

Jose Malmberg

**Ihmisen ominaisuudet ihmisen ja koneen automatisoidussa
erottamisessa**

Tietotekniikan kandidaatintutkielma

28. huhtikuuta 2016

Jyväskylän yliopisto

Tietotekniikan laitos

Tekijä: Jose Malmberg

Yhteystiedot: jose.j.malmberg@student.jyu.fi

Ohjaaja: Hannakaisa Isomäki

Työn nimi: Ihmisen ominaisuudet ihmisen ja koneen automatisoidussa erottamisessa

Title in English: Human characteristics in telling humans and computers apart automatically

Työ: Kandidaatintutkielma

Sivumäärä: 23+0

Tiivistelmä: Tietotekniikan kehityksen myötä yhä useammat aiemmin ihmisen tehtävänä olleet työt ovat siirtyneen tietokoneiden vastuulle. Monissa tehtävissä kone kykeneekin jo nyt korvaamaan ja jopa ylittämään ihmisen kyvykkyyden. Tässä tutkielmassa selvitettiin automatisoidun Turingin testin (eng. Completely Automated Public Turing Test to Tell Computers and Humans Apart, CAPTCHA) uusien toteutuksien avulla mitä ihmisen ominaisuuksia koneilla ei vielä ole. Näitä ominaisuuksia ovat monet ns. intuitiiviset kyvyt kuten luonnollisen kielen poikkeavuuksien tunnistaminen, kolmiulotteista hahmottamista tukevat syvyysnäön piirteet, päättelykyky sekä kyky tulkita epäsuoria sisäisiä merkityksiä kuten huumoria.

Avainsanat: CAPTCHA, ihminen, ominaisuus, ominaisuudet, Turingin testi

Abstract: Many tasks formerly conducted by humans have been transferred to be executed by computers, along with the development of the information technology. At many tasks the computers are able to replace and even to exceed human abilities at the moment. This paper examines the new implementations of the Completely Automated Public Turing Test to Tell Computers and Humans Apart (CAPTCHA) to find out what human abilities computers still do not have. This kind of abilities are for example many so called intuitive abilities such as the identification of anomalies in the natural language, some supporting properties of the stereoscopic vision, the reasoning ability and the ability to interpret indirect internal meanings, such as humor.

Keywords: CAPTCHA, human, characteristic, characteristics, Turing test

Kuviot

Kuvio 1. Esimerkki tekstipohjaisesta captchasta.....	6
Kuvio 2. Esimerkki kuvapohjaisesta captchasta	8

Sisältö

1	JOHDANTO	1
2	IHMISEN JA TIETOKONEEN AUTOMATISOITU EROTTAMINEN.....	3
2.1	Automatisoidun erottamisen toimintaperiaate	3
2.2	Automatisoidun erottamisen käyttökohteet	4
3	YLEISIMMÄT TOTEUTUKSET IHMISEN JA KONEEN EROTTAMISEKSI.....	6
3.1	Tekstipohjainen toteutus	6
3.2	Äänipohjainen toteutus.....	7
3.3	Kuvapohjainen toteutus	7
4	UUDET TOTEUTUKSET IHMISEN OMINAISUUKSIEN PERUSTEELLA.....	10
4.1	Hahmotuskykyyn perustuvat toteutukset	10
4.2	Syvyysnäköön perustuva toteutus	12
4.3	Poikkeavuuden tunnistamiseen perustuva toteutus	13
4.4	Päätelykykyyn perustuva toteutus	14
4.5	Huumorintajuun perustuva toteutus	15
5	JOHTOPÄÄTÖKSET.....	16
	LÄHTEET	17

1 Johdanto

Tietokoneiden tehtävänä on alusta asti ollut korvata aiemmin vain ihmisille mahdollisia, mutta rutiininomaisia ja paljon toistoa vaativia tehtäviä. Teknologian kehittyessä ihmiset ovat onnistuneet ratkaisemaan koneella yhä vaativampia ongelmia ja tämän seurauksena koneiden vastuulle on uskottu yhä useampia ja aiempaa monimuotoisempia tehtäviä. Monilla osalualueilla kone kykeneekin jo nyt korvaamaan tai jopa ylittämään ihmisen kyvykkyyden.

Kaikkia ihmisen ominaisuuksia ei silti vielä kyetä koneellisesti korvaamaan, joten näistä vielä jäljellä olevista ominaisuuksista on hyvä olla tietoinen. Kun tiedostetaan ihmisen ja koneen erilaisuus ja molempien vahvuudet, voidaan näitä molempia yhdistää kunkin ongelmatilanteen vaatimalla tavalla parhaan lopputuloksen aikaansaamiseksi. Erityisesti näin voidaan välttää se, ettei konetta turhaan yritetä käyttää johonkin, minkä suorittamiseen sillä ei tällä hetkellä ole edellytyksiä. Vastaavasti ihmistä on usein epäedullista käyttää tehtäviin, jotka voitaisiin suorittaa yhtä hyvin koneellisestikin.

Tässä kirjallisuuskatsauksessa aihetta lähestytään siitä näkökulmasta mitä ihmisen ominaisuuksia hyväksikäyttäen ihminen ja tietokone voidaan erottaa toisistaan. Ihmisen ja koneen erottamisella on olemassa omat tärkeät sovelluskohteensa, mutta samalla sen avulla voidaan pohtia sitä missä tilanteissa olisi edelleen hyödyksi soveltaa ihmisen ominaisuuksia. Katsauksen tavoitteena on kaiken kaikkiaan siis verrata ihmistä ja tekoälyä eli tietokoneohjelmaa keskenään, joten parhaiten tähän tarkoitukseen sopii Turingin testi (Turing 1950).

Tänä päivänä yleisin Turingin testin sovellus on täysin automatisoitu julkinen Turingin testi ihmisten ja tietokoneiden erottamiseksi (eng. Completely Automated Public Turing Test to Tell Computers and Humans Apart, CAPTCHA). Turingin testiin perustuen CAPTCHA on suunniteltu siten, että sen läpäiseminen edellyttää ratkaisijaltaan sellaisia inhimillisiä ominaisuuksia, joita tietokoneohjelmilla ei ainakaan vielä ole (Ahn ym. 2003). Näin ollen tutkinnalla uusimpia CAPTCHA:n toteutuksia voidaan käytännönläheisesti kartoittaa mitkä ihmisen ominaisuudet ovat vielä tällä hetkellä ohjelmien saavuttamattomissa.

