho, kdy útočník zkopíroval hlas jeho syna, za kterého se následně vydával. Falešný syn svému otci sdělil, že způsobil autonehodu v opilosti a byl umístěn do cely předběžného zadržení. Volající po vysvětlení situace informoval otce, že jej zastupuje veřejný ochránce Barry Goldstein a tento právník ho bude kontaktovat. Podvodník se v dalším telefonátu vydával za Barryho Goldsteina a vyžadoval po panu Garym, aby za syna zaplatil 10 % z kauce, tedy cca 9 000 dolarů prostřednictvím Bitcoinu. Případ naštěstí dopadl dobře, jelikož pan Gary zavolal manželce svého syna, aby se dozvěděl více informací. K jeho překvapení se situace s nehodou vůbec nestala. Tento případ byl jeden ze 195 případů podvodů se syntetickým hlasem, které byly v roce 2023 v USA reportovány. [3][27]

1.4.4 DEEPFAKE

Princip deepfake spočívá v záměně reálné osoby za osobu digitálně upravenou a celý proces je založen na hlubokém učení (HU). To znamená, že UI z dostupných zdrojů nasnímá tvář požadované osoby a následně může promítat falešně vytvořenou projekci na tvář útočníka. Deepfake se nemusí používat pouze pro záměnu obličeje, ale pro celkovou úpravu videa

nebo fotek. Prostředí okolo nebo i hlas samotného útočníka může být upraven tak, jak si sociotechnik zamane. [28][29]

Deepfake lze použít dvěma způsoby. U prvního způsobu se upravuje již existující obsah tak, že jsou do úst oběti vkládána slova, která nebyla řečena nebo jsou vytvářeny činnosti, které nebyly provedeny. U druhého způsobu je útočníkova tvář nebo i celá postava zcela nahrazena za jinou synteticky vytvořenou postavu, kterou chce útočník zdiskreditovat. [29]

Využití deepfake generovaného obsahu lze zneužít pro dobré, ale i špatné účely. Mezi ty dobré se řadí CGI efekty pro filmy, umění, podpora support center a další. Pokud je řeč o špatné aplikaci deepfake, je na místě zmínit vytváření nekonsensuální ženské pornografie, která aktuálně tvoří 96 % generovaného deepfake obsahu na internetu. Mimo sexuální hrozby je deepfake zneužitelný pro páchání podvodů, ovlivnění politické situace nebo vytváření falešných důkazů. [29][30]

Nejbližší dobře zdokumentovaný případ zneužití deepfake nástroje a syntetického hlasu pro personalizovaný útok byl na Slovensku. Majitel firmy GymBeam, Dalibor Cicman informoval na svém LinkedIn profilu o zkušenosti jeho zaměstnance. Uvedl, že jeho zaměstnance kontaktovali přes MS Teamsy dva podvodníci. Jeden z podvodníků byl digitálně upraven do podoby pana Cicmana. Druhý podvodník byl představen jako externí právník, který po většinu času mluvil a snažil se získat informace o bankovním kontě firmy. Podvod byl prozrazen digitální kopií pana Cicmana, který uvedl, že je 2 týdny mimo kancelář. V tu chvíli byl podvod prozrazen, jelikož se zaměstnanec s panem Cicmanem v daný den setkal. Samotný zaměstnanec k tomuto dodává, že kopie vypadala velice věrohodně a podezřelá byla pouze forma komunikace přes MS Teamsy, protože firma komunikuje převážně na jiných platformách. [31]

1.5 Shrnutí

V první kapitole byly shrnuty základní pojmy a popsána závažnost útoku SI, které jen v USA mají každý rok za následek ztráty v jednotkách miliard dolarů. Následně byl podrobněji popsán vliv psychologie v oblasti SI a model útoku podle Moutona, který poskytl náhled na celý průběh útoku. V poslední části kapitoly byly představeny techniky SI, které jsou vylepšeny UI. Překvapením bylo, že UI nejenže vylepšuje stávající techniky, ale na jejím základě vznikají i úplně nové hrozby jako Indirect Prompt Injection. V následující kapitole bude popsána základní terminologie UI a hojně rozšířené modely jako ChatGPT, Gemini atd.

2 UMĚLÁ INTELIGENCE

Tato kapitola bude pojednávat o základní kategorizaci a pojmech týkající se umělé inteligence. Jsou zde řešeny pojmy jako strojové učení, hluboké učení a neuronové sítě. Kapitola je převážně zaměřena na modely, které generují textový, obrázkový nebo audio obsah.

2.1 Definice a základní pojmy umělé inteligence

UI představuje jedno z nejrychleji rostoucích odvětví moderní informatiky. Tento dynamický a rozmanitý obor přináší nejen technologické inovace, ale také klade otázky ohledně etiky, soukromí a bezpečnosti. UI lze tedy definovat jako:

„podobor informatiky, který se zabývá konstrukcí strojů schopných simulovat lidskou inteligenci.“ [32]

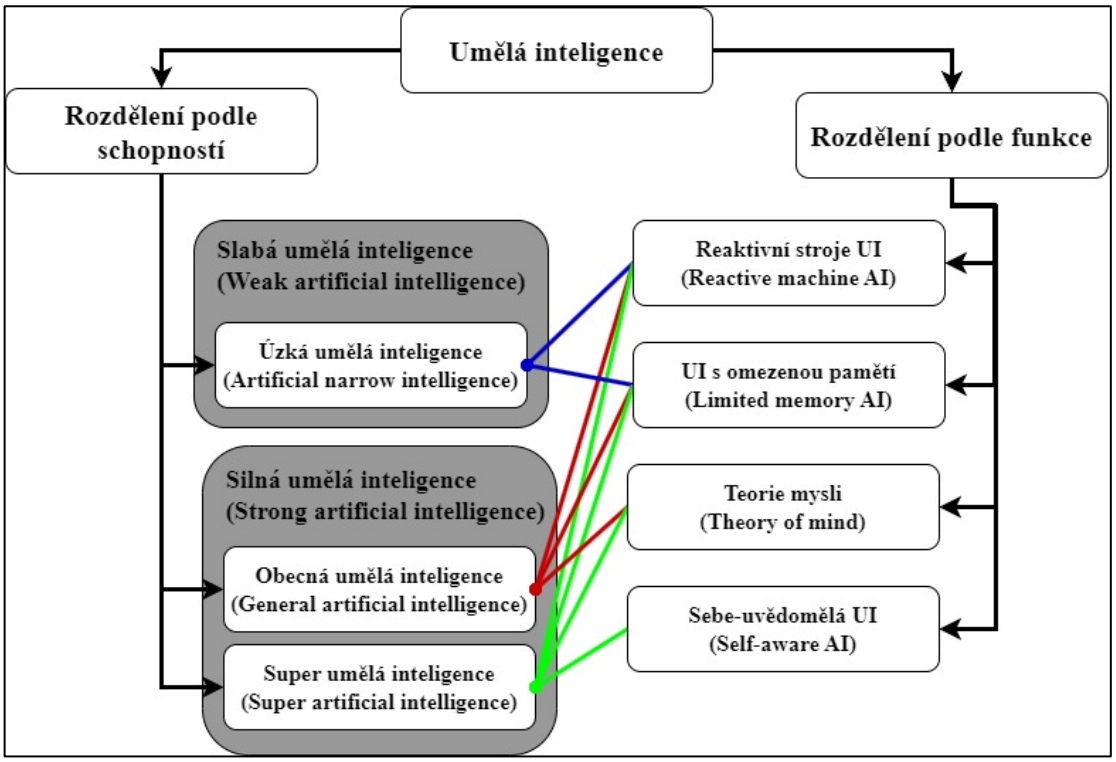
Úkolem UI je tedy umožnit počítačům provádět činnosti napodobující lidské chování jako je čtení, tvoření, psaní, učení a další. UI lze klasifikovat dle schopnosti zvládat napodobování lidského chování a podle funkcí, které do dané problematiky přináší. [33]

Rozdělení na základě schopností UI:

- Úzká UI – V dnešní době nejrozšířenější typ UI. Tato UI se nazývá úzká, protože se trénuje a zaměřuje na specifické úkoly, které dovede zautomatizovat a vykonat rychleji než člověk. Příkladem takové inteligence můžou být autonomní auta, hlasový asistenti a chatboti. [34]
- Obecná UI – typ UI, která se zvládne naučit a plnit jakoukoliv intelektuální činnost, kterou zvládne lidská bytost. Tento model je v současnosti zatím hypotetický, jelikož se lidstvo nenachází ve fázi, kdy by bylo schopné naprogramovat veškeré funkce mozku. [34]
- Super UI – Pokud se obecná UI rovná té lidské, tak tzv. super UI ji ve všem převyšuje. Super UI si je vědoma sama sebe, cítí a prožívá emoce a taky má potřebu naplňovat své potřeby a touhy. Stroje vybaveny super inteligencí jsou schopny předčít člověka v jakékoliv oblasti od medicíny po jadernou fyziku. Super UI je obdobně jako předešlý typ UI pouze teoretická. [34][35]

Předešlé rozdělení se často nahrazuje pojmenováním slabá a silná UI. Tohle rozdělení také referuje právě na schopnost UI řešit úkoly. Pokud je zmíněna slabá UI, je tím myšlena úzká UI, která zvládá řešit pouze úzkou oblast problémů. Do silné UI poté spadá obecná a super

UI, které už řeší komplexní úkoly na úrovni rovné nebo převyšující lidskou inteligenci. Základní rozdělení na schopnosti a funkce UI je na Obrázku 5. [36]



Obrázek 5: Rozdělení UI. [36][37]

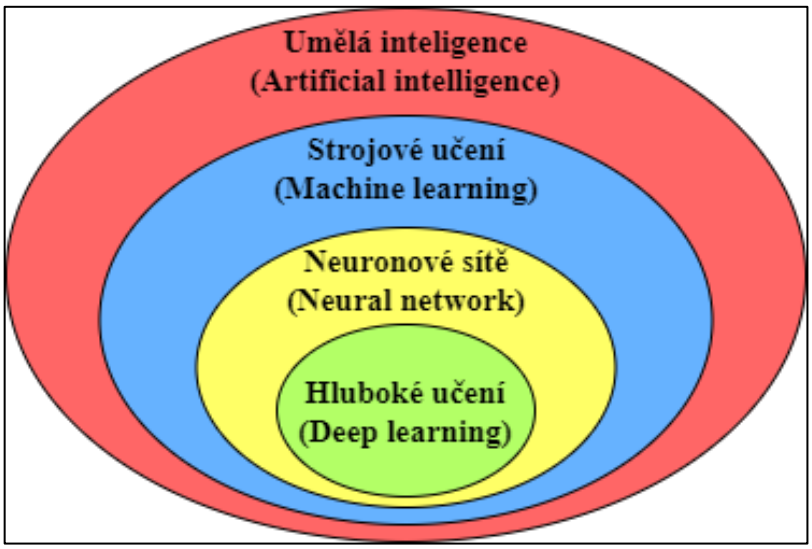
Rozdělení na základě funkce UI:

- **Reaktivní stroje UI** – Ke své funkci tyto systémy nevyužívají žádnou paměť, ze které by byly čerpány předchozí rozhodnutí nebo výsledky. Pracuje se zde s aktuálními daty, které jsou prostřednictvím statistické matematiky přímo vyhodnoceny do více či méně kvalitního výstupu. Tyto systémy jsou navrhované pro řešení specifických úkolů. Příkladem mohou být doporučovací systémy na webech jako Youtube nebo Netflix. [36]
- **UI s omezenou pamětí** – UI pracuje s omezenou pamětí, ve které dočasně ukládá předchozí výsledky svých rozhodnutí. Minulá data se současnými jsou použita jako základ pro dosažení uspokojivých výsledků v určité oblasti. Stále ale není dosaženo kvalit lidské inteligence, příkladem může být učení cizích jazyků. Lidský jedinec se po naučení druhého jazyka zvládne naučit další v mnohem kratším čase. U UI s omezenou pamětí toto neplatí. Pro každý úkol je třeba UI od začátku učit na velkém množství dat jednotlivé jazyky. S tímto typem UI se lze setkat u aktuálně velmi rozšířených generativních modelů, autonomních aut a taky chatbotů. [36][37]

- **Teorie mysli –** Teoretická funkce UI, která poskytuje možnost porozumět pocitům jiných entit. UI by mohla vytvářet vztahy podobné těm lidským, jelikož by zvládla vhodně reagovat a porozumět emočnímu stavu jedince. Porozumění by se vztahovalo i na umělecká díla, texty, hudbu a další obsah, který by mohla UI s theory of mind chápat. Aktuální UI jen generuje obsah, ale nemá vůbec ponětí o tom, čím je ten obsah unikátní. [36][37]
- **Sebe-uvědomělá UI** – Sebe-uvědomělá UI si je plně vědoma svých emocí, tužeb, přesvědčení a potřeb, ale zároveň chápe i emoce druhých. Taková UI by se plně vyrovnala lidskému jedinci, ale v současnosti ji nelze kvůli omezení hardwaru a neexistujících algoritmů vytvořit. [36]

2.2 Strojové učení, hluboké učení a neuronové sítě

V této podkapitole jsou probrány tři podobory UI, a to strojové učení (SU), hluboké učení (HU) a umělé neuronové sítě (dále jen UNS). Tyto pojmy jsou často mezi sebou zaměňovány, ale každá z těchto disciplín má svá vlastní specifika, která ji odlišují. Názornou ukázkou je následující Obrázek 6, který popisuje vztahy mezi jednotlivými disciplínami. [38]

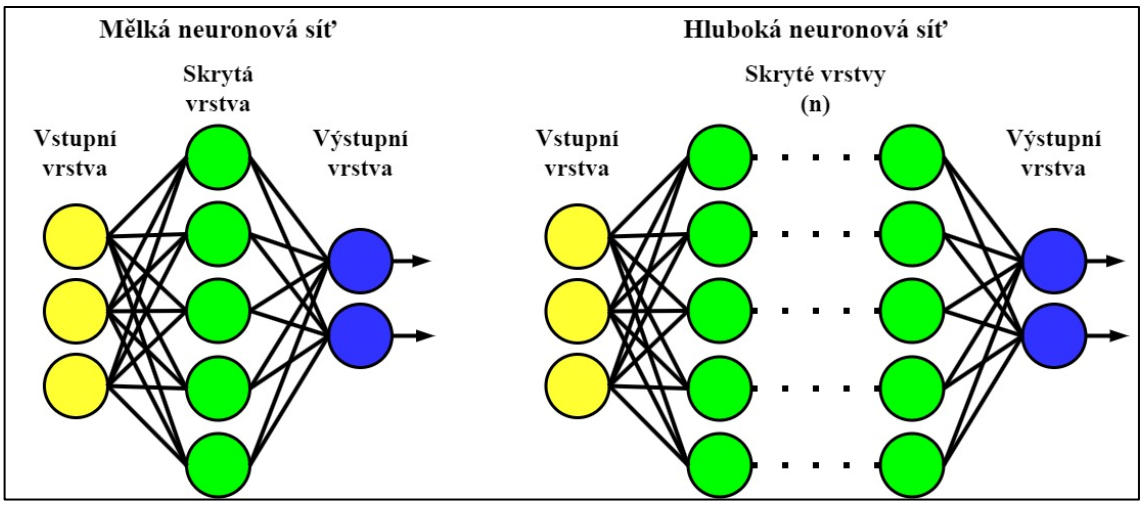


Obrázek 6: Souvislosti mezi obory. [38]

SU je podoborem UI, který umožnil strojům samostatně se učit z velkého objemu dat bez průběžného programování. Za pomoci statistických algoritmů se v datech hledají opakující se vzory, které jsou shlukovány a tříděny do odpovídajících výstupů. K těmto účelům byly vyvinuty algoritmy jako rozhodovací strom, k-nejbližší soused, neuronové sítě a další. Dále jsou popsány jednotlivé metody SU: [39]

- Učení s učitelem – je učení, které pracuje na označených data setech. Do modelu jsou vložena vstupní data a průběžnou úpravou pravděpodobnosti jednotlivých složek data setu, se model snaží dosáhnout předem stanovených vzorových dat. [39]
- Učení bez učitele – je učení, které pracuje na neoznačených data setech. Modelu jsou představeny náhodná data, ve kterých jsou hledány podobnosti a opakující se vzorce. [39]
- Kombinované učení – je učení, které pracuje s menším označeným a větším neoznačeným data setem. Označený data set je vodítkem, které je nápomocné při hledání a třídění dat v neoznačeném data setu. [39]
- Zpětnovazebné učení – je podobné jako učení s učitelem, ale místo vzorových dat se model učí ze špatných a dobrých rozhodnutí. Model je poté odměňován za správná rozhodnutí, aby byl zapamatován postup k získání výsledku. [39]

UNS jsou jedním z algoritmů SU, které hledají paterny v označených, ale převážně ve velkém množství neoznačených dat. Název a celá myšlenka vychází z chování neuronů v lidském mozku, při kterém virtuálně vytvořený umělý neuron obdobně zpracovává mnoho vstupů, ale dále posílá pouze jeden výstup. Neuronové sítě jsou složeny ze vstupních, skrytých a výstupních vrstev umělých neuronů. Vstupní vrstva slouží pro nahrání dat a výstupní vrstva pro reprezentaci už upravených vstupních dat. Skrytá vrstva je mezivrstva určena pro veškeré výpočty a oproti předešlým vrstvám pracuje nezávisle na uživateli. [40][41]



Obrázek 7: Mělká a hluboká neuronová síť. [38]

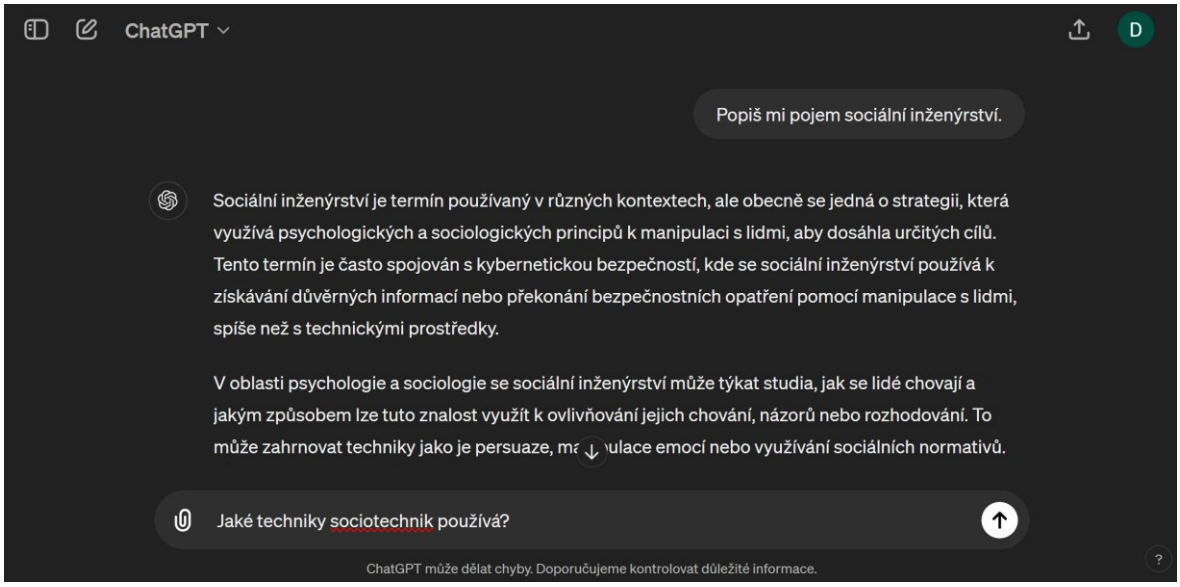
Neuronové sítě lze dělit podle počtu skrytých vrstev na mělké neuronové sítě a na hluboké neuronové sítě (Obrázek 7). Rovná-li se počet skrytých vrstev jedné, považuje se za mělkou UNS. Pokud je počet skrytých vrstev roven nebo je vyšší než dvě, jedná se o hluboké neuronové sítě. Slovo „hluboké“ odkazuje právě na **hluboké učení**. Tedy každá UNS, která splňuje podmínku počtu skrytých vrstev, je považována za algoritmus hlubokého učení. Algoritmus HU je podoborem SU a liší se tím, že používá ke své funkci mnohem větší sety dat, učí se ze svých chyb, je přesnější a není třeba velkého zásahu člověka. Dále jsou uvedeny některé struktury hlubokých neuronových sítí: [38][42]

- Vícevrstvý percepton
- Konvulenční neuronové sítě
- Rekurentní neuronová síť (Obousměrné RNS, Dlouhá krátkodobá paměť RNS atd.)
- Generativní adverzní síť
- Auto-Enkodéry (Kontraktivní Autoencoder, Variační Autoencoder atd.)
- Kombinace jednotlivých sítí [38]

2.3 Modely umělé inteligence

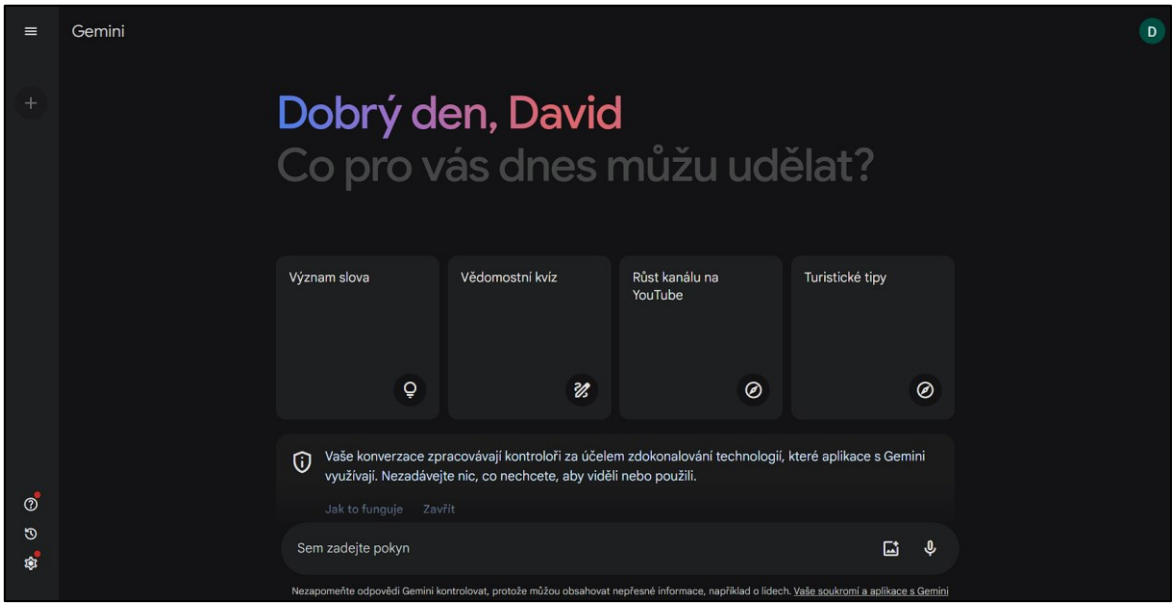
Tato kapitola je věnována modelům UI, které se staly každodenním nástrojem pro běžného člověka. Především jsou zde popsány textové (chatGPT a Gemini), obrázkové (Dall-E) a hlasové (ElevenLabs) modely. Nejpoužívanější modely byly vybrány podle návštěvnosti za měsíc únor 2024. Mezi sebou bylo porovnáno více nástrojů a ten s větší četností návštěv byl vybrán jako nejpoužívanější. Výsledný výběr byl složen z chatGPT (1,7 miliardy návštěv), Gemini (316 miliónů návštěv), Dall-E (1,7 miliardy návštěv), ElevenLabs (18,1 milionů návštěv) a VoiceMod (3,4 miliónů návštěv). Jelikož je Dall-E implementováno do ChatGPT byla brána pro modely stejná návštěvnost. [43]

Prvním popsaným modelem v oblasti UI je **ChatGPT** od společnosti OpenAI. Jedná se o chatbota, který zvládá z malého textového vstupu vygenerovat rozsáhlé texty na výstupu (Obrázek 8). Model je vycvičen na velkém objemu textových dat, které umožnily zpracování přirozeného jazyka, tedy jak jednat a reagovat v jakémkoliv naučeném jazyce. ChatGPT je zkratkou pro Generative Pre-Trained Transformer, což odkazuje na transformační neuronovou síť. Tyto sítě jsou schopny najít podobnosti mezi slovy, lépe pochopit význam textu a vybrat v něm důležité prvky. Nejnovější verze ChatGPT 4 zvládá nově, kromě práce s textem, zpracovat a analyzovat obrázkové vstupy, a analyzovat ručně psaný text. Zároveň je v něm implementován Dall-E 3 a je o cca 40 % přesnější než předchozí verze. [43][45]



Obrázek 8: Ukázka prostředí a generovaného textu nástroje ChatGPT. [45]

Gemini, dříve Bard, je textový model od společnosti Google (Obrázek 9). Tento model je multimodální, což znamená, že dokáže trénovat nejen na textových datech, ale jsou použita i velká množství obrázkových a zvukových dat. Tento styl učení umožnil odpovídat na dotazy i jiným způsobem než textovou formou. Model je založen na transformačních neuronových sítích a nové technice zvané cross-modal attention. Technika poskytuje lepší pochopení jemných nuancí mezi různými typy dat. [46][47]



Obrázek 9: Prostředí modelu Gemini. [48]

Dall-E 3 je model pro generování obrázkového obsahu, který byl vyvinut společností OpenAI. Model je schopen z textového vstupu vygenerovat jakýkoliv obrázek na výstupu.

Jedná se o další z modelů, který pro svou funkci používá transformační a auto-enkondéry neuronové sítě. Auto-enkondéry segmentují vstupní text pro vykreslení všech podrobností v požadavku. Jiné modely pracují pouze s průměrováním pravděpodobnosti, což má za následek vynechání některých požadavků (Obrázek 10). Další modely, které se specializují na generování obrázků, jsou MidJourney, Stable Diffusion nebo Dream. [49]

Ukázka komplikovaného požadavku na dva obrázkové modely: *Lev se zlatou korunou na hlavě, sedící na růžovém kameni. Okolo něj jsou různá zvířata, která se mu klaní a jsou oblečena v černém smokingu.*

Obrázek 10: Vlevo obrázek od Dall-E 3 a vpravo obrázek od Stable Diffusion.

ElevenLabs je platforma, která se specializuje na vytváření syntetických hlasů pro chatboty call centra, dabing nebo tvůrce zábavního obsahu na internetu. Model poskytuje obsáhlou knihovnu hlasů, kterými lze nahradit originální hlas, ale je možné vytvořit i syntetický hlas podle vlastní předlohy. Model se může naučit jakýkoliv hlas prostřednictvím funkce instantní klonování hlasu nebo profesionální klonování hlasu. První varianta vyžaduje cca 30 sekund požadovaného hlasu a profesionálnímu klonování je potřeba poskytnout nejméně 30 minut, v ideálním případě až 3 hodiny zvukového záznamu. Rozdíl mezi funkcemi je především v kvalitě výsledného výstupu. Model poskytuje převedení textu na hlas, ale nově je možné převádět hlas uživatele na jakýkoliv hlas dle výběru. Převedení tzv. “z hlasu na hlas“ slouží pro přesnější zachycení různých výšek hlasu. Oproti modelu “text na hlas“ se odstraní výrazné pauzy mezi slovy a hlasu je ponechána určitá realističnost a emocionální zabarvení hlasu. Dalšími modely pro tvorbu syntetických hlasů jsou třeba Speechify. [50]

Poslední nástroj se týkal taky audia a jednalo se o nástroj **VoiceMod**, který oproti předešlému nástroji umožňuje převádět hlas na syntetický hlas v reálném čase. Oproti předešlému modelu ElevenLabs odpadá potřeba vytváření nahrávek a je možné převádět hlas na jiný s minimálním zpožděním (cca 500 ms). Rychlost převodu je ale značně ovlivněna výpočetní technikou a pokud není dostatečně výkonná, hrozí dlouhé prodlevy a sekání hlasu. Nástroj poskytuje knihovnu hlasů, funkci klonování hlasu a možnost si vytvořit unikátní hlas prostřednictví manuální syntézy. Především poslední funkce je zajímavá, protože poskytuje možnost hlasu přidat větší hloubku nebo jej přetvořit do úplně jiné podoby. [51]

2.4 Shrnutí

V poslední části byly vymezeny základní pojmy UI a čím jsou charakteristické. Na konec byly popsány hojně používané modely a jejich základní funkcionality. Další kapitola je již počátkem praktické části a představí metodiku výzkumu a návrh sociálního experimentu.

