BİLDİRİLER KİTABI

PROCEEDINGS BOOK

***Editors*/Editörler**

Prof. Dr. Şeref Sağıroğlu
Prof. Dr. Mustafa Alkan
Doç. Dr. Sedat Akleylek

ISBN: **978-605-86904-8-6**

ISC TURKEY 2018

11. ULUSLARARASI
BİLGİ GÜVENLİĞİ
ve **KRİPTOLOJİ**
KONFERANSI

11th INTERNATIONAL CONFERENCE
ON INFORMATION
SECURITY &
CRYPTOLOGY

17 - 18 Ekim - *October* 2018• ANKARA BTK MERKEZ BİNASI • *ICTA HEADQUARTER*

BİLGİGÜVENLİĞİ
DERNEĞİ

PROCEEDINGS OF
11th INTERNATIONAL CONFERENCE ON INFORMATION SECURITY AND CRYPTOLOGY (ISCTURKEY 2018)

11. ULUSLARARASI BİLGİ GÜVENLİĞİ VE KRİPTOLOJİ KONFERANSI
BİLDİRİ KİTABI

ISC TURKEY 2018
BİLDİRİLER KİTABI
PROCEEDINGS

17-18 Ekim / ***October 2018***
ANKARA BTK MERKEZ BİNASI / ANKARA ICTA HEADQUARTER

Ankara, TÜRKİYE / *TURKEY*

***Editors*/Editörler**
Prof. Dr. Şeref Sağıroğlu
Prof. Dr. Mustafa Alkan
Doç. Dr. Sedat Akleylek

www.iscturkey.org
ISBN 978-605-86904-8-6

ABOUT / HAKKINDA

This book compromises the proceedings of ISCTURKEY 2018. The articles in the proceedings reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by ISCTURKEY 2018 Organising Committee.

Bu bildiriler kitabında yer alan bildirin tam metinleri konferans konu başlıklarına uygun olarak yazarlar tarafından hazırlanmıştır. Bildiri özetleri yazarların kendi fikirlerini yansıtır ve herhangi bir değişiklik yapılmadan aynı şekilde basılmıştır. Bu bildiri kitabında yayımlanan görüşler yazarlara ait olup bu görüşlerinden ISCTURKEY 2018 Düzenleme Kurulu sorumlu tutulamaz.

Contact to / İrtibat:

Bilgi Güvenliği Derneği

Adres : Maltepe Mahallesi Tuncer Sok. No.4/8 - Çankaya 06570 - Ankara - Türkiye
Tel : +90 312 231 1810
Fax : +90 312 231 1810
Eposta : bilgi@bilgiguvenligi.org.tr

ORGANISERS / ORGANİZASYON

BİLGİ GÜVENLİĞİ DERNEĞİ

GAZİ ÜNİVERSİTESİ 1926

İTÜ

İSTANBUL TEKNİK ÜNİVERSİTESİ 1773

ODTÜ

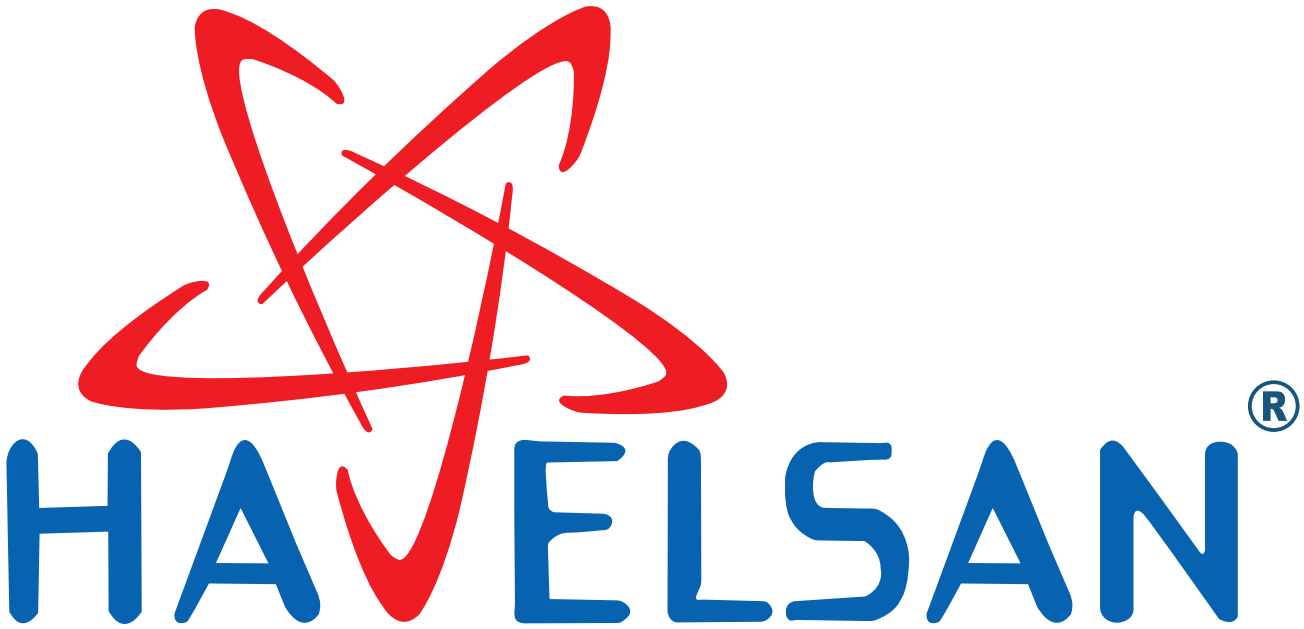
T.C.
Ulaştırma ve Altyapı
Bakanlığı

BİLGİ TEKNOLOJİLERİ VE İLETİŞİM KURUMU

SPONSORS / SPONSORLAR

THANKS TO / TEŞEKKÜR EDERİZ…

MAIN SPONSOR
ANA SPONSOR



SPONSORS / SPONSORLAR

THANKS TO / TEŞEKKÜR EDERİZ…

PLATININUM SPONSORS
PLATİN SPONSORLAR

CHOMAR

HUAWEI

GOLD SPONSOR
ALTIN SPONSOR

NETAS

SILVER SPONSORS
GÜMÜŞ SPONSORLAR

FORTINET

STM

Türk Telekom

vodafone

Verify

aselsan

TURKCELL

ARISTA

HIKVISION

SPONSORS / SPONSORLAR

THANKS TO / TEŞEKKÜR EDERİZ…

PANEL SPONSORS
PANEL SPONSORLARI

TÜRKSAT atarlabs BİZNET BİLİŞİM

STAGE SPONSOR
SAHNE SPONSORU

ayesaş

WEB SITE SPONSOR
WEB SİTESİ SPONSORU

Fonksiyon Yazılım

STANDS SPONSORS
STAND SPONSORLARI

Labris NETWORKS · eset ENJOY SAFER TECHNOLOGY™ · net iletişim · atarlabs

BİZNET BİLİŞİM · BARİKAT Bilişim Güvenliği · METRİC · epati BİLİŞİM TEKNOLOJİLERİ

BACKDROP SPONSOR
SAHNE SPONSORU

ayesaş · HUAWEI

SPONSORS / SPONSORLAR

THANKS TO / TEŞEKKÜR EDERİZ…

NOTEBOOK AND PEN SPONSOR
NOT DEFTERİ VE KALEM SPONSORU

turknet

NAME BADGE SPONSOR
YAKA KARTI SPONSORU

TURKTRUST

MEDIA SPONSORS
MEDYA SPONSORLARI

CyberMag
TÜRKİYE'NİN İLK VE TEK SİBER GÜVENLİK DERGİSİ

faselis

EVENT MANAGEMENT SPONSOR
ETKİNLİK YÖNETİMİ SPONSORU

sencron

COMMITTEES / KURULLAR

Onur Kurulu / *Honory Committee*

Dr. Ömer Fatih **SAYAN**, *BTK Başkanı*

Prof. Dr. Mehmet **KARACA**, *Ulaştırma ve Altyapı Bakanlığı, Bakan Yardımcısı*

Prof. Dr. İbrahim **USLAN**, *Gazi Üniversitesi Rektörü*

Prof. Dr. Mustafa Versan **KÖK**, *ODTÜ Rektörü*

Düzenleme Kurulu / *Organising Committee*

Mustafa **ALKAN**, Eş Başkan/ CoChair, *Gazi Üniversitesi /Gazi University*

Ferruh **ÖZBUDAK**, Eş Başkan*/ CoChair,* Orta Doğu Teknik Üniversitesi*/Middle East Technical University*

Ertuğrul **KARAÇUHA**, Eş Başkanı*/CoChair,* İstanbul Teknik Üniversitesi*/İstanbul Technical University*

Şeref **SAĞIROĞLU**, Eş Başkan/CoChair, *Bilgi Güvenliği Derneği /Information Security Association of Turkey*

Ali **YAZICI**, *Bilgi Güvenliği Derneği/Information Security Association*

Abdullah Raşit **GÜLHAN**, *Bilgi Güvenliği Derneği/Information Security Association*

Burhanettin **AL**, *Bilgi Güvenliği Derneği/Information Security Association*

Mehmet **GÜLŞEN**, *Bilgi Güvenliği Derneği/Information Security Association*

Mehmet Ali **İNCEEFE**, *Bilgi Güvenliği Derneği/Information Security Association*

Mustafa **ÜNVER**, *Bilgi Güvenliği Derneği/Information Security Association*

Bilgehan **ARSLAN**, *Gazi Üniversitesi/Gazi University*

Duygu **SİNANÇ**, *Gazi Üniversitesi/Gazi University*

Enver **ÖZDEMIR**, *İstanbul Teknik Üniversitesi/İstanbul Technical University*

Fatih **DEMİRHAN**, *Orta Doğu Teknik Üniversitesi/Middle East Technical University*

Lütfiye **ATA DURAK**, *İstanbul Teknik Üniversitesi/İstanbul Technical University*

Merve Sedef **GÜNDÜZ**, *Gazi Üniversitesi/Gazi University*

Murat **CENK**, *Orta Doğu Teknik Üniversitesi/Middle East Technical University*

Oğuzhan **KÜLEKÇİ**, *İstanbul Teknik Üniversitesi / İstanbul Technical University*

Ramazan **TERZİ**, *Gazi Üniversitesi/Gazi University*

Sebahattin **EKER**, *İstanbul Teknik Üniversitesi / İstanbul Technical University*

Sedat **AKLEYLEK**, *Ondokuz Mayıs Üniversitesi/Ondokuz Mayis University*

Yavuz **CANBAY**, *Gazi Üniversitesi/Gazi University*

Bilim Kurulu / *Program Committee*

A. Naci **ÜNAL**, *Bahçeşehir Üniversitesi, Bahçeşehir University*

A. Nurdan **SARAN**, *Çankaya Üniversitesi/Çankaya University*

Ahmet **KOLTUKSUZ**, *Yaşar Üniversitesi/Yaşar University*

Ahmet **ÖZMEN**, *Sakarya Üniversitesi/Sakarya University*

Akın **MARSAP**, *Aydın Üniversitesi/Aydın University*

Albert **LEVI**, *Sabancı Üniversitesi/Sabancı University*

Ali Aydın **SELÇUK**, *TOBB ETÜ/TOBB University of Economics and Technology*

Ali **DOĞANAKSOY**, *Orta Doğu Teknik Üniversitesi/METU*

Ali **İNAN**, *Adana Bilim ve Teknoloji Üniversitesi*

Ali **ŞENTÜRK**, *Mersin Üniversitesi/Mersin University*

Ali **YAZICI**, *Atılım Üniversitesi/Atılım University*

Ali Ziya **ALKAR**, *Hacettepe Üniversitesi/Hacettepe University*

Alisher **KHOLMATOV**, *Sabancı Üniversitesi/Sabancı University*

Alok **TONGAONKAR**, *Symantec*

Alper **ÖZBİLEN**, *Bilgi Güvenliği Derneği/Information Security Association of Turkey*

Alper **UĞUR**, *Pamukkale Üniversitesi/Pamukkale University*

Alptekin **KÜPCÜ**, *Koç Üniversitesi/Koc University*

Ammar **DAŞKIN**, *İstanbul Medeniyet Üniversitesi/ İstanbul Medeniyet University*

Asaf **VAROL**, *Fırat Üniversitesi, Fırat University*

Atila **BOSTAN**, *Atılım Üniversitesi/Atılım University*

Atilla **ELÇİ**, *Aksaray Üniversitesi/Aksaray University*

Atilla **ÖZGİT**, *Orta Doğu Teknik Üniversitesi/METU*

Aydın **ALATAN**, *Orta Doğu Teknik Üniversitesi/METU*

Ayşe **BAŞAR BENER**, *Boğaziçi Üniversitesi/Boğaziçi University*

Barış Bülent **KIRLAR**, *Süleyman Demirel Üniversitesi/Süleyman Demirel University*

Bedri **ÖZER**, *Fırat Üniversitesi/Fırat University*

Berkant **USTAOĞLU**, *İzmir Teknoloji Enstitüsü/İzmir Institute of Technology*

Berna **ORS YALÇIN**, *İstanbul Teknik Üniversitesi/İstanbul Technical University*

Berrin **YANIKOĞLU**, *Sabancı Üniversitesi/Sabancı University*

Berry **SCHOENMAKERS**, *Eindhoven University of Technology*

Bimal **ROY**, *Indian Statistical Institute*

Bülent **ÖRENCİK**, *İstanbul Teknik Üniversitesi /İstanbul Technical University*

Bülent **TUĞRUL**, *Ankara Üniversitesi/Ankara University*

Cebrail **ÇİFTLİKLİ**, *Erciyes Üniversitesi/Erciyes University*

Cevat **SENER**, *Orta Doğu Teknik Üniversitesi/METU*

Cihangir **TEZCAN**, *Ortadoğu Teknik Üniversitesi/METU*

Cüneyt **BAZLAMAÇCI**, *Orta Doğu Teknik Üniversitesi/METU*

Çağdaş **ÇALIK**, *National Institute of Standards*

Debasis **GIRI**, *Haldia Institute of Technology*

Deniz **TAŞKIN**, *Trakya Üniversitesi/Trakya University*

Derviş **KARABOĞA**, *Erciyes Üniversitesi/Erciyes University*

Ecir Uğur **KÜÇÜKSİLLE**, *Süleyman Demirel Üniversitesi/Süleyman Demirel University*

Eiji **OKAMOTO**, *University of Tsukuba*

Elif **SAYGI**, *Hacettepe Üniversitesi/Hacettepe University*

Emin **ANARIM**, *Boğaziçi Üniversitesi/Boğaziçi University*

Emin İslam **TATLI**, *İstanbul Medipol Üniversitesi/İstanbul Medipol University*

Emir **DİRİK**, *Uludağ Üniversitesi/Uludağ University*

Emrah **ÇAKÇAK**, *Orta Doğu Teknik Üniversitesi/METU*

Engin **AVCI**, *Fırat Üniversitesi/Fırat University*

Engin **KIRDA**, *ISECLAB*

Enis **KARAARSLAN**, *Muğla Üniversitesi/Muğla University*

Ercan **BULUŞ**, *Namık Kemal Üniversitesi/Namık Kemal University*

Erdal **IRMAK**, *Gazi Üniversitesi/Gazi University*

Erdem **ALKIM**, *Ondokuz Üniversitesi/Ondokuz Mayıs University*

Erdoğan **DOĞDU**, *TOBB Ekonomi ve Teknoloji Üniversitesi/TOBB University of Economics and Technology*

Erkan **AFACAN**, *Gazi Üniversites/Gazi University*

Erkan **BEŞDOK**, *Erciyes Üniversitesi/Erciyes University*

Erkay **SAVAŞ**, *Sabancı Üniversitesi/Sabancı University*

Ersan **AKYILDIZ**, *Orta Doğu Teknik Üniversitesi/METU*

Ersin **ELBAŞI**, *TÜBİTAK/The Scientific and Technological Reseach Council of Turkey*

