Mikko Pirilä

KÄYTTÄJÄN MANIPULAATION EHKÄISY

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2019

TIIVISTELMÄ

Pirilä, Mikko
Käyttäjän manipulaation ehkäisy
Jyväskylä: Jyväskylän yliopisto, 2019, 29 s.
Tietojärjestelmätiede, kandidaatintutkielma
Ohjaaja: Koskelainen, Tiina

Tässä tutkielmassa käydään läpi käyttäjän manipulaatiota, (engl. *social engineering*) tarkemmin sanottuna miten se ilmenee käytännössä ja miksi se on niin tehokas tapa tehdä tietoturvahyökkäys, sekä miten siihen voidaan varautua. Käyttäjän manipulaatio -hyökkäyksessä pyritään pääsemään käsiksi esimerkiksi organisaation tietokantaan tai yksityisen ihmisen tietoihin. Hyökkäys kohdistuu tietokoneiden ja muiden laitteiden sijasta käyttäjään, jota hyökkääjä pyrkii manipuloimaan saadakseen haluamansa. Nämä manipulaatiohyökkäykset perustuvat psykologiaan ja ihmisten perusominaisuuksiin, tehden niistä tehokkaita ja vaikeita ja joissain tilanteissa mahdottomia havaita ennen kuin on myöhäistä. Tutkielmassa kuvaillaan muun muassa käyttäjän manipulaatioon liittyviä erilaisia psykologisia tekijöitä, hyökkäysmetodeja ja hyökkäysten peruspiirteitä. Tämä tutkielma pyrkii avaamaan syitä sille, miksi käyttäjän manipulaatio on suuri ja verrattain tuntematon uhka. Käyttäjän manipulointi on suuri ongelma, koska uhri voi olla yksityinen henkilö tai yritys ja menetykset kummassakin tapauksessa voivat olla huomattavia. Esimerkiksi internet on pullollaan käyttäjän manipulointiin pyrkiviä pop up -mainoksia ja huijaussähköposteja. Yksi suurimmista ongelmista liittyen käyttäjän manipulaation on kyseisten hyökkäysten/huijausten tunnistaminen. Siksi onkin tärkeää tutkia aihetta ja ottaa selville, miten käyttäjän manipuloinnista syntyviä menetyksiä ja haittoja saadaan vähennettyä, sekä millaiset hyökkäykset ovat kaikista menestyksekkäimpiä. Tämän kirjallisuuskatsauksen tuloksena on katsaus käyttäjän manipulaatiosta, sekä siitä miksi se on tehokas hyökkäystapa ja miten mahdollisesti sitä voitaisiin ehkäistä. Tutkielman tuloksista voidaan päätellä, että käyttäjän manipulaatio ei saa sen vaatimaa huomiota tietoturvan suhteen, sekä, että sen aiheuttamat vahingot vuosittain ovat massiiviset. Ainoa tapa päästä näistä hyökkäyksistä eroon on loppujen lopuksi valistaa ihmisiä niiden vaaroista, mutta hyökkäysten monimuotoisuuden vuoksi se on käytännössä mahdotonta.

Asiasanat: käyttäjän manipulaatio, tietoturva, psykologinen vaikuttaminen, kalastelu, verkkohuijaukset

ABSTRACT

Pirilä, Mikko
The prevention of social engineering
Jyväskylä: University of Jyväskylä, 2019, 29 p.
Information Systems, Bachelor's Thesis
Supervisor: Koskelainen, Tiina

In this report, we will go through social engineering and how it manifests itself in practice and why it is such an effective way to attack as well as how can one prepare against such attack. In social engineering attacks the attacker aims to acquire access for example to an organisations database or private citizens personal information. Instead of focusing on computers or other electronic devices, the attack targets the user, which the attacker attempts to manipulate to acquire what (s)he wants. These social engineering attacks are based on human psychology and basic human characteristics, making them effective and difficult and in some cases impossible to detect before it is too late. This report describes the different psychological factors related to social engineering as well as different attack methods and different features that the attacks hold. This report aims to clarify, why social engineering is such an immediate and rather unknown threat. For example, the internet is filled with pop up ads and spam that aim to manipulate the user. One of the biggest problems regarding social engineering is the detection of the attacks and scams. Thus, it is important to research and find out how the losses accumulating from social engineering can be reduced and what kinds of attacks are the most successful. The result of this report is an examination of social engineering and why it is such an effective avenue of attack as well as how it could be prevented. From the results we can conclude, that social engineering does not get the attention it needs regarding information security and that the losses it causes are massive. In the end the only real way to get rid of these attacks is to educate people about the dangers of social engineering, but due to the complexity and multiple different methods, it is impossible in practice.

Keywords: social engineering, IT-security, psychological influencing, phishing, internet scams

KUVIOT

TAULUKOT

SISÄLLYS

1 JOHDANTO

Perinteiset tietoturvauhat, kuten virukset ja hakkerit on otettu huomioon virustentorjuntaohjelmia luodessa, mutta käyttäjän manipulaation kautta tapahtuvat hyökkäykset ovat jääneet vaille suurempaa huomiota, eikä niitä vastaan ole mitään konkreettista puolustusta. Toki useat sähköpostiohjelmat automaattisesti siirtävät epäilyttävät viestit roskapostiin ja käyttöjärjestelmät, kuten Windows kysyvät erikseen halutaanko tuntematon tiedosto suorittaa. Näistä huolimatta käyttäjän manipulaation aiheuttamat menetykset ovat vuosittain huomattavat. Tässä kirjallisuuskatsauksessa käydään läpi yleisesti käyttäjän manipulaatiota ja sen piirteitä, sekä siihen liittyvää psykologiaa ja ehkäisymenetelmiä. Myöskin yleisimpiin hyökkäysmetodeihin luodaan pikainen katsaus esimerkkeineen.

Tutkimusta on rajattu siten, että se ei koske täysin fyysisessä maailmassa tapahtuvia toimia. Pääosin tutkielmassa käsitellään käyttäjän manipulaatiota yritysten ja organisaatioiden näkökulmasta, mutta viitataan tarpeen vaatiessa yksityisiä ihmisiä koskeviin seikkoihin. Näkökulmana tässä kirjallisuustutkielmassa on tunnistaminen ja ehkäisy.

Tutkimukselle aiheesta on tarvetta, sillä ottaen huomioon käyttäjän manipulaation yleisyyden ja ongelmallisuuden on se verrattain vähän esillä ollut asia, sekä ihmisten yleistieto kyseisestä aiheesta on lähdekirjallisuuden mukaan hyvin vähäistä. Käyttäjän manipulaatiosta aiheutuvat menetykset ovat myös yksi syy tämänkaltaiselle tutkimukselle, sillä 60 % yrityksistä joutui käyttäjän manipulaation kohteeksi vuonna 2016 tietoturvayritys Agarin 2016 tekemän tutkimuksen mukaan (The Cyber-Security source, 2016).

Tutkielman keskeisiä käsitteitä ovat: Käyttäjän manipulaatio, psykologinen vaikuttaminen, hyökkäysmetodi. Käyttäjän manipulaatiolla viitataan tässä tutkielmassa rikollisiin toimiin käyttäjän manipuloimiseksi siten, että hyökkääjä pyrkii hyötymään siitä. Psykologisella vaikuttamisella tarkoitetaan ihmisten psykologisten tekijöiden hyväksikäyttöä manipulaatio hyökkäyksissä. Hyökkäys-metodilla viitataan erilaisiin lukuisiin käyttäjän manipulaation käytännön sovelluksiin.

Tutkimuksen tavoite on selvittää mikä tekee käyttäjän manipulaatiosta niin tehokkaan tavan hyökätä ja ohittaa tietoturva, sekä miten käyttäjän manipulaatiota voidaan ehkäistä ja miten näiden hyökkäysten varalta voidaan valmistautua. Tämä tutkimus pyrkii vastaamaan seuraaviin tutkimuskysymyksiin:

- Miksi käyttäjän manipulaatio on niin tehokas hyökkäystapa?
- Miten käyttäjän manipulaatiota voidaan ennaltaehkäistä?

Tämä työ on toteutettu kirjallisuuskatsauksena. Lähteitä on etsitty käyttäen Google Scholaria, JYKDOK-tietokanta, sekä muutamaa luotettavaksi todettua verkkosivua. Lähteitä on etsitty muun muassa hakusanoilla social engineering, social engineering psychology, pretexting, phishing. Tutkielman tiedonhaussa on päädytty Google Scholariin, koska suomenkielistä lähdekirjallisuutta oli vain vähän saatavilla aiheesta ja kyseisestä tietokannasta saatujen julkaisujen kautta on päässyt muihin aihetta käsitteleviin teoksiin. Tiedonhaussa käytettiin myöskin Google-hakukonetta, mutta kyseistä kautta saadun informaation luotettavuus arvioitiin tarkkaan.

Tutkielman keskeisimmät tulokset olivat erilaiset psykologiset tekijät ja niiden vaikutukset käyttäjien alttiuteen joutua käyttäjän manipulaation uhriksi. Myöskin kävi ilmi, että kouluttaminen ja valistaminen aiheesta ei välttämättä ole tehokasta pitkällä aikavälillä, mutta mahdolliset jokapäiväiset muistutukset, kuten julisteet ja avaimenperät parantavat käyttäjän manipulaation vaarojen muistamista.

Tutkielman ensimmäisessä kappaleessa käydään läpi käyttäjän manipulaatiota yleisesti, sekä neljää ympäristöä, joissa nämä hyökkäykset tapahtuvat. Ensimmäisessä kappaleessa listataan myös yleisimpiä hyökkäysmetodeja ja niille ominaisia piirteitä, sekä esimerkkejä. Toisessa kappaleessa käsitellään käyttäjän manipulaatioon olennaisesti liittyvää psykologiaa pääosin perustuen Cialdinin (2001) kuuteen vaikuttamiseen liittyvään psykologiseen tekijään. Näistä kuudesta tekijästä käy ilmi, miten ja miksi hyökkääjät hyväksikäyttävät ihmisten tapaa toimia samalla tavalla samoissa tilanteissa. Myöskin Fan, Lwaktare ja Rong (2017) listaamia ihmisten perusominaisuuksia verrataan kuhunkin kuudesta Cialdinin (2001) kategoriasta, täten osoittaen mihin jokaiselta ihmiseltä löytyvään perusominaisuuteen mikäkin kategoria liittyy. Viimeisessä kappaleessa käsitellään käyttäjän manipulaation tunnistamista listaamalla hyökkäysten yleisiä piirteitä, sekä sitä, miten niitä voidaan ehkäistä. Kappaleessa käydään läpi muun muassa empiirinen tutkimus valistamisen toimivuudesta, sekä toimintamalli käyttäjän manipulaatiota epäillessä.

