AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE
WYDZIAŁ FIZYKI I INFORMATYKI STOSOWANEJ

KATEDRA INFORMATYKI STOSOWANEJ I FIZYKI KOMPUTEROWEJ

Praca dyplomowa

Anonimowość w sieci - analiza rozwiązań i ich wpływu na działanie aplikacji i serwisów internetowych

Anonymity in the network - analysis of solutions and their impact on the operation of applications and websites

Autor:	Jędrzej Łukasz Szostak
Kierunek studiów:	Informatyka Stosowana
Opiekun pracy:	dr inż. Ewa Olejarz-Mieszaniec

Kraków, 2024

Spis treści

1. Wstęp

1.1. Wprowadzenie

W dzisiejszych czasach ciężko uzyskać częściową, a co dopiero pełną anonimowość w sieci. Każda nasza aktywność online jest gdzieś zapisywana i agregowana. Wszystkie serwisy internetowe zbierają informacje o swoich użytkownikach, czy to w celach targetowania behawioralnego czy zwyczajnego świadczenia usług. W wielu przypadkach są to dane krytyczne ze względu na prywatność, takie jak dane medyczne lub płatnicze. Oczywiście niemożliwe jest nieprzesyłanie żadnych danych, ale warto ograniczyć ich ilość w sytuacjach, które tego nie wymagają. Wiele z informacji udostępniamy sami, nieświadomi jakie zagrożenia może to za sobą nieść. W sieci czyha wiele osób, które czekają na zdobycie wykorzystanie ujawnionych danych. Może to wynikać z nieuwagi, albo niepoprawnego przechowywania tych zasobów, a w konsekwencji wycieku. Idealnym przykładem jest ostatni głośny wyciek z bazy laboratoriów diagnostycznych *ALAB* [1]. Oprócz osób prywatnych czy tak zwanych hakerów mających złe intencje, również dla organizacji rządowych nasze dane mogą być łakomym kąskiem. Czy to w celach inwigilacji (patrz Chiny), czy wpływania na nastroje społeczne i targetowanie osób na odpowiednich stanowiskach. W sieci trzeba być bardzo ostrożnym, gdyż pozornie niegroźne i codzienne czynności mogą udostępniać o nas dane krytyczne i wrażliwe takie jak lokalizacja, dane adresowe, dane płatnicze oraz zdjęcia czy historię przeglądania. Wielu użytkowników i deweloperów na szczęście zdaje sobie sprawę z tych zagrożeń. Dzięki temu powstaje sporo narzędzi mających na celu ograniczenie ekspozycji naszych danych wrażliwych, a użytkownicy już teraz mają dostęp do wielu technologi takich jak trybu prywatnego w wyszukiwarkach, rozwiązania typu *VPN* i *Proxy*, *zFone*, trasowanie cebulowe *TOR*, kryptowaluty, tymczasowe adresy e-mail, *adblocki*, szyfrowane komunikatory i dedykowane przeglądarki skupiające się na bezpieczeństwie. Nie tylko dobór odpowiednich dedykowanych rozwiązań ma znaczenie. Różnice pod kątem zbierania danych mogą występować nawet pomiędzy pozornie wręcz identycznymi i mało znaczącymi narzędziami takimi jak przeglądarki czy wyszukiwarki internetowe. Niestety, niektóre z nich nie informując o tym wprost, generują pokaźne przychody ze zbierania i przetwarzania naszych danych. Pomimo kuszących zapewnień producentów, trzeba mieć na uwadze, na ile te rozwiązania są skuteczne i jaki mają wpływ na komfort użytkowania serwisów internetowych. Choć poszukiwanie prywatności w internecie jest działaniem pozytywnym, to warto mieć na uwadze, że może być również wykorzystywane przez przestępców w celu ukrycia swojej tożsamości lub przeprowadzenia nielegalnych operacji takich jak odbywają się w *darknet* [2].

1.2. Cel pracy

Celem niniejszej pracy jest przyjrzenie się i przeprowadzenie kompleksowej analizy rożnego rodzaju technologii mających gwarantować anonimowość i bezpieczeństwo w sieci. Założeniem projektu jest przedstawienie rozwiązań, które można wykorzystać w codziennej pracy przez zwykłego użytkownika. Z tego powodu głównie będą analizowane narzędzia dostępne dla szerszego grona odbiorców i w większości darmowe, co pozwoli na wprowadzenie ich w życie w kontekście osób fizycznych, a nie wielkich korporacji posiadających wręcz nieograniczony budżet. Przegląd rozpocznie się od egzegezy aktów prawnych regulujących te zagadnienia i określających granice pozyskiwania oraz przetwarzania danych. Kolejnym krokiem będzie omówienie teoretyczne wszystkich rozpatrywanych w pracy technologii (w większości niekomercyjnych). Następnie, wybrane z nich zostaną przetestowane w dwóch aspektach:

– wydajności – jak ich użytkowanie wpływa na prędkość wykonywania zapytań i odpowiedzi oraz zużycie krytycznych zasobów komputera. Jest to kluczowe, gdyż rozwiązanie może zapewniać bardzo dużą ochronę (całkowita, jak autor postara się udowodnić w tej pracy jest niemożliwa), ale jednocześnie wpływać na komfort użytkowania, co sprawi że ostatecznie będzie nieefektywne.

– skuteczności – na ile deklaracje producentów pokrywają się z rzeczywistością oraz w jakim stopniu są to rozwiązania sprawiające rzeczywistą różnicę w szeroko pojętym zakresie bezpieczeństwa.

Do tego celu zostanie stworzona aplikacja agregująca i przetwarzająca uzyskane dane statystyczne. Będzie napisana w języku Python i działać z poziomu wiersza poleceń (*command-line*). Oprócz zbierania (za pomocą wysyłania zapytań) i przetwarzania statystyk, pozwoli również na raportowanie i wygodną prezentację uzyskanych rezultatów. Następnie uzyskane wyniki będą analizowane i porównywane między sobą. Na tej podstawie autor postara się dokonać podsumowania uzyskanych rezultatów i zaproponować optymalne rozwiązanie dające się w prosty sposób wprowadzić w życie.

2. Bezpieczeństwo danych

2.1. Gromadzenie danych

W dobie cyfryzacji i powszechnego dostępu do Internetu, dane użytkowników stały się jednym z najcenniejszych zasobów dla firm i organizacji na całym świecie. Codzienne korzystanie z Internetu przez miliony osób generuje ogromne ilości danych, które są systematycznie zbierane, analizowane i wykorzystywane w różnorodnych celach. W jaki jednak sposób są zbierane? Poniżej przedstawiono najczęściej stosowane metody.

2.1.1. Pliki cookies

Pliki *cookies*, czyli tzw. ciasteczka, to niewielkie pliki tekstowe, które są przechowywane na urządzeniu użytkownika przez przeglądarkę internetową na potrzeby strony internetowej. Pliki te są odczytywane przez serwer lub przeglądarkę podczas kolejnych wizyt na tej samej stronie. Mogą przechowywać różnorodne dane. Do najczęściej zapisywanych informacji należą [3]:

- Identyfikator sesji – unikalny numer przypisywany do sesji użytkownika, który pozwala na identyfikację jego działań podczas wizyty na stronie.
- Preferencje użytkownika – ustawienia wybrane przez użytkownika, takie jak język strony, wielkość czcionki czy układ strony.
- Dane logowania – informacje niezbędne do uwierzytelnienia użytkownika, takie jak nazwa użytkownika i zaszyfrowane hasło.
- Historia przeglądania – informacje o odwiedzanych stronach i klikniętych linkach.
- Dane koszyka zakupowego – informacje o produktach dodanych do koszyka, które pozwalają na kontynuowanie zakupów przy kolejnej wizycie na stronie.
- Informacje o reklamach – dane dotyczące wyświetlanych i klikanych reklam.

Pliki *cookies* można skategoryzować ze względu na różne czynniki [4]:

1. Ze względu na czas przechowywania:

- *Session cookies* – przechowywane tymczasowo w pamięci przeglądarki i usuwane automatycznie po zamknięciu przeglądarki. Używa się ich głównie do przechowywania informacji o aktywnej sesji użytkownika, takich jak zawartość koszyka.
- *Persistent cookies* – przechowywane na urządzeniu użytkownika przez określony czas (określony w parametrach) lub do momentu ich ręcznego usunięcia. Umożliwiają zapamiętanie preferencji użytkownika, danych logowania i innych ustawień.

2. Ze względu na pochodzenie:
   - *First-party cookies* – ustawiane przez odwiedzaną stronę internetową w celu przechowywania preferencji użytkownika i informacji niezbędnych do poprawnego działania strony.
   - *Third-party cookies* – ustawiane przez zewnętrzne serwisy (np. reklamy, analitykę) i używane głównie do celów marketingowych w celu personalizacji reklam.
3. Ze względu na funkcję:
   - *Essential cookies* – konieczne do prawidłowego funkcjonowania strony internetowej poprzez zapewnianie podstawowych funkcji, takich jak nawigacja po stronie czy dostęp do bezpiecznych obszarów strony.
   - *Functional cookies* – umożliwiają zapamiętywanie wyborów dokonanych przez użytkownika (np. nazwa użytkownika, język) i personalizację interfejsu. Ułatwiają korzystanie ze strony i poprawiają jej funkcjonalność.
   - *Performance cookies* – zbierają informacje o sposobie korzystania ze strony internetowej przez użytkowników, takie jak najczęściej odwiedzane strony czy komunikaty o błędach. Pomagają w poprawie działania strony, analizy wydajności i optymalizacji treści.
   - *Advertising cookies* – używane do dostarczania reklam, które są bardziej dostosowane do użytkownika i jego zainteresowań. Śledzą aktywność użytkownika na różnych stronach i tworzą profile zainteresowań, które są wykorzystywane do targetowania reklam.
   - *Social media cookies* – umożliwiają integrację ze społecznościami i udostępnianie treści na platformach społecznościowych (np. *Facebook*, *Twitter*). Mogą również służyć do śledzenia aktywności użytkownika na platformach społecznościowych.

Pliki *cookies* są najpopularniejszą metodą zbierania danych. Mimo że przynoszą wiele korzyści, ich użycie wiąże się również z wyzwaniami prywatności i bezpieczeństwa.

2.1.2. Śledzenie za pomocą JavaScript

JavaScript jest językiem programowania wykorzystywanym głównie do tworzenia interaktywnych elementów na stronach internetowych. Skrypty *JavaScript* są osadzane w kodzie*HTML* strony i wykonywane przez przeglądarkę internetową. Dzięki swojej wszechstronności, język ten pozwala na dynamiczną manipulację elementami strony oraz na zbieranie danych w czasie rzeczywistym. Skrypty *JavaScript* mogą zbierać szeroką gamę danych o użytkownikach i ich aktywności na stronach internetowych. Od kliknięć, ruchów myszki i przewijania co pozwala na rejestracje interakcji z przyciskami

i linkami oraz znalezienie najczęściej odwiedzanych części serwisu po zapisywaniu danych wprowadzanych do różnych formularzy i zbieranie czasu spędzonego na stronie. Jak widać, stwarza to ogromne możliwości oraz zagrożenia. Do najpopularniejszych metod zastosowanych do tego celu należą:

1. *Google Analytics*

   Google Analytics to jedno z najczęściej używanych narzędzi do analizy ruchu na stronach internetowych. Integracja *Google Analytics* pozwala na zbieranie szczegółowych danych o odwiedzinach, demografii i zachowaniach użytkowników oraz efektywności kampanii marketingowych.

2. *Tag Manager*

   Google Tag Manager (GTM) to narzędzie umożliwiające zarządzanie i wdrażanie tagów śledzących na stronie internetowej bez konieczności bezpośredniej modyfikacji kodu. Za pomocą GTM można łatwo dodawać i konfigurować różnorodne tagi, takie jak śledzenie kliknięć, scrollowania czy interakcji z formularzami.

3. *Heatmapy*

   Narzędzia do tworzenia *heatmap* takie jak *Hotjar* czy *Crazy Egg* wykorzystują *JavaScript* do wizualizacji interakcji użytkowników z elementami strony. *Heatmapy* pokazują, które części strony są najczęściej klikane, przewijane lub najeżdżane myszką.

4. Rekordery sesji

   Rejestrują pełne sesje użytkowników, umożliwiając przeglądanie nagrań z ich aktywności na stronie. Dzięki temu można dokładnie zobaczyć, jak użytkownicy poruszają się po stronie, co klikają i jakie problemy napotykają.

5. *A/B testing*

   A/B testing polega na porównywaniu dwóch wersji strony internetowej (wersji A i wersji B) w celu sprawdzenia, która z nich przynosi lepsze rezultaty. Narzędzia do *A/B testingu*, wykorzystują JavaScript do losowego przydzielania użytkowników do różnych wersji strony i monitorowania ich zachowań.

Mimo swojego szerokiego zastosowania stanowi to wciąż bardzo duże zagrożenie i ingerencję w prywatność użytkowników.

2.1.3. Logi serwerowe

Logi serwerowe to pliki, które rejestrują wszystkie zdarzenia i działania mające miejsce na serwerze internetowym. Każda interakcja użytkownika z serwerem, taka jak odwiedzenie strony internetowej, pobranie pliku czy wysłanie formularza, jest w nich zapisywana. Są one kluczowym narzędziem do monitorowania i analizy ruchu, diagnozowania problemów oraz zapewnienia bezpieczeństwa, dlatego są często wymagane przez regulatorów prawnych. Najczęściej przechowywane w nich dane obejmują:

– Adres IP użytkownika

- Data i godzina zdarzenia (*timestamp*), umożliwiająca śledzenie aktywności w czasie.
- Rodzaj zapytania HTTP, takie jak GET, POST, PUT, DELETE.
- Adres URL strony lub zasobu, do którego uzyskano dostęp.
- Kod odpowiedzi serwera, informujący o wyniku zapytania.
- Adres URL strony, z której użytkownik został przekierowany na bieżącą stronę.
- Informacje o przeglądarce internetowej i systemie operacyjnym użytkownika.
- Wielkość danych przesyłanych w odpowiedzi na zapytanie.
- Czas potrzebny na przetworzenie zapytania przez serwer.

Pomimo wielu korzyści, logi serwerowe wiążą się również z pewnymi wyzwaniami. Zawierają one dane, które mogą być uznane za dane wrażliwe. Konieczne jest zapewnienie, że dane te są przechowywane i przetwarzane zgodnie z przepisami oraz są zabezpieczone przed nieautoryzowanym dostępem.

2.1.4. Fingerprinting

Jest to technika zbierania danych o urządzeniu i przeglądarce użytkownika w celu jednoznacznej identyfikacji tego urządzenia w sieci. W odróżnieniu od plików cookies, które są przechowywane na urządzeniu użytkownika, *fingerprinting* polega na analizie właściwości sprzętowych i programowych urządzenia, co pozwala na tworzenie unikalnego „odcisku palca” użytkownika. Technika ta jest często stosowana w analizie ruchu sieciowego, marketingu, bezpieczeństwie.

Fingerprinting wykorzystuje różnorodne dane dostępne podczas przeglądania stron internetowych. Najczęściej są to:

- Typ i wersja przeglądarki, jej język, zainstalowane wtyczki, czcionki, rozdzielczość ekranu, strefa czasowa.
- Typ i wersja systemu operacyjnego, ustawienia regionalne.
- Model i producent urządzenia, rozdzielczość ekranu, głębokość koloru.
- Adres IP, typ połączenia (Wi-Fi, sieć komórkowa), dostawca usług internetowych.
- Dane dotyczące procesora, karty graficznej, liczby rdzeni, poziomu naładowania baterii.
- Historia przeglądania, cookies, dane sesji.

Fingerprinting można przeprowadzać na różne sposoby. Zaliczają się do nich generowanie unikalnych obrazów i sprawdzanie subtelnych różnic w renderowaniu zależnych od urządzeń, analogicznie może być do tego wykorzystany dźwięk, dostępne fonty oraz API przeglądarki. Dlatego rozwiązanie to budzi poważne obawy dotyczące prywatności. Technika ta pozwala na śledzenie użytkowników bez ich wiedzy i zgody, co może prowadzić do nadużyć i naruszeń bezpieczeństwa.

Warto zaznaczyć, że *Fingerprinting* nie jest niezawodny i może być mniej skuteczny w przypadku urządzeń o podobnych konfiguracjach. Ponadto, użytkownicy mogą korzystać z narzędzi i technik zapobiegających *fingerprintingowi*, takich jak przeglądarki chroniące prywatność (np. Tor) lub rozszerzenia go blokujące [5].

2.1.5. Zbieranie danych z mediów społecznościowych

Media społecznościowe, takie jak *Facebook*, *Twitter*, *Instagram*, *LinkedIn*, i *TikTok*, stały się integralną częścią życia codziennego ludzi na całym świecie. Platformy te umożliwiają komunikację, wymianę informacji i interakcję społeczną, generując jednocześnie ogromne ilości danych o użytkownikach. W przeciwieństwie do innych sposobów, w tym często to użytkownicy z własnej woli udostępniają te dane, które można podzielić na kilka kategorii:

– Dane profilowe – Informacje podane przez użytkowników podczas tworzenia konta, takie jak imię, nazwisko, wiek, płeć, lokalizacja, wykształcenie, zatrudnienie, zainteresowania.

– Dane o aktywności – Informacje dotyczące działań użytkowników na platformie, takie jak publikowane posty, komentarze, polubienia, udostępnienia, subskrypcje i obserwowanie innych użytkowników.

– Dane behawioralne – Informacje o sposobie korzystania z platformy, takie jak czas spędzany na stronie, kliknięcia, przeglądane treści, interakcje z reklamami.

– Dane sieciowe – Informacje o połączeniach między użytkownikami, takie jak listy znajomych, obserwujących i obserwowanych kont.

– Dane multimedialne – Zdjęcia, filmy, nagrania audio i inne multimedia udostępniane przez użytkowników.

– Dane lokalizacyjne – Informacje o lokalizacji, zbierane na podstawie tagów geograficznych i ustawień lokalizacji.

Media społecznościowe są źródłem ogromnych ilości danych o użytkownikach. Wiele osób niestety nie zdaje sobie sprawy, że ich dane są gromadzone i analizowane [6].

2.2. Kategoryzacja danych

W sieci mamy do czynienia z bardzo dużą ilością danych, wiele z nich ma niewielkie znaczenie, ale są też takie, które nazywamy danymi wrażliwymi. Warto je zidentyfikować. Dane wrażliwe obejmują:

– PII (*ang. Personally Identifiable Information*) – są to dane pozwalające na identyfikację osoby. Zwykle są sankcjonowane prawnie.

– Dane osobowe – odnoszą się do osoby, ale nie zaliczają się do PII. Mogą to być informacje dotyczące ich zainteresowań, przekonań, lokalizacji oraz aktywności online i offline.

– Dane zastrzeżone i poufne – odnoszą się do informacji kontraktowych lub biznesowych. Ich ujawnienie mogłoby zagrozić firmie lub umowie.

PII to adres e-mail, płeć, imię, nazwisko, data urodzenia, adres zamieszkania, adres IP, profil w mediach społecznościowych, numer telefonu, numer ubezpieczenia społecznego lub inny numer identyfikacyjny, dane karty kredytowej, dane medyczne lub biometryczne. Ich cechą szczególną jest to, że są specyficzne dla danej osoby, mogą być używane samodzielnie lub w połączeniu z innymi informacjami do zidentyfikowania tej osoby. Należą do wrażliwych i regulowanych danych, ponieważ można je traktować jako unikalny identyfikator.

Nie jest to jednak kwestia trywialna. Klasyfikacja może w dużym stopniu zależeć od kontekstu. Weźmy na przykład lokalizację. O ile w domu taka informacja jest wrażliwa, bo ujawnia adres zamieszkania, to w pracy, może, już taka nie być (szczególnie, że wiele firm wymaga dostępu do lokalizacji pracowników). Dodatkowo, istotnym czynnikiem są preferencje dotyczące prywatności. Dla jednej osoby jej poglądy polityczne czy przekonania religijne są informacjami, które wolałyby zachować dla siebie, a inna będzie je codziennie udostępniać na portalach społecznościowych. Może to się tyczyć wielu danych, takich jak muzyka, zainteresowania, posiadane urządzenia, lista połączeń i wiele innych.

Dlatego tak potrzebne są regulacje, które dają użytkownikom wybór, które z ich danych są wrażliwe i nie chcą się nimi dzielić poprzez wybór w preferencjach prywatności.

Warto zaznaczyć, że istnieje ryzyko prywatności generowane pośrednio. Możemy zgadzać się na udostępnianie naszej lokalizacji, ale gdy dane te są agregowane i łączone w jednym miejscu z ich analizy można wyciągnąć informacje takie jak adres zamieszkania, miejsce pracy, znajomi czy nawet ulubiona kawiarnia.

Podobnie wykazano, że analiza polubień na Facebooku może być wykorzystana do określenia wielu informacji takich jak płeć, orientacja seksualna i przekonania polityczne, a nawet da się na ich podstawie stworzyć cały profil psychologiczny. Są to oczywiście pewne przewidywania, ale wiadomo jest, że działania w Internecie i mediach społecznościowych pozostawiają unikalne ślady, identyfikując wzorce, które ujawniają osobiste cechy danej osoby. Pokazuje to niestety, że wszystkie dane generowane przez użytkowników w internecie dają możliwość identyfikacji [7].

Z tego powodu termin „dane wrażliwe"może oznaczać wszelkie dane osobowe, niezależnie od tego, czy bezpośrednio umożliwiają one identyfikację czy nie. Jak wyjaśniono wcześniej, zwłaszcza w dużych ilościach lub zgrupowane, mogą niestety do tego prowadzić.

Ostatnią kategorią danych wrażliwych są dane, które są zastrzeżone lub poufne z powodów niezwiązanych z osobami fizycznymi. Mogą to być tajemnice handlowe, poufne informacje o firmie lub niektórych produktach oraz informacje podlegające klauzulom poufności, które ze względu na swój charakter muszą być utrzymywane w tajemnicy. W wypadku ich wycieku, mogą zaszkodzić firmie lub dać nieuczciwą przewagę konkurencji.

Jak widać, nie jest to zagadnienie proste i posiada wiele niuansów. Dopiero po zidentyfikowaniu danych jako wrażliwych, można określić, jak najlepiej je chronić [8].

2.3. Zagrożenia

Dlaczego bezpieczeństwo jest to tak ważne? Co sprawia, że przez ostatnie kilkanaście lat systematycznie zwiększał się udział cyberprzestępstw w ogólnej liczbie wykroczeń [9]? Bez wątpienia kluczowy wpływ na to miała znaczna informatyzacja wielu usług i przeniesienie baz danych do Internetu. Pomimo sztywnych reguł co do hostowania i zabezpieczeń tego typu zasobów cały czas zdarzają się wycieki i naruszenia danych. Nie tylko hackerzy czyhają na nasze dane. W coraz bardziej polaryzującym się świecie również dla instytucji rządowych czy wielkich korporacji są to informacje na wagę złota. W ostatnich latach szczególnie po agresji rosyjskiej na Ukrainę przez rosyjski wywiad są przeprowadzane kampanie mające na celu identyfikację osób na krytycznych stanowiskach. Do tego zaliczają się nawet pozornie nieistotne zdjęcia wrzucane na media społecznościowe jakiegoś wyimaginowanego żołnierza z podpisem, aby inni komentowali wraz z identyfikacją swoich jednostek. Dla obcych wywiadów są to informacje na wagę złota. Równocześnie wiele przedsiębiorstw o charakterze krytycznym takich jak elektrociepłownie, przedsiębiorstwa transportowe (tutaj przykład paraliżu białoruskiej kolei) czy produkcyjne stało się celem cyberataków [10]. Oczywiście mogą być to ataki bezpośrednio na infrastrukturę, lecz wciąż bardzo często wektorem ataku jest człowiek i jego słabość. Dlatego tak istotne jest, aby nie ujawniać zbyt wielu informacji, które mogą w takich atakach pomagać. Patrząc od drugiej również cyberprzestępcy dbają o anonimowość tak żeby było ich jak najciężej zidentyfikować czy powiązać z daną organizacją czy instytucją.

