Hyökkäykset I2P-verkkoa ja sen käyttäjiä vastaan

LuK-tutkielma
Turun yliopisto
Tietotekniikan laitos
Tietojenkäsittelytiede
2024
Henri Haapanen

TURUN YLIOPISTO
Tietotekniikan laitos

Henri Haapanen: Hyökkäykset I2P-verkkoa ja sen käyttäjiä vastaan

LuK-tutkielma, 23 s.
Tietojenkäsittelytiede
Helmikuu 2024

Kasvaneen yritys- ja valtiotason tarkkailun sekä sensuurin vuoksi on monilla ihmisillä herännyt huolia sananvapauden ja yksityisyyden jatkuvuudesta. Tämän seurauksena anonyymien salaverkkojen suosio kasvanut, sillä ne tarjoavat käyttäjilleen salatun kommunikaatioväylän ja mahdollisuuden ylläpitää palveluita ilmaiseksi verkossa. Tämän vuoksi salaverkkoja vastaan suoritetaan yhä enemmän hyökkäyksiäi, joilla pyritään paljastamaan käyttäjien henkilöllisyys tai toiminta. Hyökkäyksiä kohdistetaan myös itse salaverkkoja kohtaan, muun muassa palvelunestohyökkäyksinä.

Tämä tutkielma on toteutettu kirjallisuuskatsauksena ja sen tavoitteena on luoda katsaus hyökkäyksistä, joita voidaan käyttää valkosipulireititykseen pohjautuvaa I2P-salaverkkoa ja sen käyttäjiä vastaan. Jokaisen esitellyn hyökkäystyypin kohdalla on pyritty tuomaan esiin sen toiminta, vakavuus ja mahdolliset suojautumiskeinot.

Tutkielmassa osoitetaan, että suuri osa matalan kynnyksen heikoista hyökkäyksistä on suoraviivaista mitätöidä ja laajat verkkotason hyökkäykset ovat muuttuneet verkon koon ja tietoturvapäivitysten myötä vaikeiksi toteuttaa. Sen sijaan kohdistettu tietoliikenneanalyysi, jonka avulla yksittäisen käyttäjän toiminta verkossa voidaan selvittää, on pysynyt mielekkäänä hyökkäyskeinona. Tor-verkkoon kohdistettu tutkimus on huomattavasti laajempi kuin I2P-verkkoa käsittelevä tutkimus, ja sen takia olisikin hyödyllistä tarkastella sitä, miten näiden kahden salaverkon tutkimusta voisi tuoda yhteen. Hyökkäyksiä Tor-verkkoa vastaan on tutkittu laajasti, jäljelle jää tutkittavaksi näiden hyökkäysten soveltuvuus I2P-verkkoa vastaan.

Asiasanat: I2P, salaverkko, tietoturva, anonyymi verkko, pimeä verkko, anonymiteetti, vertaisverkko

Sisällys

1 Johdanto

Internetin muututtua jokapäiväiseksi palveluksi informaatioyhteiskunnan tueksi on monilla sen käyttäjistä herännyt huoli yritys- ja valtiotason tarkkailusta. Vastauksena tähän kasvavaan kysyntään on syntynyt useita yksityisyyskeskeisiä vaihtoehtoja, niin kutsutut anonyymit verkot tai salaverkot kuten, Tor (The Onion Router) ja I2P (Invisible Internet Project), joista Tor alun perin suunniteltiin tiedusteluyhteisölle. Nämä salaverkot ovat hupenevan sananvapauden ja lisääntyneen sensuurin myötä kasvattaneet suosiotaan [1][2]. I2P-verkossa oli helmikuussa 2024 arviolta noin 80 000[1] vertaista eli reititintä, Tor-verkon kotisivujen arvio päivittäisistä käyttäjistä liikkuu miljoonissa[2]. Nämä kaksi lukua eivät edusta täysin samaa asiaa, mutta kokoero on kiistaton. Näiden ohella myös tunnelointipalvelut eli VPN:t ovat nousseet suosituiksi. Digitaalimaailma on siis lujittanut otettaan jokapäiväisessä elämässä ja edellä mainitut salaverkot palvelevat kasvavaa käyttäjäkuntaa, joka tiedostaa tämän sekä tietotalouden tuomat lieveilmiöt ja haittapuolet nykypäivänä ja etenkin tulevaisuudessa [3]. Nämä salaverkot tarjoavat yksilöille suojatun kommunikaatioväylän ja ilmaisia verkko-osoitteita palveluiden ylläpitäjille, verkoissa tapahtuu tarjotun yksityisyyden ansiosta myös huomattavasti laitonta toimintaa ja valuuttavirrat kasvavat vuosi vuodelta [4].

Kiinnostuksen kasvu on johtanut myös lukuisiin yrityksiin tutkia ja löytää haavoittuvaisuuksia salaverkkojen toiminnasta ja etenkin Torin sipulireitityst ä on tut-

[1]Pääkehittäjän, *zzz*, sivulta: http://stats.i2p/cgi-bin/total_routers_month.cgi

[2]https://metrics.torproject.org/userstats-relay-table.html

kittu jo laajemmin, sen lippulaivastatuksen ja helppokäyttöisyyden ansiosta. Huomattava osa tuhoisimmista hyökkäyksistä nojaa suuren vertaisparven luomiseen ja sen hallintaan. Vertaisparvea hyödyntämällä hyökkääjä voi valvoa tietyn verkkoalueen, osaverkon, liikennettä ja vertaisia; yrittää tulkata miten usein, kenen kanssa ja minkä tyyppisiä viestejä lähetetään. Vertaisparvea voidaan myös käyttää palvelunestohyökkäyksissä. Salaverkon koon kasvaessa tämäntyylinen lähestyminen käy kalliiksi ja analysointiparvi voi herättää ulkopuolisten huomion, mutta hyökkääjä voi onnistuessaan paikantaa kohteensa sijainnin verkossa ja lopulta myös yhdistää sen tiettyyn IP-osoitteeseen. Tämän seurauksena organisaation tai yksilön fyysinen sijainti tai asuinpaikka paljastuu [5].

Tutkielman tarkoituksena on luoda katsaus salaverkkojen toimintaan, syventyä eritoten I2P-verkon toimintaan ja esitellä erilaisia hyökkäyksiä, joiden kautta on yritetty murtaa I2P-verkon yksityisyys ja tietoturva. I2P:n tapauksessa suuri osa hyökkäyksistä on teoreettisia eikä niitä ole onnistuneesti suoritettu ainakaan julkisesti. Suojautumiskeinoja ja suosituksia on myös yritetty tuoda esiin valottamaan hyökkäysten konkreettisten haittojen laajuutta. Tarkastellaan siis mitkä niistä ovat jatkuvia uhkia ja onko niihin suoraviivaisia ratkaisuja. Salaverkkojen toimintaa tutkitaan sipulireitityksen kautta, koska sekä Tor että I2P käyttävät sitä pohjanaan. Toteutukset ovat kuitenkin erilaiset ja I2P:n hyödyntämää versiota kutsutaan tarkemmin valkosipulireititykseksi, mutta käsitteet ja komponentit ovat pääosin rinnastettavissa [6][7].