Tässä tutkielmassa keskitytään teknisessä, sekä sosiaalisessa ympäristössä tapahtuviin käyttäjän manipulointi -hyökkäyksiin, mutta ei oteta kantaa täysin fyysisessä maailmassa tapahtuviin hyökkäyksiin, kuten esimerkiksi roskisten penkominen.

2 KÄYTTÄJÄN MANIPULAATIO JA SEN YMPÄRISTÖT

Perinteisesti yritykset ja yksityiset tahot parantavat tietoturvaansa hankkimalla parempia virustentorjunta -ohjelmia ja käyttämällä turvalliseksi todettuja ohjelmistoja, mutta perinteiset, tekniset suojakeinot eivät auta, jos tietoturvahyökkäys pyrkii vaikuttamaan suoraan käyttäjiin. Social engineering (tästä eteenpäin käyttäjän manipulointi/manipulaatio) on tietoturvahyökkäys, joka kohdistetaan käyttäjään tietoturvajärjestelmien ohittamiseksi (Safa, Von Solms, Furnell 2016).

2.1 Käyttäjän manipulaatio käytännössä

Koska käyttäjän manipulaatio on helposti automatisoitavissa, ovat suuren mittakaavan hyökkäykset helposti toteutettavissa (Krombholz, Hobel, Huber, Weippl 2015). Esimerkkejä käyttäjän manipulaation uhreiksi joutuneista isoista yrityksistä ovat muun muassa Google vuonna 2009, RSA (yhdysvaltalainen tietoturvayritys) vuonna 2011 ja Facebook, sekä New York Times vuonna 2013. Käyttäjän manipulaatio-hyökkäykset kohdistuvat niin sanottuihin tietotyöläisiin, joiden suurin vara on tieto itsessään. Käyttäjän manipulaation motivaationa voi toimia muun muassa halu välttää teknisempää hyökkäystapaa, oppimisen/kokeilunhalu, rahallinen hyöty, kosto, ulkoinen paine, toisen hyökkääjän matkiminen tai idealistiset syyt (Chantler & Broadhurst, 2006). Hadnagyn (2010) mukaan käyttäjän manipulaation syvin tarkoitus on saada käyttäjä toimimaan tavalla, joka ei välttämättä ole heidän kannaltaan paras mahdollinen. Käyttäjän manipulaatio -hyökkäys pyrkii vaikuttamaan ihmisten sisäisiin haluihin ja ominaisuuksiin, kuten ystävyyteen, yksinäisyyteen, ahneuteen jne. ja hyväksikäyttämällä näitä asioita hyökkääjä pyrkii saamaan kohteen toimimaan haluamallaan tavalla.

Jatkuvasti globalisoituvassa maailmassa työtiimit eivät välttämättä ole ikinä nähneet toisiaan henkilökohtaisesti, vaan kommunikaatio tapahtuu sähköisesti. Tämä, sekä henkilökohtaisten laitteiden käyttäminen työelämässä ovat luoneet runsaasti uusia mahdollisuuksia käyttäjän manipulaatio -hyökkäyksille (Krombholz ym. 2015).

Seuraavassaluvussa käydään läpi, mitä käyttäjän manipulaatio on ja missä asiayhteyksissä ja ympäristöissä sitä tapahtuu, sekä listataan erilaisia metodeja, joita hyökkääjät käyttävät. Perinteisestä tietoturvahyökkäyksestä poiketen käyttäjän manipuloinnissa saatetaan toimia fyysisessä tai sosiaalisessa ympäristössä digitaalisen ympäristön lisäksi, tai mahdollisesti kaikissa kolmessa, riippuen mihin kyseinen hyökkäys kohdistuu ja mitä sillä halutaan saavuttaa (Krombholz ym. 2015). Hyökkäykset käyttävät hyväkseen tiettyjä ihmisten psykologisia haavoittuvuuksia, kuten esimerkiksi hyväntahtoisuutta,

luottamusta sekä mukavuudenhalua (Cialdini 2001). Tarkoitus on tätä kautta saada käyttäjä luovuttamaan tietoja mitä he eivät muuten luovuttaisi. Käyttäjän manipulointi käytännössä tapahtuu erilaisten huijausten ja tiedonhankinnan kautta, kuten esim. phishing -sähköpostien, puhelin huijausten tai roskisten penkomisen kautta. Myöskin sosiaalisen median kautta saatava informaatio käyttäjistä voi olla arvokasta tietoa hyökkääjälle (Teirivaara 2016). Käyttäjän manipulaatio -hyökkäykset eivät rajoitu pelkästään yrityksiin ja valtiollisiin tahoihin, vaan myös yksityiset ihmiset ovat usein näiden hyökkäysten kohteina. Yleensä yksityisiin ihmisiin kohdistuvat hyökkäykset pyrkivät identiteettivarkauteen (Uebelacker & Quiel 2014).

2.2 Käyttäjän manipulaation ympäristöt

Käyttäjän manipulaatiohyökkäyksiin liittyy Krombholz ym. (2015) mukaan neljä eri ympäristöä, jotka ovat: tekninen-, fyysinen-, sosiaalinen ja sosio-tekninen ympäristö. Ympäristöllä tässä tutkielmassa viitataan siihen, mitä kautta hyökkääjä toimii kullakin hetkellä, sillä Krombholz ym. (2015) mukaan käyttäjän manipulaatiohyökkäykset ovat monitahoisia ja sisältävät eri vaiheita, joissa hyödynnetään eri ympäristöjä. Tässä tutkielmassa keskitytään tekniseen- ja sosiaaliseen ympäristöön, sekä niiden välimuotoon, sosio-tekniseen ympäristöön ja sivutaan tarpeen mukaan fyysistä ympäristöä. Jokaista metodia ei välttämättä voida kategorisoida pelkästään yhteen ympäristöön kuuluvaksi, sillä esimerkiksi Fan ym. (2017) kategorisoivat houkuttelu -metodin tekniseksi lähestymistavaksi, kun taas Krombholz ym. (2015) mukaan se laskettaisiin sosio-tekniseksi. Voidaan siis päätellä, että kategorisointi tulee tehdä käytännön toteutuksen perusteella, sillä joillain metodeilla saattaa olla useampi toteutustapa tai/ja vaihe, kuten esimerkiksi edellä mainittu houkuttelu -metodi.

2.2.1 Tekninen ympäristö

Internet, sähköposti ja muut digitaaliset alustat lasketaan tekniseksi ympäristöksi. Grangerin (2001) mukaan internetissä tapahtuvat salasanojen kalastelu on tehokasta, koska käyttäjillä on tapana käyttää samaa, yksinkertaista salasanaa useilla eri verkkosivuilla ja mahdollisesti jopa työpaikan järjestelmissä. Fan ym. (2017) mukaan teknisessä ympäristössä tapahtuvat käyttäjän manipulaatio -hyökkäykset saattavat olla mainoksia verkkosivujen reunoilla, tai erilaisia huijaussähköposteja. Kalastelu-, sekä taukopaikka -hyökkäykset ovat esimerkkejä digitaalisessa ympäristössä toimivista metodeista. Myöskin hyökkäyksen kohteista saatava informaatio esimerkiksi erilaisten sosiaalisten medioiden ja hakukoneiden kautta voidaan laskea kuuluvan tähän kategoriaan. Teknisessä ympäristössä toimiessa voidaan hyökkäykset usein automatisoida siten, että hyökkäys tavoittaa mahdollisimman monta käyttäjää (Krombholz ym. 2015).

2.2.2 Fyysinen ympäristö

Fan ym. (2017) mukaan käyttäjän manipulaatiossa fyysisellä ympäristöllä viitataan reaalimaailmassa tapahtuviin toimiin, joiden suorittamiseksi hyökkääjän on fyysisesti tehtävä jotain halutun informaation hankkimiseksi. Tällaisia, vähemmän hienostuneita toimitapoja voivat olla esimerkiksi roskisten penkominen (engl. *dumpster diving*) asiakirjojen, tai muun arkaluontoisen materiaalin toivossa, sekä perinteiset murtovarkaudet ja olan yli vilkuileminen salasanojen toivossa.

2.2.3 Sosiaalinen ympäristö

Krombholz ym. (2015) mukaan sosiaalinen ympäristö on käyttäjän manipulaatio hyökkäyksen tärkein ympäristö. Tässä lähestymistavassa hyökkääjä on sosiaalisessa kanssakäymisessä uhrinsa kanssa esimerkiksi puhelimen välityksellä. Krombholz ym. (2015) mukaan tyypillisesti hyökkääjät pyrkivät muodostamaan suhteen uhriin ennen varsinaista hyökkäystä, jotta uhrin luottamus hyökkääjää kohtaan olisi suurempi, kasvattaen hyökkäyksen onnistumisen todennäköisyyttä.