Jednak jeśli wiadomo, że takie zagrożenia występują to warto je również zidentyfikować i skategoryzować. Oto najważniejsze z nich [11]:

1. Cyberprzestępstwa

   Zalicza się do nich nie tylko kradzieże, w tym niszczenie danych czy phishing, ale także nieuczciwą manipulację, oszustwa, bezpośredni lub pośredni przymus i inne działania niezgodne z prawem. Istotnym w kontekście prywatności jest ostatnio popularna kradzież tożsamości.

2. Inwigilacja

   Polega na nieuprawnionym monitorowaniu danych. Może być prowadzona zarówno przez instytucje państwowe (np. przy pomocy systemów wykrywania twarzy) lub przez podmioty prywatne czy osoby fizyczne (podglądanie mieszkań czy wynajmowanych obiektów za pomocą smart urządzeń, gromadzenie danych o geolokalizacji w celach marketingowych).

3. Profilowanie

   Jest to coś nieuniknionego w dzisiejszym świecie. Teoretycznie niegroźne, ale jak zostało wspomniane wyżej często wiąże się z łączeniem danych, dzięki czemu potrafi stanowić spore zagrożenie dla prywatności. Tutaj brane pod uwagę jako realizowane bez poszanowania praw i wbrew woli użytkownika.

4. Ujawnienie informacji

   Mogłoby się zaliczać do cyberprzestępstw, ale w tej kategorii zaliczamy takie, które wynikają z nieumyślnego działania. Może to być awaria systemu, czy brak odpowiedniej świadomości o bezpieczeństwie danych.

5. Zagrożenia związane z wykorzystaniem sztucznej inteligencji

   Dane wchodzące w skład datasetów uczących powinny być sprawdzane pod kątem zawierania informacji wrażliwych. W innym wypadku model może je wykorzystać, albo co gorsze zwrócić przypadkowo lub pod wpływem ataku nieuprawnionej osobie. Jest to szczególnie istotne w modelach językowych, które uczą się na bieżąco. W najlepszym wypadku udostępniamy dane firmie lub osobom odpowiedzialnym za dane narzędzie.

Jak można zaobserwować, czyhających zagrożeń jest bardzo dużo. Z tego powodu rządy zaczęły się im przyglądać i opracowywać odpowiadające na nie i mające chronić obywateli akty prawne.

2.4. Aspekt legalny i obowiązujące akty prawne

W celu przeciwdziałania wyżej wymienionym zagrożeniom wiele państw, w różnym czasie zaczęło wprowadzać prawa mające na celu ochronę oraz zabezpieczenie użytkowników przed wykorzystywaniem ich danych. Choć początki polegały przede wszystkim na definicji słów kluczowych i umożliwienie ścigania tego typu przestępstw to obecnie na całym świecie mamy jasno zdefiniowane reguły według, których przetwarzane i procesowane są nasze dane. Podstawa prawna do karania za naruszenie tych praw została jasno zdefiniowana. Niestety, nie są to akty perfekcyjne. Nie wszystko da się ująć w ramach prawniczego żargonu, a i potrafi on być często zagmatwany i niezrozumiały dla zwykłego zjadacza chleba. Podstawa prawna potrafi też znacznie różnić się w zależności od rejonu występowania co prowadzi nieraz do problemów natury logistycznej i implementacyjnej. Poniżej autor postara się zwięźle omówić rozwiązania obowiązujące w różnych częściach świata.

2.4.1. Unia Europejska

Jednym z pierwszych międzynarodowych aktów prawnych w dziedzinie przepisów dotyczących bezpieczeństwa i ochrony danych osobowych była Konwencja 108 Rady Europy z 28 stycznia 1981 roku. Celem Konwencji było zapewnienie każdej osobie fizycznej, bez względu na narodowość i miejsce zamieszkania, poszanowania jej praw i podstawowych wolności, zwłaszcza prawa do prywatności, w kontekście automatycznego przetwarzania danych osobowych. Elastyczny charakter Konwencji pozwalał jej sygnatariuszom na pewną swobodę w implementacji odpowiednich standardów ochrony prawnej jednostki w prawodawstwie krajowym. Jednak ze względu na rosnące zagrożenia dotyczące ochrony danych osobowych, zostały podjęte działania na rzecz ujednolicenia regulacji w tym zakresie. Efektem tych działań było uchwalenie Dyrektywy 95/46/WE Parlamentu Europejskiego i Rady Europejskiej z dnia 24 października 1995 [12] roku dotyczącej ochrony osób fizycznych w zakresie przetwarzania danych osobowych oraz swobodnego przepływu tych danych. W tamtym okresie była uznawana za przełomową. Skupiała się na obszernym zdefiniowaniu pojęcia danych osobowych i ich przetwarzania. Nie obejmowała przetwarzania danych przez osoby fizyczne w celach osobistych oraz działalności wykraczającej poza kompetencje prawne Wspólnoty, np. w zakresie bezpieczeństwa publicznego. Początkowo nie regulowała przetwarzania danych osobowych przez instytucje unijne.Określone zostały zasady legalnego przetwarzania. Było ono zgodne z prawem jedynie w przypadkach, gdy:

– osoba, której dane dotyczą, wyraziła na nie zgodę,

– było ono konieczne do realizacji umowy,

– wymagało tego prawo,

– służyło ochronie żywotnych interesów osoby,

– było niezbędne do wykonania zadania publicznego lub władzy publicznej,

– wynikało z uzasadnionych interesów administratora danych lub osoby trzeciej.

Dodatkowo nakładała obowiązek informacyjny wobec osób, których dane były przetwarzane. Dawała im możliwość dostęp do danych i sprzeciwu wobec ich przetwarzania. Dyrektywa zakładała poufność przetwarzania danych. W kontekście technologicznym, nakładała obowiązek stosowania odpowiednich środków technicznych i organizacyjnych w celu ochrony danych przed nieuprawnionym dostępem lub uszkodzeniem. Transfer danych do krajów trzecich było dozwolone przy odpowiednim poziomie ochrony, z wyjątkami, obejmującymi wyrażenie zgody lub gdy było to konieczne do realizacji umowy.

W 2012 roku rozpoczęto proces modernizacji Dyrektywy 95/46/WE, która coraz częściej była postrzegana jako wymagająca dostosowania do zmieniających się realiów technologicznych, takich jak rozwiązania chmurowe, rozwój Internetu i międzynarodowy przepływy danych. Z tego powodu Komisja Europejska przedstawiła projekt dwóch aktów prawnych, które stanowią podwaliny dzisiejszego ustawodawstwa:

1. Rozporządzenie w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych oraz swobodnego przepływu takich danych (Ogólne rozporządzenie o ochronie danych w skrócie RODO).
2. Dyrektywa w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez właściwe organy do celów zapobiegania przestępstwom, ich wykrywania, prowadzenia postępowań przygotowawczych, ścigania przestępstw oraz wykonywania kar, a także swobodnego przepływu takich danych.

Zmiany te weszły w życie w roku 2016, a państwa członkowskie miały czas na dostosowanie do nich krajowych przepisów do 2018 roku. RODO [13] wprowadziło szerokie zmiany w zakresie ochrony danych osobowych na terenie Unii Europejskiej. Główne z nich to:

– Ujednolicenie w państwach członkowskich – RODO w przeciwieństwie do Dyrektywy 95/46/WE jest rozporządzeniem, co oznacza, że jest bezpośrednio stosowane we wszystkich państwach członkowskich UE. Pozwala to zniwelować różnice ustawodawcze na terenie Unii.

– Zakres terytorialny – rozszerza zakres stosowania na podmioty spoza UE, które oferują towary lub usługi na terenie wspólnoty.

– Zasady przetwarzania danych – ustanawia siedem podstawowych zasad przetwarzania danych: legalność, rzetelność, przejrzystość, ograniczenie celu, minimalizacja danych, prawidłowość, ograniczenie przechowywania, integralność i poufność, oraz rozliczalność.

– Prawa osób, których dane dotyczą – wprowadza nowe prawa, takie jak prawo do przenoszenia danych, prawo do bycia zapomnianym, oraz rozszerza prawa dotyczące dostępu do danych i ich sprostowania.

– Zgoda na przetwarzanie danych: wymaga, aby zgoda była wyrażona jednoznacznie i świadomie, oraz aby była łatwa do wycofania.

– Inspektor ochrony danych (DPO) – wprowadza obowiązek powołania *Inspektora Ochrony Danych* w sytuacjach, takich jak przetwarzanie lub monitorowanie na dużą skalę danych szczególnych kategorii.

– Ocena skutków dla ochrony danych (DPIA) – wprowadza obowiązek przeprowadzania DPIA w przypadku przetwarzania danych mogącego powodować wysokie ryzyko naruszenia praw lub wolności osób fizycznych.

– Zgłaszanie naruszeń ochrony danych – wymaga zgłaszania naruszeń ochrony danych osobowych organowi nadzorczemu w ciągu 72 godzin od ich stwierdzenia oraz powiadamiania osób, których to dotyczy.

– Kary za naruszenia – określa surowe kary finansowe za naruszenia przepisów, sięgające do 20 milionów euro lub do 2 bądź 4% całkowitego rocznego obrotu przedsiębiorstwa, w zależności od rodzaju i skutków przewinienia. Poprzednio sankcje były regulowane przez przepisy krajowe.

– Podejście oparte na ryzyku – hołduje podejściu opartemu na ryzyku, które wymaga od administratorów i podmiotów przetwarzających ewaluację ryzyka związanego z przetwarzaniem danych i stosowania odpowiednich środków ochronnych.

– Międzynarodowe transfery danych – stosuje surowe zasady dotyczące transferów danych do państw trzecich, wymagając odpowiedniego poziomu ochrony danych lub zastosowania odpowiednich zabezpieczeń. Wprowadza znacznie bardziej szczegółowe wymogi i procedury.

RODO stanowi znaczący krok naprzód w zakresie ochrony danych osobowych, odpowiadając na wyzwania związane z globalizacją, postępem technologicznym i rosnącą cyfryzacją. Nowe przepisy wprowadzają bardziej szczegółowe i surowe wymagania, kładąc większy nacisk na odpowiedzialność administratorów danych oraz na prawa osób, których danych to dotyczy.

2.4.1.1. Polska

Ochrona danych osobowych w Polsce ma swoje początki w Ustawie o ochronie danych osobowych, która weszła w życie 30 kwietnia 1998 roku. Ustawa ta była efektem implementacji przepisów uchwalonych 29 sierpnia 1997 roku i miała na celu dostosowanie polskiego prawa do międzynarodowych standardów, w szczególności do dyrektywy 95/46/WE Parlamentu Europejskiego. Przyjęcie ustawy nastąpiło w czasie, gdy Polska intensywnie przygotowywała się do akcesji do Unii Europejskiej, jako krok w celu wdrożenia odpowiednich regulacji prawnych.

Pierwsza ustawa była stosunkowo zwięzła, liczyła dziewięć stron i wprowadzała kluczowe definicje związane z ochroną danych osobowych, takie jak dane osobowe, administrator danych, zbiór danych,

przetwarzanie oraz katalog danych wrażliwych. Określała także przesłanki legalności przetwarzania danych oraz szczegółowe obowiązki informacyjne zależne od źródła pozyskania danych. Ustawa ustanowiła GIODO, który miał nadzorować przestrzeganie przepisów o ochronie danych osobowych. GIODO miał uprawnienia do prowadzenia kontroli, nakazywania usunięcia uchybień, udostępniania danych oraz ich zabezpieczania. Nie miał jednak możliwości nakładania administracyjnych kar pieniężnych. Ustawa wymagała rejestracji zbiorów danych osobowych, co często skutkowało przesyłaniem kompletnych baz danych do GIODO. Kolejnym istotnym krokiem w polskim prawodawstwie była implementacja Rozporządzenia Ogólnego o Ochronie Danych Osobowych (RODO), które weszło w życie w 2018 roku. UODO (które zastąpiło GIODO) zyskał możliwość nakładania administracyjnych kar pieniężnych, co znacząco zwiększyło jego zdolność do egzekwowania przepisów.

Ochrona danych osobowych w Polsce przeszła długą drogę od wprowadzenia pierwszej ustawy w 1998 roku, poprzez implementację RODO w 2018 roku, aż po dzisiejsze wyzwania związane z dynamicznie zmieniającymi się realiami technologicznymi i społecznymi. Chociaż świadomość i ochrona praw osób fizycznych znacząco się poprawiły, przed UODO wciąż stoi wiele wyzwań, a proces adaptacji do nowych realiów pozostaje ciągły [14].

2.4.2. USA

W Stanach Zjednoczonych ochrona danych osobowych i prywatności w sieci jest zdecentralizowana, co oznacza, że istnieją zarówno federalne, jak i stanowe regulacje, które mogą się różnić pod pewnymi względami [15]. Najważniejsze z nich to:

- *Children's Online Privacy Protection Act* – COPPA ma na celu ochronę prywatności dzieci poniżej 13. roku życia podczas zbierania ich danych osobowych przez firmy online. Wymaga ona zgody rodziców na zbieranie danych, określa, jakie informacje mogą być zbierane i jak powinny być przechowywane. Instytucją nadzorującą jest Federalna Komisja Handlu (*Federal Trade Commission* - FTC).

- *Health Insurance Portability and Accountability Act* – HIPAA koncentruje się na ochronie prywatności i bezpieczeństwa danych zdrowotnych pacjentów. Określa, jakie dane mogą być przechowywane i przetwarzane oraz jakie środki bezpieczeństwa muszą być stosowane przez organizacje zdrowotne. Nadzór sprawuje Departament Zdrowia i Opieki Społecznej (*Department of Health and Human Services - HHS*).

- *Gramm-Leach-Bliley Act* – GLBA chroni dane osobowe klientów instytucji finansowych. Wymaga od instytucji finansowych utrzymywania polityk prywatności, zapewnienia bezpieczeństwa danych i informowania klientów o praktykach dotyczących prywatności. Instytucje nadzorujące to federalne instytucje finansowe, w tym Federalna Komisja Handlu (FTC).

- *Electronic Communications Privacy Act* – ECPA chroni prywatność komunikacji elektronicznej, w tym e-maili, rozmów telefonicznych i danych przesyłanych przez Internet. Określa warunki, w jakich rząd i inne podmioty mogą uzyskiwać dostęp do takich danych. Nadzór nad tymi przepisami sprawuje Departament Sprawiedliwości (*Department of Justice* - DOJ).

– *Fair Credit Reporting Act* - FCRA chroni prywatność informacji zawartych w raportach kredytowych i innych raportach konsumenckich. Określa, jakie informacje mogą być zawarte w raportach kredytowych, jak długo mogą być przechowywane oraz jakie są uprawnienia konsumentów w zakresie dostępu do swoich danych. Nadzór sprawuje Federalna Komisja Handlu (FTC).

– *State Data Breach Notification Laws* wymaga zgłaszania naruszeń bezpieczeństwa danych osobowych klientów lub mieszkańców danego stanu. Określa, jakie organizacje muszą zgłaszać naruszenia, kiedy i jakie są wymogi dotyczące powiadomienia osób dotkniętych naruszeniem. Nadzór zazwyczaj sprawują agencje ochrony konsumentów na poziomie stanowym.

– *California Online Privacy Protection Act* – CalOPPA chroni prywatność użytkowników Internetu, zwłaszcza w Kalifornii. Wymaga umieszczania przez strony internetowe polityk prywatności i informowania użytkowników o zbieraniu ich danych osobowych. Nadzór sprawuje Generalny Prokurator Kalifornii.

Przenoszenie danych osobowych między Unią Europejską a Stanami Zjednoczonymi wiąże się z licznymi wyzwaniami i problemami prawnymi. W UE obowiązuje Ogólne Rozporządzenie o Ochronie Danych, które nakłada surowe wymogi na przetwarzanie danych osobowych i ich transfer poza granice UE. W 2020 roku Trybunał Sprawiedliwości Unii Europejskiej unieważnił umowę *Privacy Shield*, która umożliwiała transfer danych między UE a USA, uzasadniając to obawami dotyczącymi ochrony danych przed dostępem ze strony amerykańskich służb wywiadowczych.

Obecnie transfer danych między UE a USA może odbywać się na podstawie standardowych klauzul umownych (SCC) lub innych mechanizmów zgodnych z RODO. Jednak te rozwiązania również budzą kontrowersje i są przedmiotem skarg i postępowań sądowych. Prawo amerykańskie nie oferuje takiej samej ochrony danych osobowych jak prawo unijne, co budzi obawy o naruszenie prywatności danych. Dodatkowo, tak jak w przypadku *Privacy Shield* istnieje ryzyko, że dane przekazywane do USA mogą być dostępne dla amerykańskiego wywiadu, co jest sprzeczne z zasadami RODO dotyczącymi ochrony danych. Z tego powodu firmy muszą spełniać złożone wymagania prawne i administracyjne, aby zapewnić zgodność z przepisami zarówno UE, jak i USA, co generuje dodatkowe koszty i ryzyko [16].

2.4.3. Chiny

Najważniejsze akty prawne w Państwie Środka dotyczące prywatności i bezpieczeństwa [17]:

– *Personal Information Protection Law* – PIPL uchwalona w 2021 roku, ma na celu ochronę danych osobowych obywateli Chin. Określa zasady zbierania, przetwarzania, przechowywania i udostępniania danych osobowych. Wymaga zgody na zbieranie danych osobowych, określa obowiązki podmiotów przetwarzających dane oraz prawa jednostek do ochrony ich danych osobowych. Instytucją nadzorującą jest *Cyberspace Administration of China* (CAC).

– *Cybersecurity Law* reguluje ochronę danych w sektorze telekomunikacyjnym oraz ogólnie w Internecie. Wymaga od operatorów sieci i dostawców usług internetowych przestrzegania zasad ochrony danych osobowych i danych użytkowników. Nadzór również sprawuje *Cyberspace Administration of China* (CAC).

– Ustawa o ochronie danych medycznych – Zapewnia ochrona prywatności danych medycznych pacjentów. Określa, jakie dane zdrowotne mogą być przechowywane i przetwarzane przez instytucje medyczne oraz jakie środki bezpieczeństwa należy stosować. Instytucją nadzorującą jest Ministerstwo Zdrowia Chin.

– Ustawa o ochronie danych finansowych ma na celu ochronę danych osobowych klientów instytucji finansowych. Określa zasady zbierania, przetwarzania i przechowywania danych osobowych w sektorze finansowym. Nadzór sprawuje Narodowy Bank Ludowy Chin (*People's Bank of China*).

– Ustawa o ochronie danych osobowych w sektorze edukacyjnym zapewnia ochronę danych osobowych studentów i pracowników w nauczaniu. Określa, jakie informacje mogą być zbierane i przetwarzane przez placówki edukacyjne oraz jakie są obowiązki ochrony danych. Instytucja nadzorująca - Ministerstwo Edukacji Chin.

– Ustawa o ochronie danych osobowych w sektorze transportowym reguluje ochronę danych osobowych w sektorze transportowym, w tym informacje o podróżujących. Określa zasady przetwarzania danych osobowych w transporcie publicznym i prywatnym. Nadzór sprawuje Ministerstwo Transportu Chin.

Dopiero w ostatnich latach Chiny wprowadziły bardziej szczegółowe i zharmonizowane przepisy dotyczące ochrony danych osobowych, co jest reakcją na rosnące wyzwania związane z rozwojem środowiska cyfrowego. Transfer danych osobowych między Unią Europejską a Chinami wiąże się z licznymi wyzwaniami i problemami prawnymi. Chociaż Chiny wprowadziły PIPL, standardy ochrony danych nie są równie rygorystyczne jak te w UE. Dodatkowo istnieje istotne ryzyko, że dane przekazywane do Chin mogą być dostępne dla chińskich władz, co jest sprzeczne z zasadami RODO. Warto zaznaczyć, że w Chinach instytucje nadzorujące, takie jak CAC, mają szerokie uprawnienia w zakresie kontroli i egzekwowania przepisów dotyczących ochrony danych, co może prowadzić do nieprzewidywalnych interwencji i wymogów.

2.4.4. Reszta świata

Warto nadmienić, że ochrona danych osobowych nie ogranicza się jednak do wymienionych państw. Jest to zagadnienie pilnie analizowane i regulowane na całym świecie. Nie tylko w celu dostosowania się do przepisów unijnych czy amerykańskich, aby umożliwić prowadzenie z tymi państwami interesów, ale również ze względu na ochronę swoich obywateli na arenie krajowej. Wymienienie jednak i opisanie wszystkich tego typu aktów prawnych jest zadaniem wręcz niemożliwym, także warto pamiętać, że zagadnienie to jest znacznie szersze niż opisano powyżej.

3. Przegląd rozwiązań zapewniających anonimowość

3.1. Przeglądarki internetowe

W celu korzystania z sieci potrzebnych jest kilka rzeczy. Jedną z najważniejszych jest przeglądarka internetowa. Jest to narzędzie, które uruchamiamy jako pierwsze, chcąc mieć dostęp do Internetu. Działanie przeglądarki jest bardzo proste. Oprogramowanie komunikuje się z serwerem za pomocą protokołu *HTTP* lub szyfrowanego *HTTPS*, po wpisaniu określonego adresu strony internetowej. W tym momencie system pobiera stronę internetową w fizycznej formie, wraz z zawartością. Pliki są automatycznie otwierane i wyświetlane na ekranie docelowego sprzętu. Strony internetowe można wyświetlać też w formie tekstowej (kod *HTML*), zapisywać pliki na dysku twardym, zapamiętywać hasła i loginy do wielu stron itd. Przeglądarka może też przesyłać dane na serwer (np. pliki w chmurę albo cookies zawierające informacje o użytkowniku). Oznacza to, że w praktyce przeglądarka jest niezbędna do sprawnego korzystania z Internetu. Nawet pobranie przeglądarki wymaga dostępu do przeglądarki (oczywiście jest to również możliwe z poziomu *command-line*, ale to dla bardziej zaawansowanych użytkowników). Jest to też jeden z podstawowych programów, który dołączony jest do systemów operacyjnych takich jak Windows od firmy Microsoft czy MacOS dla urządzeń firmy Apple [18].

Przeglądarki potrafią istotnie różnić się lukami w zabezpieczeniach (co można dostrzec choćby po częstotliwości wypuszczania łatek), wpływem na wydajność urządzenia (każda przeglądarka inaczej obciąża procesor czy pamięć) oraz szybkością ładowania stron. Nie wspominając o subiektywnych odczuciach i kwestii wizualnej.

Aby dostarczać lepsze i bardziej spersonalizowane usługi, przeglądarki zbierają różnorodne dane o swoich użytkownikach. Proces ten obejmuje różne techniki i mechanizmy, które mogą być bardziej lub mniej widoczne dla użytkownika. Nie każdemu może się to podobać. Z tego powodu istnieją również rozwiązania takie jak przeglądarka *Brave*, które według zapewnień producentów ograniczają agregację danych o przeglądających. Na ile skuteczne są takie rozwiązania i czy rzeczywiście zwiększają prywatność?

3.2. Tryb prywatny/incognito

Tryb prywatny (znany również jako tryb *incognito* w *Google Chrome*) to funkcja dostępna w większości nowoczesnych przeglądarek internetowych, która umożliwia użytkownikom przeglądanie internetu

bez zapisywania pewnych danych. Innymi słowy jest to tryb przeglądania, w którym przeglądarka nie zapisuje lokalnie danych sesji przeglądania. Dane, które nie są zapisywane lub są usuwane po zakończeniu sesji, obejmują:

- Historia przeglądania – strony internetowe odwiedzone podczas sesji nie są zapisywane w historii przeglądarki.
- Pliki *cookie* i dane stron – wszystkie pliki *cookie* i dane stron są usuwane po zamknięciu okna trybu prywatnego.
- Formularze i dane logowania – informacje wpisywane w formularzach i dane logowania nie są zapisywane.
- Pamięć podręczna (*cache*) – pliki i zasoby ładowane ze stron internetowych nie są przechowywane w pamięci podręcznej przeglądarki.