Tutkimuskysymykset

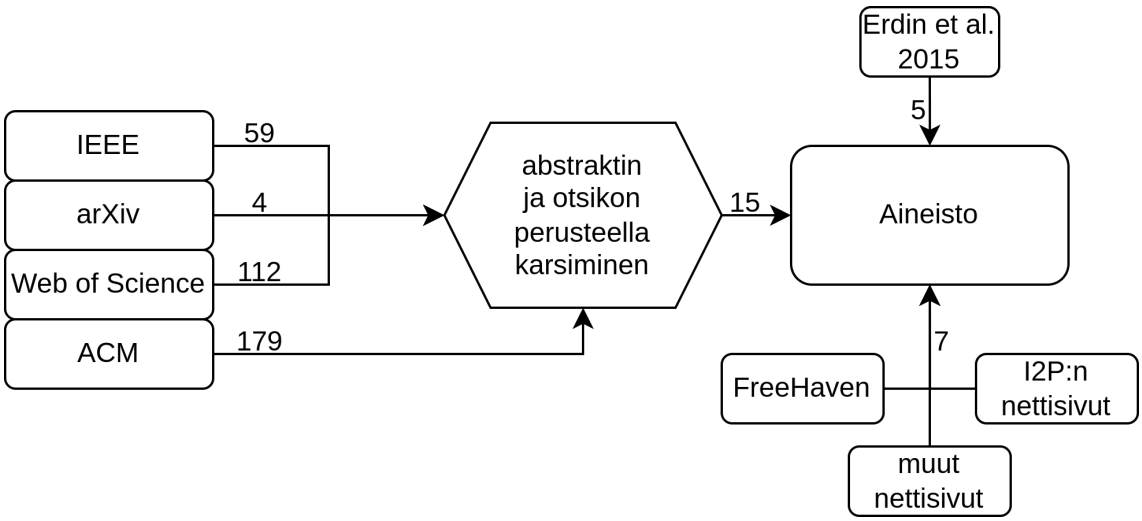
TK1: Millä menetelmillä I2P:n tietoliikennettä on yritetty murtaa ja analysoida?

TK2: Millaisia keinoja on käytetty näiden hyökkäysten estämiseksi?

TK3: Mitkä näistä hyökkäyksistä ovat haitallisimpia tai vaikeimmin torjuttavia?

Aineisto ja menetelmät

Aineisto koottiin aluksi seuraavista tietokannoista: IEEE, arXiv, ACM ja Springer. Hakuterminä oli pelkkä ”I2P”, sillä kaikki testatut lisäehdot rajasivat tuloksia liikaa. Tämän jälkeen sitä laajennettiin hakemalla alkuaineiston viitteistä asiaankuuluvia julkaisuja, näiden lisäksi viitteitä ja tietoa etsittiin erinäisistä opinnäytetöistä ja näin saatiin laajempi, kattavampi kokonaiskuva kuin hakutietokannat alun perin antoivat ymmärtää. FreeHaven-projektin ylläpitämä julkaisukokoelma[3] ja I2P:n kotisivut olivat erittäin hyödyllisiä kapea-alaisen, syvän tietonsa ansiosta. Erdin et al. 2015 julkaisu ”How to Find Hidden Users: A Survey of Attacks on Anonymity Networks” oli myös tärkeä, sillä se tarjoaa kattavan kuvan hyökkäyksistä ja miten näitä erilaisia hyökkäyksiä on mahdollista soveltaa sekä Tor- että I2P-verkossa.



Kuva 1.1: Aineistokaavio.

Tutkielma on jaettu kolmeen osaan, ensiksi käydään läpi sipulireititys, siihen liittyvät käsitteet ja sen tunnetuimmat toteutukset, jotta lukija saa tarvittavat pohjatiedot. Toisessa osassa käydään läpi eri hyökkäyksiä, joita on kehitetty valkosipulireititykseen pohjautuvaa I2P-salaverkkoa vastaan. Hyökkäykset luokitellaan, niiden

[3]https://www.freehaven.net/anonbib/

toiminta käydään läpi ja jokaisen hyökkäyksen kohdalla on selitetty niiden tehokkuus, vaativuus sekä mahdolliset suojautumiskeinot niitä vastaan. Lopuksi kootaan havainnot ja käydään läpi johtopäätökset.

2 Salaverkot

Hyökkäyksien ymmärtämisen kannalta tässä luvussa käydään läpi salaverkkojen toimintaperiaatteet ja niiden menestyneimmät ilmentymät. Kokonaiskuvan kannalta tärkeiksi asioiksi nousee myös miten salaverkot laajentuvat tavallisen Internetliikenteen päälle ja miten niiden toteutukset eroavat toisistaan sekä miten nykypäivänä on mahdollista reitittää verkkoliikennettä siten, ettei kukaan muu kuin sen haluttu vastaanottaja kykene selvittämään sen sisältöä eikä verkkovirran päätepisteiden osapuolia kyetä tunnistamaan ulkopuolelta.

2.1 Sipulireititys

Sipulireititys on yleiskäyttöinen tiedonsiirtotapa, joka kehitettiin 1990-luvun jälkipuoliskolla ja jota soveltamalla voidaan luoda tiedonsiirtokerros, joka mahdollistaa salaisen viestimisen julkisessa verkossa. Se tarjoaa salatun yhteyden, jota ei kyetä helposti salakuuntelemaan ja jonka tuottamaa verkkoliikennettä on vaikea analysoida. Yhteydet ovat kaksisuuntaisia ja reaaliaikaisia, vastaanottaja voi siis käyttää samaa reittiä lähettämiseen kuin vastaanottamiseenkin korkeintaan muutaman sekunnin viiveellä [7]. Jotta tiedonsiirto kerroksessa onnistuu, on jokaisen osapuolen käytettävä siihen suunniteltua ohjelmaa. Tälläisen järjestelmän toiminta on seuraavanlainen: käyttäjän ohjelmainstanssi luo *sipulin* eli kerroksittaisen tietorakenteen, jossa jokainen kerros edustaa jotain verkossa olevaa ohjelmainstanssia, jonka läpi viesti halutaan kuljettaa. Kun ohjelma on saanut reitin selville vertaisprofiloinnin

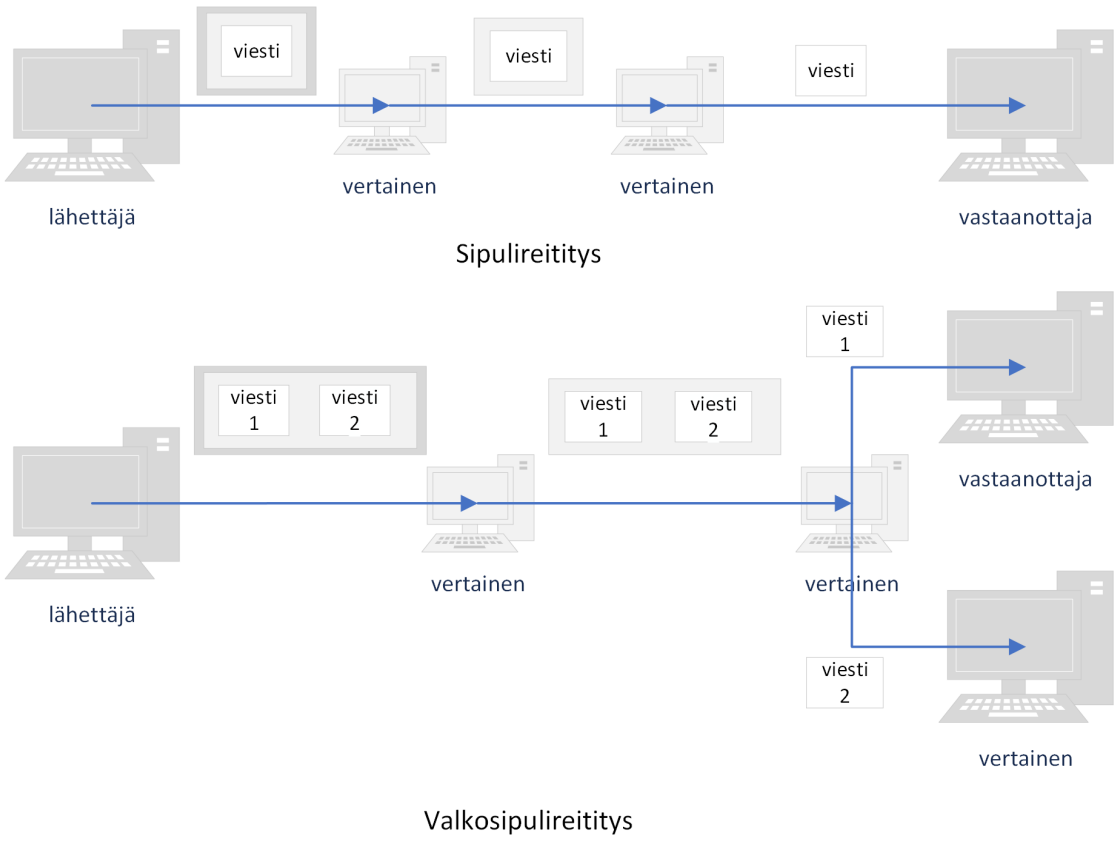
avulla, se kokoaa matkan varrella olevien välipysäkki-instanssien tarvitsemat tiedot kerroksiin ja salaa jokaisen kerroksen siten että jokainen välipysäkki kykenee avaamaan vain juuri sille instanssille tarkoitetun kerroksen. Näin ollen jokaiselle instanssille paljastuu vain ne reititysohjeet, joita se tarvitsee edelleenlähettääkseen jäljelle jääneen sipulin. Tämä prosessi toistuu kunnes sipulin ydin saapuu päätepisteen vastaanottajalle. Reitillä olevat instanssit eivät tiedä datapaketin koko reittiä eivätkä sen sisältöä, jolloin toteutuu sipulireitityksen salatun ja anonyymin viestinnän tavoitteet eli sekä sisältö että osapuolet pysyvät salassa. [8]