II. PRAKTICKÁ ČÁST

3 METODIKA A NÁVRH SOCIÁLNÍHO EXPERIMENTU

Výzkum celé praktické části je zaměřen na využití syntetického hlasu rodinného příslušníka, přátel nebo vedoucího při telefonních podvodech. Tato nově vzniklá hrozba představuje velké nebezpečí, protože v oběti vzbuzuje okamžitou důvěru, aniž by útočník musel budovat dlouhodobý vztah s obětí. První kapitola praktické části je rozčleněna na metodiku a návrh sociálního experimentu.

3.1 Metodika

V metodice praktické části jsou zmíněny veškeré postupy a metody využité ve výzkumu. Je zde předložen hlavní cíl práce a dílčí cíle, které byly v rámci práce zkoumány. V rámci metodiky byl uveden výzkumný soubor, výzkumné metody a stanoveny hypotézy.

3.1.1 Cíl praktické části

Cílem práce bylo analyzovat nově vzniklou hrozbu v podobě podvodných telefonátů se syntetickým hlasem a na základě zjištění z analýzy vytvořit protiopatření. Dílčí cíle byly zvoleny dva:

a) Dozvědět se prostřednictvím ankety, jak je aktuálně hrozba vnímána ve společnosti.
b) Zjistit za pomoci sociálního experimentu, zda je hrozba v současnosti použitelná v České republice.

3.1.2 Výzkumný soubor

Zkoumaný soubor byl tvořen dvěma skupinami respondentů. První skupina byla získána samovýběrem prostřednictvím sdílené ankety, která cílila na širokou veřejnost. Jelikož anketa není zaměřena na určitou skupinu lidí, bylo možné sdílet anketu v digitálním prostoru a získat odpovědi od široké škály respondentů. Anketa byla umístěna na sociální sítě, streamovací platformy atd., kde byla očekávána velká rozmanitost lidí. Po ukončení sběru dat bylo získáno celkem 203 respondentů. Následnou filtrací nevhodných odpovědí byla první skupina nakonec tvořena 180 respondenty, z jejichž odpověďmi se dále pracovalo.

V druhé skupině se nacházeli přátelé a rodinní příslušníci, na kterých bylo možné vyzkoušet sociální experiment se syntetickým hlasem. Tito respondenti byli vybráni na základě záměrného výběru. Celkem bylo vybráno 6 respondentů, ale ve výsledku se experimentu zúčastnilo pouze 5 respondentů. Na respondentech byl proveden sociální experiment a následně uskutečněn polostrukturovaný rozhovor.

3.1.3 Výzkumné metody

Výzkum v praktické části byl založen především na kvantitativní metodě, ale pro podporu výsledků byl ještě vytvořen sociální experiment, ve kterém se pracovalo s kvalitativními metodami. Sociální experiment byl podpůrnou metodou, ve které se zjišťovalo, zda je tato hrozba využitelná i na území České republiky a podrobněji je rozepsána v kapitole „Návrh sociálního experimentu“. Dále je zmiňována pouze anketa a polostrukturovaný rozhovor.

Anketa byla zvolena jako zástupce kvantitativní metody. Oproti dotazníkovému šetření je tato metoda aplikovatelná na širokou škálu respondentů a není nutné cílit na specifická uskupení. To umožnilo cílit na mladší i starší generaci a dozvědět se informace a pohledy lidí na danou problematiku z velkého vzorku respondentů. Anketa je navíc tvořena menším počtem jednoduchých otázek, které bývají většinou uzavřené, ale podle potřeby byly použity i otevřené otázky. Proto byla vhodnější volbou pro získání potřebných dat k tomuto tématu. [52]

Otázky byly respondentům v rámci ankety předkládány v podobě online dotazníku. Dotazník byl sdílen na streamovacích platformách, sociálních sítích a e-mailovou komunikací ve firmách, které byly ochotné sdílet dotazník mezi starší generaci.

Dotazník byl konstruován tak, aby podal náhled na danou problematiku a zda jsou respondenti s touto hrozbou obeznámeni či nikoliv. Celkově byla anketa složena z dvanácti otázek, které se týkaly problematiky vishingu pohaněného syntetickým hlasem. První tři otázky jsou tvořeny identifikačními položkami, které byly zaměřené na pohlaví, věk a vzdělání respondenta. Zbytek otázek se již týkal problematiky a až na otevřené otázky č.10 a č.12 byly všechny otázky uzavřené. Dále byla snaha zjistit, zda je syntetický hlas generovaný v češtině rozpoznatelný lidským uchem. Tato informace byla ověřena za pomoci jednoduchého testu v otázce č.6. Na základě dotazníku byly stanoveny hypotézy, které byly vyhodnoceny za pomoci testu nezávislosti chí-kvadrát, který pomáhá hledat významné vazby mezi daty. Celé znění dotazníku je k nahlédnutí v **PŘÍLOZE P I**. [52][53]

Stanovené hypotézy

- **H1:** Předpokládám, že respondenti nad 35 let jsou méně informováni o existenci podvodných telefonátů se syntetickým hlasem než mladší populace.
- **H2:** Předpokládám, že respondenti s maturitou lépe rozpoznají syntetický hlas od reálného.
- **H3:** Předpokládám, že existuje významný rozdíl ve znalostech o opatřeních proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví.

Další zvolenou metodou pro výzkum byl polostrukturovaný rozhovor. Cílem rozhovoru bylo podpořit sociální experiment získáním výpovědí jednotlivých obětí a zjištěním informací, které jsem jako útočník nemusel zaznamenat. V následujících odrážkách jsou uvedeny otázky do rozhovoru a v případě nejasností byly položeny doplňující dotazy.[54]

- Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?
- Co ti na rozhovoru přišlo podezřelé?
- Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?
- Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodů v blízké budoucnosti?

Otázky byly v předstihu poslány respondentům i s informacemi, jak bude rozhovor probíhat. Rozhovory byly provedeny osobně v místech, kde to respondentům vyhovovalo nebo byla navázána komunikace online. Před započetím rozhovoru byly upřesněny veškeré dotazy a respondenti byli ujištěni, že jejich podobizna, hlas a další údaje budou anonymizovány.

Pro přepis rozhovoru byl využit selektivní protokol. Tento protokol umožňuje vynechat nepotřebné informace a taky chování respondenta, které nemá v tomto výzkumu žádnou výpovědní hodnotu. Není tedy potřeba doslovného přepisu a je možné zapsat pouze důležité informace. Veškeré rozhovory byly převedeny do textové podoby a jsou uvedeny v **PŘÍLOZE P II, III, IV, V** a **VI.** Všechen text byl podroben metodě kontrastů a srovnání, při které se v textu hledaly podobnosti s ostatními respondenty a zajímavé informace, které mi mohly uniknout. [54]

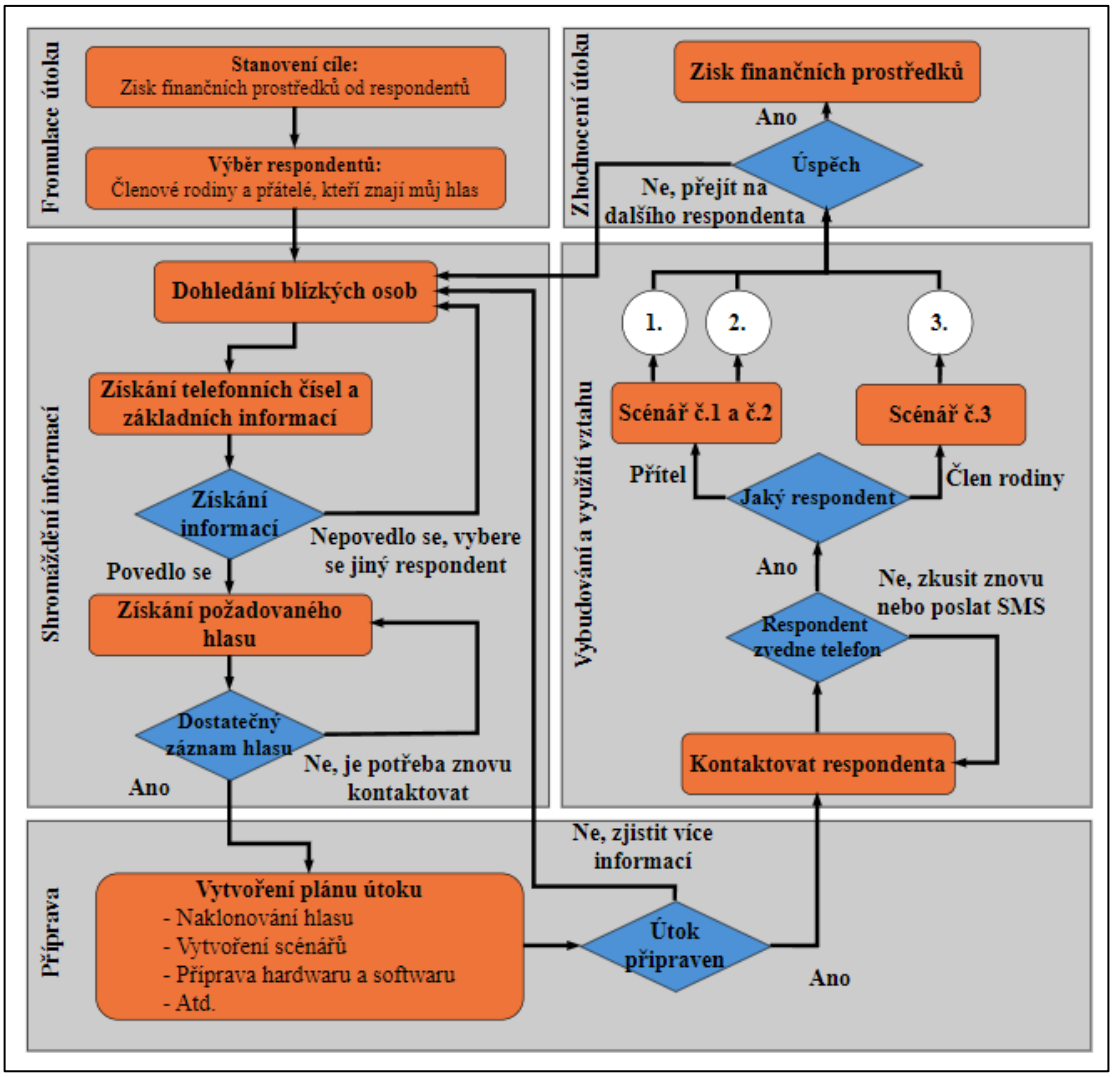
3.2 Návrh sociálního experimentu

Sociální experiment je zpracován pro ověření, zda vishingové útoky, které zneužívají syntetický hlas, jsou v aktuální formě nebezpečné pro občany ČR. Dalším důvodem je zisk informací z průběhu experimentu, které by byly využitelné pro tvorbu protiopatření. Proto je celý experiment koncipován jako útok SI nebo spíše podvod, aby byly získány poznatky přímo z průběhu útoku.

Vyhodnocení celého experimentu vychází z úspěšného/neúspěšného provedení sociálního experimentu na respondentech. Následně je proveden rozhovor, aby byl zjištěn i úhel pohledu respondentů na danou problematiku.

3.2.1 Model útoku

Celý experiment je založen na technice zvané vishing, která využívá syntetický hlas blízké osoby k podvodným telefonátům. Používaný hlas pro útok je můj vlastní a je s ním cíleno na rodinné příslušníky a přátelé, kteří můj hlas dobře znají. Postup útoku si bere inspiraci z modelu pana Moutona, který byl pro potřeby útoku upraven (Obrázek 11). [54]



Obrázek 11: Blokové schéma postupu v sociálním experimentu.

Nejprve byl **formulován útok,** stanovil se cíl útoku a výběr obětí. Ve fázi **určení cíle** byl specifikován cíl útoku jako „Získání finančních prostředků od respondentů“. Bylo usouzeno, že takového cíle lze dosáhnout pravděpodobněji, než kdyby se podvod zaměřoval na citlivé údaje jako jsou hesla nebo platební údaje, které by mohly vyvolat větší podezření. Respondenti byli vybráni z okruhu rodiny a blízkých přátel, jejichž vazby byly poměrně jednoduše dohledatelné prostřednictvím sociálních sítí.

I když to nebylo v rámci práce vyžadováno, byl vyzkoušen zjednodušeně i krok **shromáždění informací**. Tento krok byl zjednodušen kvůli časové náročnosti, která se pojí s vytěžováním informačních zdrojů a dohledáváním informací o dané osobě. Pokud byly dohledány blízké osoby za pomoci veřejných příspěvků, ve kterých byla má osoba označena, bylo následně nutné dohledat jejich telefonní čísla. Tento krok byl složitý, a proto byl zjednodušen a čísla se získala převážně z uložených kontaktů. I tak bylo získáno pár čísel v rámci profilů obětí na sociálních sítích a jedno číslo bylo získáno za pomoci SMS z cizího čísla. Získání čísla skrze SMS proběhlo tak, že nejprve bylo zjištěno ze sociálních sítí telefonní číslo blízké osoby. Tato osoba nebyla využita pro experiment, ale posloužila k získání čísla jednoho z potřebných respondentů právě skrze požadavek o telefonní číslo v SMS zprávě (Obrázek 12).

Obrázek 12: Falešná SMS pro získaní čísla.

Kromě zmíněných informací nebylo nutné shromažďovat další velké objemy dat o respondentech. Scénáře pro experiment byly totiž navrženy tak, aby byl řešen přímo vymyšlený problém a nedocházelo ke zbytečnému povídání.

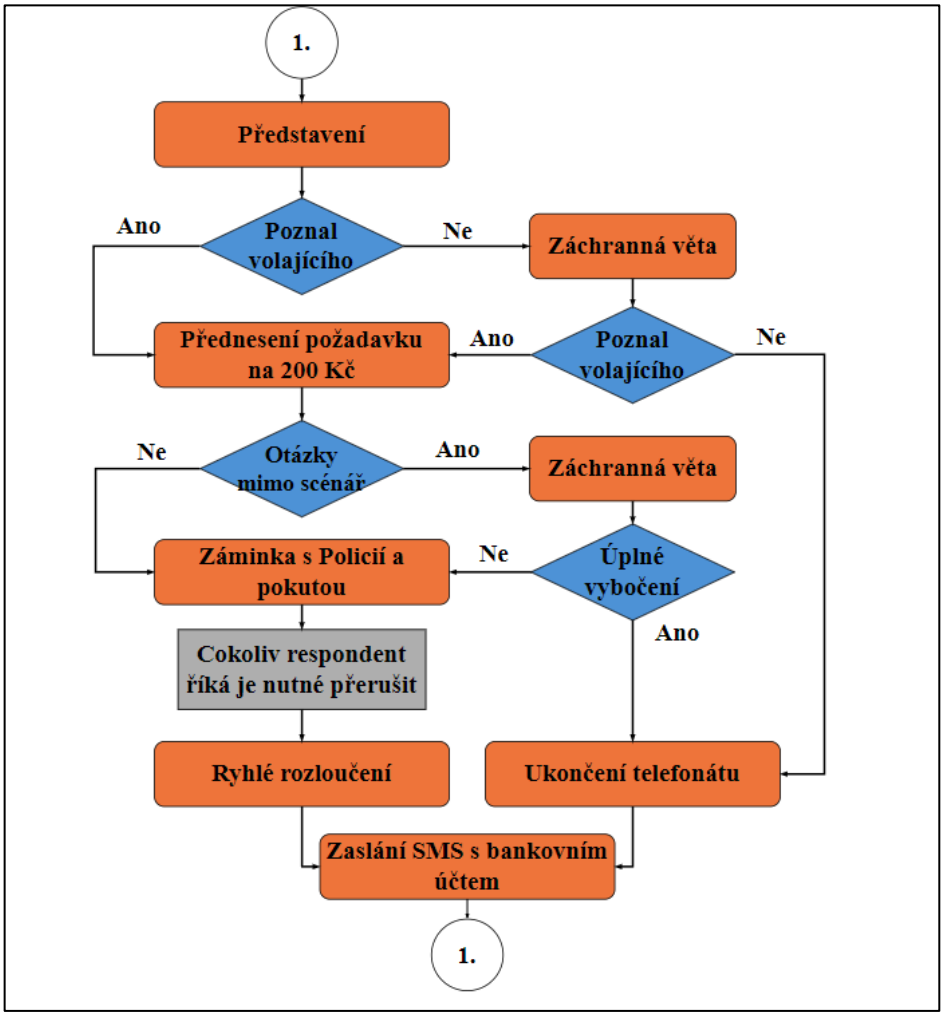
Důležitým prvkem celého podvodu bylo získání nahrávky požadované osoby. Jelikož byl podvod založen na mém hlasu, byla jednoduše poskytnuta 30s nahrávka pro okamžité naklonování hlasu. Sociotechnik ale může získat hlas mnoha způsoby, třeba skrze telefonát, videa na sociálních sítích nebo nahráním rozhovoru během osobního setkání. Ze získaných nahrávek je možné extrahovat hlas oběti v potřebné délce a naklonovat jej.

Ve fázi **příprava** byl stanoven postup útoku a všechny náležitosti, které by mohly pomoci s úspěšným experimentem. Pro podvod byla zvolena technika v podobě vishingu se syntetickým hlasem. Dále na základě získaných informací byly vytvořeny tři scénáře, které sloužily jako návod pro provedení podvodu. Ve scénářích byly zakomponovány principy ovlivnění jako apel na autoritu a jelikož můj hlas je respondentům dobře znám, byl využit i princip

oblíbenosti. Z jiných manipulačních technik byl využit třeba časový nátlak. Scénáře byly namluveny syntetickým hlasem vytvořeným aplikací ElevenLabs, která poskytovala nahrávky naklonovaného hlasu. Více o nástroji ElevenLabs v kapitole „Nástroje pro sociální experiment“. Tyto nahrávky byly následně pouštěny při telefonátu s respondentem.

Scénář č.1:

Ve scénáři je vymyšlena záminka, ve které se svým syntetickým hlasem zavolám přátelům s požadavkem na půjčení finančního obnosu (Obrázek 13). Požadavek se týká velmi nízké částky 200 Kč, abych mohl zaplatit pokutu za rychlost, kterou jsem zrovna dostal od Policie ČR. V telefonátu apeluji na nedostatek času, tíživou situaci a taky na autoritu, která po mě pokutu vyžaduje okamžitě zaplatit, jinak přijdu i o body na řidičském průkazu. Věty, které se musely připravit dopředu jsou uvedeny níže. Celý scénář je ještě podpořen zvuky jedoucích aut v pozadí, aby v obětech evokovaly, že jsem opravdu někde u cesty.



Obrázek 13: Blokové schéma pro první scénář.

Přichystaný postup pro první scénář:

1. **Respondent:** (Předpokládá se představení)
   1.1. Autor: „*Čau, to jsem já Dejv, volám ti z druhého čísla“*
2. **Respondent: (**Po představení je očekáváno, že se respondent zeptá „*Co potřebuji*“ nebo „*Jak se mám*“. Proto byli nachystány dvě odpovědi)
   2.1. Autor: „*Prosím tě potřeboval bych nutně půjčit 200 korun.“*
   2.2. Autor: „*Jo dobrý, jen bych potřeboval nutně půjčit 200 korun.“*
3. **Respondent:** „*K čemu to potřebuješ?“* (Víceméně cokoliv, co respondent odpoví, povede k přednesení záminky, popřípadě k záchranným větám.)
   3.1. Autor: „*Hele chytli mě policisti za rychlost… no a buď můžu kartou zaplatit pokutu teď nebo na stanici, ale to by mi sebrali i body. Takže bych to chtěl vyřešit rovnou na místě, ale chybí mi těch 200 korun. Jakmile mi dojde výplata, tak ti zaplatím celou částku i s poplatky za okamžité odeslání, ano?*“
4. **Respondent**: (Zde mohlo dojít buď k potvrzení požadavku nebo bylo očekáváno, že respondent bude zmatený a bude se doptávat. Proto byly vytvořeny odpovědi s poděkováním a rychlým ukončením rozhovoru.)
   4.1. Autor: *„To je parádní. Děkuji moc. Hele ještě tady potřebuji něco vyřešit s policisty a už se budu muset rozloučit. V SMS ti pošlu číslo účtu, a ještě jednou děkuji moc a někdy se vidíme. Čau.“*
   4.2. Autor: „*Pane Vítku, můžete na chviličku*.“ (Hlas falešného policisty) „*Ano strážníku? Hele promiň, už budu muset letět. Pošlu ti číslo účtu v esemesce a budu ti moc vděčný, když mi to co nejrychleji pošleš. Zatím se měj a děkuji.“*

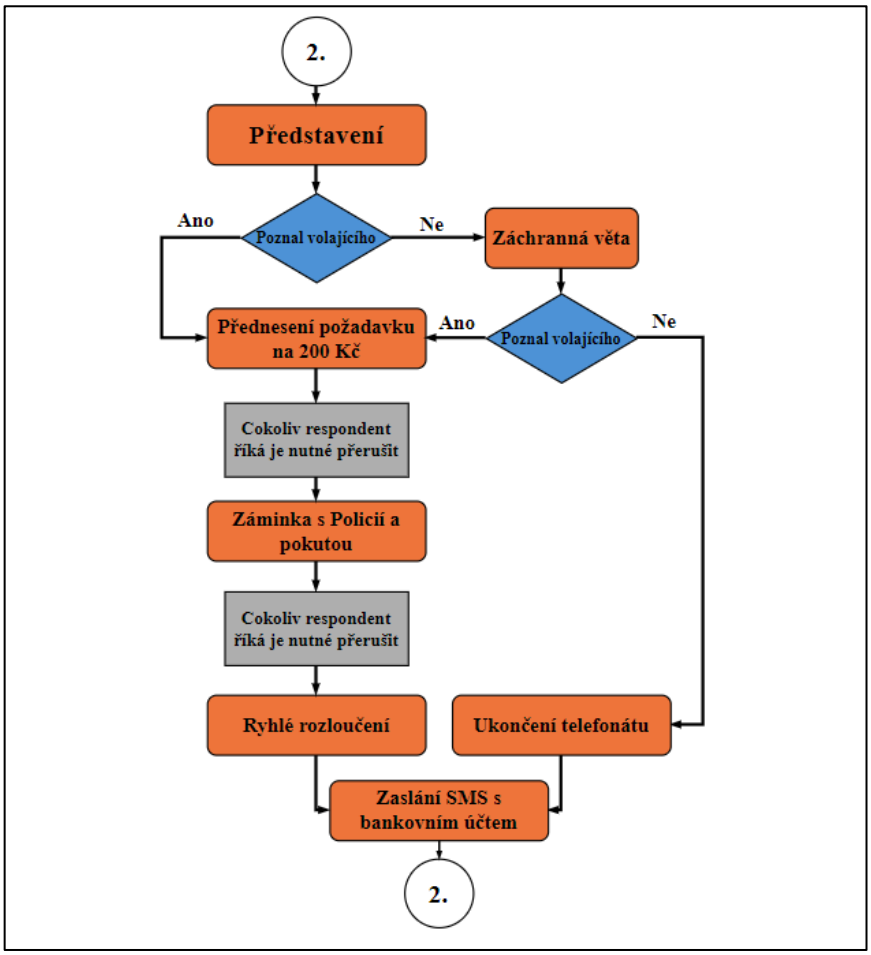
Záchranné věty: *„Ano“, „Ne“, „V SMS ti pošlu bankovní účet.“, „David Vítek“*

Po ukončení telefonátu byla oběti zaslána SMS, ve které byly uvedeny informace k bankovnímu účtu a pro jistotu i částka, která byla po oběti požadována. Zpráva SMS by byla poslána i v případě, kdyby telefonát nedopadl dobře. Ve zprávě by bylo možné ještě dokončit scénář a dovysvětlit respondentovi nejasnosti. Stejně tomu bylo i u všech ostatních scénářů.

Scénář č.2:

Tento scénář je víceméně stejný jako první scénář, ale byl koncipován do podoby jednodušších odpovědí s ponecháním základních prvků ze scénáře č.1 (Obrázek 14). Odpovědi byly předělány do podoby, která zajistí návrat k požadavku o peníze, když respondent vybočí

ze scénáře. Taky zde byla snaha o snížení využití záchranných vět. Změna scénáře byla nutná, protože se třetí respondent často snažil přerušit rozhovor a doptávat se na detaily.



Obrázek 14: Blokové schéma ke scénáři č.2.

Předpokládaný průběh telefonátu:

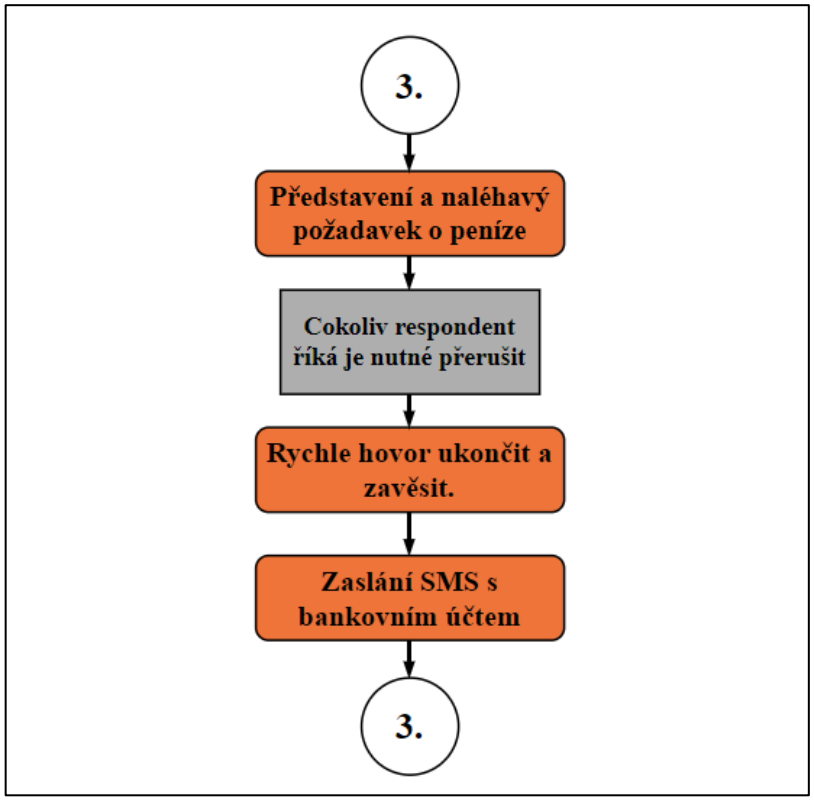
1. **Respondent:** (Předpokládá se představení)
   1.1. Autor: *„Ahoj, tady Dejv, volám ti z druhého čísla.“*
2. **Respondent:** (Víceméně cokoliv, co respondent odpoví, povede ke stejné odpovědi útočníka, popřípadě k záchranným větám.)
   2.1. Autor: *„Hele, potřeboval bych co nejrychleji půjčit nebo spíše poslat na účet 200 korun.“*
3. **Respondent:** (Víceméně cokoliv, co respondent odpoví, povede ke stejné odpovědi útočníka, popřípadě k záchranným větám.)
   3.1. Já: *„Hele nemám moc času ti to vysvětlovat dopodrobna. Hrajou tam prostě roli policajti a pokuta. Pak bych ti to vysvětlil, jak se uvidíme nebo jak ti to budu vracet, tak ti to… Tak ti řeknu, co se stalo.“*

4. **Respondent:** (Víceméně cokoliv, co respondent odpoví, povede ke stejné odpovědi útočníka)
   - **4.1. Autor:** *„Už budu muset jít, potřebuji tady ještě něco vyřešit s těma policajtama. Číslo účtu ti pošlu v té smsce teda… a budeš moc hodná, když mi to pošleš. Díky moc, zatím čau.“*

Záchranné věty: *„Ano“, „Ne“, „V SMS ti pošlu bankovní účet.“, „David Vítek“*

Scénář č.3:

V tomto případě byl vyzkoušen scénář, který okamžitě přechází k prosbě o peníze a vůbec se nezaměřuje na vysvětlení situace (Obrázek 15). Byla zde snaha vytvořit co nejnaléhavější situaci, aby respondent neměl moc času přemýšlet nad tím, co se děje a mít dotazy mimo scénář. Proto zde nebylo ani počítáno se záchrannými větami.