Esra **YOLAÇAN**, *Osmangazi Üniversitesi/Osmangazi University*

Eşref **ADALI**, *İstanbul Teknik Üniversitesi/ITU*

Ertan **ONUR**, *Orta Doğu Teknik Üniversitesi/METU*

Eyüp Burak **CEYHAN**, *Bartın Üniversitesi/Bartın University*

Faruk **GÖLOĞLU**, *ESAT-COSIC*

Fatih **SULAK**, Atılım Üniversitesi/*Atılım University*

Fatma **BÜYÜKSARAÇOĞLU SAKALLI**, *Trakya Üniversitesi/Trakya University*

Fatoş **YARMAN VURAL**, *Orta Doğu Teknik Üniversitesi/METU*

Ferruh **ÖZBUDAK**, *Orta Doğu Teknik Üniversitesi/METU*

Gökay **SALDAMLI**, *Boğaziçi Üniversitesi/Boğaziçi University*

Gökhan **DALKILIÇ**, *Dokuz Eylül Üniversitesi/Dokuz Eylül University*

Guangzhi **QU**, *Oakland University*

Hacer K**ARACAN**, *Gazi Üniversitesi/Gazi University*

Hakan **TEKEDERE**, *Gazi Üniversitesi/Gazi University*

Halil İbrahim **BÜLBÜL**, *Gazi Üniversitesi/Gazi University*

Hamdi Murat **YILDIRIM**, *Bilkent Üniversitesi/Bilkent University*

Harold **BAIER**, *TU DARMSTADT*

Hayri **SEVER**, *Hacettepe Üniversitesi/Hacettepe University*

Hidayet **TAKÇI**, *Cumhuriyet Üniversitesi/Cumhuriyet University*

Hüseyin **DEMİRCİ**, *TÜBİTAK/The Scientific and Technological Reseach Council of Turkey*

Hüseyin **HIŞIL**, *Yaşar Üniversitesi/Yaşar University*

Hüsrev Taha **SENCAR**, *TOBB ETÜ/TOBB University of Economics and Technology*

Ion **TUTANESCU**, *University of Pitesti*

İbrahim Alper **DOĞRU**, *Gazi Üniversitesi/Gazi University*

İbrahim **SOĞUKPINAR**, *Gebze Yüksek Teknoloji Enstitüsü/Gebze Institute of Technology*

İlhami **ÇOLAK**, *Nişantaşı Üniversitesi/Nisantasi University*

İlkay **ULUSOY**, *Orta Doğu Teknik Üniversitesi/METU*

İsmail **GÜLOĞLU**, *Doğuş Üniversitesi/Doğuş University*

İsmail **SAN**, *Anadolu Üniversitesi/Anadolu University*

İzzet Gökhan **ÖZBİLGİN**, *HAVELSAN Akademi Direktörü - Gazi Üniversitesi*

Jianying **ZHOU**, *ASTAR Institute for Infocomm Research*

John A. **CLARK**, *University of York*

Jongsub **MOON**, *Korea University*

Jorge **NAKAHARA**, *Universite' Libre de Bruxelles (ULB), Belgium*

Kasım **ÖZTOPRAK**, *KTO Karatay Üniversitesi/Karatay University*

Katerina **MITROKOTSA**, *Delft University of Technology*

Kemal **BIÇAKCI**, *TOBB Ekonomi ve Teknoloji Üniversitesi/TOBB University of Economics and Technology*

Kerem **KAŞKALOĞLU**, *Özyeğin Üniversitesi/Özyeğin University*

Kıvanç **DİNÇER**, *Hacettepe Üniversitesi/Hacettepe University*

Kıvanç **MIHÇAK**, *Boğaziçi Üniversitesi/Boğaziçi University*

Koray **KARABINA**, *Florida Atlantic University*

Leyla **BERBER**, *Bilgi Üniversitesi/Bilgi University*

Mehmet **AKTAŞ**, *TÜBİTAK Bilgem, BTE/The Scientific and Technological Reseach Council of Turkey*

Mehmet **DEMİRCİ**, *Gazi Üniversitesi/Gazi University*

Mehmet **KİRAZ**, *TÜBİTAK-UEKAE/The Scientific and Technological Reseach Council of Turkey*

Mehmet **TEKEREK**, *KSU Üniversitesi/KSU University*

Mehmet Emin **DALKILIÇ**, *Ege Üniversitesi/Ege University*

Mehmet Ufuk **ÇAĞLAYAN**, *Yaşar Üniversitesi/Yaşar University*

Melek D. **YÜCEL**, *Orta Doğu Teknik Üniversitesi/METU*

Melissa **DANFORD**, *California State University*

Meltem **SÖNMEZ TURAN**, *National Institute of Standards and Technology (NIST)*

Mert **ÖZARAR**, *Konya Gıda ve Tarım Üniversitesi*

Mine **AKKAN**, *9 Eylül Üniversitesi/9 Eylül University*

Muhammet Ali **AYDIN**, *İstanbul Üniversitesi/İstanbul University*

Muhammet **ÜNAL**, *Gazi Üniversitesi/Gazi University*

Muhiddin **UĞUZ**, *Ortadoğu Teknik Üniversitesi/METU*

Murat **AK**, *Akdeniz Üniversitesi/Akdeniz University*

Murat **AŞKAR**, *İzmir Ekonomi Üniversitesi/İzmir University of Economics*

Murat **AYDOS**, *Hacettepe Üniversitesi/Hacettepe University*

Murat **CENK**, *Orta Doğu Teknik Üniversitesi/METU*

Murat **KARAKAYA**, *Atılım Üniversitesi/ Atılım University*

Mustafa **ALKAN**, *Gazi Üniversitesi/Gazi University*

Nazife **BAYKAL**, *Orta Doğu Teknik Üniversitesi/METU*

Oğuz **YAYLA**, *Hacettepe Üniversitesi/Hacettepe University*

Orhun **KARA**, *TÜBİTAK-UEKAE/The Scientific and Technological Reseach Council of Turkey*

Osmanbey **UZUNKOL**, *FernUniversität in Hagen, Germany*

Ömer Faruk **BAY**, *Gazi Üniversitesi/Gazi University*

Özgür **AKAN**, *Orta Doğu Teknik Üniversitesi/METU*

Peter **COOPER**, *Sam Houston State University*

Qinghan **XIAO**, *Defence Research and Development Canada*

Resul **DAŞ**, *Fırat Üniversitesi/Fırat University*

Sedat **AKLEYLEK**, *Ondokuz Mayıs Üniversitesi/Ondokuz Mayis University*

Selçuk **BAKTIR**, *Bahçeşehir Üniversitesi/Bahçeşehir University*

Selçuk **KAVUT**, *Balıkesir Üniversitesi/Balıkesir University*

Selim **KINACI**, *SSMDB, EGM/Turkish National Police*

Serap **ŞAHİN**, *İYTE/Izmir Institute of Technology*

Serdar **BOZTAŞ**, *RMIT Üniversitesi/RMIT University*

Serdar Süer **ERDEM**, *GYTE/Gebze Institute of Technology*

Sevil **ŞEN**, *Hacettepe Üniversitesi/Hacettepe University*

Shahram **RAHIMI**, *Southern Illinois University*

Suat **ÖZDEMİR**, *Gazi Üniversitesi/Gazi University*

Subhamoy **MAITRA**, *Indian Statistical Institute*

Süleyman **ÖZARSLAN**, *Orta Doğu Teknik Üniversitesi/METU*

Şaban **EREN**, *Maltepe Üniversitesi/Maltepe University*

Şeref **SAĞIROĞLU**, *Gazi Üniversitesi/Gazi University*

Taner **ALTUNOK**, *Türk Hava Kurumu Üniversitesi / Turkish Aviation Association University*

Tarık **YERLİKAYA**, *Trakya Üniversitesi/Trakya University*

Tekin **MEMİŞ**, *Kadir Has Üniversitesi/Kadir Has University*

Tolga **MATARACIOĞLU**, *Tübitak Bilgem Siber Güvenlik Enstitüsü*

Tolga **SAKALLI**, *Trakya Üniversitesi/Trakya University*

Tolga **YALÇIN**, *Konya Gıda ve Tarım Üniversitesi*

Tuğba **TAŞKAYA TEMİZEL**, *Ortadoğu Teknik Üniversitesi/METU*

Tuğkan **TUĞLULAR**, *İzmir Yüksek Teknoloji Enstitüsü/Izmir Institute of Technology*

Tuğrul **YANIK**, *Celal Bayar Üniversitesi/Celal Bayar University*

Türksel Kaya **BENSGHİR**, *TODAİE*

Umut **ULUDAĞ**, *TÜBİTAK UEKAE/The Scientific and Technological Reseach Council of Turkey*

Vasif **NABİYEV**, *Karadeniz Teknik Üniversitesi/Karadeniz Technical University*

Veysel **ASLANTAŞ**, *Erciyes Üniversitesi/Erciyes University*

Volkan **ATALAY**, *Orta Doğu Teknik Üniversitesi/METU*

Yadigar **IMAMVERDIYEV**, *Institute of Information Technology, Azerbaijan National Academy of Sciences*

Yurdahan **GÜLER**, *Ortadoğu Teknik Üniversitesi/METU*

Yusuf **İPEKOĞLU**, *Orta Doğu Teknik Üniversitesi/METU*

Yusuf Murat **ERTEN**, *İzmir Yüksek Teknoloji Enstitüsü/Izmir Institute of Technology*

Yücel **SAYGIN**, *Sabancı Üniversitesi/Sabancı University*

Ziya **AKTAŞ**, *Çankaya Üniversitesi/Çankaya University*

Zülfükar **SAYGI**, *TOBB ETÜ/TOBB University of Economics and Technology*

Danışma Kurulu / *Advisory Board*

A. Neşe **SAYARI**, *Biznet*

Abdullah Raşit **GÜLHAN**, *SİNERJİTÜRK*

Ahmet Hamdi **ATALAY**, *HAVELSAN*

Ali **YAZICI**, *ASELSAN*

Batuhan **TOSUN**, *ISSA Türkiye*

Bilal **ÖNAL**, *BGD*

Burak **ÇİFTER**, *BOA TEKNOLOJİ*

Burhanettin **AL**, *Turkcell*

Cem **AKOYMAK**, *TÜRK TELEKOM*

Cemal **AKYEL**, *Akyel Online*

Doğan Ufuk **GÜNEŞ**, *YASAD*

Emine Yazıcı **ALTINTAŞ**, *UDHB*

Faruk **ECZACIBAŞI**, *TBV*

Ferhat **YEŞİLLİ**, *BİH Grup*

Füsun Sarp **NEBİL**, *TİD*

Gökhan **ÖZBİLGİN**, *HAVELSAN*

Gürcan **GÜRSÜ**, *ALBERK QA TECHNIC*

Hanzade **SARIÇİÇEK**, *ODTÜ Teknokent*

Huzeyfe **ÖNAL**, *BGA*

İlker **TABAK**, *TBD*

Kadriye Yıldız **BARLAS**, *BGD*

Kemal **CILIZ**, *TÜBİSAD*

Mehmet Ali **İNCEEFE**, *BGD*

Mesut **DEMİRBİLEK**, *Vodafone*

Metin **TARAKÇI**, *ÇMD*

Muhterem **İLHAN**, *Vodafone*

Mustafa **AKGÜL**, *İNETD*

Mustafa **MACAR**, *BGD*

Mustafa **YANARTAŞ**, *TÜBİFED*

Nahit **GÖK**, *SABİDER*

Orhan **TURAN**, *BGD*

Selim **ÜLKÜ**, *BGD*

Öner **DEMİRKOL**, *TürkTrust*

TOPICS / KONULAR

Siber Güvenlik

- Kurumsal Sistem Güvenliği
- Dağıtık ve Yaygın Sistem Güvenliği
- Donanım Tabanlı Güvenlik
- Olay İşleme ve Penetrasyon Testi
- Yasal Sorunlar
- Multimedya ve Belge Güvenliği
- İşletim Sistemleri ve Veritabanı Güvenliği
- Gizlilik sorunları
- SCADA ve Gömülü Sistem Güvenliği
- Güvenli Yazılım Geliştirme
- Bulut Bilişim Güvenliği
- Büyük Veri Güvenliği
- Sosyal Ağlarda Güvenlik
- Web Tabanlı Uygulamalar ve Hizmetlerin Güvenliği
- Güvenlik Protokolleri
- VOIP, Kablosuz ve Telekomünikasyon Ağ Güvenliği

Dijital Adli Bilişim

- Siber Suçlar
- Karşı-Adli Bilişim ve Karşı-Karşı Adli Bilişim Teknikleri
- Veri sızıntısı ve Veri Koruma
- Veritabanında Adli Bilişim
- İçerik Filtreleme
- Dosya Sistemi ve Bellek Analizi
- Sanal ve Bulut Ortamlarında Adli Tıp
- Bilgi Gizleme
- Multimedia Adli Bilişimi
- İçeriden Saldırıların İncelenmesi
- Büyük Ölçekli Araştırmalar
- Malware Adli Bilişimi ve Anti-Malware Teknikleri
- Ağ Adli Bilişimi ve Trafik Analizi
- Donanım Hassasiyeti ve Cihazların Adli Bilişimi
- Yeni Tehditler ve Geleneksel Olmayan Yaklaşımlar

Bilgi Güvencesi ve Güvenlik Yönetimi

- İş Sürekliliği ve Felaket Kurtarma Planlaması
- Kurumsal Yönetim
- Kritik Altyapı Koruma
- Dijital Haklar Yönetimi ve Fikri Mülkiyet Koruması
- Güvenlik Ekonomisi
- Dolandırıcılık Yönetimi
- Kimlik Yönetimi
- Kanun ve Yönetmelikler
- Güvenlik Politikaları ve Güven Yönetimi
- Tehditler, Güvenlik Açıkları ve Risk Yönetimi

Siber Savaş ve Fiziki Güvenlik

- Gözetleme Sistemleri
- Biyometri Uygulamaları
- Siber Savaş Eğilimleri ve Yaklaşımlar
- Elektronik Pasaportlar, Ulusal Kimlik ve Akıllı Kart Güvenliği
- Sosyal Mühendislik
- Kimlik ve Erişim Kontrol Sistemleri
- Biyometri Standartları
- Yeni Teori ve Algoritmalar

IoT Destekli Teknolojiler

- 5G Ağlar ve IoT
- Yazılım Tanımlı Ağ (SDN) ve IoT
- Sensör ve Aktüatör Ağları
- Ultra-düşük güç IoT Teknolojileri ve Gömülü Sistem Mimarileri
- Giyilebilir Cihazlar, Vücut Algılayıcı Ağlar, Akıllı Taşınabilir Aygıtlar
- IoT Cihazlar ve Sistemleri için Tasarım Uzayı Keşif Teknikleri
- Heterojen Ağlar
- IoT Protokolleri (IPv6, 6LoWPAN, RPL, 6TiSCH, W3C)
- Iot için Adlı Veri Ağı (NDN)
- Nano Şeylerin Interneti
- Sensör Veri Yönetimi, IoT Madenciliği ve Analitiği
- Adaptif Sistemler
- Dağıtık Depolama
- Veri Füzyonu

- Yönlendirme ve Kontrol Protokolleri
- Kaynak Yönetimi, Erişim Kontrolü
- Kimlik Yönetimi ve Nesne Tanıma
- Yerini Belirleme Teknolojileri
- Uç Nokta Bilişimi, Sis Bilişimi ve IoT
- Makineler Arası Haberleşme (M2M) ve IoT
- Endüstriyel IoT

IoT Uygulama ve Hizmetleri

- Siber-fiziksel sistemler
- İşbirlikçi Uygulamalar ve Sistemler
- Servis Deneyimleri ve Analizi
- Akıllı Şehirler, Akıllı Kamu Yerleri, Akıllı Ev/Bina
- e-Sağlık, Yaşam Desteği,
- Akıllı Ulaşım
- Akıllı Şebekeler, Enerji Yönetimi
- Tüketici Elektroniği
- Kırsal Hizmetler ve Üretim
- Endüstriyel IoT Servis Oluşturma ve Yönetimi
- Kitle Kaynaklı Algılama, İnsan Merkezli Algılama
- Büyük Veri ve IoT Veri Analitiği
- Semantik Teknolojiler
- Mobil Bulut Bilişim ve IoT
- IoT için Yatay Uygulama Geliştirme
- IoT Uygulama Geliştirme için Tasarım Prensipleri ve En İyi Uygulamalar