2.2 Valkosipulireititys ja I2P

Valkosipulireititys esiteltiin vuonna 2000 ja se rakentuu sipulireititysmallin päälle muutamin parannuksin. Sipulireitityksessä määriteltiin yksi viesti joka piilotettiin yhteen sipuliin yhdelle vastaanottajalle, kun taas valkosipulireitityksessä yhdessä *valkosipulissa* voi olla monta viestiä, erillisin reititysohjein, eri vastaanottajille. Kuvassa 2.2 tätä eroa on selkeytetty, vastaavia kuvia on esitetty useissa artikkeleissa [9]–[11]. Usean viestin sisällyttäminen yhden valkosipulin sisään tarkoittaa sitä, että jokaisella niin sanotulla kynnellä täytyy olla omat reititysohjeensa, joiden avulla ne voidaan edelleenlähettää määränpäihinsä, sen jälkeen kun ensimmäinen vastaanottaja tai muu reititykseen osallistuva instanssi on vastaanottanut valkosipulin ytimen. Etuna tässä monimutkaisemmassa menettelyssä ovat toimituksen luotettavuus ja kestävyys lisääntyneiden reititysmahdollisuuksien myötä. Datapaketti voi mennä montaa eri reittiä kohteeseensa. I2P:n toteutuksessa kaikki verkkoon osallistuvat ohjelmainstanssit ovat vertaisia eli jokainen lähettää, välittää sekä vastaanottaa näitä valkosipuleita. Mikäli yksi reitin varrella oleva vertainen kaatuu tai ei vastaa, niin viesti menee muiden vertaisten kautta perille. [6][12]

I2P:n pääpiirteinä voidaan pitää sen hajautettua ja suljettua verkkomallia. Keskitetty infrastruktuuri ja yhteydet ulkoisiin verkkoihin on minimoitu. Tämä salattu

vertaisverkko on erillinen normaalista Internetistä, ja toisin kuin Tor, I2P tarjoaa yhtenäisemmän käyttäjäkokemuksen, sillä se on suunniteltu olemaan oma verkkonsa; ohjelman mukana tulee alkeellinen sähköpostiohjelma, BitTorrent-sovellus sekä WWW-palvelin ohjeineen [1]. Sovellukset I2P-verkossa voivat kommunikoida keskenään OSI-mallin mukaisen kuljetuskerroksen päällä käyttäen TCP:n kaltaista NTCP-protokollaa tai UDP:n kaltaista SSU-protokollaa. I2P-reititin eli käyttäjän paikallinen ohjelmainstanssi tulkkaa nämä viestit ja hoitaa datapakettien lähettämisen *I2P-tunneleiden* kautta määränpäihinsä [5]. Tunneli muodostuu vertaisista, joiden kautta voidaan lähettää tai vastaanottaa datapaketteja toisen osapuolen kanssa. Se on siis ohjelmainstanssien muodostama polku verkon läpi ja jokainen polku on luotu joko vastaanottamiseen tai lähettämiseen. I2P-reititin luo tunneleita ja sisältää rajapinnan muille ohjelmille, jonka ansiosta ulkoisten ohjelmien on suoraviivaista lähettää viestejä verkkoon ja vastaanottaa niitä myös sieltä [1].



Kuva 2.1: Sipulireititys ja valkosipulireititys. Pohjautuu ”Empirical Measurement and Analysis of I2P Routers” (Liu et al. 2014) julkaisussa esitettyyn kuvaan.

NetDB ja vertaisten profilointi

Pelkkien tunneleiden luonti ei riitä kuvaamaan I2P:n tiedonsiirtoprosessia, joten nyt selvitetään miten nämä ohjelmainstanssit löytävät toisensa, jotta yhteyksiä voidaan ylipäätään luoda. I2P-verkko ylläpitää tähän tarkoitukseen hajautettua tietokantaa, joka rakentuu Kademlia-algoritmin ja hajautetun tiivisteen päälle (DHT, *Distributed hash table*), ja tämä *netDB* sisältää verkossa olevista vertaisista kahdenlaista dataa: ohjelmainstanssin uniikin verkkotunnisteen (*RouterInfo*) ja reittiohjeita verkossa oleviin palveluihin (*LeaseSet*) [13][14][15]. Jälkimmäinen sisältää joukon tunneleita ja vastaanottajan julkisen salausavaimen, joiden avulla voi muodostaa sala-

tun yhteyden niiden määrittämään palveluun. Pieni osa I2P-verkon vertaisista, joilla on suuri kaistanleveys, ylläpitää tätä hajautettua tietokantaa. Näitä vertaisia kutsutaan *floodfill*-reitittimiksi. NetDB varmistaa edellä mainitun metadatan tasaisen jakautumisen verkon eri osiin. Tämän hajautetun rakenteen ansiosta verkon skaalautuvuus ei ole ongelma I2P:n mallissa. [12] Mahdollisimman laadukkaiden tunneleiden luontia varten jokaisen solmun tulee täyttää tietyt kriteerit. Tämän vuoksi *RouterInfossa* on mukana siihen liittyvän reitittimen asettamia lippuja, jotka kertovat muun muassa sen kaistanleveydestä ja nykyisestä taakasta. Näiden avulla tehdään vertaisvalintaa, mutta koska kyseessä ei ole naiivi ensimmäisen sukupolven salaverkko, ei reitittimien itseraportoituun tietoon voida luottaa ja sitä kohdellaan vain suuntaa antavana; jokainen reititin pitää itse kirjaa vertaistensa suorityskyvystä ja käyttää omia havaintojaan lopullisessa päätöksessä. [15]

3 Salaverkkojen murtaminen

Luvussa esitellään hyökkäyksien kirjo, muodot ja niihin liittyvä tutkimus. Suurin osa tutkimuksesta on keskittynyt Torin ja yleisesti sipulireitityksen murtamiseen, I2P:n osuus on verrattaen pieni mutta kuten edellisessä luvussa todettiin, on näiden kahden salaverkkototeutuksen jakama perusta laaja. Tästä seuraa se, että huomattava osa Torin tietoturvatutkimuksesta kattaa myös I2P:n [16].

3.1 Hyökkäyksistä

Kun tarkoituksena on luoda laaja ja selkeä kartta erilaisista tavoista, joilla kohteen identiteetti on yhdistetty reaalimaailmaan, kohteen viestien määränpää saatu selville tai tämän pääsy salaverkkoon on estetty, on selkeyden vuoksi hyvä jakaa hyökkäykset omiin kategorioihinsa. I2P:n kaltaisen vertaisverkon tapauksessa hyökkäyksiä kohdistetaan muun muassa vertaisten valintaan ja uhrin verkkokyselyiden estämiseen, nämä edellä mainitut toteutetaan keinotekoisella vertaisparvella joka mahdollistaa kohteen eristämisen muusta verkosta sen valvomiseksi ja sen tuottaman tietoliikenteen analysoimiseksi [5]. Muita hyökkäyksiä on sovelluskerroksessa (*application layer*) eli ei itse I2P-verkon toteutuksessa vaan sen päällä toimivissa kolmannen osapuolen ohjelmissa, esimerkiksi metadatan vuodattaminen tekemällä virheellisiä kyselyitä kohteen palvelimelle tai houkuttelemalla kohde vierailemaan hyökkääjän nettisivulla, joka on suunniteltu kiertämään I2P-verkon suoja erilaisten käyttäjäpuolella ajettavien pienoisohjelmien kautta. Yleisimmin tämä tapahtuu asia-

kaspuolen Javascript-koodin avulla, joka suoritetaan käyttäjän selaimessa [17]. Tämäntyyppiset hyökkkäykset voidaan lukea yleisemmin sovelluskerroksen hyökkäyksiksi ja edelleen selainpohjaisiksi (*browser-based*) hyökkäyksiksi niiden menetelmistä riippuen.