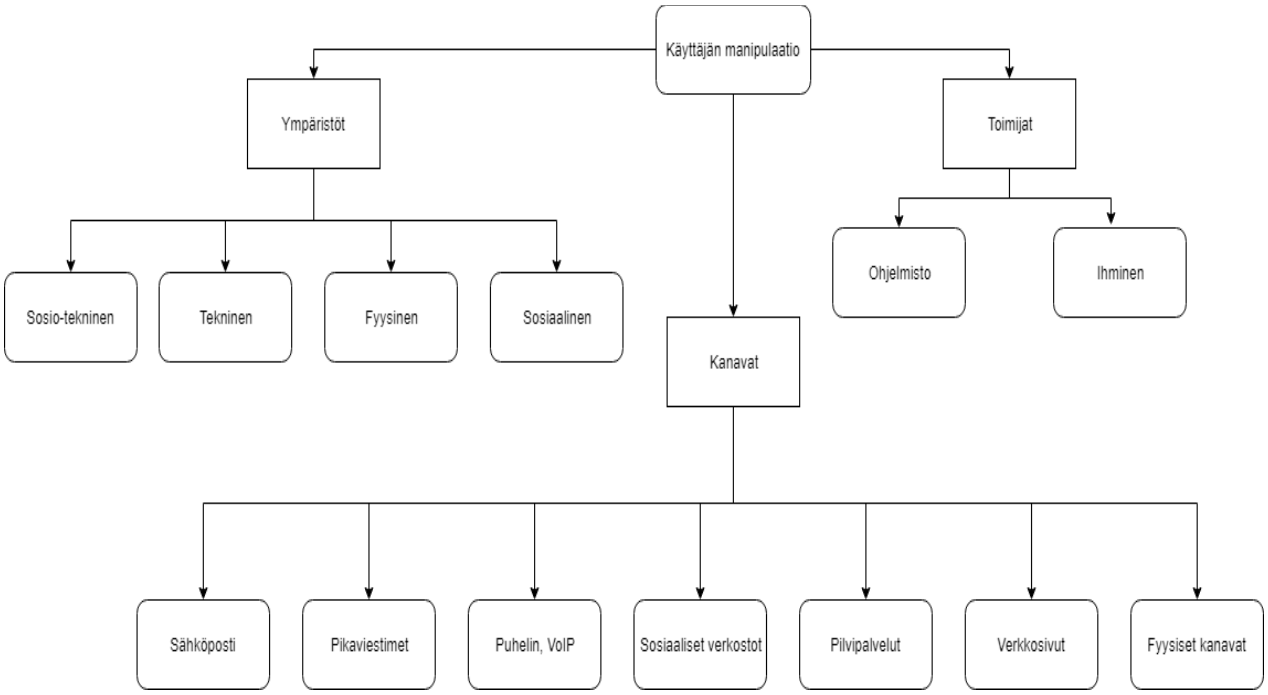
Cialdini (2001) listaa seuraavat kuusi psykologista tekijää, joihin käyttäjän manipulaatio hyökkäykset perustuvat: auktoriteetti, vastapalveluksen halu, johdonmukaisuus, sosiaalinen validointi, pidettävyys ja niukkuus. Ghafir, Prenosil, Alhejailan sekä Hammoudeh (2016) listaavat edellä mainittujen lisäksi vielä halun auttaa, sitoutumisen sekä vähäisen osallistumisen ominaisuuksiksi, joita hyökkääjät käyttävät hyväksi. Näitä psykologisia aspekteja käydään läpi tämän tutkielman myöhemmässä kappaleessa.

2.2.4 Sosio-tekninen ympäristö

Sosio-tekniset hyökkäykset voidaan nähdä Huber, Mulazzani Schrittwieser ja Weippl (2010) mukaan yhdistelmänä perinteisiä käyttäjän manipulaatio metodeja, sekä teknologiaa. Vaikka käyttäjän manipulaatio -hyökkäykset usein hyödyntävät kaikkia edellä mainittuja ympäristöjä, on sosio-tekninen ympäristö Krombholz ym. (2015) mukaan luonut tehokkaimmat aseet hyökkääjien käyttöön. Sosio-teknisissä hyökkäyksissä Huber ym. (2010) mukaan perehdytään uhriin, joko manuaalisesti tai automatisoitujen bottien avulla ja käytetään saatua informaatiota hyökkäyksessä hyödyksi. Esimerkiksi perinteisistä roskaposti-, sekä kalastelu huijauksista kehittyneemmät, sosio-tekniset versiot Huber ym. (2010) mukaan ovat tietoisia kohteesta jollain tavalla, esimerkiksi sosiaalisista verkostoista löydetyn tiedon avulla. Toimintaperiaate näillä hyökkäyksillä on kuitenkin sama kuin niiden normaali versioissa, mutta ne ovat Huber ym. (2010) mukaan erittäin paljon uskottavampia ja tehokkaampia huijauksia. Toinen esimerkki sosio-teknisestä hyökkäyksestä on seuraavassa kappaleessa esiteltävä houkuttelu- metodi, jossa hyödynnetään käyttäjien uteliaisuutta.

2.3 Esimerkkejä hyökkäysmetodeista

Käyttäjän manipulaatio -hyökkäyksissä käytetään useita erilaisia metodeja haluttuun päämäärään pääsemiseksi. Hyökkäys voi tapahtua muun muassa puhelimitse, sähköpostitse, valheellisten internet -mainosten tai haittaohjelmien avulla. Krombholz ym. (2015) listaa seitsemän käyttäjän manipulaatiohyökkäyksissä käytettävää kanavaa taksonomiassaan seuraavasti: sähköposti, pikaviestimet puhelimet/VoIP -puhelimet, sosiaaliset verkostot, pilvipalvelut, verkkosivut ja erilaiset fyysiset kanavat. Hyökkäyksen toimijana voi olla joko ihminen tai ohjelmisto. Tätä Krombholz ym. (2015) luomaa taksonomiaa havainnollistetaan kuviossa 1. Seuraavassa osiossa listataan yleisimpiä hyökkääjien käyttämiä metodeja ja käydään läpi niiden toimintaperiaatteita, sekä tunnuspiirteitä.



KUVIO 1 Käyttäjän manipulaation taksonomia (Krombholz ym., 2015, s. 116).

2.3.1 Phishing - Kalastelu

Phishing eli kalastelu hyökkäyksissä uhri ohjataan väärennetylle verkkosivulle aidon näköisen linkin kautta, siinä toivossa, että he eivät huomaisi eroa ja yrittäisivät esimerkiksi kirjautua oikeilla tunnuksillaan väärennetylle sivustolle. Kalastelu -hyökkäykset ovat yksi yleisimmistä huijauskeinoista ja Gupta, Singhal ja Kapoorin (2016) mukaan ne leviävät tyypillisimmin sähköpostin ja pikaviestinten kautta. Ferreira, Coventry ja Lenzini (2015) kategoriovat kalastelu

-sähköpostit kolmeen eri kategoriaan niiden tarkoitusten mukaan. Tietovarkaus, jossa tarkoituksena on saada haltuun käyttäjän tietoja muun muassa pankkiin, tai jonkin tilin deaktivointiin liittyvien sähköpostien avulla. Haittaohjelmia sisältävien sähköpostien tarkoitus on saastuttaa uhrin laite haittaohjelmalla sähköpostissa olevan liitteen, tai linkin kautta. Kolmanneksi kategoriaksi Ferreira ym. (2015) listaa petossähköpostit, joissa usein luvataan suuria summia rahaa vastineeksi siitä, että käyttäjä maksaa esimerkiksi rahoihin liittyvät käsittelykulut.

Tiettyyn kohteeseen, esimerkiksi yritykseen kohdennetusta kalasteluhyökkäyksestä käytetään englanninkielistä nimitystä spear phishing. Spear phisingissä hyökkääjät etsivät kohdekäyttäjistä olennaista sisäpiirin informaatiota ja lähettävät tekaistuja sähköpostiviestejä imitoiden esimerkiksi tuttuja yrityksiä (Caputo, Pfleeger, Freeman, Johnson 2013). Nämä kohdennetut hyökkäykset ovat moninkerroin tehokkaampia, kuin massoittain lähetettävät sähköpostit. Caputo ym. (2013) mukaan Ciscon (2011) tekemässä turvallisuusraportissa kävi ilmi, että miljoonasta lähetetystä perinteisestä kalastelusähköpostista kahdeksan toivat tuottoa hyökkääjälle, keskimäärin 2000 dollaria per uhri. Kohdennetuissa sähköposteissa onnistumisprosentti oli huomattavasti suurempi, tuhannesta lähetetystä kohdennetusta sähköpostista kaksi toimi, tuoden keskimäärin 80000 dollarin tuoton hyökkääjälle. Spear phishingillä oli siis tässä esimerkissä suhteellisesti 250 kertaa suurempi todennäköisyys onnistua, kuin massoittain lähetetyllä huijaussähköpostilla. Taulukko 1 havainnollistaa edellä käsiteltyä Caputo ym. (2013) raporttia Ciscon (2011) tutkimuksesta.

TAULUKKO 1 Ciscon (2011) raporttiin perustuva vertailutaulukko kalasteluviestien ja kohdennettujen kalasteluviestien välillä (Caputo ym., 2013, s. 29)

	Massoittain lähetetyt huijausviestit	Kohdennetusti lähetetyt viestit
Lähetetyt viestit	1000000	1000
Estoprosentti	99 %	99 %
Avausprosentti	3 %	7 %
Klikkausprosentti	5 %	50 %
Uhreja	8 kpl	2 kpl
Tuottoa per uhri	2,000$	80,000$
Kokonaistuotto	16,000$	160,000$
Kokonaiskustannukset	2,000$	10,000$
Nettotuotto	14,000$	150,000$

2.3.2 Houkuttelu

Houkuttelu (engl. *baiting*) -metodissa jätetään haittaohjelman sisältävä tallennusväline uhrin läheisyyteen, siinä toivossa, että se kytkettäisiin kiinni uhrin laitteeseen (Krombholz ym. (2015). Baiting hyökkäys voidaan toteuttaa muun muassa USB-tikuilla tai QR-koodeja sisältävillä lehtisillä, jotka ohjaavat haitalliselle sivustolle tai saastuttavat laitteen haittaohjelmalla. Esimerkki tällaisesta hyökkäyksestä on Operation Olympic games, joka oli Yhdysvaltojen ja Israelin salainen ja myöntämätön kyberoperaatio, jonka tarkoituksena oli sabotoida Iranin ydinohjelma stuxnet viruksella. Karnouskosin (2011) mukaan Stuxnetin uskotaan levinneen järjestelmään työntekijän kytkettyä USB-muistitikun, tai piilohallintaohjelmiston (engl. *rootkit*) sisältäneen tietokoneen kiinni järjestelmään joko vahingossa, tai tahallaan. Houkutteluun liittyvällä latinankielisellä termillä quid pro quo, joka tarkoittaa vaihtokauppaa, on Teirivaaran (2016) mukaan käyttäjän manipulaation kontekstissa samankaltainen merkitys. Hyökkääjä voi esimerkiksi tarjota palveluita kohteelle informaatiota vastaan.