Rozwiązanie to charakteryzuje się prostotą użytkowania. W celu odpalania wystarczy, aby użytkownik otworzył nowe okno trybu prywatnego/incognito z menu przeglądarki. Przeglądarka wyświetli nową sesję przeglądania w oddzielnym oknie, często oznaczonym jako tryb prywatny. Strony odwiedzone podczas sesji trybu prywatnego są traktowane jako nowe wizyty, więc użytkownik jest wylogowany z wszelkich kont, a strony traktują go jako nowego użytkownika. Wszystkie pliki *cookie* są tymczasowo przechowywane tylko na czas trwania sesji. Formularze i dane logowania nie są zapisywane automatycznie w przeglądarce. Przeglądarka nie zapisuje historii przeglądania ani pamięci podręcznej. Po zamknięciu okna trybu prywatnego przeglądarka usuwa wszystkie tymczasowe dane sesji. Nie pozostają żadne ślady historii przeglądania, plików *cookie* ani danych formularzy [19].

Tryb prywatny bywa przydatny również innych aspektach niż (przynajmniej teoretyczne) zapewnienie prywatności. Jednym z zastosowań jest możliwość testowania stron internetowych bez wpływu na pliki *cookie* i pamięć podręczną przez programistów i twórców. Co ciekawe, wiele stron potrafi manipulować cenami i ofertami w zależności od tego, które odwiedzenie strony przez użytkownika ma miejsce (często oglądamy jakiś konkretny lot to znaczy, że nam na nim zależy, czyli można podnieść cenę, bo i tak go kupimy). Tryb prywatny potrafi zniwelować ten problem i zapewnić przeglądanie ofert oraz cen bez wpływu na sugestie i reklamy oparte na historii przeglądania. Oczywiście wszystko to zostanie dalej przetestowane w pracy.

3.3. AdBlock

AdBlock to popularne rozszerzenie do przeglądarek internetowych, które umożliwia użytkownikom blokowanie reklam wyświetlanych na stronach internetowych. Dostępne są różne wersje *AdBlocka*, takie jak *AdBlock Plus*, *uBlock Origin*, a także oryginalny *AdBlock*. Rozwiązanie to działa na zasadzie filtrowania zawartości strony internetowej, aby ukryć lub usunąć elementy reklamowe.

Instalacja *AdBlocka* jest stosunkowo prosta. Wystarczy zainstalować go jako rozszerzenie do swojej przeglądarki (np. *Google Chrome, Mozilla Firefox, Microsoft Edge*). Po zainstalowaniu automatycznie zaczyna działać podczas przeglądania stron internetowych. Narzędzie korzysta z list filtrowania, które

zawierają reguły blokowania reklam. Listy te są tworzone i aktualizowane przez społeczności użytkowników i organizacje zajmujące się prywatnością. Kiedy użytkownik odwiedza stronę internetową, *AdBlock* analizuje kod strony w poszukiwaniu elementów pasujących do reguł blokowania zawartych w listach filtrowania. Elementy reklamowe, takie jak bannery, *pop-upy*, śledzące skrypty i inne formy reklam, są ukrywane lub usuwane, zanim strona zostanie wyświetlona użytkownikowi. Użytkownicy mogą dostosowywać działanie *AdBlocka*, dodając własne reguły blokowania lub białe listy (*whitelist*) stron, na których reklamy mają być wyświetlane. Możliwe jest również blokowanie konkretnych elementów na stronie poprzez interaktywne narzędzie dostępne w *AdBlocku*. W teorii istnieje wiele korzyści korzystania z *AdBlocka* (oczywiście zostaną one w dalszej części pracy poddane analizie):

– Szybsze ładowanie stron – blokowanie reklam zmniejsza ilość pobieranych danych, co przyspiesza ładowanie stron internetowych.

– Mniejsza liczba przeszkadzających reklam: Użytkownicy mogą przeglądać strony internetowe bez uciążliwych reklam, które mogą przeszkadzać w czytaniu treści.

– Ochrona prywatności – blokowanie śledzących skryptów i trackerów reklamowych pomaga chronić prywatność użytkowników przed niechcianym śledzeniem online.

– Oszczędność danych – mniej reklam oznacza mniej danych do pobrania, co jest korzystne dla użytkowników z ograniczonym dostępem do internetu lub korzystających z mobilnych planów danych.

Blokowanie reklam może wpływać na dochody twórców treści internetowych, którzy polegają na przychodach z reklam do finansowania swojej działalności. Niektóre wersje *AdBlocka*, takie jak *AdBlock Plus*, mają program *„Acceptable Ads"*, który pozwala na wyświetlanie nieinwazyjnych reklam od partnerów, aby wspierać twórców treści. Dodatkowo niektóre strony internetowe wykrywają użycie *AdBlocka* i proszą użytkowników o jego wyłączenie lub wprowadzają systemy subskrypcji dla dostępu do treści bez reklam.

AdBlock to narzędzie, które pozwala użytkownikom na blokowanie reklam i ochronę prywatności podczas przeglądania internetu. Dzięki wykorzystaniu list filtrowania i reguł blokowania, *AdBlock* usuwa reklamy i śledzące skrypty z przeglądanych stron. Choć oferuje wiele korzyści, takich jak szybsze ładowanie stron i mniejsza ilość reklam, ma również swoje wyzwania, zwłaszcza w kontekście wpływu na twórców treści internetowych.

3.4. Proxy

Proxy (serwer *proxy*) to pośrednik między urządzeniem użytkownika a docelowym serwerem internetowym. Działa jako pośrednik w przekazywaniu żądań *HTTP* (lub innych protokołów) między klientem (np. przeglądarką internetową) a serwerem docelowym (np. stroną internetową). *Proxy* może być zarówno serwerem sprzętowym, jak i oprogramowaniem działającym na serwerze. Działa jako brama, przekazując żądania od użytkownika do zasobów internetowych i z powrotem. Serwery *proxy* są

używane z różnych powodów, takich jak zwiększenie prywatności, bezpieczeństwa, wydajności sieci, a także do obejścia ograniczeń geograficznych lub cenzury.

Sposób działania *Proxy* jest następujący:

- Użytkownik wysyła żądanie do serwera *proxy* zamiast bezpośrednio do docelowego serwera internetowego.
- Serwer *proxy* odbiera żądanie i przekazuje je do docelowego serwera.
- Serwer *proxy* odbiera odpowiedź od docelowego serwera i przekazuje ją z powrotem do użytkownika.

Wyróżniamy kilka rodzajów *Proxy*:

- *Forward Proxy* – działa jako pośrednik między użytkownikiem a serwerem internetowym. Jest używany do ukrywania tożsamości użytkownika i zapewniania dostępu do zablokowanych treści.
- *Reverse Proxy* – działa jako pośrednik między internetem a serwerami wewnętrznymi organizacji. Jest używany do równoważenia obciążenia, buforowania i zwiększania bezpieczeństwa.
- *Transparent Proxy* przekazuje żądania bez ukrywania tożsamości użytkownika. Jest często używany do filtrowania treści i monitorowania ruchu internetowego.
- *Anonymous Proxy* – ukrywa tożsamość użytkownika, ale ujawnia, że żądanie pochodzi z serwera *proxy*.
- *High Anonymity Proxy* – ukrywa zarówno tożsamość użytkownika, jak i fakt, że żądanie pochodzi z serwera proxy.

Proxy to potężne narzędzie, które pozwala użytkownikom na zwiększenie prywatności, bezpieczeństwa i wydajności podczas korzystania z internetu. Działa jako pośrednik między urządzeniem użytkownika a docelowym serwerem internetowym, oferując różnorodne funkcje, takie jak ukrywanie adresu IP, omijanie ograniczeń geograficznych i cenzury. Mimo wielu korzyści, *proxy* może mieć również pewne ograniczenia, takie jak potencjalne spowolnienie prędkości internetu i ryzyko związane z korzystaniem z niezaufanych serwerów.

3.5. VPN

VPN (*Virtual Private Network*) to usługa, która pozwala użytkownikom na stworzenie zaszyfrowanego połączenia do serwera *VPN*, który następnie łączy się z internetem w ich imieniu. Dzięki temu rzeczywisty adres IP użytkownika jest ukrywany, a wszystkie dane przesyłane między jego urządzeniem a internetem są zaszyfrowane.

Dokładne działanie *VPN* wygląda następująco. Użytkownik uruchamia oprogramowanie *VPN* na swoim urządzeniu (komputerze, smartfonie, tablecie). Oprogramowanie tworzy bezpieczne połączenie

z serwerem *VPN*, który może znajdować się w dowolnym miejscu na świecie. Wszystkie dane przesyłane między urządzeniem użytkownika a serwerem *VPN* są szyfrowane. To oznacza, że nawet jeśli ktoś przechwyci te dane, nie będzie w stanie ich odczytać. Po połączeniu z serwerem *VPN*, adres IP użytkownika jest zastępowany adresem IP serwera *VPN*. To utrudnia śledzenie aktywności użytkownika w sieci i umożliwia dostęp do zasobów ograniczonych geograficznie. *VPN* pozwala na ominięcie cenzury internetowej w krajach, gdzie dostęp do pewnych treści jest ograniczony przez rząd. Zaszyfrowane dane są przesyłane przez serwer *VPN* do docelowego serwera internetowego (np. strony internetowej). Odpowiedź z tego serwera jest również przesyłana przez serwer *VPN*, szyfrowana, a następnie odszyfrowywana na urządzeniu użytkownika. Jest również często używany przez firmy do zapewnienia bezpiecznego połączenia zdalnego dla pracowników, umożliwiając im dostęp do zasobów firmowych z dowolnego miejsca na świecie.

VPN ukrywa rzeczywisty adres IP użytkownika, co utrudnia śledzenie jego aktywności online przez strony internetowe, dostawców usług internetowych (*ISP*) i potencjalnych hakerów. *VPN* chroni dane użytkownika przed przechwyceniem przez osoby trzecie, szczególnie w przypadku korzystania z publicznych sieci *Wi-Fi*, które są bardziej podatne na ataki.

Można wyszczególnić następujące typy *VPN*:

– *Remote Access VPN* – umożliwia użytkownikom zdalne łączenie się z prywatną siecią (np. firmową siecią lokalną) za pomocą zaszyfrowanego połączenia.

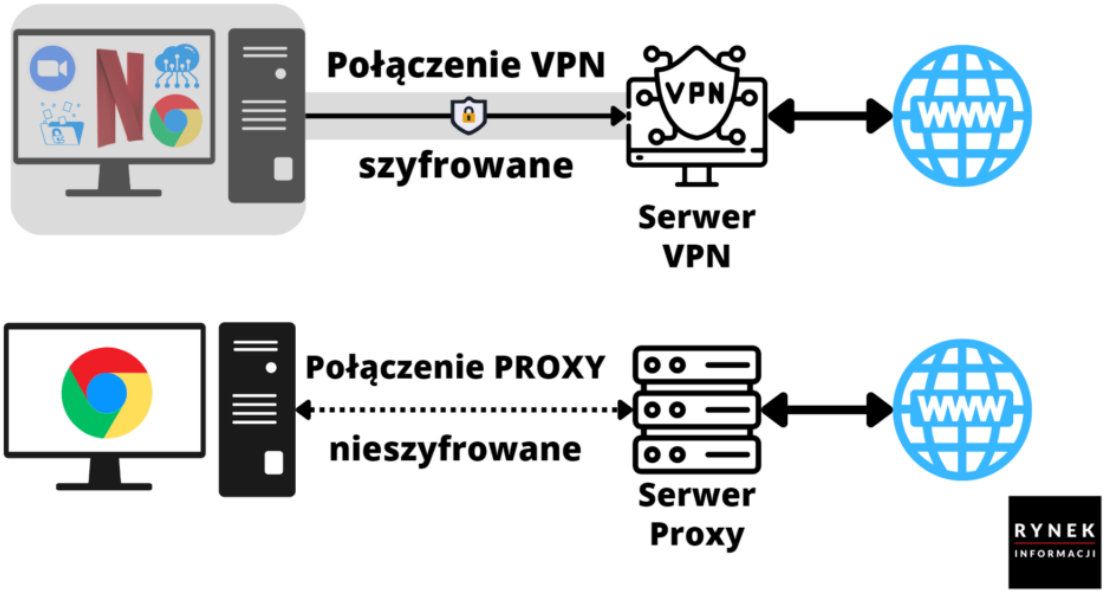
– *Site-to-Site VPN* – łączy dwie sieci lokalne w różnych lokalizacjach, umożliwiając bezpieczną komunikację między nimi.

– *Client-to-Site VPN* – pozwala pojedynczym użytkownikom na połączenie się z siecią firmową zdalnie, zwykle za pomocą oprogramowania *VPN* na ich urządzeniach.

3.5.1. Różnica między VPN, a Proxy

Choć oba rozwiązania wydają się podobne, to istnieje miedzy nimi parę różnic. Oba rozwiązania służą do nawiązywania połączenia między klientem a serwerem poprzez przekierowanie ruchu przez węzeł pośredni. Różnią się nieco sposobem obsługiwania połączenia i przesyłania danych. Warto zapoznać się z tymi różnicami zwłaszcza w kontekście prywatności i bezpieczeństwa.

VPN oferuje kilka kluczowych funkcji bezpieczeństwa. Po pierwsze, użytkownik zyskuje unikalny adres IP oraz bezpieczny tunel do komunikacji między urządzeniem a serwerem, co umożliwia szyfrowanie ruchu dla wszystkich protokołów internetowych. *VPN* jest łatwy w obsłudze, ponieważ można go szybko włączyć i wyłączyć, przechodząc tym samym z szyfrowanego połączenia na standardowe połączenie internetowe. Dzięki *VPN* dane użytkownika są chronione przed przechwyceniem, nawet podczas korzystania z otwartych sieci *WiFi*. Szyfrowanie *AES-256* oraz stosowanie certyfikatów zapewniają wysoki poziom bezpieczeństwa i prywatności, chroniąc dane przed dostępem dostawców internetu, instytucji państwowych i organów ścigania. *VPN* pozwala również omijać blokady geograficzne oraz uzyskiwać dostęp do treści niedostępnych w danej lokalizacji.



Rys. 3.1. Graficzne porównanie technologii *Proxy* oraz *VPN* [20]

Z drugiej strony, serwery *proxy* działają jako pośrednicy w połączeniach internetowych, ukrywając adres IP użytkownika, ale nie szyfrują przesyłanych danych. Wykorzystują one standardowe protokoły *HTTP* lub *SOCKS*. *Proxy* również umożliwia dostęp do stron blokowanych ze względu na lokalizację użytkownika.

VPN jest idealny do pracy zdalnej, bezpiecznego przeglądania internetu z publicznych hotspotów, korzystania z serwisów zablokowanych w danym kraju, a także do bezpiecznych połączeń *peer-to-peer*. Proxy natomiast dobrze sprawdzi się do omijania blokad narzuconych przez dostawców usług internetowych, ochrony infrastruktury wewnętrznej, przyspieszenia ładowania stron dzięki lokalnej *pamięci cache*, a także do ukrywania adresu IP przed odwiedzanymi stronami.

Podsumowując, *VPN* zapewnia wyższy poziom bezpieczeństwa dzięki szyfrowaniu danych, podczas gdy proxy oferuje jedynie anonimizację ruchu co wyraźnie widać na rysunku 3.1. Oba rozwiązania mają swoje wady i zalety, dlatego ważne jest, aby wybrać odpowiednie narzędzie w zależności od potrzeb. Należy pamiętać, że zarówno mało zaufane sieci *VPN*, jak i publiczne serwery proxy mogą stanowić zagrożenie dla bezpieczeństwa, a *VPN* wpłynąć na prędkość połączenia.

3.6. Trasowanie cebulowe - TOR

Trasowanie cebulowe, znane również jako *Tor* (*The Onion Router*), to technologia i sieć, która umożliwia anonimowe przeglądanie internetu. *Tor* jest zaprojektowany tak, aby chronić prywatność i anonimowość użytkowników, ukrywając ich lokalizację i działania w sieci przed monitorowaniem i analizą ruchu.

Tor działa poprzez wielowarstwowe szyfrowanie i przesyłanie danych przez serię wolontariuszy (serwerów), znanych jako węzły *Tor*. Proces ten jest analogiczny do warstw cebuli, stąd nazwa „trasowanie cebulowe".

Dane użytkownika są wielokrotnie szyfrowane zanim opuszczą jego urządzenie. Każda warstwa szyfrowania jest usuwana przez kolejny węzeł Tor, ujawniając instrukcje dotyczące następnego węzła, przez który mają przejść dane. Następnie dane przechodzą przez serię węzłów Tor (zwykle trzy węzły: wejściowy, pośredni i wyjściowy). Każdy węzeł odszyfrowuje tylko jedną warstwę szyfrowania, znając jedynie poprzedni i następny węzeł w sieci. Węzeł wejściowy zna jedynie prawdziwy adres IP użytkownika, ale nie zna ostatecznego celu danych. Węzeł wyjściowy zna tylko ostateczny cel danych, ale nie zna prawdziwego adresu IP użytkownika. Ponieważ każdy węzeł zna tylko swojego bezpośredniego poprzednika i następnika, żadna pojedyncza jednostka nie ma pełnego obrazu trasy danych. To utrudnia śledzenie aktywności użytkownika w sieci [21].

Korzyści z Używania Tor:

– Anonimowość – umożliwia użytkownikom ukrycie swojej tożsamości online, chroniąc ich przed śledzeniem przez rządy, dostawców usług internetowych i strony internetowe.

– Dostęp do zasobów ograniczonych geograficznie – umożliwia użytkownikom dostęp do stron internetowych i usług, które mogą być ograniczone w ich regionie.

– Ochrona prywatności – chroni przed inwigilacją i analizą ruchu, utrudniając identyfikację i śledzenie użytkowników.

– Bezpieczeństwo komunikacji – pomaga chronić komunikację przed przechwyceniem przez osoby trzecie, co jest szczególnie istotne dla dziennikarzy, aktywistów i osób żyjących w reżimach autorytarnych.

Warto zwrócić uwagę, że rozwiązanie to może negatywnie wpływać na wydajność. Sieć *Tor* może być wolniejsza niż tradycyjne połączenia internetowe, ze względu na wielokrotne szyfrowanie i przesyłanie danych przez liczne węzły. Dodatkowo, niektóre strony internetowe mogą blokować ruch pochodzący z węzłów *Tor*, ograniczając dostęp użytkowników do pewnych zasobów. Chociaż *Tor* chroni przed wieloma formami inwigilacji, nie jest odporny na wszystkie ataki. Węzły wyjściowe mogą być potencjalnie złośliwe i monitorować ruch, choć nie mogą łatwo zidentyfikować użytkowników. *Tor* jest czasem używany do ukrywania nielegalnej działalności, co przyciąga negatywną uwagę i może prowadzić do prób deanonimizacji sieci przez organy ścigania (choćby poprzez kontrolę pewnej liczby węzłów wyjściowych) [22].

Najprostszym sposobem korzystania z *Tor* jest użycie *Tor Browser*, zmodyfikowanej wersji przeglądarki *Firefox*, która jest skonfigurowana do korzystania z sieci Tor. Zaawansowani użytkownicy mogą skonfigurować swoje urządzenia do korzystania z sieci Tor za pomocą oprogramowania, takiego jak *Tor Project* i narzędzi do konfiguracji sieci. *Tor* umożliwia również dostęp do ukrytych usług, które są dostępne tylko w sieci *Tor*, oferując dodatkowy poziom anonimowości zarówno dla dostawców usług, jak i ich użytkowników.

3.7. Komunikatory

Komunikatory internetowe stanowią szeroko stosowane narzędzia do komunikacji tekstowej, głosowej i wideo. Istnieje wiele różnych komunikatorów, z których niektóre są uznawane za bardziej bezpieczne niż inne, z uwagi na różne funkcje zapewniania prywatności i bezpieczeństwa.

Czym charakteryzują się bezpieczne komunikatory [23]:

– Szyfrowanie *end-to-end* – zapewnia, że treść wiadomości jest zaszyfrowana na urządzeniu nadawcy i odszyfrowana tylko na urządzeniu odbiorcy.

– Brak zbierania metadanych – komunikatory bardziej bezpieczne starają się ograniczyć zbieranie metadanych o użytkownikach, takich jak czas trwania rozmowy czy częstotliwość kontaktów.

– Mechanizmy autoryzacji i uwierzytelniania – zapewniają, że tylko uprawnione osoby mogą uzyskać dostęp do konta i przesyłać wiadomości.

Popularne komunikatory:

– *WhatsApp* – popularny komunikator należący do *Mety*, który oferuje szyfrowanie *end-to-end*, ale gromadzi metadane o użytkownikach.

– *Facebook Messenger* – bardzo popularny komunikator posiadający integrację z *Facebookiem*.

– *Telegram* – kontrowersyjny ze względu na swoje rosyjskie pochodzenie i wykorzystanie przez grupy terrorystyczne.

Bezpieczne komunikatory są szczególnie ważne w dobie wzrastającej świadomości prywatności cyfrowej i troski o ochronę danych osobowych. Wybór odpowiedniego komunikatora zależy od indywidualnych potrzeb bezpieczeństwa i funkcjonalności.

3.8. E-mail

E-maile są powszechnym narzędziem komunikacji w dzisiejszym świecie cyfrowym. Każdego dnia miliony ludzi wysyłają i odbierają wiadomości e-mail, często zawierające prywatne informacje, dane osobowe czy ważne dokumenty. Porządne rozwiązania używają protokołu szyfrowania, takiego jak *SSL/TLS*, aby zabezpieczyć transmisję między klientem poczty a serwerem. Istotnym aspektem bezpieczeństwa skrzynek są podejrzane wiadomości e-mail i linki. Dobrą praktyką jest nigdy nie otwierać załączników ani nie klikać na linki od nieznanych nadawców. Z innych podstawowych zasad nie wolno przesyłać wrażliwych danych, takich jak numery kont bankowych czy hasła, w niezabezpieczonych wiadomościach e-mail. Wybór renomowanego klienta poczty, który oferuje zaawansowane funkcje bezpieczeństwa, takie jak filtrowanie spamu i ochrona antywirusowa może również ochronić przed nieprzewidzianymi problemami.

Bezpieczeństwo e-maili jest kluczowe dla ochrony prywatności i danych osobowych. Stosowanie silnych haseł, dwuskładnikowej weryfikacji, szyfrowania wiadomości oraz ostrożność w otwieraniu podejrzanych wiadomości to podstawowe kroki, które pomogą zabezpieczyć Twoje konto e-mail przed atakami i nieautoryzowanym dostępem. Regularne aktualizacje oprogramowania i świadomość zagrożeń cybernetycznych są również ważne dla utrzymania bezpieczeństwa w korzystaniu z e-maili.

3.8.1. Tymczasowy adres e-mail

Temp maile, znane również jako adresy jednorazowe lub adresy e-mail z terminem ważności, są narzędziem używanym do ochrony prywatności użytkownika w Internecie. Umożliwiają one tworzenie tymczasowych adresów e-mail, które można wykorzystać do rejestracji na stronach internetowych, subskrypcji newsletterów, czy w sytuacjach, gdzie nie chcemy ujawniać swojego głównego adresu e-mail.

Działanie tymczasowych adresów e-mail wygląda następująco:

– Użytkownik odwiedza stronę internetową lub korzysta z aplikacji oferującej usługę tymczasowego e-maila. Tam generowany jest unikalny adres e-mail, który można od razu wykorzystać.

– Wszystkie wiadomości wysłane na ten tymczasowy adres są dostępne na stronie dostawcy usługi lub mogą być przekazywane na główny adres e-mail użytkownika. Wiadomości są zazwyczaj przechowywane przez krótki okres, np. od kilku minut do kilku godzin.

– Po upływie określonego czasu tymczasowy adres e-mail oraz wszystkie wiadomości do niego wysłane są automatycznie usuwane. To eliminuje ryzyko długoterminowego przechowywania danych.

Tymczasowe adresy e-mail pomagają chronić główny adres e-mail użytkownika przed spamem, phishingiem oraz innymi zagrożeniami związanymi z ujawnieniem adresu w Internecie. Dzięki krótkoterminowej naturze, ryzyko przechwycenia lub nieautoryzowanego dostępu do wiadomości jest minimalizowane. Jest idealnym rozwiązaniem do rejestracji na stronach internetowych, które mogą wysyłać dużą ilość niechcianych wiadomości. Używanie tymczasowych adresów pozwala na zachowanie anonimowości, co jest szczególnie przydatne przy korzystaniu z usług, które mogą wymagać weryfikacji e-maila, ale nie są zaufane.