Näiden lisäksi on suuri joukko hyökkäyksiä, jotka paljastavat käyttäjiä niiden tuottaman tietoliikenteen ajoituksia tarkkailemalla. Nimittäisin niitä *ajoitushyökkäyksiksi* (*Timing attack*). Tähän luokkaan kuuluvat hyökkäykset ovat aina olleet suuri huolenaihe matalan viiveen salaverkoissa, koska on paljon suoraviivaisempaa paikantaa käyttäjiä tietoliikenneanalyysillä, kun tiedetään heidän lähettävän viestejä niin nopeasti kuin pystyvät. Paljon vaikeampaa on puolestaan analysoida tietoliikennettä, mihin ei tällaisia oletuksia voida asettaa. Korkean viiveen sähköpostiviestintä on hyvä esimerkki palvelusta, joka voidaan toteuttaa siten että sen analysoiminen on käytännössä mahdotonta. Samoja tekniikoita ei tosin voida soveltaa pikaviestintään tai verkkoselaamiseen, sillä ne nojaavat välittömyyteen [18][19].

Tietoliikenteen analysointia (*traffic analysis*) ja luokittelua (*traffic classification*) on aikaisemmin toteutettu tilastollisilla menetelmillä ja tekoälyn yleistymisen jälkeen myös syväoppimisen avulla. Syväoppimisen menetelmillä on saavutettu hyviä tuloksia pienellä virhemäärällä. Analyysillä tarkoitetaan tietoliikenteen ominaisuuksien eristämistä ja vertailua, voidaanko osapuolista tai sisällöstä saada selkoa ja voidaanko esimerkiksi yhden käyttäjän tuottama tietoliikenne tunnistaa.

Vertaisverkossa anonymiteetille tärkeitä seikkoja, lukuisten muiden ohella, ovat verkon koko eli uniikkien vertaisten lukumäärä, ohjelman käyttämät kryptograafiset algoritmit ja niiden oikeaoppinen toteuttaminen ja itse ohjelmainstanssin piilotettavuus eli se, pystytäänkö ohjelman tietoliikenne naamioimaan. Toisin sanoen jos liikennettä tarkkailee joku, huomaako hän siinä toistuvia osioita tai kaavoja, jotka merkkaavat verkkoliikenteen I2P-verkolle kuuluvaksi. I2P:n tapauksessa verkko on kasvanut tasaisesti, mikä tekee monista hyökkäyksistä epäkäytännöllisiä resurssi-

vaatimusten vuoksi. Sitä vastoin I2P-reititin ei tällä hetkellä piilota olemassaoloaan täydellisesti eli ulkopuolinen, jolla on kyky valvoa verkkoyhteyttä, voi todeta käynnissä olevan ohjelmainstanssin sen tuottaman tietoliikenteen kautta. Tämä korostaa edellistä, verkon koosta tehtyä huomiota ja sen tärkeyttä. Kehitystiimi on kuitenkin ilmaissut jatkavansa protokolliensa kehittämistä, ainakin edellä mainittujen ongelmien saralla [20].

I2P:n uhkamalli

I2P:n uhkamalli ottaa kantaa useisiin hyökkäyksiin, mutta ei ole kokonaisuudessaan I2P:n toteutukselle spesifi vaan peilaa osin Torin ja muiden matalan viiveen salaverkkojen malleja, ottaen kantaa tunnettuihin hyökkäyksiin ja niitä varten jo kehitettyihin turvatoimiin [5][20].

3.2 Hyökkäystyypit

Taulukossa 3.1 esitellään erilaisia hyökkäystapoja, joita on käsitelty salaverkkoihin liittyvässä kirjallisuudessa. Lihavoinnilla korostetut tutkimukset toteutettiin suoraan I2P-verkossa ja niistä kaikki paitsi Hoang et al. esittelevät konkreettisen hyökkäyksen. Timpanaro et al. (ei taulukoitu) ja Hoang et al. suorittivat monitorointia ja tietoliikenteen luokittelua, Hoang et al. kuitenkin esittelivät myös mahdollisen palvelunestohyökkäyksen ja siihen tarvittavat resurssit [14][21]. Muut taulukoidut artikkelit esittelevät hyökkäyksiä sipulireititystä vastaan Tor-verkossa, näitä voidaan teoriassa soveltaa myös I2P-verkkoon niiden yhteneväisyyksien ja esiteltyjen hyökkäysten yleismaallisuuden vuoksi [16]. Keskeisimmät hyökkäykset tämän tutkielman kannalta on esitetty taulukossa 3.1 lihavoituina, poislukien Hoang et al.

Taulukko 3.1: Valikoima I2P:n tietoturvatutkimuksesta (**TK1, TK2**)

	Fingerprinting	Peer selection	Sybil	Peer blocking	Browser-based	Timing
Abbott et al., 2007 [19]					x	x
Wang et al., 2007 [22]						x
Herrmann et al., 2009 [23]	x					
Shi ja Matsuura, 2009 [24]	x					
Herrmann ja Grothoff, 2011 [25]		x				
Crenshaw, 2011 [17]					x	
Panchenko et al., 2011 [26]	x					
Cai et al., 2012 [27]	x					
Egger et al., 2013 [5]			x			
Hoang et al., 2018 [21]				x		

3.2.1 Fingerprinting

Fingerprinting tai *web page fingerprinting* [23], [24], [26], [27] tarkoittaa tässä tutkielmassa kohteen ja verkkopalvelun, yleisimmin verkkosivun, välisen verkkoliikenteen tallentamista ja tilastollisten menetelmien kohdistamista kyseiseen tallenteeseen. Näin hyökkääjä saa, vaihtelevalla menestyksellä, selville sen verkkosivun tai palvelun jolla kohde vierailee. Yleisin lähestymistapa tämäntyylisen hyökkäyksen toteuttamiseen on luoda tietokanta, mihin tallennetaan kohdesivujen nimet ja niiden tuottama verkkoliikenne niillä vierailtaessa. Näitä tallenteita verrataan eri menetelmin niihin mitä kohde tuottaa vieraillessaan eri sivustoilla, yritetään siis selvittää ovatko kohteen ja tietyn palvelimen väliset keskustelut samanlaisia kuin hyökkääjän käymät keskustelut saman palvelimen kanssa. Tämän jälkeen on hyökkääjän

suoraviivaista yhdistää verkkojalanjälki tietokannassa olevaan verkkosivuun.