Tutkimuksen metodina on käytetty systemaattista kirjallisuuskartoitusta Flinkin mallin mukaan, kuten Salminen (2011) sen esittää. Hakujen tietokantoina on käytetty ensisijaisesti

IEEE:n ja ACM Digital Libraryn tietokantoja, ja tutkimuskysymys on ollut muotoa ”Mitä ihmisen ominaisuuksia tietokoneella ei vielä ole?” Näistä tietokannoista saadut tulokset rajattiin seulonnan kautta vain CAPTCHAa koskeviin artikkeleihin. Löydettyjen artikkeleiden perusteella kartoitusta jatkettiin edelleen tutkijoiden Levy ja Ellis (2006) määrittelemiä forward ja backward search – metodeja käyttäen.

Tässä katsauksessa luvussa 2 kuvaillaan CAPTCHA:n teoreettista taustaa ja sen käyttökohteita, minkä jälkeen luvussa 3 kerrotaan tarkemmin yleisimmistä CAPTCHA:n toteutusratkaisuista. Luvussa 4 selostetaan käytännön tutkimusten ja esimerkkien kautta sitä millä tavalla ihmisen tiettyjä ominaisuuksia ja vastaavasti tekoälyn heikkoutta hyödynnetään uusimmissa vielä toimivissa CAPTCHA toteutuksissa ja lopuksi luvussa 5 kootaan yhteen tutkimuksen johtopäätökset ja suuntaa mahdollisille jatkotutkimuksille.

2 Ihmisen ja tietokoneen automatisoitu erottaminen

Ihmisen ja tietokoneen automatisoituun erottamiseen käytettävästä testistä käytetään nimitystä CAPTCHA (eng. Completely Automated Public Turing test to tell Computers and Humans Apart). Tämä voi olla mikä tahansa automaattisesti generoitavissa oleva testi, jonka suurin osa ihmisistä läpäisee, mutta jota nykyiset tietokoneohjelmat eivät kykene läpäisemään. Testi perustuu Turingin testin toimintaperiaatteeseen, mutta alkuperäisestä järjestyksestä poiketen CAPTCHAN toteutuksessa ihmisen ja koneen erottamisen tuomarina toimii tietokone. Termin "CAPTCHA" lisäksi testistä on aiemmin käytetty myös nimitystä käänteinen Turingin testi (eng. Reverse Turing Test, RTT). Koska tämä voi kuitenkin viitata myös testiin, jossa molemmat osallistujat pyrkivät esiintymään tietokoneohjelmina, automatisoituun ihmisen tunnistamiseen viitattaessa käyttöön on vakiintunut termi CAPTCHA¹. (Ahn, Blum ja Langford 2004)

2.1 Automatisoidun erottamisen toimintaperiaate

Captchan taustalla on ajatus luoda niin vaikeita inhimillistä älykkyyttä vaativia ongelmia, etteivät kullakin hetkellä parhaatkaan tietokoneohjelmat kykene niitä ratkaisemaan. Kyse on siis kryptografiassa käytetystä metodista, jossa tarkoituksena on salata haluttu tieto laskennallisesti niin vaativalla ongelmalla, että vaikka ratkaisumalli on kaikkien saatavilla, salauksen purkaminen ei tästä huolimatta ole mahdollista kohtuullisessa ajassa. Samalla tavalla captchan ratkaisun ollessa suurimmalle osalle ihmisistä ilmeinen, ratkaisun kehittäminen olisi periaatteessa mahdollista kenelle tahansa asiansa osaavalle yksilölle, mikäli ratkaiseminen koneellisesti suinkin olisi mahdollista. Tällä tavoin eri toimijoiden yrittäessä ratkaista kulloinkin kyseessä olevaa captchaa ohjelmallisesti, tullaan samalla edistäneeksi tekoälyn kehittymistä. Parhaimmillaan näin saadaan yhdistettyä sekä tekoälyn- ja turvallisuuden tutkijoiden että jopa pahantahtoisten ohjelmoijien työpanos tekoälyn kehittämiseksi. (Ahn ym. 2003)

Määritelmästä seuraavan ihmisen ja koneen erottamista koskevan vaatimuksen lisäksi Klue-

1. Tästä eteenpäin CAPTCHA kirjoitetaan pienin kirjaimin ("captcha") luettavuuden helpottamiseksi.

ver ja Zanibbi (2009) ovat esittäneet captchan generoimiselle seuraavat tavoiteltavat ominaisuudet:

- Automatisoitu (Automated)
- Avoin (Open)
- Käytettävä (Usable)
- Turvallinen (Secure)

Automatisoitu tarkoittaa, että tehtävien automaattisen generoimisen ja vastauksien analysoimisen tulisi olla tuomarina toimivalle ohjelmalle helppoa. Avoimuus taas tarkoittaa sitä, että testin luomiseen käytettyjen algoritmien ja datan tulisi olla julkista eli kaikille saatavilla. Edelleen siis samoin kuin kryptografiassa, ohjelman täytyy olla suunniteltu siten, että se säilyy turvallisena vaikka olisi kokonaisuudessaan julkinen. Käytettävyys taas määritellään ihmisen näkökulmasta helppoutena ratkaista annettu ongelma, riippumatta käyttäjän kielestä, fyysisestä sijainnista, koulutuksesta tai hahmotuskyvystä.