Obrázek 15: Blokové schéma ke scénáři č.3.

Očekávaný průběh telefonátu:

1. **Respondent:** (Předpokládá se představení)
   - **1.1. Autor:** *„Čau, rodinný příslušníku, tady David, hele potřebuji nutně poslat 200 korun. Neptej se proč, pak ti to všechno vysvětlím. Nemám teďka moc čas. Pošlu ti číslo účtu v SMS a prosím co nejrychleji je tam pošli.“*

2. **Respondent:** (Cokoliv zde respondent začne řešit, je přerušeno následujícím textem)
   2.1. **Autor:** „*Hele fakt nemám čas ti to vysvětlovat. Prostě potřebuji to nutně poslat. Pošlu ti číslo účtu v SMS. Už musím končit, fakt promiň. Prostě mi to pošli, jo? Díky moc. Zatím čau.“*

Na závěr je nutné dodat, že předem připravené scénáře nejsou flexibilní a nelze je upravovat během experimentu. Jakékoliv větší vybočení mimo věty a záchranné fráze může mít za následek neúspěšný útok.

Ve fázi **vybudování a využití vztahu** bylo očekáváno projevení síly UI v plném rozsahu. Naklonováním mého hlasu je sociotechnikovi umožněno zneužít již vybudované vazby mezi přáteli nebo rodinou. Proto vytváření důvěry je možné z postupu úplně vynechat a pouze využít již vybudovaný vztah. Vyhodnocení úspěšnosti je uvedeno v kapitole „Vyhodnocení sociálního experimentu“

3.2.2 Nástroje pro sociální experiment

V této podkapitole jsou rozepsány a zdůvodněny veškeré hardwarové a softwarové nástroje, které byly během experimentu využity. Především byla pozornost věnována softwarovým nástrojům, které byly pro experiment zásadnější.

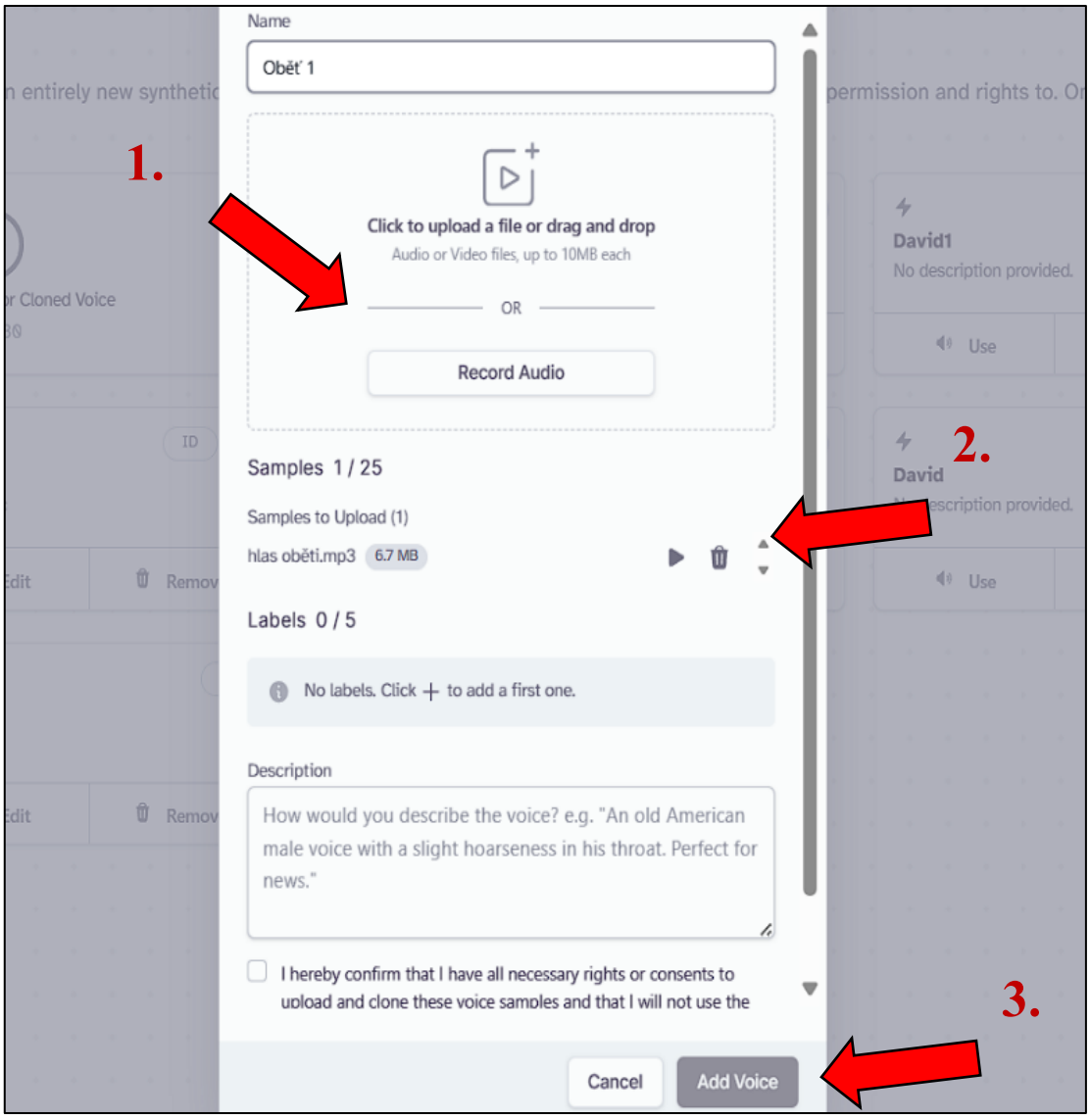
Nejprve bylo v plánu využít **softwarové nástroje** UI, které by byly schopny převádět skutečný hlas na syntetický v reálném čase. Bylo vyzkoušeno několik aplikací, které přinesly zjištění, že tyto modely mají velké výpočetní nároky na techniku. Převáděný hlas byl často zasekaný a taky zde byla velká prodleva v syntéze hlasu na syntetickou podobu. Modely byly vyzkoušeny i na stolním počítači se silnější grafikou (Gigabyte GTX 1650), ale problémy stále přetrvávaly. Další počítač, který by byl ještě výkonnější, nebylo možné v rámci mých kontaktů sehnat. Proto se přistoupilo k jiné variantě a zvolil se model skutečný hlas na syntetický hlas, který nepřevádí hlas v reálném čase, ale vytváří pouze nahrávky se syntetizovaným hlasem.

Srdcem celého experimentu byl softwarový nástroj ElevenLabs, se kterým lze vytvořit syntetický hlas jakéhokoliv člověka. Tato webová aplikace byla zvolena, protože nabízí generování hlasu v češtině, modely text/hlas na syntetický hlas a služby jako instantní a profesionální klonování hlasu. Podrobnější popis funkcí byl zmíněn v teoretické části. V experimentu byla zvolena funkce instantní klonování, jelikož byla pro tento typ podvodu mnohem vhodnější. Funkce zvládne z relativně krátkých nahrávek vytvořit přesvědčivou kopii hlasu,

což je pro podvody, při kterých sociotechnik nemá dostatek záznamů, mnohem vhodnější způsob klonování. Samozřejmě platí pravidlo, že čím více nahrávek je aplikaci poskytnuto, tím přesněji je hlas naklonován. [50]

Webová aplikace ElevenLabs je velice intuitivní. Po registraci a zaplacení vybraného plánu je zobrazeno jednoduché prostředí (Obrázek 16). Naklonování hlasu probíhá v záložce *Voices,* kde je možné přidat nově naklonovaný hlas tím, že zvolíme instantní klonování a následně umístíme nahrávky člověka, kterého chceme napodobit. [50]

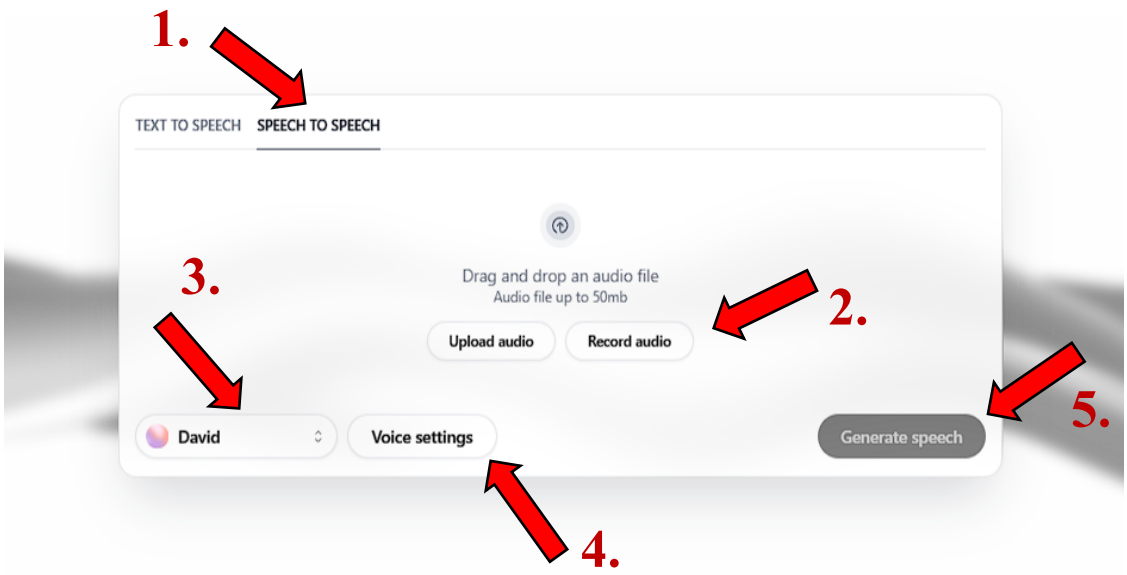
1. Pole pro vložení nahrávek klonovaného hlasu člověka.
2. Zobrazení nahrávek a jejich počet.
3. Tlačítko pro vygenerování klonovaného hlasu.



Obrázek 16: Prostředí pro instantní klonování hlasu. [50]

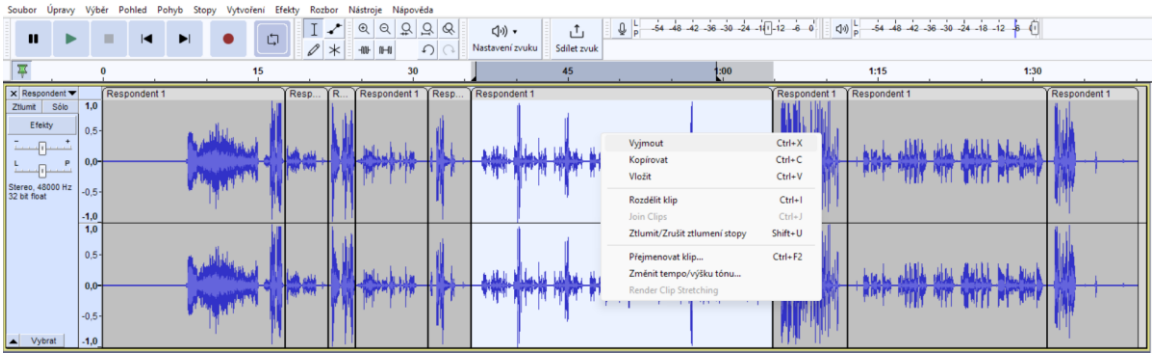
Nově naklonovaný hlas je možné okamžitě využívat v záložce *Speech*. V záložce je zobrazeno jednoduché prostředí, ve kterém je možné pracovat s hlasem a upravit jej na syntetický hlas vybraného člověka (Obrázek 17). [50]

1. Možnost výběru modelu text na syntetický hlas nebo hlas na syntetický hlas.
2. Pole pro nahrání potřebného audia.
3. V záložce si lze zvolit jaký hlas chceme napodobovat.
4. Možnost úpravy generovaného hlasu především jeho stability, rychlosti a dalších funkcí.
5. Tlačítko pro vygenerování hlasu.



Obrázek 17: Prostředí pro převedení hlasu na syntetický hlas. [50]

Dalším velice důležitým nástrojem z řad softwaru je program Audacity. Software je uzpůsoben k nahrávání a k základní úpravě zvukové stopy. Program byl v experimentu použit pro nahrání kvalitního audia a taky pro odstranění nežádoucích zvuků v nahrávce jako třeba šum nebo praskání. Upravené nahrávky jsou následně poskytnuty jako základ pro klonování mého hlasu. V audacity je možné nahrávat zvuk z mikrofonu, ale taky přímo ze systému. To znamená, že je možné telefonát uskutečnit skrze počítač a nahrát mnohem kvalitnější zvuk své oběti, než kdyby se telefonát nahrával přes mikrofon. Poté stačí vystříhat nežádoucí hlas pro získání požadovaného hlasu (Obrázek 18).



Obrázek 18: Ukázka práce v Audacity a proces odstranění nežádoucího hlasu.

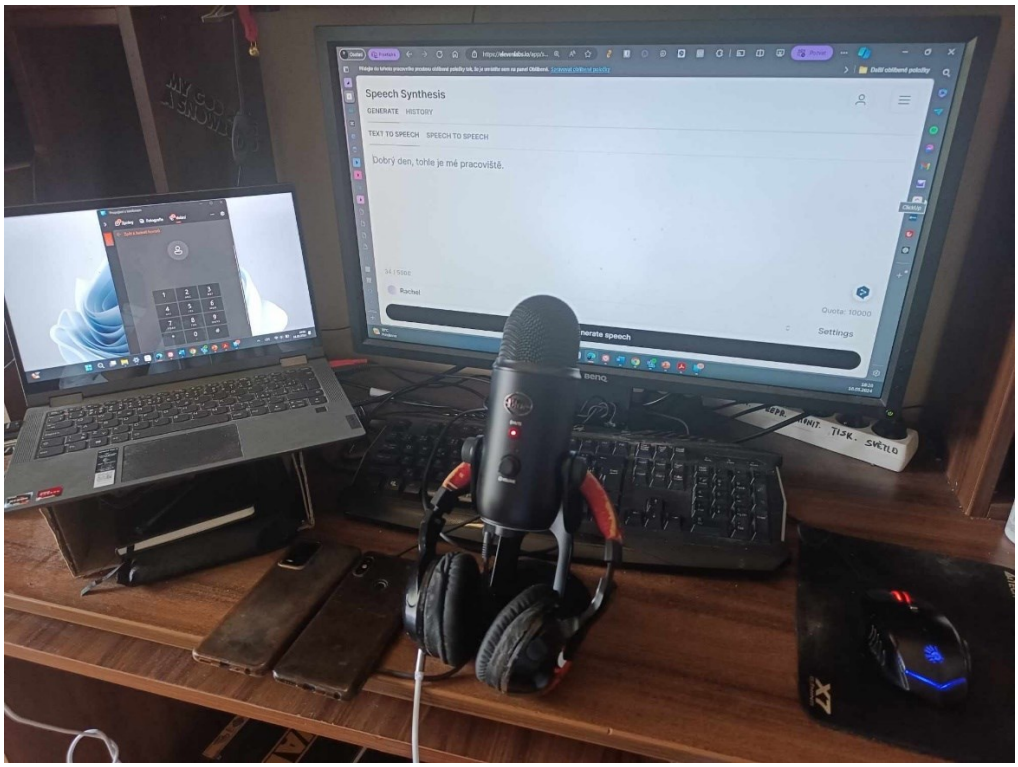
V plánu bylo využití i spoofing aplikace, ale vzhledem k jejich nefunkčnosti a ztrátě finančních prostředků u některých vyzkoušených poskytovatelů, bylo od tohoto záměru upuštěno. Pro experiment bylo využito druhé telefonní číslo, které respondenti nemohli znát.

Hardwarové nástroje hrají při provedení experimentu spíše vedlejší roli a hlavní místo zaujímají softwarové nástroje. Zvolený model od ElevenLabs totiž nemá velký vliv na výkon počítače, takže postačí jakýkoliv počítač s připojením k internetu. Úspěšnost experimentu více závisí na použitém mikrofonu než na zvoleném počítači. Pokud je zvolen model, při kterém se převádí originální hlas na umělý hlas, je mikrofon důležitý k přesnému zaznamenání hlasu. Bez kvalitního mikrofonu může záznam obsahovat šum, praskání a jiné chyby, které brání ve vytvoření uvěřitelného syntetického hlasu. Poslední položkou byly dva telefony, jeden se SIM kartou s novým telefonním číslem a druhý telefon pro přehrávání záznamu se syntetickým hlasem. Pro experiment byly použity tyto hardwarové nástroje a celá sestava je vidět na obrázku 19:

- Notebook – Lenovo IdeaPad Flex 5
- Mikrofon - Logitech G Blue Yeti USB, Blackout
- 2x Mobilní telefon

Základní myšlenkou bylo propojit první telefon s novým číslem a s počítačem za pomoci Bluetooth. Následně byly z druhého telefonu pouštěny nahrávky do externího mikrofonu napojeného na počítač. V rámci konverzace s první respondentkou a dalšího testování bylo zjištěno, že zde byl problém a hlas se při hovoru sekal. Proto byl pro další pokusy zvolen postup, při kterém byly nahrávky pouštěny z jednoho telefonu do druhého. Existuje ještě jedna možnost v podobě zakoupení příslušenství pro VoIP telefonování jako náhlavní souprava, specializovaný software pro VoIP a VoIP tarif. Celý hovor by pak mohl být proveden skrze počítač. Tato možnost by byla finančně náročnější, pokud by byly započítány i

dosavadní investice v podobě mikrofonu a služeb ElevenLabs. Navíc i samotná implementace by zabrala hodně času.



Obrázek 19: Pracovní prostředí pro experiment.

3.3 Shrnutí

První kapitola praktické části byla věnována metodice výzkumu a návrhu sociálního experimentu. Účelem této kapitoly bylo představení metod a postupu při výzkumu, aby bylo možné výzkum v budoucnu zopakovat a popřípadě vylepšit. Hlavní úkolem metodické části bylo stanovení hlavního cíle a dílčích cílů. Základní metodou výzkumu byla anketa, která měla za cíl zjistit, jak je hrozba vnímána veřejností. Spíše podpůrnou metodou byl polostrukturovaný rozhovor s respondenty, kteří byli podrobeni sociálnímu experimentu. V návrhu sociálního experimentu byly vytvořeny tři scénáře, podle kterých byl veden celý experiment. V neposlední řádě byly stanoveny a popsány nástroje využité pro experiment. V následující kapitole dochází k prezentování a interpretování veškerých získaných dat.

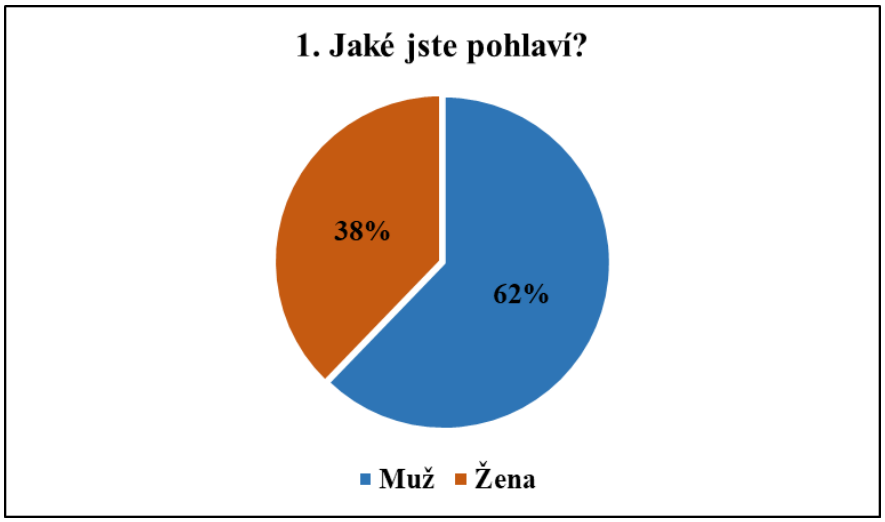
4 VYHODNOCENÍ ANKETY

Ve čtvrté kapitole proběhlo vyhodnocení online dotazníku. V rámci ankety je provedena důkladná analýza všech odpovědí získaných od respondentů. Po prozkoumání odpovědí jsou popsány limity ankety a ověřeny hypotézy stanovené v kapitole „Metodika“.

Anketa byla vyplněna celkem **203** respondenty, z toho **23** respondentů bylo vyřazeno a finální zůstatek byl tvořen **180** respondenty. Z vyjmutých **23** respondentů bylo **16** respondentů odstraněno, protože nebrali dotazník vážně. Zbylých **7** respondentů uvedlo jistou skepsi ohledně 9. otázky, což je více probráno v „Limitech ankety“.

Otázka č. 1: Jaké jste pohlaví?

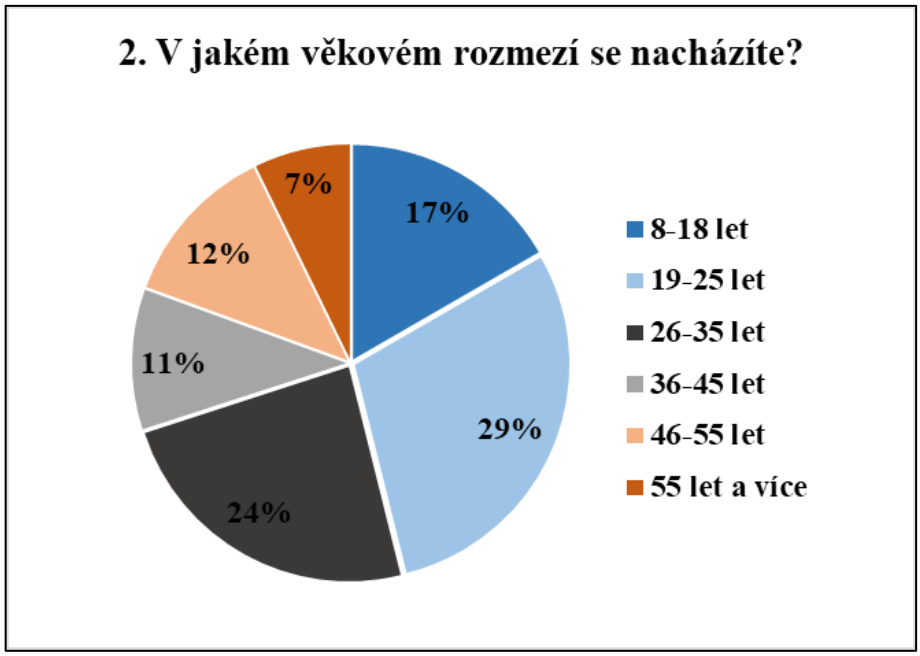
U identifikační otázky č. 1 bylo zastoupeno ženské pohlaví **68 (38 %)** ženami a mužské zastoupení tvořilo **112 (62 %)** mužů.



Graf 2: Graf k otázce č.1.

Otázky č. 2: V jakém věkovém rozmezí se nacházíte?

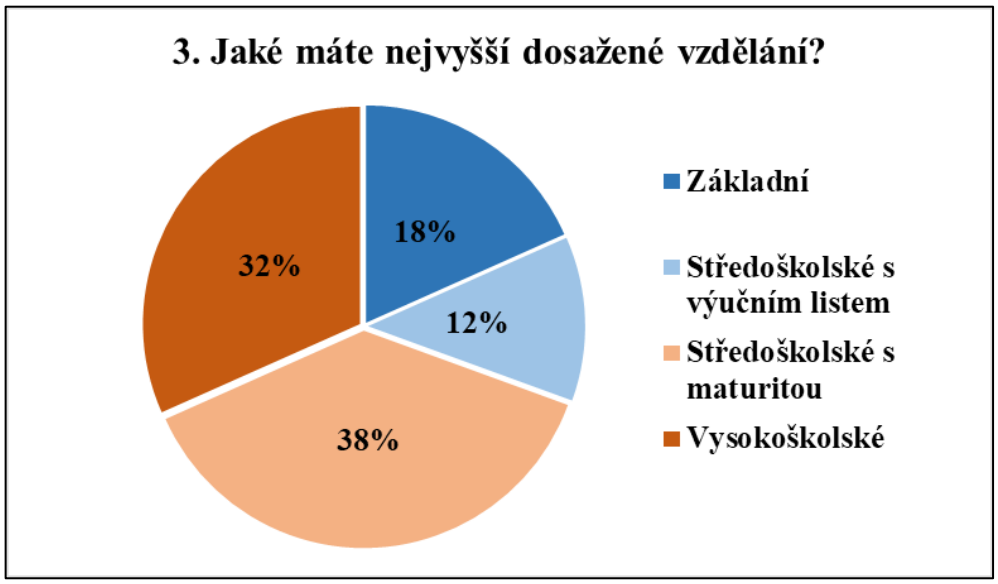
U otázky č. 2 byla nejvíce zastoupena mladší populace do 35 let. Přesněji se jednalo o **30 (17 %)** respondentů mezi 8-18 lety, **43 (24 %)** respondentů mezi 26-35 lety a nejpočetnější skupina **53 (29 %)** respondentů v rozmezí 19-25 let. Starší populace už nebyla tolik zastoupena, ale i přesto byl získán dostatečný vzorek respondentů mezi 36-45, 46-55 a nad 55 let. V absolutní a relativní hodnotě byly věkové kategorie zastoupeny **19 (11 %), 22 (12 %)** a **13 (7 %)**.



Graf 3: Graf k otázce č.2.

Otázka č. 3: Jaké máte nejvyšší dosažené vzdělání?

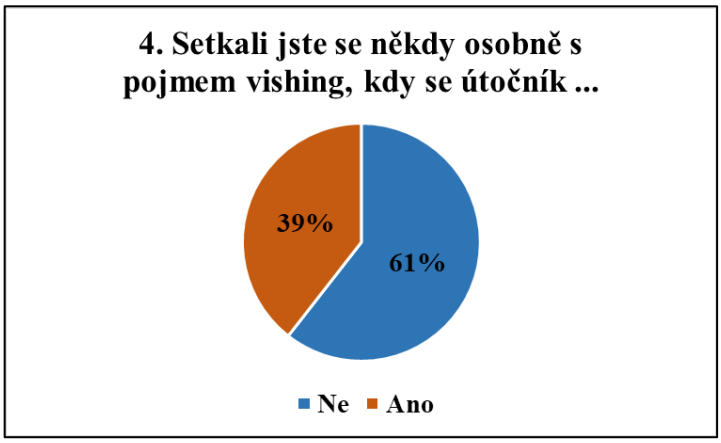
Nejvíce respondentů uvedlo, že jejich nejvyšší dosažené vzdělání je středoškolské vzdělání s maturitou, a to v **68 (38 %)** případech. Vysokoškolské vzdělání uvedlo **57 (32 %)** respondentů. V menší míře byli zastoupeni respondenti se základním vzděláním v počtu **33 (18 %)** a středoškolským vzděláním s výučním listem **22 (12 %)** respondentů.