Tämän luokan hyökkäyksiä vastaan puolustautuminen koostuu verkkoliikenteen täyttämisestä satunnaisella tai harmittomalla datalla, mutta myös muut vaikeasti ennustettavat datamuunnokset käyvät. Selaimen välimuisti on yksi esimerkki, mutta sen arvattavuus tekee siitä erittäin huonon suojan. Yleinen puolustuskeino on lisätä satunnaisesti valittu määrä tavuja jokaiseen datapakettiin oikean datan rinnalle, mutta tämän menetelmän tehokkuus vaihtelee, eikä usein tarjoa kokonaista suojaa. Toinen vähemmän tunnettu menetelmä, jota voidaan soveltaa edellisen kanssa samaan aikaan, on monen eri sivun lataaminen samaan aikaan. Se voitaisiin toteuttaa verkkolaajennuksen kautta, joka aina verkkosivua pyydettäessä tekisi toisen pyynnön taustalla, yhdelle tai useammalle satunnaiselle sivulle. Tällainen naamioimiskeino on mieleinen, koska sen voi toteuttaa kokonaan käyttäjäpuolella. [26]

3.2.2 Peer selection

Peer selection [25], tai suomeksi vertaisten valinta, käsittää hyökkäykset jotka hyväksikäyttävät I2P-ohjelmainstanssin reitin, eli tunneleiden, luontiprosessia. Käsitteenä tämä kattaa kaikki vertaisverkot missä osalliset tai vertaiset suorittavat jatkuvaa profilointia laadukkaiden yhteyksien ylläpitämistä ja luomista varten. I2P:n tapauksessa on yksi käytännöllinen esimerkki, jonka päämääränä on salaverkossa toimivan palvelimen tai sivuston sijainnin paljastaminen. Hyökkääjä tarvitsee vertaisparven, joka on hajautettu moneen eri /16 aliverkkoon. Kohteen paljastamiseksi tulee hyökkääjän päästä osaksi kohteen tunneleita sekä lähetys- että vastaanottopuolelle. Usean aliverkon vaatimus johtuu I2P:n toteutuksesta; se ei salli usean tunnelivertaisen valitsemista samasta aliverkosta eli kohteen paikantaminen ei käytännössä ole mahdollista vain yhdestä aliverkosta käsin.

Vertaisjoukon luonnin jälkeen hyökkääjän tulee kysellä netDB:ltä reittiohjeita kohdesivulle, tämä paljastaa ne vertaiset jotka, tarjoamansa kaistanleveyden ansios-

ta, sillä hetkellä reitittävät kohdesivun liikennettä ja osallistuvat sen tunneleihin. Kun nämä tunnelivertaiset on tiedossa, voidaan niitä vastaan aloittaa palvelunestohyökkäys, tämän ansiosta kohde ei enää valitse kyseisiä vertaisia tunneleihinsa suorituskykyprofiloinnin perusteella. Hyökkääjä toistaa tätä prosessia kunnes hyökkääjän omia vilpillisiä monitorivertaisia aletaan kohteen toimesta valita sen tunneleihin. Kun hyökkääjällä on näitä monitorivertaisiaan kohteen lähetys- ja vastaanottotunneleiden osina, hyökkääjän pää- tai hallintainstanssi lähettää ajoitettuja pyyntöjä kohdesivulle. Edellä mainitut monitorivertaiset kuuntelevat hallintainstanssin lähettämiä pyyntöjä, mikäli sellainen havaitaan kahden monitorivertaisen välistä niin voidaan olettaa kohdesivun sijaitsevan tuossa väliin jäävässä I2P-ohjelmainstanssissa.

Tämän hyökkäyksen estäminen ei ole suoraviivaista, sillä hyökkääjän ajoitettujen pyyntöjä voi naamioida monella eri tavalla, niiden tunnistamisen vaikeuttamiseksi. Toinen tulokulma koskee suorituskykynsä vuoksi valittuja vertaisia, joita vastaan hyökkääjä suorittaa palvelunestohyökkäyksiä. I2P-ohjelmainstanssin tulisi välttää antamasta tietoa näistä vertaisista, valitsemalla myös satunnaisesti vertaisia osaksi reititystään. Tällaisessa tilanteessa hyökkääjä ei tiedä tarkkaan niitä vertaisia, keihin kohdistaa palvelunestohyökkäys. [25]

3.2.3 Sybil

Sybil [5] käsittää verkkohyökkäykset, joiden päätavoite on suuren vertaisparven avulla toteutettu auktoriteetin tai päätäntävallan keskittäminen hyökkääjälle. Käytännössä hyökkääjä haluaa tehdä itsestään mahdollisimman suuren osan verkkoa ja hallita sen tietoliikennettä. I2P:n tapauksessa tämäntyylisiä hyökkäyksiä voidaan kohdistaa NetDB:n hajautukseen, joko kohdistetusti vain uhrin läheisyyteen tai netDB:n koko avainavaruuteen. Kohdistettua versiota nimitetään *Eclipse*-hyökkäykseksi ja vaatii 8-15 ohjelmainstanssia, tämänhetkistä NetDB-toteutusta vastaan. Koko avainavaruuden kattava versio, *floodfill takeover*, vaatii noin 2% verkon vertaisten

lukumäärästä. Hyökkääjä luo mahdollisimman monta ohjelmainstanssia ja asettaa ne manuaalisesti *floodfill*-tilaan, jolloin hyökkääjän rooli netDB:n hajauttamisessa ja jakelussa kasvaa. Kun nämä hyökkääjän hallitsemat vertaiset ovat ainoita, jotka jakelevat tiettyä netDB:n osaa, on hyökkääjä täysin vastuussa sen osan levikistä. Näin hyökkääjä voi suorittaa palvelunestohyökkäyksen ja eristää uhrin muusta verkosta.

Vakavuutensa vuoksi tämäntyylisiin hyökkäyksiin on jo puututtu kehittäjien toimesta, *floodfill*-reitittimien lukumäärää on nostettu, jolloin jokainen yksittäinen *floodfill*-reititin palvelee pienempää osaa avainavaruudesta, vaikeuttaen sen kokonaista haltuunottoa. Toinen toteutettu suojakeino on netDB-kyselyiden muuttaminen siten, että vain yksi *floodfill*-reititin voi vastata jokaisesta /16 aliverkosta[1] vertaisen kyselyyn. Tämä tarkoittaa sitä, että hyökkääjän täytyy hajauttaa operoimansa vertaiset useaan eri verkkoon pilvipalveluita hyödyntämällä, mikä lisää hyökkäyksen vaatimia resursseja sekä työmäärää. NetDB:n DHT-toteutuksessa käytettävän Kademlia-algoritmin vaihtamisesta R^{5}N-algoritmiin[2] on myös keskusteltu, mutta tämä vaihto ei ole toistaiseksi toteutunut. [5]

3.2.4 Peer blocking

Peer blocking [21] tarkoittaa tässä tutkielmassa palvelunestohyökkäystä, jonka avulla hyökkääjä voi estää tuntemansa kohteen pääsyn I2P-verkkoon. Hyökkääjä tarvitsee 6-20 ohjelmainstanssia, joiden avulla kerätään listaa vertaisista, joita uhri tulee todennäköisimmin käyttämään päästäkseen verkkoon. 20 ohjelmainstanssia hyödyntämällä tutkijat kykenivät tunnistamaan 95% uhrin käyttämistä vertaisista. Kun vertaiset ovat tiedossa, voidaan niiden IP-osoitteet estää perinteisin menetelmin hyökkääjän toimesta, mikä vaatii pääsyn uhrin verkon asetuksiin. Hyökkääjänä voisi siis toimia uhrin palveluntarjoaja tai erilaiset valtiotahot. Mikäli yli 90% näistä vertai-

[1] https://github.com/i2p/i2p.i2p/commit/5eba38a24ef3b2cc356809057d418905dc2f974c

[2] https://lsd.gnunet.org/lsd0004/

sista kyetään estämään, on I2P-verkon käyttäminen käytännössä mahdotonta, sillä 95-100% verkkopyynnöistä keskeytyvät eli ne aikakatkaistaan.