Kluever ja Zanibbi (2009) esittävät turvallisuuden viittaavan vain siihen kuinka testin tulisi olla algoritmillisesti (i. automatisoidusti) vaikeasti ratkaistava. Alkuperäisen tekstipohjaisen captchan toteutuksen virallisessa ohjeistuksessa turvallisuus jaetaan edelleen tarkempiin yksittäisiin osa-alueisiin: kuvien turvallisuuteen, sovelluskoodin turvallisuuteen ja turvallisuuden säilymiseen laajamittaisenkin käyttöönoton jälkeen. Tässä listauksessa turvallisuus on siis määritelty ennen kaikkea testiohjelman toteutuksen näkökulmasta, ja sen tavoitteena on minimoida ohjelman haavoittuvuus. Tällä tavalla halutaan taata, että käyttäjälle kulloinkin annetun ongelman onnistunut ratkaiseminen perustuu todella ratkaisijan inhimillisiin ominaisuuksiin eikä vain testin toteuttamiseksi käytetyn ohjelman heikkouksiin. (University 2010)

2.2 Automatisoidun erottamisen käyttökohteet

Captcha käytetään silloin, kun palvelun tarjoajan kannalta on olennaista pitää jokin palvelun osa vaivattomasti ihmisten saatavilla mutta samalla rajoittaa automaattisten ohjelmien pääsyä kyseiseen osaan. Tällaisia tilanteita on yleisimmin kahta tyyppiä: sisällön tuottamisen estäminen ja jo olemassa olevan sisällön suojaaminen.

Sisällön tuottamisen estämisessä perinteisin ja muutoinkin yleisin captchan käyttökohde on sähköpostin tai muiden ilmaisten käyttäjätilien luomisen rajaaminen vain ihmisten saataville. Vastaavalla tavalla captchaa voidaan käyttää varmistamaan, että vain ihmiset voivat lisätä kommentteja blogiin tai muuhun verkkojulkaisuun. Näiden lisäksi erityisesti verkkopohjaisissa äänestyksissä tai vastauslomakkeissa on tärkeää varmistua siitä, että vain todellisilla henkilöillä on mahdollisuus antaa äänensä, eivätkä pahantahtoiset toimijat pääse vaikuttamaan tuloksiin tuottamalla suurta määrää keinotekoisia vastauksia. (University 2010)

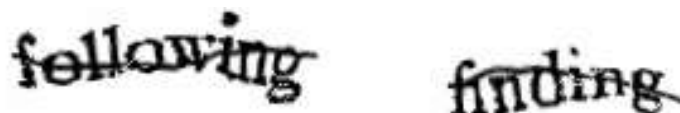
Captchaa voidaan käyttää sisällön suojaamiseen esimerkiksi tilanteessa, jossa ei tahdota julkisen sähköpostiosoitteen tulevan kerätyksi automaattisesti roskapostittajan tai muun ei-halutun toimijan listalle. Tällöin osoite voidaan piilottaa siihen asti kunnes käyttäjä on onnistuneesti ratkaissut annetun captcha-testin. Vastaavalla tavalla captcha-testin läpäisemistä voidaan vaatia salasanasuojatuissa palveluissa aina salasanan syöttämisen yhteydessä, ja tällä tavoin ehkäistä palveluun kohdistuvia sanakirjahyökkäyksiä. Joissain tilanteissa internetsivustontarjoaja voi haluta pitää sivustonsa hakukoneiden indeksoinnin ulkopuolella ja tällöinkin voidaan käyttää captchaa estämään automaattisten indeksoijien pääsy sivustolle. (University 2010)

3 Yleisimmät toteutukset ihmisen ja koneen erottamiseksi

Captchan määritelmä ei ota kantaa siihen millainen testin tehtävän käytännössä tulisi olla ihmisen ja koneen erottamiseksi. Toteutustapa on siis täysin kunkin toteuttajan valittavissa niin pitkään kuin tehtävä vain on suurimmalle osalle ihmisistä suhteellisen helposti ratkaistavissa, mutta koneelle vaikeaa. Vaikeuden toteutumiseksi kirjallisuudessa on määritelty, että koneen tulee kyetä ratkaisemaan testi korkeintaan 1 %:ssa yrityksistä (Bursztein, Martin ja Mitchell 2011). Tässä luvussa esitellään kolme yleisintä captchan toteutusta pääpiirteissään ja toimintaperiaatteineen sekä otetaan kantaa niiden toimivuuteen.

3.1 Tekstipohjainen toteutus

Ensimmäiset käytännössä toteutetut captchat olivat tekstipohjaisia ja yhä tälläkin hetkellä tekstipohjainen captcha on kaikkein yleisin captchan variaatio. Toteutukseltaan ja toimintaperiaatteeltaan tekstipohjainen captcha on varsin yksinkertainen ja perustuu eri tavoin vääristellyn tekstin lukemiseen. Käytännössä captchan generoiva ohjelma valitsee tietokannastaan yhden tai useamman sanan ja tuottaa näistä kuvan, jossa sanat esiintyvät jollain tavalla vääristellyssä muodossa (ks. Kuvio 1). Käyttäjän tehtävänä on syöttää alkuperäiset sanat takaisin ohjelmalle. (Ahn, Blum ja Langford 2004)



Kuvio 1. Esimerkki tekstipohjaisesta captchasta. ¹

Tämä toteutuksen toimivuus perustui siihen, että konenäkö ja tekstin hahmottaminen olivat 2000-luvun alussa tietokoneohjelmille hyvin vaikeita tehtäviä. Ihmisille taas hieman vääristellyn tekstin lukeminen ei juuri tuota vaikeuksia tehtävän arkipäiväisyyden ja sen kautta kertyneen laajan kokemuksen tähden. Konenäkö ja tekstin tunnistaminen ovat kuitenkin kehittyneet viimeisen parin vuosikymmenen aikana niin paljon, että tekoälyä tutkiva ja kehittä-

1. Lähde: BMaurer at English Wikipedia - Own work, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=2201550>

vä yritys Vicarious ilmoitti vuonna 2013 kykenevänsä läpäisemään tekstipohjaisen captchan automaattisesti 90%:ssä yrityksistä (Vicarious 2013). Tämä ylittää huomattavasti captchan murtamiseksi määritellyn 1 %:n tarkkuuden, joten käytännössä tekstipohjainen toteutus ei enää täytä ihmisen ja koneen erottamiseksi asetettuja vaatimuksia.