Graf 4: Graf k otázce č.3.

U otázky č. 4: Setkali jste se někdy osobně s pojmem vishing, kdy se útočník v telefonním hovoru vydává třeba za policistu nebo jinou osobu a snaží se z Vás získat finanční prostředky nebo jiné údaje?

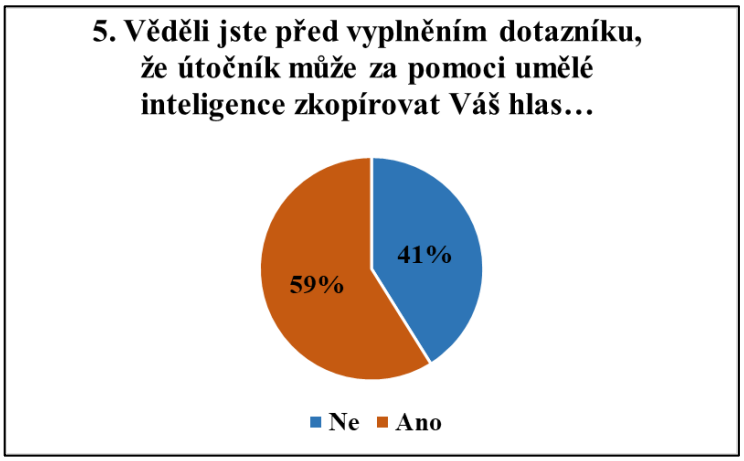
V otázce č. 4 bylo cílem zjistit, zda se lidé již někdy osobně setkali s hrozbou v podobě podvodných telefonátů. K překvapení bylo zjištěno, že osobní setkání s vishingem zažilo **71 (39 %)** respondentů ze 180. Ze zkoumaného vzorku byl tedy hrozbě vystaven každý třetí člověk. Zbylých **109 (61 %)** respondentů se s vishingem nesetkalo.



Graf 5: Graf k otázce č.4.

Otázka č. 5: Věděli jste před vyplněním dotazníku, že útočník může za pomoci umělé inteligence zkopírovat Váš hlas a vytvořit jeho syntetickou (umělou) kopii a podpořit tak věrohodnost vishingových útoků?

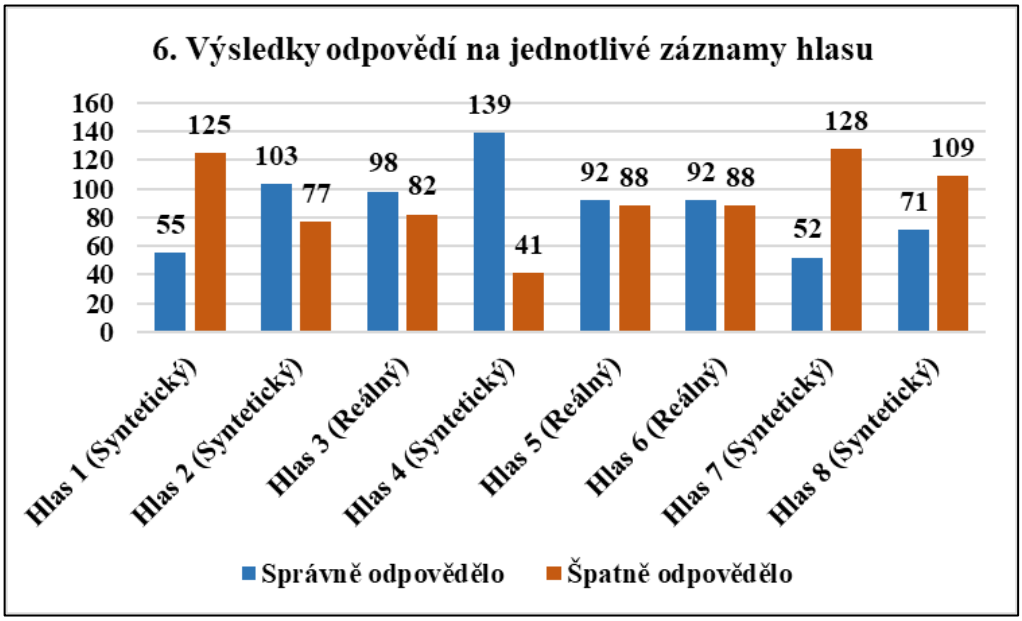
Jak moc je populace obeznámena s podvody se syntetickým hlasem bylo zkoumáno v otázce č. 5. O možnosti naklonovat hlas rodinného příslušníka, přítele atd. doposud nevědělo **74 (41 %)** respondentů. Respondentů, kteří o dané hrozbě věděli, bylo **106 (59 %)**.



Graf 6: Graf k otázce č.5.

Otázka č.6: Tato otázka se zaměřuje na Vaši schopnost rozeznat syntetický hlas od reálného hlasu. Níže máte video, které obsahuje 8 záznamů hlasu. Pusťte si jej a pokuste se určit, zda je hlas reálný či vytvořený umělou inteligencí. (Cesta k videu je v **PŘÍLOZE P VII)**

Cílem této otázky bylo zjistit za pomoci testu, zda jsou lidé schopni rozpoznat reálný hlas od syntetického hlasu a jak na tom technologie syntézy hlasu v ČR je. Test byl sestaven z 8 hlasů a správné odpovědi byly následující: Hlas 1 – syntetický, Hlas 2 – syntetický, Hlas 3 – reálný, Hlas 4 – syntetický, Hlas 5 – reálný, Hlas 6 – reálný, Hlas – 7 syntetický a Hlas 8 – syntetický. Za každou správně zodpovězenou otázku získal respondent 1 bod.



Graf 7: Graf k otázce č.6.

U hlasu 1, 7 a 8 bylo zajímavé, že i když byly syntetické, tak respondentům dělaly největší obtíže. Respondenti často zaměňovali hlas za reálný. To mohlo být způsobeno tím, že v těchto hlasech byla snaha o zachování přirozených pauz, mlaskání nebo používání vyplní jako „ehmm“ nebo „nooo“. U syntetického hlasu 2 a 4 bylo využito lehkého smíchu a změny v intonaci hlasu během mluvení, což třeba pro hlas č.4 bylo nejspíš jedním z odhalitelných prvků, protože tento hlas zvládlo nejvíce respondentů rozpoznat. U hlasů 3, 5 a 6 převažovaly spíše správné odpovědí, ale oproti syntetickým nebylo možné sledovat tak razantní nepoměr v celkovém počtu odpovědí.

Tabulka 1: Statistika z výsledků jednotlivých testů.

Průměr	Medián	Modus	Směrodatná odchylka
3,9	4	4	± 1,8

Z tabulky 1 je patrné, že počet bodů získaných z testu u celého zkoumaného souboru se pohyboval průměrně okolo 4 bodů. To potvrdila i hodnota medián, která není tolik náchylná na vliv extrémních hodnot jako průměr. Byl stanoven i Modus, který určil nejčetnější počet bodů. Směrodatná odchylka určila rozmezí 2,1 – 5,7 bodů od průměrné hodnoty. Podle tohoto rozmezí se stanovilo jednoduché hodnocení. Respondenti do 2 bodů mají horší rozpoznávací schopnost než průměrný respondent a měli by být do budoucna obezřetnější. Respondenti nacházející se v rozmezí od 3 do 5 bodů mají průměrné rozlišovací schopnosti. Respondent, který získal 6 a více bodů, má při aktuálním vývoji technologie velkou šanci rozpoznat klonovaný hlas. Samozřejmě je potřeba počítat s tím, že při časovém nátlaku během útoku a dalších aspektech dané situace, je schopnost rozpoznat hlas značně omezená.

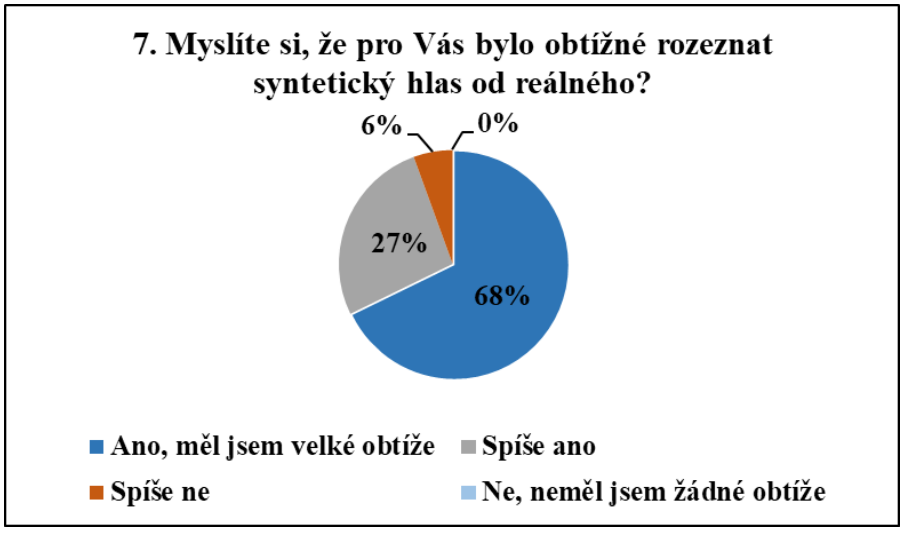
Tabulka 2: Respondenti seřazení podle zisku bodů.

Počet bodů:	Počet respondentů s daným bodovým ziskem.
0 z 8	4
1 z 8	13
2 z 8	22
3 z 8	34
4 z 8	41
5 z 8	32
6 z 8	23
7 z 8	7
8 z 8	4

Z tabulky 2 je patrné, že bodový zisk odpovídá normálnímu rozložení na Gaussově křivce. V rámci zkoumané skupiny v tabulce 2 bylo zjištěno, že **107 (60 %)** respondentů dosáhlo průměrných hodnot. Nejlepší výsledky měla pouze skupina **34 (18 %)** respondentů. K velkému překvapení bylo zjištěno, že zde bylo i **39 (22 %)** respondentů, kteří dosahovali podprůměrných až velmi nízkých výsledků. Z testu jako takového je patrné, že generování hlasu v češtině je aktuálně na úrovni, kdy je obtížné rozeznat reálný hlas od klonovaného.

Otázka č.7: Myslíte si, že pro Vás bylo obtížné rozeznat syntetický hlas od reálného?

Po skončení testu byl respondent dotázán na otázku č.7. Nejvíce respondentů, tedy **122 (68 %)** uvedlo, že mělo velké obtíže s rozpoznáváním. „Spíše ano“ uvedlo jako svou odpověď **48 (27 %)** respondentů a „Spíše ne“ odpovědělo jen **10 (6 %)** respondentů. Většina respondentů neměla potřebu přeceňovat své schopnosti a dokonce i **33** respondentů s výsledky od 6 bodů a výše uvedlo, že měli obtíže rozpoznat reálný a syntetický hlas. Nikdo z respondentů, ani s nižším skóre, neuvedl, že by s tím neměl vůbec žádné obtíže.



Graf 8: Graf k otázce č.8.

Otázka č.8: Pokud by po Vás útočník vyžadoval citlivé údaje (telefonní číslo, údaje na platební kartě) nebo finanční prostředky, jak silnou důvěru by u Vás vzbudil s hlasem uvedeným v možnostech?

Otázka č.8 byla konstruována jako Likertova škála, ve které se zjišťovala úroveň důvěry v jednotlivé položky za pomoci tříbodové stupnice. Respondenti volili možnost na škále 0-2, kdy možnosti byly následující: 0 - Žádná důvěra a silné podezření, 1 - Důvěra, ale slabé podezření, 2 - Silná důvěra bez podezření. Jednotlivé odpovědi respondentů se následně sečetly a do tabulky 3 byl uveden celkový počet pro danou možnost a položku. Navíc byla zjištěna nejčastěji volená možnost za pomoci funkce modus.

Tabulka 3: Absolutní hodnoty pro otázku č.8.

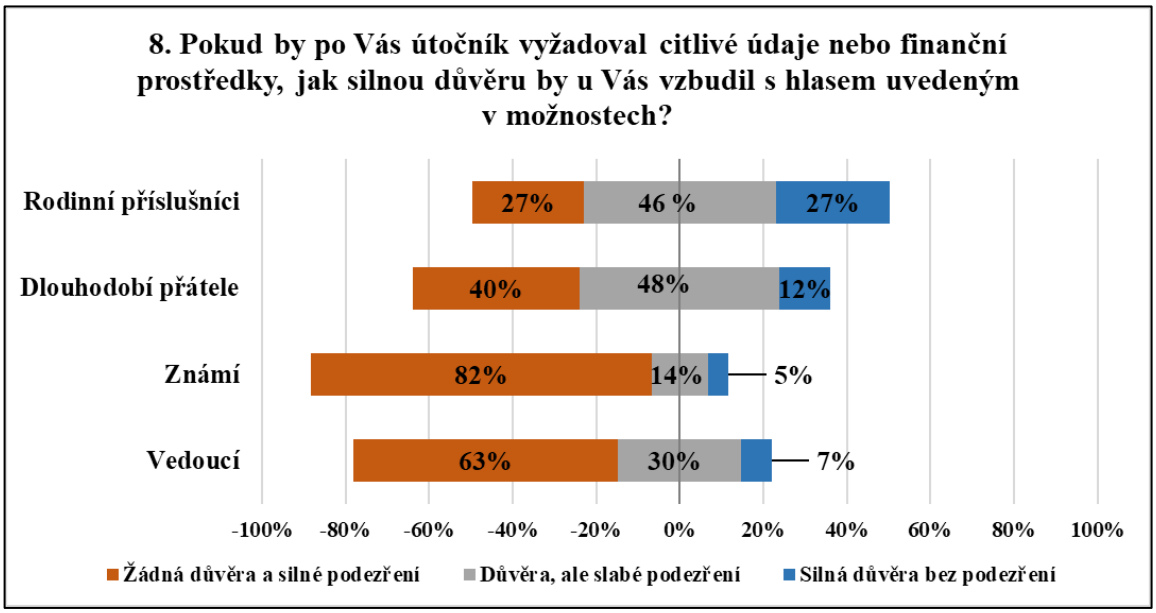
	Žádná důvěra a silné podezření	Důvěra, ale slabé podezření	Silná důvěra bez podezření	Modus
Rodinní příslušníci	48	83	49	**1**
Dlouhodobí přátele	72	86	22	**1**
Známí	147	24	9	**0**
Vedoucí	114	53	13	**0**

Z grafu č.9 je patrné, že největší důvěru měl hlas rodinného příslušníka. Respondenti vyjádřili silnou důvěru ve **49 (27 %)** případech. Celkem **83 (46 %)** respondentů zvolilo jako svou odpověď důvěru, ale už se slabým podezřením. Zbytek respondentů **48 (27 %)** odpovědělo, že by neměli žádnou důvěru, ani pokud by útočník napodobil hlas rodinného příslušníka.

U dlouhodobých přátel měli respondenti silnou důvěru pouze u **22 (12 %)** případů, ale u důvěry se slabým podezřením se oproti předešlé možnosti počet respondentů navýšil. Respondenti v počtu **86 (48 %)** uvedli, že by měli důvěru se slabým podezřením, kdyby jim zavolal dlouhodobý přítel s nějakým zvláštním požadavkem. Co se týče žádné důvěry a silného podezření, zde počet respondentů narostl na **72 (40 %)**.

Hlasu známé osoby, se kterou se respondenti setkávají spíše občas a jedná se například o kolegy z práce nebo občasné přátele, věřili respondenti nejméně, přesněji se k odpovědi „Žádná důvěra a silné podezření“ přiklonilo **147 (82 %)** respondentů. Menší část souboru by důvěřovala, a to v **9 (5 %)** případech silně a ve **24 (14 %)** případech se slabou nedůvěrou.

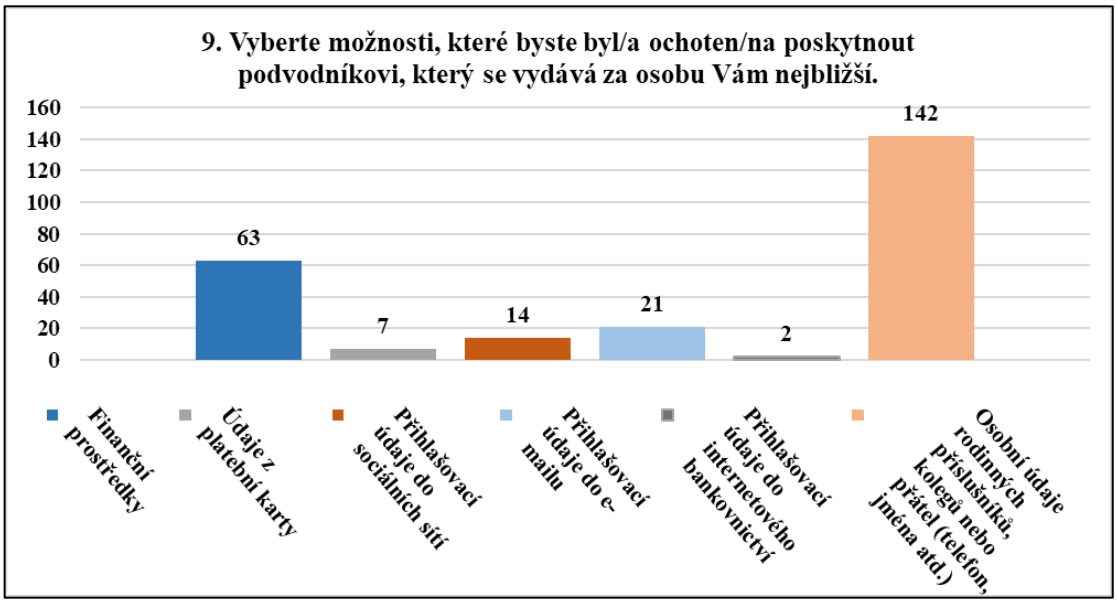
Silnou důvěru v hlas vedoucího vyslovilo **13 (7 %)** respondentů. V druhé možnosti odpovědělo **53 (30 %)** respondentů, že by důvěru měli, ale se slabým podezřením. Poměrně velká část respondentů a to **114 (63 %)** by vedoucímu nedůvěřovala a měla by při požadavku silné podezření, že se něco děje.



Graf 9: Graf k otázce č.8.

Otázka č.9: Vyberte možnosti, které byste byl/a ochoten/na poskytnout podvodníkovi, který se vydává za osobu Vám nejbližší.

Otázka byla zaměřena na to, jaké údaje jsou respondenti schopni poskytnout blízké osobě, za kterou se vydává útočník. Největší zastoupení zde měly odpovědi „Osobní údaje dalších osob“ **142**, „Finanční prostředky“ by poskytlo **63** respondentů, **21** lidem by nevadilo poskytnout „Přihlašovací údaje do e-mailu“, **14** do „Sociálních sítí“, **7** by jich poskytlo „Platební údaje z karty“ a jen **2** osoby „Údaje k internetovému bankovnictví“.

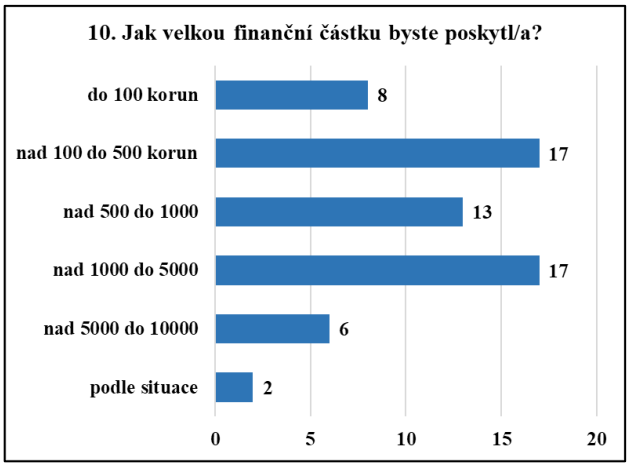


Graf 10: Graf k otázce č. 9.

Otázka č.10: Jak velkou finanční částku byste poskytl/a?

Otázka č.10 byla otevřená a každý mohl napsat libovolnou částku, kterou by byl ochotný poskytnout.

Otázka navazovala na otázku č.9, pokud respondent označil finanční prostředky. V jiném případě otázku nemusel vyplňovat. Ze **63** osob, které by poskytly finanční prostředky, by se **8 (13 %)** respondentů rozhodlo poslat částku do 100 Kč a **17 (27 %)** respondentů by zaslalo částku nad 100 Kč do 500 Kč. Částky v hodnotě nad 500 Kč do 1000 Kč se rozhodlo poskytnout **13 (21 %)** respondentů. S poskytnutím částky nad 1000 Kč do 5000 Kč by souhlasilo **17 (27 %)** respondentů a **6 (10 %)** respondentům by nevadilo poskytnout částku nad 5000 Kč do 10000 Kč. Objevili se i **2 (3 %)** respondenti, u kterých by záleželo na situaci.



Graf 11: Graf k otázce č.10.

Otázka č.11: Myslíte si, že vishing se syntetickým hlasem blízké osoby je nebezpečnější než běžný vishing?

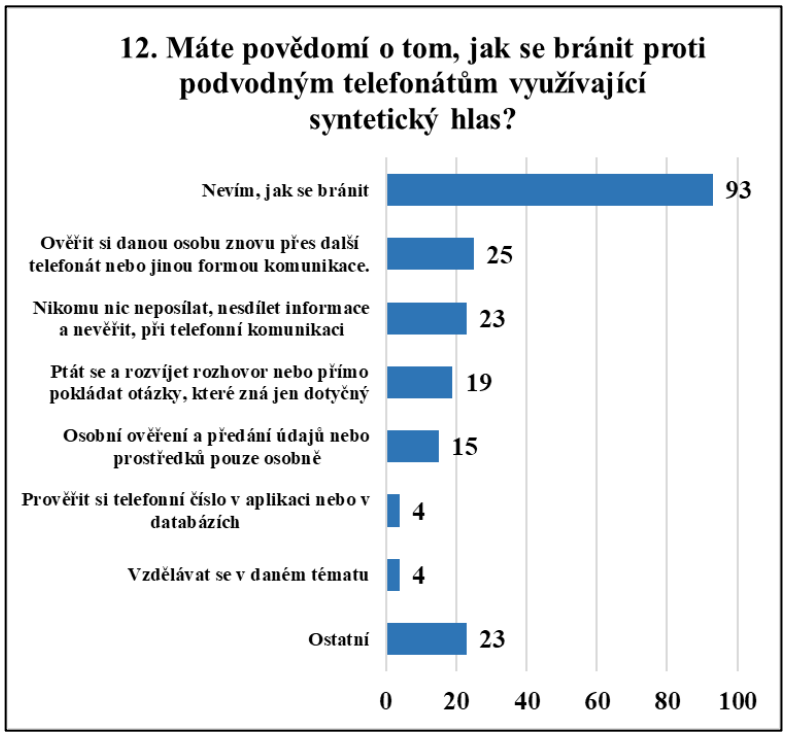
Cílem otázky bylo zjistit, jak je vnímána nová hrozba napříč zkoumaným souborem. Jestli bude respondenty podceňována oproti běžným telefonním podvodům nebo nikoliv. Respondenti se v **161 (90 %)** případech shodli, že vishing útoky se syntetickým hlasem jsou nebezpečnější než běžné telefonní podvody. Jen **2 (1 %)** respondenti si myslí, že není nebezpečnější a **17 (9 %)** staví běžný vishing a vishing se syntetickým hlasem na stejnou úroveň.

Graf 12: Graf k otázce č.11.

Otázka č.12: Máte povědomí o tom, jak se bránit proti podvodným telefonátům využívající syntetický hlas?

Poslední otázka č.12 byla otevřená, kde respondenti mohli psát opatření nebo různé postupy, jak se bránit proti podvodným telefonátům se syntetickým hlasem. Veškeré odpovědi byly analyzovány a přiřazeny do kategorie, ve které byla stejná nebo podobná opatření.

V otázce č.12 bylo **93 (52 %)** respondenty nejčastěji vyplněno „Nevím, jak se bránit“. Zbytek respondentů **87 (48 %)** uvedlo aspoň nějaký způsob, jak by se mohla veřejnost bránit. Respondenti ve **25** případech uvedli, že by si oběť měla „ověřit osobu přes telefon nebo jinou formu komunikace“. Mezi pesimističtější způsob, jenž byl podpořen **23** respondenty, patřilo „Nikomu nic neposílat, nesdílet informace a nevěřit při telefonní komunikaci“. Na toto opatření navazuje druhé s **15** respondenty, kteří vyžadují „Osobní ověření a předání údajů a prostředků pouze osobně“. Poslední opatření, které mělo jedno z větších zastoupení bylo „Ptát se a rozvíjet rozhovor nebo přímo pokládat otázky, které zná jen dotyčný“, jenž napsalo **19** lidí. Mezi menší zastoupení patřili **4** respondenti, kteří poukazovali na potřebu „Vzdělávání v daném tématu“. Další **4** respondenti doporučili „Prověřit si telefonní číslo v aplikaci nebo v databázích“. Do poslední kategorie “Ostatní” byla shrnuta veškerá opatření, která neměla velké zastoupení v rámci respondentů.



Graf 13: Graf k otázce č.12.

V položce “Ostatní” byla uvedena opatření jako zavolat policii, strach z telefonování, paranoia, blokování čísla, spoléhat na nové technologie, snažit se o rozpoznání syntetického hlasu podle chyb v hlasu, nahrát hovor, ignorovat hovory ze záznamníku, heslo, zavěsit telefon, nezvedat cizí čísla a na konec zpozornět, pokud volající vyžaduje peníze nebo jiné údaje.

4.1 Limity ankety

Před ověřováním hypotéz jsou zmíněny některé limity a možná i "kontroverze", které byly zjištěny během vyplňování a vyhodnocení dotazníku. Jednalo se o chyby ve zpracování dotazníku, které by se v dalším výzkumu mohly opravit a rozšířit tak budoucí výzkum o nová zjištění.

Jedna z prvních "kontroverzí" se týká otázky č.9, u které 7 respondentů uvedlo, že by nikdy nic ze zmíněných informací neposkytli. V této otázce bylo pracováno s tím, že při telefonátu můžeme vždy prozradit nějaké informace a pokud by telefonát probíhal s blízkou osobou, o to víc informací by se mohlo poskytnout. Proto zde byla koncipována odpověď ohledně osobních údajů jiné osoby, kam může patřit lokace, jméno, kontakt a mnohé další informace, které by útočníkovi byly nápomocné. Respondenti, kteří odmítli cokoliv poskytnout, jsou uvedeni v tabulce 4. Z tabulky vyplývá, že prvních 5 respondentů v rozpoznávacím testu bylo pod průměrnou hodnotou a 3 z nich své schopnosti lehce přeceňovali. Zbylí dva respondenti se sice dostali na průměrnou hodnotu, ale vůbec do té doby netušili, že útočník může napodobit hlas rodinného příslušníka. Nebylo by vhodné zde uvádět závěry typu poslali by/neposlali by přes telefon nějaké informace, ale předložené informace jsou zrcadlem daným respondentům, že jejich odolnost nemusí být stoprocentní. Otázka č.9 by mohla být do budoucna koncipována jako polouzavřená, aby se podařilo vyhnout těmto problémům a nasbírat více informací pro podobné porovnání, jako tomu je v tabulce č.9. U respondentů by bylo zkoumáno, zda své schopnosti nepřeceňují.

Tabulka 4: Respondenti, kteří byli v rozporu s otázkou č.9.