Palveluntarjoajan tai valtiotahon aiheuttama palvelunesto on lähtökohtaisesti vaikea ohittaa, mutta joitakin ratkaisuja löytyy. I2P:n tuottaman verkkoliikenteen naamioiminen on yksi tapa, joka on myös kehittäjien tiedossa, mutta sen toteuttaminen ei ole suoraviivaista. Toinen vaihtoehto perustuu siihen seikkaan, ettei hyökkääjän keräämä tieto vertaisista päivity yhtä nopeasti kuin uhrin käytettävissä olevat vertaiset. Tästä seuraa se, että juuri I2P-verkkoon liittyneet vertaiset eivät ole välittömästi hyökkääjän estolistalla ja kykenevät siksi kommunikoimaan uhrin kanssa esteettä, kunnes nekin havaitaan ja estetään. Näitä juuri liittyneitä vertaisia voisi siis hyödyntää uhrin verkkoliikenteen reitittämiseen. Ongelmana tässä on jatkuva uusien, estolistalta puuttuvien, vertaisten tarve. Uusien vertaisten lisäksi voitaisiinkin hyödyntää palomuurin takana olevia vertaisia, joilla ei ole julkista IP-osoitetta, eivätkä sen takia päädy ennustettavasti estolistalle. [21]

3.2.5 Browser-based

Browser-based [17], [19] eli selainpohjaiset hyökkäykset keskittyvät selaimen antamien tietojen analysointiin, hyökkäyksiä voisi myös luonnehtia HTTP-pohjaisiksi. Tämän luokan hyökkäyksiä voidaan soveltaa sekä käyttäjiin että palveluihin. Käyttäjän selaimen ominaisuuksista voidaan luoda yksilöivä sormenjälki, ja vastaavasti HTTP-palvelimen antamia vastauksia voidaan yksilöidä ja vertailla, esimerkiksi otsaketietoja (*headers*) tarkastelemalla. Etenkin jos palvelin toimii I2P-verkossa ja tavallisessa Internetissä samaan aikaan, tarjoten samoja palveluita. Näiden menetelmien toimivuus ei kuitenkaan ole itsestäänselvää, tehokkuus nojaa useimmiten selaimen käyttäjän tai palvelun ylläpitäjän tekemiin virheisiin. Palveluiden analysoinnissa tutkitaan myös verkkoliikennettä, jonka pienet vaihtelut voivat paljastaa kyseisen TCP/IP-pinon toteutuksen ja sitä kautta käyttöjärjestelmän.

Suojautuminen yksinkertaisimpia HTTP-hyökkäyksiä vastaan on suoraviivaista palvelinpuolella. Seuraavat keinot ovat vain osa kokonaisuutta mutta tehoavat heikoimpia hyökkäyksiä vastaan: ohjeista HTTP-palvelin lähettämään mahdollisimman vähän otsaketietoja ja määritä I2P:hen suunnatuille verkkosivuille kuunteluosoitteeksi pelkästään I2P-ohjelmainstanssin osoite[3]. Ne otsaketiedot, joita ei ohjelman asetuksien kautta kyetä poistamaan, voidaan HTTP-palvelimella pyörivän verkkosivun ohjelmakoodissa muuttaa. Apachen `apache2.conf`-tiedostosta voidaan seuraavat asetukset muuttaa:

```
ServerTokens Prod
ServerSignature Off
```

Ja vastaavasti PHP:ssa:

```
header_remove('X-Powered-By');
```

PHP-esimerkki pitää sisällyttää osaksi kaikkia verkkosivun tarjoilemia sivuja, esimerkiksi osana projektin `head.php` tai `header.php` -tiedostoa, muuten turva jää vajavaiseksi.

Käyttäjäpuolella hyökkäyksiä voidaan estää selaimen asetuksia muuttamalla, haluttu lopputulos on mahdollisimman yleinen asetusten kokoonpano, eli sormenjälki. Tässä muutamia esimerkkejä: Tor-selaimen käyttäminen I2P:n selaamiseen, selainikkunan koon jättäminen oletusasetukseen, selaimessa suoritettavan Javascriptin kytkeminen pois, evästeiden estäminen tai I2P-verkkoa varten luotujen selainlaajennusten asentaminen[4].

[3]https://httpd.apache.org/docs/2.4/vhosts/

[4]https://github.com/eyedeekay/I2P-in-Private-Browsing-Mode-Firefox

3.2.6 Timing

Timing [19], [22] eli ajoitusten analyysi tarkoittaa salatun verkkoliikenteen datapakettien välisten aikaleimojen tutkimista. Hyökkääjä luo siis halutulle verkkovirralle oman tunnistettavan leiman tai jalanjäljen muuttamalla siihen kuuluvien datapakettien kulun nopeutta hieman. Esimerkkinä yhteys verkkosivun ja käyttäjän välillä, missä verkkosivu itse upottaa tunnistettavan leiman käyttäjälle menevään paluuliikenteeseen eli vastaukseen. Hyökkääjä voi varmistaa, jo epäillyn, käyttäjän ohjelmainstanssin tämän jälkeen monitoroimalla tälle kohdistettua verkkoliikennettä. Mikäli siinä havaitaan sama leima, voidaan päätellä että käyttäjä vieraili hyökkääjän verkkosivulla.

I2P-verkossa viestien käsittelyyn ja lähettämiseen kuluva aika vaihtelee suuresti, tämän takia ajoitushyökkäykset eivät ole niin vakava uhka kuin suoraviivaisemmissa verkoissa. Uhka se kuitenkin on ja sitä vastaan voidaan käyttää muutamia tekniikoita. Sovellusten tuottamaan verkkoliikenteeseen voitaisiin tehdä muutoksia ennen lähettämistä, esimerkiksi verkkokyselyn tuottamat datapaketit voidaan järjestää uudelleen, niiden viivettä voidaan kasvattaa eli niitä ei lähetettäisi heti kun kyetään, satunnaisgeneroituja datapaketteja voidaan lisätä luonnollisten väliin ja toisiinsa liittymättömistä datapaketeista voisi muodostaa ryppäitä[5]. Edellä mainittuja suojakeinoja ei ole toistaiseksi toteutettu. [20]

[5]https://geti2p.net/en/get-involved/todo#batching

4 Pohdinta

Aiemmin käsitellyt hyökkäykset jakautuvat verkko- sekä sovellustason hyökkäyksiin ja niiden välillä on monia eroja: osa vaatii suuren vertaisparven, osa pienen mutta tarkasti sijoitetun vertaisparven ja osa nojaa käyttäjän tai ylläpitäjän virheisiin. Vertaisparven ylläpitämiseen ja tarkkaan hallitsemiseen nojaavat hyökkäykset, kuten ”Practical Attacks Against The I2P Network” julkaisussa esitetyt, eivät ole helppoja toteuttaa vaikka ovatkin käytännöllisiä. Niiden kustannukset kasvavat jatkuvasti, koodimuutosten ja verkon koon kasvun myötä, sen lisäksi ne voidaan havaita vilpittömien vertaisten toimesta. Näiden lisäksi alati vaihtuva verkon rakenne vaikeuttaa tämäntyylisiä hyökkäyksiä, hyökkääjän täytyy tuntea I2P-verkon toteutuksesta muutamia yksityiskohtia ja ottaa ne huomioon hyökkäyksen toteutuksessa. I2P-verkossa enemmistö verkkosivuista on vilpittömiä, joten ilmiselvät verkkotason hyökkäykset herättävät myös nopeasti julkisuutta ja tiukempia turvatoimia; on helpompaa saada julkista kannatusta, esimerkiksi kryptologian asiantuntijalta tai libertaristilta, sananvapaudelle kuin anonyymeille huumeiden kauppapaikoille. Näihin seikkoihin ja taustamateriaaliin nojaten, on mielestäni hyökkääjälle paljon edullisempaa keskittyä laajojen hyökkäysten sijaan keskitettyihin ja hyvin kohdistettuihin hyökkäyksiin.

Sovellustason hyökkäykset palvelimia vastaan, eritoten selainpohjaiset sellaiset, pysyvät hyökkääjälle mieleisinä niiden helppokäyttöisyyden ja monipuolisuuden ansiosta. HTTP-sovellukset ovat yleisiä ja huonosti toteutetut sellaiset siis myös varsin yleisiä, jolloin kokemattomatkin hyökkääjät voivat hyödyntää automatisoituja työ-

kaluja, hyväksikäyttääkseen sovelluksen ylläpitäjän tai toteuttajan tekemiä virheitä. Käyttäjiä vastaan kohdistetut selainpohjaiset hyökkäykset ovat nykyään vaikeampia toteuttaa selainten kehittyneiden ja monimuotoisten turvamekanismien vuoksi, kuten aiemmin todettiin aineiston pohjalta. Syynä tähän on se, että pieni ja kokenut ohjelmistokehittäjien ryhmä yleensä toteuttaa yhtenäiset käyttäjätyökalut, joilla suuri käyttäjäkunta pääsee yhteyteen tuhansien kehittäjien sovelluksiin, monet näistä sovellusten kehittäjistä tai ylläpitäjistä ovat puolestaan vain harrastelijoita.