3.2 Äänipohjainen toteutus

Äänipohjainen captcha kehitettiin hyvin pian tekstipohjaisen testin kehittämisen jälkeen (Kochanski, Lopresti ja Shih 2002). Äänipohjainen captcha ei kuitenkaan ole yleistynyt yhtä nopeasti kuin tekstipohjainen eikä levinnyt aivan yhtä laajalti, mutta nykyisin varsin useat suuret palveluntarjoajat tarjoavat myös äänipohjaista captchaa vaihtoehtona tekstipohjaisen rinnalla. Äänipohjaista captchaa käytetään erityisesti silloin kuin käyttäjän tunnistaminen jollain palvelun ulkopuolisella metodilla (puhelinsoitto, sähköposti) ei ole kannattavaa, mutta palvelun saatavuus halutaan taata myös visuaalisesti rajoittuneille käyttäjille. Suurimmassa osassa nykyisiä palveluita tilanne on edellä kuvatus kaltainen, jolloin äänipohjainen captcha on ainoa varteen otettava vaihtoehto. (Bigham ja Cavender 2009)

Äänipohjaisessa captchassa käyttäjän tehtävänä on tunnistaa äänen lausuttuja kirjaimia taustamelua/häiriöitä sisältävältä äänitteeltä. Toimintaperiaate on siten käytännössä sama kuin tekstipohjaisessakin captchassa, sillä erotuksella, että nauhalle mahdollisesti lisätty taustamelu ja kunkin kirjaimen ääntämisasun luonnollinen vaihtelu vastaavat tekstin vääristämistä. Toimintaperiaatteen samankaltaisuudesta johtuen äänipohjainen captcha on siten altis teoreettisesti samankaltaiselle murtamiselle kuin tekstipohjainenkin. Bursztein ym. (2011) ovatkin osoittaneet, että pahimmassa tapauksessa tietokone kykenee ratkaisemaan äänipohjaisen captchan jopa paremmin kuin ihminen. Näin ollen äänipohjainenkin captcha on toteutukseltaan vanhentunut.

3.3 Kuvapohjainen toteutus

Kuviin perustuva captcha kehitettiin teksti- ja äänipohjaisen captchan jälkeen vaihtoehtona, jolla välttää näissä esiintyviä heikkouksia. Kirjoitettujen tai lausuttujen kirjaimien sijaan kuvapohjainen toteutus perustuu kuviin ja niiden sisällön tunnistamiseen sekä semanttiseen









analysointiin. Yksinkertaisimmillaan tämä tarkoittaa esimerkiksi oikeasisältöisten kuvien valitsemista useiden kuvien joukosta (ks. Kuvio2). Myös monitasoisempia useiden kuvien semanttisten sisältöjen tunnistamista ja yhdistämistä vaativia toteutuksia on esitelty (Vikram, Fan ja Gu 2011). Kuvapohjaisen captchan käyttämiä kuvia ei välttämättä ole millään tavalla vääristelty tunnistamisen vaikeuttamiseksi, vaan sisällön tulkitsemisen katsotaan olevan jo itsessään tarpeeksi haasteellista nykyisille ohjelmille.

Create your account

Username

Password

Security check
 Select the image that is different from this group

[Create account](#)

Kuvio 2. Esimerkki kuvapohjaisesta captchasta.²

2. Lähde: By Pau Giner - The image includes other images in the public domain:Friedrich Herlin, Reading Saint Peter (1466); Beard_004aMajor; 00"NATO"(Pražský studentský summit, 2011); Picture_of_Levi_Sullivan%22_Oct 2010; Gatica%27_negra04; James_Wiley_Magoffin%27_by_Henry_Cheever_Pratt; John_Coats_Browne%27,_painting_by_the_Joseph_Wright; Almeida_J%C3%BAnior_-_Caipira_Pitando%2C_1895; CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=20500953>

Tehtävään luonteen vuoksi kuvapohjainen captcha vaatii aina jonkinlaisen tunnistetiedoilla varustetun kuvatietokannan. Koska tällainen tietokanta on väistämättä rajallinen ja tehtävä on yleensä jonkinlainen luokitteluongelman erityistapaus, ohjelmien on koneoppimista hyväksikäyttäen mahdollista ennen pitkää tunnistaa kunkin tietokannan kuvien sisällöt. Koneoppimisen ollessa jatkuvasti kehittyvä tutkimuksen ala, kuvapohjaisenkin captchan vanheneminen on vain ajan kysymys. (Golle 2008)

Kuvapohjaisen captchan erityistapaus, ja siitä edelleen kehitetty variaatio, on videopohjainen captcha. Kluever ja Zanibbi (2009) ovat esitelleet erään tällaisen toteutuksen, jossa ihmisen ja koneen erottaminen tapahtuu nimeämällä näytetyn videon sisältöön sopivia avainsanoja. Toimintaperiaate on siis samanlainen kuin monissa yksittäisiin kuviin perustuvissa captcha toteutuksissa. Koneellisen ratkaisemisen kannalta suurin ero näissä kahdessa onkin vain vaadittavan laskentatehon määrässä. Koneiden laskentatehon kasvaessa videon tulkitseminen ja siitä oppiminen on tulevaisuudessa yhtä vaivatonta kuin kuvien tulkitseminen tällä hetkellä. Kluever ja Zanibbi (2009) toteavatkin artikkelissaan heidän toteutuksensa ratkeavan jo tuolla hetkellä jopa 13%:ssa hyökkäyksistä.

4 Uudet toteutukset ihmisen ominaisuuksien perusteella

Tietokoneiden kehittyessä ja saadessa yhä uusia aiemmin ihmisen yksinoikeutena olleita ominaisuuksia vanhat captchat eivät enää täytä tehtäväänsä kuten tulisi. Niinpä edellisessä luvussa esiteltyjen perinteisten ja yleisimpien captcha-toteutuksien jälkeen on esitelty useita ehdotuksia, jotka hyödyntävät ihmisen ominaisuuksia ja koneen heikkouksia uusin tavoin. Tässä luvussa esitellään ihmisen ominaisuuksien perusteella lajiteltuna muutamia vuoden 2010 jälkeen esiteltyjä captcha-konsepteja, jotka tämän tutkielman kirjoittamishetkellä kykenevät erottamaan ihmisen ja koneen toisistaan.