IoT Toplumsal Etkileri

- IoT'da İnsan Rolü, Sosyal Hizmetler
- Değer Zinciri Analizi
- IoT için Yeni İnsan-Aygıt Etkileşimleri
- Sosyal Modeller ve Ağlar
- Yeşil IOT: Sürdürülebilir Tasarım ve Teknoloji
- Kent Dinamikleri ve Kitle Kaynaklı Hizmetler
- IoT Sürdürülebilirliği ve ROI için Ölçümler ve Değerlendirmeler

IoT için Güvenlik ve Gizlilik

- IoT Gizlilik ve Güvenlik Endişeleri
- Kimlik Saptama ve Kimlik Doğrulama Sorunları
- IoT Güvenliği için Kablosuz Sensör Ağı
- IoT'da Saldırı Tespiti
- IoT için kriptografi, anahtar yönetimi ve yetkilendirme
- IoT'da Fiziksel / MAC / Ağ Saldırıları
- IoT'da Çapraz Katmanlı Saldırılar
- IoT'da QoS Optimizasyonu ile Güvenlik
- IoT'da Gizlilik Tabanlı Kanal Erişimi
- IoT Adli Bilişimi
- IoT'da Büyük Veri ve Bilgi Bütünlüğü
- IoT'da Haberleşme Güvenliği
- IoT'da Güvenlik Standartları

IoT Deneysel Sonuçlar ve Dağıtım Senaryoları

- Araştırma ve Uygulama Arasındaki Boşluğu Kapama
- Deneysel Prototipler ve Sınama Ortamları
- Çok amaçlı IOT Sistem Modelleme ve Analiz
- IOT Ara Bağlantı Analizi
- Gerçek Vaka Dağıtım Senaryoları ve Sonuçları
- Standardizasyon ve Düzenleme

PREFACE /ÖNSÖZ

Bilgi güvenliği ve siber güvenlik alanında, ulusal ve uluslararası boyutta bilimsel, teknik, sosyal ve kültürel çalışmalar yürüterek kişisel, kurumsal ve ulusal farkındalığın oluşması ve ortak akıl ile çözüm önerilerinin geliştirilmesi amacı ile 2007 yılında kurulan Bilgi Güvenliği Derneği (BGD) her yıl Uluslararası Bilgi Güvenliği ve Kriptoloji (ISCTURKEY) Konferansı düzenlemektedir. Bu konferansın onuncusu, Gazi Üniversitesi, İstanbul Teknik Üniversitesi ve Ortadoğu Teknik Üniversitesi işbirliğiyle ve T.C. Başbakanlık, T.C. Ulaştırma, Denizcilik ve Haberleşme Bakanlığı ve Bilgi Teknolojileri ve İletişim Kurumu'nun destekleriyle 17-18 Ekim 2018 tarihlerinde BTK Kongre Merkezinde gerçekleştirilmiştir.

Uluslararası ISCTURKEY Konferansı, düzenlendiği ilk yıldan beri Türkiye'nin bilgi güvenliği alanındaki bilimsel ve sektörel çalışmalarının paylaşıldığı, üniversite-kamu-endüstri işbirliğinin geliştirildiği, kamuoyunun bilgilendirildiği, eğitildiği, ulusal ve uluslararası tüm bilim insanları, araştırmacılar ve sektörel uygulayıcılar arasında bilgi alışverişinin sağlandığı, ülkemizin bu alandaki en önemli etkinliğidir. Bu etkinlik ile bilgi güvenliği alanında, toplumun her kesiminin farkındalığının arttırılması, bilimsel bilgi birikimine katkı sağlanması, kurumlar ve sektörler arasında işbirliği imkânlarının oluşturulması ve en önemlisi bunu uluslararası boyutta yaparak uluslararası işbirliğinin arttırılması hedeflenmiştir.

ISCTURKEY 2018 Konferansı Avrupa Ağ ve Bilgi Güvenliği Ajansı (ENISA) tarafından da desteklenmiş ve Avrupa Birliği'nin her yılın Ekim ayı olarak belirlediği "Avrupa Siber Güvenlik Ayı" etkinlikleri kapsamına alınmıştır.

ISCTURKEY 2018 Konferansının bu yılki ana teması "Siber Güvenlik ve Blok Zincir Teknolojisi" olarak belirlenmiştir. Milli güvenliğin önemli bir parçası olan siber güvenlik konusunda zafiyet gösterilmemesi için hem nitelikli siber güvenlik uzmanları yetiştirilmesi hem de gerek donanım gerek yazılım alanında milli ve yerli çözümler üretilmesinin şart olduğu düşüncesinden hareketle ISCTURKEY 2018 Konferans programı oluşturulmuştur.

ISCTURKEY 2018 Konferansına, bu yıl 1500'ün üzerinde kişi elektronik kayıt yaptırmıştır. Konferans programında; 4 panel, 2 Kurul Toplantısı, 3 akademik oturum, 1 davetli konuşmacı, 4 eğitim, 3 firma ve ürün tanıtım oturumu gerçekleştirilmiştir.

Konferans açılış konuşmalarını; Bilgi Teknolojileri ve İletişim Kurumu Başkanı Ömer Abdullah Karagözoğlu, Ulaştırma ve Altyapı Bakanlığı, Bakan Yardımcısı Dr. Ömer Fatih Sayan yapmışlardır.

Konferansa sunulmak üzere gönderilen bildiriler, Konferans Bilim Kurulu tarafından incelenmiş ve sunulması önerilen bildiriler, akademik oturumlarda sunulmuş ve bu kitapçıkta basılmıştır.

Bu yıl onuncusunu yaptığımız bu uluslararası konferansın başta ülkemiz ve kurumlarımız olmak üzere tüm katımcılarına faydalı olmasını dileriz.

Saygılarımızla.

Prof. Dr. Şeref **SAĞIROĞLU**, *Konferans Eş-Başkanı*

Prof. Dr. Mustafa **ALKAN**, *Konferans Eş-Başkanı*

Prof. Dr. Ersan **AKYILDIZ**, *Konferans Eş-Başkanı*

Prof. Dr. Ferruh **ÖZBUDAK**, *Konferans Eş-Başkanı*

17 EKİM 2018, ÇARŞAMBA - *17 OCTOBER 2018, WEDNESDAY*

08:30 - 09:00	**KAYIT**
09:00 - 11:00	**AÇILIŞ KONUŞMALARI / ANA SALON / MAIN HALL** • Ahmet Hamdi ATALAY - *Bilgi Güvenliği Derneği YK Başkanı* • Dr. Ömer Fatih SAYAN - *Ulaştırma ve Altyapı Bakanlığı, Bakan Yardımcısı* • Mustafa VARANK - *T.C. Sanayi ve Teknoloji Bakanı* • Mehmet Cahit TURHAN - *T.C. Ulaştırma ve Altyapı Bakanı* • Bilgi Güvenliği Derneği - *VTR Gösterimleri* • Bilgi Güvenliği Derneği - Siber Güvenlik Hizmet Ödülleri Töreni • Fatih Baştan, *HUAWEI Kıdemli Strateji ve İş Geliştirme Yöneticisi*
11:30-12:30	**Davetli Konuşmacı:** • **Prof. Dr. Ali Aydın SELÇUK,** TOBB ETU, Blok Zincir Teknolojisi ve Yaygınlaşması Önündeki Problemler
12:30-13:30	**ÖĞLE YEMEĞİ ARASI / *LUNCH BREAK***
13:30 – 15:30	**PANEL - 1 / ANA SALON / MAIN HALL** ***"Siber Güvenlik ve Blok Zincir Teknolojisi"*** ***Panel Yöneticisi:*** • Prof. Dr. Ertuğrul KARAÇUHA, *İTÜ Bilişim Enstitüsü Müdürü* ***Panelistler:*** • Doç. Dr. Mehmet Sabır KİRAZ, *TÜBİTAK Blok Zincir Araştırma Laboratuvarı Direktörü* • Dr. Öğretim Üyesi Pelin ANGIN, *ODTÜ Öğretim Üyesi* • Gökhan SEÇKİN, *Kimlic Blockchain Yazılım Teknoloji Genel Müdürü* • S. Bilgehan ÜSTÜNDAĞ, *CHOMAR Antivirüs CEO* • İlker İMAMOĞLU, *FORTİNET Türkiye Teknik Müdürü* • Fatma Hacıoğlu DOĞAR, *NETAŞ Siber Güvenlik Servisleri Direktörü*
15:30 – 16:00	**İLETİŞİM ARASI / *BREAK TIME***
16:00 - 18:00	**PANEL - 2 / ANA SALON / MAIN HALL** ***"Ulusal Güvenlik Açısından Siber Güvenlik"*** ***Panel Yöneticisi:*** • Ahmet Hamdi ATALAY, *HAVELSAN Genel Müdürü* ***Panelistler:*** • Ömer KORKUT, *STM Genel Müdür Yardımcısı* • Mahmut KÜÇÜK, *Siber Güvenlik Direktörü* • Mehmet Ali ORTAYATIRTMACI, *TÜRKSAT Kurumsal Bilgi ve Siber Güvenlik Yönetimi Direktörü* • Haydar Erdem YILMAZ, *VODAFONE Bilgi Teknolojileri Operasyon Direktörü* • M. Feridun AKTAŞ, *TURKCELL Teknoloji Yönetişimi ve Güvenlik Direktörü*

18 EKİM 2018, PERŞEMBE - *18 OCTOBER 2018, THURSDAY*

09:00-09:15	**Açılış Konuşması** • **Mehmet Fatih KACIR,** T.C. Sanayi ve Teknoloji Bakanlığı Bakan Yardımcısı
09:15 - 10:30	**PANEL - 3 / ANA SALON / MAIN HALL** ***"Siber Güvenlik Sanayi ve Kümelenmesi"*** ***Panel Yöneticisi:*** • Mustafa ÖZÇELİK, *SSB Siber Güvenlik ve Bilişim Sistemleri Grup Başkanı* ***Panelistler:*** • Doç. Dr. İzzet Gökhan ÖZBİLGİN, *HAVELSAN Ar-Ge, Teknoloji ve Ürün Yönetimi Direktörü* • Burak KIRIMER, *TÜRKTRUST Ar-Ge Merkezi Müdürü* • Murat TORA, *ATAR Labs Kurucu Ortağı* • Fatih Baştan - *HUAWEI Kıdemli Strateji ve İş Geliştirme Yöneticisi* • Serdar YOKUŞ, *Biznet Bilişim Genel Müdürü*
10:30 - 11:00	**İLETİŞİM ARASI / *BREAK TIME***
11:00 - 12:30	**PANEL - 4 / ANA SALON / MAIN HALL** ***"Siber Güvenlikte Eğitim ve İnsan Kaynağı Yetiştirme Politikaları"*** ***Panel Yöneticisi:*** • Prof. Dr. Şeref SAĞIROĞLU, *Gazi Üniversitesi Bilgisayar Mühendisliği Bölüm Başkanı* ***Panelistler:*** • Ali Kemal YURTSEVEN, *HAVELSAN Siber Güvenlik Grup Müdürü* • Prof. Dr. Türksel KAYA BENSGHIR, *Hacı Bayram Veli Üniversitesi, YÖK Siber Güvenlik Çalışma Grubu Üyesi* • Doç. Dr. Sedat AKLEYLEK, *19 Mayıs Üniversitesi, YÖK Siber Güvenlik Çalışma Grubu Üyesi* • Prof. Dr. Ferruh ÖZBUDAK, *ODTÜ UME Kriptografi ABD Başkanı* • Zafer POLAT, *ARISTA NETWORKS Ülke Müdürü*
12:30 - 13:30	**İLETİŞİM ARASI / *BREAK TIME***
13:30 – 15:00	**ANA SALON** Siber Güvenlik Eğitimi Oturum 1 / Training on Cyber Security Session 1 **"Blok Zincir Teknolojileri"** ***Eğitmenler:*** • Emre YÜCE & Duygu ÖZDEN, *HAVELSAN*
15:00 - 15:30	**İLETİŞİM ARASI / *BREAK TIME***
13:30 – 15:00	**ANA SALON / MAIN HALL** Siber Güvenlik Eğitimi Oturum 2 / Training on Cyber Security Session 2 **"Pardus İşletim Sistemi ve Kamuda Kullanımı"** ***Eğitmenler:*** • Artur MEHMET & Ali Orhun AKKIRMAN, *HAVELSAN*
15:00 - 15:30	**İLETİŞİM ARASI / *BREAK TIME***
17:30 - 18:00	**KAPANIŞ KONUŞMALARI / *CLOSING REMARKS:*** **ANA SALON**
	• Prof. Dr. Şeref Sağıroğlu, *ISC TURKEY 2018 Konferansı Eş Başkanı* • Prof. Dr. Mustafa Alkan, *ISC TURKEY 2018 Konferansı Eş Başkanı* • Prof. Dr. Ertuğrul Karaçuha, *ISC TURKEY 2018 Konferansı Eş Başkanı* • Prof. Dr. Ferruh Özbudak , *ISC TURKEY 2018 Konferansı Eş Başkanı*

ARTICLES PRESENTED IN ISCTURKEY 2018

ISCTURKEY 2018'DE SUNULAN BİLDİRİLER

Medikal Verilerin Blok Zinciri Mimarisiyle Güvenliğinin Sağlanması

Securing Medical Data with Blockchain Architecture

Ömer KASIM
Kütahya Dumlupınar Üniversitesi
Simav Teknoloji Fakültesi Elektrik Elektronik Mühendisliği
Kütahya, Türkiye
omer.kasim@dpu.edu.tr

Özet— Günümüz teknolojik cihazlarının ve sosyal medyanın hayatımıza girmesiyle kişisel verilerde müthiş bir artış olmuştur. Veri miktarındaki bu artış, depolama ve yönetim süreçlerinde bulut ortamını önemli hale getirmektedir. Kişiye ait medikal veriler, hassas verilerdir. Bu verilerinin bulut ortamında saklanması ve korunması kritik öneme sahiptir. Bu problemin çözümünde farklı yaklaşımlar olsa da Blok Zinciri mimarisi, verilerin bloklar halinde saklanmasını sağlayarak bir denetim sürecini etkin kılmaktadır. Çalışmada oluşturulan blok zinciri ile medikal veriler, blok zinciri içerisinde tutularak kayıtların güvenli bir şekilde oluşturulması, erişilmesi ve paylaşılmasını kontrol bloğu ile mümkün hale getirmektedir.

Anahtar Kelimeler—Kişisel Veri, Medikal Veri, Blok Zinciri, Güvenlik, Bulut Ortamı

***Abstract*— The current technological devices and social media are integrated to our lives so that there has been a tremendous increase in personal data. This increase in data volume makes the cloud environment important during storage and management processes. The medical records of the personal data are sensitive. The storage and protection of these data in the cloud environment is critical. Though there are different approaches to solve this problem, blockchain architecture makes an audit process effective by storing the data in blocks. The medical data generated in this study are kept within the blockchain, making it possible to securely create, access and share medical records with control block.**

Index Terms— Personal Data, Medical Data, Block Chain, Security, Cloud Environment

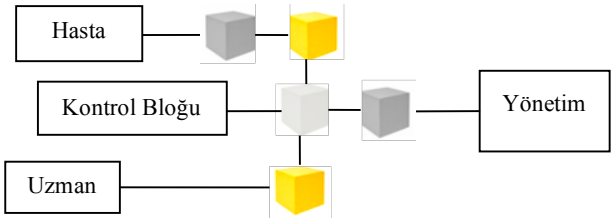
I. GİRİŞ

Günümüz teknoloji dünyasındaki üretilen veri miktarındaki müthiş artış hızı, bulut ortamında verilerin saklanmasını gerekli kılmaktadır. Verilerin bulut ortamında saklanması bir bulut veri kaynağı yaratılması ile başlamaktadır. Bu veri kaynağı, bulut ortamında bulunan veri nesnesinde gerçekleştirilen "veri oluşturma" ve işlemlerin geçmişini kaydeden "meta verileri" bileşenlerinden oluşmaktadır [1]. Verilerin bulut ortamında saklanması güvenlik sürecinin etkin kullanımını gerektirmektedir. Kişiye ait hassas verilerin güvenliğinin sağlanması önemli hale gelmektedir.