Näitä jo mainittuja hyökkäystyyppejä hyödyllisempänä pitäisin tietoliikenteen analyysiä sekä verkkosivujen ja tietovirtojen sormenjäljentämistä. Näiden tekniikoiden hyödyntäminen ei vaadi jatkuvaa valvomista ihmisen toimesta eivätkä ne jätä jälkiä, koska ne ovat passiivisia hyökkäyksiä. Kohteen ja menetelmien määrittelemisen jälkeen hyökkäys voidaan jättää keräämään tietoa. Ne ovat siis täysin huomaamattomia, eivätkä ne ole aikakriittisiä. Tämäntyylisiä hyökkäyksiä vastaan puolustautuminen on paljon vaikeampaa kuin niiden toteuttaminen, ja koneoppimismenetelmien kehittyessä uskoisin niiden roolin kasvavan entisestään. Aiemmin käsiteltiin sitä, miten monen sivun samanaikainen lataaminen voisi välttää tämäntyylisen tiedonkeruun, mutta aineiston perusteella menetelmät kehittyvät niin nopeasti, että koneoppiminen tulee murtamaan juuri tämänlaiset puolustuskeinot nopeammin kuin niitä vastaan ehditään kehittämään ja toteuttamaan suojakeinoja. Tor-verkon tutkimuksessa on jo esitelty tämän kuvauksen toteuttavia tekniikoita sekä hyökkäys- että puolustusnäkökulmasta. Aiheesta löytyy lisätietoa hakusanoilla ”multi-tab fingerprinting” [28]–[35].

5 Yhteenveto

Tämän tutkielman tavoitteena oli luoda eri lähteitä hyödyntävä kokonaiskuva hyökkäyksistä, joita on sovellettu sipulireititykseen perustuvia salaverkkoja vastaan. Keskipisteenä oli I2P-salaverkko, joka rakentuu valkosipulireitityksen ympärille. Hyökkäykset verkkoa ja sen käyttäjiä vastaan luokiteltiin, niiden toiminta ja vakavuus käytiin läpi, lisäksi mahdollisia suojautumismenetelmiä käsiteltiin jokaisen hyökkäystyypin kohdalla. Tutkielman motivaationa oli vähäinen I2P:tä koskevan tutkimuksen määrä, etenkin nykyaikaisten, 2020-luvulla laadittujen, kirjallisuuskatsausten puute.

Tämän tutkielman perusteella voidaan todeta, että I2P-verkkoa vastaan on suoritettu ja voidaan jatkossakin suorittaa onnistuneita hyökkäyksiä. Nämä hyökkäykset voivat pahimmassa tapauksessa paljastaa kohteena olevan palvelun ylläpitäjän tai tavallisen käyttäjän IP-osoitteen, mikä usein johtaa myös heidän henkilöllisyytensä paljastumiseen. Kryptografisia hyökkäyksiä ei ole verkkoa vastaan onnistuneesti suoritettu eli tietoliikenteen tarkka sisältö pysyy salassa. Hyökkääjä voi kuitenkin tietoliikennettä analysoimalla saada, pahimmassa tapauksessa, lähes varmaa tietoa siitä, mitä tutkittava tietoliikenne koskee. Verkkosivut ja I2P-verkkoa hyödyntävät sovellukset voidaan tunnistaa niiden tuottaman tietoliikenteen kautta, jolloin tietyssä mielessä tiedetään mitä kohde I2P-verkossa tekee vaikka viestien sisältö pysyykin salattuna.

Tähän uhkakuvaan liittyy joitakin tarkennuksia, joiden avulla voidaan perustella I2P-verkon tietoturvan ja anonymiteetin olevan pääosin eheä. Huomattava osa

matalan kynnyksen hyökkäyksistä nojaa käyttäjän tai ylläpitäjän tekemiin virheisiin, eikä itse I2P-verkon suunnitteluun tai toteutukseen. Käyttäjäpuolen työkaluja ja käyttöohjeita parantamalla voidaan nämä niin kutsutut heikot hyökkäykset mitätöidä, lopullinen vastuu jää kuitenkin käyttäjälle. Sen sijaan itse verkkoa vastaan toteutetut tuhoisemmat hyökkäykset hyväksikäyttävät I2P-verkon rakennetta, eikä niitä vastaan ole niin suoraviivaista suojautua. Kehittäjät ovat toteuttaneet suojakeinoja tämäntyylisiä hyökkäyksiä vastaan itse ohjelmakoodiin ja niiden suorittaminen on näiden muutosten myötä hankaloitunut. Jatkuva verkon koon kasvu on myös tärkeä suojakeino näitä tuhoisia hyökkäyksiä vastaan. 3% vertaisista ei ole ylitsepääsemätön määrä jos verkon vertaisten lukumäärä on esimerkiksi alle tuhat, mutta verkon nykyinen koko on jo kasvanut kymmeniintuhansiin vertaisiin, jolloin hyökkäyksen toteuttaminen vaikeutuu monikymmenkertaisesti.

Jatkotutkimusaiheita on monia tämän katsauksen perusteella. Syvällisempi Tor-verkkoon kohdistuneen tutkimuksen läpikäynti ja sen soveltuvuuden tutkiminen ja testaaminen I2P-verkkoa vastaan, erityisesti Tor-verkon palveluita vastaan toteutettava ”Web Fingerprinting” eli *WF* on laaja ja aktiivinen tutkimusalue. Tässä tutkielmassa pyrittiin tuomaan näiden kahden salaverkon tutkimusta yhteen, mutta laajuutensa vuoksi ei koko syvyyttä kyetty sisällyttämään tähän yhteyteen. Empiirinen tutkimus tässä tutkielmassa läpikäytyjen hyökkäysten pohjalta antaisi konkreettisen pohjan I2P-verkon uhkamallille, mikä vuorostaan antaisi kehittäjille tietoa siitä, mitkä ovat suurimmat uhkat I2P-verkolle reaalimaailmassa. Aiheen rajaaminen pelkästään tietoliikenteen analyysiin ja sitä vastaan suojautumiseen. Esimerkiksi I2P:n verkkoprotokollien tuottaman tietoliikenteen naamioiminen, verkkosivujen sekä muiden verkkosovellusten tuottaman sormenjäljen hävittäminen tai satunnaistaminen yhteyskohtaiseksi. Tutkimuksessa voitaisiin ottaa selvää siitä, mitkä nykyaikaisista menetelmistä olisivat sopivimmat juuri I2P:n tapauksessa.

Lähdeluettelo

[1] ”The Invisible Internet Project”. (lokakuu 2023), url: https://geti2p.net/.

[2] ”Tor Project History”. (lokakuu 2023), url: https://www.torproject.org/about/history/.

[3] ”Freedomhouse: Freedom on the Net”. (2023), url: https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence.

[4] ”The Truth About The Dark Web”. (syyskuu 2019), url: https://www.imf.org/en/Publications/fandd/issues/2019/09/the-truth-about-the-dark-web-kumar.

[5] C. Egger, J. Schlumberger, C. Kruegel ja G. Vigna, ”Practical Attacks against the I2P Network”, teoksessa *Research in Attacks, Intrusions, and Defenses*, S. J. Stolfo, A. Stavrou ja C. V. Wright, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, s. 432–451, ISBN: 978-3-642-41284-4.

[6] R. R. Dingledine, ”The free haven project: Design and deployment of an anonymous secure data haven”, osio 8.1.1 'Garlic Routing', tutkielma, Massachusetts Institute of Technology, 2000.

[7] D. M. Goldschlag, M. G. Reed ja P. F. Syverson, ”Hiding Routing information”, teoksessa *Information Hiding*, R. Anderson, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 1996, s. 137–150, ISBN: 978-3-540-49589-5.

