Kanditutkielma

Tietojenkäsittelytieteen kandiohjelma

Kontekstiriippuvaiset tiivistefunktiot haittaohjelmien torjuntamenetelmänä

Valtteri Varvikko

2.11.2022

Matemaattis-luonnontieteellinen tiedekunta

Helsingin yliopisto

Yhteystiedot

PL 68 (Pietari Kalmin katu 5)
00014 Helsingin yliopisto

Sähkopostiosoite: info@cs.helsinki.fi
URL: http://www.cs.helsinki.fi/

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET – UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Study programme
Faculty of Science	Batchelor's Programme in Computer Science

Tekijä — Författare — Author

Valtteri Varvikko

Työn nimi — Arbetets titel — Title

Kontekstiriippuvaiset tiivistefunktiot haittaohjelmien torjuntamenetelmänä

Ohjaajat — Handledare — Supervisors

prof. Nikolaj Tatti

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Bachelor's thesis	November 2, 2022	7 pages

Tiivistelmä — Referat — Abstract

Sint numquam id et nobis consectetur et (Kornblum, 2006). Optio accusamus maiores vel (Naik et al., 2019). Illum id sed asperiores aliquid explicabo necessitatibus et. Incidunt voluptatum rerum ut aut officiis ut fugiat. Ex amet corrupti est. Repellat veniam iusto qui. Non quaerat mollitia iure velit corporis recusandae vel nihil. Neque et aut dignissimos nemo soluta illum consequatur.

Et minus doloremque harum tenetur. Nemo architecto aliquam atque voluptatem facere praesentium temporibus. Quo non dolorem natus nihil et rem. Et fugit tempora nulla quia velit non sit. Qui enim sint id veniam dolor. Et voluptas eligendi impedit adipisci aut unde nisi. Sint exercitationem id assumenda sed fugiat quis. Eum perferendis rerum quo et deserunt qui.

Beatae nemo facere doloribus modi voluptas veritatis perspiciatis eius. Quibusdam et doloremque qui et accusantium corporis. Architecto tenetur harum architecto eveniet enim. Perferendis corporis dolores qui veniam ut quasi et neque. Aut ut aspernatur consectetur omnis consequatur nesciunt corporis. Non pariatur amet dignissimos sed quidem. Consectetur et cum dolor. Tempore et optio autem ipsam est cupiditate maiores.

ACM Computing Classification System (CCS)
Security and privacy → Cryptography → Symmetric cryptography and hash functions → Hash functions and message authentication codes
Security and privacy → Intrusion/anomaly detection and malware mitigation → Malware and its mitigation

Avainsanat — Nyckelord — Keywords

algorithms, hash functions

Säilytyspaikka — Förvaringsställe — Where deposited

Helsinki University Library

Muita tietoja — övriga uppgifter — Additional information

Sisällys

1 Johdanto

Tietojenkäsittelyssä käsitellään valtavia määriä dataa.

Tiedon esittämislle kompaktimmassa muodossa on laajalti tarvetta, milloin tallennuskapasiteetin säästämisen, milloin datan tehokkaan tunnistamiseen vuoksi.

Laskentakyvyn kasvaessa myös tietoliikenne ja tallennuskapasiteetti kehittyvät.

Rikolliset tahot hyödyntäväþ tietojenkäsittelyä [MISSÄ].

Tietojenkäsittelyn kehitys luo uusia, rikollisia intressejä tyydyttäviä mahdollisuuksia.

Kasvava datavirta helpottaa haittaohjelmien huomaamatonta leviämistä.

Sekä tekninen rikostutkinta että kyberrikollisuus ovat riippuvaisia tietotekniikan kehityksestä. Innovaatiot luovat mahdollisuuksia kehittää rikollisia menetelmiä sekä toisaalta ratkaisuja haitallista toimintaa vastaan.

Tiivistystä voidaankin hyödyntää keinona haittaohjelmien havaitsemisessa suuresta datamäärästä.

2 Tiivisteet

2.1 Tiivisteet haittaohjelmien torjunnassa

H	e	i		m	a	a	i	l	m	a	!
7b			7f			e5			53		

Taulukko 2.1: Paloittain määritelty hajautusarvo merkkijonolle.

h	e	i		m	a	a	i	l	m	a	?
38			7f			e5			09		

Taulukko 2.2: Palottain määritelty hajautusarvo merkkijonolle, jonka merkkejä on vaihdettu.

Minus optio similique similique voluptatem placeat accusantium impedit voluptates. Consequatur fugiat soluta cupiditate consequuntur quis fugiat. Beatae placeat soluta dolores consequatur molestias in fuga. Ut et accusantium et atque illo. Aut doloribus labore corporis magnam est natus quaerat deserunt.

H	e	i	,		m	a	a	i	l	m	a	!
7b			e9			54			03			e7

Taulukko 2.3: Palottain määritelty hajautusarvo merkkijonolle, johon on lisätty uusi merkki.

2.2 Samankaltaiset tiedostot

2.3 Paikallisherkkyys

3 Kontekstiriippuvaiset tiivistefunktiot

3.1 Paloittain määritely tiiviste

3.2 Kontekstiriippuvaisuus

3.3 Toteutuksista

3.3.1 ssdeep

3.3.2 SDHASH

4 Kontekstiriippuvaisten tiivisteiden soveltaminen haittaohjelmien torjunnassa

4.1 Edellytykset

4.2 Tulosten hyödyntäminen

4.3 Skaalautuvuus

5 Haavoittuvuudet ja hyökkäykset

5.1 Harhaanjohtaminen dataa manipuloimalla

5.1.1 Aiheettomat osumat

5.1.2 Naamiointi osumien välttämiseksi

6 Yhteenveto

Lähteet

Kornblum, J. (2006). ”Identifying almost identical files using context triggered piecewise hashing”. *Digital Investigation* 3. The Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS ’06), s. 91–97. ISSN: 1742-2876. DOI: https://doi.org/10.1016/j.diin.2006.06.015. URL: https://www.sciencedirect.com/science/article/pii/S1742287606000764.

Naik, N., Jenkins, P. ja Savage, N. (2019). ”A Ransomware Detection Method Using Fuzzy Hashing for Mitigating the Risk of Occlusion of Information Systems”. Teoksessa: *2019 International Symposium on Systems Engineering (ISSE)*, s. 1–6. DOI: 10.1109/ISSE46696.2019.8984540.