Odpovědi:				
Pořadí	Otázka č.4	Otázka č.5	Bodové hodnocení z otázky č.6	Otázka č.7
1	Ano	Ano	3	Ano, měl jsem velké obtíže
2	Ne	Ano	2	Spíše ano
3	Ne	Ano	3	Spíše ano
4	Ne	Ano	2	Spíše ano
5	Ano	Ano	2	Ano, měl jsem velké obtíže
6	Ano	Ne	4	Ano, měl jsem velké obtíže
7	Ne	Ne	4	Spíše ano

Další zajímavost se naskytla při vyhodnocení otázky č.12, ve které 7 respondentů uvedlo jako opatření „Nezvedat cizí čísla" nebo „Prověřit si telefonní číslo v aplikaci nebo

v databázích“. Tato opatření nejsou špatná, ale postačí pouze v případech, kdy útočník volá opravdu z cizího čísla. Pokud by útočník využil tzv. spoofing ID volajícího, mohl by se na obrazovce volaného zobrazit, jakkoliv by si zamanul. Při vytváření dotazníku bylo předpokládáno, že ohledně spoofingu je již dostatečná informovanost. [55][56] Bohužel i v rámci 180 respondentů se našlo 7 osob, které nejsou s touto technikou obeznámeny. Proto by bylo vhodné v budoucím výzkumu toto téma zkoumat a zaměřit se i na otázky ohledně spoofingu.

4.2 Ověření hypotéz

Podkapitola se věnuje ověření hypotéz, které byly stanoveny v části „Metodika“. Ke každé hypotéze je stanovena nulová hypotéza (H0) a alternativní hypotéza (HA). V každé hypotéze jsou zobrazena data, se kterými se pracovalo, popsán postup a provedeno zhodnocení.

Hypotéza 1:

Tato hypotéza byla stanovena dle statistik z reportu IC3. Statistika ukazuje rapidní nárůst finančních ztrát u jedinců, kteří jsou starší 30 let a čím vyšší věk, tím byla oběť náchylnější na podvod. Z další statistiky od Microsoftu bylo zjištěno, že nejvíce si ze vzorku 17 000 respondentů osvojuje UI převážně mladší populace ve věku 18-24 let (56 %) a čím starší populace, tím zájem o UI klesá. Proto byla snaha zjistit, zda informovanost ohledně existence hrozby v podobě syntetického hlasu neklesá s přibývajícím věkem. [11][57]

H1: Předpokládám, že respondenti nad 35 let jsou méně informováni o existenci podvodných telefonátů se syntetickým hlasem než mladší populace.

H_10: Předpokládám, že mezi věkovými skupinami nad 35 let a pod 35 let **neexistuje** významný rozdíl v informovanosti o podvodných telefonátech se syntetickým hlasem.

H_1A: Předpokládám, že mezi věkovými skupinami nad 35 let a pod 35 let **existuje** významný rozdíl v informovanosti o podvodných telefonátech se syntetickým hlasem.

Tabulka 5: Kontingenční čtyřpolní tabulka k hypotéze č.1.

		Obeznámenost s podvodnými telefonáty se syntetickým hlasem (otázka č.5)		
		Ne	**Ano**	**Celkem:**
Věk	**Do 35 let**	35	91	**126**
	36 let a více	39	15	**54**
	Celkem:	**74**	**106**	**180**

V tabulce 5 jsou seřazeny získané hodnoty z dotazníku na základě věku a informovanosti o používaní syntetického hlasu. Jelikož bylo potřeba otestovat pouze závislosti mezi věkovým rozsahem nad 35 let a pod 35 let, byly hodnoty dle této hranice syntetizovány do dvou skupin podle tabulky 5. Vznikla tak čtyřpolní tabulka (χ^2_{2x2}), ve které se chí-kvadrát test (χ^2) počítá podle vzorce 1: [53]

$$\chi^2 = n * \frac{(n_{11} * n_{22} - n_{12} * n_{21})^2}{(n_{11} + n_{12})(n_{21} + n_{22})(n_{11} + n_{21})(n_{12} + n_{22})} \quad (1)$$

Hodnota *n* značí celkový počet respondentů, popřípadě počet respondentů v dané buňce. Po dosazení hodnot a vypočítání χ^2 byla získána hodnota $\chi^2_{2x2} = 30{,}84$. Dalším krokem bylo zjistit kritickou hodnotu, podle které by bylo možné zamítnout nebo nezamítnout nulovou hypotézu. Kritická hodnota byla získána ze stupně volnosti (*df*) a hladiny významnosti (*p*). Stupeň volnosti se vypočítá jako součin počtu řádků (r) zmenšeného o jedna s počtem sloupců (c) zmenšeným o jedna dle vzorce 2: [53]

$$\mathrm{df} = (r - 1) * (c - 1) \quad (2)$$

Hodnota *p* se nemusí počítat a ve většině případech se stanovuje na 0,05 (5 %). Dle získaných parametrů *df* a *p* vybíráme ze statistické tabulky kritických hodnot pro požadovanou hodnotu $\chi^2_{df(p)}$. Celkově byly získány tyto hodnoty: $(\chi^2_{2x2}) = 30{,}84$; $df = 1$; $\chi^2_1(0{,}05) = 3{,}84$. [53]

Nepřijímáme nulovou hypotézu, jelikož $\boldsymbol{\chi^2_{2x2} > \chi^2_1(0{,}05)}$. V rámci zkoumaného souboru existuje významný rozdíl v informovanosti ohledně podvodných telefonátů se syntetickým hlasem mezi skupinami nad 35 let a pod 35 let.

Hypotéza 2:

V hypotéze bylo zkoumáno, zda má dosažené vzdělání nějaký vliv na rozpoznávání generovaného audio obsahu. Hypotéza byla stavěna na předpokladu, že lidé s vyšším vzděláním mohou lépe odhalit nedokonalosti nebo jiné chyby syntetického hlasu.

H2: Předpokládám, že respondenti s maturitou rozpoznají syntetický hlas od reálného lépe než respondenti bez maturity.

$H_2$0: Předpokládám, že ve schopnosti rozpoznat syntetický hlas od reálného **neexistuje** významný rozdíl v závislosti na stupni dosaženého vzdělání.

H_2A: Předpokládám, že ve schopnosti rozpoznat syntetický hlas od reálného **existuje** významný rozdíl v závislosti na stupni dosaženého vzdělání.

Tabulka 6: Kontingenční tabulka k hypotéze č.2.

		Schopnost rozpoznat syntetický a reálný hlas (otázka č.6)			
		Nejhorší výsledky	**Průměrné výsledky**	**Nejlepší výsledky**	**Celkem**
Vzdělání	**Respondenti bez maturity**	13	31	11	**55**
	Respondenti s maturitou	26	76	23	**125**
	Celkem:	**39**	**107**	**34**	**180**

Aby byla zajištěna dostatečná četnost hodnot pro jednotlivé položky, došlo k syntéze proměnných (Tabulka 6). Proměnná "vzdělání" byla sloučena do dvou skupin, které byly naplněny podle toho, zda respondent měl nebo neměl středoškolské vzdělání s maturitou. Proměnná ve sloupcích odkazuje na dosažené výsledky v testu, kde platí následující: 0-2 body z 8 = nejhorší výsledky; 3-5 bodů z 8 = průměrné výsledky; 6-8 bodů z 8 = nejlepší výsledky. [53]

V kontingenční tabulce 6 jsou uvedeny aktuální hodnoty získané od respondentů (n). Pro výpočet bylo důležité ještě dopočítat hodnoty očekávané (m) a poté se dle vzorce 3 spočítal χ^2. [53]

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(n_{ij} - m_{ij})^2}{m_{ij}} \tag{3}$$

Na závěr se pro porovnání, stejně jako v předešlé hypotéze, dohledala kritická hodnota za pomoci *df* a *p*. Výsledky z χ^2 jsou následující (χ^2_{2x2}) = 0,85; *df* = 2; $\chi^2_2(0,05)$ = 5,99. [53]

Po zjištění, že $\boldsymbol{\chi^2 < \chi^2_2(0,05)}$, byla **přijata nulová hypotéza.** Tedy je možné uvažovat o tom, že ve schopnosti rozpoznat syntetický hlas od reálného neexistuje významný rozdíl v závislosti na stupni dosaženého vzdělání.

Hypotéza 3:

Při stanovení hypotézy se vycházelo z reportu od Phoenix Group, který zmiňoval, že za rok 2021 ve Velké Británii utržili muži větší finanční ztráty než ženy. U žen se částka vyšplhala na 1 133 liber a muži měli ztrátu 2 780 liber, což je více jak dvojnásobek. Proto byla

pozornost věnována tomu, jestli odolnost vůči podvodům a znalost opatření má nějakou souvislost s pohlavím. [58]

H3: Předpokládám, že existuje významný rozdíl ve znalostech o opatřeních proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví.

H30: Předpokládám, že **neexistuje** významný rozdíl ve znalostech o opatřeních proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví.

H3A: Předpokládám, že **existuje** významný rozdíl ve znalosti ohledně opatření proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví.

Tabulka 7: Kontingenční tabulka k hypotéze č.3.

		Znalost opatření (otázka č.12)		
		Neznají opatření	**Znají opatření**	**Celkem:**
Pohlaví	**Muž**	52	60	112
	Žena	41	27	68
	Celkem:	93	87	180

Opět byla využita čtyřpolní tabulka pro test nezávislosti dvou proměnných. Jedna z proměnných byla tvořena pohlavím a druhá proměnná byla syntetizována na kategorie, zda respondenti znají či neznají opatření (Tabulka 7). Mezi respondenty znající opatření spadali ti, kteří uvedli aspoň nějaké opatření v otázce č.12. Hypotéza stála na předpokladu, že by jedno z pohlaví mohlo být náchylnější na novou hrozbu, skrze minimální povědomí o tom, jak se proti hrozbě bránit. [53]

Postup byl totožný jako u první hypotézy, kdy se podle vzorce 1 vypočítala hodnota χ^2 = 2,71. Následně se určil *df* a *p*. Po dohledání kritické hodnoty byly získány tyto hodnoty (χ^2_{2x2}) = 2,71; *df* = 1; $\chi^2_1(0,05)$ = 3,84. [53]

Po porovnání hodnoty χ^2_{2x2} a $\chi^2_1(0,05)$, *kdy platí, že* $\boldsymbol{\chi^2_{2x2} < \chi^2_1(0,05)}$), je z výsledků možné **přijmout nulovou hypotézu**. To znamená, že neexistuje statisticky významný rozdíl ve znalostech o opatřeních proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví.

4.3 Shrnutí

Provedením a vyhodnocením ankety byl naplněn první dílčí cíl praktické části. V rámci oslovené populace bylo zjištěno, že 106 respondentů o dané hrozbě v podobě podvodných telefonátů se syntetickým hlasem ví. Z těchto 106 respondentů věděla o hrozbě převážně mladší populace, což bylo ověřeno i v první zkoumané hypotéze, která uvedla, že existuje rozdíl v informovanosti mezi věkovým rozmezím nad 35 let a pod 35 let. V rámci opatření je tedy možné pracovat s předpokladem, že je potřeba více informovat starší populaci o aktuální hrozbě.

Další pohled na problematiku poskytl test, kterým se posuzovala schopnost rozpoznat reálný a syntetický hlas. Výsledky testu od zkoumaného souboru byly spíše průměrné. Pouze 34 respondentů zvládlo vyplnit test nad 5 bodů. Zbylých 107 respondentů mělo průměrné výsledky a 39 respondentů mělo nejhorší výsledky. Prostřednictvím druhé hypotézy bylo zkoumáno, zda má dosažené vzdělání souvislost se schopností rozpoznávat reálný a syntetický hlas, což bylo v hypotéze přijato jako statisticky nevýznamné tvrzení. Pozitivní bylo, že žádný z respondentů nepřeceňoval své schopnosti v rozpoznávání a byli si vědomi, že jim činí rozpoznávání velké obtíže. Poznatky, které byly zjištěny během testu, jsou využity při tvorbě opatření. Jelikož 82 % výsledků bylo průměrných až podprůměrných, není potřeba se zabývat tím, jak rozpoznat syntetický hlas. Pozornost by měla směřovat spíše k základním pravidlům a tato pravidla vytvářet způsobem, který by oběti dokázal pomoci bez ohledu na hlas volajícího.

Hrozba vishingu se syntetickým hlasem je vnímána 161 respondenty jako nebezpečnější než běžný vishing. V kontrastu s touto informací bylo zjištěno, že 93 respondentů nemá tušení, jak se proti této hrozbě bránit, i když o hrozbě vědí a jsou si vědomi i její nebezpečnosti. Na druhou stranu zde bylo 87 respondentů, kteří poskytli sice pouze základní postupy, jak se bránit, ale v mnoha případech jsou tyto postupy dostačující. Pozitivní bylo, že jen dva respondenti uvedli snahu rozpoznat syntetický hlas během telefonátu. Ostatní respondenti uvedli praktické opatření jako ověření si osoby přes jiné médium, ptát se na otázky, které zná jen daná osoba atd. Tyto metody jsou mnohem účinnější než snaha o rozpoznání syntetického hlasu. Při ověřování poslední hypotézy bylo zjištěno, že neexistuje rozdíl v informovanosti o opatření proti podvodným telefonátům se syntetickým hlasem v závislosti na pohlaví. Pohlaví v rámci zkoumaného souboru nemá vliv na schopnost se lépe bránit. Dále následuje kapitola, ve které je vyhodnocen sociální experiment.

5 VYHODNOCENÍ SOCIÁLNÍHO EXPERIMENTU

V této kapitole je vyhodnocen experiment a analyzovány rozhovory s respondenty. Na závěr jsou zmíněny limity, které byly během experimentu zjištěny. Veškeré nahrávky podvodů jsou umístěny v přiloženém souboru a seznam souborů je uveden v **PŘÍLOZE P VII**. Kvůli citlivosti situace jsou hlasy respondentů upraveny, aby nebylo možné rozpoznat osobu na nahrávce a byla jí zaručena anonymizace.

Sociální experiment byl zamýšlen provést se 6 respondenty. Bohužel jeden respondent vůbec nereagoval na cizí číslo a ani po urgování skrze SMS zprávy neodpovídal. Ozval se až za několik hodin, kdy už nešlo experiment provést. Proto byl z vyhodnocení vyřazen a soubor byl sestaven pouze z 5 respondentů. I když byl vzorek takto malý, poskytl plnohodnotná data a postřehy, které byly nápomocné při tvorbě opatření. Samozřejmě zde byla i jistá limitace, která je více rozvedena v kapitole „Limity sociálního experimentu“.

Sociální experiment byl úspěšný ve třech případech z pěti. To znamená, že 3 respondenti uvěřili syntetickému hlasu a zaslali na bankovní účet 200 Kč. Zbylým respondentům přišel telefonát podezřelý a radši si jej ověřili. Tím byl u těchto respondentů experiment vyhodnocen jako neúspěšný. V rámci respondentky č.1 (R1) nebyly žádné potíže, okamžitě uvěřila syntetickému hlasu a zámince, která byla předložena. Následně byl s respondentkou proveden rozhovor, aby byly zjištěny nějaké informace k budoucím podvodům. Bylo mi řečeno, že se hlas často zasekával, a i po mém vyzkoušení bylo zjištěno, že je nutné způsob provedení předělat. Proto byla technická podoba podvodu následně upravena, což je zmíněno i v kapitole „Nástroje sociálního experimentu“. U respondenta č.2 (R2) byl využit třetí scénář, při kterém na něj bylo silně naléháno, aby zaslal peníze. Samotný rozhovor byl bez potíží, ale kvůli špatně zvolené zámince byl R2 podezřívavý a zkontroloval si, zda jsem opravdu volal já. Nejzajímavější úspěch byl u respondenta č.3 (R3). Po úspěchu s první respondentem byl očekáván snadný průběh, ale nastal pravý opak. Respondent byl velice zvídavý a neustále se na něco doptával. Dokonce se stalo, že respondentovi byla puštěna jiná nahrávka než ta, která měla navazovat. Naštěstí i přes chyby ve scénáři a celkovou pochybnost celého hovoru byl podvod úspěšný. Převážně za to byla zodpovědná dobře zvolená záminka. Ta zajistila u respondenta představu, že během hovoru bylo ještě jednáno s policisty, a proto nebylo možné správně reagovat na jeho dotazy. Kvůli častému vybočení byl scénář č.1 upraven do kratší podoby, aby se předešlo dlouhým rozhovorům, u kterých by mohl respondent nabít podezření. Nový scénář č.2 byl následné vyzkoušen na respondentce č.4 (R4), ale neúspěšně.

R4 hned ze začátku hovoru měla pochyby, zda se jednalo o mně a po zbytek telefonátu byla velice podezřívavá. Jediné, co k tomu R4 uvedla, že jí neseděla intonace hlasu. Bohužel příčina mohla být ve spoustě věcech jako špatně nahraný hlas, špatný signál nebo v hlasitosti nahrávky. Poslední respondent č.5 (R5) byl vystaven scénáři č.3 a v tomto případě oproti R2 byl podvod úspěšný. I když byl respondent značně zmaten, nízká částka a syntetický hlas ho nakonec přesvědčily peníze zaslat.

5.1 Analýza rozhovorů

Po podvodném telefonátu byly provedeny polostrukturované rozhovory pro zjištění zajímavostí a souvislostí mezi názory i ze strany respondentů. Především bylo potřeba zjistit, jak byl respondentem vnímán syntetický hlas, jestli nebylo něco podezřelého nebo naopak měli v něj plnou důvěru.

Otázka č.1: *„Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?“*

Respondenti byli dotazováni na jejich prvotní pocity, které měli při zjištění, že se jednalo o podvod. R1 a R3 byli překvapení při zjištění, že hlas nebyl reálný, ale syntetický. Oba uvedli, že pro ně bylo důležité v první řadě pomoci příteli a následné zjištění pozadí telefonátu je šokovalo.

R3 k tomuto dodal: „*Bylo by to určitě jiný, kdyby mi zavolal cizí člověk, tak to by mi přišlo jako podezřelejší, ale pokud ti zavolá člověk, kterého dobře znáš a je to někdo blízký, tvůj kamarád, nebo prostě takhle… Tak dalo, by se říct, že mu vždycky chceš, nebo prostě máš takovou potřebu, mu vždycky pomoct, si myslím.“*

Setkání s touto situací bylo pro R1 podnětem k úvahám, jak sebe i své blízké proti hrozbě bránit. Jediná R4 měla radost, protože podvod dokázala rozeznat, ale dodala, že pokud by byl podvod lépe zpracovaný a odpovědi by lépe navazovaly, tak možná by byl příště úspěšnější.

Nejstaršího respondenta (R5) mrzelo, že bez dalšího ověření danou částku poslal, i když byla malá. Uvedl: *„No pocit… Pocit to ve mně zanechalo špatný, jelikož ten hlas byl velmi, velmi podobný tobě, ale jelikož se jednalo o malou částku, tak jsem nějak ztratil ostražitost.“*

R1, R2 a R4 byli překvapení, že je již něco takového možné v ČR. R1 nečekala, že je možné vytvořit tak realistickou kopii něčího hlasu v češtině. Předpokládala, že klonování hlasu je možné pouze v zahraničí.

R1: „*Takže mě to rozhodně překvapilo. Brala jsem to tak, že asi jako pokud bude něco až takhle úplně vyspělé, tak to bude furt ještě v zahraničí a ne u nás.“*

I přestože byl R2 obeznámen s technologií deepfake, syntetickým hlasem i podvody s těmito technologiemi spjatými, byl taktéž překvapen, že se něco takového může stát na území ČR.

Otázka č.2: *„Co ti na rozhovoru přišlo podezřelé?“*

Druhá otázka byla zaměřená na zjištění, zda respondenti neměli během telefonátu nějaké pochybnosti ohledně mého hlasu. K mému překvapení čtyři respondenti uvedli, že jim přišel podezřelý spíše obsah rozhovoru než samotný hlas. Jediné R4 neseděla intonace a hlas jí přišel hluboký, ale nedokázala rozvést proč jí to tak přišlo. Navíc jí bylo divné, že o půjčení peněz byla požádána zrovna ona, když jsme si takto peníze nikdy nepůjčovali. Všichni respondenti, na které byl vyzkoušen scénář s policií, zmiňovali, že jim nebyla divná monotónnost v hlase, protože podle nich mohl mít na hlas vliv stres z policie.

R1: „*Monotónnost. No brala jsem to tak, že v momentě, kdy člověka chytnou policisti a řeší takto nepříjemnou věc. Tak je to takové, že rozhodně nebude člověk vysmátý nebo podobně.“*

Dokonce u R3, jak už bylo zmíněno, byla záminka s policii velmi vhodně zvolená, protože dokázala ospravedlnit dlouhé prodlevy a špatně volené a nenavazující odpovědi. Ohledně obsahu záminky bylo ještě podezřelé dle R1 a R3, že policie měla platební terminál a hovor nebyl uskutečněn z mého telefonního čísla. I v těchto případech bylo respondenty vytvořeno ospravedlnění, že mohlo dojít ke ztrátě nebo vybití telefonu, popřípadě modernizaci platebních metod u policie.

R1: „*Co mě jako překvapilo, je vlastně že voláš z cizího čísla, ale to mohl být vybitý telefon, nemusel si mít kredit cokoliv.“*

R2 a R5 můj hlas nepřišel nijak divný a okamžitě jej poznali, ale pochybnosti měli u způsobu sdělení. Jednalo se o naléhavý scénář č.3 připravený pro členy rodiny. Oběma přišlo podezřelé, že byly požadovány peníze bez řádného vysvětlení a pod velikým časovým nátlakem.

R5 sdělil: „*No intonace… to mě ani tak nepřišlo. Spíš rychlost mluvy, protože v normálním životě, když se spolu bavíme o nějakých problémech, tak prostě nemáš takový hlas, teda hlas máš, ale nemluvíš takto naléhavě.“*

Podezření bylo také hodně ovlivněno požadovanou částkou. Tři respondenti uvedli, že částka byla hodně nízká, tím pádem jim nedělalo takový problém peníze zaslat a i R2, který nebyl

podveden, zmiňoval, že mu částka přišla nízká. Jediná R4 nechtěla zaslat ani tak nízkou částku a jak bylo zmíněno, bylo jí divné, že jí o to vůbec žádám.

Otázka č.3: „*Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?“*

Třetí otázka byla zaměřená na zjištění, zda zážitek ovlivnil důvěru ve schopnost respondentů rozpoznat podvod. Popřípadě jestli hodlají být do budoucna obezřetnější a jakým způsobem. Na schopnost R1 měl zážitek negativní dopad. R1 nedělalo v minulosti problém rozeznat smishing nebo vishing podvody, ale jakmile to byla blízká osoba, hned se jí snažila pomoc.

R1: „*No, ale v tomto případě je to spíš takové těžší v tom, že všichni chceme nějak vždycky pomoct kamarádům, tak asi nejsme hned paranoidní.“*

Na doplňující otázku ohledně větší paranoi v budoucích telefonátech R1 uvedla, že je běžně paranoidní, ale svým kamarádům se bude snažit vždy pomoci, i když už o dané hrozbě ví. Do budoucna ale podobné požadavky hodlá více ověřovat a o dané hrozbě si více zjistit.

R2 byl rád, že situaci zhodnotil dobře a podivný telefonát si ověřil. K doplňující otázce ohledně jeho opatření dodal, že vůbec nezvedá cizí čísla. Naštěstí byl R2 v zahraničí a pro jistotu bral i cizí čísla, kdyby se jednalo o nějaký problém v ČR. Jinak čeká až se daná osoba třeba ověří skrze SMS. Tahle odpověď evokovala otázku ohledně spoofingu, ale R2 zachoval chladnou hlavu, a přestože to byla pro něj novinka, byl schopný popsat správný postup obrany.

R2: *„ …že tím, jak mi ten rozhovor přišel podezřelý tou rychlostí. Tak bych prostě šel jako do telefonu a zavolal na číslo, které mám uložené pod tebou a zase ti zavolal, jestli jako všechno v pohodě.“*

R3 k třetí otázce podotkl: *„ … že už vlastně v dnešní době nemůžeš věřit ani svým dobrým kamarádům, …“*

Tímto navazoval na skutečnost, že útočník si v dnešní době může zcizit hlas jakéhokoliv člověka a zvládne z lidí vymámit finanční prostředky. K doplňující otázce ohledně paranoi zmínil, že jakékoliv transakce bude řešit osobně. Dále uvedl, že pokud by osobní setkání nebylo možné, zkusil by se pravdy dopátrat skrze rodinou skupinu na sociálních sítích. Popřípadě by byl po této zkušenosti více podezřívavý k naléhavým a zkratkovitým zprávám.

U R4 měl zážitek pozitivní vliv na její rozpoznávací schopnost. Uvedla, že je přirozeně nedůvěřivá a každé cizí číslo si nejdříve prověřuje skrze veřejné databáze, ale pokud v nich není, tak jej vezme.

R4 znovu uvedla: „*Ale myslím si, že kdyby ten tón byl stejný, tak bych to možná nepoznala. Řekla bych si, jooo, tak asi ti to nikdo nezvedal, tak si zkusil mě.“*

Následně se ale nad tímto tvrzením zamyslela a zmínila, že by si hovor stejně ověřila přes jiné médium. Poté byla položena doplňující otázka ohledně její paranoi týkající se budoucích telefonátů a taky byla informována o existenci spoofingu. R4 byla překvapená, že je spoofing možný.

R4: *Jako asi nad tím budu možná teďka víc přemýšlet, když vím, že se mi to může stát. No asi teďka budu taková opatrnější, že když po mě třeba někdo něco bude chtít, tak si to nějak jako dvakrát ověřím.*

R5 k dané otázce uvedl pouze: „*No tak, jelikož jak se říká ta umělá inteligence, je velmi, velmi zákeřná. Tak to prostě, i když se bude jednat o malé částky a podobnost hlasu. Přesto si budu ověřovat věci no.“*

Otázka č.4: „*Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti?“*

Čtvrtá otázka byla zaměřena na návrhy respondentů ohledně přístupu společnosti k podvodným telefonátům se syntetickým hlasem. Všichni respondenti by uvítali nějaký druh školení, informační videa sdílená v médiích nebo na sociálních sítích, a to hlavně pro starší populaci. Kromě respondenta č.5 bylo všemi respondenty uvedeno, že by bylo potřebné o problematice informovat i mladší populaci a zahrnout výuku o moderních hrozbách do učebních plánů, popřípadě aspoň vytvářet přednášky nebo programy na toto téma. Od respondentů bylo svědomité, že v potaz brali i jinou věkovou kategorii než právě seniory. Opravdu výstižné bylo vyjádření respondenta č.3.

R3: „*Já jsem si třeba myslel, že na tohle se můžou napálit jenom lidi ve starším věku, kteří mají tu důvěru i k cizímu člověku. Na mě ale teďka jde vidět, že se může napálit i člověk, který má obecné znalosti o této problematice a zajímá se o novinky v oboru IT.“*

R1, R2 a R5 doplnili, že je nutné informovat nejen o existenci hrozby, ale navíc zmínit i opatření, která by byla proti ní účinná nebo způsob, jak syntetický hlas rozeznat.

R2: *„Asi nějaký jakoby výcuc metod, které by mi pomohly tady ty podvodné telefony odhalit třeba jako, jak postupovat krok za krokem, když mám podezření, že se něco takového děje.“*

5.2 Limity sociálního experimentu

Jak bylo zmíněno na začátku kapitoly, jednou z limitací byl nízký počet respondentů. Práce s takto malým vzorkem respondentů byla zapříčiněna podobou sociálního experimentu. Jelikož byl experiment konstruován od začátku jako podvod, bylo zde silné morální a etické hledisko, které nedovolovalo vyzkoušet experiment na více lidech. Nemohl jsem experiment cílit na kohokoliv, koho jsem si usmyslel, ale bylo nutné opatrně vybírat osoby, u kterých bylo očekáváno pochopení a hlavně odpuštění, že jsem si k nim tohle dovolil. Nesmí se opomenout ani právní hledisko, které zabraňovalo zneužít hlas cizí osoby a vydávat se za ni. Jelikož by se jednalo o zcizení osobních údajů v podobě hlasu a po následném klonování by byla zcizena i celá identita dané osoby. Pokud by se stalo, že daná osoba by dala souhlas s propůjčením hlasu, stále by zde bylo morální a etické hledisko, které by mi nedovolilo zneužít důvěry cizí osoby, která se mnou nemá nic společného. Kvůli této právní stránce byl použit pouze můj hlas. I tak v rámci zákona nebyl experiment zcela beztrestný, ale bylo pracováno s myšlenkou, že pokud je pochopeno, jak podvod funguje a jaké má limity, je dosaženo vědomosti, jak se proti hrozbě bránit. Na daných zjištění tedy závisela tvorba opatření, která mohou pomoci více lidem, než kterým bylo uškozeno. Proto věřím, že i ohledně problémů s právním hlediskem získám pochopení. Velký vliv měla na experiment i časová náročnost, protože bylo potřeba vymyslet funkční scénáře a vytvořit nahrávky s mým syntetickým hlasem pro každý scénář. Nahrávky navíc byly vícekrát namluveny, aby zněly uvěřitelněji, čímž byla délka zpracování taktéž ovlivněna. V dalším výzkumu by mohl být sociální experiment proveden jako penetrační testování v rámci firem. Tím by se eliminovaly výše zmíněné problémy a byl by poskytnut dostatek respondentů.