2.3.3 Taukopaikka -hyökkäys

Fan ym. (2017) mukaan taukopaikka (engl. *waterholing*) hyökkäystyypissä otetaan selvää, millä verkkosivuilla hyökkäyksen kohteet käyvät ja pyritään saastuttamaan kyseinen sivusto haittaohjelmalla siten, että myös kyseisten uhrien laitteet myös saastuisivat. Etuna tällaisessa hyökkäyksessä on hyökkääjän kannalta se, että käyttäjät, jotka havaitsisivat kalastelu sähköpostit ja muut helpommin havaittavat huijaukset, eivät välttämättä osaa odottaa haitallista linkkiä tutulla ja turvallisella verkkosivulla, täten tehden heistä haavoittuvaisia hyökkäykselle.

2.3.4 Valmistelu

Valmistelu (engl. *pretexting*) ei varsinaisesti ole metodi, vaan se tarkoittaa taustatarinan tai hyökkäystilanteen valmistelua, siten että hyökkäys olisi mahdollisimman uskottava uhrin näkökulmasta. Hadnagy (2010) määrittelee valmistelun toiminnaksi, jossa luodaan skenaario uhrin suostuttelemiseksi, siten, että uhri luovuttaa informaatiota, tai tekee jotain, mitä hyökkääjä haluaa. Hadnagy (2010) mukaan valmistelun laatu riippuu täysin siitä, mitä informaatiota uhrista saadaan esille. Esimerkiksi erilaiset henkilötiedot, sekä uhrin mieltymyksen kohteet ja perustiedot, kuten perheenjäsenten nimet voivat olla arvokasta informaatiota hyökkääjälle. Valmisteluvaihe on olennainen osa jokaista manipulaatiohyökkäystä tavalla tai toisella, sillä sen avulla luodaan hyökkäyksen taustatilanne.

2.3.5 Käänteinen manipulaatio

Nelson (2008) mukaan tyypilliseen käänteiseen käyttäjän manipulaatioon (engl. *reverse social engineering*) kuuluu kolme vaihetta: sabotaasi, mainostus, sekä avustus. Käyttäjän käänteinen manipulaatio on tilanne, jossa käyttäjä hakee apua ongelmaansa, tai hyökkääjän aiheuttamaan ongelmaan, itse hyökkääjältä. Tämän seurauksena käyttäjän ja hyökkääjän välille syntyy korkeampi luottamussuhde, kuin muissa käyttäjän manipulaatio tilanteissa, joissa hyökkääjä ottaa ensin yhteyttä käyttäjään. (Krombholz ym. 2015). Käänteinen manipulaatio voi esimerkiksi toimia siten, että hyökkääjä on aiheuttanut ongelman käyttäjän tietokoneeseen ja käyttäjää kehotetaan ottamaan yhteyttä IT-tukeen, joka tosiasiassa onkin hyökkääjä. Hyökkääjä voi "mainostaa" itseään joko erilaisten internet -mainosten, tai virheilmoitusten avulla.

Esimerkki sosiaalisessa mediassa tapahtuvasta käänteisestä manipulaatiosta on Irani, Balduzzi, Balzarotti, Kirda sekä Pu (2011) tutkimus siitä, miten voidaan automatisoida käyttäjän manipulaatio hyödyntämällä tässä tapauksessa muun muassa Facebookin ystävän ehdotus toimintoa siten, että käyttäjät päätyvät lähettämään ystäväpyyntöjä hyökkääjän luomille tileille, koska heillä oli paljon yhteisiä ystäviä. Tutkimuksen aikana tämä tekaistu profiili saavutti suurimman sallitun määrän ystäviä, sekä ystäväpyyntöjä Facebookissa.

3 KÄYTTÄJÄN MANIPULAATIOON LIITTYVÄ PSYKOLOGIA

Tässä luvussa käydään läpi käyttäjän manipulaatioon liittyvää psykologiaa ja sitä, miten hyökkääjät käyttävät ihmisten eri ajattelu- ja käyttäytymistapoja hyväksi. Ihmisten kyky tehdä täysin tietoisia ja rationaalisia päätöksiä ei ole aina optimaalinen. Bezuidenhout, Mouton sekä Venter (2010) mukaan päätöksen teon rationaalisuuteen vaikuttavat muun muassa älykkyys, heuristiikat (esimerkiksi päätöksentekoon liittyvät henkiset oikotiet, jotka saattavat johtaa arviointivirheisiin), henkilökohtaiset preferenssit ja alttius manipulaatiolle muiden toimesta.

3.1 Vaikutuskeinot

Koska käyttäjän manipuloinnin on tarkoitus tunkeutua järjestelmään sen käyttäjien kautta, kohdistuu hyökkäys aina ihmisiin, jotka ovat tietoturvan heikoin lenkki. Ghafir ym. (2016) mukaan käyttäjän manipulaation psykologinen puoli perustuu hyökkääjän ja uhrin väliseen dialogiin, joten enemmän psykologiaa hyödyntävät metodit vaativat enemmän taustatutkimusta kohteesta, kuten esimerkiksi sosiaaliset lähestymistavat verrattuna massoittain lähetettäviin kalastelu sähköposteihin.

Cialdini (2001) on listannut kuusi psykologista tekijää, joita hyväksikäyttämällä ihmisten toimiin voidaan vaikuttaa: auktoriteetti, vastapalveluksen halu, johdonmukaisuus, sosiaalinen validointi, pidettävyys ja niukkuus. Myöskin Fan ym. (2017), sekä Ghafir ym. (2016) ja Bezuidenhout ym. (2010) listaavat samoja ominaisuuksia kuin Cialdini (2001), sekä lisäävät uusia ominaisuuksia, joita hyökkääjät hyväksikäyttävät, mutta jotka voidaan katsoa johdetuiksi Cialdinin (2001) kuudesta tekijästä. Esimerkkejä näistä uusista vaikutustavoista ovat muun muassa ylikuormitus, jossa käyttäjää vaivataan jatkuvilla pyynnöillä siten, että käsiteltävän informaation määrä on niin suuri, että käyttäjä ei kerkeä arvioida sitä täysin, johtaen hätiköityihin päätöksiin (Fan ym. 2017). Välinpitämättömyys, jota hyökkääjä yleensä Ghafir ym. (2016) mukaan hyödyntää siten, että manipulaatiohyökkäys kohdistuu työläisiin, joilla ei ole kiinnostusta kyseistä informaatiota tai toimintaa kohtaan ja joita kohtaan hyökkääjän on helppo vedota auktoriteettiin. Esimerkkeinä alemman tason työntekijät, kuten reseptionistit, siivoojat ja vartijat. Huolimatta hyökkäystyypistä, tai mitä psykologista tekijää hyväksikäytetään, on perimmäinen tarkoitus hämätä tai harhaanjohtaa käyttäjää. Lähes kaikissa käyttäjän manipulaatio -hyökkäyksissä hyökkääjä haluaa käyttäjän uskovan, että hän on joku, joka hän ei ole (Scheeres, 2008). Hyökkäyksen onnistuminen kuitenkin riippuu paljolti myös siitä, missä kontekstissa hyökkäys suoritetaan (Bullée, Montoya, Pieters, Junger, Hartel 2015).

Krombholz, ym. (2015) mukaan käyttäjät keskimäärin uskovat olevansa hyviä havaitsemaan käyttäjän manipulaatio hyökkäyksiin liittyviä seikkoja, vaikka esimerkiksi Qin ja Burgoon (2007) tutkimus on osoittanut, että yleisesti ihmisten kyky havaita valheita ja vilppiä on heikko. Tämä saattaa johtua optimismiharhasta, jonka Weinstein (1980) määrittelee Junger, Montoya ja Overink (2017) mukaan uskomukseksi, jossa ihmiset ajattelevat, että negatiiviset asiat tapahtuvat todennäköisemmin muille kuin itselle.

3.2 Hyökkäyksen kohteen psykologiset tekijät ja niihin vaikuttaminen

Fan ym. (2017) määrittelevät myös erilaisia ihmisten sisäisiä negatiivisia ja positiivisia ominaisuuksia, joihin erityyppiset käyttäjän manipulaatio -hyökkäykset perustuvat. Fan ym. (2017) mukaan positiivisia ominaisuuksia ovat siveys, maltillisuus, hyväntekeväisyys, ahkeruus, kärsivällisyys, kiltteys ja nöyryys. Negatiivisia ominaisuuksia taas ovat himo, ylensyönti, ahneus, laiskuus, viha, kateus ja ylimielisyys. Seuraavassa osiossa käydään läpi Cialdinin (2001) kuusi kategoriaa auktoriteetti, vastapalveluksen halu, johdonmukaisuus, sosiaalinen validointi, pidettävyys, sekä niukkuus ja miten niitä käytetään hyväksi käyttäjän manipuloinnissa, sekä listataan mitä negatiivisia- ja positiivisia ominaisuuksia Fan ym. (2017) ovat kuhunkin kategoriaan liittäneet. Yhteyksiä näiden ominaisuuksien ja kategorioiden välillä havainnollistetaan taulukossa 2 sivulla 17. Taulukon vasemmassa reunassa listataan edellä mainitut negatiiviset-, sekä niiden alla positiiviset ominaisuudet. Näihin ominaisuuksiin vaikuttavat Cialdinin (2001) kuusi psykologista tekijää on listattu taulukon yläosassa. Taulukko kuvaa, mihin ihmisen perusominaisuuteen mikäkin psykologinen tekijä vaikuttaa. Taulukkoa tulkitessa tulee ottaa huomioon, että Fan ym. (2017) eivät eksplisiittisesti kategorioineet ominaisuuksia, vaan kuhunkin kategoriaan voi liittyä useampi ominaisuus, kuin mitä on listattu.