4.1 Hahmotuskykyyn perustuvat toteutukset

Ensimmäiset tekstipohjaiset captchat perustuivat ihmisen kykyyn tunnistaa jollain tavalla vajaita tai vääristeltyjä kirjainmerkkejä. Vaikka konenäkö onkin noista ajoista kehittynyt huomattavasti tavallisen tekstin tunnistamisen suhteen, monet muut ihmisen hahmotuskykyyn piirteet ovat edelleen ohjelmille vaikeita. Näistä piirteistä eräs tärkeä ja hyvin arkipäiväinen on epätäydellisen havainnon täydentäminen (eng. amodal completion). Ihminen kykenee hahmottamaan näkemänsä kappaleen kokonaisuutena vaikka se olisi osittain peittyneet muiden kappaleiden taakse. Esimerkiksi suurin osa ihmisistä tunnistaa puun takaa näkyvän auton yhdeksi kokonaisuudeksi vaikka todellisuudessa itse havainto koostuu vain kahdesta puun eri puolilla näkyvästä auton puolikkaasta. Tätä vielä koneelle ongelmallista ihmisen ominaisuutta ovat käyttäneet hyväkseen esimerkiksi Mori, Uda ja Kikuchi (2012), Kunitani ja Uda (2013) sekä Sawada ja Uda (2016), joiden esittämät toteutukset esitellään seuraavassa tarkemmin.

Mori, Uda ja Kikuchi (2012) ovat kehittäneet videopohjaisen captcha toteutuksen, jossa käyttäjän tehtävänä on tunnistaa 10s kestävältä ja 30 kuvaa sekunnissa sisältävältä videolta yhteensä neljä numero- tai kirjainmerkkiä. Suurimman osan videon kestosta merkit ovat nähtävillä, mutta vääristettynä ja erikokoisten liikkuvien esteiden peittämänä, niin että niiden tunnistaminen on mahdotonta. Vuorollaan jokainen merkki näkyy yhden kuvan ajan vääristämättömänä, mutta edelleen osittain peitettynä, siten että ihmiselle sen tunnistaminen on

suhteellisen vaivatonta.

Tätä toteutusta ovat kehittäneet edelleen Kunitani ja Uda (2013). Heidän toteutuksessaan tunnistettavat merkit ovat samalla tavoin jatkuvasti nähtävillä, mutta kaikki merkit ovat kolmiulotteisia, esteiden peittämiä ja liikkuvat sekä pyörivät satunnaisesti. Videon kesto ja kuvien määrä sekunnissa ovat tässä samat kuin edellisessä.

Kyseisten captcha-toteutuksien murtamisen näkökulmasta osittain peitetyn kirjainmerkin täydentäminen ja tunnistaminen algoritmillisesti on mahdollista, mutta laskennallisesti vaativaa. Lisäksi kun kyseessä on video, joka sisältää yhteensä 300 yksittäistä kuvaa ($10s * 30fps$), ja muutoin kaksiulotteiseen merkkiin lisätään vielä kolmas ulottuvuus, tunnistamisen laskennallinen vaativuus kasvaa niin suureksi, ettei tehtävän ratkaiseminen raa'alla laskentateholla ole käytännössä kannattavaa. Edellä mainituilla tavoilla toteutettujen captcha toteutuksien tarkkaa laskennallista vaativuutta ei kuitenkaan vielä ole selvitetty, joten koneiden laskentatehon kasvun myötä nämäkin ratkaisut saattavat osoittautua toimimattomiksi.

Myös Sawada ja Uda (2016) ovat perustaneet toteutuksena samaan ideaan, jonka Mori, Uda ja Kikuchi (2012) esittelivät, mutta alkuperäisen toteutuksen yhteydessä esiteltyjen toimenpiteiden lisäksi he hyödyntävät ihmisen näköhavainnon jälkikuvaa. Heidänkin toteutuksessaan yksittäinen merkki on osittain peitetty tunnistamisen vaikeuttamiseksi, mutta jokainen merkki on jaettu 3×3 ruudukkoon, joista käyttäjä näkee aina vain yhden ruudun kerrallaan. Nämä kaikki 9 ruutua näytetään käyttäjälle kukin vuorollaan tietystä järjestyksessä yhden sekunnin aikana. Koska kuvat näytetään hyvin lyhyen aikavälin sisällä, niin näköaistin viiveen takia kuva hahmotetaan yhtenä kokonaisuutena. Tämän seurauksena ihminen kykenee tunnistamaan useassa osassa esitettävän osittain peitetyn merkin lähes yhtä helposti kuin silloin jos kokonaisuus näytettäisiin kokonaan yhdellä kertaa.

Murtamisen kannalta jälkikuvaan perustuva toteutus edellyttää koneelta useiden yksittäisten kuvien yhdistämistä yhdeksi analysoitavaksi kuvaksi. Tämän lisäyksen seurauksena tutkijoiden Sawada ja Uda (2016) esittelemä captcha-toteutus on vääjäämättä koneelle vaativampi kuin alkuperäinen toteutus. Koska alkuperäisenkään tarkka laskennallinen vaativuus ei ole tiedossa, tämänkään toteutuksen todellista vaativuutta ei tiedetä ja sen vanhentuminen tulevaisuudessa on sikäli mahdollista.

4.2 Syvyyšnäköön perustuva toteutus

Hahmotuskyvyn ja havainnon täydentämisen lisäksi ihmisellä on syvyyšnäöksi kutsuttu kyky tulkita ympäristöään kolmiulotteisena kaksiulotteisen havainnon perusteella. Osa tästä kyvystä perustuu fyysikaaliseen aistihavaintoon siitä kuinka eri etäisyydellä olevat kappaleet näkyvät kumpaankin silmään hieman eri kulmasta. Monissa tilanteissa tämä ero on kuitenkin niin pieni, että ihmisen täytyy käyttää syvyystvaikutelman saavuttamiseen muita syvyyshivihjeitä, kuten tuttujen kappaleiden koon intuitiivista arvioimista. Käytännössä tämä ei lopulta edellytä edes konkreettisen kappaleen näkemistä, vaan jo kappaleen representaation perusteella osaamme arvioida kuinka suuri todellinen kappale olisi suhteessa muihin tuntemiimme kappaleisiin (Nejati ym. 2014).