Další limitace byla zjištěna ve formě provedení sociálního experimentu jako takového. I když byl experiment ze 60 % úspěšný, je nutné dodat, že zvolená forma předem nahraných odpovědí se syntetickým hlasem není tou nejlepší cestou, kterou by se podvodníci mohli vydat. Způsob provedení obsahoval spoustu problémů v podobě neflexibilních odpovědí, nenávaznosti na otázky a dlouhých pomlk. Tyto problémy vzbuzovaly v respondentech nedůvěru, kterou musela často zachraňovat dobře zvolená záminka. I přesto by předem nahrané odpovědi mohly útočníkům sloužit jako podpůrná metoda na podobném principu, jako tomu bylo u příkladu pana Garyho nebo pana Cicmana v kapitole „Vishing“ a „Deepfake“.

Sociotechnik by na podobném principu nejprve pustil nahrávku blízké osoby v tíživé situaci a následně by se vložil do konverzace jako policista nebo jiná autorita. Budoucnost těchto podvodů je možné spatřit v převodech hlasu v reálném čase, ale jak již bylo zmíněno, pro sociotechnika by to znamenalo velkou počáteční investici do výpočetní techniky. Na druhou stranu by záleželo na tom, zda by takový podvod dosahoval požadované návratnosti na pokrytí vstupních nákladů. Poslední omezení, které bylo odhaleno při přípravě, se týkalo dohledání telefonních čísel. Pokud byli dohledáni blízcí lidé, bylo náročné dohledat osobní telefonní čísla. Proto je vishing se syntetickým hlasem vhodnou technikou pro útoky pro dlouhodobé kampaně, které trvají někdy i několik měsíců až let, aby dokázaly nashromáždit veškeré informace o oběti.

5.3 Shrnutí

Provedením sociálního experimentu byl splněn i druhý dílčí cíl praktické části. Na základě zjištění, kterých bylo dosaženo během přípravy experimentu, práce se syntetickým hlasem a z rozhovoru s respondenty, je možné tvrdit, že podvody se syntetickým hlasem jsou aktuálně realizovatelné na území České republiky. Technologie syntézy hlasu je na takové úrovni, že bylo možné kopírovat hlas v češtině s opravdu velkou podobností k originálu. Proveditelnost podvodu potvrzují i 3 respondenti, kteří syntetickému hlasu uvěřili a zaslali finanční prostředky.

Pro respondenty byl tento zážitek zajímavou zkušeností, která u některých vyvolala překvapení, že už se mohou takové podvody dít i v ČR. Nejdůležitější zjištění z rozhovoru bylo, že 4 respondenti poznali a projevili důvěru v syntetický hlas a z toho 3 respondenti podvodu plně podlehli. Pochybnosti, které zazněly nejčastěji, se vztahovaly k návaznosti odpovědí, monotónnosti v hlase, cizímu číslu, vlastnictví platebního terminálu policií a u scénáře č.3 byl problém s akutností požadavku. Zde bylo zajímavé zejména to, že pro úspěch experimentu nestačilo pouze mluvit se syntetickým hlasem, ale i dobře zvolená záminka měla zásadní dopad na výsledek. Syntetický hlas a záminka u podvedených respondentů vyvolala snahu si dané chyby jakýmkoliv způsobem ospravedlnit. Přikládat je třeba stresu z policie, vybitému telefonu nebo modernizací policejního vybavení atd. U třetí otázky bylo pozorováno uvědomění respondentů, kteří do budoucna budou obezřetnější a podivné požadavky si budou raději ještě jednou ověřovat. Od respondentů byl v poslední otázce projeven velký zájem o informovanost, především je zajímala opatření proti hrozbě nebo co všechno UI

dokáže. V následující kapitole jsou na základě zjištění ze čtvrté a páté kapitoly vytvořena opatření.

6 OPATŘENÍ

V následující kapitole jsou navržena opatření, která mohou posílit obranu jedince proti vishing podvodům se syntetickým hlasem. Opatření jsou vytvořena na základě informací, které byly zjištěny během návrhu a provedení sociálního experimentu. V potaz jsou brány i opatření, která zmiňovali respondenti v dotazníku. Jelikož je celý podvod se syntetickým hlasem založen na běžném vishingu, je možné využít již existující opatření, která jsou nápomocná i proti vishingu se syntetickým hlasem. Opatření jsou navržena způsobem, aby zvládla pokrýt i podvody využívající spoofing.

Opatření, která zde vznikla, jsou mířena na běžného člověka, který si z těchto opatření může odnést základní i pokročilejší informace, jak se bránit proti vishingu se syntetickým hlasem. Podklady mohou být užitečné pro státní i soukromý sektor. Například soukromé firmy mohou implementovat upravená pravidla dle svých potřeb do své bezpečnostní politiky.

6.1 Základní pravidla

Nejprve bych chtěl zmínit, že bezpečnost každého jedince je silně ovlivněna tím, kolik informací dá o sobě vědět v digitálním prostoru. Čím méně osobních informací je útočníkovi poskytnuto, tím menší manipulační prostor dostává během útoku: [59]

- Skrýt nebo odstranit veškeré osobní i citlivé informace na sociálních sítích, i před přáteli. Nenechat útočníka zjistit jakékoliv informace jako jsou datum narození, telefonní čísla, údaje o přátelích a rodině ze seznamu přátel a mnohé další.
- Soukromý účet na sociálních sítích spravovat způsobem, aby se informace o osobě dostaly pouze k jeho nejbližším. Využívat funkci blízkých přátel a sdílet jakékoliv fotky nebo informace pouze mezi ně.
- Stáhnout aplikace, které v rámci možností odhalují a informují o nebezpečných číslech. O těchto aplikacích více v podkapitole „Technická opatření“.
- Aktivace vícefaktorového ověření. Základem by mělo být internetové bankovnictví, sociální sítě, e-maily, hesloví manažeři a antivirové programy. Dle specifik jedince rozšířit na další. Ověřovací kód nikdy nikomu nesdělovat.
- Aplikace pro komunikaci typu WhatsApp, Telegram, Signal atd. nastavit do režimu, ve kterém není možné dohledat profilový účet podle telefonního čísla, ale pouze podle jména. Popřípadě umožnit dohledat účet podle čísla jen kontaktům, které už jsou uloženy v telefonu.

- Neustále se vzdělávat o nových hrozbách.

Tohle byla pravidla preventivního charakteru, tedy rady, jak se co možná nejlépe vyhnout nevyžádanému telefonátu. Pokud už útok probíhá, je potřeba přejít k mitigačním opatřením, které zvládnou útočníka se syntetickým hlasem efektivně odrazit: [59][60]

- Je-li v telefonátu vyvíjen, i nepatrně, jakýkoliv časový nebo emocionální nátlak a nutnost něco sdělit nebo poslat, okamžitě zpozornět a být obezřetný. Dokonce i když jde o rodinu, přátelé nebo vedoucího v práci a je voláno z jejich osobních čísel.
- Vždy, když v telefonátu volající vyžaduje finanční částky nebo informace, je nutné telefon okamžitě zavěsit a v kontaktech vyhledat číslo volající osoby a zavolat dané osobě zpět. Další možností je ověření přes sociální sítě nebo osobně.
- Nikdy nepoužívat telefonní čísla, která diktuje volající. Vždy si kontaktní údaje o blízké osobě dohledat skrze kontakty v telefonu nebo získat kontaktní údaje přímo od organizace skrze webové stránky nebo jiným způsobem.
- Cizí čísla, která volají na osobní telefonní číslo, nezvedat, dokud se v SMS nebo jiným způsobem daná osoba neprokáže. Popřípadě využít k ověření čísla aplikaci pro zjištění volajícího.
- Pokud útočník využil spoofing nebo syntetický hlas, okamžitě informovat osobu, které číslo nebo hlas patří, aby mohla podniknout patřičné kroky. Nejlépe informovat všechny známé a informovat patřičné orgány.
- Reportovat podvodné telefonáty Policii ČR, telefonním operátorům a taky zapisovat čísla do aplikací určených k tomuto účelu. Důležité je zmínit, že útočník pracuje se syntetickým hlasem nebo využívá spoofingu čísla známé osoby k podvodům.
- Číslo, které bylo zcizeno spoofingem okamžitě přestat používat a pořídit si jiné.

Nejlepší možným řešením je teda telefonní hovor okamžitě ukončit a zkontrolovat si volajícího jiným způsobem. Pokud osoba nevykazuje zájem o žádné specifické požadavky, ale jsou zde nějaké nejasnosti třeba v hlase, je možné na volajícího vyvíjet nátlak: [59]

- Nenechat vést útočníka rozhovor. Je-li rozhovor v útočníkových rukách, je nutné vybočit z připraveného rozhovoru.
- Konfrontovat jej z mnoha dotazy, které může znát jen pravý volající nebo říct lež, kterou by pravý volající odhalil a sledovat útočníkovu reakci.
- Využívat potvrzovacího hesla mezi volajícími. Toto opatření je vhodné spíše pro menší okruh lidí, jako jsou právě rodinní příslušníci.

6.2 Hlavní znaky telefonátu se syntetickým hlasem

Z ankety vyplynulo, že odhalení útoku snahou rozpoznat syntetický hlas od reálného je velice obtížné. A pokud se k dobře naklonovanému hlasu přidá časový nátlak a další faktory, je rozpoznání hlasu až nemožné. Proto je nutné používat raději základní pravidla než se snažit o rozpoznání hlasu. I přesto jsou zde uvedeny hlavní rysy, které mohou u nekvalitních útoků pomoci s rozpoznáním.

Prvně je potřeba brát na vědomí, že rozpoznatelnost syntetického hlasu hodně záleží na zvoleném modelu. Pokud je zvolen text na syntetický hlas, je v nahrávkách často slyšitelná výrazná mezera mezi slovy a tento model se dá lehce odhalit. Větší problém nastává u modelu hlas na syntetický hlas, který zvládá chybu výrazných mezer eliminovat. Při práci se syntetickým a normálním hlasem bylo zjištěno, že generovaný hlas má často hlubší tóninu, což může být taky jedna z možností pro rozpoznání. U obou modelů je nutné vytvořit nahrávky klonovaného hlasu předem, což v rámci telefonátu způsobí, že útočník nereaguje na položené otázky nebo reaguje opožděně.

- Dávat pozor jak na dlouhé pomlky mezi slovy, tak i na opožděné odpovědi.
- Brát na zřetel výšku hlasu a zpozornět při výskytu pouze hlubšího tónu hlasu.
- Všímat si monotónnosti v hlase během rozhovoru a chybějícího emočního zabarvení hlasu (hlas bez smíchu, smutku atd.).
- Věty volajícího nenavazují na otázky volaného.

Tyto znaky bohužel neplatí na modely, které jsou schopny převádět hlas útočníka na syntetický v reálném čase bez nutnosti vytvářet nahrávky hlasu předem. Převody v reálném čase zatím nejsou tolik rozšířené, protože potřebují k funkci výkonné počítače, a i tvorba hlasového modelu je časově náročnější. Pokud by byl útočník vyzbrojen tímto modelem, byl by jeho hlas měněn s minimálním zpožděním na hlas jakéhokoliv člověka. Opatření proti tomuto modelu naštěstí zůstává stejné a hlavními zásadami jsou zpozornět u požadavků o údaje nebo peníze, nic nesdělovat, zavěsit a ověřit si požadavek ještě jednou přes jiný komunikační kanál.

6.3 Technologické opatření

Pokud už je nutné mít veřejně umístěné telefonní číslo na internetu, je na místě využívat dvou telefonů nebo, pokud to zařízení umožňuje, používat dvě SIM. Takhle si osoba může vytvořit osobní číslo pouze pro rodinu a přátelé a pracovní číslo pro zákazníky a komunikaci

s veřejnými subjekty. Nejmodernějším způsobem je využívání e-SIM, což je SIM karta, která je implementována přímo v obvodech telefonu. Výhodou této technologie je možnost využívat více operátorů a s tím i více čísel. Uživatel tedy může každému číslu přiřadit účel, ať už pro osobní nebo pracovní využití. Tímto rozdělením lidí na osobní kontakty, pracovní atd. je docíleno toho, že by se útočník dostal pravděpodobně jen k pracovnímu číslu. Kdyby se chtěl následně vydávat za člena rodiny, měl by volaný možnost včasně odhalit podvodníka, jelikož by útočník nevolal na osobní číslo. [61]

Mezi aplikace, které jsou nápomocné proti nevyžádaným telefonátům, patří například Call Insider, Můžu to zvednout? nebo Kdo mi volal?. Tyto aplikace pracují s databází záznamů, ve které se schraňují informace o nebezpečných číslech. Pokud se na telefonu zobrazí neznámé číslo, je okamžitě otestováno na shodu v databázi. Jedná-li se o nebezpečná čísla, marketingová sdělení atd. je na obrazovce telefonu zobrazeno hodnocení a recenze k danému číslu. U vishingu se syntetickým hlasem je důležité psát recenze, které by další uživatele informovaly, že útočník může využít tuto hrozbu. Pokud by útočník zvládl vytvořit i spoofing čísla blízké osoby, měla by osoba tuto skutečnost zmínit v komentářích a informovat ostatní, že byla vystavena spoofingu. Číslo si následně změnit a používat jiné. [62]

Do budoucna vidím potenciál v aplikacích generující unikátní čísla, která lze napojit na aktuální hlavní číslo v telefonu. Aplikace Cloaked, která je zatím dostupná pouze v beta verzi na území USA, dokáže zakrýt hlavní telefonní číslo náhodně generovaným číslem. Unikátních čísel může být neomezeně a každé může být poskytováno k jinému účelu od komunikace s rodinou po pracovní telefonáty. Uživatel aplikace může vcelku přesně určit volajícího, jelikož aplikace zobrazuje, z kterého vygenerovaného čísla je voláno. Odhalení útočníka se syntetickým hlasem by bylo zase o něco jednodušší, jelikož by se volanému útočník neozýval na osobní číslo, ale na úplně jiné. [63]

Na druhou stranu zmíněné aplikace jsou spravované třetí stranou, která může mít přístup k telefonním číslům a dalším datům. Další věcí je, že pokud se útočník dostane do účtu přes chybu uživatele, dostane se automaticky ke všem datům. Každý z nás si tedy musí vyhodnotit klady a zápory a říct si, zda je ochoten data sdílet s třetí stranou či nikoliv.

V návaznosti na podvody se syntetickým hlasem a deepfake obsah začaly organizace vyvíjet i nástroje pro obranu, které využívají UI. První vyvíjená technologie se jmenuje AntiFake a její princip spočívá v znesnadnění syntézy hlasu cizích osob. Technologie mírně zkresluje zaznamenaný zvukový signál tak, že stále zní správně lidským posluchačům, ale pro UI je

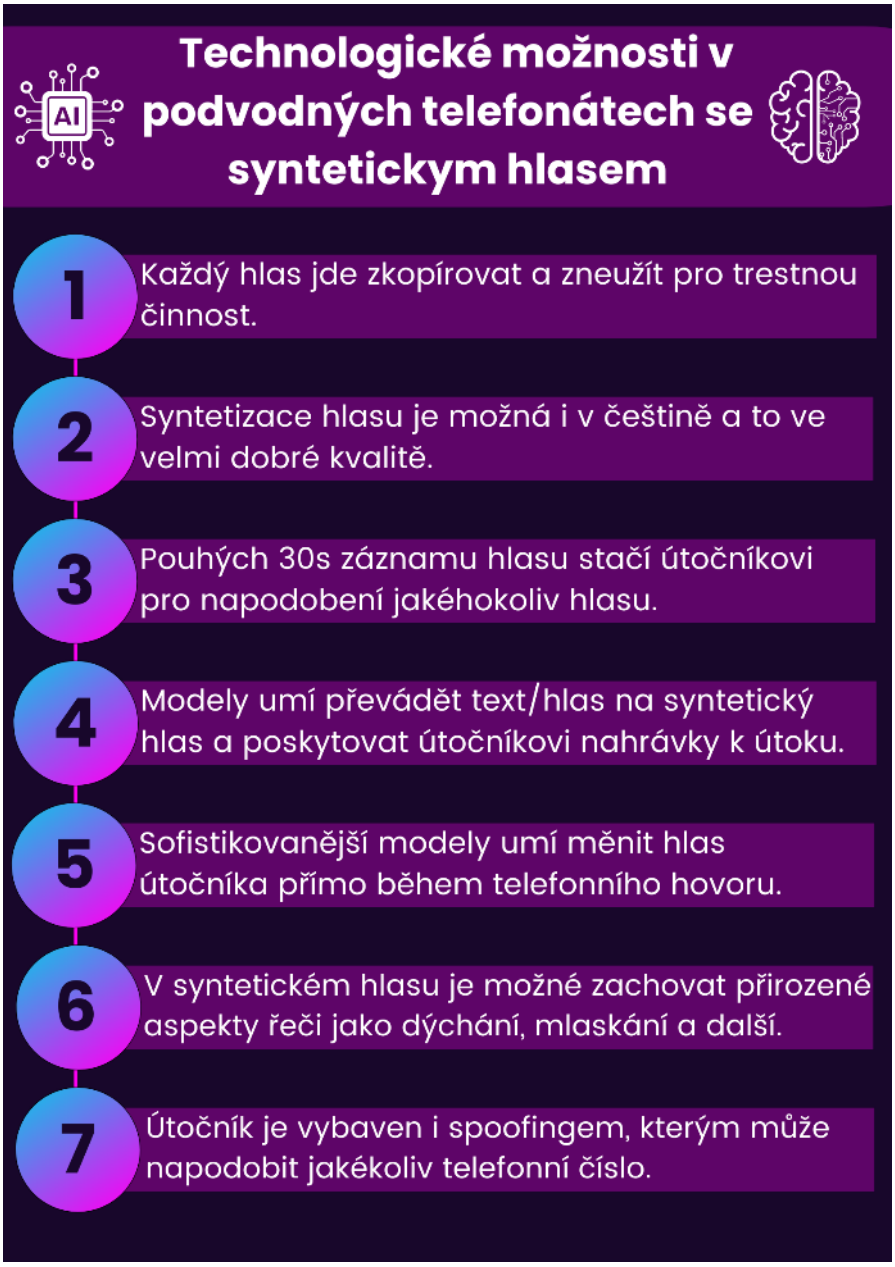
zcela odlišný. Tímto způsobem ztěžuje kriminálníkům využití hlasových dat k syntéze hlasů a k jejich napodobení. Dále existují nástroje pro detekci syntetického hlasu, které skrze behaviorální analýzy, detekce tzv. živosti hlasu atd. dokážou odhalit syntetizovaný hlas. Mezi takové technologie patří AI Voice Detector, ElevenLabs nebo deepfakedetector. [64][65][66]

6.4 Způsoby informovanosti

Jako nedílnou součástí boje proti podvodům je shledána informovanost, kterou vyžadovali i všichni respondenti, co okusili hrozbu na vlastní kůži. Informovanost v podobě školení, informačních materiálů, výstrah od Policie ČR, telekomunikačních společností atd. může mít opravdu velký vliv na snížení počtu podvedených v budoucnosti.

V rámci informovanosti o útoku by se mohla využít stávající práce Policie ČR. Ta na svých sociálních sítích aktivně informuje o podvodech, které společnost aktuálně sužují. Jedná se o velice jednoduchý způsob, jak informovat desítky tisíc lidí v krátkém čase. Jak bylo zmíněno, Policie ČR informuje pouze o aktuálních podvodech, které za sebou nechaly již větší množství obětí. To je problém, protože se nijak nebrzdí začínající hrozba, ale již plně rozjetý podnik s podvody, který si útočníci v případě potřeby upraví. Policie by tedy mohla směrovat informovanost více preventivně a sledovat trendy spjaté s podvody v zemích jako USA. Nejzávažnější hrozby by mohla reflektovat v rámci sociálních sítí i se základními protiopatřeními.

Pro informování obyvatelstva byly vytvořeny dva informační materiály v podobě jednoduché infografiky s textem. Dle odpovědí respondentů z rozhovoru byly vyžadovány převážně informace o tom, co útočník s UI dokáže a jak se bránit proti podvodným telefonátům se syntetickým hlasem. První infografika by měla být představena především starším lidem, aby měli základní povědomí o tom, jaké technologické možnosti má útočník vyzbrojen s UI a spoofingem (Obrázek 20).



Obrázek 20: Technologické možnosti.

Druhá infografika obsahuje ty nejdůležitější kroky, které dokážou efektivně odrazit útok se syntetickým hlasem (Obrázek 21). Oba materiály je možné do budoucna spojit s informacemi o dalších hrozbách poháněných UI a vytvořit ucelený a komplexní soubor informací a opatření. Tento soubor prezentovat široké veřejnosti a nastínit jim jaké hrozby v sobě UI skrývá. Cesty k oběma infografikám jsou uvedeny v **PŘÍLOZE P VII**.

Obrázek 21: Základní kroky obrany.

6.5 Shrnutí

Hrozba byla důkladně zanalyzována získáním odpovědí od 180 respondentů z ankety a z praktických zkušeností při tvorbě a provedení sociálního experimentu. Na základě získaných poznatků byla vytvořena opatření, čímž bylo dosaženo hlavního výzkumného cíle praktické části a tím byla zkompletována celá diplomová práce.

V poslední kapitole byl vytvořen seznam opatření, která mohou pomoci běžnému člověku, ale i organizacím být odolnější vůči podvodným telefonátům se syntetickým hlasem. Byly stanoveny základní pravidla obrany a z nich vybráno pět nejdůležitějších pravidel pro vytvoření informačního materiálu. Dále byla zmíněna technologická opatření jako využití dvou SIM karet, eSIM karty, aplikace Cloaked atd. Na závěr kapitoly byly uvedeny způsoby informování o hrozbě a vytvoření informačního materiálu v podobě dvou plakátů.

ZÁVĚR

Diplomová práce se zabývá využitím umělé inteligence v oblasti sociálního inženýrství. Cílem diplomové práce je analyzovat nově vzniklou hrozbu v podobě podvodných telefonátů se syntetickým hlasem a na základě zjištění z analýzy vytvořit protiopatření.

V první kapitole teoretické části byly popsány základní pojmy spjaté se SI, vliv psychologie v oblasti SI, model útoku SI a techniky SI, které je aktuálně možné podpořit UI. Zajímavé bylo, že UI nejenže vylepšuje stávající techniky, ale vytváří úplně nové hrozby jako Indirect Prompt Injection. V druhé kapitole teoretické části byly vymezeny základní pojmy UI a na závěr byly popsány hojně používané modely UI pro generování různého obsahu.

První kapitola praktické části byla věnována metodice výzkumu a návrhu sociálního experimentu. Účelem této kapitoly bylo představení hlavního cíle práce, dílčích cílů, metod a předvést nástroje a návrh sociálního experimentu.

Analyzováním ankety v další kapitole byl naplněn první dílčí cíl praktické části. Ze zkoumaného souboru vyplynulo, že o hrozbě se syntetickým hlasem je informována převážně mladší populace pod 35 let, což bylo ověřeno i v první zkoumané hypotéze. Další pohled na problematiku poskytl test, kterým se posuzovala schopnost rozpoznat reálný a syntetický hlas. Z průměrných až podprůměrných výsledků testu bylo patrné, že rozpoznání syntetického hlasu dělá respondentům potíže. V rámci ověřování druhé hypotézy nebyla zjištěna žádná významná souvislost mezi dosaženým vzděláním a schopností rozpoznat reálný a syntetický hlas. Pozitivní bylo, že respondenti nepřeceňovali své schopnosti a byli si vědomi, že jim činí rozpoznávání velké obtíže. Kvůli průměrným výsledkům bylo usouzeno, že v rámci obrany není nutné se snažit rozpoznat syntetický hlas. Hrozba vishingu se syntetickým hlasem byla vnímána jako nebezpečnější než běžný vishing. V kontrastu s touto informací bylo zjištěno, že většina respondentů nemá tušení, jak se proti této hrozbě bránit. Zbylí respondenti uváděli spíše základní kroky, ale nic komplexního. Poslední hypotéza se zaměřovala na opatření proti vishingu a na to, zda existují rozdíly ve znalostech těchto opatření v závislosti na pohlaví. Výsledky výzkumu tuto hypotézu nepotvrdily a znalosti respondentů se tedy neliší v závislosti na pohlaví.

Provedením sociálního experimentu v páté kapitole byl splněn i druhý dílčí cíl praktické části. Na základě zjištění, kterých bylo dosaženo během provádění experimentu a z rozhovorů s respondenty, je možné tvrdit, že podvody se syntetickým hlasem jsou realizovatelné na území České republiky. Nejdůležitější zjištění v rámci rozhovorů bylo, že 4 z 5

respondentů nepoznali, že se jedná o syntetický hlas. Pochybnosti, které zazněly nejčastěji, byly monotónnost v hlase, špatná návaznost na odpovědi, cizí číslo, vlastnictví platebního terminálu policií a u scénáře č.3 byla podezřelá velká akutnost požadavku. Překvapivé bylo, že pro úspěch experimentu nestačilo pouze mluvit se syntetickým hlasem, ale i dobře zvolená záminka měla zásadní dopad na výsledek. Respondenti do budoucna hodlají být více obezřetní a podivné požadavky si budou ověřovat. Od respondentů byl v poslední otázce projeven velký zájem o informovanost, především je zajímala opatření proti hrozbě nebo co všechno UI dokáže.

Poslední kapitolou byla zkompletována celá diplomová práce a naplněn hlavní cíl práce. Na základě zjištěních informací a dalších zdrojů byl vytvořen seznam opatření, která mohou pomoci v obraně proti podvodným telefonátům se syntetickým hlasem. Byly stanoveny základní pravidla obrany a z těch vybráno pět nejdůležitějších pravidel. Dále byla zmíněna technologická opatření a na závěr kapitoly byly uvedeny způsoby informování o hrozbě a vytvoření informačního materiálu v podobě dvou plakátů.

V rámci zlepšení poznatků by bylo vhodné rozšířit daný výzkum do budoucna o zjištěné limity z ankety a sociálního experimentu. Především by se mohl v dalším výzkumu využít a prozkoumat převod hlasu v reálném čase, který je den ode dne propracovanější a tím i nebezpečnější. Výzkum by se mohl obohatit i o jiné techniky SI využívající UI a na základě zjištění tvořit komplexní protiopatření proti tomuto nebezpečí.

SEZNAM POUŽITÉ LITERATURY

[1] SJOUWERMAN, Stu. *How AI Is Changing Social Engineering Forever*. Online. Forbes. 2023. Dostupné z: https://www.forbes.com/sites/forbestechcouncil/2023/05/26/how-ai-is-changing-social-engineering-forever/. [cit. 2024-05-05].

[2] PUIG, Alvaro. *Scammers use AI to enhance their family emergency schemes*. Online. Consumer Advice. 2023. Dostupné z: https://consumer.ftc.gov/consumer-alerts/2023/03/scammers-use-ai-enhance-their-family-emergency-schemes. [cit. 2024-05-04].