TAULUKKO 2 Negatiivisten- ja positiivisten ominaisuuksien vaikuttamisen kategorisointi (Fan ym., 2017; Cialdini, 2001.)

	Auktoriteetti	Vastapalveluksen halu	Johdonmukaisuus	Sosiaalinen validointi	Pidettävyys	Niukkuus
Negatiiviset ominaisuudet						
Himo			X		X	X
Ylensyönti			X			X
Ahneus			X		X	X
Laiskuus			X		X	
Viha					X	
Kateus			X		X	X
Positiiviset ominaisuudet						
Siveys					X	
Maltillisuus					X	
Hyväntekeväisyys		X		X	X	
Ahkeruus					X	X
Kärsivällisyys	X				X	
Kiltteys		X		X	X	
Nöyryys	X	X		X	X	

3.2.1 Auktoriteetti

Auktoriteetti (engl. *authority*) Fan ym. (2017) mukaan vaikutuskeinona ilmenee siten, että käyttäjät myöntyvät hyökkääjän pyyntöihin helpommin, jos pyytäjä on korkeammassa asemassa käyttäjään nähden. Auktoriteetin uskottavuus vaikuttaa myös siihen, kuinka helposti uhri on halukas luovuttamaan tietoja. Auktoriteetin uskottavuutta hyökkääjän kannalta voivat parantaa muun muassa erilaiset univormut, virkamerkit ja henkilökohtaiset tittelit. Cialdinin (2001) mielestä auktoriteetti vaikutuskeinona on tehokas siksi, että ihmisille on pienestä asti opetettu, että auktoriteetin uhmaaminen on väärin ja totteleminen oikein. Ghafir ym. (2016) mukaan auktoriteetti toimii pelotteena erityisen tehokkaasti uusia- ja matalan tason työntekijöitä vastaan. Yleisin esimerkki auktoriteetin käytöstä manipulaatiohyökkäyksissä on Ghafir ym. (2016) mukaan on tilanne, jossa hyökkääjä kysyy käyttäjän salasanaa tekeytymällä turvallisuus-, IT-, tai muuksi korkean tason henkilöksi. Auktoriteetti vaikuttaa Fan ym. (2017) mukaan muun muassa ihmisten kärsivällisyyteen ja nöyryyteen.

3.2.2 Vastapalveluksen halu

Vastapalveluksen halulla (engl. *reciprocation*) Cialdini (2001) viittaa siihen, että ihmiset tuntevat olevansa velkaa niille, jotka ovat heitä aiemmin auttaneet. Ghafir ym. (2016) mukaan vastapalveluksen halua hyödynnetään varsinkin käänteisessä käyttäjän manipulaatiossa, jossa esimerkiksi hyökkääjä auttaa

käyttäjää ratkaisemaan jonkin ongelman, jonka jälkeen käyttäjä on mielestään kiitollisuudenvelassa hyökkääjälle. Sekä Cialdini (2001), Ghafir ym. (2016) ja Fan ym. (2017) mukaan vastapalvelukseen vetoaminen toimii käyttäjän manipulaatiossa, koska se on normi kaikissa yhteiskunnissa ja sen jättämättä tekemistä katsotaan pahalla. Haluun tehdä vastapalvelus riippuu muun muassa henkilön nöyryydestä, kiltteydestä ja hyväntekeväisyyden halusta (Fan ym. 2017).

3.2.3 Johdonmukaisuus

Johdonmukaisuus (engl. *consistency*) nähdään usein positiivisena piirteenä ja epäjohdonmukaisuutta taas yleensä pidetään ei-toivottuna piirteenä. Johdonmukaisuus viittaa loogiseen ajatteluun, sekä terveeseen järjenkäyttöön, eikä ilman sitä ihminen pystyisi toimimaan. Käyttäjän manipulaatio -hyökkäyksissä käytetään hyväksi johdonmukaisuudesta syntyviä ennalta-arvattavuuksia ja ihmisten liiallista sitoutumista omiin päätöksiinsä.

Chantler ja Broadhurst (2006) mukaan ihmiset pyrkivät seuraamaan sitoumuksia työpaikoilla, vaikka nämä sitoumukset vaikuttaisivat alunperinkin epävarmoilta. Työntekijät haluavat luonnostaan tulla nähdyksi luotettavina ja sitoutuneina, joten hyökkääjä voi pelotella työntekijää rangaistuksilla, tai muilla negatiivisilla seurauksilla, jos hän ei tee niin kuin hyökkääjä sanoo (Ghafir ym., 2016). Cialdinin (2001) mukaan ihmiset luontaisesti seisovat itsepäisesti päätöksiensä takana, vaikka jälkikäteen voitaisiin osoittaa, että päätös on ollut huono. Haluun olla johdonmukainen vaikuttaa Fan ym. (2017) mukaan muun muassa himo, ylensyönti, ahneus, laiskuus ja kateus.

3.2.4 Sosiaalinen validointi

Sosiaalinen validointi (engl. *social validation*) Anttilan (2016) mukaan tarkoittaa sitä, että ihmiset muuttavat käyttäytymistään ja uskomuksiaan samankaltaisiksi muiden ihmisten kanssa, koska he näkevät nämä toimet sosiaalisesti hyväksyttävinä. Cialdini (2001) väittää, että päättelemme, onko jokin oikein ottamalla selvää, mitä muut ajattelevat siitä ja pidämme jotain asiaa oikeana tapana toimia, jos näemme, että muutkin tekevät sitä. Tämä onkin Cialdinin (2001) mielestä sosiaalisen validoinnin suurin vahvuus, sekä heikkous, koska jos kaikki tekevät tietyn asian väärin, on kyseistä käyttäytymistä hyvin vaikea korjata suuren henkilömäärän vuoksi.

3.2.5 Pidettävyys

Pidettävyydellä (engl. *liking*) on suuri vaikutus siihen, miten vastaanotamme meille annettuja pyyntöjä ja käskyjä. Jos pyynnön tekevä taho tai henkilö on mielestämme pidettävä, on paljon todennäköisempää, että suostumme hänen pyyntöihinsä. Cialdini (2001) mukaan asioita, joilla on vaikutus pidettävyyteen

ovat mm. samankaltaisuus, ulkoinen viehättävyys, kohteliaisuus/kehuminen, yhteistyön tekeminen ja samaan ryhmään kuuluminen.

Uebelacker ja Quiel (2014) mukaan pidämme enemmän ihmisistä, joilla on samankaltaisia mielipiteitä, eritysiesti epäselvissä tilanteissa. Hyökkääjä voi Ghafir ym. (2016) esimerkin mukaan aloittaa keskustelun kohteen kanssa satunnaisesta aiheesta tarkoituksena luoda henkilökohtainen yhteys tämän henkilön kanssa. Tämän tarkoitus on tehdä hyökkääjästä tuttavallinen ja pidettävä uhrin näkökulmasta esimerkiksi omaamalla samoja mielenkiinnon kohteita tai olemalla kotoisin samalta alueelta.

3.2.6 Niukkuus

Ihmisillä on tapana pitää harvinaisia ja rajallisia asioita suuressa arvossa, joten jonkin asian niukkuuteen (engl. *scarcity*) vetoaminen on yleinen tapa saada käyttäjä luovuttamaan tietoja Teirivaaran (2016) mukaan niukkuuteen vetoamalla saadaan kohteessa aikaan kiireellisyyden tunne, jonka tarkoitus on heikentää kriittistä ajattelukykyä, koska mahdollisuus tehdä, tai saada jotain harvinaista on rajallinen. Fan ym. (2017) mielestä jonkin asian niukkuus vetoaa muun muassa ihmisten ahneuteen, ylensyöntiin, himoon, kateuteen, sekä ahkeruuteen.

4 KÄYTTÄJÄN MANIPULOINNIN EHKÄISYKEINOT

Suurin käyttäjän manipulaatioon liittyvä ongelma on sen havaitseminen ajoissa, sillä usein se havaitaan vasta, kun vahinko on aiheutettu ja monessa tapauksessa näitä hyökkäyksiä ei havaita lainkaan. Näiden hyökkäysten onnistumisprosenteista on kuitenkin hyvin vähän kirjallisuutta. Tässä kappaleessa käsittelemme käyttäjän manipulaation ehkäisyä ja havaitsemista, sekä Bezuidenhout ym. (2010) laatimaa SEADM-toimintamallia (Kuvio 2, sivulla 25) käyttäjän manipulaation tunnistamiseksi. Tilastokeskuksen (2018) tekemän tutkimuksen mukaan 82 % kaikista suomalaisista käytti internettiä päivittäin tai lähes päivittäin vuonna 2018 ja ikäluokissa 16-44, (noin kaksi miljoonaa suomalaista (Findikaattori (2018)) päivittäinen käyttö oli lähes 100 %. Toisin sanoen noin kaksi miljoonaa suomalaista altistuu mahdolliselle käyttäjän manipulaatiolle joka päivä käyttäessään verkkoa. Vaikka voitaisiin ajatella, että perinteiset ja paikoin ilmiselvät huijaussähköpostit menevät suoraan roskapostikansioon ja ovat muutenkin helppoja havaita ja sivuuttaa, on silti riski, että massoittain lähetettynä ne aiheuttavat jonkinlaisia vahinkoja.