Wszystkie wiadomości są usuwane po określonym czasie, co może być problematyczne, jeśli potrzebujemy dostępu do wiadomości w przyszłości. Tymczasowe adresy e-mail zazwyczaj nie oferują zaawansowanych funkcji, takich jak filtry wiadomości, foldery czy synchronizacja z innymi usługami e-mail. Niektóre tymczasowe adresy e-mail są publicznie dostępne. Każdy, kto zna adres, może odczytać otrzymane wiadomości.

3.9. Hasła i ich przechowywanie

Hasła jak zostało wspomniane wyżej są kluczowym elementem ochrony cyfrowej tożsamości użytkowników. Są pierwszą linią obrony przed nieautoryzowanym dostępem do kont online, systemów komputerowych i danych osobowych. Jednakże, wraz z rosnącą liczbą usług wymagających uwierzytelniania, zarządzanie i bezpieczne przechowywanie haseł stało się wyzwaniem. Poniżej omówiono najlepsze praktyki i technologie związane z hasłami oraz ich przechowywaniem.

Tworzenie silnych haseł wymaga spełnienia paru kluczowych punktów. Cechy silnego hasła [24]:

- Długość – minimum 12 znaków, im dłuższe, tym lepsze.
- Złożoność – użycie kombinacji małych i wielkich liter, cyfr oraz znaków specjalnych.
- Brak powtarzalności – unikanie łatwo przewidywalnych wzorców, słów ze słownika, imion czy dat urodzenia.
- Unikalność – każde konto powinno mieć unikalne hasło, aby zminimalizować ryzyko, że jedno złamane hasło da dostęp do wielu kont.

Oprócz stworzenia porządnego hasła istotne jest przechowywanie haseł, nawet najmocniejsze hasło nie ma znaczenia w sytuacji, gdy zostanie ono ujawnione. Z tego powodu ważne jest stosowanie najlepszych praktyk przechowywania haseł:

- Unikanie przechowywania w formie tekstu jawnego – nigdy nie przechowuj haseł w niezabezpieczonych plikach tekstowych.
- Używanie menedżerów haseł – narzędzia te przechowują hasła w zaszyfrowanej bazie danych, dostępnej za pomocą głównego hasła.
- Haszowanie haseł – serwery przechowujące hasła powinny używać algorytmów skrótu (np. *bcrypt*, *scrypt*, *Argon2*) zamiast przechowywania haseł w postaci jawnej.

Jednak pomimo zastosowania wszystkich powyższych punktów wyciek może się zdarzyć. Niekoniecznie z winy użytkownika, a dostarczyciela usług czy osób trzecich. W tym celu zabezpiecza się hasła. Do najpopularniejszych metod należą:

- Haszowanie – proces przekształcania hasła w unikalny ciąg znaków za pomocą funkcji mieszającej. Nawet mała zmiana w haśle powoduje całkowicie inny wynik haszowania.
- *Salting* – dodawanie unikalnego, losowego ciągu znaków (*salt*) do hasła przed jego haszowaniem. Chroni przed atakami słownikowymi i *rainbow tables*. *Salt* powinien być unikalny dla każdego hasła i przechowywany wraz z haszem.
- *Peppering* – dodawanie dodatkowego, tajnego ciągu znaków (*pepper*) do hasła przed haszowaniem. *Pepper* jest tajny i nie przechowuje się go razem z haszem.

Dodatkowo warto pamiętać o bezpiecznych metodach używania haseł:

– Dwuskładnikowe uwierzytelnianie (2FA) – dodanie drugiej warstwy zabezpieczeń, zwykle w postaci kodu wysłanego na telefon komórkowy lub generowanego przez aplikację (np. *Google Authenticator*).

– Regularna zmiana haseł – regularne zmienianie haseł, szczególnie w przypadku podejrzenia naruszenia bezpieczeństwa.

– Monitorowanie bezpieczeństwa kont – korzystanie z usług monitorujących, które informują o potencjalnych wyciekach haseł.

3.10. Kryptowaluty

Kryptowaluty to cyfrowe lub wirtualne waluty, które wykorzystują kryptografię do zabezpieczania transakcji, kontrolowania tworzenia nowych jednostek oraz potwierdzania transferu aktywów. Najbardziej znaną kryptowalutą jest *Bitcoin*, który został wprowadzony w 2009 roku przez osobę lub grupę osób pod pseudonimem Satoshi Nakamoto. *Blockchain* to rozproszona księga (*ledger*), która rejestruje wszystkie transakcje w sieci kryptowaluty. Jest to baza danych współdzielona przez uczestników sieci, zapewniająca transparentność i bezpieczeństwo. *Blockchain* składa się z bloków, które są połączone w łańcuch za pomocą kryptograficznych haszów. Sposób działania *blockchainu* jest intrygujący i zdecydowanie odmienny od standardowego systemu bankowego. Każdy blok w *blockchainie* zawiera listę transakcji. Blok składa się z nagłówka bloku i transakcji. Nagłówek bloku zawiera metadane, takie jak *hash* poprzedniego bloku, znacznik czasu i *nonce*. Każdy blok jest identyfikowany przez unikalny *hash*, który jest generowany przy użyciu kryptograficznej funkcji haszującej. *Hash* poprzedniego bloku jest zawarty w nagłówku kolejnego bloku, tworząc łańcuch bloków. Sieć *blockchain* opiera się na konsensusie w celu zatwierdzania nowych bloków. W przypadku *Bitcoina* stosowany jest algorytm *Proof of Work* (PoW), który wymaga, aby górnicy rozwiązywali skomplikowane problemy matematyczne, aby dodać nowy blok do łańcucha. Raz dodany blok nie może być zmieniony bez modyfikacji wszystkich kolejnych bloków, co wymaga konsensusu większości uczestników sieci. To zabezpiecza *blockchain* przed manipulacjami. Dzięki rozproszonej naturze i kryptograficznemu zabezpieczeniu *blockchain* jest bardzo odporny na ataki i oszustwa. Wszystkie transakcje są publicznie dostępne i mogą być weryfikowane przez uczestników sieci. Transakcje zapisane w *blockchainie* są trwałe i niezmienne, co zapewnia integralność danych. *Blockchain* eliminuje potrzebę centralnych pośredników, takich jak banki [25].

Problemem kryptowalut są wciąż niejasne i kształtujące się regulacje. Rządy na całym świecie pracują nad regulacjami, co może wpłynąć na ich przyszłość. Regulacje mogą wprowadzić stabilność, ale także ograniczyć anonimowość i decentralizację. Aktualne *blockchainy* mają problemy ze skalowalnością, co oznacza, że mogą obsługiwać ograniczoną liczbę transakcji na sekundę. Projekty takie jak *Ethereum 2.0* i rozwój sieci *Lightning* dla *Bitcoina* starają się rozwiązać te problemy. Pomimo rosnącej popularności, adopcja kryptowalut przez użytkowników końcowych i przedsiębiorstwa jest nadal w stosunkowo wczesnym etapie. Rozwój infrastruktury, takiej jak portfele kryptowalutowe i giełdy, jest

kluczowy dla dalszej adopcji. Pomimo wysokiego poziomu bezpieczeństwa, *blockchainy* nie są całkowicie odporne na ataki. Znane są przypadki włamań do giełd kryptowalutowych oraz ataków typu *„51% attack"*, gdzie złośliwy aktor przejmuje kontrolę nad większością mocy obliczeniowej sieci.

Kryptowaluty i *blockchain* mają potencjał do zrewolucjonizowania wielu aspektów naszego życia, od finansów po zarządzanie łańcuchem dostaw. Pomimo wyzwań, jakie stoją przed tymi technologiami, ich rozwój i adopcja są dynamiczne i pełne możliwości.

Warto jednak wyróżnić jedną niewymienioną szczególna kryptowalutę. Jest nią Monero. Oczywiście może istnieć wiele jeszcze lepszych rozwiązań, ale ta, ze względu na swoją popularność, a co za tym idzie w świecie kryptowalut, użytkowość, jest warta omówienia.

Monero (XMR) to kryptowaluta, która skupia się na zapewnieniu pełnej prywatności i anonimowości transakcji. Wprowadzona w 2014 roku, *Monero* różni się od wielu innych kryptowalut poprzez swoje zaawansowane technologie kryptograficzne, które ukrywają nadawców, odbiorców oraz kwoty transakcji.

Do kluczowych funkcji *Monero* należą:

– *Ring Signatures* (podpisy pierścieniowe) – ta technologia miesza podpis cyfrowy nadawcy z podpisami innych użytkowników *Monero*, co sprawia, że nie można zidentyfikować, kto faktycznie zainicjował transakcję. Dzięki temu, nadawca transakcji pozostaje anonimowy.

– *Stealth Addresses* (ukryte adresy) – ukryte adresy są jednorazowymi adresami generowanymi dla każdej transakcji, co uniemożliwia powiązanie transakcji z konkretnym odbiorcą. Nawet jeśli ktoś zna adres publiczny odbiorcy, nie jest w stanie zobaczyć, jakie transakcje zostały do niego wysłane.

– *RingCT (Ring Confidential Transactions)* – technologia *RingCT* ukrywa kwoty transakcji, zapewniając dodatkową warstwę prywatności. Kwoty transakcji są widoczne tylko dla stron biorących udział w transakcji, co eliminuje możliwość śledzenia przepływu środków.

– *Dandelion++* – protokół ten poprawia prywatność na poziomie sieci, ukrywając źródłowy adres IP użytkownika, który inicjuje transakcję. Jest to istotne zabezpieczenie przed atakami, które mogłyby śledzić lokalizację użytkownika.

Monero jest oparty na protokole *CryptoNote*, który został zaprojektowany z myślą o prywatności. Monero używa kryptografii krzywych eliptycznych (ECC) oraz algorytmu *Proof of Work*(PoW), aby zabezpieczyć swoją sieć i umożliwić wydobycie (*mining*).

Działanie *Monero* wygląda następująco. Kiedy użytkownik wysyła *Monero*, jego portfel tworzy ukryty adres dla odbiorcy i generuje podpis pierścieniowy, aby zamaskować tożsamość nadawcy. Kwota transakcji jest ukrywana za pomocą *RingCT*. Górnicy w sieci *Monero* weryfikują transakcje, rozwiązując skomplikowane problemy matematyczne. Podpisy pierścieniowe i *RingCT* są używane, aby zapewnić, że transakcja jest ważna, bez ujawniania szczegółów dotyczących nadawcy, odbiorcy i kwoty. Zatwierdzone transakcje są dodawane do *blockchaina Monero*, który jest publiczny, ale dzięki zaawansowanym technologiom kryptograficznym, szczegóły transakcji pozostają ukryte [26].

3.11. zFone

ZFone, znany również jako *ZRTP* (*Zimmermann Real-Time Protocol*), to protokół kryptograficzny służący do uzgadniania kluczy, zaprojektowany w celu zabezpieczenia komunikacji głosowej przez IP (*VoIP*). Został opracowany przez Phila Zimmermanna, twórcę *PGP* (*Pretty Good Privacy*), i jest specjalnie ukierunkowany na zapewnienie szyfrowania *end-to-end* dla połączeń VoIP.

Voice over Internet Protocol to technologia przesyłania głosu przez internet, zamieniająca dźwięk na cyfrowe pakiety danych. Umożliwia prowadzenie rozmów bez tradycyjnych linii telefonicznych, oferując tańszą i bardziej elastyczną komunikację. Często jest zintegrowana z platformami do wideokonferencji, takimi jak *Zoom* czy *Microsoft Teams*.

Kluczowe cechy technologii *ZFone* obejmują [27]:

- Szyfrowanie *End-to-End* – *ZFone* zapewnia, że komunikacja między dwiema stronami jest szyfrowana *end-to-end*, co oznacza, że tylko zamierzone strony mogą odszyfrować i zrozumieć treść komunikacji.
- Negocjacja kluczy – protokół umożliwia bezpieczną negocjację kluczy między stronami uczestniczącymi w rozmowie, zapewniając, że klucze szyfrujące są wymieniane bez narażenia na przechwycenie lub manipulację.
- Ochrona przed atakami *Man-in-the-Middle* – *ZFone* jest zaprojektowany, aby zapobiegać atakom typu *man-in-the-middle*, gdzie trzecia strona przechwytuje i potencjalnie zmienia komunikację między dwiema stronami.
- Otwarty standard – *ZFone* opiera się na otwartych standardach kryptograficznych, co oznacza, że jego specyfikacje i implementacje są otwarte do publicznej weryfikacji i analizy.
- Kompatybilność – *ZFone* może być zintegrowany z różnymi aplikacjami i platformami *VoIP*, co czyni go wszechstronnym rozwiązaniem dla różnych potrzeb komunikacyjnych.
- Uwierzytelnianie – zawiera mechanizmy weryfikacji tożsamości stron uczestniczących w komunikacji, pomagając zapewnić, że komunikacja nie jest fałszowana ani przechwytywana przez nieuprawnione strony.

4. Aplikacja do gromadzenia i prezentowania danych

4.1. Wymagania wstępne

Celem aplikacji było stworzenie narzędzia do mierzenia wskaźników wydajności oraz określania poziomu zbieranych informacji w trakcie korzystania z internetu przy pomocy technologii mających zapewnić zwiększony poziom ochrony prywatności. W założeniu miała to być aplikacja konsolowa, która oprócz sprawdzenia wyników, zapisuje je do lokalnej bazy danych oraz potrafi je graficznie zaprezentować.

4.2. Wykorzystane technologie

4.2.1. Python

Python to wszechstronny język programowania, który zyskał ogromną popularność dzięki swojej prostocie, czytelności oraz szerokim możliwościom zastosowania. Jego składnia jest intuicyjna, co czyni go idealnym wyborem zarówno dla początkujących programistów, jak i doświadczonych deweloperów. Dzięki bogatej bibliotece standardowej oraz ogromnej liczbie dostępnych pakietów, *Python* znajduje zastosowanie w takich dziedzinach jak analiza danych, uczenie maszynowe, rozwój aplikacji webowych, automatyzacja zadań, a nawet tworzenie gier. Jego otwartość i rozbudowane wsparcie społeczności sprawiają, że *Python* jest regularnie aktualizowany i rozwijany, co gwarantuje dostęp do najnowszych technologii i narzędzi. Istotnym atutem przy jego wyborze do napisania tej aplikacja była dostępność biblioteki *matplotlib*, która w prosty sposób pozwala na wizualizację danych oraz innych bibliotek umożliwiających analizę zapytań [28].

4.2.2. Selenium

Selenium to popularne narzędzie do automatyzacji testów aplikacji webowych, które umożliwia programistom oraz testerom automatyzację interakcji z przeglądarką internetową. Dzięki *Selenium* można symulować działania użytkownika, takie jak klikanie, wprowadzanie danych do formularzy czy nawigacja między stronami, co pozwala na kompleksowe testowanie aplikacji w różnych warunkach. *Selenium* wspiera wiele przeglądarek, takich jak *Chrome*, *Firefox* czy *Safari*, a także działa na różnych systemach

operacyjnych, co czyni je wszechstronnym rozwiązaniem dla zespołów deweloperskich. Kluczową zaletą *Selenium* jest możliwość integracji z różnymi językami programowania, w tym *Pythonem*, *Javą*, *C++* i innymi, co pozwala na łatwe włączenie testów automatycznych do istniejących projektów. Pozwala na wielopłaszczyznowe sprawdzenie zapytań i analizę odpowiedzi [29].

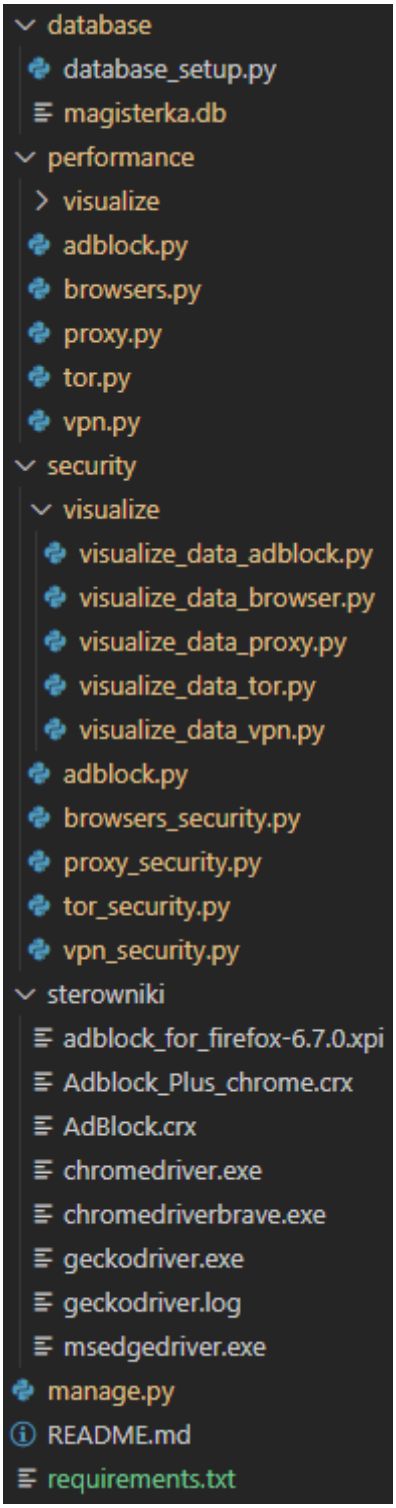
4.2.3. SQLite

SQLite to lekka, wbudowana baza danych SQL, która jest powszechnie stosowana w aplikacjach, gdzie potrzeba zarządzania danymi nie wymaga pełnoskalowego serwera bazodanowego. *SQLite* jest samodzielną biblioteką, która przechowuje całą bazę danych w postaci pojedynczego pliku na dysku, co czyni ją wyjątkowo prostą w instalacji, użytkowaniu i przenoszeniu. Dzięki niewielkim wymaganiom zasobów oraz szybkiemu działaniu, *SQLite* jest idealnym rozwiązaniem dla aplikacji mobilnych, stron internetowych, narzędzi analitycznych, a także prototypownia i testowania. Baza danych *SQLite* obsługuje pełny zestaw funkcji SQL, takich jak transakcje, zapytania, indeksowanie i integrację z językami programowania, co sprawia, że mimo swojej prostoty, jest niezwykle wszechstronnym narzędziem. Dodatkowo, *SQLite* jest rozwiązaniem *open source*, co oznacza, że jest darmowy do użycia zarówno w projektach komercyjnych, jak i niekomercyjnych [30]. Prostota tego narzędzia była głównym kryterium doboru do projektu.

4.3. Architektura

Implementacja aplikacji zawiera się w trzech katalogach, *database*, *security* i *sterowniki*, mieszczących się w katalogu głównym, w którym znajduje się również plik *manage.py*, służący do wywoływania aplikacji. Importuje on wszystkie skrypty testujące. Zaimplementowano w nim wybór odpowiedniego narzędzia i przystosowanego do niego testu przy pomocy prostych opcji dodawanych w trakcie uruchamiania pliku.

Katalog *database* zawiera bazę danych oraz plik konfiguracyjny w celu jej stworzenia *database_setup.py*. W katalogu *security* umieszczone są pliki do testowania wszystkich technologi pod kątem zbieranych informacji, analogicznie w folderze *performance* pod kątem wydajności. Oba katalogi zawierają podkatalogi *visualize* w których znajdują się pliki do tworzenia graficznych prezentacji wyników. Dodatkowo folder *sterowniki* przechowuje wszystkie rozszerzenia i sterowniki *selenium* dla przeglądarek. Całą strukturę można zaobserwować na rysunku 4.1.



Rys. 4.1. Struktura plików projektu

4.3.1. Testowanie

Na początku każdego skryptu przeprowadzana jest podobna operacja. Ustawiane są *webdrivery* dla odpowiednich przeglądarek. W tym kroku dokonywane jest również dołączanie rozszerzeń, czy wybranie trybu *incognito*. Dla każdej przeglądarki ustawiamy odpowiadający jej predefiniowany *webdriver*

```
driver = None
try:
    if browser_name == 'chrome':
        options = webdriver.ChromeOptions()
        if incognito:
            options.add_argument("--incognito")
        service = ChromeService(executable_path=driver_path)
        options.add_argument("--headless")
        options.add_argument('--log-level=3')
        driver = webdriver.Chrome(service=service, options=options)
```

Rys. 4.2. Przykładowa konfiguracja *webdrivera*

z biblioteki *selenium*. Następnie ustawiamy opcje, tak jak na rysunku 4.2 wymuszenie trybu *incognito* lub ustawienie portu na, którym sterownik ma nasłuchiwać. W celach optymalizacyjnych dodawana jest opcja *headless*, która sprawia, że przeglądarka nie jest odpalana za każdym razem w trybie graficznym. Dodatkowo, poziom logów jest ustawiany na bardzo wysoki, gdyż przy takiej liczbie żądań jest zwracanych bardzo dużo danych co zaburza obraz i odbiór kluczowych informacji.

Omawianie można podzielić na dwie podkategorie, w ramach których mechanizm wygląda bardzo podobnie.

4.3.2. Wydajność

Przed przystąpieniem do sprawdzania tworzony jest proces, który reprezentuje działanie aplikacji. Na jego podstawie pobierane są informacje. W tym pakiecie skryptów sprawdzane było pięć wskaźników:

- czas ładowania strony

  W tym celu zapisywany jest czas tuż przed wysłaniem żądania i zaraz po jego zakończeniu. Następnie te wartości są odejmowane od siebie.

- zużycie CPU (procesora)

  Zapisywane są wartości użycia CPU przed i po wykonaniu zapytania. Na ich podstawienie wyliczany jest wynik.

- obciążenie pamięci RAM

  Analogicznie jak powyżej, sprawdzane jest obciążenie pamięci podręcznej przed i zaraz po zakończeniu procesu.

– zapis i odczyt z dysku

  Ponownie rejestrujemy stan początkowy i końcowy i na ich podstawie wyznaczamy ostateczne wartości.

Wszystkie powyższe wyliczenia można zaobserwować na rysunku 4.3.

```
process = psutil.Process()
cpu_start = process.cpu_times().user
memory_start = process.memory_info().rss
disk_start = psutil.disk_io_counters()

driver.get('about:blank')
start_time = time.time()
driver.get(url)
end_time = time.time()

load_time = end_time - start_time

cpu_end = process.cpu_times().user
memory_end = process.memory_info().rss
disk_end = psutil.disk_io_counters()

cpu_usage = cpu_end - cpu_start
memory_usage = memory_end - memory_start
disk_read = disk_end.read_bytes - disk_start.read_bytes
disk_write = disk_end.write_bytes - disk_start.write_bytes
```

Rys. 4.3. Implementacja zbierania danych o wydajności

Po wyznaczeniu metryk, wszystkie zostają zapisane do wcześniej stworzonej bazy danych.

4.3.3. Prywatność

W celu przetestowania wyznaczonych wartości, potrzebne było przeprowadzenie jednej istotnej zmiany. Mianowicie zamiana *webdriver* z *selenium* na te pochodzące z biblioteki *seleniumwire*. Pozwoliło to na uzyskanie dostępu do zmiennej *request* (jest to widoczne na rysunku 4.4) co było kluczowe w dalszym procesowaniu programu i de facto pozwoliło uzyskać dostęp do poszukiwanych danych.

W tej podgrupie sprawdzane były trzy kluczowe wskaźniki:

– Liczba plików cookies przesyłanych w trakcie przetwarzania żądań.

  W tym celu wykorzystana została metoda *get_cookies* dostępna dla *drivera* z *seleniumwire*.

– Liczba żądań zapisujących dane do pamięci lokalnej.

  Otrzymywane poprzez wywołanie kodu *JavaScript*, który zwraca lokalną pamięć.

– Całkowita liczba ruchu sieciowego wygenerowanego w trakcie obsługi zapytania.

  Uzyskane dzięki zapisywaniu wszystkich odpowiedzi i mapowaniu ich na docelowo pożądany format.