Bulut ortamında tutulacak hassas verilerden birisi de medikal verilerdir. Medikal verilerin güvenliği açısından bakıldığında kişisel verilere ait gizlilik sorunlarını çözümleyen çeşitli çözüm önerileri literatürde sunulmuştur. Bu çözümlerden birisi veri anonimleştirme olarak isimlendirilmiştir. Veriler anonim hale getirilerek tanımlanabilir bilgilerin korunması sağlanmaktadır [2]. Bu süreçte kişiye ait veriler, hassas verilerden ayrı saklanarak veriye erişim, kontrollü hale getirilmektedir [3]. Hassas verilere erişim kontrolü ise hassas verilerin dağılımını belirleyen "yakınlık derecesi" parametresi ile belirlenmektedir [4]. Bir diğer yaklaşım ise verileri paylaşmadan önce hesaplama işlemine tabi tutarak şifreleme esasına dayanmaktadır. Şifrelenen verinin sahip olduğu örüntüyü bilen kullanıcılar, kullanıcı haklarına göre belirlenen belirli sorguları çalıştırabilmektedir [5].

Alınan bu önlemler güvenlik noktasında denetim mekanizması sağlasa da medikal verileri içeren meta verilerdeki hassas bilgilerin korunması gerekmektedir. Bulut ortamında oluşturulacak veriler, sahteciliğine karşı savunmasız olduğu ilgili çalışmada tespit edilmiştir [6]. Bu durum veri güvenilirliğini önemli hale getirmektedir. Bu verilerin güvenliğinin sağlanarak bulut ortamında saklanabilmesi günümüzde blok zinciri mimarisiyle mümkün olabilmektedir [7].

Blok zinciri, finans sektöründe kullanılan eşler arası dağıtılmış blok teknolojisi üzerine inşa edilmiştir. Bir kullanıcının kimliğinin bir ağ içinde nasıl tanımlandığına bağlı olarak izin verilen ve izinsiz blok zinciri olarak iki farklı türde tasarlanabilmektedir. İzinsiz bir sistem tasarımında katılımcıların kimliğinin sahte veya anonim olması önem arz etmemektedir. Bu tasarımda her kullanıcı blok zinciri mimarisine yeni bir blok ekleyebilmektedir [8]. Diğer taraftan izin verilen bir blok zincir tasarımında ise bir kullanıcının kimliği, bir kimlik sağlayıcı tarafından kontrol edilmektedir. Bu tasarımda kimlik sağlayıcısının rolü kritik öneme sahiptir. Bu sağlayıcı, ağ içinde erişim kontrolünü ve kullanıcının uzlaşmaya katılma haklarını koruma görevini üstlenmektedir. Ayrıca yeni bir bloğu onaylamak için güvenilir olması zorunluluğu bulunmaktadır [9].



Şekil 1: Algoritmanın Akış Diagramı

Blok zinciri teknolojisini kullanarak oluşturulacak bir bulut ortamı, veri kaynağı ile verilerin gizliliğini ve kullanılabilirliğini güvenli hale getirilebilmektedir [10]. Bu süreç ile proaktif siber güvenliğe katkı sağlanmaktadır [11].

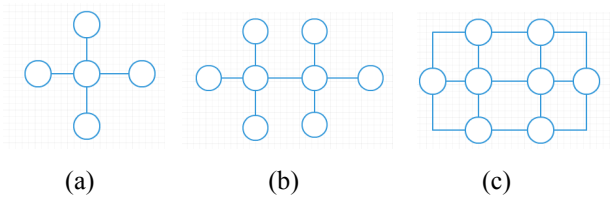
Çalışmada medikal verileri içeren bir blok zinciri mimarisi geliştirilmiştir. Geliştirilen yapı şekil 1'de ifade edilmiştir. Şekil 1'de blok zincirinde her bir hastaya ait veriler bir blok içerisinde tutulmaktadır. Verilerin eklenmesi ya da güncellenmesi durumunda blokların "hash" içerikleri değişmektedir. Bu durum içeriği değişen bloktan sonraki zincir içerisindeki blokların "hash" bilgilerinin güncellenmesi ile son bulması gerekmektedir. Blok zinciri mimarisinde Madencilik süreci olarak isimlendirilen işlem ile zincirdeki bloklar taranmakta ve "hash" bilgileri SHA algoritması ile onaltılık kodlara dönüştürülerek güncellenmektedir. Bu süreç ile blok zinciri doğrulanması yapılmaktadır. Doğrulama süreci ve madencilik işlemi orta noktadaki kontrol bloğu ile sağlanmaktadır. Bu blok üzerinden birbirine bağlı olan blokların hash kodları eşlenerek zincir oluşmaktadır. Kontrol bloğu ile aynı zamanda veri eklenmesi ve güncellenmesi yapılmaktadır. Ayrıca kontrol bloğu sayesinde doğrulama işlemi ile araya eklenebilecek ya da saldırı amaçlı içeriği güncellenecek blokların doğrulaması yapılmayacak olduğundan bulut ortamındaki verilerin korunması noktasında siber güvenliğe destek olunmaktadır. Ayrıca kontrol bloğu ile blok içerisindeki verilerin okunması hasta, uzman ya da yönetim rollerine göre erişim hakkı sağlanmaktadır. Blok zincirine özel bir başlangıç noktası olduğundan bu bilgiyi bilmeyen bir kullanıcı, zincire müdahale edememektedir. Ayrıca bloğa ait zaman damgası bilgisi ile hash içeriği güncellendiğinden verilerin oluşma zamanına göre de bir güvenlik süreci geliştirilen yöntem ile sağlanmaktadır.

II. Medikal Veri İşlemleri

Uzun süreli tedavi ve hastanın ömür boyu izlemesini gerektiren ciddi bir tıbbi rahatsızlık içeren hastalıklar ve bu hastalılara ait tedavi süreçleri hassas bilgileri içermektedir. Bu nedenle, hastanın tıbbi geçmişini muhafaza etmesi, tedavi ve tedavi sonrası izleme sırasında tıbbi verilerine erişebilmesi veya araştırma amaçlı paylaşılması önem arz etmektedir. Bir hastanın hareketliliği nedeniyle, her hastanın ziyareti sırasında üretilen verilerin yönetimi, özellikle sağlık verilerinin hassas doğası göz önüne alındığında zor bir süreci içermektedir. Blok zinciri mimarisi olmadan yapılacak bir tasarımda hassas verilerden herhangi birinin Hastane 1'den Hastane 2'ye aktarılması gerekiyorsa belirli bir protokol takip edilmektedir. Bu protokolün işlemesi sürecinde hasta veya onun resmi temsilcisi bir rıza anlaşması imzalamak zorundadır. Bu anlaşmanın içeriğinde aktarılacak verileri belirten ve verilerin alıcısıyla ilgili bilgileri bulunmaktadır. Bu anlaşma sonrasında hastaya ait veriler bir başka sağlık kuruluşuna aktarılabilmektedir. Her bir aktarımda hasta ile ilgili olan tahlil ve tedavi süreçlerini içeren verilerin güvenliğinden dolayı tekrar tanımlama yapılması gerekmektedir. Tahlillerin tekrar yapılması hastaya ait klinik işlemlerin tekrarlanması gibi sonuçları olan bu süreçte zaman kaybı yaşanmaktadır. Ayrıca bu yaklaşımla, hastanın verilerinin herhangi bir erişim kontrolünü sürdürmesi ve verileri tam olarak görebilmesi oldukça zor bir süreçtir.

III. Blok Zincir Mimarisi

Blok zincir mimarisi, tek nokta üzerinden merkezi olarak oluşturulmuş güven sistemi yerine verilerin bloklar ile ifade edilmesi sürecine dayanmaktadır. Bu durum sistemin şifreli kayıt defteri biçiminde oluşarak daha verimli çalışmasını sağlamaktadır.

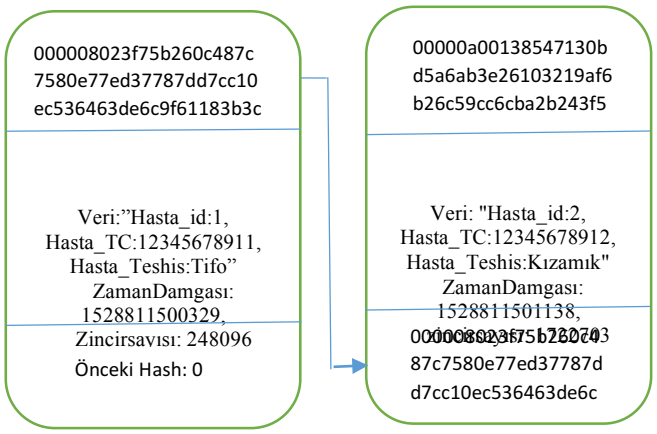


Şekil 2: Blok Zinciri Mimarisi Yapısı
a) Merkezi Blok zinciri Tasarımı
b) Sorumluluğun Dağıtıldığı Blok Zinciri Tasarımı
c) Dağıtık Blok Zinciri Tasarımı

Üç farklı biçimde blok zinciri oluşturulabilmektedir. Üç türe ait sistem yapısı şekil 2'de gösterilmiştir. Şekil 2a'da görüldüğü üzere merkezi esas alan bir blok zinciri tasarımı görülmektedir. Bu tasarımda merkez düğümdeki blok üzerinden tüm süreç idare edilmektedir. Sorumluluğun paylaşılması ile farklı düğümdeki bloklardaki giriş seviyeleri ile bloklar arası iletişim süreci sağlanmaktadır. Bu durum Şekil 2b'de ifade edilmiştir. Herkesin eşit sorumluluğa ve haklara sahip olduğu dağıtık yapıda ise Şekil 3c'de görüldüğü üzere bloklar arası iletişim ile herkes her bloğa ulaşabilmekte ve müdahale edebilmektedir. Her bir tasarım sürecinde izinli ya da izinsiz bir süreç üzerinden yapılan bilgi güncelleme ve veri ekleme işlemleri ile hash bilgileri güncellenmektedir. Madencilik işleminin ardından blok zinciri doğrulaması yapılarak bulut ortamında veriler saklanmaktadır.

IV. Geliştirilen Yöntem

Blok zinciri yapısı, birbirine bağlı bloklardan oluşan bir mimariye sahiptir. Zincirde yer alan her bir blok kendine ait



Şekil 3: Medikal Verinin Blok Zinciri Mimarisindeki Tasarımı

sayısal bir imzaya sahiptir. Bu imza kendisinden önceki bloğun sahip olduğu imza ve blok içerisindeki veriye göre belirlenmektedir. İmzanın blok zinciri sürecindeki ismi "hash" olarak isimlendirilmektedir. Hash sadece kendisinden önceki bloğun hash verisine sahip değildir. Önceki bloğun hash içeriği, blok içerisindeki veri, zaman damgası ve işlenen zincir sayısı bilgileri kullanılarak bloğun kendine ait hash içeriği hesaplanmaktadır. Bir bloktaki verinin değişmesi hash içeriğinin değişmesine de neden olacaktır. Bu durum hash içeriği değişen bloktan sonraki bloklara yansıtılması gerekmektedir. Dolayısıyla değişiklik yapılan bloktan sonraki blokların hash içeriklerinin güncellenmesi gerekmektedir. Hash içerikleri hesaplanarak bir bloğun geçerli ya da geçersiz olduğu kararı Madencilik işlemi ile verilmektedir.

Geliştirilen sistem şekil 2a'da ifade edilen merkezi blok zinciri mimarisine sahiptir. Bu mimaride ortada yer alan blok kontrol bloğu olarak tasarlanmıştır. Veri girişi, güncellemesi ve madencilik işlemleri bu blok üzerinden yapılmaktadır. Bu süreç JAVA platformunda geliştirilmiştir. Oluşturulan veri blokları ise dizi listesi biçiminde olduğu için hastane içerisindeki lokal bir bulut üzerine aktarılmaktadır.

Kontrol bloğunda üç farklı kullanıcı tipi bulunmaktadır. Her bir kullanıcı sisteme bağlanırken kullanıcı rolünü seçmektedir. Seçilen rol doğrultusunda blok üzerinde görebileceği bilgiler süzülmektedir. Bilgiler metin dosyasında saklandığından veri okuma işlemi JAVA dilinin metin dosyası satır numarası okuma işlemi ile gerçekleştirilmektedir. Hasta kullanıcı adıyla sadece birinci ve ikinci satırda yer alan "hastatc" ve "hastaid" alanları görülebilmektedir. Uzman kullanıcı yetkisiyle hasta yetkisine istinaden ek olarak teşhis bilgisi görülebilmektedir. Yönetim yetkisi ile giriş yapıldığında ise zaman damgası ve işlenen zincir sayısı satırları görülebilmektedir.

Şekil 3'te oluşturulan blok zinciri yapısı içeresindeki hash içerikleri Sha256 algoritması ile şifrelenmektedir. Zincirdeki ilk bloğun hash bilgisi 0 olarak çalışmada belirlenmiştir. Bu değer ile zincire giriş yapılması mümkün olacaktır. Bu sayı üzerinden hash içerikleri oluşturulmaktadır. Bu değeri sadece zincire veri ekleyecek kullanıcıların bilmesi güvenlik açısından önemlidir. Hash içeriği oluşturulurken önceki blok, hash içeriğinin yanı sıra zaman damgası ve çözülen zincir sayısı bilgisi de sürece dâhil edilmektedir. Madencilik sürecinde eğer farklı bir içeriğe sahip blok sisteme eklenmiş ise blok geçersiz olacaktır.

Blok 1
Hash: "000008023f75b260c487c7580e77ed3778
7dd7cc10ec536463de6c9f61183b3c",
oncekiHash: "0",
Veri: "Hasta_id:1,Hasta TC:12345678911 Hastalığı:Tifo",
Zaman Damgası: 1528811500329,
Çözülen Zincir Sayısı: 248096

Blok 2
Hash: "00000a00138547130bd5a6ab3e261
03219af6b26c59cc6cba2b243f58400d5e9",
oncekiHash: "000008023f75b260c487c7580e77ed
37787dd7cc10ec536463de6c9f61183b3c",
Veri: "Hasta_id:2,Hasta TC:12345678912 Hastalığı:Kızamık",
Zaman Damgası: 1528811501138,
Çözülen Zincir Sayısı: 1722703

Blok 3
Hash: "00000a8f1e4daa6b472bc414ec3d800
f3c4e4f11fb93606d2465af218b60e3ec",
oncekiHash: "00000a00138547130bd5a6ab3e26103
219af6b26c59cc6cba2b243f58400d5e9",
Veri: "Hasta_id:3,Hasta TC:12345678913 Hastalığı:HIV",
Zaman Damgası: 1528811506083,
Çözülen Zincir Sayısı: 3050138

Blok 4
Hash: "00000a04d97fcebb1ddfae25d2a1917
c3052c01d06ce577767f8cf82465804fd",
oncekiHash: "00000a8f1e4daa6b472bc414ec3d800
f3c4e4f11fb93606d2465af218b60e3ec",
Veri: "Hasta_id:4,Hasta TC:12345678914 Hastalığı:Sağlıklı",
Zaman Damgası: 1528811514665,
Çözülen Zincir Sayısı: 260346

Blok 5
Hash: "000005a84cc4a5cf00cb8df5214f911
fbcadf84218ba781fe79a47658f4178c3",
oncekiHash: "00000a04d97fcebb1ddfae25d2a1917
c3052c01d06ce577767f8cf82465804fd",
Veri: "Hasta_id:5,Hasta TC:12345678915 Hastalığı:Sağlıklı",
Zaman Damgası: 1528811515417,
Çözülen Zincir Sayısı: 170066

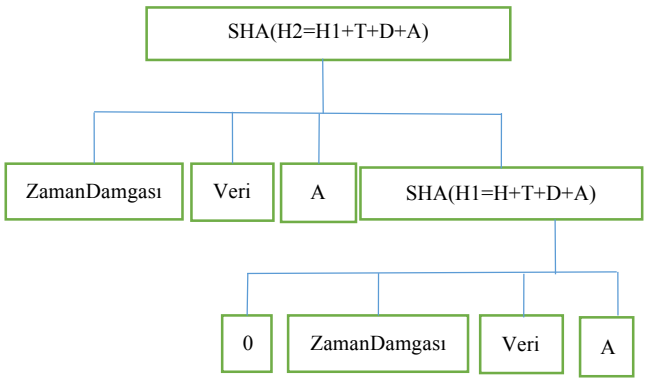
Şekil 4: Medikal Verinin Blok Zinciri Mimarisindeki Tasarımı

Geliştirilen programda her bir blok dizi listeleri üzerinde saklanmaktadır. Oluşturulan dizi listelerinde blok ve blok bağlantıları içerik olarak saklanmaktadır. Medikal verilerin olduğu her bir blok içerisinde barındırdığı verilerin yanı sıra kendi hash bilgisine de sahip olmaktadır. Zaman damgası ve

çözülen zincir sayısı bilgileri ile birlikte önceki hash kodlarıyla eşleştirilerek blok zinciri oluşmaktadır. Çalışmada elde edilen 5 bloğa ait sürece ait bilgiler şekil 4'te ifade edilmiştir. Her bir blokta Hash bilgisi, önceki hash bilgisi, medikal veri, zaman damgası ve çözülen zincir sayısı bilgileri bulunmaktadır. Bloktaki hash bilgisi bloğun kendi imzasını belirlemektedir. Geliştirilen ilk bloğun hash bilgisi 64 hex karaktere sahip olan *"000008023f75b260c487c7580e 77ed37787dd7cc10ec536463de6c9f61183b3c"* bilgisini içermektedir. Bu bilgi SHA256 şifreleme algoritması ile elde edilmektedir. Birbirine bağlı 5 blok hem kendi hash bilgisini hem de kendinden önceki bloğun hash bilgisini tutmaktadır. Bu süreç sonucu blok zinciri yapısına benzer bir mimariyi oluşturmamıza olanak sağlamaktadır.