4.1 Tyypillisiä piirteitä ja estäminen

Riippuen hyökkäysmetodista, on manipulaatiohyökkäyksillä tiettyjä tunnuspiirteitä. Tässä osiossa käydään läpi yleisimpiä tunnuspiirteitä ja käsitellään sitä, miten manipulaatiohyökkäyksiltä voidaan suojautua. Koska käyttäjän manipulaatio -hyökkäykset perustuvat suurilta osin samoihin psykologisiin tekijöihin, on niillä usein samanlaisia piirteitä, riippuen käytetystä hyökkäysmetodista. Esimerkkejä näistä tunnuksenomaisista piirteistä ovat muun muassa puhelinkeskustelu -tilanteessa ilmenevä painostus ja kiireellisyyden tunne, sekä lupaukset rahallisesta hyödystä erilaisissa huijausviesteissä. Yleisesti ottaen, jos jokin on liian hyvää ollakseen totta, se luultavasti ei ole totta. Myöskin vastakkainen tilanne, jossa käyttäjää uhataan sanktioilla ja muilla seuraamuksilla saattaa olla pelkkä veruke käyttäjän manipuloimiseksi. Loppukädessä Bezuidenhout ym. (2010) mukaan käyttäjän tulisi luottaa omaan vaistoonsa. Koyun ja Al Janabi (2017) listaavat seuraavanlaisia tyypillisiä piirteitä liittyen käyttäjän manipulaatioon. Suunnattoman kiireen tunteen luominen, sellaisen informaation pyytäminen, mihin kysyjällä ei pitäisi olla pääsyä, tai joka kysyjän pitäisi tietää ja liian hyvältä kuulostavat asiat, kuten lotossa voittaminen. Nelsonin (2008) mukaan käyttäjän manipulaatio toimii hyökkäystapana, koska kaikilla ihmisillä on samankaltaiset psykologiset haavoittuvuudet, joita voidaan hyväksikäyttää. Esimerkiksi vastuun jakaminen siten, että uhri ei koe olevansa itse täysin vastuussa saa uhrin antamaan tietoja helpommin.

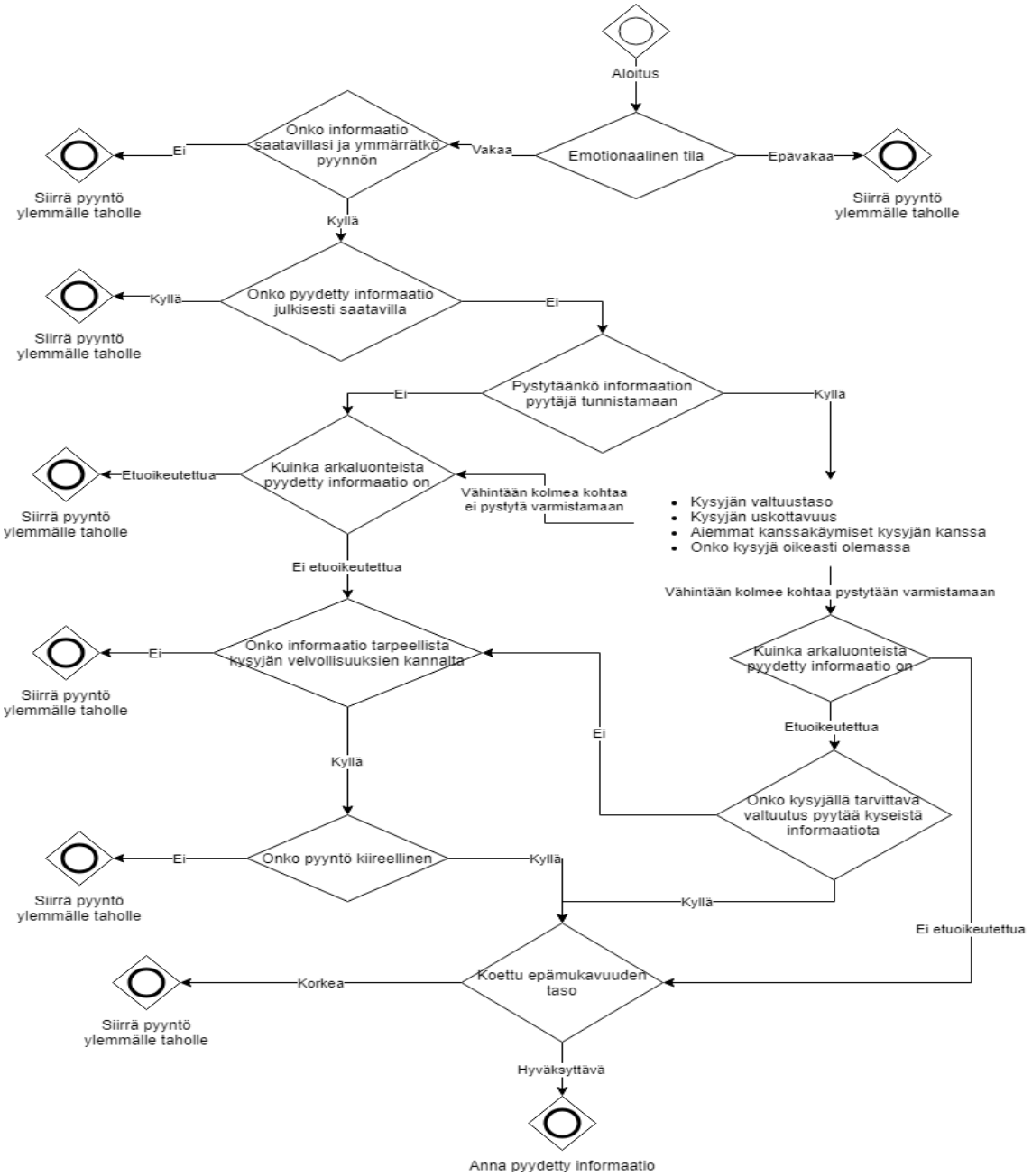
Scheeres (2008) mukaan organisaatiot käyttävät haavoittuvuuden arviointi -tiimejä tekemään testejä simuloidakseen oikeaa käyttäjän manipulaatio -hyökkäystä täten selvittäen organisaation haavoittuvuuksia. Näiden testien suorittajien haastatteluista on Scheeres (2008) mukaan käynyt ilmi, että kyseisten haavoittuvuustestien tekijöiden onnistumisprosentti informaatiota hankkiessa, tai verkostoon tunkeutuessa käyttäjän manipulaatiota hyökkäystapana käyttäen oli 100. Bullée ym. (2015) mielestä kyseiset penetraatiotestit ovat kuitenkin kontrolloimattomia ja niin sanotusti neuloja heinäsuovassa harvinaisuutensa puolesta. Myöskin erilaisilla menettelytavoilla voidaan varautua käyttäjän manipulaatiota vastaan, esimerkiksi epäilyttävässä tilanteessa soittajalle takaisin soittaminen, kuten myös kulunvalvonnan noudattaminen, sekä selkeä käsitys siitä kenellä on oikeus mihinkin informaatioon (Koyun & Al Janabi, 2017). Koyun ja Al Janabi (2017) mukaan käyttäjän manipulaatiota epäillessä tulee välittömästi katkaista yhteys hyökkääjään ja poistaa epäilyttävät sähköpostit, sekä työpaikalla informoida IT-turvallisuudesta vastaavia henkilöitä.

Nelsonin (2008) mukaan paras tapa ehkäistä käyttäjän manipulaatiota on yksinkertaisesti kouluttaa ja valistaa työntekijöitä aiheesta, mutta Bezuidenhout ym. (2010) mukaan varsinkin työelämässä, nämä koulutukset usein unohdetaan suhteellisen nopeasti. Varsinkin Ghafir ym. (2016) mainitsemien alemman tason työntekijöiden valistaminen käyttäjän manipulaation vaaroista on tärkeää, sillä he ovat haavoittuvaisin kohde käyttäjän manipulaatiolle. Fan ym. (2017) listaamat, aiemmin mainitut seitsemän positiivista ja negatiivista, sisäistä ominaisuutta voidaan liittää manipulaatiohyökkäyksiin, sillä esimerkiksi kalasteluviestien lupauksen rahasta voidaan katsoa vetoavan muun muassa ahneuteen ja pelote seuraamuksista taas voidaan katsoa ainakin laiskuuteen ja kiltteyteen vetoavaksi.

Bullée ym. (2015) tekemän empiirisen tutkimuksen mukaan työntekijöiden valistamisella käyttäjän manipulaation riskeistä oli suuri positiivinen vaikutus. Twenten yliopiston henkilökunnasta koottua otosta. Heille annettiin myös aiheeseen liittyvä avaimenperä, sekä humoristinen juliste muistuttamaan aiheesta. Hyökkäystilanteessa hyökkääjä pyrki verukkeen avulla saamaan kunkin käyttäjän avaimen haltuunsa. Lopputuloksena 37 % koulutuksen saaneista luovutti avaimensa hyökkääjälle, kun taas kontrolliryhmässä, joka ei ollut saanut minkäänlaista koulutusta aiheesta, vastaava luku oli 62,5 %. Valistuksen ja testauksen välillä oli yksi viikko, joten tutkimuksesta ei käy ilmi miten hyvänä vastakeinona koulutus toimii pidemmällä aikavälillä. Gulas ja Weinberger (2011) mukaan huumori toimii tapana muistaa informaatiota paremmin, samoin kuten pienet muistutukset Gisquet-Verrier ja Riccio (2012) mielestä, tässä esimerkissä avaimenperän muodossa.

4.2 SEADM -toimintamalli

Tässä luvussa käydään läpi Bezuidenhout ym. (2010) ehdottamaa toimintamallia käyttäjän manipulaation havaitsemiseksi. SEADM (Social Engineering Attack Detection Model) -malli on luotu helpoksi aputyökaluksi siinä tilanteessa, jos yksittäinen työntekijä epäilee mahdollista manipulaatiohyökkäystä. Kuvio 2 havainnollistaa SEADM -mallin toimintaa.



KUVIO 2 Toimintamalli yksittäisille työntekijöille käyttäjän manipulaation välttämiseksi. (Bezuidenhout ym., 2010, s. 4)

Tässä kappaleessa käydään kyseinen malli lyhyesti läpi. Mallissa etuoikeutetulla tarkoitetaan informaatiota, johon ei ole vapaata pääsyä ja pyynnön siirtämisellä ylöspäin viitataan pyynnön siirtämistä esimiehelle, jolla on enemmän tietoa asiasta.