Zapisywanie ciasteczek, pamięci oraz wszystkich odpowiedzi w surowym formacie byłoby mało informatywne i wymagało bardzo dużo manualnej analizy, dlatego w tej pracy uwaga została zwrócona na ilościowy wymiar tych danych.

```python
def get_cookies(driver):
    return driver.get_cookies()

def get_local_storage(driver):
    return driver.execute_script("return window.localStorage")

def analyze_traffic(driver):
    traffic_data = []
    for request in driver.requests:
        if request.response:
            entry = {
                'url': request.url,
                'method': request.method,
                'status_code': request.response.status_code,
                'request_headers': dict(request.headers),
                'response_headers': dict(request.response.headers)
            }
            traffic_data.append(entry)
    return traffic_data
```

Rys. 4.4. Implementacja zbierania danych o prywatności

4.3.4. Realizacja w bazie danych

Jak zostało wyżej wspomniane, do stworzenia bazy danych została wykorzystana technologia *SQLite*. Schemat jest bardzo prosty, każda technologia posiada odpowiadające sobie dwie tabele. Jedna dla wskaźników wydajnościowych, a druga dla bezpieczeństwa. Różnice w tych schematach są niewielkie, mimo to została podjęta decyzja o podzieleniu ich dla każdego rozwiązania osobno. Jak widać na rysunku 4.5, ich liczba jest znacząca. Orientacyjnie każda tabela wygląda następująco:

1. Dla wydajności:
   – *memory_usage* – zawiera zużycie pamięci podręcznej. Typ danych to *Real*.
   – *cpu_usage* – przechowuje zużycie procesora. Typ danych to *Real*.
   – *load_time* – reprezentuje czas ładowania strony. Typ danych to *Real*.
   – *name* – zawiera nazwę przeglądarki. Typ danych *Text*.

– *page* – przechowuje adres url sprawdzanej strony, lub ogólnie przypisany jej typ. Typ danych to *Text*.

– *id* – reprezentuje unikatowe ID przypisane do konkretnego testu. Typ danych to *Integer*.

– *disk_write* – zawiera dane o zapisie na dysku. Typ danych to *Real*.

– *disk_read* – przechowuje dane o odczycie z dysku. Typ danych to *Real*.

– *mode* (opcjonalne) – uwzględnia informacje o tym w jakim trybie wykonywane jest zapytanie, prywatnym lub normalnym. Typ danych to *Text*.

2. Dla prywatności:

– *cookies* – zawiera liczbę plików cookies tworzonych i zapisywanych w trakcie zapytania. Typ danych to *Integer*.

– *local_storage* – przechowuje liczbę danych zapisanych do pamięci lokalnej. Typ danych to*Integer*.

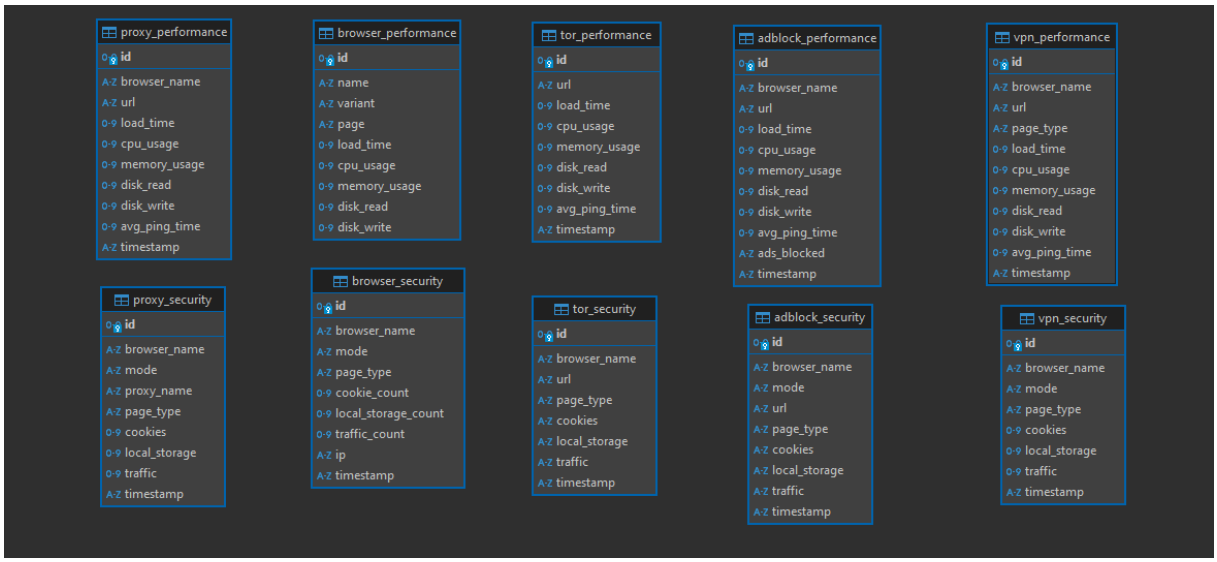
– *traffic* – reprezentuje całkowitą liczbę wygenerowanego ruchu sieciowego. Typ danych to*Integer*.

– *name* – zawiera nazwę przeglądarki. Typ danych to *Text*.

– *page* – przechowuje adres url sprawdzanej strony, lub ogólnie przypisany jej typ. Typ danych to *Text*.

– *id* – reprezentuje unikatowe ID przypisane do konkretnego testu. Typ danych to *Integer*.

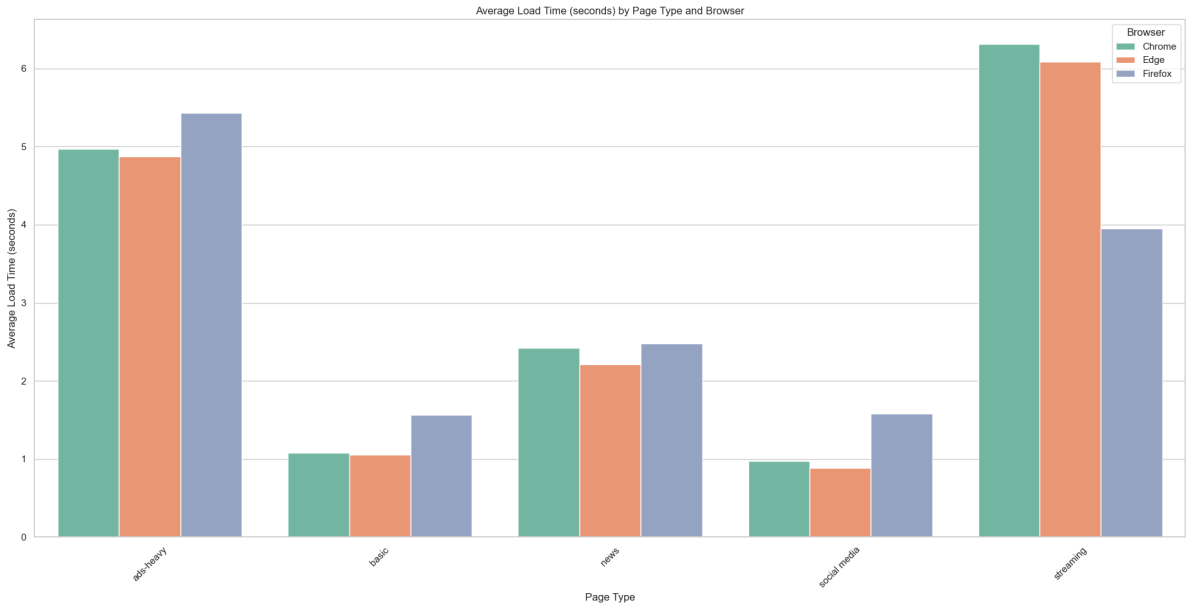
– *mode* (opcjonalne) – uwzględnia informacje o tym w jakim trybie wykonywane jest zapytanie, prywatnym lub normalnym. Typ danych to *Text*.



Rys. 4.5. Wizualizacja bazy danych i wszystkich zawierających się w niej tabel

4.3.5. Prezentacja wyników

Prezentacja wyników zrealizowana została w osobnym skrypcie wywoływanym z pliku testującego. Pobiera dane z bazy danych a następnie przy wykorzystaniu biblioteki *matplotlib* rysuje wyniki na ekran w postaci wykresów słupkowych. Robi to osobno dla każdego ze wskaźników. Co istotne, nie wyświetla jedynie rezultatów dla ostatniego testu, ale dodaje je do wszystkich rezultatów i na ich podstawie wylicza średnią. Przykładowy obrazek wygenerowany przy jej pomocy można zaobserwować na rysunku 4.6. Oczywiście takie wykresy są prezentowane dla każdego ze sprawdzanych wskaźników.



Rys. 4.6. Przykładowy wykres otrzymany za pomocą skryptu rysującego

4.4. Wymagania sprzętowe aplikacji

Aplikacja wymaga komputera z systemem operacyjnym Windows zdolnego do uruchomienia przeglądarki i zapasem pamięci pozwalającym na instalację czterech przeglądarek i sterowników *selenium* do nich. Dodatkowo konieczne będzie załączenie przedstawionych na rysunku 4.7 zależności bibliotek, z pliku *requirements.txt*. Aplikacja była testowana dla poniższych wersji przeglądarek:

- *Chrome* 128.0.6613.85
- *Edge* 128.0.2739.42
- *Firefox* 127.0
- *Brave* 1.69.160
- *Tor browser* 13.5.2

Niestety, kompatybilność przeglądarek z *webdriverami* stanowi bardzo często problem przez co reprodukcja wyników może być trudna do uzyskania. Możliwe jest napotkanie wielu błędów.

```
bleach==6.1.0
blinker==1.7.0
cryptography==41.0.5
googleapis-common-protos==1.62.0
matplotlib==3.9.2
numpy==1.26.0
pandas==2.2.2
pillow==10.4.0
ping3==4.0.8
psutil==5.9.5
pycryptodome==3.19.0
pyOpenSSL==24.2.1
pyparsing==3.1.2
PySocks==1.7.1
requests==2.31.0
seaborn==0.13.2
selenium==4.21.0
selenium-wire==5.1.0
setuptools==72.1.0
SQLAlchemy==1.4.17
urllib3==2.0.6
websocket-client==1.6.4
```

Rys. 4.7. Biblioteki wymagane do zainstalowania w celu uruchomienia skryptu

4.5. Instalacja i uruchomienie

Jest to aplikacja konsolowa, więc wszystko odbywa się z poziomu konsoli. Przed pierwszym odpaleniem należy zainstalować wszystkie wymagane biblioteki. Można to zrobić, wykonując komendę:

pip install -r requirements.txt

Kolejnym krokiem jest przejście do folderu *database* i uruchomienie pliku konfiguracyjnego bazy danych.

python database_setup.py

Następnie trzeba wrócić do folderu głównego projektu. Z tego poziomu można już wywoływać program i rozpocząć testowanie. Każde wywołanie wygląda następująco:

python manage.py {opcja1} {opcja2}

gdzie *opcja1* określa, pod jakim kątem testujemy technologię. Z kolei *opcja2* określa jakie narzędzie będzie testowane. Wszystkie opcje są przedstawione poniżej:

– *opcja1* – {security, performance}

– *opcja2* – {browsers, adblock, tor, vpn, proxy}

Warto zaznaczyć, że przy testowaniu przeglądarek automatycznie uruchamiany jest test dla trybu prywatnego. Kolejnym istotnym punktem jest to, że należy pamiętać o włączeniu *VPN-a* przed przystąpieniem do jego testów. Po uruchomieniu testu dla którejkolwiek technologii należy chwilę poczekać i na ekranie pokażą się wykresy przedstawiające otrzymane rezultaty.

Oczywiście jest to opis dla jednostkowego testowania. W celu otrzymania bardziej miarodajnych rezultatów potrzebna była automatyzacja testowania. Tak, aby nie było potrzebne każdorazowe manualne uruchamianie narzędzia. Byłoby to szczególnie uciążliwe biorąc pod uwagę czas wykonywania każdego testu. W tym celu przy wywoływaniu skryptu została dodana opcja *iterations* zarządzająca pętla w pliku startowym (rysunek 4.9), której domyślnie przypisana jest wartość jeden. Ostatecznie na rysunku 4.8 widać wszystkie wywołania aplikacji z parametrami, które pozwoliły uzyskać ostateczne wyniki.

```
python manage.py security browsers --iterations 100
python manage.py security adblock --iterations 100
python manage.py security vpn --iterations 100
python manage.py security proxy --iterations 100
python manage.py security tor --iterations 100
python manage.py performance browsers --iterations 100
python manage.py performance adblock --iterations 100
python manage.py performance vpn --iterations 100
python manage.py performance proxy --iterations 100
python manage.py performance tor --iterations 100
```

Rys. 4.8. Wszystkie wykorzystane wywołania w celu otrzymania wyników

```
group_scripts = scripts.get(args.group)
if group_scripts:
    selected_script = group_scripts.get(args.script)
    if selected_script:
        for _ in range(args.iterations):
            selected_script.main()
    else:
        print(f"Error: Script '{args.script}' not found in group '{args.group}'.")
        sys.exit(1)
else:
    print(f"Error: Group '{args.group}' not found.")
    sys.exit(1)
```

Rys. 4.9. Automatyzacja testów przy użyciu prostej pętli *for*

5. Prezentacja i analiza uzyskanych wyników

Przed przystąpieniem do prezentacji wyników uzyskanych po przeprowadzeniu testów, zostanie przedstawiona specyfikacja warunków, w jakich zostały one wykonane.

Charakterystyka maszyny, na której były wykonywane testy:

- system operacyjny – Windows 10 Home 22H2
- procesor – AMD Ryzen 3 3200G with Radeon Vega Graphics 3.60 GHz
- zainstalowana pamięć RAM – 32,0 GB
- karta graficzna – NVIDIA GeForce GTX 1650
- karta sieciowa – TP-Link Wireless USB Adapter

Charakterystyka sieci:

- protokół – Wi-Fi 5 (802.11ac)
- pasmo – 5 GHz
- typ zabezpieczeń – WPA2-Personal
- szybkość łącza (odbieranie/przesyłanie) – 433/433 (Mbps)

Przetestowano następujące rodzaje stron:

- *https://www.example.com* – podstawowa statyczna strona, o bardzo ograniczonej funkcjonalności
- *https://www.youtube.com* – strona streamingowa
- *https://www.wp.pl* – strona zawierająca informacje ze świata
- *https://www.thepiratebay.org* – strona z dużą ilością reklam i o szemranej reputacji
- *https://www.reddit.com* – strona reprezentująca media społecznościowe

Wybór został dokonany w taki sposób, aby objąć różne rodzaje witryn, ale skupiając się na takich, które są często odwiedzane przez użytkowników.

5.1. Wydajność

W porównaniu do aplikacji zaszły pewne zmiany. Ze względu na to, że przy testowaniu pojedynczych stron internetowych pewne wskaźniki wydajnościowe miały niewielkie wartości i nieznacznie się od siebie różniły, oraz były podatne działania w tle, podjęta została decyzja o zmianie metody testowania. Wyniki oprócz czasu ładowania zostały otrzymane poprzez odpalenia wszystkich testowanych stron jednocześnie i sprawdzenie wyników w systemowym menadżerze zadań. Dało to lepszą izolację testów od innych czynników działających w tle na komputerze i dokładniejsze oddanie rzeczywistego obciążenia. Rezultaty zostały przedstawione dla stu prób.

Jednym z kluczowych celów pracy było sprawdzenie, w jaki sposób rozwiązania mające zapewnić anonimowość wpływają na wydajność i komfort korzystania z internetu. W tym podpunkcie zostanie to dokładnie przeanalizowane.

5.1.1. Porównanie przeglądarek i trybu prywatnego

Do tego testu wybrane zostały przeglądarki:

- *Chrome,*
- *Firefox,*
- *Brave,*
- *Edge.*

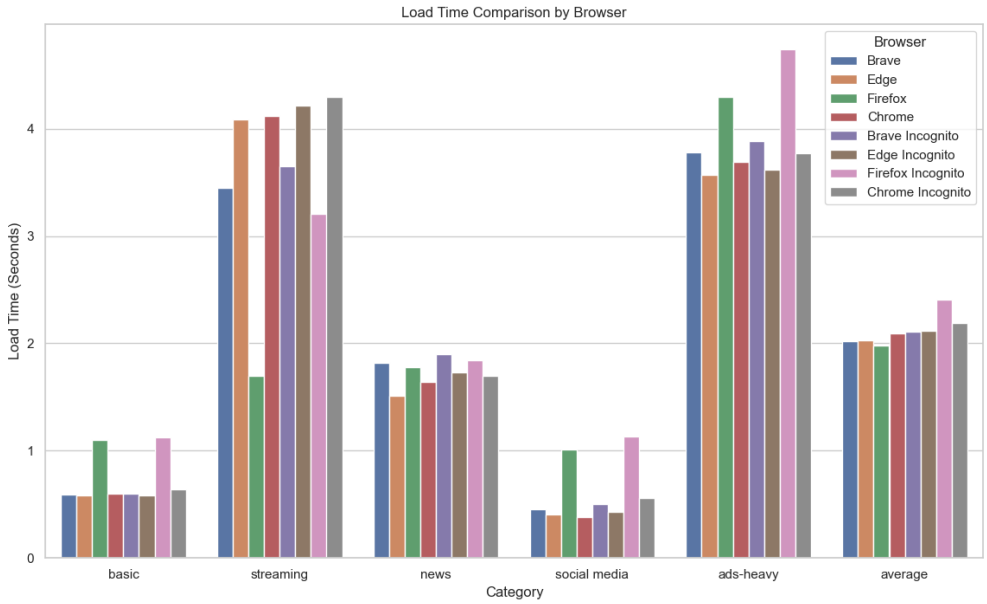
Wybór został wykonany na podstawie popularności. Brakuje tu oczywiście *Safari*, które wyraźnie zajmuje pod tym względem drugie miejsce [31], niestety jednak *Apple* nie podtrzymuje wsparcie dla tej przeglądarki na systemy *Windows*. Dodatkowo wybrana została przeglądarka *Brave*, która reklamuje się jako ta zapewniająca zwiększoną prywatność. Pozostałe technologie w celu uniknięcia zbyt dużego *biasu* jednej przeglądarki był testowane na trzech przeglądarkach: *Chrome*, *Edge*, *Firefox*.

Dokładne wersje przeglądarek można znaleźć w rozdziale 4.4 tej pracy. Dla każdej z przeglądarek testowany był zarówno tryb zwykły jak i *incognito*.

Tabela 5.1 oraz rysunek 5.1 przedstawiają wyniki dla każdej z nich:

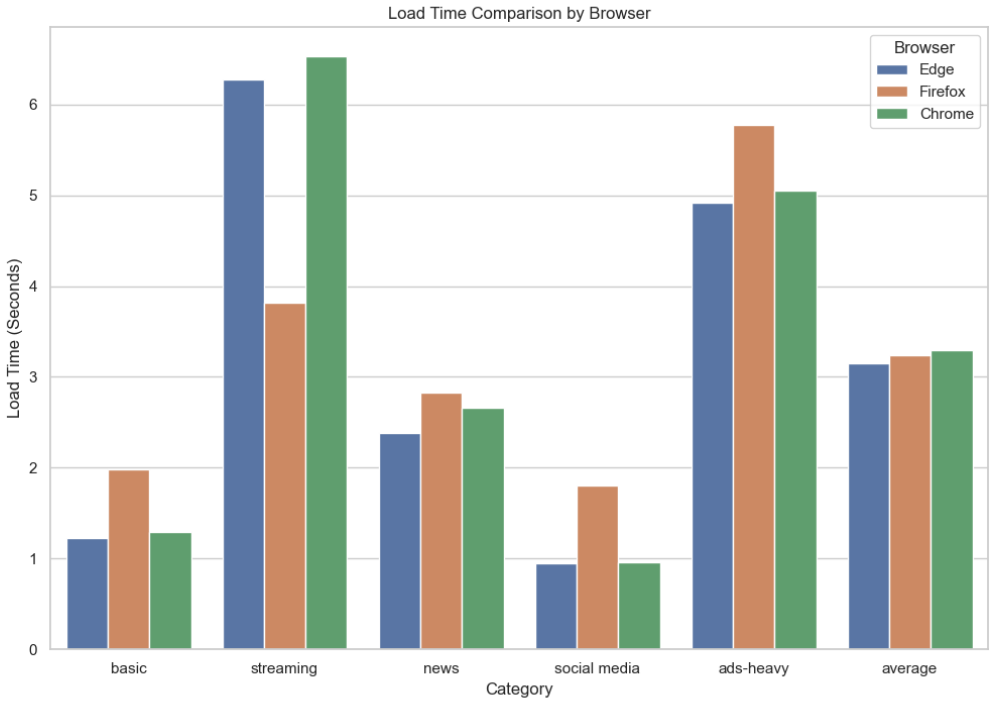
	Brave	**Edge**	**Firefox**	**Chrome**	**Brave Incognito**	**Edge Incognito**	**Firefox Incognito**	**Chrome Incognito**
CPU (%)	3.71	8.42	4.57	5.00	3.75	8.83	4.71	5.25
RAM (MB)	573	514	777	329	598	576	832	405
Disk (MB/s)	0.42	0.07	0.07	0.07	0.40	0.05	0.04	0.05
Network (Mbps)	0.92	3.14	2.71	0.08	1.23	3.75	3.11	2.40

Tabela 5.1. Wyniki metryk wydajnościowych dla otwartych wszystkich testowanych stron jednocześnie dla poszczególnych przeglądarek w trybie normalnym i prywatnym



Rys. 5.1. Czas ładowania w zależności od strony, przeglądarki i trybu

5.1.2. VPN



Rys. 5.2. Czas ładowania w zależności od strony dla technologii *VPN*

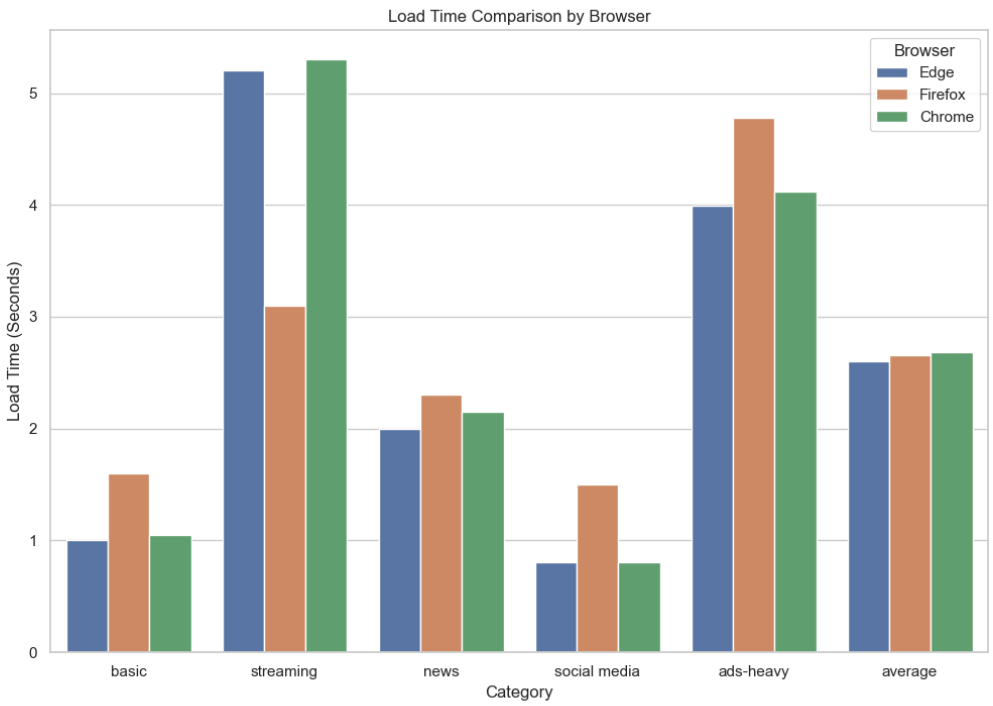
Testy były wykonywane przy wykorzystaniu rozwiązania *Virtual Private Network ProtonVPN*. Głównie z tego powodu, że jest to darmowe i sprawdzone rozwiązanie. Co istotne, technologia *VPN* nie jest automatycznie odpalana ze skryptu, także przed wykonaniem kodu należy każdorazowo upewnić się, że *VPN* został uruchomiony manualnie. Wyniki można zaobserwować w tabeli 5.2 oraz rysunku 5.2.