[3] *FBI Warns of Scammers Targeting Senior Citizens in Grandparent Scams and Demanding Funds by Wire, Mail, or Couriers*. Online. IC3. 2023. Dostupné z: https://www.ic3.gov/Media/Y2023/PSA231117. [cit. 2024-05-05].

[4] MOUTON, Francois; LEENEN, Louise; MALAN, Mercia M. a VENTER, H. S. Towards an Ontological Model Defining the Social Engineering Domain. Online. In: KIMPPA, Kai; WHITEHOUSE, Diane; KUUSELA, Tiina a PHAHLAMOHLAKA, Jackie (ed.). *ICT and Society*. IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, s. 266-279. ISBN 978-3-662-44207-4. Dostupné z: https://doi.org/10.1007/978-3-662-44208-1_22. [cit. 2023-11-19].

[5] *What is social engineering?* Online. IBM. Dostupné z: https://www.ibm.com/topics/social-engineering. [cit. 2024-04-29].

[6] *Sociální inženýrství v akci: příběhy a lekce z fyzických penetračních testů*. Online. NGSS. Dostupné z: https://www.ngss.cz/clanek/socialni-inzenyrstvi-v-akci-pribehy-a-lekce-z-fyzickych-penetracnich-testu-2023-08-16. [cit. 2024-04-29].

[7] MITNICK, Kevin D. a SIMON, William L. *Umění klamu*. Gliwice: Helion, 2003. ISBN 83-7361-210-6.

[8] *Sociální inženýrství*. Online, Bakalářská práce. Praha: Ambis Vysoká škola, a.s., 2021. Dostupné z: https://is.ambis.cz/th/n7zmu/BP_ZivnaJana.pdf. [cit. 2024-04-29].

[9] *Meaning of social engineer in English*. Online. English meaning - Cambridge Dictionary. Dostupné z: https://dictionary.cambridge.org/dictionary/english/social-engineer. [cit. 2024-05-04].

[10] OZKAYA, Erdal. *Learn Social Engineering*. Packt Publishing, 2018. ISBN 978-1-78883-792-7.

[11] *Internet crime report 2023.* Online. Ic3. 2023. Dostupné z: https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf?trk=public_post_comment-text [cit. 2023-11-19].

[12] *2023 Data Breach Investigations Report (DBIR*). Online. Verizon. 2023. Dostupné z: https://www.verizon.com/business/resources/T253/reports/2023-data-breach-investigations-report-dbir.pdf. [cit. 2023-11-19].

[13] KAHNEMAN, Daniel. *Myšlení: rychlé a pomalé*. Pod povrchem. V Brně: Jan Melvil, 2012. ISBN 978–80–87270–42–4.

[14] UEBELACKER, Sven a QUIEL, Susanne. The Social Engineering Personality Framework. Online. In: *2014 Workshop on Socio-Technical Aspects in Security and Trust.* Vídeň: IEEE, 2014, s. 24-30. ISBN 978-1-4799-7901-1. ISSN 2325-1697. Dostupné z: https://doi.org/10.1109/STAST.2014.12. [cit. 2023-10-18].

[15] CIALDINI, Robert B. *Zbraně vlivu: manipulativní techniky a jak se jim bránit*. Žádná velká věda. V Brně: Jan Melvil, 2012. ISBN 978-80-87270-32-5.

[16] HADNAGY, Christopher. *Social engineering: The Art of Human Hacking*. 2. Vyd. Wiley Publishing, 2018. ISBN 978-1119433385.

[17] *PĚTIFAKTOROVÝ MODEL OSOBNOSTI, PROSOCIÁLNÍ CHOVÁNÍ A EMPATIE U STUDENTŮ STŘEDNÍCH A VYSOKÝCH ŠKOL*. Online, Výzkumná studie. Ostrava: Ostravská Univerzita v Ostravě, 210n. l. Dostupné z: https://psychkont.osu.cz/fulltext/2010/Mlcak_2010_2.pdf. [cit. 2023-11-19].

[18] MOUTON, Francois; MALAN, Mercia M.; LEENEN, Louise a VENTER, H.S. Social engineering attack framework. Online. In: *2014 Information Security for South Africa*. Johannesburg: IEEE, 2014, s. 1-9. ISBN 978-1-4799-3384-6. ISSN 2330-9881. Dostupné z: https://doi.org/10.1109/ISSA.2014.6950510. [cit. 2023-10-18].

[19] SALAHDINE, Fatima a KAABOUCH, Naima. Social Engineering Attacks: A Survey. Online. *Future Internet*. 2019, roč. 11, č. 4, s. 17. ISSN 1999-5903. Dostupné z: https://doi.org/10.3390/fi11040089. [cit. 2023-10-18].

[20] The Intersection of Artificial Intelligence and Social Engineering: Next-Generation Threats. Online. Versprite. 2023. Dostupné z: https://versprite.com/blog/the-

intersection-of-artificial-intelligence-and-social-engineering-next-generation-threats/. [cit. 2023-11-19].

[21] BERNSTEIN, Jeffrey. *AI increasing sophistication of social engineering attacks*. Online. Kaufman Rossin Multisite Website. 2023. Dostupné z: https://kaufmanrossin.com/blog/ai-increasing-sophistication-of-social-engineering-attacks/. [cit. 2024-05-04].

[22] NOVAK, Chris. *The Role Of AI In Social Engineering*. Online. Forbes. 2023. Dostupné z: https://www.forbes.com/sites/forbestechcouncil/2023/07/05/the-role-of-ai-in-social-engineering/. [cit. 2024-05-04].

[23] *11 Types of Phishing Tips to Prevent Phishing Attacks*. Online. Panda Security Mediacenter. 2024. Dostupné z: https://www.pandasecurity.com/en/mediacenter/types-of-phishing/. [cit. 2024-05-04].

[24] PATEL, Vinay. *Cybercriminals using AI chatbots like ChatGPT to craft credible-looking phishing emails*. Online. International Business Times UK. 2023. Dostupné z: https://www.ibtimes.co.uk/cybercriminals-using-ai-chatbots-like-chatgpt-craft-credible-looking-phishing-emails-1714721. [cit. 2024-04-29].

[25] *Indirect Prompt Injection Threats*. Online. GitHub. Dostupné z: https://greshake.github.io/. [cit. 2024-05-04].

[26] KRISHNAN, Ashwin. *Generative AI is making phishing attacks more dangerous*. Online. TechTarget Security. 2023. Dostupné z: https://www.techtarget.com/searchsecurity/tip/Generative-AI-is-making-phishing-attacks-more-dangerous. [cit. 2024-05-04].

[27] MCFARLANE, Lydia. *AI driving more sophisticated scams, tech scholars tell lawmakers* . Online. The Hill. 2023. Dostupné z: https://thehill.com/policy/technology/4313640-ai-sophisticated-scams-tech-scholars-lawmakers/. [cit. 2024-05-04].

[28] MRÓZ, Mateusz. *Co je to deepfake, jak ho vytvořit a zda je nebezpečný?* Online. Botland. 2023. Dostupné z: https://botland.cz/blog/co-je-to-deepfake-jak-ho-vytvorit-a-zda-je-nebezpecny/. [cit. 2024-05-04].

[29] BARNEY, Nick. *Deepfake AI (deep fake)*. Online. TechTarget WhatIs. 2022. Dostupné z: https://www.techtarget.com/whatis/definition/deepfake. [cit. 2024-05-04].

[30] ADEE, SALLY. *What Are Deepfakes and How Are They Created?* Online. IEEE Spectrum. 2020. Dostupné z: https://spectrum.ieee.org/what-is-deepfake. [cit. 2024-05-04].

[31] MAŇÁKOVÁ, Magdaléna a KASÍK, Pavel. *Podvodníci vás zkopírují jako loutku. Klon zavolá babičce nebo do banky*. Online. Seznam Zprávy. 2023. Dostupné z: https://www.seznamzpravy.cz/clanek/tech-ai-umela-inteligence-tady-sef-zlodeji-kopiruji-lidi-jako-loutky-klon-zavola-babicce-i-do-banky-235819. [cit. 2024-05-04].

[32] KLINGLER, Nico. *The Ultimate Guide to Understanding and Using AI Models (2024)*. Online. Viso.ai. 2024. Dostupné z: https://viso.ai/deep-learning/ml-ai-models/. [cit. 2024-05-04].

[33] *What is artificial intelligence (AI)?* Online. IBM. 2023. Dostupné z: https://www.ibm.com/topics/artificial-intelligence. [cit. 2024-05-04].

[34] S. GILLIS, Alexander. *4 main types of artificial intelligence: Explained*. Online. TechTarget Enterprise AI. 2023. Dostupné z: https://www.techtarget.com/searchenterpriseai/tip/4-main-types-of-AI-explained. [cit. 2024-05-04].

[35] SHALAMANOV, Jennifer. *The Different Types of AI: A Quick Overview*. Online. Udacity. 2021. Dostupné z: https://www.udacity.com/blog/2021/06/the-different-types-of-ai-a-quick-overview.html. [cit. 2024-05-04].

[36] *Understanding the different types of artificial intelligence*. Online. IBM Blog. 2023. Dostupné z: https://www.ibm.com/blog/understanding-the-different-types-of-artificial-intelligence/ [cit. 2024-05-04].

[37] *Types of Artificial intelligence - What are the 7 types of AI?* Online. Atlearner: Learn Science. 2019. Dostupné z: https://www.atlearner.com/2019/10/types-of-artificial-intelligence.html. [cit. 2024-05-04].

[38] SARKER, Iqbal H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. Online. *SN Computer Science*. 2021, roč. 2, č. 6. ISSN 2662-995X. Dostupné z: https://doi.org/10.1007/s42979-021-00815-1. [cit. 2023-10-19].

[39] *What is machine learning (ML)?* Online. IBM. Dostupné z: https://www.ibm.com/topics/machine-learning. [cit. 2024-05-04].

[40] ANTONIADIS, Panagiotis. *Hidden Layers in a Neural Network*. Online. Baeldung on Computer Science. 2024. Dostupné z: https://www.baeldung.com/cs/hidden-layers-neural-network. [cit. 2024-05-04].

[41] *AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the difference?* Online. IBM Blog. 2023. Dostupné z: www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks. [cit. 2024-05-04].

[42] *Deep Learning vs. Machine Learning: A Beginner's Guide*. Online. Coursera. 2024. Dostupné z: https://www.coursera.org/articles/ai-vs-deep-learning-vs-machine-learning-beginners-guide. [cit. 2024-05-04].

[43] *Effortlessly Analyze Your Competitive Landscape*. Online. Similarweb. 2024. Dostupné z: https://www.similarweb.com/. [cit. 2024-05-10].

[44] *GPT-3.5, GPT-4: Zjistěte, jaký je mezi nimi rozdíl*. Online. Talkai.info. 2023. Dostupné z: https://talkai.info/cs/blog/gpt3_gpt4_difference/. [cit. 2024-05-04].

[45] *Get answers. Find inspiration. Be more productive*. Online. OpenAI. Dostupné z: https://openai.com/chatgpt. [cit. 2024-05-04].

[46] MITTAL, Aayush. *Google's Multimodal AI Gemini – A Technical Deep Dive*. Online. Unite.AI. 2023. Dostupné z: https://www.unite.ai/googles-multimodal-ai-gemini-a-technical-deep-dive/. [cit. 2024-05-04].

[47] *Learn more about Gemini, our most capable AI model*. Online. Blog.google. 2023. Dostupné z: https://blog.google/technology/ai/gemini-collection/. [cit. 2024-05-04].

[48] *Gemini*. Online. Gemini. Dostupné z: https://gemini.google.com. [cit. 2024-05-05].

[49] KAPLER, Tomáš. *DALL-E 3 vrací úder – kvalitní generování obrázků jednoduchým textovým zadáním*. Online. Kapler o AI. 2023. Dostupné z: https://www.kapler.cz/predstaveni-opeai-dall-e-3/. [cit. 2024-05-05].

[50] *Change Your Voice With Speech To Speech*. Online. ElevenLabs. Dostupné z: https://elevenlabs.io. [cit. 2024-05-05].

[51] *Free Real-Time Voice Changer*. Online. Free Real Time Voice Changer and Modulator - Voicemod. 2024. Dostupné z: https://www.voicemod.net/. [cit. 2024-05-10].

[52] *Anketa*. Online. Sociologická encyklopedie. Dostupné z: https://encyklopedie.soc.cas.cz/w/Anketa. [cit. 2024-05-04].

[53] ANDĚL, Jiří. *Základy matematické statistiky*. 2., opr. vyd. Praha: Matfyzpress, 2007. ISBN 80-7378-001-1.

[54] MIOVSKÝ, Michal. *Kvalitativní přístup a metody v psychologickém výzkumu*. Psyché (Grada). Praha: Grada, 2006. ISBN 978-80-247-1362-5.

[55] *Vishing a spoofing*. Online. Policie České republiky. 2021. Dostupné z: https://www.policie.cz/clanek/vishing-a-spoofing.aspx. [cit. 2024-05-05].

[56] HRON, Lukáš. *Podvrhnout vaše telefonní číslo není složité. Ihned se obraťte na operátora*. Online. IDNES.cz. 2022. Dostupné z: https://www.idnes.cz/mobil/mobilni-operatori/podvod-zahranici-hovor-spoofing-telefonni-cislo-mobilni-operator.A221130_120608_mobilni-operatori_LHR. [cit. 2024-05-05].

[57] GREGOIRE, Courtney. *Increased uptake of generative AI technology brings excitement and highlights the importance of family conversations about online safety, says new research from Microsoft*. Online. Microsoft On the Issues. 2024. Dostupné z: https://blogs.microsoft.com/on-the-issues/2024/02/05/generative-ai-online-safety-day-global-survey/. [cit. 2024-05-05].

[58] *Men lose more than twice as much money to scammers than women, new research reveals*. Online. Phoenix Group. 2021. Dostupné z: https://www.thephoenixgroup.com/news-views/men-lose-more-than-twice-as-much-money-to-scammers-than-women-new-research-reveals/. [cit. 2024-05-05].

[59] WEITZMAN, Cliff. *How to protect yourself from AI voice scams*. Online. Speechify. 2023. Dostupné z: https://speechify.com/blog/protect-yourself-from-ai-voice-scams. [cit. 2024-05-16].

[60] ROGERS, Reece. *Get WIRED*. Online. WIRED. 2024. Dostupné z: https://www.wired.com/story/how-to-protect-yourself-ai-scam-calls-detect/. [cit. 2024-05-16].

[61] *Používání dvou SIM karet s eSIM*. Online. Apple Support. 2024. Dostupné z: https://support.apple.com/cs-cz/109317. [cit. 2024-05-05].

[62] *Identifikace a hodnocení neznámých telefonních čísel*. Online. Identifikace a hodnocení neznámých telefonních čísel. Dostupné z: https://www.callinsider.cz/. [cit. 2024-05-05].

[63] *Cloaked protects your personal identity from everyone*. Online. Cloaked. 2023. Dostupné z: https://www.cloaked.com/. [cit. 2024-05-05].

[64] *Defending Your Voice Against Deepfakes*. Online. Technology Org. 2023. Dostupné z: https://www.technology.org/2023/11/28/defending-your-voice-vs-deepfakes/. [cit. 2024-05-18].

[65] *Elevating voice security with advanced liveness detection*. Online. Pindrop. 2024. Dostupné z: https://www.pindrop.com/. [cit. 2024-05-18].

[66] *Top 49 AI Voice Detector Alternatives*. Online. TopAI.tools. 2024. Dostupné z: https://topai.tools/alternatives/ai-voice-detector. [cit. 2024-05-18].

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

HU Hluboké učení

R1 Respondent č 1

R2 Respondent č 2

R3 Respondent č 3

R4 Respondent č 4

R5 Respondent č 5

SI Sociální inženýrství.

SU Strojové učení

UI Umělá inteligence.

UNS Umělé neuronové sítě

VoIP Voice over Internet Protocol

SEZNAM OBRÁZKŮ

SEZNAM TABULEK

SEZNAM GRAFŮ

SEZNAM PŘÍLOH

PŘÍLOHA P I: OTÁZKY V DOTAZNÍKU

1. Jaké jste pohlaví?

- Muž
- Žena

2. V jakém věkovém rozmezí se nacházíte?

- 8-18 let
- 19-25 let
- 26-35 let
- 36-45 let
- 46-55 let
- 55 let a více

3. Jaké máte nejvyšší dosažené vzdělání?

- Základní
- Středoškolské s výučním listem
- Středoškolské s maturitou
- Vysokoškolské

4. Setkali jste se někdy osobně s pojmem vishing, kdy se útočník v telefonním hovoru vydává třeba za policistu nebo jinou osobu a snaží se z Vás získat finanční prostředky nebo jiné údaje?

- Ano
- Ne

5. Věděli jste před vyplněním dotazníku, že útočník může za pomoci umělé inteligence zkopírovat Váš hlas a vytvořit jeho syntetickou (umělou) kopii a podpořit tak věrohodnost vishingových útoků?

- Ano
- Ne

6. Tato otázka se zaměřuje na Vaši schopnost rozeznat syntetický hlas od reálného hlasu. Níže máte video, které obsahuje 8 záznamů hlasu. Pusťte si jej a pokuste se určit, zda je hlas reálný či vytvořený umělou inteligencí. (Přiložené video)

	Reálný hlas	Syntetický hlas
Hlas 1	○	○
Hlas 2	○	○
Hlas 3	○	○
Hlas 4	○	○
Hlas 5	○	○
Hlas 6	○	○
Hlas 7	○	○
Hlas 8	○	○

7. Myslíte si, že pro Vás bylo obtížné rozeznat syntetický hlas od reálného?

- Ano, měl jsem velké obtíže
- Spíše ano
- Spíše ne
- Ne, neměl jsem vůbec žádné obtíže

8. Pokud by po Vás útočník vyžadoval citlivé údaje (telefoní číslo, údaje na platební kartě) nebo finanční prostředky, jak silnou důvěru by u Vás vzbudil s hlasem uvedeným v možnostech?

	Žádná důvěra a silné podezření	Důvěra, ale slabé podezření	Silná důvěra bez podezření
Rodinní příslušníci	○	○	○
Dlouhodobí přátelé	○	○	○
Známí lidé (kolegové v práci, přátelé atd., se kterými občas promluvíte)	○	○	○
Vedoucí, šéf (útočník hlasem vedoucího vyžaduje zaplatit fakturu nebo poskytnout citlivé údaje)	○	○	○

9. Vyberte možnosti, které byste byl/a ochoten/na poskytnout podvodníkovi, který se vydává za osobu Vám nejbližší.

- Finanční prostředky
- Údaje z platební karty
- Přihlašovací údaje do sociálních sítí
- Přihlašovací údaje do e-mailu
- Přihlašovací údaje do internetového bankovnictví
- Osobní údaje rodinných příslušníků, kolegů nebo přátel (telefon, jména atd.)

10. Jak velkou finanční částku byste poskytl/a?

- *„Otevřená odpověď“*

11. Myslíte si, že vishing se syntetickým hlasem blízké osoby je nebezpečnější než běžný vishing?

- Ano, je nebezpečnější
- Ne, není nebezpečnější
- Stejně nebezpečný

12. Máte povědomí o tom, jak se bránit proti podvodným telefonátům využívající syntetický hlas?

(Pokud neznáte odpověď, napište pouze "Ne")

- *„Otevřená odpověď“*

PŘÍLOHA P II: ROZHOVOR S RESPONDENTEM Č.1

R1 je zkratka pro respondenta č.1

Autor: *Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?*

R1: *Určitě mě šokovalo to, že jsem nedokázala rozpoznat, že na druhé straně nejsi ty, ale je tam nějak uměle nahraný hlas. Na druhou stranu se jednalo o dvě stovky, takže to nebylo tak jakože bych si hrozně sypala popel na hlavu, ale brala jsem to tak, že jsem v ten moment chtěla pomoct kamarádovi. A že jsem udělala dobrý skutek, i když to byl podvod. No a samozřejmě jsem přemýšlela, co bych mohla víc udělat, jak bych se tomu mohla bránit. Jenomže vzhledem k tomu, jak ten hlas je fakt jako realistický, tak asi jako rozpoznání toho hlasu tady nepřipadá moc v úvahu. Takže spíš taková ta… tak než pocity, tak spíš takový podnět k tomu se zamyslet, jak tomu můžu nejenom u sebe, ale i třeba v mém okolí zabránit.*

Autor: *Překvapilo tě, že si se s tímto setkala v České republice?*

R1: *Upřímně překvapilo, protože jsem vždycky měla za to, že máme umělou inteligenci takovou jako zastaralou. Nečekala bych, že to dokáže až takhle napodobit něčí hlas. A sama, když jsem si umělou inteligencí pomáhala v rámci textu, tak jsem to brala, že to je taková… Já jsem tomu říkala zkušební verze, kdy jsme měli zaplacenou tu nejvyšší verzi. Sama jsem musela hodně často opravovat i faktické chyby. Takže mě to rozhodně překvapilo. Brala jsem to tak, že asi jako pokud bude něco až takhle úplně vyspělé, tak to bude furt ještě v zahraničí, a ne u nás. U nás bych to asi neočekávala.*

Autor: *Co ti na rozhovoru přišlo podezřelé?*

R1: *Když jsem se nad tím pak zpětně, nebo i teď, když se nad tím zpětně zamýšlím, tak přímo na rozhovoru asi takhle na té struktuře a na tom hlasu nic. Co mě jako překvapilo, že voláš z cizího čísla, ale to mohl být vybitý telefon, nemusel si mít kredit, cokoliv. Co mě teda ještě jako překvapilo nebo překvapuje teď, že vlastně policie by nechtěla hotovost na místě, že vím, že tam bývala nějaká možnost zaplacení pokuty na místě nebo se mohlo přijít na stanici. Samozřejmě, jak to vidíme i na umělé inteligenci, tak se doba vyvíjí a můžou mít u sebe terminály. Takže asi jenom to číslo a Policie.*

Autor: *A hlas ti nepřišel podezřelý ohledně třeba monotónnosti, intonace nebo nějaké emocionální nestability, že jsem jako mluvil bez emocí, víceméně?*

R1: *Jo, jako co tam bylo takové divné… Ne, že divné, ale vypadával si občas během našeho telefonátu, ale já mám doma problém se signálem, kdy máme horší pokrytí pro moji telefonní*

síť. Takže jsem to brala tak, že tím, že nevolám z venku, kde mám stoprocentní jistotu, že to málokdy vypadne, tak jsem to přikládala k tomu, že jsem prostě doma a může to vypadávat. Mám fakt s tím zkušenosti takové, takže nad tím jsem se vůbec upřímně nezamýšlela. Monotónnost. No brala jsem to tak, že v momentě, kdy člověka chytnou policajti a řeší takto nepříjemnou věc, tak je to takové, že rozhodně nebude člověk vysmátý nebo podobně. Nebo třeba bych do tebe nějak rýpla během telefonátu, tak asi se nebudeš moct chtít smát, když tam řešíš nějaký problém s policií. Takže jsem to přikládala k tomu, že vlastně si s tou policii a v tenhle moment fakt jako asi za mě člověk reaguje jinak než v běžném životě, když je úplně uvolněný a nic neřeší.

Autor: *Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?*

R1: *No jako hodně negativní v tom smyslu, že jsem to nedokázala rozpoznat. Na druhou stranu, pokud dám příklad, kdy mi došla smska, že česká pošta mi nemůže doručit balíček, protože nebylo něco zaplaceno a mám přes odkaz zaplatit nějakou částku, tak tam jsem na první dobrou poznala, že se jedná o podvod. Samozřejmě i díky tomu, že jsem v té době nečekala vůbec žádný balíček a věděla jsem, že mi nemá co přijít. Takže tam to bylo spíše intuice, že nic nečekám. Nicméně u těch telefonních podvodů, i když jsou to třeba různé sbírky a podobně, tak to odbývám a tam předpokládám, že je to podvod. No, ale v tomto případě je to spíš takové těžší v tom, že všichni chceme nějak vždycky pomoct kamarádům, tak asi nejsme hned paranoidní. Tady je to už potom opravdu náročnější kor, když ta umělá inteligence je až tak vyspělá, takže v tomhle to byla určitě negativní zkušenost.*

Autor: *Staneš se tedy do budoucna třeba více paranoidní vůči takovýmto telefonátům?*

R1: *Paranoidní… Já osobně si myslím, že občas jsem paranoidní vůči tomu, že když třeba neznám eshop, tak třeba neplatím kreditkou u nich a zaplatím radši dobírku. Ale tady je to, pokud zavolá kamarád, tak vždycky budou chtít pomoc. Maximálně, co mě tak napadá, tak určitě si to příště víc ověřím. Jenomže druhá věc je to, že pokud člověk stojí s policisty, tak není moc času na ověřování. Takže. Asi upřímně si o tom budu muset více nastudovat.*

Autor: *Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti?*

R1: *Upřímně by se hlavně společnost měla zaměřit na takové ty hodně zranitelné skupiny, což vnímám jako děti okolo 12 - 13 let a poté seniory. Za mě obě tyhle skupiny jsou hodně náchylné a stačí například, že se za nás může někdo vydávat a zavolat našim prarodičům, že se stalo to a to a nutně potřebuje poslat peníze, tak věřím, že každý prarodič bude chtít*

pomoct a zase u té menší skupiny u těch školáčků vnímám, že stačí, aby jim je nevím, nabídli různé takové ty coiny do her, nebo jak se to jmenuje a oni ty finance pošlou, nebo pokud budou hrát nějaké online hry, tak to budou posílat za nějaké takové ty výhody. Takže určitě u té mladší skupiny by se měla postarat škola o nějakou takovou finanční gramotnost a měli by mít pravidelně přednášky a učit je nejen jak s financemi hospodařit, ale právě jak se vyhnout i těmhle podvodným telefonátům a dalším hrozbám. Pro seniory určitě by měly být různé přednášky v televizi nebo v různých domovech důchodců, nebo jak jsou takové ty kluby seniorů. Furt je vzdělávat o tom, co už dokáže umělá inteligence, aby na to byli připraveni. No a taková ta všeobecná informovanost, protože upřímně jako mám pocit, že hodně sleduju dění ohledně umělé inteligence. I v rámci práce řešíme umělou inteligenci, ale že bychom někdy řešili takovéto syntetické hlasy, to ne. Takže si myslím, že bychom měli mít více jako informovanost celkově v médiích v televizi kdekoliv.