Tätä kykyä hyödyntää DeepCAPTCHA, jonka ovat kehittäneet Nejati ym. (2014). Tässä toteutuksessa käyttäjän tehtävänä on erottaa ja järjestää joukko arkipäiväisiä kappaleita esittäviä 3D-malleja suuruusjärjestykseen. Kappaleiden koneellisen tunnistamisen vaikeuttamiseksi kunkin kappaleen katselukulma, valaistus, väritys ja tausta satunnaistetaan joka kerta testiä generoitaessa. Näin pyritään varmistamaan testin toimivuus myös tilanteessa, jossa kappaleiden tietokanta pääsisi vuotamaan julki tai oppiva ohjelma pyrkisi opettelemaan kaikki kappaleet ulkoa. (Nejati ym. 2014)

Automaattinen kappaleen tunnistaminen on konenäön osa-alue, joka käyttää kahdenlaisia tunnistamiseen kahdenlaisia metodeja: väriperustaisia ja muotoperustaisia. Näistä metodeista tunnistamiseen voidaan käyttää molempia tai vain toista tilanteesta riippuen. Tämän vuoksi DeepCAPTCHA:ssa testiä generoitaessa pyritään tietoisesti vääristämään juuri näitä attribuutteja niin paljon kuin mahdollista kuitenkin tekemättä tunnistamisesta ihmiselle mahdollontta. Nejati ym. (2014) ovat myös testanneet toteutuksena turvallisuutta useilla tunnetuilla luokitteluohjelmistoilla, jolle syötettiin 60% testiä varten generoiduista malleista esimerkeinä. Koulutusvaiheen jälkeen luokittelijoille annettiin loput 40% malleista tunnistamista varten. Tutkimus osoitti, että ulkoasun vääristämisestä johtuen ohjelmat epäonnistuivat sekä malleja esittävien kappaleiden tunnistamisessa että niiden koon arvioimisessa. (Nejati ym. 2014)

4.3 Poikkeavuuden tunnistamiseen perustuva toteutus

Ihmisellä on luontainen kyky tunnistaa poikkeavuutta sellaista kohdatessaan. Arkisemmin tästä käytetään myös nimitystä ”outous” (eng. ”strangeness”). Tämä havaitseminen on pitkälti intuitiivista ja perustuu ihmisen olemassaolonsa aikana kohtaamiin tilanteisiin ja niiden kautta muodostuneisiin sisäisiin malleihin. Hyvä esimerkki tästä on oikeakielisuus ja luonnollisen (l. ihmisten käyttämän) kielen tunnistaminen. Suurin osa ihmisistä kykenee tunnistamaan jos jonkun muun tuottama kieli ei ole kielioppisääntöjen tai konventioiden mukaista, vaikka eivät välttämättä osaa nimetä mitään erityistä virheellistä piirrettä.

Yamamoto, Tygar ja Nishigaki (2010) ovat perustaneet Strangeness in Sentences (SS-CAPTCHA) nimisen captchansa juuri tähän luonnollisen kielen ja sen poikkeamien tunnistamiseen. Heidän toteutuksessaan käyttäjän tehtävänä on erottaa ihmisten tuottamat ”oikeat” lauseet ja ohjelman toisesta kielestä kääntämät ”keinotekoiset” lauseet toisistaan. Vaikka kääntäjäohjelmien tavoitteena on tuottaa kielellisesti mahdollisimman luonnollisia lauseita, ihmiset usein kokevat niissä olevan jotain outoa. Ohjelmat taas pitävät ohjelmallisesti käännettyä lausetta täysin luonnollisena, sillä muutoin jo lauseen kääntänyt ohjelma olisi tunnistanut tuottamansa lauseen virheelliseksi ja muokannut sitä paremmaksi. (Yamamoto, Tygar ja Nishigaki 2010)

Tällä hetkellä ainut tunnettu tapa, jolla kone voisi tämän captchan ratkaista, on lauseiden vertaileminen jo olemassa oleviin lauseisiin. Ensimmäinen vaihtoehto on kääntää annettuja lauseita uudestaan eri kielten välillä ja tutkia kuinka paljon lauseet muuttuvat suhteessa kääntökertojen lukumäärään. Periaatteessa jo kerran koneella käännettyjen lauseiden muuttumisen tulisi vähentyä nopeammin kuin alun perin kääntämättömien lauseiden, minkä perusteella kone voisi erottaa lauseet toisistaan. Toinen vaihtoehto on syöttää annettujen lauseiden osia hakukoneeseen ja verrata hakutulosten määriä siinä toivossa, että internetissä esiintyy suhteessa vähemmän koneellisesti käännettyjä kuin ihmisten tuottamia lauseita. Ensimmäinen hyökkäystapa voidaan kuitenkin ehkäistä suorittamalla sama analyysi jo tehtävää generoitaessa, jolloin mukaan valitaan vain lauseita joiden muutosnopeus suhteessa kääntökertojen lukumäärään on sama. Toinen tapa taas edellyttäisi koneelta käytännössä koko lauseen tutkimista pala palalta raa’alla voimalla, mikä tekisi hyökkäyksestä kannattamattoman. (Yamamoto, Tygar ja Nishigaki 2010)

4.4 Päätelykykyyn perustuva toteutus

Ihmisillä on luontainen taipumus ratkaista eteen tulevia kysymyksiä ja ongelmia. Tämän perusteella Saha, Nag ja Dasgupta (2015) ovat esitelleet kirjallisiin kysymyksiin perustuvan captchan. Vielä konseptin asteella olevan testin on määrä sisältää kysymyksiä, joihin vastaaminen edellyttää esimerkiksi matemaattista ajattelua tai luonnollisen kielen ymmärtämistä, ja joihin on olemassa yksikäsitteinen vastaus. Johonkin tiettyyn alaan keskittymisen sijaan kysymysten on määrä olla yleisluontoisia kolmesta syystä: kysymykset sopivat paremmin kaikille ihmisille taustaan katsomatta, mahdollisten kysymysten lukumäärää on huomattavasti suurempi ja vastauksien löytäminen hakukonetta käyttäen on vaikeampaa. Lisäksi vastausyriyksiä kutakin kysymystä kohti on rajattu. (Saha, Nag ja Dasgupta 2015)

Esimerkiksi testin esittämä kysymys voisi olla muotoa:

- Liisa ja Paavo ovat kävelyllä yhdessä. Liisalla on mukana omena ja Paavolla makkara. He alkavat syödä. Kumpi syö hedelmää?