Blok içerisinde yer alan ilk değer 0 (sıfır) olarak belirlenmiştir. Bu değere zaman damgası yani verinin oluşturulma zamanı bilgisi eklenmektedir. Verinin oluşturulma zamanı bilgisi verinin güvenliğinin sağlanması için gerekmektedir.

Kendi içerisinde yer alan veri, verinin oluşturulma ya da güncellenme zamanı bilgisi ve çözülen zincir sayısı bilgilerini de içerin bir kod olan hash kod ile verinin güvenliği sağlanmaktadır. Hash kod üretme süreci şekil 5'te ifade edilmiştir. Şekildeki A bilgisi çözülen zincir sayısını, T bilgisi zaman damgası bilgisini, D bloktaki hastalıkla ilgili olan veriyi ve H'de bloktaki güncellemeden önceki hash bilgisini ifade etmektedir. Dizi listesi içerisindeki hash içerikleri karşılaştırılarak geçerli ya da geçersiz blok kararı, zincir içerisinde denetim sonucu verilmektedir. Çalışmada elde edilen doğrulama süreci, şekil 6'da ifade edilmiştir. Bloklar arası eşleşme hash kodları ile sağlanmaktadır. Her bir bloğa yeni veri eklenmesi ya da yeni bir blok eklenmesi ile hash kodları güncellenmektedir. Bu güncelleme sonucu eşleşen bloklar, blok zincirini oluşturmaktadır. Bu mimari dışında bir blok sürece dahil edilse bile doğrulama adımından geçilemediği için veriler güvende kalmaktadır [12].



Şekil 5: Hash Kodunun Oluşum Evreleri

İşlenen Blok 1
Yeni Hash: 000008023f75b260c487c7580
e77ed37787dd7cc10ec536463de6c9f61183b3c

İşlenen Blok 2
Yeni Hash: 00000a00138547130bd5a6ab3
e26103219af6b26c59cc6cba2b243f58400d5e9

İşlenen Blok 3
Yeni Hash: 00000a8f1e4daa6b472bc414ec3
d800f3c4e4f11fb93606d2465af218b60e3ec

İşlenen Blok 4
Yeni Hash: 00000a04d97fcebb1ddfae25d
2a1917c3052c01d06ce577767f8cf82465804fd

İşlenen Blok 5
Yeni Hash: 000005a84cc4a5cf00cb8df52
14f911fbcadf84218ba781fe79a47658f4178c3

Blok zinciri geçerliği: Doğrulandı.

Şekil 6: Medikal Verilerin Blok Zinciri Mimarisindeki Hash Kodlarının Üretilmesi Süreci

Bir kullanıcı, blok zinciri mimarisi içerisindeki bloklardan daha fazlasını ekleyebilirse blok zincirine sahip olma olasılığı bulunmaktadır [13]. Bu süreci engellemek amacıyla çalışmada yerel sistem üzerinde üretilen veriler, metin dosyaları ile zincire aktarılmaktadır. Bu metin dosyası belirli bir örüntü içerisinde oluşturulmaktadır. Bu örüntüye sahip olan dosyalar zincire eklenmeden önce denetlenmektedir. Bu doğrulama adımı sağlanmaz ise zincire erişim engellenmektedir. Aynı zamanda güncelleme yapılacağı zaman aynı süreç geçerlidir. Programdaki metot ile zincire eklenecek veriler, metin dosyalarından çekilerek denetim mekanizmasının ardından zincire blok olarak eklenmektedir.

V. SONUÇLAR

Medikal verilerin gizliliği ve güvenliğinin sağlanması amacıyla bu çalışmada, kullanıcı korunabildiği sağlayan blok zinciri mimarisine sahip bir bulut veri kaynağı süreci tasarlanmıştır. Blok zinciri süreci içerisinden yer alan ve değiştirilemeyen zaman damgasıyla hastaya ait verilerin blok zincirine kayıt süreci gerçekleştirilmiştir.

Sistemin tasarımında yer alan kontrol bloğu ile zincire eklenecek veri bloğu hastane ortamında güvenilir bir kaynak üzerinden beslenmesi veri güvenliği açısından önemlidir. Bu kaynak dışında bir veri eklemesi yapılırsa Madencilik süreci sonunda doğrulama yapılamayacağı için bulut ortamına eklenen veri bloğunun farklı bir kaynaktan eklendiği tespit edilecektir. Bu süreç kişiye ait hassas verilerin korunmasını sağlamaktadır. Veri ekleme süreci metin dosyası ile yapılmakta olup belirli bir ekleme stiline sahiptir. Eklenecek veriler bu stile uygun olarak zincire blok olarak eklenmektedir. Farklı bir stilde eklenecek

veriler yine Madencilik sürecinden geçemeyecektir. Bu sistem yapısında çözülen blok zinciri sayısı bilgisi ve hash içeriği oluşturulmaktadır. Bir blok içerisinde saklanan bu veriler ile oluşturulan blok zinciri ile hastaya ait verilerin doğrulanması Madencilik süreci ile gerçekleştirilmektedir. Doğrulama işleminin ardından siber saldırı ile blok zinciri içerisine sızılsa bile blok doğrulamadan geçilmediği için verilerin güvenliği sağlanmaktadır. Aynı zamanda veri okuma işlemi de kontrol bloğu üzerinden sağlanmaktadır. Hasta, yönetim ya da uzman rolünde belirlenen kullanıcılar kontrol bloğunda tespit edilerek veri bloklarına okuma amacıyla erişebileceklerdir. Bu süreç her bir kullanıcıya ait parola yardımıyla yapılmaktadır.

Çalışmada geliştirilen bu yaklaşımla, hastaya ait verilerin bulut ortamına blok zinciri mimarisinde alınması sağlanmaktadır. Bulut ortamına saf veri olarak alınması durumunda veri eklemesi, güncellemesi ve kişiye ait verilerin manipüle edilmesi söz konusudur. Bu problem blok zinciri mimarisi ile mümkün olmamaktadır. Ayrıca veriyi okuyacak kullanıcılar rollerine göre hasta, uzman ya da yönetim olarak bağlanıp sadece kendi ilgi alanlarındaki blok satırlarına ulaşabilmektedir. Bu süreç kontrol bloğu ile hastane merkezinde yer alan bir sunucu ile sağlanmaktadır. Bu sunucu sistemi Madencilik sürecini yine yerel bulut ortamına bağlanarak yapmakta ve bulut veri bloklarını doğrulamanın ardından güncellemektedir.

Hastanın farklı sağlık kuruluşlarından hizmet alması ya da bilgilerinin aktarılması gerektiğinde hastaya ait verilerinin kurumların arasında geçişi önlenmiş olacaktır. Bu süreç kontrol bloğuna erişim izni yönetim ya da uzman olarak verildiğinde gerçekleştirilmektedir. Yapılan tahlil ve tedavi bilgileri de kolay ve güvenli bir şekilde kurumlar arasında taşınmış olacaktır. Bu süreç ile özellikle hastanın bütün yaşamı boyunca sahip olduğu medikal verilerin tek bir blokta saklanması sağlanmış olmaktadır.

KAYNAKLAR

[1] Y. L. Simmhan, B. Plale, D. Gannon, “A survey of data provenance in e-science,” ACM Sigmod Record, vol. 34, no. 3, 2005: pp. 31–36.

[2] L. Sweeney, "k-anonymity: A model for protecting privacy," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems vol:10, no:05 2002: pp. 557-570.

[3] A. Machanavajjhala, D. Kifer, Johannes Gehrke, Muthuramakrishnan Venkitasubramaniam, “L-diversity: Privacy beyond k-anonymity,” ACM Transactions on Knowledge Discovery from Data (TKDD), vol:1, no:1, 2007:pp.1-52.

[4] N. Li, T. Li, S. Venkatasubramanian. "t-closeness: Privacy beyond k-anonymity and l-diversity," IEEE 23rd International Conference on Data Engineering, 2007:pp.106–115.

[5] C. Gentry, "Fully homomorphic encryption using ideal lattices,” Proceedings of the 41st annual ACM symposium on Symposium on theory of computing-STOC'09. Vol. 9, 2009:pp. 169–178.

[6] B. Lee, A. Awad, M. Awad, “Towards secure provenance in the cloud: A survey,” in 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC), 2015:pp. 577–582.

[7] Lin, Iuon-Chang, Tzu-Chun Liao, "A Survey of Blockchain Security Issues and Challenges," International Network Security Vol:19, No:5, 2017:pp:653-659.

[8] Junqueira, Flavio P., Benjamin C. Reed, Marco Serafini, "Zab: High-performance broadcast for primary-backup systems," IEEE/IFIP 41st International Conference on Dependable Systems and Networks (DSN), 2011: pp. 245-256.

[9] T. Swanson, "Consensus-as-a-service: a Brief Report on the Emergence of Permissioned," Distributed Ledger Systems 2015:pp:1-66.

[10] D. Tosh, S. Sengupta, C. A. Kamhoua, and K. A. Kwiat, “Establishing evolutionary game models for cyber security information exchange (cybex),” Journal of Computer and System Sciences, vol:98, 2016:pp.27-52.

[11] D. K. Tosh, S. Sengupta, S. Mukhopadhyay, C. Kamhoua, and K. Kwiat, “Game theoretic modeling to enforce security information sharing among firms,” in IEEE 2nd International Conference on Cyber Security and Cloud Computing (CSCloud), 2015: pp. 7–12.

[12] I. Eyal, E.G. Sirer, “Majority is not enough: Bitcoin Mining is vulnerable, in: Financial Cryptography and Data Security”, 18th International Conference, in: Lecture Notes in Computer Science, vol:8437, 2014:pp. 436–454.

[13] Li, X., Jiang, P., Chen, T., Luo, X., Wen, Q., “A survey on the security of blockchain systems.” Future Generation Computer Systems, 2017: http://dx.doi.org/10.1016/j.future.2017.08.020.

Yapay Zekâ ve Hukuk
Artificial Intelligence and Law

Göksu Hazar ERDİNÇ
İstanbul Teknik Üniversitesi Bilişim Enstitüsü
İstanbul, Türkiye
erdincg@itu.edu.tr

Ertuğrul KARAÇUHA
İstanbul Teknik Üniversitesi Bilişim Enstitüsü
İstanbul, Türkiye
karacuhae@itu.edu.tr

***Özet*— Endüstri 4.0 devrimi neticesinde siber güvenlik alanında yapay zekâ uygulamaları artmıştır. Yapay zekânın günlük hayatta kullanımın artmasıyla beraber bu konunun yasal olarak düzenlenmesi ihtiyacı da doğmuştur ve beraberinde birtakım tartışmalar da getirmiştir. Özellikle Fikri Mülkiyet Hukuku ve Ceza Hukuku bakımından düzenlemelerin ne şekilde yapılabileceği konusunda çeşitli görüşler ortaya atılmıştır. Bu çalışma kapsamında yapay zekâ kavramı irdelenerek bu konuya ilişkin hukuksal boşluklar özellikle Fikri Mülkiyet ve Ceza Hukuku kapsamında değerlendirilecektir. Mevcut hukuki boşluklara ilişkin yasal düzenlemelerin yapılması için ne gibi adımlar atılabileceği konusunda önerilere yer verilecektir.**

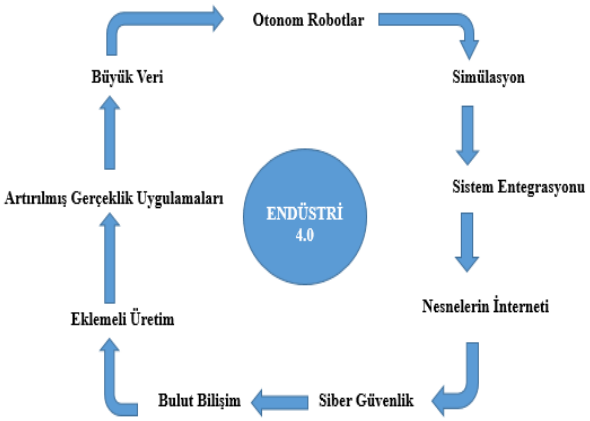
***Anahtar Kelimeler*—düzenleme, hukuk, robot, siber güvenlik, yapay zekâ**

***Abstract*- As a result of the Industry 4.0 revolution, artificial intelligence applications in the field of cyber security have increased. With the increasing use of artificial intelligence in daily life, the need for regulating the artificial intelligence has arisen, and some arguments were made. There are various views on how the regulations can be made especially in terms of Intellectual Property Law and Criminal Law. Within the scope of this study, the concept of artificial intelligence will be examined and the legal gaps related to this subject will be evaluated especially within the scope of Intellectual Property and Criminal Law. Recommendations will be made on what steps can be taken to make legal regulations regarding the existing legal gaps.**

***Keywords*- artificial intelligence, cyber security, law, regulation, robot**

I. GİRİŞ

Dördüncü sanayi devrimi olarak nitelendirilen Endüstri 4.0 ile bilişim sistemi daha otonom hale getirilmiş; yapay zekâ kavramı ön plana çıkartılmıştır. Endüstri 4.0 ile siber güvenliğe yönelik tehditler de boyut değiştirmiştir ve yapay zekânın siber güvenlik alanında kullanılması daha önemli hâle gelmiştir. [1]. Yapay zekâ ("artificial intelligence – AI"), bilgi toplumunda en önemli rol oynayan kavramlardan bir tanesidir. Nesnelerin interneti ("IoT"), büyük veri gibi alanlarda çalışmaların artmasıyla yapay zekâya verilen önem de artmıştır. Yapay zekâya ilişkin herhangi net bir tanım bulunmamakla birlikte, "bir dijital bilgisayarın veya bilgisayar kontrollü robotun, akıllı varlıklar ile yaygın olarak ilişkili görevleri yerine getirebilmesi" olarak tanımlanabilir [2]. Bu kavram, insanların akıl yürütme, anlamını keşfetme, genelleme ya da geçmiş deneyimlerden öğrenme gibi, entelektüel süreçlerin karakteristiklerine dayanarak geliştirilen sistemleri ifade etmektedir [2]. Yapay zekânın temelinde büyük veri, gelişmiş algoritmalar, nesnelerin interneti gibi çeşitli alanlar bulunmaktadır.