Bezuidenhout ym. (2010) mukaan ihmisen on vaikeampi tehdä rationaalisia päätöksiä, jos tilanteeseen liittyy paljon tunteita, esimerkiksi pelkkä huono päivä voi olla ratkaiseva tekijä huijatuksi tulemisen kannalta. Jos informaatio ei ole käyttäjän saatavilla, tai hän ei ymmärrä pyyntöä, siirretään kyseinen pyyntö taholle, jolla on pääsy tähän informaatioon. Jos informaatio on julkisesti saatavilla esimerkiksi yrityksen verkkosivuilla, pitäisi yksilöillä Bezuidenhout ym. (2010) mielestä olla ymmärrys tästä ja antaa kyseinen informaatio. Jos pyydetty informaatio ei ole julkista, tulee kysyjä tunnistaa ja selvittää neljä seikkaa: riittääkö kysyjän valtuustaso pyydettyyn informaatioon. Henkilön uskottavuus, joka voidaan selvittää ennalta määrättyjen kysymysten avulla. Aiemmat kanssakäymiset pyytäjän kanssa helpottavat selvittämään mitä informaatiota kysyjälle voidaan antaa, mutta esimerkiksi puhelimen ja sähköpostin kautta toimiessa tulisi kysyjä tunnistaa erikseen. Neljäntenä tulisi selvittää onko informaation kysyjä oikeasti olemassa yrityksessä tai sen sidosryhmissä. Jos vähintään kolme edellä mainituista neljästä kohdasta pystytään varmistamaan ja kysyjällä on valtuutus pyydettyyn informaatioon, eikä toimihenkilö koe tilanne epämukavaksi, voidaan informaatio luovuttaa kysyjälle. Jos taas ei pystytä varmistamaan vähintään kolmea edellä mainittua kohtaa, tulee selvittää, kuinka arkaluontoista pyydetty informaatio on ja onko kysyjällä valtuudet siihen aiempien vastausten perusteella. Tulee myös selvittää, onko pyydetty informaatio tarpeellista kysyjän velvollisuuksien kannalta ja ottaa huomioon olisiko informaation luovuttamisesta jotain haittaa kyseisessä tilanteessa. Jos tilanne ei ole kiireinen ja pyyntöä voidaan tarkemmin konsultoida korkeammalta taholta tulisi pyyntö ohjata eteenpäin. Kiireisessä tilanteessa, kuten hätätapauksissa, tulee toimihenkilön arvioida omaa epämukavuuden tasoaan ja luottaa omaan vaistoonsa, esimerkiksi jos pyytäjä on hyvin aggressiivinen ja hyökkäävä tilanteessa aiheuttaen ahdistusta toimihenkilölle tulee pyyntö siirtää ylemmälle taholle.

Kyseistä mallia voitaisiin käyttää esimerkiksi suurissa yrityksissä, joissa on paljon alemman tason työntekijöitä, joita ei syystä tai toisesta voida kouluttaa käyttäjän manipulaatiota vastaan. Kyseistä mallia on helppo soveltaa, eikä sen ymmärtäminen vaadi varsinaista koulutusta tai vie resursseja. Bezuidenhout ym. (2010) mukaan mallin etu on sen kyky pilkkoa päätöksentekoprosessi pienempiin ja helpommin ymmärrettäviin osiin. On havaittu, että koulutuksen teho käyttäjän manipulaatiosta oikeassa työelämässä laskee ajan myötä, joten malli toimii myös muistutuksena käyttäjän manipulaation vaaroista.

5 YHTEENVETO

Käyttäjän manipulaatio on tietojärjestelmään tunkeutumista käyttäjän kautta, ohittaen virusturvat ja muut suojatekijät. Käyttäjän manipulaatio -hyökkäyksen havaitseminen on usein hyvin vaikeaa ja yleensä se käy ilmi, kun vahinko on tapahtunut ja joskus ei ollenkaan. Hyökkääjät käyttävät hyväksi ihmisille ominaisia psykologisia piirteitä, joiden perusteella heitä pystytään manipuloimaan ja toimimaan hyökkääjälle edullisella tavalla. Verrattuna perinteisempään hakkerointiin, on käyttäjän manipuloiminen usein helpompi tapa päästä käsiksi haluttuun informaatioon, sillä se ei vaadi läheskään yhtä paljon tietoteknistä osaamista. Lähdekirjallisuudessa yleinen mielipide olikin, että käyttäjän manipulaatio ei saa tarpeeksi huomiota, verrattuna siihen, miten suuri ja yleinen uhka se käytännössä on. Varsinkin nykyaikana yleistyvä omien laitteiden käyttö työympäristössä, sekä työtiimien jakautuminen useaan eri maahan tai maanosaan heikentävät suojaa käyttäjän manipulaatiota vastaan ja pahentavat ongelmaa (Krombholz ym. 2015)

Ensimmäisessä luvussa käytiin lyhyesti läpi käyttäjän manipulaatiota yleisesti ja eri ympäristöjä missä sitä tapahtuu. Hyökkäykset voivat toimia fyysisessä-, teknisessä-, sosiaalisessa- ja sosio-teknisessä ympäristössä, riippuen miten hyökkäys toteutetaan, mutta usein käyttäjän manipulaatio -hyökkäyksessä toimitaan useammassa ympäristössä (Krombholz ym. 2015). Ensimmäisessä luvussa listattiin myös yleisiä hyökkäysmetodeja, sekä avattiin niiden toimintaperiaatteita. Käyttäjän manipulaation ympäristöjen, sekä metodien pohjalta voidaan päätellä, että käyttäjän manipulaation ehkäisemiseksi ei riitä mitään tiettyä metodia tai tapaa vastaan kouluttaminen, vaan yleisesti valistaminen asiasta saattaisi olla menestyksekkäämpää, sillä erilaisia hyökkäystapoja on hyvin paljon. Ensimmäisessä kappaleessa listattiin myös esimerkkejä yleisistä hyökkäysmetodeista. Nämä metodit eivät suinkaan ole ainoita, vaan tapoja manipuloida on lukuisia ja käytännön sovelluksissa rajana on vain mielikuvitus.

Toisessa luvussa käsiteltiin käyttäjän manipulaatioon liittyvää psykologiaa, perustuen pääosin Cialdinin (2001) kuuteen psykologiseen tekijään, joiden avulla voidaan vaikuttaa ihmiseen. Näiden psykologisten tekijöiden hyväksikäyttö on suuri syy siihen, miksi käyttäjän manipulaatio on niin tehokasta ja suhteellisen helppoa toteuttaa verrattuna teknisesti toteutettuihin hyökkäyksiin. Myöskin Fan ym. (2017) listaamat ihmisen sisäiset ominaisuudet, kuten esimerkiksi ahneus, kateus ja kiltteys ovat syynä siihen, että ihmisiin on helppo vaikuttaa psykologisesti. Koska kaikki ihmiset pohjimmiltaan omaavat samanlaiset psykologiset ominaisuudet, on hyökkääjän helppo ennakoida, miten uhri tulee suurin piirtein toimimaan tietyissä tilanteissa. Tämä onkin yksi käyttäjän manipulaation etu verrattuna tekniseen hyökkäykseen, joissa tiettyyn järjestelmään kohdennettu hyökkäys vaatii enemmän tietoa ja taitoa hyökkäykseen kohteen suhteen. On kuitenkin huomioitava, että hyökkäys ei perustu pelkästään näihin psykologisiin tekijöihin, vaan myös esimerkiksi

hyökkäyksen kontekstilla, taustatyöllä ja hyökkääjän näyttelijänlahjoilla on suuri vaikutus hyökkäyksen onnistumiseen.

Viimeisessä luvussa käytiin läpi käyttäjän manipulaation torjumista ja tunnusmerkkejä, sekä perehdyttiin Bullée ym. (2015) tekemään empiiriseen tutkimukseen, sekä Bezuidenhout ym. (2010) laatimaan SEADM -malliin, joka on suunniteltu yksittäisten työntekijöiden avuksi arvioimaan informaation pyytäjän uskottavuutta. Bullée ym. (2015) tutkimuksesta kävi ilmi, että julisteiden avulla ja valistamalla ihmisiä käyttäjän manipulaation vaaroista saadaan hyviä tuloksia, mutta tutkimus ei vastaa Bezuidenhout ym. (2010) väitteeseen siitä, miten kauaskantoista tämä koulutus loppujen lopuksi on. Tämän tutkielman tarkoitus oli vastata seuraaviin tutkimuskysymyksiin: miksi käyttäjän manipulaatio on niin tehokas hyökkäys tapa, sekä miten käyttäjän manipulaatiota voidaan ennaltaehkäistä?

Käyttäjän manipulaatio on tehokas hyökkäystapa, koska se on teknisessä ympäristössä toimiessa helposti automatisoitavissa erilaisten bottien avulla, tehden siitä kustannustehokkaan ja vaivattoman hyökkääjän näkökulmasta. Myöskin erilaisten hyökkäysmetodien korkea määrä vaikeuttaa käyttäjien valistamista, koska näiden metodien toimintatavat ovat erilaisia käytännössä. Psykologia on myös suuri tekijä näiden hyökkäysten onnistumisessa, joten ihmisten luontainen tapa toimia saattaa heidät alttiiksi käyttäjän manipulaatiolle. Kaikkien ihmisten samankaltaiset psykologiset ominaisuudet auttavat hyökkääjiä luomaan huijauksia ja tilanteita, jotka usein vaikuttavat uskottavilta uhrin näkökulmasta katsottuna. Nämä psykologiset ominaisuudet myös tarkoittavat, että ihmiset toimivat suurin piirtein samalla tavalla tietyissä tilanteissa, joka mahdollistaa saman hyökkäyksen soveltamista lähes loputtomasti.

Käyttäjän manipulaatiota voidaan ehkäistä pääosin kouluttamalla ja valistamalla, mutta lähdekirjallisuus osoittaa, että koulutuksen teho laskee ajan myötä varsinkin työelämässä. Erityisen haavoittuvaisia ovat organisaatioiden alimpien tasojen työntekijät, joilla ei välttämättä ole suurempia sitoumuksia organisaatioon tehden näistä välinpitämättömiä ja helposti manipuloitavia. Manipulaatio hyökkäykset usein kohdistuvatkin näihin alemman tason työntekijöihin, joten heidän kouluttamisensa, sekä valistaminen voitaisiin nähdä erityisen tärkeänä. Esimerkiksi julisteet ja muut muistutukset saattaisivat virkistää työntekijöiden muistia käyttäjän manipulaation vaaroista. Myöskin erilaisten toimintatapojen ja protokollien noudattaminen arkaluontoista materiaalia käsiteltäessä, sekä epäilyttävissä tilanteissa auttaa ehkäisemään käyttäjän manipulaatiosta aiheutuvia ongelmia. Yritysten tietoturvakäytäntöjen testaus on myös yksi tapa mitata organisaation valmiutta, mutta kyseisiä testejä suorittavat yritykset ovat lähdekirjallisuuden mukaan harvassa, vaikka testaus itsessään on tehokas tapa löytää ja korjata heikot kohdat tietoturvassa.