(esimerkki Saha, Nag ja Dasgupta (2015))

Tämän ihmisen kognitioon ja päätelykykyyn perustuvan captchan murtamiseksi ohjelman tulisi ensinnäkin kyetä prosessoimaan kysymyksen esittämiseen käytetty luonnollinen kieli. Tämä tarkoittaa erisnimien, subjektien, verbien ja muiden sanaluokkien erottamista toisistaan ja itse kysymyksen löytämistä tekstistä. Tämän jälkeen ohjelman tulisi löytää esitettyyn kysymykseen oikea vastaus tehtävän sisältämiä tietoja, ulkoisia tietoja (ns. yleistietoa) tai molempia edellä mainittuja hyväksi käyttäen. Koneelle tämä tarkoittaisi kaikkien mahdollisten sanojen välisten suhteiden läpikäymistä, analysoimista ja vertailua kaikkiin muihin olemassa oleviin sanoihin. Käytännössä tehtävän murtaminen raa'alla voimalla on siten kannattamatonta vastauskertojen määrän ollessa rajattu. Kysymysten sisältämien aihepiirien laajuuden vuoksi on hyvin vaikeaa myöskään luoda tarpeeksi suurta tietokantaa, jonka avulla kone voisi oppia oikeat vastaukset. (Saha, Nag ja Dasgupta 2015).

4.5 Huumorintajuun perustuva toteutus

Teksti, puhe, kuvat, videot ja tarinat sisältävät sisäisiä merkityksiä, joita ei voida mitenkään johtaa vain kyseessä olevasta havainnosta. Tällaisia merkityksiä ovat esimerkiksi konteksti, korostukset ja huumori. Näistä erityisesti huumorin ymmärtämistä pidetään yhtenä ihmisen kehittyneimmistä kognitiivisista kyvyistä ja tähän perustuen tutkijat Yamamoto, Suzuki ja Nishigaki (2010) ovat esitelleet neliruutuisista sarjakuvista koostuvan captchan. Tehtävässä käyttäjälle esitetään neliruutuinen sarjakuva, jonka ruutujen järjestys on sekoitettu. Käyttäjän tehtävänä on laittaa ruudut oikeaan järjestykseen, jotta kuvista syntyy sarjakuvan tekijän alun perin tarkoittama jollain tavalla hauska tarina.

Koneelle sarjakuvan hahmottaminen tarkoittaa ensinnäkin ruutujen sisältämän luonnollisen kielen ja kuvien sisällön prosessoimista ja näiden yhdistämistä toisiinsa. Sarjakuviin liittyy näkyvissä olevien kuvien lisäksi paljon kontekstin tuntemista, ennakointia, hiljaista tietoa ja ihmisillekin usein vain intuitiivista ilmapiirin tuntemista. Kaikkien näiden seikkojen summana koneen on todella haastavaa raa'alla voimalla laskemalla tai muullakaan keinolla käsittää mikä kulloinkin muodostuvassa sarjakuvassa olisi ihmisen mielestä humoristista. (Yamamoto, Suzuki ja Nishigaki 2010) Yamamoto, Suzuki ja Nishigaki (2010) väittävätkin, että tekniikan kehittymisestä huolimatta ohjelmille tulee olemaan lähes mahdotonta saavuttaa tasoa, jolla ne kykenisivät ymmärtämään huumorin kaltaisia sisäisiä merkityksiä.

5 Johtopäätökset

Tässä kirjallisuuskatsauksena toteutetussa tutkimuksessa tutkittiin mitkä ihmisen ominaisuudet ovat vielä tietokoneen saavuttamattomissa. Tähän tutkimuskysymykseen etsittiin vastausta kartoittamalla automatisoidun Turingin testin, captchan, uusia toteutuksia ja selvittämällä mitä ihmisen ominaisuuksia ne hyödyntävät ihmisen ja koneen erottamiseksi.

Puhtaasti hahmotuskykyyn perustuvat vielä 2000-luvun alussa toimineet teksti-, ääni- ja kuvapohjaiset captcha toteutukset ovat tietokoneiden laskentatehon kasvun myötä osoittautuneet murretuiksi ja siten vanhentuneiksi. Myöhemmätkin vuoden 2010 jälkeen esitellyt epätäydellisen havainnon täydentämiseen perustuvat captchat toimivat ihmisen ja koneen erottamisessa tällä hetkellä ennen kaikkea tehtävien laskennallisen vaativuuden tähden. Odotettavissa siis on, että mikäli laitteistojen suorituskyky jatkaa kasvuaan kuten tähänkin saakka, kaikki inhimilliseen hahmotuskykyyn perustuvat captcha -toteutukset tulevat murtumaan ennen pitkää. Ihmisen hahmotuskyvyn saavuttaminen ja ohittaminen ei siis näyttäisi olevan koneelle mahdotonta.

Muiden esiteltyjen captcha -toteutuksien yhteydessä tutkijat mainitsevat kulloinkin kyseessä olevan kyvyn intuitiivisuuden. Intuitiivisilla kyvyillä näissä tutkimuksissa tarkoitetaan kykyjä, joita ei osata täsmällisesti analysoida. Tällaisia kykyjä ei siten ole myöskään mahdollista toteuttaa algoritmillisesti. Tämän tutkielman perusteella tällaisia kykyjä ovat ihmisten väliseen kommunikointiin käytettävän kielen erottaminen koneellisesti tuotetusta, kolmiulotteista hahmottamista tukevat syvyysnäön opitut piirteet, elämäkokemuksen kautta saavutettu päättelykyky sekä viimeisimpänä kyky tulkita epäsuoria sisäisiä merkityksiä kuten huumoria.