	Edge	**Firefox**	**Chrome**
CPU (%)	10.21	5.94	6.27
RAM (MB)	649	1071	625
Disk (MB/s)	0.07	0.07	0.07
Network (Mbps)	5.78	5.33	2.07

Tabela 5.2. Wyniki metryk wydajnościowych dla otwartych wszystkich testowanych stron jednocześnie dla technologii *VPN*

5.1.3. Proxy

Aplikacja była uruchamiana dla różnych serwerów proxy dostępnych na stronie *FineProxy* [32]. Analogicznie jak dla *VPN*, testowanie odbyło się na trzech przeglądarkach. Wyniki przedstawiono na rysunku 5.3 i tabeli 5.3.



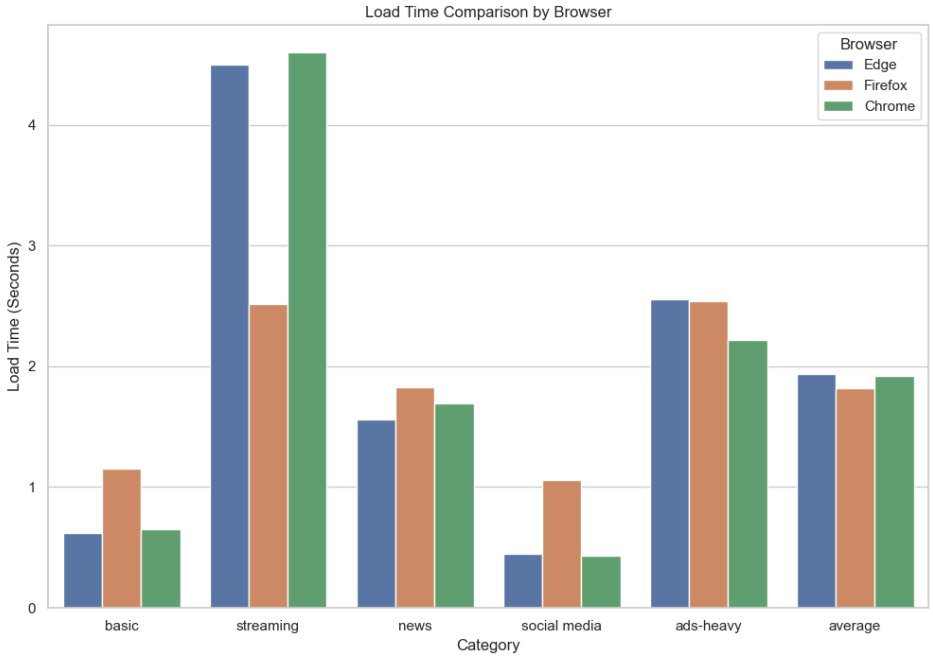
Rys. 5.3. Czas ładowania w zależności od strony dla technologii *proxy*

	Edge	Firefox	Chrome
CPU (%)	8.61	4.74	5.22
RAM (MB)	544	856	500
Disk (MB/s)	0.07	0.07	0.07
Network (Mbps)	4.75	4.31	1.67

Tabela 5.3. Wyniki metryk wydajnościowych dla otwartych wszystkich testowanych stron jednocześnie dla technologii *proxy*

5.1.4. AdBlock

Dla każdej z przeglądarek: *Chrome*, *Edge*, *Firefox* zostało dołączone rozszerzenie *AdBlock Plus* poprzez ustawienie odpowiedniej opcji *webdrivera*. Pliki rozszerzeń dla odpowiednich przeglądarek znajdują się w folderze sterowniki. Co ciekawe *Edge* i *Chrome* operują na tych samych plikach rozszerzeń (co nie musi dziwić ze względu na oparcie przeglądarki *Edge* na technologii *chromium*) *crx*, z kolei *Firefox* używa plików *xpi*. Wyniki dla technologii *AdBlock* można zaobserwować na rysunku 5.4 i tabeli 5.4.



Rys. 5.4. Czas ładowania w zależności od strony dla technologii *adblock*

5.1.5. TOR

Wyniki dla trasowania cebulowego można zobaczyć w tabelach 5.5 i 5.6.

	Edge	Firefox	Chrome
CPU (%)	8.49	4.64	5.07
RAM (MB)	521	784	335
Disk (MB/s)	0.08	0.08	0.08
Network (Mbps)	3.22	2.78	0.14

Tabela 5.4. Wyniki metryk wydajnościowych dla otwartych wszystkich testowanych stron jednocześnie dla technologii *adblock*

Load Time (s)	**TOR**
Basic	1.50
Streaming	7.85
News	2.94
Social Media	1.16
Ads-Heavy	6.12
Average	3.85

Tabela 5.5. Czas ładowania różnych stron dla przeglądarki *TOR*

All Tabs Opened (Manual Test)	**TOR**
CPU (%)	23.45
RAM (MB)	680
Disk (MB/s)	0.60
Network (Mbps)	13.97

Tabela 5.6. Wyniki metryk wydajnościowych dla otwartych wszystkich testowanych stron jednocześnie dla technologii *TOR*

5.1.6. Analiza uzyskanych wyników

Wszystkie poniższe wykresy zostały wykonane przez odpowiednie skrypty w module *visualize*, które trzeba wywołać osobno, a nie poprzez domyślne sprawdzenie technologii. Należą do nich pliki z przedrostkiem *summary*. Do ich stworzenia został wykorzystany skrypt wyciągający dane dla każdej kategorii z każdej tabel i prezentujący je na jednym wykresie. Schemat działania można zaobserwować na rysunku 5.5.

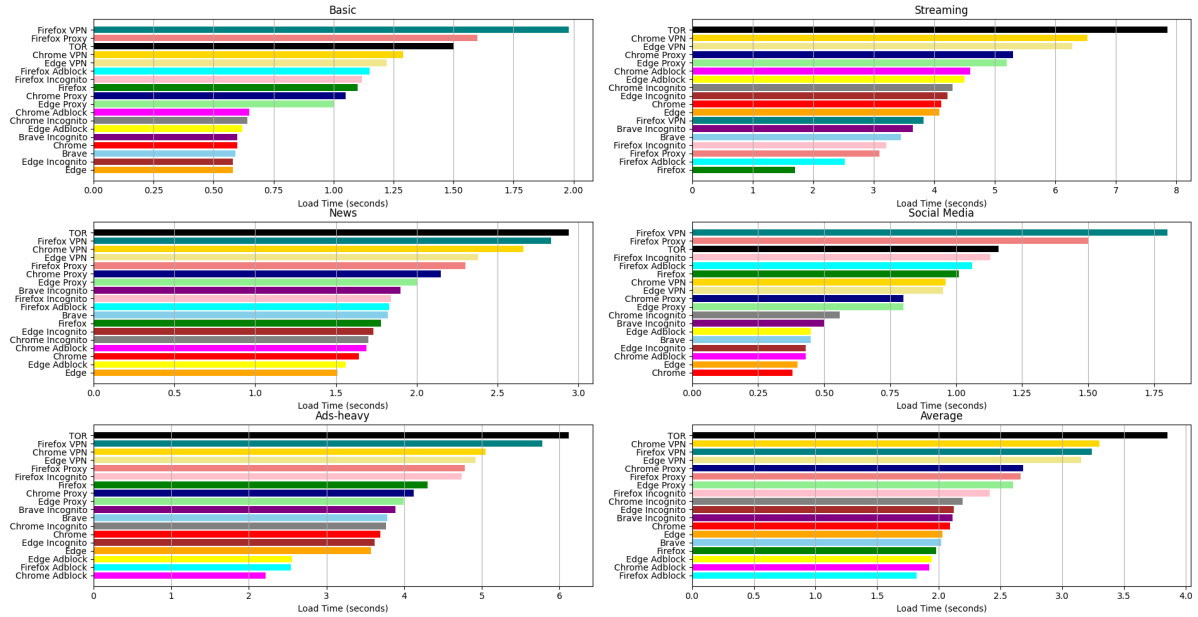
5.1.6.1. Czas ładowania

Czas ładowania jest jednym z kluczowych czynników wpływających na odczucia podczas korzystania z Internetu. Dokładne wyniki podzielone ze względu na konkretne rodzaje stron internetowych widać na rysunku 5.6.

```
all_data = {
    'TOR': tor_data,
    **{f'VPN {key}': value for key, value in vpn_data.items()},
    **{f'Proxy {key}': value for key, value in proxy_data.items()},
    **{f'Adblock {key}': value for key, value in adblock_data.items()},
    **{f'Browser {key}': value for key, value in browsers_incognito_data.items()}
}

# Colors for different categories
colors = {
    "TOR": 'black',
    "VPN Edge": 'red', "VPN Firefox": 'blue', "VPN Chrome": 'green',
    "Proxy Edge": 'orange', "Proxy Firefox": 'purple', "Proxy Chrome": 'cyan',
    "Adblock Edge": 'pink', "Adblock Firefox": 'brown', "Adblock Chrome": 'lightgreen',
    "Browser Brave": 'yellow', "Browser Edge": 'lightblue', "Browser Firefox": 'lightcoral'
    "Browser Brave Incognito": 'gold', "Browser Edge Incognito": 'lightgray',
    "Browser Firefox Incognito": 'teal', "Browser Chrome Incognito": 'silver'
}

# Creating individual bar charts for each metric

for i, metric in enumerate(metrics):
    # Sorting data for each metric
    sorted_data = sorted(all_data.items(), key=lambda x: x[1][metric])
    labels, values = zip(*[(label, data[metric]) for label, data in sorted_data])

    # Create individual figure for each metric
    plt.figure(figsize=(10, 6))
    plt.barh(labels, values, color=[colors[label] for label in labels])
    plt.title(metric)
    plt.xlabel(metric)
    plt.grid(True, axis='x')
    plt.show()
```

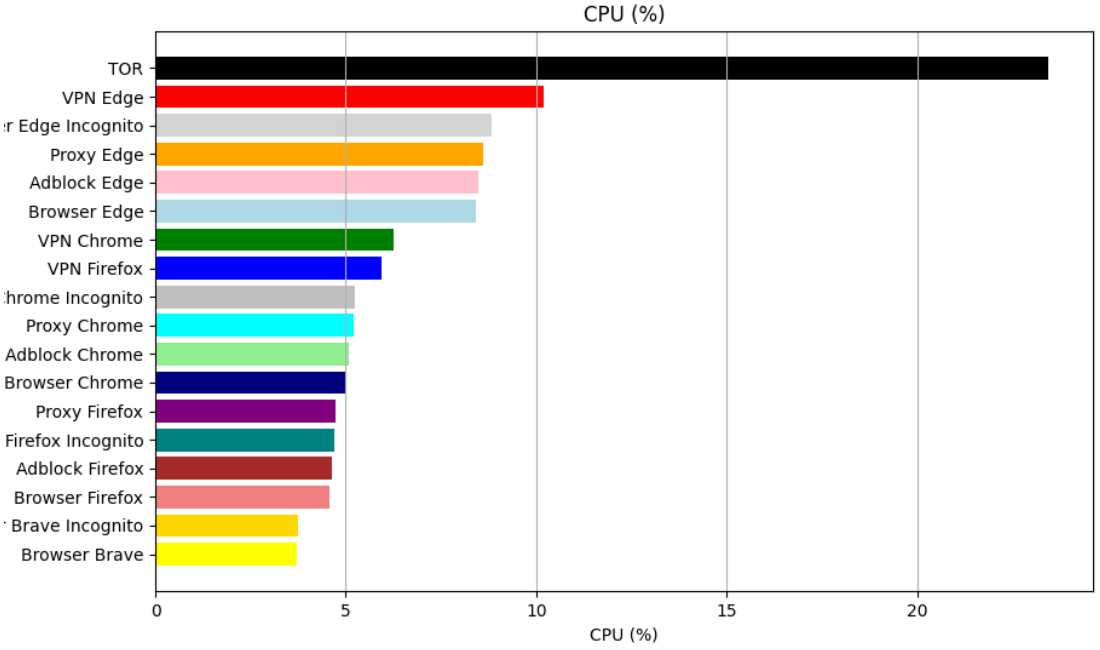
Rys. 5.5. Generowanie wykresów porównania dla metryk wydajnościowych

Można zaobserwować, że *TOR* prawie dla wszystkich kategorii wypada najgorzej. Oprócz kategorii *Social Media* i *Basic*, gdzie zdecydowanie prowadzi *Firefox*. Nie jest to dziwne, ze względu na charakterystykę narzędzia jakim jest trasowanie cebulowe. Wymaga wielu przekierowań co jest czasochłonne a dodatkowo na początku transmisji zużywany jest czas na wielokrotne szyfrowanie połączenia. Ogólnie można podzielić dane na dwie kategorie. Te przekierowujące ruch czyli *proxy*, *VPN* oraz *TOR* oraz te zmieniające jedynie coś po stronie przeglądarki - *incognito*, *Adblock*, tryb zwykłego przeglądania. Różnice pomiędzy nimi widać na rysunku 6.1. Pierwsza grupa odznacza się w stosunku do drugiej większym czasem potrzebnym do załadowania strony. Spośród przeglądarek w trybie normalnym najszybszy okazał się *Firefox* z kolei najwolniejszy *Chrome*, a w *incognito* najwolniejszy, ustępując pierwszego miejsca przeglądarce *Brave*. Nie są to jednak duże różnice. Dla trybu normalnego 5.56%, a *incognito* 14.22%. W ramach jednej technologii średnie różnice pomiędzy przeglądarkami są nieznaczne. Całościowo rozwiązaniem, które najmniej negatywnie wpływa na czas ładowania okazał się *AdBlock*, a nawet lekko poprawia wyniki, co nie dziwi ze względu na ograniczenie czasu potrzebnego na ładowanie reklam.



Rys. 5.6. Porównanie czasów ładowania stron wraz ze średnim czasem dla różnych technologii

5.1.6.2. Obciążenie CPU

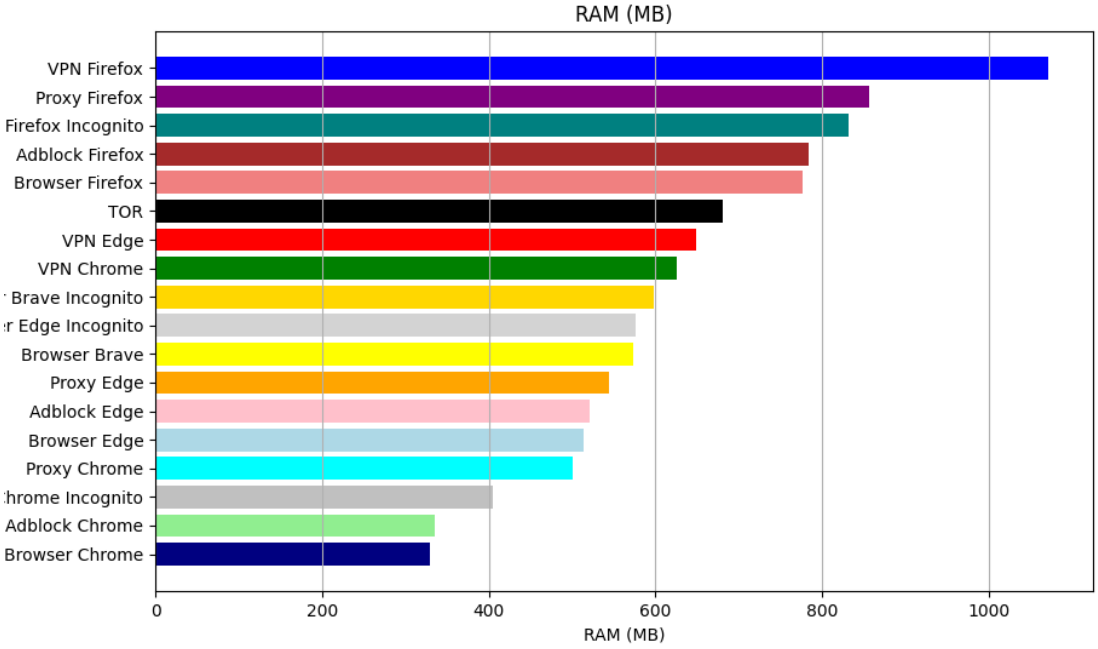


Rys. 5.7. Porównanie zużycia CPU dla różnych technologii

Pomijając trasowanie cebulowe, w przypadku procesora przeglądarka ma kluczowe znaczenie (większy od sprawdzanej technologii). Wyraźnie można uszeregować ich wpływ od największego do najmniejszego: *Edge*, *Chrome*, *Firefox*, *Brave*. O ile różnice między trzema ostatnimi nie są duże to pod

tym względem *Edge* wyraźnie odstaje. Na rysunku 5.7 można też zaobserwować ogromne wykorzystanie procesora przez technologie trasowania cebulowego sięgające, aż 23.45%. Dzieje się tak z powodu tego, że *TOR* przed rozpoczęciem wysłania żądania wykorzystuje procesor do przeprowadzenia szyfrowania połączenia stąd taki wynik. Widać też, że początkowe szyfrowanie komunikacji w wypadku *VPN* umiarkowanie zwiększa obciążenie CPU.

5.1.6.3. Pamięć RAM



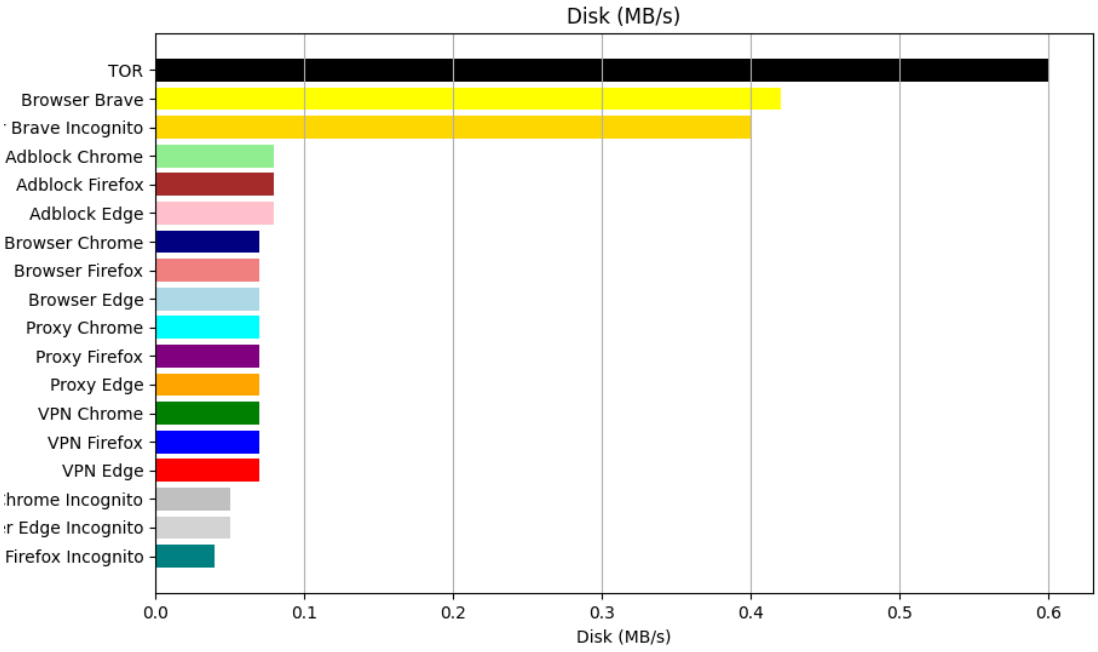
Rys. 5.8. Porównanie obciążenia pamięci podręcznej dla różnych technologii

Analizując zużycie pamięci RAM w różnych scenariuszach przeglądania, można dostrzec kilka wyraźnych wzorców. Jest to kolejna metryka, na która, główny wpływ ma przeglądarka, co widać wyraźnie na rysunku 5.8. W przypadku normalnego przeglądania, *Firefox* konsekwentnie zużywa najwięcej pamięci RAM, osiągając maksymalnie 1071 MB podczas korzystania z *VPN*, podczas gdy *Chrome* zazwyczaj wykorzystuje najmniej pamięci, osiągając niski poziom 335 MB w przypadku używania *Adblocka*. Tor, w tym kontekście, pokazuje umiarkowane zużycie RAM, mieszcząc się między wartościami obserwowanymi w konfiguracjach *VPN*, *Proxy* i *Adblock*. Przechodząc do trybu *Incognito*, *Firefox* pozostaje najbardziej obciążający, osiągając 832 MB, podczas gdy *Chrome* ponownie wykazuje niższe zużycie pamięci, z minimalnym poziomem 329 MB w trybie standardowym i 405 MB w trybie *Incognito*. *Brave* i *Edge* wykazują nieco wyższe zużycie RAM w trybie *Incognito* w porównaniu do trybu standardowego. Ogólnie rzecz biorąc, *Firefox* zużywa więcej pamięci RAM w różnych scenariuszach w porównaniu do innych przeglądarek, podczas gdy *Chrome* jest konsekwentnie bardziej efektywny pod względem zużycia pamięci. Dlatego, jeśli priorytetem jest niskie zużycie pamięci, *Chrome* jest preferowanym wyborem, szczególnie w trybie *Incognito*, podczas gdy *Firefox* może być wybierany ze względu na dodatkowe

funkcje, mimo wyższego zużycia RAM. Co może zaskakiwać w porównaniu do poprzednich metryk, *TOR* w tym wypadku szczególnie nie wyróżnia się negatywnie na tle innych rozwiązań.

5.1.6.4. Obciążenie dysku

Patrząc na zużycie dysku w różnych scenariuszach przeglądania widać kilka ciekawych rzeczy.



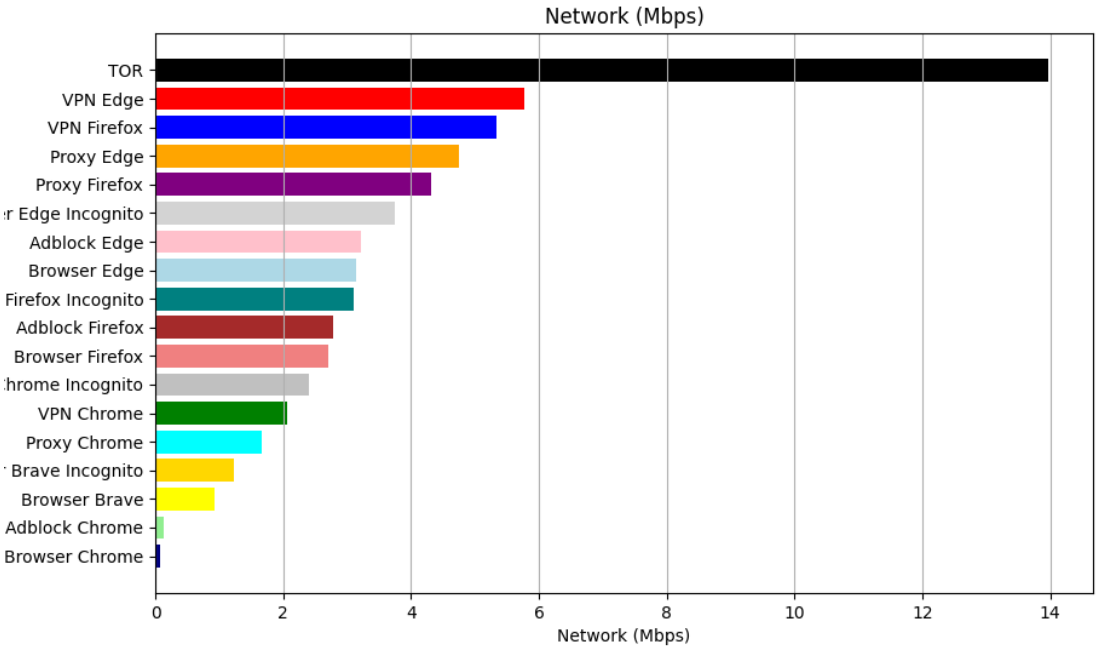
Rys. 5.9. Porównanie obciążenia dysku dla różnych technologii

W przypadku normalnego przeglądania, wszystkie przeglądarki pokazują stosunkowo niskie zużycie dysku, z większością wartości wynoszących 0,07 MB/s, z wyjątkiem *Tor*, który zużywa 0,6 MB/s (rysunek 5.9). W trybie *Incognito* zużycie dysku zmienia się nieznacznie, z większością przeglądarek zużywających od 0,05 do 0,42 MB/s. Najwyższe zużycie dysku odnotowano w przypadku przeglądarki *Brave* (0,42 MB/s), podczas gdy inne przeglądarki, takie jak *Edge* i *Firefox*, mają wartości bliższe dolnej granicy tego zakresu. Ogólnie rzecz biorąc, zużycie dysku przez przeglądarki jest stosunkowo niskie we wszystkich scenariuszach, z wyraźnymi różnicami w przypadku Tor i *Brave*, które wykazują wyższe wartości. Choć warto mieć na uwadze, że są to wciąż relatywnie niewielkie wartości.