Şekil 1: Endüstri 4.0 Oluşumu [1]

Yapay zekâ, özelliklerine göre genel olarak dört kategoride toplanmaktadır. Bunlar ise başlıca:

- Tamamen Reaktif: Yapay zekânın en basit hâlidir ve bu türde hafızaya herhangi bir şey kaydedilmez, saklanmaz veya geçmiş deneyimlerin karar mekanizmasında etkisi de söz konusu olamaz. Yapay zekâ, burada sadece durumu algılar ve duruma göre direkt tepki geliştirir. Daha geniş bir dünya algısı mevcut değildir; sadece tek bir alanda uzmanlaşabilir [3].
- Sınırlı Hafıza: Tamamen reaktif yapay zekâya göre daha gelişmiştir. Bu tür, geçmişe ait bilgileri önceden programlanmış dünya algısına ekler. Doğru kararı vermek ve bu doğrultuda hareket edebilmek için yeterince deneyimi ve hafızası mevcuttur [3].
- Zihin Teorisi: Zihin teorisi özellikli yapay zekânın insan davranışlarını etkileyen duygular ve düşünceleri anlama kapasitesi vardır. Yapay zekâ burada daha sosyaldir; yani insanlarla sosyalleşebilir ve çeşitli motivasyonları, duyguları ve niyetleri kavrayabilir [3].

- Öz Farkındalığı (Bilinci) Olan: Bilinç özellikli yapay zekâ, kendini ifade edebilme yetisine sahiptir. Yapay zekânın bilinci vardır ve hangi durumda olduğunun farkındadır. Başkalarının da duygularını öngörerek düşünceler üretip çeşitli çıkarımlarda bulunabilir, sonuç öngörebilir. Bu türü en üst düzey yapay zekâ olarak da tanımlamak mümkündür; duygusal, duyarlı/bilinçli ve süper akıllı özelliklere sahiptir [3].

Yapay zekâ alanında yapılan çalışmaların ilerlemesiyle çeşitli akıllı makineler ve eşyalar geliştirilmiştir. Yapay zekâ, hukuk alanında da kullanılmaya başlanmıştır. Hukuk alanında yapay zekâ destekli yazılımlar, yasal kullanım için doküman analizinin verimliliğini artırır ve makineler belgeleri inceleyip belirli bir durumla ilgili olarak önemli noktaları işaretleyebilir. Başka belgeler gösterildiğinde makine öğrenimi algoritmaları kullanılarak benzer şekilde alakalı olan diğer belgeleri bulmak açısından işaretlemeleri işe yarayabilir [4]. Makineler, belgeleri sıralamakta insanlara göre çok daha hızlıdır ve istatistiksel olarak doğrulanabilecek çıktılar ve sonuçlar üretebilir. Sistematik olarak makineler yapacak olsa bile, araştırmaların zamanında ve kapsamlı bir şekilde yapılması önemlidir. Bu amaçla yapılan ROSS, yapay zekâ sistemleri belgelerin analiz edilmesine yardımcı olmak için doğal dil işlemeyi kullanır ve insanların daha kolay ve hızlı iş yapabilmesine yardımcı olur [4]. ROSS'un eş kurucularından olan Andrew Arruda'ya göre yapay zekâ dört teknoloji unsurundan oluşmaktadır. Bunlar sırasıyla makine öğrenmesi ("machine learning"), ses tanıma ("speech recognition"), doğal dil işleme ("natural language processing") ve görüntü algılama ("image recognition") olarak sıralanmıştır [5].

Yapay zekânın en fazla kullanıldığı teknolojilerden bir tanesi de robotlardır. Robotlar büyük ölçüde otomotiv sektöründe montaj hattı gibi alanlara ilişkin tehlikeli, matlaştırma ve benzeri gibi tekrarlayan görevleri yapmak için imalat işlerinde kullanılmıştır [6]. Günümüzde ise robotlar artık yüksek çözünürlüklü kameralar, dokunmatik sensörler gibi önemli donanımlara sahip olup bilgisayar beyinleri tarafından yönetilen gelişmiş makineler hâline gelmiştir [6]. Konuşan, dans eden, mimikleri okuyabilen, sözel komutlara karşılık verebilen robotlar geliştirilmiştir. Araba süren, yaşlıların bakımını üstlenen, Uluslararası Uzay İstasyonu'nda çalışan, teröristleri öldürebilen robotlar mevcuttur ve bunlar nesnelerin interneti kullanılarak da (akıllı telefon uygulamaları, internet ve benzeri vasıtalar ile) kontrol edilebilmektedir [6].

II. YAPAY ZEKÂ ALANINDAKİ HUKUKİ BOŞLUKLAR

Hukuk, insanların karşı karşıya kaldığı sorunların çözülmesi amacıyla bu olaylara yönelik düzenlemeleri içeren normlardır [7]. Genellikle hukuk kuralları meydana gelen zararlı sonuçlardan ortaya çıkarak kurallaşmıştır. Örneğin dolandırıcılığa maruz kalan insanların maddi ve manevi kayba uğramaları nedeniyle dolandırıcılık suçu düzenlenmiştir. Yapay zekânın sınırlarını belirlemek mümkün olmadığı ve teknoloji hızla geliştiği için yapay zekâ ile ilgili düzenlemelerin ne şekilde olması gerektiği konusunda çeşitli tartışmalar bulunmaktadır. Hukukçular açısından yapay zekânın tanımlanamaması hususu problem teşkil etmektedir; çünkü sınırları çizilmeyen kavramlara ilişkin normlar uygulanırken uygulamayla çelişebilir veyahut kuralların uygulanmasında keyfiyete sebebiyet verebilir. Aynı şekilde, tanımlanamayan bir kavram hakkında düzenleme yapılması, temel problemlerin ortaya konulması mümkün olmamaktadır [8]. Yapay zekâ, özellikle fikri mülkiyet hukuku ve ceza hukuku açısından tartışmalara konu olmaktadır. Örnek olarak, Amerika'da mevcut fikri mülkiyet hukukuna ilişkin kurallar mahkemelerce uygulanmakla beraber yapay zekânın patente konu olması meselesi de mahkemelerce değerlendirilmektedir. Yapay zekânın tescil edilmesine veya yapay zekânın uygulanmasından kaynaklanacak muhtelif sorunlara ilişkin herhangi bir düzenleme bulunmamaktadır [9]. Amerika'daki Telif Hakları Bürosu ("Copyright Office"), "makinelerden ve salt otomatik veya rastgele mekanik süreçlerden ortaya çıkan işlerin herhangi bir insan müdahalesi olmadan yapıldığını kabul etmenin mümkün olmayacağını" belirterek yapay zekâ ile ortaya çıkan işlerin telif haklarına konu olamayacağını belirtmiştir [9]. Yapay zekânın temeli algoritmalara dayanmaktadır. Bu algoritmalar insanlar tarafından yazıldığı gibi kimi zaman da yapay zekâ becerisine bağlı olarak bilgisayar sistemleri de kendi algoritmalarını yazabilmektedir [5]. Makinelerin kendi algoritmalarını yazmaları neticesinde ortaya çıkan işler de özgün olabilir ya da telif haklarına konu olabilir. Bu açıdan Amerika'daki Telif Hakları Bürosu'nun görüşleri de yetersiz kalmaktadır.

Yapay zekâ düzenlemelerine ön ayak olması amacıyla 2016 yılının Ekim ayında Amerika hükümeti yapay zekânın kullanımının artması üzerine strateji raporu hazırlamak için toplanmıştır. Bu çerçevede, "*Yapay Zekânın Geleceği İçin Hazırlık (Preparing for the Future of Artificial Intelligence*)" isimli bir rapor ve "*Ulusal Yapay Zekâ Araştırma ve Geliştirme Strateji Plânı (National Artificial Intelligence Research and Development Strategic Plan*)" isimli bir strateji plânı yayınlanarak yapay zekâ düzenlemelerine ilişkin temelin oluşmasına zemin hazırlanmıştır [9].

Yapay zekâ düzenleme tartışmalarına Elon Musk ve Mark Zuckerberg gibi isimler de dâhil olmuştur. Elon Musk'a göre yapay zekâ düzenlemeleri reaktif normları ile değil; proaktif yani önlem alıcı normlarla kurulmalıdır [9]. Yapay zekâ düzenlemelerinin yapılması için artık geç olduğunu belirten Musk, genelde toplumun başına kötü bir şey gelince düzenlemelerin yapıldığını ve yapay zekâ konusunda da kötü olaylar yaşanmadan önlem alıcı düzenlemeler yapılması gerektiğini savunmuştur [9]. Mark Zuckerberg ise yapay zekânın tamamen zararsız olduğunu savunmuştur. Zuckerberg'in yapay zekâyı savunmasının üzerinden kısa bir süre geçtikten sonra Facebook'taki yapay zekâ sistemlerinin kapatıldığı haberi pek çok yerde yapıldı. Facebook'taki yapay zekâ sistemlerinin kapatılmasının sebebi olarak da İngilizce dilini kullanmak suretiyle birbirleriyle haberleşen yapay zekâ rehber robotlarının ("chatbot") aniden kendi iletişim biçimlerini geliştirmeleri ve insanlar tarafından anlaşılmayan bir dille haberleşmeleri gösterilmiştir [9]. Yaşanan bu durum aslında yapay zekânın limitinin neredeyse sınırsız olduğuna ve insan kontrolü dışında birtakım eylemlerde bulunabileceklerine örnek olarak gösterilebilir.

Yapay zekâ, "kamu düzeni", "yasal" ve "etik" alanlarına ilişkin birtakım sorunları da beraberinde getirmektedir. Yapay zekâlara ilişkin hukuki düzenlemelerin yapılabilmesi için sorumluluk hukuku çerçevesinde hangi süjelerin sorumlu olabileceği konusunda özellikle ceza hukuku kapsamında çeşitli tartışmalar mevcuttur. Genel kabul gören görüş doğrultusunda yapay zekâ yapay zekânın *kullanıcıları, üreticileri* veya *sahipleri* arasında üçlü bir ayrım yapılarak sorumluluk değerlendirilmektedir [5]. Yapay zekânın hukuki statüsüne ilişkin olarak herhangi bir tanım olmamakla beraber, "*Sorun yapay zekânın kendisinden kaynaklanıyor ise yapay zekâ da sorumlu tutulabilir mi?*" sorusuna herhangi bir yanıt verilememektedir. Örnek olarak, robotik cerrahi ile yapılan ameliyat esnasında hasta ölürse kimin sorumlu olacağı konusu genellikle belirsizdir. "Ölen hastanın yakınları robota mı, robot yazılımını yapan şirkete mi; yoksa hatalı hekim uygulamasından (malpractice) dolayı cerraha mı dava açmalıdır?" sorusu yanıt beklemektedir; sorunun cevabı şartlara göre değişkenlik gösterebilir; herhangi bir yeknesaklık mevcut değildir [6]. Aynı şekilde, otonom veya yarı-otonom araçların kaza yapması durumunda kimin kusurlu kabul edileceği ve sorumluluğun kime ait olacağı hususu tartışmalıdır. Arabayı yapan şirket, arabanın ve navigasyonun yazılımını yapan şirket ve arabada bulunan yolcu arasından kimlerin ne şekilde sorumlu tutulacağı tartışmalıdır. Örnek olarak, Google'ın insansız aracı Kaliforniya'da kaza yapmıştır ve bu kazaya ilişkin rapor çerçevesinde, "*aracın algoritmasında kum sebebiyle sapmanın meydana geldiği ve aracın otobüse çarptığı"* belirtilmiştir [10]. Kaza sonucunda herhangi bir ölüm ya da yaralanma meydana gelmemiştir; yani herhangi bir zarar oluşmamıştır. Herhangi bir zararın meydana gelmesi durumunda sorumluluk hukuku açısından kimin sorumlu olacağı, sigorta konusu bakımından sigortanın kapsamında kabul edilip edilmeyeceği, ne tür yaptırımların uygulanabileceği gibi sorular ortaya çıkacaktır [10].

Bir başka tartışmalı olan husus ise insanların dışarıdan müdahalesi ile yapay zekâdan beklenilmeyen sonuçların alınması, beklenmeyen tepkilerin verilmesi gibi durumlarda söz konusu olacak cezai sorumluluktur. Bu tür dış faktörlere karşı da gerekli önlemlerin alınması gerektiği gibi, yapay zekâya herhangi bir müdahalede bulunan ve farklı sonuçların oluşmasına neden olan kişilerin cezai sorumluluğu da söz konusu olmalıdır. Örnek olarak, 1981 yılında Japonya'daki bir motosiklet fabrikasında çalışan 37 yaşındaki bir işçi yakınında çalıştığı robot tarafından robotun işini engellediği gerekçesiyle çalışan makineye itilerek öldürülmüştür [11]. Hukuken failin cezalandırılabilmesi için suçun hem maddi hem de manevi unsurlarının bulunması gerekmektedir; yani hem hukuka aykırı bir fiilin gerçekleşmesi sonucunda bir zarar meydana gelmeli, hem de bu hukuka aykırı eylemin bilerek ve istenerek yerine getirilmesi gerekmektedir. Yapay zekâ kendisine yüklenen bilgiler doğrultusunda hareket edeceği için yapay zekânın iradesinin mevcut olup olmayacağı, bağımsız hareket edip etmeyeceği, suç işleme saikinin bulunup bulunmayacağı konular da tartışmaya açıktır. Suç işleme saiki kavramının tartışılması çerçevesinde ilk görüşü 1942 yılında Isaac Asimov robotlar açısından "Runaround" isimli kısa hikâyesinde ele almıştır ve ilk kez Robot Bilimi ("Robotics") kavramını ortaya atarak Robotiğin Üç Kanunu'nu vurgulamıştır. Bu çerçevede [5];

- Birinci Kanun: Robot herhangi bir insana zarar veremez, ya da herhangi bir insanın zarar görmesine izin veremez.
- İkinci Kanun: Robot, insanın emirlerine uymak zorundadır yeter ki emirin içeriği birinci kanunda belirtilen husus ile çelişmesin.
- Üçüncü Kanun: Birinci ve İkinci Kanun ile çelişmediği müddetçe robot kendi varlığını korumalıdır.

Asimov'un görüşleri robotların hukuki statüsünü belirlemek açısından faydalı olsa da; robotların zarar kavramını her yönüyle değerlendirerek zarar verebilecek eylemleri tespit edebilme yetisine sahip olduğunu kabul etmek yerinde olmayacaktır [5]. Her ne kadar insan gibi davranabilen robotlar geliştirilmiş olsa da robotların özellikle etik konusunda robotların insan ile eşdeğer seviyede düşünme yetisine sahip olacağı düşüncesini kabul etmek pek mümkün değildir.

Yapay zekânın kullanım alanlarının artmasıyla beraber kişisel verilerin de bu sistemler aracılığıyla işlenmesi ve saklı tutulması söz konusu olmuştur. Kişisel verilerin korunması bakımından da yapay zekâ konusunda düzenlemelerin yapılması önem arz etmektedir. Örneğin, Avrupa Birliği Veri Koruma Tüzüğü ("GDPR") kapsamında Avrupa Birliği ("AB") vatandaşları şirketlerden kendileriyle ilgili bilgi talep edebilmektedir. Bir AB vatandaşının kredi talebi reddedilmişse, bunun sebepleriyle ilgili kişinin bilgi talebi olabilir. Kişiye kredi verilmemesi sonucuna ulaşan yapay zekânın bu sonuca ulaşırken hangi algoritmaları ve hangi verileri kullandığı, hata yapıp yapmadığı gibi konular gündeme gelecektir [8]. Yapay zekâ kapsamında büyük veriler tutulabildiği için bu verilerin güvenliği, başka yerlere sızmaması gibi konular önem teşkil etmektedir ve yapay zekânın bu bakımdan kontrol edilebilmesi gerekmektedir. Kişisel verilerin ihlâli durumunda ağır yaptırımlar söz konusu olmaktadır ve yapay zekâların bu konuda dikkatli kullanılması gerekmektedir.

Yapay zekâ tek bir tanımı olmamakla beraber akıllı sözleşmeler, nesnelerin interneti, büyük veri, gelişmiş algoritmalar gibi uçsuz bucaksız kavramları içerisinde barındırmaktadır. Hukuk kurallarının oluşması için öncelikle tanımın yapılması büyük önem arz etmektedir. Yapay zekâya ilişkin tek bir tanım yapmaya çalışmak yapay zekânın kullanım alanlarını sınırlandırabilir ve kullanım kapasitesini de azaltabilir. Bu nedenle, yapay zekâya hukuki düzenlemeler getirilmesi konusunda teknoloji ile uğraşan şirketlerin de birtakım çekinceleri söz konusudur. Düzenlemelerin yapay zekâ kullanımını sınırlandırabileceği, mahkemelerin ve yasa koyucuların kuralları katı bir şekilde yorumlayabileceği nedenleriyle şirketler, düzenleme yapılmasını tercih etmemektedirler [9].