Autor: *Napadá tě k tomuto ještě něco, co bys chtěla dodat?*

R1: *Asi doufám, jakože takovým podvodům se delší dobu vyhnu. Rozhodně na to upozorním své okolí, no a taky mě zajímá, jak se to bude dále vyvíjet.*

Autor: *No mě ještě tak napadlo v rámci té informovanosti, co by tě tam nejvíce zajímalo? Třeba nejnovější hrozby nebo nějaká opatření?*

R1: *Když to řeknu špatně, tak u nás vnímám, že k nám všechny, ať už je to v rámci IT, v rámci marketingu, v rámci umělé inteligence, že k nám to jde až jako nakonec takové ty vymoženosti. Tak, rozhodně bych chtěla více v médiích, to co dokáže uměla inteligence, protože pochybuju, že třeba moje mamka, když to řeknu hnusně rozpozná fotku, která je vytvořená umělou inteligencí a která je reálný obraz. Takže určitě by měly být za mě nějaké seriály na internetu nebo tak s ukázkami, aby přímo ukazovali na co se mají lidé zaměřit v rámci těch podvodů.*

PŘÍLOHA P III: ROZHOVOR S RESPONDENTEM Č.2

R2 je zkratka pro respondenta č.2

Autor: *Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?*

R2: *Tak, byl jsem překvapený, upřímně, že podobné věci už se dějí i u nás v Česku. Já teda jako z poslední doby znám třeba deepfake videa politiků z instagramu, kde se jedná spíše jenom o pobavení nebo jako vím, že to jsou kanály, které mají lidi pobavit. Takže je to nějaký jako vtipný kontent. Ale tohle mě jako překvapilo, no.*

Autor: *Ještě jsi se jako nikdy nesetkal třeba se syntetickým hlasem, že by si sám něco zkoušel?*

R2: *Ne, vůbec.*

Autor: *Takže jenom s deepfake videí, kde byli politici?*

R2: *Ano. Ono to bylo vlastně jako reálné video, někde třeba ze zasedání parlamentu a jenom jako změnili... Vlastně ano, přes tady ten syntetický hlas změnili, co říká aby to nedávalo prostě smysl.*

Autor: *Jak tě překvapilo, že už se dá syntetický hlas využít i pro tyto podvody?*

R2: *Jakože vím, že se tady tohle děje, ale ve výsledku, když se to stalo mně osobně, že mě takhle jako zavolá bratr, byla to jako rychlá zpráva poplašná, že se jako něco děje a potřebuje peníze, tak mě to jako zarazilo. Prostě je to jako rodina, něco se děje, takže jsem zpozorněl.*

Autor: *Co ti na tom rozhovoru přišlo podezřelé?*

R2: *Tak určitě to, že to bylo rychlé a akutní. Nebyly tam moc prostory pro otázky. Rozhovor potom i rychle skončil. A zároveň třeba i vím, když si volám s bratrem, jakým způsobem on komunikuje a že tam chyběly takové ty drobné věci, které já třeba znám od něho, když se spolu bavíme.*

Autor: *Ohledně toho hlasu, krom tempa, tak ten hlas ti přišel jako v pohodě? Přišel ti věrohodný nebo nevěrohodný?*

R2: *Tím, jak byl krátký ten rozhovor, tak mi přišel v pohodě, jakože v tom v prvním momentu bych ani neřekl, že to je syntetický hlas.*

Autor: *A třeba že jsem po tobě chtěl jako peníze, nepřišlo ti to divné?*

R2: *No jako přišlo mi to divné, přišlo mi divné ty dvě stovky.*

Autor: *Jakože to byla moc velká částka nebo moc malá?*

R2: *Že spíše jako malá, že bych čekal tady u těch podvodů, že to bude vyšší.*

Autor: *Takže částka tam nehrála moc roli a přišlo ti divné jen tempo?*

R2: *To tempo i vlastně jako, že po mně chceš peníze, protože se mi ještě takhle nestalo, aby po mě bratr chtěl peníze a už vůbec ne, aby je po mně chtěl jako tímhle způsobem. Nestalo se mi, aby mi zavolal rychle přes telefon a chce prostě ze mě jenom rychle dostat peníze, že nemá čas to řešit a kdesi cosi. No spíš se jako osobně potkáme, nebo tak.*

Autor: *Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?*

R2: *Tak myslím si že dobrou nebo jako že jsem zhodnotil situaci dobře, že není něco úplně v pořádku. A potom jsem i vlastně ještě jako zavolal tobě, abych si ověřil, že to jsi opravdu ty.*

Autor: *Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti? Popřípadě mi můžeš i sdělit nějaká opatření, které ty máš v rámci třeba podvodných telefonátů.*

R2: *Tak určitě by měla začít osvěta už třeba na základních školách v rámci nějakých výukových programů, protože je to něco, co jako v dnešní době příchodem AI bude mít větší a větší dopad na lidi. A ne vlastně jen ve škole, protože se to dotýká všech věkových skupin. Takže i různě na sociálních sítích by to mohlo kolovat nebo v televizi nějaké rozhovory s odborníky, co se tím zabývají nebo dělají výzkum v této oblasti.*

Autor: *V rámci těch školení, co by ti nejvíce pomohlo?*

R2: *Asi nějaký jakoby výcuc metod, které by mi pomohly tady ty podvodné telefony odhalit třeba jako, jak postupovat krok za krokem, když mám podezření, že se něco takového děje. Prostě takhle, aby to třeba jednoduchý člověk, který nemá takové zkušenosti s internetem, prostě starší generace. Tak, aby byli schopni si to sami nějakým způsobem ověřit jenom tím, že budou postupovat podle nějakého jednoduchého manuálu.*

Autor: *Dobře a co se týče jako tvých opatření, máš nějaké proti podvodným telefonátům?*

R2: *No, já mám nějakou smůlu na to, že mě prostě často volají operátoři s nějakýma nabídkama, takže já už ze zásady jako třeba neberu cizí čísla. Teďka to bylo výjimečně, protože jsem na stáži v zahraničí, takže jsem nevěděl, jestli se třeba něco neděje v práci. Jako v Praze*

kolegové, jestli něco nepotřebují nebo sousedi, jestli se něco nestalo na bytě. A jinak třeba prostě i spamové emaily rovnou mažu ani neotevírám ve schránce. A snažím se být jako všeobecně prostě obezřetný na tyhle ty věci.

Autor: *Takže jako vůbec nebereš cizí čísla, tak to jsem měl štěstí.*

R2: *Ale to bylo fakt tím, že jsem prostě v zahraničí. No že kdybych byl v Praze, tak to nejspíš ten telefon… že spíš jako čekám, když to je někdo, kdo mě fakt potřebuje zkontaktovat, tak čekám třeba až napíše SMS. Čekám až se nějak představí, a že je to tedy ten člověk, ať mu třeba zavolám, ale jinak ty cizí čísla spíš ignoruju.*

Autor: *Ale co kdyby ti třeba volalo stejné číslo, které mám já, protože existuje třeba spoofing, kterým si člověk může napodobit telefonní číslo.*

R2: *Tohle to je pro mě třeba novinka, že něco takového jde udělat, ale asi bych postupoval stejně, že tím, jak mi ten rozhovor přišel podezřelý tou rychlostí. Tak bych prostě šel jako do telefonu a zavolal na číslo, které mám uložené pod tebou a zase ti zavolal, jestli jako všechno v pohodě.*

Autor: *Jo, tak to si udělal správnou věc. To je nejlepší, co můžeš udělat jít právě do kontaktů a ověřit si to. Jestli tě ještě k tomuto tématu něco napadá, tak klidně můžeš říct jinak to můžeme ukončit.*

R2: *No, já myslím, že jsme tak nějak všechno řekl, co bylo třeba, že hlavní je prostě ta osvěta, nebo informovat, že se tady to děje.*

Autor: *Dobře, kdybych měl ještě nějaké otázky, tak bych se tě potom doptal.*

PŘÍLOHA P IV: ROZHOVOR S RESPONDENTEM Č.3

R3 je zkratka pro respondenta č.3

Autor: *Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?*

R3: *No, pocity jsem měl potom takové smíšené, že jsem tomu nejdřív nechtěl věřit, protože jsem takový důvěřivý, prostě důvěřivá osoba. Bylo by to určitě jiný, kdyby mi zavolal cizí člověk, tak to by mi přišlo jako podezřelejší, ale pokud ti zavolá člověk, kterého dobře znáš, a je to někdo blízký, tvůj kamarád, nebo prostě takhle… Tak dalo, by se říct, že mu vždycky chceš, nebo prostě máš takovou potřebu, mu vždycky pomoct si myslím. Teda aspoň já to tak mám nastavené. Tudíž mě to potom hodně překvapilo, že to vlastně nebylo pravý. Jooo, jestli tady už můžu zmínit i to, že možná troška toho podezření tam bylo v tom, že jsi mluvil jako velice zmatený člověk nebo to je další otázka?*

Autor: *To se potom dostaneme v další otázce, ale můžeš klidně mluvit. Nebudeš přerušován. Klidně to dokonči a pak se na to zeptám ještě jednou kdyžtak.*

R3: *Aha OK. Většinou totiž člověk, který je zastavený takhle Policii a už má jen pár bodů, tak se to snaží vyřešit co nejšetrněji a nejrychleji, aby o ty body nepřišel. A to jsem si potom uvědomil až později, když to vezmu na ty blokové pokuty… Tak si myslím, že to byla dobrá věc na vyzkoušení tohoto experimentu, ale potom jak jsem dotelefonoval, tak mi vlastně došlo, skrze to, že většina těch policejních hlídek u sebe terminály nemají. I když v dnešní době je to čím dál rozšířenější. No a na druhou stranu jsem si říkal, že se to dá všechno vyřešit na stanici, že ti dají blokovku a ty musíš do určité doby zaplatit. Tudíž mi přišlo trošku podezřelý, že to je potřeba hned.*

Autor: *Takže říkáš, že během toho útoku tě to vůbec nenapadlo? Nepřemýšlel jsi nad tím, prostě si chtěl jenom pomoci kamarádovi?*

R3: *Ano, přesně tak bych to řekl.*

Autor: *Po jak dlouhé době tě to napadlo?*

R3: *No ono mi to možná už trošku napadlo při tom rozhovoru, protože jsem ti tam dával potom ještě nějaké otázky, na které si nebyl schopný odpovědět. Což mi přišlo podezřelý, ale bral jsem to ve výsledku tak, že třeba ti dali policisté možnost zavolat jenom z jejich čísla, stály za tebou a ty si nemohl být v rozhovoru osobnější.*

Autor: *Takže říkáš, že mě trošku zachránila i ta situace, kterou jsem uměle vytvořil?*

R3: *To si myslím, že asi jo.*

R3: *Ještě dodám jednu věc, že pokud by si ten rozhovor koncipoval tak, že potřebuješ třeba půjčit peníze na rychlo a bylo by to řádově v tisících, tak už já třeba osobně bych to neposlal. Neposlal bych to, protože kdybychom se o tom nebavili nějak víc, ale když to bylo v řádu stovek a na jednu stranu vždycky když… Jelikož se známe prostě dlouho, tak si pamatuju, že si tím účtem nedisponoval a nosil si sebou jenom hotovost. Takže to byl ještě další aspekt, ale to je prostě to, že toho člověka znáš a víš, jak se choval.*

Autor: *To, že jsem neměl v minulosti bankovní účet, ale jenom peníze v hotovosti, hrálo spíš ve prospěch podvodu nebo v neprospěch?*

R3: *No potom jsme se bavili, že sis zakládal bankovní účet. Takže to hrálo potom v prospěch, protože už jsem věděl, že sis ten účet založil. Člověk, když si založí účet a dá se všude platit kartou nebo telefonem, tak už hotovost u sebe moc nenosíš. Může se stát, že nemusíš mít potřebnou částku na tom účtu a než si prostě nějak zdlouhavě převádět finance třeba ze spořicího účtu. Což někdy může být zdlouhavý. Někdy to může být rychlým převodem, ale nad tímto prostě můžeš jen polemizovat, jestli ano nebo ne.*

Autor: *Co ti na rozhovoru přišlo podezřelé? To už jsi trošku zmínil. Můžeš to zopakovat znova tedy? Já bych se doptal, kdyžtak.*

R3: *Podezřelé mi určitě přišlo to, že si mi zavolal a sdílel si se mnou tu informaci. Já jsem zvyklý vždycky jako ten rozhovor rozvíjet. Něco se dozvědět nového, jak taková normální konverzace dvou kamarádů… Tak to mi přišlo podezřelý, ale vzal jsem v potaz, že za tebou je policejní hlídka a ty se prostě nemůžeš nějak vybavovat a musíš toto vyřídit co nejdřív.*

Autor: *Přišlo ti něco zvláštního na tom umělém hlase? Neber v potaz, jak jsem odpovídal, ale jestli ti to přišlo realistické, popřípadě slyšel si tam nějaké chyby jako monotónnost nebo nevyjádření emocí?*

R3: *No to jako určitě, ale největší chybou, kterou bych chtěl vytknout… Byla možná ta hloubka hlasu. Já jsem tě v tom sice poznal. Poznal jsem toho člověka, který mi volal. Sice z neznámého čísla, ale hnedka, jak jsem ten telefon vzal, tak jsem věděl prostě, že seš to ty. To je zase o tom, že toho člověka znám, ale prostě ta hloubka hlasu se mi tam zdála taková pofidérní, ale dalo by se říct, že jsem to bral tak, že seš prostě ve stresu a stres udělá hodně dle mého názoru.*

Autor: *Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?*

R3: *Určitě jsem se nad tím velice zamyslel, že už vlastně v dnešní době nemůžeš věřit ani svým dobrým kamarádům, samozřejmě. Což vlastně tam nejde o to, že bys jako nevěřil kamarádům, ale že ten hlas, že ten útočník může využít hlasu cizího člověka… Když mu prostě zcizí telefonní seznam a ví o tom člověku aspoň něco. Tak si myslím, že jsi schopný vymámit z těch lidí nějaký částky, ale jak jsem říkal, kdyby to byly řády tisíců, tak už bych se zamyslel víc. Tady to bylo prostě v řádu dvou stovek, což je pro dnešní společnost malý peníz.*

Autor: *Budeš do budoucna pracovat s větší paranoiou? Když ti někdo zavolá a bude chtít po tobě peníze a bude to třeba tvůj kamarád, rodinný příslušník, ale bude volat třeba z cizího čísla, nebo ze stejného čísla, které má za pomocí spoofingu.*

R3: *Určitě je to pro mě zajímavá zkušenost. Takhle to poznat úplně z druhé strany... Jo, kdy už opravdu v dnešní době, troufám si říct, nemůžeš věřit na sociálních sítích i přes telefonní hovory vlastně nikomu a jako na mě to celkem udělalo dojem. Určitě se budu nad tím zamýšlet do budoucna. A já třeba osobně jako nerad půjčuji takhle peníze, když ty lidi neznám jo, tady to bylo prostě něco jiného. Ten hovor byl udělaný tak, že toho člověka znám. Na druhou stranu si myslím, že to ve mně nechá ten dojem… Prostě příště, když bude člověk nebo kdokoliv z mého okolí potřebovat půjčit jakoukoliv finanční částku. Určitě si s ním domluvit osobní setkání a neřešit to přes telefon, tak přes jakoukoliv sociální síť.*

Autor: *Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti? Ty jsi tady zmínil tedy osobní setkání, ale co když ta možnost osobního setkání nebude? Třeba budeš otec, tvůj syn bude studovat vysokou v Praze a útočník bude potřebovat právě jeho hlasem půjčit třeba 2000 korun. To osobní setkání v tuto chvíli nebude možné.*

R3: *No to je otázka, jakože kdybychom mluvili o tomhle případu, tak kdybych měl syna jako na vysoké škole, tak bych ho určitě podporoval nějakou částkou. Posílal bych mu měsíčně nějakou částku a kdyby půl roku studoval na té škole a stačila mu ta částka… Pak by z ničeho nic mi takhle zavolal, že potřebuje poslat nějakou vyšší částku, na jakékoliv výdaje, tak bych se nad tím určitě zamyslel po této zkušenosti. A zase na druhou stranu bych přemýšlel o tom, proč by mi to třeba nesdělil už při té příležitosti, kdyby třeba dojel domů.*

Autor: *A třeba ověření přes jiná média?*

R3: *Určitě, z vlastní zkušenosti máme prostě třeba v rodině WhatsApp skupiny nebo komunikuješ přes messenger. Určitě bych se i snažil dopátrat vlastně k tomu, jak to je, i přes jiné médium.*

Autor: *Zmínil jsi teďka ty rodinné skupiny, což mi zní jako opravdu skvělá možnost, jak se bránit proti útokům s hlasem rodinného příslušníka. Jestli to chápu dobře, tak tam máte jenom rodinu, takže tam by sis to dokázal bez problémů ověřit?*

R3: *Jo já si myslím, že jo. No, a ještě je tady jedno morální hledisko, že kdyby mi zavolal vlastní syn, mluvil takhle ve zkratkách a takhle divně, tak už mně by to přišlo velice podezřelý. Přemýšlel bych jestli ho nikdo nevydírá nebo není v nějaké takové situaci a začal bych to řešit jinak, kdyby mi takhle zavolal syn… Prostě řekl větu a já bych se s ním chtěl začít normálně bavit, jak jsem normálně zvyklý se bavit. V tomhle případě by mi to přišlo hodně podezřelý, protože si myslím, že není normální, aby ti zavolal syn a řekl, že potřebuje nutně půjčit peníze. Nad tímto bych se zamyslel. Kdyby sis s tím člověkem obden nebo co 3 dny volal. Tak kdyby ti najednou prostě zavolal a ve zkrácené verzi dialogu by ti řekl jenom co chce, neodpovídal by ti na ty tvoje otázky, tak bych měl určitě větší podezření. Možná bych měl i strach, jestli právě není terčem nějakého vydírání, nebo nějakého jiného útoku.*

Autor: *Mělo by se o této problematice více informovat, popřípadě napadá tě jak?*

R3: *Tak já si myslím, že všichni víme, jak je AI v dnešní době na vysoké úrovni. Myslím si, že tahle problematika spadá do kyberbezpečnosti. Nebylo by tedy vůbec od věci vytvářet nějaké semináře nebo přednášky právě o této problematice. Já jsem si třeba myslel, že na tohle se můžou napálit jenom lidi ve starším věku, kteří mají tu důvěru i k cizímu člověku. Na mě ale teďka jde vidět, že se může napálit i člověk, který má obecné znalosti o této problematice a zajímám se o novinky v oboru IT. Určitě by teda, jak už jsem se o tom zmínil, nebylo od věci právě udělat preventivní programy na toto téma. Třeba v rámci mého oboru, tak máme často přednášky od BESIP, který nám přednáší o různých rizicích na cestách atd. V rámci vývoje umělé inteligence, tak si myslím, že určitě by to nebylo od věci vytvořit preventivní programy.*

PŘÍLOHA P V: ROZHOVOR S RESPONDENTEM Č.4

R4 je zkratka pro respondenta č.4

Autor: *Jaké pocity v tobě zanechalo zjištění, že se jednalo o podvod?*

R4: *No, takže byla jsem ráda, že jsem teda na to přišla, jakože je to podvod. Na druhou stranu mě hodně překvapilo, že vlastně jsem se s něčím takovým setkala, protože jsem myslela, že do Česka nic takového ještě nedošlo. Jako bylo to dobře provedené jenom by to chtělo asi teda do budoucna možná trochu víc upravit, aby to možná potom vyšlo i na mě.*

Autor: *Takže zlepšit reakce na tvoje otázky?*

R4: *Ano*

Autor: *Co ti na rozhovoru přišlo podezřelé?*

R4: *No tak ten hlas, ale až v té druhé odpovědi, protože v té první odpovědi to bylo jako v pohodě, ale v té druhé odpovědi jsem si říkala ježišmarja, je to tenhle Dejv, není to jiný Dejv. No a pak jako taky jsem si říkala, že kdybys chtěl peníze, proč bys volal mě? Jakože ne že bychom nebyli kamarádi, ale zas jakože abychom si říkali o peníze to se nestává...*

Autor: *Přišel ti ten hlas jiný kvůli tomu, že jsme se třeba teďka 8 měsíců nebavili, nebo fakt ti to přišlo úplně jiné, třeba ohledně výšky hlasu.*

R4: *Hmm asi ta intonace toho hlasu. Nooo, jako už jsem tě po telefonu slyšela víckrát, takže noo… Tak, jako mohlo to být třeba klidně, když je člověk ve stresu tak změní trochu tón hlasu, ale ten půl-tón zůstává stejný, ale teďka to bylo takové trochu jako… Bylo to jiné.*

Autor: *Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod?*

R4: *Je to dobrý, když jsem na to přišla, ale je to dáno i tím, že si pokaždé ověřím neznámé telefonní číslo. Dávám si to do Googlu a podle toho, jestli mi vyjede nebo nevyjede nějaký výsledek, tak jednám. Teďka nevyjel žádný výsledek. Tak jsem zavolala zpátky, protože jsem si říkala, že je to někdo známý. Většinou mi ale vyjede, že to je reklama nebo volá někdo z banky, a tak to nezvedám a rovnou si to popisuji nezvedat nebo tak nějak. Ale myslím si, že kdyby ten tón byl stejný, tak bych to možná nepoznala. Řekla bych si, jooo, tak asi ti to nikdo nezvedal, tak si zkusil mě.*

Autor: *Takže by mě mohla zachránit ta připravená situace?*

R4: *No jako nejspíš, ale asi bych měla trochu víc otázek, než bych něco udělala. A myslím si, že stejně bych ti zkusila zavolat na tvůj telefon.*

Autor: *To si víceméně udělala.*

R4: *Jo jakože tím, že tohle bylo jiné číslo… nebo bych ti napsala na messenger nebo na Facebook ať pošleš třeba… Ale zas ten Facebook je asi taky lehce prolomitelný, ten je možná jako těžko říct no.*

Autor: *Tady mě napadá otázka ohledně tvojí paranoii do budoucích telefonátů, když ti budou volat přátelé nebo rodina. Kdy existuje i spoofing, kterým si můžou útočníci vzít stejné číslo, jaké má tvoje rodina. Jak ohledně takovýchto telefonátů?*

R4: *Jako asi nad tím budu možná teďka víc přemýšlet, když vím, že se mi to může stát. No asi teďka budu taková opatrnější, že když po mě třeba někdo něco bude chtít, tak si to nějak jako dvakrát ověřím. Teď mě třeba napadlo, že bych mohla položit nějakou úplně nesmyslnou otázku a uvidím, co na to protějšek odpoví. Nebo nějakou lež ve stylu, že jsme se včera domluvili, že se dneska potkáme a pozorovala bych, jak na to ten druhý bude reagovat.*

Autor: *Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti?*

R4: *Myslím si, že tady tohle je nejrizikovější pro starší generace, které jednak třeba vůbec o nějakém takovém podvodu neví a ani neví, že nějaká umělá inteligence existuje. Takže si myslím, že by měla být nějaká prevence, aby se naučili buď lépe používat telefon, anebo prostě, aby jim ta rodina vysvětlila, co se může stát. Vytvořit nějaké video preventivně, které by se pouštělo, které by to tak nějak ukazovalo, co se všechno může stát. Vlastně aby pochopili, že to není jenom tak zvednout telefon a s někým si popovídat, když to nese nějaké ty rizika.*

Autor: *Chtěla by si to předat jenom do rukou těch rodin anebo by to mělo být i v jiných…*

R4: *Tak určitě i ve školách by se to mělo asi jako učit, aby to už byla tak nějak součást té výuky, protože tohle už s námi bude asi do budoucna pořád. Už se toho nikdy nezbavíme, takže je určitě dobré, aby o tom prostě všichni věděli a nějaká ta prevence se zavedla no.*

Autor: *Napadá tě ještě něco, co bys chtěla k tomuto tématu říct?*

R4: *...*

Autor: *Mě ještě jen ohledně toho hlasu napadlo. Nebylo ti podezřelé, jak jsem reagoval na tvoje otázky nebo odpovědi?Nepřišlo ti to moc divné právě v tomto?*

R4: *No tak ta poslední, že tam už si neodpověděl ani na tu otázku, co jsem se ptala, tak už to bylo takové, jakože rychle to ukončit. Takže to už bylo takové podezřelé, ale jinak ze začátku určitě ta připravenost byla.*

Autor: *Ty jsi mi tam potom hned zavolala na mé reálné číslo. Jak tě to napadlo?*

R4: *Hmmm.*

Autor: *Prostě automaticky nějaký pud?*

R4: *No asi jo, já nevím ani. Pro mě je totiž normální, když si i něco čtu, tak si to ověřím třeba z více zdrojů. Nebylo to nic pro mě nepřirozeného, že bych si to zkusila ověřit.*

Autor: *Máš takovou povahu prostě jako zvídavou?*

R4: *Spíš nedůvěřivou.*

Autor: *Nedůvěřivá, takže jsem narazil na špatnou oběť. Chápu.*

PŘÍLOHA P VI: ROZHOVOR S RESPONDENTEM Č.5

R5 je zkratka pro respondenta č.5

Autor: *Jaké pocity v tobě zanechalo zjištění, že se jedná o podvod?*

R5: *No pocit… Pocit to ve mně zanechalo špatný, jelikož ten hlas byl velmi, velmi podobný tobě, ale jelikož se jednalo o malou částku, tak jsem nějak ztratil ostražitost.*

Autor: *Takže mě zachránila v rámci toho rozhovoru, že jsem chtěl malou částku, takže ti to kvůli tomu nebylo nápadné?*

R5: *Ano, protože mě byl sice podezřelý ten nátlak. Tak se totiž v normálním životě nechováš, ale říkal jsem si, že se může stát cokoliv, a navíc se jednalo o malou částku, tak jsem to nijak neřešil.*

Autor: *Co ti na rozhovoru přišlo podezřelé?*

R5: *No podezřelé… Jako úplně podezřelé mě to přišlo až později, až když jsem začal nad tím přemýšlet. Ten nátlak a rychlost hovoru a všechno honem, honem.*

Autor: *A co se týče třeba mého hlasu, když jsem na tebe mluvil, měl jsem tam delší pasáže. Tak nepřišlo ti třeba na tom hlasu něco divné, že by sis řekl, tohle není úplně jako můj syn.*

R5: *No intonace… to mě ani tak nepřišlo. Spíš rychlost mluvy, protože v normálním životě, když se spolu bavíme o nějakých problémech, tak prostě nemáš takový hlas, teda hlas máš, ale nemluvíš takto naléhavě.*

Autor: *Takže spíš, než ten syntetický hlas, tak tě zarazila ta situace vytvořená mnou až tedy potom co jsme dotelefonovali?*

R5: *Ano*

Autor: *Jaký dopad měl tento zážitek na tvoji důvěru ve vlastní schopnost rozpoznat podvod? Popřípadě můžeš asi zmínit i jak se hodláš do budoucna proti tomu bránit.*

R5: *No tak, jelikož jak se říká ta umělá inteligence, je velmi, velmi zákeřná. Tak to prostě, i když se bude jednat o malé částky a podobnost hlasu. Přesto si budu ověřovat věci no.*

Autor: *A jakým způsobem by sis to ověřoval třeba?*

R5: *Zavolat zpátky na telefonní číslo, které mám v seznamu nebo na jiné aplikace třeba Viber nebo Whatsapp.*

Autor: *Takže si ověřit informaci skrze… Nevím, jestli si to uvedl moc dobře, ale chtěl si asi říct, že skrze kontaktní údaje, co máš v kontaktech, tak přes ty telefonní čísla ověřovat?*

R5: *Ano, vesměs vytáčím hlasem z telefonního seznamu.*

Autor: *Jak si myslíš, že by se měla společnost zabývat tímto druhem podvodu v blízké budoucnosti?*

R5: *No mělo by probíhat velmi, velmi jako rozsáhlejší… třeba v médiích nějaké školení, já bych neřekl školení, ale nějaké prostě informační sdělení v televizi na internetu víceméně.*

Autor: *Co by bylo pro tebe, ve tvém věku, přínosné v rámci toho školení nejvíce?*

R5: *Ta informace, jak se vůči těmto věcem bránit.*

Autor: *Takže nejen informovat, ale i nějaká protiopatření by si uvítal?*

R5: *Tak, tak ano. Protiopatření, jestli potom se bude dát do telefonu stáhnout nějaká aplikace, která. by mohla tu umělou inteligenci rozpoznat?*

Autor: *Napadá tě ještě něco, co bych si k tomuto tématu chtěl dodat?*

R5: *Hmm, to je asi tak všechno. Říkám hlavně teďka už člověk bude víc přemýšlet po téhle zkušenosti.*

PŘÍLOHA P VII: SEZNAM PŘILOŽENÝCH SOUBORŮ

\respondent_1.mp3	Zvukový záznam z průběhu podvodu s respondentem č.1.
\respondent_2.mp3	Zvukový záznam z průběhu podvodu s respondentem č.2.
\respondent_3.mp3	Zvukový záznam z průběhu podvodu s respondentem č.3.
\respondent_4.mp3	Zvukový záznam z průběhu podvodu s respondentem č.4.
\respondent_5.mp3	Zvukový záznam z průběhu podvodu s respondentem č.5.
\video_pro_dotaznik.mov	Video vytvořené pro dotazník.
\plakaty.pdf	Oba informační materiály vytvořené v rámci práce.