5.1.6.5. Użycie sieci

Analizując zużycie sieci z rysunku 5.10, można zauważyć kilka wyraźnych wzorców. W przypadku normalnego przeglądania, *Tor* zużywa najwięcej zasobów sieciowych, osiągając 13,97 Mbps, co jest nawet kilkukrotnie wyższe niż w przypadku innych przeglądarek. Przeglądarki w trybie *VPN* wykazują niższe zużycie sieci, z *Firefox* na poziomie 5,33 Mbps, *Edge* na 5,78 Mbps, i *Chrome* na 2,07 Mbps. W trybie *Proxy* i *Adblock* zużycie sieci jest jeszcze niższe, przy czym *Chrome* osiąga minimalne wartości odpowiednio 1,67 Mbps i 0,14 Mbps. W trybie *Incognito*, przeglądarki również pokazują zróżnicowane

zużycie sieci, z *Edge Incognito* na najwyższym poziomie 3,75 Mbps, a *Chrome Incognito* zużywa najmniej – tylko 0,08 Mbps. Ogólnie rzecz biorąc, *Tor* wyróżnia się najwyższym zużyciem sieci, natomiast *Chrome*, zarówno w trybie zwykłym, jak i *Incognito*, wykazuje najniższe zużycie, co czyni go bardziej efektywnym pod względem tej metryki.



Rys. 5.10. Porównanie zużycia sieci dla różnych technologii

5.2. Bezpieczeństwo

Dane zbierane w tej sekcji zostały dokładnie opisane w podrozdziale 4.3.3. Testowane zestawienia nie różnią się niczym w porównaniu do sprawdzania wydajnościowego, dlatego opisywanie konfiguracji zostało w tym punkcie pominięte.

5.2.1. Porównanie przeglądarek i tryby prywatnego

Wyniki prywatnościowe w kolejności pliki *cookies*, zapisywana pamięć lokalna i generowany ruch sieciowy zostały przedstawione w tabelach 5.7 - 5.9.

5.2.2. VPN

Analogicznie jak powyżej w tabelach 5.10 - 5.12 można odnaleźć wyniki generowane dla technologii *VPN*.

Cookies	Brave	Edge	Firefox	Chrome	Brave Incognito	Edge Incognito	Firefox Incognito	Chrome Incognito
Basic	0	0	0	0	0	0	0	0
Streaming	6	4	4	4	6	4	4	4
News	15	16	16	16	15	16	16	16
Social Media	1	1	7	1	1	1	7	1
Ads-Heavy	0	0	6	0	0	0	6	0
Average	4.4	4.2	6.6	4.2	4.4	4.2	6.6	4.2

Tabela 5.7. Generowane pliki *cookies* dla przeglądarek w trybie normalnym i *incognito*

Local Storage	Brave	Edge	Firefox	Chrome	Brave Incognito	Edge Incognito	Firefox Incognito	Chrome Incognito
Basic	6	6	6	6	6	6	6	6
Streaming	11	11	11	11	11	11	11	11
News	9	9	9	9	9	9	9	9
Social Media	7	6	7	6	7	6	7	6
Ads-Heavy	6	15	15	16	6	15	15	16
Average	7.8	9.4	9.6	9.6	7.8	9.4	9.6	9.6

Tabela 5.8. Generowana pamięć lokalna dla przeglądarek w trybie normalnym i *incognito*

Traffic	Brave	Edge	Firefox	Chrome	Brave Incognito	Edge Incognito	Firefox Incognito	Chrome Incognito
Basic	47	18	20	1	47	13	20	1
Streaming	195	186	197	153	197	183	187	156
News	134	125	134	110	154	121	135	110
Social Media	48	16	300	2	48	11	295	2
Ads-Heavy	55	44	76	28	55	39	76	27
Average	95.8	77.8	145.4	58.8	100.2	73.4	142.6	59.2

Tabela 5.9. Generowany ruch sieciowy dla przeglądarek w trybie normalnym i *incognito*

Cookies	Chrome	Firefox	Edge
Basic	0	0	0
Streaming	4	4	4
News	16	16	16
Social Media	1	7	1
Ads-Heavy	0	0	0
Average	4.2	5.4	4.2

Tabela 5.10. Liczba generowanych plików cookies dla technologii *VPN*

Local Storage	Chrome	Firefox	Edge
Basic	6	6	6
Streaming	11	11	11
News	9	9	9
Social Media	6	6	6
Ads-Heavy	16	15	15
Average	9.6	9.4	9.4

Tabela 5.11. Generowana pamięć lokalna dla technologii *VPN*

Traffic	Chrome	Firefox	Edge
Basic	1	17	18
Streaming	156	184	179
News	111	134	126
Social Media	2	304	16
Ads-Heavy	27	48	43
Average	59.4	137.4	76.4

Tabela 5.12. Generowany ruch sieciowy dla technologii *VPN*

5.2.3. Proxy

Wyniki dla technologii *proxy* można zaobserwować w tabelach 5.13, 5.14 i 5.15.

Cookies	Chrome	Firefox	Edge
Basic	0	0	0
Streaming	4	4	4
News	16	16	16
Social Media	1	7	1
Ads-Heavy	0	5	0
Average	4.2	6.4	4.2

Tabela 5.13. Liczba generowanych plików cookies dla technologii *proxy*

Local Storage	Chrome	Firefox	Edge
Basic	6	6	6
Streaming	11	11	11
News	9	9	9
Social Media	6	6	6
Ads-Heavy	16	14	15
Average	9.6	9.2	9.4

Tabela 5.14. Generowana pamięć lokalna dla technologii *proxy*

Traffic	Chrome	Firefox	Edge
Basic	1	20	18
Streaming	11	212	182
News	111	136	125
Social Media	2	300	16
Ads-Heavy	28	68	44
Average	30.6	147.2	77

Tabela 5.15. Generowany ruch sieciowy dla technologii *proxy*

5.2.4. AdBlock

Po skorzystaniu z rozszerzeń do przeglądarek blokujących reklamy, wyniki w stosunku do zwykłego przeglądania zmieniły się w widoczny w tabelach 5.16 - 5.18.

Cookies	Chrome	Firefox	Edge
Basic	0	0	0
Streaming	4	4	4
News	16	16	16
Social Media	1	7	1
Ads-Heavy	0	6	0
Average	4.2	6.6	4.2

Tabela 5.16. Liczba generowanych plików cookies dla technologii *adblock*

Local Storage	Chrome	Firefox	Edge
Basic	6	6	6
Streaming	11	11	11
News	9	9	9
Social Media	6	7	6
Ads-Heavy	16	15	15
Average	9.6	9.6	9.4

Tabela 5.17. Generowana pamięć lokalna dla technologii *adblock*

Traffic	Chrome	Firefox	Edge
Basic	1	19	18
Streaming	153	200	184
News	111	136	125
Social Media	2	300	18
Ads-Heavy	28	76	44
Average	59	146.2	77.8

Tabela 5.18. Generowany ruch sieciowy dla technologii *adblock*

5.2.5. TOR

Wszystkie zebrane dane pochodzą z aplikacji, oprócz danych dotyczących technologii *TOR*. Stało się tak ze względu na duże ograniczenia, jakie ta technologia nakłada na wykonywanie kodu JavaScript oraz problemy z dostępem do pamięci lokalnej. Analiza została wykonana przy pomocy programu *WireShark*, który służy do analizy ruchu sieciowego [33]. Wyniki dla tej technologii można zaobserwować w tabeli 5.19.

Category	Cookies	Storage	Traffic
Basic	0	1	54
Streaming	1	3	587
News	2	2	369
Social Media	3	2	830
Ads-Heavy	1	5	229
Average	1.40	2.60	413.80

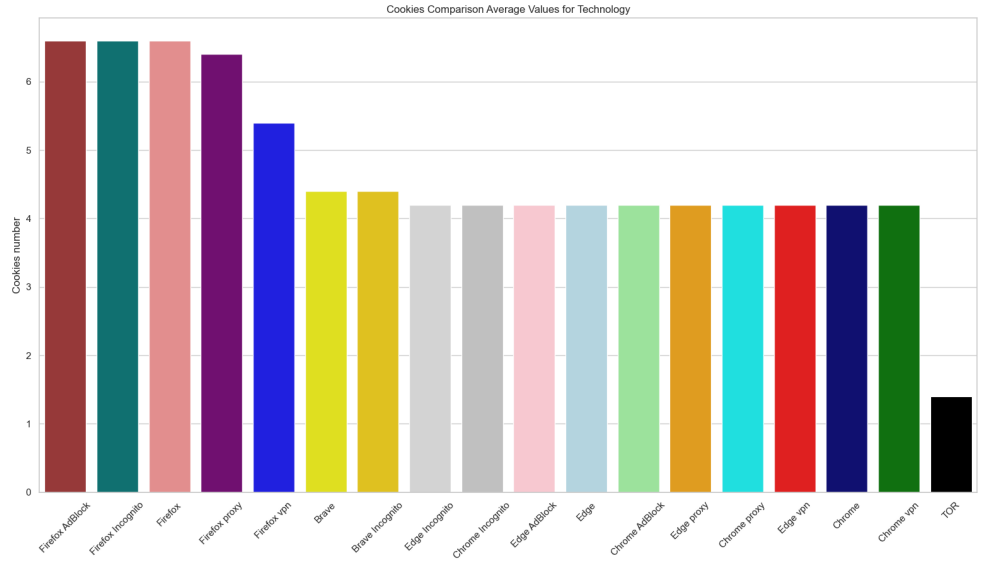
Tabela 5.19. Wyniki metryk prywatności dla technologii *TOR*

5.2.6. Analiza uzyskanych wyników

Kolejnym aspektem, pod kątem którego sprawdzane był poziom ochrony danych użytkowników i liczba generowanych o nich informacji. Dodatkowo oprócz wskaźników wymienionych wcześniej przy pomocy narzędzia *WireShark* zostało sprawdzone ujawnianie adresu IP oraz szyfrowanie połączenia. Wykresy zostały wygenerowane analogicznie jak w podrozdziale 5.1.6.

5.2.6.1. Pliki cookies

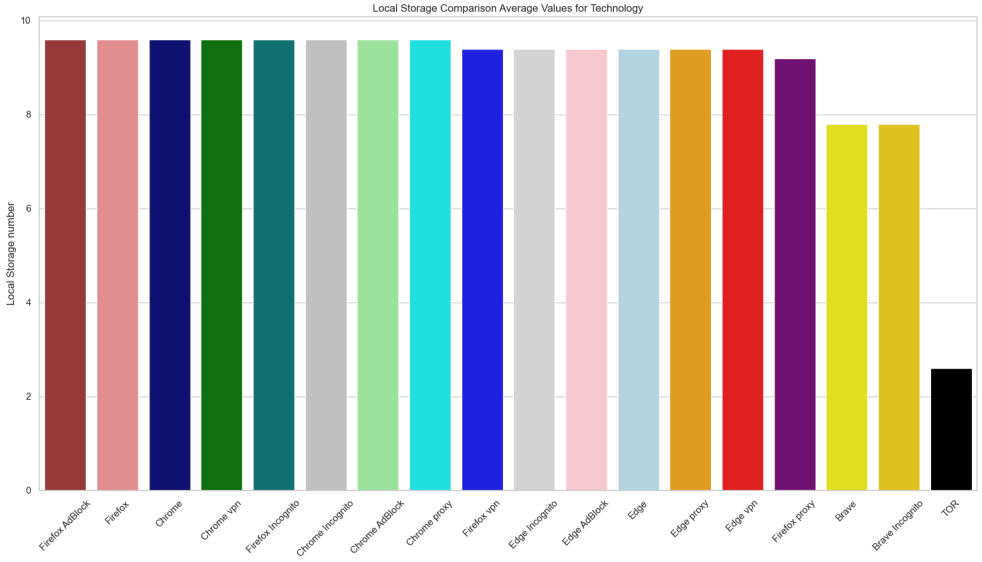
Patrząc na rysunek 5.11 można zauważyć, że wśród przeglądarek Firefox generuje najwięcej plików *cookies*, zwłaszcza na stronach informacyjnych (16) i w mediach społecznościowych (7), co daje mu najwyższą średnią (6.6) w porównaniu do innych przeglądarek. *Brave*, *Edge* i *Chrome* generują podobną liczbę *cookies*, z wartością średnią na poziomie około 4.2-4.4, przy czym *Brave* nie generuje żadnych *cookies* na stronach *„ads-heavy"*. Tryby *incognito*, wtyczki blokujące reklamy oraz użycie proxy w przeglądarkach nie mają wpływu na liczbę generowanych *cookies*. Z kolei użycie *VPN* w *Firefoxie* zmniejsza liczbę *cookies* na stronach *„ads-heavy"* do zera, co obniża jego średnią do 5.4. Przeglądarka *TOR* wyróżnia się najmniejszą liczbą generowanych cookies we wszystkich kategoriach, szczególnie na stronach *„news"* i *„social media"*, ze średnią wynoszącą zaledwie 1.4. *Chrome*, zarówno w trybie standardowym, jak i w trybach *incognito*, z *AdBlockiem* oraz z *proxy* czy *VPN*, zachowuje stałą liczbę generowanych *cookies*, co sugeruje brak różnic w zachowaniu przeglądarki w tych wariantach.



Rys. 5.11. Średnia generowana liczba *cookies* dla różnych technologii

5.2.6.2. Local Storage

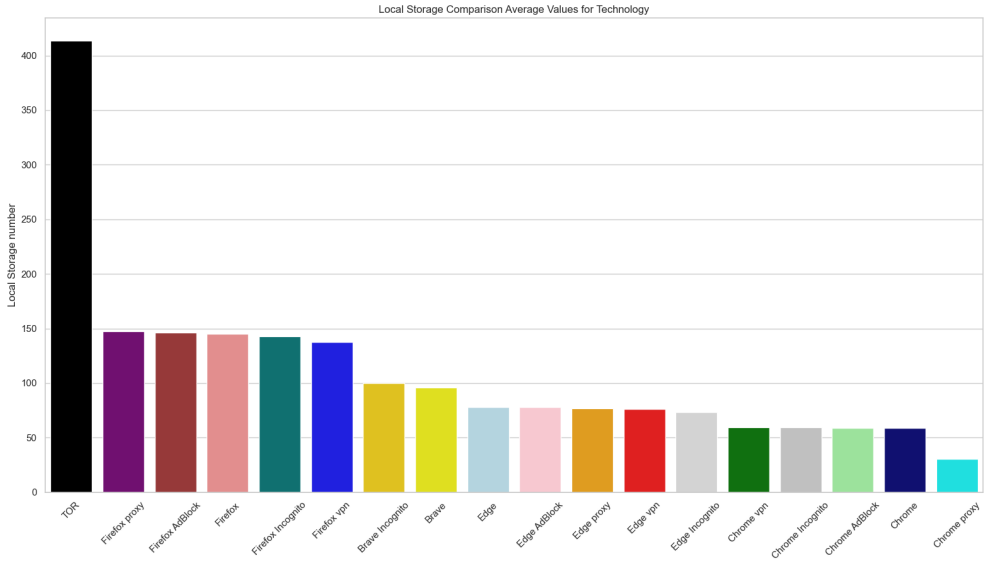
Jak wynika z rysunku 5.12, *Firefox* i *Chrome* generują najwięcej pamięci lokalnej, szczególnie na stronach *„ads-heavy"*, gdzie wartość ta osiąga 15-16 jednostek, a ich średnia wynosi 9.6. *Edge* również generuje znaczną ilość pamięci lokalnej (średnio 9.4), podobnie jak *Firefox*, z wyjątkiem sytuacji, gdy działa z proxy, gdzie średnia spada do 9.2. Przeglądarka *Brave*, zarówno w trybie standardowym, jak i *incognito*, generuje najmniej pamięci lokalnej spośród głównych przeglądarek, ze średnią 7.8.



Rys. 5.12. Średnia generowana liczba danych typu *local storage* dla różnych technologii

Tryby *incognito*, *AdBlock*, *proxy* oraz *VPN* nie mają znaczącego wpływu na ilość generowanej pamięci lokalnej, co sugeruje, że te technologie nie zmieniają sposobu, w jaki przeglądarki zarządzają *local storage* w analizowanych kategoriach stron. Oczywiście wyniki te mogą wydawać się dziwne. Wynika to z faktu, że o ile pamięć lokalna jest generowana w trakcie korzystania z Internetu w trybie *incognito* to po zakończonej sesji jest usuwana. Z kolei przeglądarka *TOR* generuje zdecydowanie najmniej pamięci lokalnej we wszystkich kategoriach, ze średnią wynoszącą jedynie 2.6, co potwierdza jej skuteczność w minimalizowaniu śladów użytkownika na stronach internetowych. Tak dobry wynik jest w stanie uzyskać dzięki ograniczeniu generowania wpisów do lokalnej pamięci analogicznie jak przy *cookies* w trakcie działania.

5.2.6.3. Ruch sieciowy



Rys. 5.13. Średni generowany ruch sieciowy dla różnych technologii

Pomijając TOR, *Firefox* generuje największy ruch sieciowy, szczególnie na stronach *„social media"* (300 jednostek) i *„ads-heavy"* (76 jednostek), co daje średnią 145.4. Podobnie jak *Firefox*, przeglądarka *Brave* również generuje dużą ilość ruchu, zwłaszcza na stronach *„news"* (134 jednostki) oraz *„streaming"* (195 jednostek), z średnią na poziomie 95.8. Z kolei *Chrome* generuje najmniejszy ruch na stronach *„basic"* i *„social media"*, z bardzo niskimi wartościami 1 jednostki i 2 jednostek odpowiednio, i średnią wynoszącą 58.8. Tryb *incognito* w przeglądarkach nie zmienia znacząco generowanego ruchu w porównaniu do trybu standardowego. Przykładowo, *Brave Incognito* generuje podobny poziom ruchu do *Brave* w trybie standardowym, ale z nieco wyższymi wartościami w kategorii *„news"* (154 jednostki). Wtyczki blokujące reklamy oraz użycie proxy wpływają na ruch sieciowy w ograniczonym zakresie, z niewielkimi różnicami w porównaniu do wersji standardowych przeglądarek. Przeglądarka *TOR*, znana z ochrony prywatności, generuje znacząco większy ruch w porównaniu do innych przeglądarek we wszystkich kategoriach, szczególnie na stronach *„social media"* (830 jednostek) i *„ads-heavy"* (229 jednostek), co skutkuje najwyższą średnią 413.8. Ten wysoki poziom ruchu widoczny na rysunku

5.13 może być wynikiem dodatkowych warstw zabezpieczeń i routingu, które *TOR* stosuje, aby zapewnić anonimowość użytkowników. W wypadku testowanej konfiguracji TOR posiadał domyślne trzy węzły i widać, że mniej więcej o tyle zwiększyło to generowanie ruchu sieciowego.

5.2.7. IP i szyfrowanie danych

Adres IP i szyfrowanie komunikacji są kluczowe dla prywatności i anonimowości w internecie. Adres IP identyfikuje i lokalizuje urządzenie w sieci, co może umożliwiać śledzenie aktywności użytkownika i narażać go na ataki cybernetyczne. Z kolei szyfrowanie komunikacji zabezpiecza dane przed przechwyceniem i nieautoryzowanym dostępem, chroniąc prywatność użytkownika i zapobiegając kradzieży tożsamości oraz innym formom cyberprzestępczości. Ukrywanie adresu IP oraz szyfrowanie połączeń, takie jak te oferowane przez *VPN* czy *TOR*, pozwala na korzystanie z internetu z większą swobodą i bezpieczeństwem, zabezpieczając użytkowników przed cenzurą, blokowaniem treści i monitorowaniem. Informacje, które rozwiązania potrafią to zapewnić znajdują się w tabeli 5.20.

Technologia	**Zmienia IP**	**Szyfruje komunikację**
Normalne korzystanie z przeglądarki	Nie	Nie
Tryb Incognito	Nie	Nie
AdBlock	Nie	Nie
Proxy	Tak	Nie (chyba że dodatkowo szyfrowane)
VPN	Tak	Tak
TOR	Tak	Tak

Tabela 5.20. Ukrywanie adresu IP oraz szyfrowanie komunikacji ze względu na technologię

W *TOR* szyfrowanie zaczyna się na urządzeniu użytkownika, gdzie dane są wielowarstwowo szyfrowane przed przesłaniem przez sieć węzłów. Każdy węzeł odszyfrowuje jedną warstwę szyfrowania, co zapewnia wysoki poziom anonimowości dzięki trasowaniu danych przez różne węzły i wielokrotnemu szyfrowaniu. Z kolei w *VPN* szyfrowanie odbywa się na urządzeniu użytkownika, a dane są przesyłane w zaszyfrowanej formie do serwera *VPN*. Serwer *VPN* odszyfrowuje dane, które następnie są przekazywane do docelowej witryny lub usługi. W tym przypadku, komunikacja między użytkownikiem a serwerem *VPN* jest zabezpieczona, ale warto mieć na uwadze, że dostarczyciel usługi *VPN* zna tożsamość użytkownika i to on odszyfrowuje dane, więc trzeba obdarzyć go pewnym zaufaniem.

5.3. Anonimowość w innych obszarach aktywności sieciowej

Chociaż praca skupia się na ukrywaniu tożsamości w przeglądarkach internetowych z różnymi możliwościami konfiguracyjnymi, jednakże temat anonimowości jest bardzo szeroko pojmowany, istotny również w innych aktywnościach użytkownika w sieci, dlatego warto przyjrzeć się jeszcze narzędziom zapewniającym bezpieczeństwo tożsamości w innych obszarach.

5.3.1. Komunikatory

Komunikatory odgrywają kluczową rolę w codziennym życiu przechowując niezliczoną ilość informacji o użytkownikach często bardzo prywatnych przeznaczonych do konkretnego odbiorcy, warto więc wybrać taki który zapewni najwyższy poziom bezpieczeństwa.

WhatsApp wykorzystuje szyfrowanie *end-to-end*, co oznacza, że wiadomości są szyfrowane od momentu ich wysłania aż do momentu ich odbioru przez adresata. Dzięki temu, nawet sama firma *WhatsApp* nie ma dostępu do treści wiadomości, co znacząco zwiększa poziom prywatności użytkowników [34].

Messenger, będący częścią ekosystemu *Facebooka* (teraz *Meta*), oferuje szyfrowanie *end-to-end*, ale tylko dla tzw. „czatów tajnych". Standardowe rozmowy na *Messengerze* są szyfrowane podczas transmisji, ale mogą być przechowywane na serwerach *Meta*, co może budzić wątpliwości dotyczące prywatności. Warto zaznaczyć, że *Facebook (Meta)* jest znany z intensywnego zbierania danych o użytkownikach, co może wpływać na postrzeganą anonimowość rozmów [35].

Telegram z kolei stosuje model szyfrowania, który różni się od podejścia *WhatsApp*. Telegram oferuje dwie opcje szyfrowania: „zwykłe" szyfrowanie dla chmurowych czatów oraz szyfrowanie *end-to-end* dla tzw. „czatów tajnych". Choć *Telegram* zapewnia szyfrowanie *end-to-end* w trybie „czatów tajnych", rozmowy w chmurze są dostępne na wszystkich urządzeniach, co może stwarzać pewne ryzyko związane z bezpieczeństwem [36].

Porównując te trzy platformy, warto zauważyć, że choć każda z nich oferuje pewien poziom szyfrowania, różnią się one w sposobie, w jaki przechowują dane i zapewniają prywatność. *WhatsApp* i *Telegram* oferują pełne szyfrowanie *end-to-end*, jednak *Telegram* stosuje różne poziomy szyfrowania w zależności od typu czatu. *Messenger*, z kolei, zapewnia szyfrowanie *end-to-end* tylko dla wybranych rozmów, co może wpływać na postrzeganą pewność ochrony prywatności. Ostateczny wybór platformy zależy więc od indywidualnych potrzeb użytkowników i ich priorytetów dotyczących bezpieczeństwa i anonimowości.