Lähtökohtaisesti näiden hyökkäysten syy piilee psykologiassa ja ihmisten käyttäytymisessä. Voidaan valmistaa vaikka kuinka hyviä tietoturvaohjelmia, mutta ne eivät auta, jos niitä ei osata hyödyntää tai jos halu toimia turvallisesti puuttuu. Mielestäni tämä kirjallisuuskatsaus osoittaa, millaisia asioita ja

käyttäytymistä tulisi yrittää huomata epäillessä mahdollista manipulaatio hyökkäystä. Koska ihmisten perus käyttäytymistä ei voida helposti muuttaa, tulisi valistus ja koulutus aiheesta tehdä sen mukaisesti.

Käytännössä on vaikeaa saada kokonaiskäsitystä hyökkäysten vaikutuksista ja niiden kohteista, sillä mahdollisesti joissain tilanteissa hyökkäystä tai edes sen vaikutuksia ei olla havaittu. Tämä johtaa siihen, että tutkimukset perustuvat suurimmaksi osaksi tilanteisiin, joissa hyökkääjä on jäänyt kiinni/hyökkäys on epäonnistunut tai seuraukset on huomattu, jättäen nämä menestyksekkäimmät hyökkäykset huomiotta. Myöskin osa lähdekirjallisuudesta oli suhteellisen vanhaa (yli 10 vuotta), mutta koska hyökkäysten perusperiaate ja siihen liittyvä psykologia on pysynyt samana en näe tätä suurena ongelmana. Tulokset ovat yleistettävissä, koska ihmiset ovat psykologisesti samankaltaisia ja ehdotetut ehkäisykeinot ovat vain yleistasolla. Esimerkiksi sitä, millainen valistus tehoaa parhaiten käyttäjän manipulaatiota vastaan olisi voitu tutkia enemmän, mutta aiheesta ei löytynyt tarpeeksi kirjallisuutta perinpohjaisesti tutkittavaksi. Yleisesti lähdekirjallisuus oli luotettavaa ja niissä tuotiin useasti ilmi samoja asioita ja johtopäätöksiä.

Vaikka lähdekirjallisuus oli yksimielistä siitä, että alemman tason työntekijät ovat haavoittuvaisin kohden organisaatioissa, ei juurikaan käynyt ilmi, mitkä ovat väestöllisesti suurimmat riskiryhmät esimerkiksi iän ja sukupuolen mukaan. Myöskin näiden hyökkäysten tunnuksenomaisista piirteistä oli yllättävän vähän tutkimusta. Tietoturvan, sekä käyttäjän manipulaation tunnuspiirteiden kartoittamiseksi olisi hyvä tutkia tarkemmin mitä yhteisiä tekijöitä näillä hyökkäyksillä on. Suurimpien riskiryhmien, sekä yleisimpien tunnuspiirteiden tarkempi tutkiminen voisivat mahdollisesti laskea käyttäjän manipulaatiosta aiheutuvia tappioita huomattavasti. Mielestäni myöskään yksityisiin ihmisiin kohdistuvat manipulaatio hyökkäykset eivät ole saaneet niiden vaatimaa huomiota

LÄHTEET

Anttila, D. (2016). *Käyttäjän manipulaatio organisaation tietoturvauhkana*. (Kandidaatintutkielma). Jyväskylän yliopisto.

Bezuidenhout, M., Mouton, F. and Venter, H. (2010). Social engineering attack detection model: SEADM. *2010 Information Security for South Africa*.

Bradley, T.(2011. 30. kesäkuuta). IT Knowledge Exchange: Cisco Report--Email Attacks: This Time It's Personal. Haettu 1.4.2019 osoitteesta https://itknowledgeexchange.techtarget.com/security-detail/cisco-report-email-attacks-this-time-its-personal/

Bullée, J., Montoya, L., Pieters, W., Junger, M., & Hartel, P. (2015). The persuasion and security awareness experiment: reducing the success of social engineering attacks. *Journal Of Experimental Criminology*, *11*(1), 97-115. doi: 10.1007/s11292-014-9222-7

Caputo, D., Pfleeger, S., Freeman, J. and Johnson, M. (2014). Going Spear Phishing: Exploring Embedded Training and Awareness. *IEEE Security & Privacy*, 12(1), pp.28-38.

Chantler, A., & Broadhurst, R. (2006). *Social Engineering and Crime Prevention in Cyberspace.* Queensland University of Technology

Cialdini, R.B. (2001). *Influence: Science and Practice.* Allyn & Bacon. Haettu Haettu 1.4.2019 osoitteesta https://www.influenceatwork.com/wp-content/uploads/2012/02/Influence_SP.pdf

Ferreira, A., Coventry, L., & Lenzini, G. (2015). Principles of Persuasion in Social Engineering and Their Use in Phishing. *Lecture Notes In Computer Science*, 36-47. doi: 10.1007/978-3-319-20376-8_4

Findikaattori (2019, 29. maaliskuuta) Väestön ikärakenne. Haettu 1.4.2019 osoitteesta https://findikaattori.fi/fi/14

Ghafir, I., Prenosil, V., Alhejailan, A. and Hammoudeh, M. (2016). Social Engineering Attack Strategies and Defence Approaches. *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*.

Gisquet-Verrier, P., & Riccio, D. (2012). Memory reactivation effects independent of reconsolidation. *Learning & Memory*, 19(9), 401-409. doi: 10.1101/lm.026054.112

Granger, S. (2001*). Social Engineering Fundamentals, Part I: Hacker Tactics.* SecurityFocus, 2001

Fan, W., Lwakatare, K., & Rong, R. (2017). Social Engineering: I-E based Model of Human Weakness for Attack and Defense Investigations. *International Journal Of Computer Network And Information Security*, *9*(1), 1-11. doi: 10.5815/ijcnis.2017.01.01

Gulas, C., & Weinberger, M. (2006). *Humor in Advertising: A Comprehensive Analysis*.

Gupta, S., Singhal, A. and Kapoor, A. (2016). A literature survey on social engineering attacks: Phishing attack. *2016 International Conference on Computing, Communication and Automation (ICCCA)*.

Hadnagy, C. (2010). *Social engineering: The art of human hacking*. John Wiley & Sons. Haettu 1.4.2019 osoitteesta https://books.google.fi/books?hl=fi&lr=&id=9LpawpklYogC&oi=fnd&pg=PT7&dq=pretexting+social+engineering&ots=vbjtFVk4VS&sig=vNujKiJ32AHcjTsRp3uzp1FByFc&redir_esc=y#v=onepage&q=pretexting&f=false

Huber, M., Mulazzani, M., Schrittwieser, S., & Weippl, E. (2010). Cheap and automated socio-technical attacks based on social networking sites. *Proceedings Of The 3Rd ACM Workshop On Artificial Intelligence And Security - Aisec '10*. doi: 10.1145/1866423.1866435

Irani, D., Balduzzi, M., Balzarotti, D., Kirda, E. and Pu, C. (2011). Reverse Social Engineering Attacks in Online Social Networks. *Detection of Intrusions and Malware, and Vulnerability Assessment*, pp.55-74.

Junger, M., Montoya, L., & Overink, F. (2017). Priming and warnings are not effective to prevent social engineering attacks. *Computers In Human Behavior*, *66*, 75-87. doi: 10.1016/j.chb.2016.09.012

Karnouskos, S. (2011). Stuxnet worm impact on industrial cyber-physical system security. *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*.

Koyun, A., & Al Janabi, E. (2017). Social Engineering Attacks. *Journal Of Multidisciplinary Engineering Science And Technology (JMEST)*, *4*(6).

Krombholz, K., Hobel, H., Huber, M. and Weippl, E. (2015). Advanced social engineering attacks. *Journal of Information Security and Applications*, 22, pp.113-122.

Nelson, R. (2008). *Methods of hacking: social engineering.* Haettu 1.4.2019 osoitteesta https://s3.amazonaws.com/academia.edu.documents/32169493/Methods_of_Hacking_-_Social_Engineering.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3

A&Expires=1554158017&Signature=%2BvoPToCRJ4Eul%2Fj1tvLj1gxF%2Flk%3D&response-content-disposition=inline%3B%20filename%3DMethods_of_Hacking_-_Social_Engineering.pdf

Perez, R. (2016). 60% of enterprises were victims of social engineering attacks in 2016. Haettu 1.4.2019 osoittesta https://www.scmagazineuk.com/60-enterprises-victims-social-engineering-attacks-2016/article/1475806

Qin, T., & Burgoon, J. (2007). An Investigation of Heuristics of Human Judgment in Detecting Deception and Potential Implications in Countering Social Engineering. *2007 IEEE Intelligence And Security Informatics*. doi: 10.1109/isi.2007.379548

Scheeres, J. (2008). *Establishing the human firewall: reducing an individual's vulnerability to social engineering attacks*. Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio.

Sohrabi Safa, N., Von Solms, R. and Furnell, S. (2016). Information security policy compliance model in organizations. *Computers & Security*, 56, pp.70-82.

Teirivaara, T. (2016*). Tietoturvan ihmiselementti: sosiaalinen manipulointi.* (Kandidaatintutkielma). Jyväskylän yliopisto.

Tilastokeskus (2018, 4. joulukuuta), Suomalaisten internetin käyttö 2018– viestintää, asiointia, tiedonhakua ja medioidenseuraamista. Haettu 1.4.2019 osoitteesta https://www.stat.fi/til/sutivi/2018/sutivi_2018_2018-12-04_kat_001_fi.html

Uebelacker, S., & Quiel, S. (2014). The Social Engineering Personality Framework. *2014 Workshop On Socio-Technical Aspects In Security And Trust*. doi: 10.1109/stast.2014.12