[8] D. Goldschlag, M. Reed ja P. Syverson, "Onion Routing", *Commun. ACM*, vol. 42, nro 2, s. 39–41, helmikuu 1999, ISSN: 0001-0782. DOI: 10.1145/293411.293443. url: https://doi.org/10.1145/293411.293443.

[9] J. Saleem, R. Islam ja M. A. Kabir, "The Anonymity of the Dark Web: A Survey", *IEEE Access*, vol. 10, s. 33 628–33 660, 2022. DOI: 10.1109/ACCESS.2022.3161547.

[10] T. de Boer ja V. Breider, "Invisible Internet Project(Report)", tutkielma, University of Amsterdam, helmikuu 2019.

[11] P. Liu, L. Wang, Q. Tan, Q. Li, X. Wang ja J. Shi, "Empirical measurement and analysis of I2P routers", *Journal of Networks*, vol. 9, nro 9, s. 2269, 2014.

[12] zzz (Pseudonym) ja L. Schimmer, "Peer Profiling and Selection in the I2P Anonymous Network", teoksessa *Proceedings of PET-CON 2009.1*, Dresden, Germany, maaliskuu 2009, s. 59–70.

[13] P. Maymounkov ja D. Mazières, "Kademlia: A Peer-to-Peer Information System Based on the XOR Metric", teoksessa *Peer-to-Peer Systems*, P. Druschel, F. Kaashoek ja A. Rowstron, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, s. 53–65, ISBN: 978-3-540-45748-0.

[14] J. P. Timpanaro, T. Cholez, I. Chrisment ja O. Festor, "Evaluation of the anonymous I2P network's design choices against performance and security", teoksessa *2015 International Conference on Information Systems Security and Privacy (ICISSP)*, 2015, s. 1–10.

[15] "The Network Database". (marraskuu 2023), url: https://geti2p.net/en/docs/how/network-database.

[16] E. Erdin, C. Zachor ja M. H. Gunes, "How to Find Hidden Users: A Survey of Attacks on Anonymity Networks", *IEEE Communications Surveys & Tutorials*, vol. 17, nro 4, s. 2296–2316, 2015. DOI: 10.1109/COMST.2015.2453434.

[17] A. Crenshaw, "Darknets and hidden servers: Identifying the true IP/network identity of I2P service hosts", teoksessa *Proceedings of Black Hat 2011*, Washington, DC, tammikuu 2011. url: http://www.irongeek.com/i.php?page=security/darknets-i2p-identifying-hidden-servers.

[18] A. Back, U. Möller ja A. Stiglic, "Traffic Analysis Attacks and Trade-Offs in Anonymity Providing Systems", teoksessa *Information Hiding*, I. S. Moskowitz, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2001, s. 245–257, ISBN: 978-3-540-45496-0.

[19] T. G. Abbott, K. J. Lai, M. R. Lieberman ja E. C. Price, "Browser-Based Attacks on Tor", teoksessa *Privacy Enhancing Technologies*, N. Borisov ja P. Golle, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, s. 184–199, ISBN: 978-3-540-75551-7.

[20] "I2P Threat Model". (lokakuu 2023), url: https://geti2p.net/en/docs/how/threat-model.

[21] N. P. Hoang, P. Kintis, M. Antonakakis ja M. Polychronakis, "An Empirical Study of the I2P Anonymity Network and its Censorship Resistance", *CoRR*, vol. abs/1809.09086, 2018. arXiv: 1809.09086. url: http://arxiv.org/abs/1809.09086.

[22] X. Wang, S. Chen ja S. Jajodia, "Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems", teoksessa *2007 IEEE Symposium on Security and Privacy (SP '07)*, 2007, s. 116–130. DOI: 10.1109/SP.2007.30.

[23] D. Herrmann, R. Wendolsky ja H. Federrath, "Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naıve-Bayes Classifier", teoksessa *Proceedings of the 2009 ACM Workshop on Cloud Computing Security*, sarja CCSW '09, Chicago, Illinois, USA: Association for

Computing Machinery, 2009, s. 31–42, ISBN: 9781605587844. DOI: 10.1145/1655008.1655013. url: https://doi.org/10.1145/1655008.1655013.

[24] Y. Shi ja K. Matsuura, "Fingerprinting Attack on the Tor Anonymity System", teoksessa *Information and Communications Security*, S. Qing, C. J. Mitchell ja G. Wang, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, s. 425–438, ISBN: 978-3-642-11145-7.

[25] M. Herrmann ja C. Grothoff, "Privacy-Implications of Performance-Based Peer Selection by Onion-Routers: A Real-World Case Study Using I2P", teoksessa *Privacy Enhancing Technologies*, S. Fischer-Hübner ja N. Hopper, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, s. 155–174, ISBN: 978-3-642-22263-4.

[26] A. Panchenko, L. Niessen, A. Zinnen ja T. Engel, "Website Fingerprinting in Onion Routing Based Anonymization Networks", teoksessa *Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society*, sarja WPES '11, Chicago, Illinois, USA: Association for Computing Machinery, 2011, s. 103–114, ISBN: 9781450310024. DOI: 10.1145/2046556.2046570. url: https://doi.org/10.1145/2046556.2046570.

[27] X. Cai, X. C. Zhang, B. Joshi ja R. Johnson, "Touching from a Distance: Website Fingerprinting Attacks and Defenses", teoksessa *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, sarja CCS '12, Raleigh, North Carolina, USA: Association for Computing Machinery, 2012, s. 605–616, ISBN: 9781450316514. DOI: 10.1145/2382196.2382260. url: https://doi.org/10.1145/2382196.2382260.

[28] Q. Yin, Z. Liu, Q. Li et al., "An Automated Multi-Tab Website Fingerprinting Attack", *IEEE Transactions on Dependable and Secure Computing*, vol. 19, nro 6, s. 3656–3670, 2022. DOI: 10.1109/TDSC.2021.3104869.

[29] X. Deng, Q. Yin, Z. Liu et al., "Robust Multi-tab Website Fingerprinting Attacks in the Wild", teoksessa *2023 IEEE Symposium on Security and Privacy (SP)*, 2023, s. 1005–1022. DOI: 10.1109/SP46215.2023.10179464.

[30] H. Ali, M. Iqbal, S. u. R. Khan et al., "Poker Face Defense: Countering Passive Circuit Fingerprinting Adversaries in Tor Hidden Services", teoksessa *2023 International Conference on IT and Industrial Technologies (ICIT)*, 2023, s. 1–7. DOI: 10.1109/ICIT59216.2023.10335871.

[31] P. Liu, L. He ja Z. Li, "A Survey on Deep Learning for Website Fingerprinting Attacks and Defenses", *IEEE Access*, vol. 11, s. 26 033–26 047, 2023. DOI: 10.1109/ACCESS.2023.3253559.

[32] M. Shen, K. Ye, X. Liu et al., "Machine Learning-Powered Encrypted Network Traffic Analysis: A Comprehensive Survey", *IEEE Communications Surveys & Tutorials*, vol. 25, nro 1, s. 791–824, 2023. DOI: 10.1109/COMST.2022.3208196.

[33] B. Sun, W. Yang, M. Yan, Y. Zhu ja Z. Bai, "A Practical Website Fingerprinting Defense Approach with Universal Adversarial Perturbations", teoksessa *2022 7th International Conference on Computer and Communication Systems (ICCCS)*, 2022, s. 752–760. DOI: 10.1109/ICCCS55155.2022.9846237.

[34] M. Jiang, B. Cui, J. Fu, T. Wang ja Z. Wang, "KimeraPAD: A Novel Low-Overhead Real-Time Defense Against Website Fingerprinting Attacks Based On Deep Reinforcement Learning", *IEEE Transactions on Network and Service Management*, s. 1–1, 2024. DOI: 10.1109/TNSM.2024.3360082.

[35] Q. Yin, Z. Liu, Q. Li et al., "An Automated Multi-Tab Website Fingerprinting Attack", *IEEE Transactions on Dependable and Secure Computing*, vol. 19, nro 6, s. 3656–3670, 2022. DOI: 10.1109/TDSC.2021.3104869.