5.3.2. E-mail i tymczasowy adres e-mail

E-mail jest jednym z najstarszych i najczęściej używanych środków komunikacji w sieci. Pomimo swojej powszechności, e-mail może stanowić istotne wyzwanie w zakresie bezpieczeństwa i prywatności. W tradycyjnych systemach e-mail, takie jak *Gmail* czy *Outlook*, dane użytkowników, takie jak treść wiadomości, adresy nadawców i odbiorców, oraz metadane wiadomości (np. godzina wysłania), są przechowywane na serwerach dostawcy usługi. Przy posiadaniu jednego adresu e-mail do wielu kluczowych serwisów może wystąpić zagrożenie agregacji danych opisywane w podrozdziale 2.1.2.

Tymczasowe adresy e-mail są rozwiązaniem, które umożliwia użytkownikom zachowanie anonimowości i ochronę prywatności podczas rejestracji na stronach internetowych lub w celu uniknięcia spamowania. Usługi takie jak *Guerrilla Mail* czy *Temp Mail* oferują użytkownikom tymczasowe adresy e-mail, które można wykorzystać do odbierania wiadomości przez krótki czas, po czym są one automatycznie usuwane. Zaletą korzystania z tymczasowych adresów e-mail jest to, że pomagają one uniknąć długoterminowego śledzenia przez strony internetowe i ograniczają ryzyko otrzymywania spamu.

Warto jednak zauważyć, że tymczasowe adresy e-mail mają swoje ograniczenia. O ile są skuteczne w ochronie prywatności w krótkim okresie, to mogą być mniej odpowiednie do długotrwałej komunikacji, a także mogą nie oferować pełnej ochrony przed bardziej zaawansowanymi formami śledzenia i phishingu. Do istotnych serwisów należy wciąż korzystać z dostarczycieli standardowych renomowanych skrzynek pocztowych.

5.3.3. Hasła i ich przechowywanie

Hasła są fundamentalnym elementem systemów zabezpieczeń, używanym do autoryzacji i ochrony dostępu do danych i usług. Skuteczne zarządzanie hasłami jest kluczowe dla zapewnienia bezpieczeństwa użytkowników i systemów komputerowych. Istnieje wiele praktyk i technologii dotyczących przechowywania i zarządzania hasłami, które mają na celu minimalizację ryzyka ich kompromitacji. Większość dobrych praktyk i sugestii zostało zawarte w rozdziale 3.9.

Nawet najbardziej zaawansowane metody haszowania mogą zostać osłabione przez ludzkie zaniedbania i niewłaściwe praktyki zarządzania hasłami.

5.3.4. ZFone

Zfone wykorzystuje kryptografię asymetryczną i symetryczną do zabezpieczenia połączeń głosowych. Szyfruje zarówno sygnały głosowe, jak i metadane, takie jak adresy IP i informacje o czasie połączenia. W przeciwieństwie do innych rozwiązań, *Zfone* jest zaprojektowany w sposób, który nie wymaga zmiany infrastruktury istniejących systemów *VoIP*, co czyni go łatwym do wdrożenia w istniejących środowiskach telekomunikacyjnych. *Zfone* jest wykorzystywane w różnych obszarach, gdzie bezpieczeństwo komunikacji głosowej jest kluczowe. Jest to szczególnie ważne w biznesie, gdzie poufność rozmów może być wymagana przez prawo lub polityki wewnętrzne firmy. Z kolei W kontekście rządowym i wojskowym, może być używane do ochrony poufnych informacji przekazywanych przez głos w ramach operacji wywiadowczych i komunikacji strategicznej. *Zfone* znajdzie też zastosowanie w organizacjach zajmujących się ochroną praw obywatelskich i dziennikarzy, którzy potrzebują zabezpieczyć swoje rozmowy przed ewentualnymi próbami podsłuchu.

Jest to wartościowe narzędzie, lecz trudne do przetestowania i implementacji. Głównie nadaje się to wykorzystywania w specjalistycznych środowiskach [37].

5.3.5. Kryptowaluty

Technologia ta zapewnia transparentność, decentralizację i bezpieczeństwo, eliminując potrzebę centralnego organu, takiego jak bank. Kryptowaluty wykorzystują kryptografię do zabezpieczania transakcji oraz do kontroli tworzenia nowych jednostek. W przypadku *Bitcoina*, transakcje są weryfikowane i zapisywane w rozproszonym *ledgerze* zwanym *blockchain*, który jest publiczny i dostępny dla każdego. Anonimowość użytkowników w kryptowalutach może być zarówno zaletą, jak i wadą. *Bitcoin*, mimo że zapewnia pewien poziom anonimowości, jest bardziej pseudo niż całkowicie anonimowy. Oznacza to,

że chociaż adresy portfeli nie są bezpośrednio powiązane z tożsamością, można je śledzić i analizować w *blockchainie*, co może prowadzić do ujawnienia tożsamości w przypadku powiązań z rzeczywistymi danymi osobowymi (jak wypłaty z banków). W przeciwieństwie do *Bitcoina*, kryptowaluty takie jak *Monero* oferują bardziej zaawansowane funkcje anonimizacji, takie jak podpisy pierścieniowe w *Monero* które utrudniają śledzenie transakcji.

Kryptowaluty mają różne zastosowania, od inwestycji i spekulacji po codzienne płatności. *Ethereum*, na przykład, umożliwia tworzenie i zarządzanie inteligentnymi kontraktami i zdecentralizowanymi aplikacjami (*DApps*), co rozszerza możliwości poza proste transakcje finansowe. Jednakże, korzystanie z kryptowalut wiąże się z ryzykiem takim jak zmienność cen, ryzyko utraty kluczy prywatnych są istotnymi problemami, które mogą wpływać decyzje na użytkowników i inwestorów.

Z drugiej strony w porównaniu do tradycyjnych systemów finansowych, takich jak niższe opłaty transakcyjne, brak potrzeby pośredników oraz globalny zasięg. Brak regulacji, ryzyko związane z cyberatakami oraz zmienność cen są istotnymi wyzwaniami. W przeciwieństwie do tradycyjnych banków, które oferują ochronę depozytów i regulacje prawne [38].

5.4. Wnioski

5.4.1. Przeglądarki internetowe

Przeglądarki radzą sobie pod kątem wydajnościowym bardzo podobnie. Co ciekawe, wbrew popularnej opinii, *Chrome* wcale nie ma najwyższego zużycia pamięci RAM, co może wpływać na wydajność systemu przy intensywnym przeglądaniu, szczególnie w kontekście otwierania wielu kart. Jednakże prawdopodobnie posiadanie pięciu otwartych kart nie jest wystarczająco intensywnym użytkowaniem, aby tę teorię potwierdzić. Niestety, korzystanie w sposób natywny z przeglądarek nie zapewnia zaawansowanej ochrony.

5.4.2. Tryb prywatny/incognito

Tryb *incognito*, co dobitnie pokazały te badania, nie wykazuje zbyt dobrych właściwości ochrony danych użytkowników. Jedyne co robi, to czyszczenie pamięci lokalnej po zamknięciu sesji. Niestety, nie zapewnia ochrony przed innymi sposobami agregacji danych. Jego działanie ogranicza się do uniemożliwienia sprawdzenia historii wyszukiwania. Zapewnia zbliżone wyniki do trybu normalnego, choć jest nieznacznie wolniejszy właśnie przez nieprzechowywanie ustawień lokalnie.

5.4.3. AdBlock

Wpływ *AdBlocka* na komfort korzystania z internetu jest trudny do zmierzenia. Potrafi zaoszczędzić sporo czasu, szczególnie na platformach streamingowych takich jak *YouTube*, pomijając reklamy (co nie zalicza się do czasu ładowania strony), sprawić że strona będzie bardziej przejrzysta, ale z drugiej

strony może się okazać, że pewne funkcjonalności lub cała strona (jeśli posiada zaimplementowane mechanizmy antyadblock'owe) stanie się niezdatna do użytku. Usuwanie reklam może ograniczać skrypty śledzące, szczególnie jeśli w tego typu reklamę ktoś ma skłonności klikać. Dodatkowo nieraz potrafi zablokować złośliwe skrypty. Ostatecznie jest to przydatne narzędzie, z którego warto korzystać.

5.4.4. Proxy

Rozwiązanie to ma jeden istotny plus, którym jest przekierowanie połączenia ukrywające adres IP. Pozwala uzyskać dostęp do treści niedostępnych w danym rejonie geograficznym, ale nie ogranicza zbierania danych o użytkowniku, może je jedynie lekko mistyfikować. W wypadku niesprawdzonego serwera ujawnia jego właścicielowi bardzo dużo informacji. Wpływa na wydajność nieznacznie mniej niż *VPN*, lecz nie posiada jego dodatkowych atutów.

5.4.5. VPN

Dużym plusem tego rozwiązania jest ukrywanie adresu IP oraz tunelowanie połączenia. Jeśli chodzi o ograniczenie zapisywanych danych o użytkowniku nie gwarantuje specjalnej poprawy w stosunku do zwykłego korzystania z Sieci. Dzięki symulowanej lokacji potrafi jednak oszukać i zapisać mylne informacje, nie do końca oddające stan faktyczny. Potrafi tak jak proxy zapewnić dostęp do treści niedostępnych w danej lokalizacji. Ma umiarkowany wpływ na wydajność, który przy korzystaniu z dobrego łącza jest niezauważalny, ale jeśli internet jest gorszej jakości to może być to dostrzegalne.

5.4.6. TOR

Trasowanie cebulowe zapewnia najwyższy poziom ochrony użytkownika (chociaż warto mieć na uwadze problem kontrolowania węzłów wyjściowych), znacznie ograniczając poziom zapisywanych informacji. Warto zaznaczyć, że Tor posiada również różne poziomy bezpieczeństwa, które można wybrać w ustawieniach. Do tego stopnia, że mogą zostać zablokowane wszystkie rzeczy niewymagane do działania strony (takie jak np.*Javascript*) [39]. Dodatkowo narzędzie to ukrywa adres IP i stosuje wielokrotne szyfrowanie, dzięki czemu w trakcie transmisji nasze dane są bezpieczne. Dodatkowo pozwala na korzystanie ze stron *.onion* niedostępnych w inny sposób. Niestety, ale nawet w minimalistycznej wersji *Tor* działa wyraźnie wolniej od innych rozwiązań.

5.4.7. Podsumowanie

Niestety, nie ma jednej idealnej recepty na wykorzystanie technologii w życiu codziennym, tak aby uzyskać pełną anonimowość. Na nic mogą się zdać nawet najlepsze systemy ochrony jeśli tak jak w przypadku twórcy super prywatnego komunikatora *Telegram*, czyjaś dziewczyna wrzuca namiętnie zdjęcia z wakacji, po których udało się go znaleźć i aresztować [40]. Ten przykład dobitnie też pokazuje jak niechcianym przez rządy może być narzędzie, nad którym mają mocno ograniczoną kontrolą. Przy

wyborze odpowiednich narzędzi trzeba szczegółowo zapoznać się z ich charakterystyką oraz zastanowić się w jakim celu chcemy z nich korzystać. Trzeba rozważyć, na jakie ustępstwa jesteśmy w stanie sobie pozwolić. Weźmy za przykład *Tor*, który jako jedyny jest zauważalnie wolniejszy od innych rozwiązań w trakcie codziennego użytkowania. Może to być lekko irytujące, ale nie uniemożliwia korzystania z internetu. Dla niektórych może to być za dużo, a inni nawet nie zauważą różnicy. Jeśli po przeprowadzonych w ramach pracy badań trzeba było ułożyć jakąś listę proponowanych narzędzi i działań dla zwykłego użytkownika nieobeznanego w szczególny sposób z kwestiami technicznymi, to zawierałaby ona:

- zainstalowanie rozszerzenia *AdBlock* (dowolnego) dla preferowanej przeglądarki,
- korzystanie nawet z darmowego *VPN*,
- zapisywanie haseł w menadżerze haseł, trzymanym na dysku wymiennym,
- korzystanie z innego komunikatora niż *Messenger*,
- ograniczenie publikowania prywatnych i osobistych informacji w social-mediach,
- do serwisów mało istotnych czy użytkowanych jednorazowo korzystanie z tymczasowego adresu e-mail.

6. Podsumowanie

W niniejszej pracy dyplomowej skupiono się na kompleksowej analizie narzędzi i rozwiązań mających na celu zapewnienie anonimowości i bezpieczeństwa w sieci. W obliczu rosnącej liczby zagrożeń związanych z prywatnością online, celem było zrozumienie, jak różne technologie wpływają na ochronę danych osobowych i jakie oferują możliwości dla przeciętnego użytkownika.

W pierwszej części pracy zaprezentowano przegląd aktów prawnych regulujących ochronę danych i prywatności, co stanowiło teoretyczne fundamenty dla dalszych analiz. Następnie, szczegółowo omówiono różne technologie i narzędzia, takie jak *VPN*, *TOR*, *adblocki*, tryb *incognito* oraz *proxy* zwracając uwagę na ich funkcje, zalety i potencjalne wady.

Przeprowadzone badania miały na celu ocenę dwóch kluczowych aspektów: wydajności oraz skuteczności wybranych rozwiązań. Analizując wpływ tych narzędzi na prędkość działania systemów oraz zużycie zasobów, oceniono ich praktyczność i komfort użytkowania. Dodatkowo, zbadano zgodność deklaracji producentów z rzeczywistym poziomem ochrony, jaki oferują te technologie.

Stworzenie aplikacji w języku *Python*, która agreguje, przetwarza i raportuje dane statystyczne, umożliwiło porównanie wyników uzyskanych z innych źródeł badawczych i ogólnodostępnych informacji w sieci. Dzięki temu możliwe było sformułowanie wniosków na temat efektywności poszczególnych rozwiązań i zaproponowanie optymalnych strategii ochrony.

Wyniki wskazują, że chociaż pełna anonimowość w sieci jest niemożliwa do osiągnięcia, odpowiednie zastosowanie dostępnych narzędzi może znacząco poprawić bezpieczeństwo i prywatność użytkowników. W szczególności, różnorodność dostępnych rozwiązań pozwala na dostosowanie poziomu ochrony do indywidualnych potrzeb i preferencji, przy jednoczesnym uwzględnieniu wpływu na wygodę użytkowania. Ostateczne wnioski i rekomendacje mają na celu wsparcie użytkowników w wyborze najbardziej odpowiednich narzędzi, które najlepiej odpowiadają ich wymaganiom i poziomowi ochrony danych, jakiego oczekują.

6.1. Perspektywy rozwoju pracy

Oczywiście nie można uznać, że badania, jakie zostały dokonane w pracy są wystarczające i wyczerpują zagadnienie anonimowości oraz rozwiązań ją zapewniających. Jest wiele innych narzędzi wartych sprawdzenia, które ze względu na ograniczone możliwości czasowe i zasobowe zostały pominięte. Kluczowym aspektem mogącym poprawić wyniki byłoby lepsze izolowanie procesów, w ramach których

testy są wykonywane tak, aby czynności czy procesy w tle miały jak najmniejsze znaczenie. Dodatkowo warto byłoby rozszerzyć sprawdzanie o filtrowanie i kategoryzację plików *cookies* ze względu na ich pochodzenie oraz funkcje. Dałoby to lepszy obraz co właściwie jest zbierane i kto to zbiera oraz mogło pomóc w analizie poprzez dodanie czynnika jakościowego oprócz jedynie sprawdzenia ilościowego. Biorąc pod uwagę zagrożenie wynikające z *JavaScriptu* znajdującego się niemal na każdej stronie internetowej, należałoby sprawdzić, jakie działania są wykonywane przez skrypty tego języka w tle i na ile są one niebezpieczne. Ciekawym aspektem jest również porównanie narzędzi od różnych dostawców w ramach tej samej technologii, jak np. sprawdzenie jaki wpływ na wyniki ma wybór dostawcy *VPN*, szczególnie w kontekście rozwiązań darmowych i prywatnych. Bardzo interesującym aspektem jedynie napomkniętym w tej pracy jest również zbieranie danych w trakcie korzystania z kont powiązanych. Szczególnie takich dostawców jak *Google*, którzy niemal wymagają tego na swoich serwisach jak *YouTube*, mogłoby to uwidocznić zalety rozwiązań takich jak tymczasowe adresy e-mail. Dodatkowo wartościowe mogłoby się okazać rozszerzenie sprawdzania aplikacji na inne systemy operacyjne. Niestety, jedną rzeczą, której nie udało się uzyskać w tej pracy była unifikacja testów ze względu na problemy z przeglądarka *Tor*. Dane zostały zebrane, ale w inny i mniej zautomatyzowany sposób, co należałoby poprawić. W przyszłości można też rozwinąć projekt o dokładne sprawdzenie komunikatorów wraz z szyfrowaniem i danymi uzyskiwanymi przez koncerny w trakcie ich użytkowania.

Bibliografia

[1] Marcin Maj Niebezpiecznik. *Wyciekły wyniki badań tysięcy Polaków, którzy oddali krew do badań w ALAB. https://niebezpiecznik.pl/post/wyciekly-wyniki-badan-tysiecy-polakow-z-firmy-alab/*. Dostęp: (12-09-2024). 2024.

[2] Gabriel Weimann. „Terrorist migration to the dark web". W: *Perspectives on Terrorism* 10.3 (2016), s. 40–44.

[3] Amir M Hormozi. „Cookies and privacy". W: *Information Security Journal* 13.6 (2005), s. 51.

[4] Robert Bond. „The EU E-Privacy Directive and Consent to Cookies". W: *The Business Lawyer* 68.1 (2012), s. 215–223. ISSN: 00076899.

[5] Randika Upathilake, Yingkun Li i Ashraf Matrawy. „A classification of web browser fingerprinting techniques". W: *2015 7th International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE. 2015, s. 1–5.

[6] Mary Madden i in. „Teens, social media, and privacy". W: *Pew Research Center* 21.1055 (2013), s. 2–86.

[7] Asadzadeh Laleh i Rahimi Shahram. „Analyzing Facebook activities for personality recognition". W: *2017 16th IEEE international conference on machine learning and applications (ICMLA)*. IEEE. 2017, s. 960–964.

[8] Katharine Jarmul. *Practical Data Privacy*. "O'Reilly Media, Inc.", 2023.

[9] Brian Greer. „The growth of cybercrime in the United States". W: *Growth* (2017).

[10] Miles Pollard i Lawson Mansell. „A Case Study of Russian Cyber-Attacks on the Ukrainian Power Grid: Implications and Best Practices for the United States". W: *Pepperdine Policy Review* 16.1 (2024), s. 1.

[11] Cezary Banasiński i in. *Cyberbezpieczeństwo*. Wolters Kluwer, 2020.

[12] Parlament Europejski i Rada. *Dyrektywa 95/46/WE Parlamentu Europejskiego i Rady z dnia 24 października 1995 r. w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych. https://eur-lex.europa.eu/legal-content/PL/ALL/?uri=CELEX\%3A31995L0046*. Dostęp: (12-09-2024). 2003.

[13] European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj*. Dostęp: (12-09-2024). 2016.

[14] Anna Pacyga. *Polityka ochrony danych osobowych w systemach e-administracji w Polsce*. url=https://ruj.uj.edu.pl/xmlui/handle/item/227860. Dostęp: (12-09-2024).

[15] Thorin Klosowski. „The state of consumer data privacy laws in the US (and why it matters)". W: *New York Times* (2021).

[16] Sagar Pol. *BRIDGING THE ATLANTIC DIVIDE: A COMPREHENSIVE STUDY OF PERSONAL DATA TRANSFER BETWEEN THE EU AND THE US THROUGH ADEQUACY DECISION*. url=https://ruj.uj.edu.pl/handle/item/328855. Dostęp: (12-09-2024).

[17] Emmanuel Pernot-Leplay. „China's approach on data privacy law: a third way between the US and the EU?" W: *Penn St. JL & Int'l Aff.* 8 (2020), s. 49.

[18] Alan Grosskurth i Michael W Godfrey. „A reference architecture for web browsers". W: *21st IEEE International Conference on Software Maintenance (ICSM'05)*. IEEE. 2005, s. 661–664.

[19] Graeme Horsman. „A process-level analysis of private browsing behavior: A focus on Google Chromes Incognito mode". W: *2017 5th International Symposium on Digital Forensic and Security (ISDFS)*. IEEE. 2017, s. 1–6.

[20] Bartłomiej Olechowski Rynek Informacji. *Różnice pomiędzy VPN i Proxy. https://rynekinformacji.pl/wp-content/uploads/2023/02/VPN-vs-proxy-1024x576.png*. Dostęp: (12-09-2024). 2020.

[21] Abid Khan Jadoon i in. „Forensic analysis of Tor browser: a case study for privacy and anonymity on the web". W: *Forensic science international* 299 (2019), s. 59–73.

[22] Bram Emmen, Christianne de Poot i Wouter Stol. „What are they doing in the dark: Police strategies and working methods in fighting crime on the Tor Network". W: *European Journal of Policing Studies* 6.4 (2023), s. 1–21.

[23] Yevhenii Tsyliurnyk, Oleksandr Tomenchuk i Grzegorz Kozieł. „Badanie wiedzy użytkowników w zakresie bezpieczeństwa komunikatorów internetowych". W: *Journal of Computer Sciences Institute* 26 (2023).

[24] Anne Adams, Martina Angela Sasse i Peter Lunt. „Making passwords secure and usable". W: *People and computers XII: proceedings of HCI'97*. Springer. 1997, s. 1–19.

[25] Satoshi Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System*. url=https://bitcoin.org/bitcoin.pdf. Dostęp: (12-09-2024).

[26] Nicolas Van Saberhagen. *CryptoNote v 2.0*. url=https://github.com/monero-project/research-lab/blob/master/whitepaper/whitepaper.pdf. Dostęp: (12-09-2024).

[27] Bayu Solehudin Arip Priyatna i Nono Heryana. „Analysis Effect of Zfone Security on Video Call Service in Wireless Local Area Network". W: *International Journal of Computer Techniques* 6 (list. 2019), s. 1–5.

[28] Python Software Foundation. *What is Python? https://www.python.org/doc/essays/blurb/?external_link=true*. Dostęp: (12-09-2024). 2024.

[29] Software Freedom Conservancy. *About Selenium. https://www.selenium.dev/about/*. Dostęp: (12-09-2024). 2024.

[30] SQLite. *About SQLite. https://www.sqlite.org/about.html*. Dostęp: (12-09-2024). 2024.

[31] StatCounter. *Browser Market Share Worldwide from May 2010 to May 2024*. Dostęp: (12-09-2024). 2024.

[32] FineProxy. *Lista darmowych proxy. https://fineproxy.org/pl/free-proxy/*. Dostęp: (12-09-2024).

[33] Wireshark Foundation. *Wireshark User's Guide Version 4.5.0. hhttps://www.wireshark.org/docs/wsug_html/*. Dostęp: (12-09-2024). 2024.

[34] Martijn Terpstra. „WhatsApp & privacy". W: *Netherlands: Radboud University Nijmegen* (2013).

[35] Erica Jaeger. „Facebook Messenger: Eroding user privacy in order to collect, analyze, and sell your personal information". W: *J. Marshall J. Info. Tech. & Privacy L.* 31 (2014), s. i.

[36] Javokhir Komiljonov. „How Safe Telegram Is to Keep Personal Data and Conversations". W: *EUROPEAN JOURNAL OF BUSINESS STARTUPS AND OPEN SOCIETY* 4.2 (2024), s. 54–57.

[37] Martin Petraschek i in. „Security and Usability Aspects of Man-in-the-Middle Attacks on ZRTP." W: *J. Univers. Comput. Sci.* 14.5 (2008), s. 673–692.

[38] Sabine Houy, Philipp Schmid i Alexandre Bartel. „Security aspects of cryptocurrency wallets—a systematic literature review". W: *ACM Computing Surveys* 56.1 (2023), s. 1–31.

[39] Tor Browser. *Tor Browser Security Settings. https://tb-manual.torproject.org/security-settings/*. Dostęp: (12-09-2024). 2024.

[40] Tribunal Judiciaire de Paris. *Komunikat prasowy na temat zatrzymania Pavela Durova. https://www.tribunal-de-paris.justice.fr/sites/default/files/2024-08/2024-08-26\%20-\%20CP\%20TELEGRAM\%20.pdf*. Dostęp: (12-09-2024). 2024.

Spis rysunków

Spis tabel