III. SONUÇ VE ÖNERİLER

- Henüz yapay zekâya ilişkin bir düzenleme bulunmadığı için yapay zekâya ilişkin çıkabilecek ihtilafların çözüm noktası çoğunlukla muğlak kalmaktadır. Ne tür hukuk

kuralları ve etik kurallarının doğru kabul edilebileceği ve bunlara kimin karar vereceği konusu tartışmaya açık konulardan bir tanesidir [12].Teknoloji sürekli ilerlediği için dinamik yapıya sahip yapay zekâya ilişkin düzenleme yapılması zorlaşmaktadır. Hızla ilerleyen teknoloji çerçevesinde getirilen düzenlemelerin yetersiz kalması söz konusu olacaktır ve sürekli güncelleme gerektirecektir.

- Yapay zekânın hukuki statüsü ve herhangi bir ihtilâf yaşandığında sorumluluğun ve yaptırımların ne şekilde olacağı konusunda mevcut düzenleme bulunmamakla birlikte, doktrinde çeşitli tartışmalar mevcuttur. Özellikle kişisel verilerin korunması, ceza hukuku ve fikri mülkiyet hukuku bağlamında yapay zekâ düzenlemeleri önem teşkil etmektedir.
- Hukuki düzenlemelerin yapılması açısından pek çok tartışma ve görüş mevcut olmakla beraber, teknolojinin de hızla ilerlemesi sebebiyle yapay zekâya ilişkin çalışmaların yapılması ve gerekli önlemlerin alınması son derecede önem arz etmektedir. Düzenlemelerin hem meydana gelebilecek zararlı sonuçları önlemesi hem de zararlı sonuçlar meydana geldikten sonra sorumluların belirlenmesi ve yaptırım yapılması amacıyla kurallar içermesi gerekmektedir.
- Teknoloji ve yapay zekâ çalışmalarının çok hızlı bir şekilde gelişmesi, insanlığın refah seviyesini artırmasının yanısıra bugün itibariyle düzenlemeleri yapılmayan konularda da birtakım hukuki sorunlar doğuracağı açıktır. Bu bağlamda, ülkemizde kamunun bu değişim ve gelişim sürecini çok yakından takip ederek gerekli düzenlemeleri yapmak için hazırlıklı olması gerekmektedir. Yapay zekâya ilişkin raporların ve strateji planlarının hazırlanması bu konuda atılabilecek önemli adımlardan bir tanesi olacaktır.
- Üniversitelerde özellikle hukuk fakültelerinde ve sosyal bilimler alanlarında bu konuya yönelik derslerin verilmesi, insan kaynağının yetiştirilmesi şimdiden çok önemli hâle gelmiştir. Yapay zekâ konusunda teknik olarak da bilgilerin verilerek işin doğası gereği ne şekilde düzenlemelerin yapılabileceği yolunda bu şekilde önemli adımlar atılabilir.
- Siber güvenliğe yönelik saldırıların giderek artması ve saldırıların boyut değiştirmesi ile yapay zekâ kavramı daha da önem kazanmıştır. Siber güvenlikte yapay zekânın kullanımına ilişkin çalışmaların artırılması özellikle saldırıların önceden tespiti ve bu saldırıların önlenmesi açısından önemlidir.
- Üçüncü kişilerin müdahalesi neticesinde yapay zekânın uygulamasına ilişkin zararlı sonuçlar doğabilir. Müdahalede bulunan kişilerin cezai sorumluluğuna ilişkin caydırıcı cezaların öngörülmesi gerekmektedir.
- Yapay zekâya ilişkin mevcut bir düzenleme olmamakla birlikte gerek etiksel olarak gerekse hukuki ve teknik olarak çeşitli görüşler ortaya atılmaktadır. Bu görüşler arasında henüz bir uzlaşma yoktur ve düzenlemenin nasıl yapılması gerektiğine ilişkin evrensel kabul gören bir çözüm yöntemi söz konusu değildir. Hem ulusal hem uluslararası boyutlarda çeşitli hukukçuların, uygulamadaki teknik bilgiye sahip kişilerin ve etik konusunda uzman kişilerin bir araya gelmesi suretiyle düzenlemelerin yapılması açısından gerekli adımlar atılmalıdır.

KAYNAKÇA

[1] "Endüstri Tarihine Kısa Bir Yolculuk", *Endüstri 4.0 Platformu*, 14 Aralık, 2017. [Çevrimiçi]. Erişim adresi: http://www.endustri40.com/endustri-tarihine-kisa-bir-yolculuk/ [Erişim tarihi: 5 Eylül 2018].

[2] B. J. Copeland, "Artificial intelligence," *Encyclopædia Britannica*, 17 Ağustos, 2018. [Çevrimiçi]. Erişim adresi: https://www.britannica.com/technology/artificial-intelligence. [Erişim tarihi: 12 Eylül 2018].

[3] A. Hintze, "Understanding the Four types of AI, from reactive robots to self-aware beings," *The Conversation*, 14 Kasım 2016. [Çevrimiçi]. Erişim adresi: http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings Erişim Tarihi: 20 Eylül 2018].

[4] B. Marr, "How AI and machine learning are transforming law firms and the legal sector," *Forbes*, 23 Mayıs 2018. [Çevrimiçi]. Erişim adresi: https://www.forbes.com/sites/bernardmarr/2018/05/23/how-ai-and-machine-learning-are-transforming-law-firms-and-the-legal-sector/#63ebb28d32c3. [Erişim tarihi: 12 Temmuz 2018**].**

[5] I. Giuffrida, F. Lederer, and N. Vermerys, "A legal perspective on the trials and tribulations of ai: how artificial intelligence,the internet of things, smart contracts, and other technologies will affect the law", *Case Western Reserve University School of Law Scholarly Commons*. 2018. [Çevrimiçi]. Erişim adresi: https://scholarlycommons.law.case.edu/caselrev/vol68/iss3/14/. [Erişim tarihi: 19 Eylül 2018].

[6] M. Goodman, *Future Crimes-Inside the Digital Underground and the Battle for Our Connected World*, UK: Penguin Random House, Corgi Books, 2016, pp.465-512

[7] R. M. Catea, "Challenges of the not-so-far future: eu robotics and ai law in business", *Challenges of the Knowledge Society.Private Law*, Vol 12, Mayıs, pp 213-216, 2018.

[8] K. Firth-Butterfield, "Artificial intelligence and the law: more questions than answers?" *Scitech Lawyer*, Vol. 14, Mayıs, pp. 28-31, 2017.

[9] J S. Azadian, and G. M. Fahy, "Artificial intelligence and the law: navigating known unknowns," *The Computer & Internet Lawyer*, Vol. 35, pp 19-23, 2018.

[10] "Yapay Zeka ve Hukuk," *Yapay Zeka Hukuku*. [Çevrimiçi]. Erişim adresi: https://yapayzekahukuku.com/practice-areas/car-accidents/. [Erişim tarihi: 13 Eylül 2018].

[11] H. Gabriel, *When Robots Kill : Artificial Intelligence Under Criminal Law*, Northeastern University Press, 2013. [E-book] Erişim adresi: ProQuest, https://ebookcentral.proquest.com. [Erişim tarihi: 17 Eylül 2018]

[12] H. Gabriel, "The criminal liability of artificial intelligence entities - from science fiction to legal social control," *Akron Intellectual Property Journal*, Vol. 4, pp. 171-201, 2010.

Siber Güvenliğin Sağlanması İçin Saldırı Simülasyonu Modeli

Attack Simulation Model for Ensuring Cyber Security

Onur AKTAŞ
Bilgisayar Mühendisliği Bölümü,
Hacettepe Üniversitesi
sec@onuraktas.net

***Abstract*— In this study, we propose a model for more effective use of cyber security systems and prevention of information security violations by using cyber attack simulation to measure human resource readiness and cyber security infrastructure. All information security measures used as software and hardware shall be adapted to the information system used for maximum benefit. This cyber security measures can be inadequate over time due to factors such as constant changes, the increasing number of cyber attacks, the diversity of the underlying infrastructures, human resources especially in large information infrastructure. With the proposed model, it is evaluated that human resources can be used more effectively in addition to security measures.**

***Index Terms*— Cyber security, cyber attack, cyber security infrastructure, cyber threat**

***Özet*— Bu çalışmada siber güvenlik sistemlerinin en etkin şekilde kullanılması ve bilgi güvenliği ihlallerinin en iyi şekilde önlenmesi için insan kaynağının ve siber güvenlik önlemlerinin siber saldırıları simülasyonları ile kontrol edilmesini önermekte olup bunun için bir model öne sürülmektedir. Yazılım ve/veya donanım olarak kullanılan tüm bilgi güvenliği önlemleri doğru entegrasyon için kullanıldığı bilişim sistemine uygun hale getirilmelidir. Uygun hale getirilen siber güvenlik önlemleri özellikle büyük bilişim sistemlerinde, sık sık yapılan değişiklikler, siber saldırıların gittikçe artması, kullanılan bilişim alt yapılarının çeşitliliği - dolayısıyla daha fazla atak vektörünün olması - , insan kaynağı gibi etkenlerden dolayı zaman içerisinde yetersiz kalabilmektedir. Önerilen model sayesinde güvenlik önlemlerinin yanında insan kaynağının da daha etkin şekilde kullanabileceği değerlendirilmektedir.**

***Anahtar Kelimeler*— Siber güvenlik, siber saldırı, siber güvenlik alt yapısı, siber tehdit**

I. Giriş

Bir kurumun güvenliği açısından siber güvenlik altyapısının sağlamlığının önemi bilinmekle birlikte bunu ölçmenin kolay bir yolu yoktur. Güvenilir ve gerçeğe yakın siber saldırı simülasyonu bunu yapmanın en kolay yoludur. Hem yapılan araştırmalar sayesinde hem de son günlerde yaşanan siber saldırılar incelendiğinde görülmektedir ki doğru yapılandırmalar ve güvenlik önlemlerinin doğru şekilde kullanılması ile siber saldırıların büyük çoğunluğu önlenebilecektir. İlgili bilgiler çerçevesinde, siber saldırıları önlemek için bu çalışmada siber saldırı simülasyonlarının kullanılmasına yönelik model önerilmektedir.

Çalışma içerisinde ikinci bölümde temel siber güvenlik önlemleri için yazılım veya donanım olarak çalışabilen sistemlerden bahsedilmiştir. Üçüncü bölümde problemin ana kaynağı olan, siber güvenlik önlemlerinin doğru kullanılmamasının nedenlerine ve konu ile alakalı literatür çalışmalarına yer verilmiştir. Dördüncü bölümde ise siber saldırganların özellikle büyük bilişim sistemlerine sızmak, kontrol etmek ve bilgi kaçırmak için kullandığı yöntemlere kısaca değinilmiştir. Sonuncu ve beşinci bölümde ise saldırı simülasyonu için önerilen modelin detayları bahsedilmiştir.

II. Güvenlik Sistemlerine Genel Bakış

Siber saldırılar ile alakalı kurum veya kuruluşlar çeşitli stratejiler, süreçler ve güvenlik önlemleri kullanmaktadır. Süreçler konusunda her ne kadar farklılıklar olsa da önleyici, tespit edici ve düzeltici olarak gruplanabilirler [1]. Önleme amaçlı kontroller, olası bir tehdidi henüz gerçekleşmeden enleme amacını, tespit etmeye yönelik kontroller siber saldırı başarı olsa bile izleme ve alarm sistemleri sayesinde tespit edebilme amacını, düzeltici önlemler ise siber saldırı sonrası işlerin devam etmesini sağlama amacını taşımaktadır [2].

Önleme amaçlı uygulanan siber güvenlik adımları güvenlik duvarı, kullanıcı erişim kontrolü, güvenlik sıkılaştırmaları, zararlı yazılım koruması ve yama yönetimi olarak sıralanabilir [3]. Bu adımlar bilgi güvenliği ihlallerinin yaşanmasını engellemek için en önemli aşamaları içermektedir. Ağ tabanlı güvenlik duvarları, web ağ geçitleri, kum havuzları ve uygulama tabanlı güvenlik duvarları siber saldırıları önleme amaçlı kullanılan sistemlerden bazılarıdır. Ağ tabanlı güvenlik duvarları siber saldırıları uygulama güvenlik duvarlarına göre daha alt katmanlarda önleme yapmaktadır.. Uygulama güvenlik duvarları kullanıcılar seviyesinde, trafik üzerine uygulamalar için özel politikalar belirlenmesi görevini getirmektedir [4]. Bu politikalar özellikle web uygulamalarına gelen siber saldırıları önleme amacıyla kullanılmaktadır. Kum havuzları şüpheli

dosyaları çalıştırarak zararlı tespiti için kullanılmaktadır. Web ağ geçitleri son kullanıcı web trafiği üzerinde kontroller yapmak için kullanılmakta bu sayede son kullanıcının zararlı web sitelerine girmesi engellenebilmektedir.

Siber saldırılar sonrası olay analizi (tespiti) için daha farklı güvenlik önlemleri bulunmaktadır. Bu güvenlik önlemleri sayesinde olayın kaynağı, saldırganların kullandığı yöntemler, etkilenen sistemler, yaşanan veri kayıpları tespit edilebilir. Yapılan bir araştırmaya göre aşağıdaki bazı güvenlik yatırımlarına ait özellikler olay tespiti sırasında en fazla öneme sahip yatırımlardır [5]. Bu yatırımlar aşağıda belirtilmektedir.

- IPS//Firewall/UTM Alarmları
- Log Analizleri
- Güvenlik ve Olay Bilgisi Yönetimi
- Web Ağ Geçidi

Düzeltici önlemler, bir siber saldırının başarılı olmasından sonraki adımları içermekte ve bu nedenle önem kazanmaktadır. Sistemlerin saldırganla bağının kesilip kesilmeyeceğinin karar verilmesi, etkilenen sistemlerin hangi yöntemlerle geri kurtulacağının tespit edilmesi düzeltici önlemlere örnek olarak verilmektedir.

III. Güvenlik Yatırımlarının Doğru Kullanılamaması

Önleyici, tespit edici veya düzeltici süreçlerin hangi aşamasında kullanılırsa kullanılsın güvenlik önlemlerinin hiçbiri dahil olduğu sistemde yüzde yüz bir güvenlik sağlayamazlar. İlgili yatırımın doğru şekilde yapılandırılması, sistemlerin takip edilmesi, başarılı ve başarısız siber saldırılarına analizi, insan kaynağının yetkin ve yeterli şekilde kullanılması gibi maddeler güvenlik yatırımlarının başarısını doğrudan etkileyen maddelerdir.

Sistemlerin bir kez başarılı bir şekilde yapılandırılması da yeterli olmamaktadır. Özellikle büyük ölçekli birçok yazılım ve donanımın dahil olduğu bilişim sistemleri birçok değişikliğin olduğu oldukça canlı sistemlerdir. Bu sistemlerde yeni yazılımların ve donanımların kurulması, yeni kullanıcı ihtiyaçları veya artan kullanıcı sayısı nedeniyle oldukça sık görülebilmektedir. Aynı zamanda, farklı takımların ağ üzerindeki farklı yerlere erişim isteği, güvenlik duvarı gibi ağ katmanında kurallar tanımlanan cihazlarda birçok kuralların değişmesine neden olmaktadır. Bunun gibi yaşanan örnekler oldukça sık değişen yapılandırmaların oluşmasına bu nedenle de güvenlik ihlallerinin önlenmesi ve tespit edilmesine yönelik eksiklikler ortaya çıkabilmektedir.