Tällä hetkellä ihmisen ominaisuuksien automatisoimisen rajoitteena vaikuttaa siis olevan oma ymmärryksemme ihmisten kykyjen toiminnasta sekä jossain määrin koneiden tapa ajatella. Mielenkiintoista onkin millä tavoin esimerkiksi neuroverkkoihin perustuvien tietokoneiden kehitys tulee muuttamaan tilannetta tietokoneiden eduksi. Uuden tyyppisten koneiden yleistymisen vaatii kuitenkin aikaa ja vielä tässä vaiheessa olisi mielenkiintoista tutkia kuinka tällä hetkellä olisi kannattavaa yhdistää ihmisen ja koneen ominaisuuksia käytännössä.

Lähteet

Ahn, Luis Von, Manuel Blum, Nicholas J. Hopper ja John Langford. 2003. "CAPTCHA: Using hard AI problems for security", 294–311. Springer.

Ahn, Luis Von, Manuel Blum ja John Langford. 2004. "Telling humans and computers apart automatically". *Communications of the ACM* 47 (2): 56–60.

Bigham, Jeffrey P., ja Anna C. Cavender. 2009. "Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use". Teoksessa *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1829–1838. ACM.

Bursztein, E., R. Beauxis, H. Paskov, D. Perito, C. Fabry ja J. Mitchell. 2011. "The Failure of Noise-Based Non-continuous Audio Captchas". Teoksessa *Security and Privacy (SP), 2011 IEEE Symposium on*, 19–31. ID: 1. ISBN: 1081-6011.

Bursztein, Elie, Matthieu Martin ja John Mitchell. 2011. "Text-based CAPTCHA Strengths and Weaknesses". Teoksessa *Proceedings of the 18th ACM Conference on Computer and Communications Security*, 125–138. CCS '11. Chicago, Illinois, USA: ACM. ISBN: 978-1-4503-0948-6. <http://doi.acm.org/10.1145/2046707.2046724>.

Golle, Philippe. 2008. "Machine Learning Attacks Against the Asirra CAPTCHA". Teoksessa *Proceedings of the 15th ACM Conference on Computer and Communications Security*, 535–542. CCS '08. Alexandria, Virginia, USA: ACM. ISBN: 978-1-59593-810-7. <http://doi.acm.org/10.1145/1455770.1455838>.

Kluever, Kurt Alfred, ja Richard Zanibbi. 2009. "Balancing usability and security in a video CAPTCHA". Teoksessa *Proceedings of the 5th Symposium on Usable Privacy and Security*, 14. ACM.

Kochanski, Greg, Daniel P. Lopresti ja Chilin Shih. 2002. "A reverse turing test using speech." Teoksessa *INTERSPEECH*.

- Kunitani, Kenta, ja Ryuya Uda. 2013. "Proposal of CAPTCHA Using Three Dimensional Objects". Teoksessa *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication*, 57:1–57:6. ICUIMC '13. Kota Kinabalu, Malaysia: ACM. ISBN: 978-1-4503-1958-4. <http://doi.acm.org/10.1145/2448556.2448613>.
- Levy, Yair, ja Timothy J. Ellis. 2006. "A systems approach to conduct an effective literature review in support of information systems research". *Informing Science: International Journal of an Emerging Transdiscipline* 9 (1): 181–212.
- Mori, T., R. Uda ja M. Kikuchi. 2012. "Proposal of Movie CAPTCHA Method Using Amodal Completion". Teoksessa *Applications and the Internet (SAINT), 2012 IEEE/IPSJ 12th International Symposium on*, 11–18. ID: 1.
- Nejati, Hossein, Ngai-Man Cheung, Ricardo Sosa ja Dawn C. I. Koh. 2014. "DeepCAPTCHA: An Image CAPTCHA Based on Depth Perception". Teoksessa *Proceedings of the 5th ACM Multimedia Systems Conference*, 81–90. MMSys '14. Singapore, Singapore: ACM. ISBN: 978-1-4503-2705-3. <http://doi.acm.org/10.1145/2557642.2557653>.
- Saha, S. K., A. K. Nag ja D. Dasgupta. 2015. "Human-Cognition-Based CAPTCHAs". ID: 1, *IT Professional* 17 (5): 42–48.
- Salminen, Ari. 2011. "Mikä kirjallisuuskatsaus". *Johdatus kirjallisuuskatsauksen tyypeihin ja hallintotieteellisiin sovelluksiin. Vaasan yliopiston julkaisuja. Opetusjulkaisuja* 62.
- Sawada, Kouta, ja Ryuya Uda. 2016. "Effective CAPTCHA with Amodal Completion and Aftereffects". Teoksessa *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*, 53:1–53:5. IMCOM '16. Danang, Viet Nam: ACM. ISBN: 978-1-4503-4142-4. <http://doi.acm.org/10.1145/2857546.2857600>.
- Turing, Alan M. 1950. "Computing machinery and intelligence". *Mind* 59 (236): 433–460.
- University, Carnegie Mellon. 2010. *The Official CAPTCHA site*. Saatavilla WWW-muodossa, <http://http://captcha.net/>, viitattu 9.4.2016.

Vicarious. 2013. *Vicarious AI passes first Turing Test: CAPTCHA - Vicarious - News*. Saatavilla WWW-muodossa, <http://news.vicarious.com/post/65316134613/vicarious-ai-passes-first-turing-test-captcha>, viitattu 9.4.2016.

Vikram, Shardul, Yinan Fan ja Guofei Gu. 2011. "SEMAGE: A New Image-based Two-factor CAPTCHA". Teoksessa *Proceedings of the 27th Annual Computer Security Applications Conference*, 237–246. ACSAC '11. Orlando, Florida, USA: ACM. ISBN: 978-1-4503-0672-0. <http://doi.acm.org/10.1145/2076732.2076766>.

Yamamoto, T., T. Suzuki ja M. Nishigaki. 2010. "A Proposal of Four-Panel Cartoon CAPTCHA: The Concept". Teoksessa *Network-Based Information Systems (NBIS), 2010 13th International Conference on*, 575–578. ID: 1. ISBN: 2157-0418.

Yamamoto, T., J. D. Tygar ja M. Nishigaki. 2010. "CAPTCHA Using Strangeness in Machine Translation". Teoksessa *Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on*, 430–437. ID: 1. ISBN: 1550-445X.