Bilişim sistemlerinde güvenliğin ancak ve ancak en zayıf halka kadar olduğu bilinen bir gerçektir. Yanlış yazılmış kurallar, unutulmuş eski kullanıcı hesapları, eksik yapılandırmalar gibi kullanıcı hataları, büyük maliyetler ile yapılan siber güvenlik yatırımlarını boşa çıkarabilmektedir. Örneğin web uygulama güvenlik duvarında yapılan küçük bir yapılandırma hatası, ilgili web uygulamasına yapılan tüm siber saldırıların güvenlik tarafından kontrol edilememesine neden olabilmektedir. Aynı şekilde ağ güvenlik duvarlarının doğru yapılandırılamamasından kaynaklı eksiklikler ele geçirilmiş sistemlerin saldırgan ile bağlantı kurabilmesine neden olmakta, dolayısıyla siber saldırının başarılı olmasına neden olmaktadır. Gartner adlı bir firmanın yaptığı çalışma içerisinde, 2020 senesindeki güvenlik duvarlarını geçmenin %99 ihtimalle yapılandırma hatasından kaynaklanacağı yer almaktadır [6]. Bu göstermektedir ki yanlış yapılandırmalar nedeniyle siber saldırıların başarılı olma oranı artmaktadır.

Özellikle tespit edici önlemler diğer önlemlere göre daha fazla insan kaynağına ihtiyaç duymaktadır. İnternet üzerinden birçok zafiyet tarama, port tarama, sömürü tetikleme, servis tespiti gibi işlemler otomatik olarak yapılmaktadır. Port tarama yazılımları 3 dakika içerisinde tüm interneti bir port için tarayabilecek seviyeyi geçmiş bulunmaktadır [7]. Son zamanlarda bulaştığı sistemlerde otomatik olarak tarama yaparak başka yerlere bulaşan zararlı yazılımlar tespit edilmiştir. Mirai olarak bilinen ve benzer işlemlere sahip zararlı yazılım, internet üzerine birçok taramayı otomatik olarak yapmaktadır [8]. Bu da göstermektedir ki bilişim sistemlerine yönelik şüpheli işlemler internet sürekli düzenli olarak yapılmaktadır. Tüm bu zararlı aktiviteler tespit edici güvenlik yatırımlarında (iz kayıtları, güvenlik olay bilgisi yönetimi) iz bırakabilmekte ve/veya alarm üretebilmektedir. Üretilen iz kayıtları ve alarmların çok yüksek boyutta olması incelemeyi zorlaştırdığı için daha önemli siber saldırıların tespit edilmesini zorlaştırmaktadır. Bu nedenle siber saldırıları kategorilendirilerek, anormal davranışları iz kayıtlarından çıkarabilecek insan kaynağının, tespit edici önlemler almak için büyük önem taşıdığı değerlendirilmektedir.

Yapılandırma hataları yapılan bir çalışmaya göre Google'a ait ana servislerin hataya düşmesinin en önemli ikinci sebebidir, ve yanlış yapılandırmaların önemli zararlar verdiğine dair çalışmalar bulunmaktadır [9] [10]. 2016 yılında yapılan bir çalışma ise siber güvenlik alanında çalışanların direnç, ilgisizlik, bilgi eksikliği, farkındalık eksikliği, yanlış çalışma ve ihmal gibi nedenlerden dolayı bilgi güvenliği ihlallerine neden olabileceği göstermektedir [11]

İster insan kaynağı isterse yanlış/eksik yapılandırmalardan dolayı oluşabilecek zafiyetlerin önceden tespit edilmesi bilgi güvenliği ihlallerini tespit etmek için büyük önem taşımaktadır. Yetkin insan kaynağı ile birlikte, yüksek maliyetlere çıkan siber güvenlik yatırımlarının en doğru şekilde kullanılmasının siber saldırılarının engellenmesi ve tespiti konusunda büyük fayda sağlayacağı değerlendirilmektedir. Her ne kadar siber güvenlik sistemlerinde, güvenlik duvarlarına yönelik politikalarına yönelik anomali tespitlerine yönelik birçok çalışma [12,13,14,15,16] ve web uygulama güvenlik duvarlarının sıkılaştırmalarına ve eğitilmesine yönelik çalışmalar [17,18,19] yapılmış olsa da siber güvenliğin bir bütün olarak değerlendirilmesi gerekmektedir. Geçmişte yapılan çalışmalar [22,23,24] bilişim sistemlerinin belirli bir alanına yapılan çeşitli siber saldırıları simüle edilmesini önermektedir. Bu kapsamda ihtiyacın giderilmesi için bilişim sistemlerini bir bütün olarak ele alınmasının ve düzenli olarak güncel zararlı yazılım ve siber saldırı tekniklerinin simüle edilmesi önerilmektedir.

IV. SİBER SALDIRILARA HIZLI BİR BAKIŞ

Siber saldırganlar hedef sisteme sızarak kendi istekleri doğrultusunda sistemi kullanabileceği gibi sistemi çalışmaz hale getirebilecek saldırılar da düzenleyebilir. Saldırganlar hedef sistemle bağlantı kurmak için komuta kontrol adı verilen uzaktan yönetilebilir bir sisteme ihtiyaç duyarlar. Komuta kontrol merkezleri ile hedef olan sistem arasında bilinen protokoller üzerinden (HTTP, FTP, IRC) veri alışverişi olabileceği gibi farklı platformlar (Twitter) veya daha alt seviyedeki OSI katmanlarını (TCP) kullanarak veri alışverişi sağlanabilir. Komuta kontrol merkezleri yalnızca bir IP adresi olabileceği gibi alan adları üzerinden de gerçekleştirilebilir. Zararlı yazılımlar içerisinde komuta kontrol merkezinin alan adını tespit eden algoritmalar kullanılabilir [20] .

Bazı durumlara veri alışverişi yerine hedef sistemden yalnızca veri kaçırma işlemi de gerçekleştirilebilir. Veri alışveriş yöntemlerine ek olarak veri kaçırma için saldırganlar bilgi kaçırmak için bazı protokolleri amacı dışında kullanarak (DNS tünelleme, ICMP tünelleme) ilgili protokolün paketleri içerisinde de veri kaçırabilmektedir. Tünelleme olarak bilinen bu yöntem güvenlik cihazlarını atlatmak veya gizlenmek için kullanılabilmektedir.

Siber saldırgan hedef sisteme bir kez erişim yaptıktan sonra yetki dahilinde sistemler üzerinde bir kullanıcının yapabileceği her işlemi ve daha fazlasını gerçekleştirilebilir. Amaçladığı saldırı için yetkisinin düşük olması durumunda yetki yükseltme için sistemleri zorlayabilir. İsteği doğrultusunda sistemi çalışmaz hale getirebilir (Stuxnet örneği [21]), sistemde kalıcı olabilir, verileri silebilir, düzenleyebilir veya ele geçirebilir.

Eğer saldırı erişimi engellemeye yönelik (servisleri çalışmaz duruma getirme) yapılıyorsa bir komuta kontrol merkezi ile veri alışverişine veya veri kaçırılmasına ihtiyaç duyulmaz. Uygulama , servis veya ağ katmanında hedef sisteme siber saldırılar düzenlenebilir.

Hedef sisteme ilk erişim oltalama saldırıları, uygulama zafiyetleri, servis bazlı zafiyetler, insan faktöründen dolayı oluşan zafiyetler, sunucu zafiyetleri, işletim sisteminde bulunan zafiyetler, mobil cihazlarda bulunan zafiyetler ya da bir başka atak vektörü kullanılarak yapılmış olabilir. Özellikle hedef sistemlerin internete açık tüm alt yapıları (web uygulamaları, eposta sistemleri) siber saldırganlar için öncelikli hedef olabilmektedir.

Kötü niyetli kullanıcılar tarafından kullanılan atak vektörlerinin, veri kaçırma/alışverişi için kullanılan yöntemlerin ve diğer zararlı aktivitelerinin bilinmesi durumunda durumunda benzer siber saldırılı simülasyonları gerçekleştirilebilir. Bu simülasyonlar ise hem insan kaynağını eğitmen hem de siber güvenlik altyapısını test etmek için kullanılabilir.

V. ÖNERİLEN SALDIRI SİMÜLASYON MODELİ

Bu bölümde önerilen saldırı simülasyon modeline yönelik ana kategoriler, saldırı senaryoları ve simülasyon iş akışı konuları anlatılmıştır.

A. Simülasyon Sistemi Ana Kategorileri

Saldırı simülasyon sistemi altı ana kategoriden oluşmaktadır. Bu kategoriler aşağıda verilmektedir.

- Ağ işlemleri veritabanı
- Zararlı aktivite veritabanı
- Ağ işlemleri simülatörü
- Zararlı aktivite simülatörü
- Sanal komuta kontrol sunucusu
- Trafik Analiz Sunucusu

Ağ işlemi veritabanı içerisinde siber saldırganlar tarafından kullanılan hedef sisteme uzaktan yapılan siber saldırılara ait senaryo bilgileri bulunmaktadır.

Zararlı aktivite veritabanı, saldırganın hedef sisteme bir kere sızdıktan sonra yapacağı işlemler barındırmaktadır.

Veritabanları içerisinde modüller halinde kod parçaları bulunmakta olup ilgili simülatör tarafından talep edilerek çalıştırılırlar. Her kod parçacığı son kullanıcı tarafında zararlı bir aktivite oluşturmakta ya da uzak hedefe siber saldırı paketleri göndermektedir.

Ağ işlemleri simülatörü, ağ işlemi veritabanı içerisinde bulunan modülleri çalıştırarak hedef sisteme yönelik saldırı senaryolarını gerçekleştirmekle ve yönetmekle sorumludur.

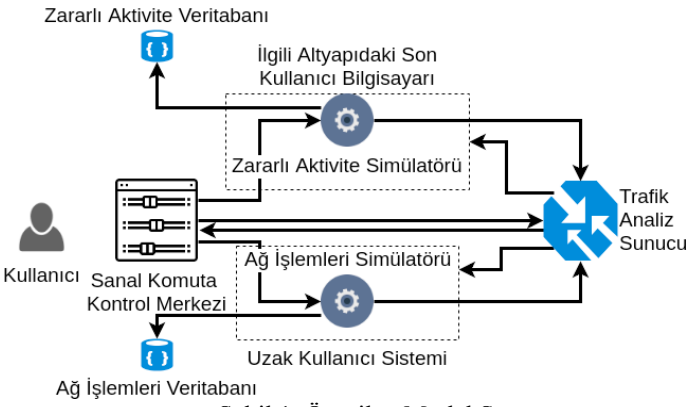
Zararlı aktivite simülatörü benzer şekilde zararlı aktivite veritabanı içerisindeki saldırı simülasyonunu hedef sistem içerisinde çalıştırır ve yönetir.

Simülatörler yalnızca sanal komuta kontrol sunucusu ile bağlantı kurup kendisine gönderilen komuta göre ilgili veritabanından modülü alarak çalıştırmakla yükümlüdür. Simülatörler son kullanıcı veya siber saldırganların kullandığı bir işletim sistemine kurulmalı ve Sanal Komuta Kontrol Sunucusu ile bağlantı sağlamalıdır. Son kullanıcı bilgisayarına kurulan simülatör ilgili sistemin iç ağına dahil edilmeli ve herhangi bir kullanıcı tarafından kullanılmayan bir sistem olmalıdır. Ağ işlemleri simülatörü ise ilgili sistemin internete açık varlıklarına yönelik saldırı simülasyonu gerçekleştireceği için dışarıda bir sisteme kurulması yeterlidir.

Sanal komuta kontrol sunucusu, sistemi kullanan kişiye bir arayüz sunar. Bu arayüzden sistemi kullanan kişi simüle etmek istediği siber saldırıları ilgili simülatörlere gönderebilir. Arayüz üzerinden aynı zamanda simüle edilen siber saldırı hakkında bilgi verilerek gerekli kontrollerin yapılmasını sağlar.

Trafik analiz sunucusu veritabanları içerisinde bulunan modüller için gerekli bilgileri sağlamakla ve hedef sistemden çıkarılan verileri kontrol etmekle yükümlüdür. Simülatörler tarafından çalıştırılan zararlı yazılım aktivitesi sonucunda bir veri alışverişi veya veri kaçırılması söz konusu ise trafik analiz sunucusu istenen veriyi iletir ya da verilerin çıkarılması için gerekli alt yapıyı oluşturup dinleme işlemi yapar. Eğer başarılı şekilde bir veri alışverişi ya da veri çıkarma işlemi yapıldıysa ilgili simülatörü ve sanal komuta kontrol sunucusunu bilgilendirir.

Tüm sisteme ait örnek şekil aşağıda paylaşılmaktadır.



Şekil 1. Önerilen Model Şeması

B. Saldırı Senaryoları

Siber saldırı simülasyon sistemi ilgili sistemde beş farklı siber saldırı senaryonu kategorisinde zararsız siber saldırı gerçekleştirilebilmektedir. Bu kategoriler Zararlı Aktivite ve Ağ İşlemleri veritabanında bulunan kod parçacıkları (modüller) sayesinde yapılabilecek saldırı simülasyonlarının kategorileridir ve aşağıda verilmiştir.

- Veri kaçırma yöntemleri
- Veri alışveriş yöntemi
- İnternete açık uygulama saldırıları
- Zararlı son kullanıcı aktiviteleri
- Zararlı iç ağ aktiviteleri

Olası bir siber saldırı sonrasının başarılı olması durumunda hedef sistem ile veri alışverişi veya hedef sistemden veri kaçırma yöntemleri simüle edilebilmektedir. Tünelleme, veri kaçırma (HTTP, FTP, Twitter, vb..) modülleri ile bu yöntem test edilebilir.

Veri kaçırma yöntemleri hedef sistemden herhangi bir veriyi kaçırma senaryoları içermektedir.

Veri alışveriş yöntemi ise hedef sisteme bir kere sızıldıktan sonra zararlı aktiviteleri yönetmek için hedef sistemdeki simülatör ile bağlantı kurmaya yönelik senaryoları barındırır. Bu senaryolar zararlı yazılımların komuta kontrol merkezleri ile yapacağı bağlantıları da kapsamaktadır.

İnternete açık uygulama saldırıları daha üst seviye de internete açık sistemlere yönelik saldırı senaryolarını kapsamaktadır.

Zararlı son kullanıcı veya zararlı iç ağ aktiviteleri ise hedef sistemde son kullanıcı veya siber saldırganlar tarafından kullanılabilecek saldırı simülasyonlarını içermektedir.

Tüm senaryolar sayesinde siber saldırıların olası yöntemleri ve etkileri ilgili sistem üzerinde test edilebilecektir.

C. Simülasyon İş Akışı

Önerilen saldırı simülasyonu yöntemine göre iş akışları aşağıda verilmiştir.

Testin yapılacağı sisteme zararlı aktivite simülatörü kurulur ve simülatörün sanal komuta kontrol merkezi ile trafik analiz sunucusuna erişim yapılması sağlanır.

Kullanıcı sanal komuta kontrol merkezinden seçtiği veri kaçırma veya veri alışverişi simülasyonunu modülüne ait bilgileri simülatöre gönderir (örneğin DNS tünelleme ile veri kaçırma)

Simülatör ilgili modüle ait kod parçacığını zararlı aktivite veritabanından talep eder.

Eğer modül içerisinde veri alışverişi veya veri kaçırılması için bir uzak sistem talep ediliyorsa trafik analiz sunucusundan bunu talep eder (örneğin DNS trafiği kontrol edilebilen alan adları).

Trafik analiz sunucusu istenen sistemi ayağa kaldırarak ilgili bilgileri simülatöre ve sanal komuta kontrol merkezine gönder (örneğin tekil olarak oluşturulmuş alt alan adları ve DNS sunucusu)

Simülatör, trafik analiz sunucusundan aldığı veriler ile kod parçacığını çalıştırır, trafik analiz sunucusuna cevap döner.

Trafik analiz sunucusu hem simülatörden aldığı verileri hem de ayağa kaldırdığı sistemi kontrol ederek veri alışverişinin veya veri kaçırma işleminin başarılı olup olmadığını kontrol eder (örneğin gelen DNS trafiğinin içerisinde simülatörden aldığı verinin olup olmadığını).

Veri alışverişi veya veri gönderme işlemi başarılı olduysa, sanal komuta kontrol merkezini uyarır.

Sanal komuta kontrol sunucusu kullanıcıya saldırı sonucunu, sıkılaştırma önerilerini ve saldırı detaylarını paylaşır.

İnternete açık uygulamalara yönelik siber saldırılar da benzer iş akışı ile çalışır. Fakat bu önerilen modelde, internetten yapılan saldırıların etkisi ilgili sistemde bulunan bir istemci ile ölçülmediği için kullanıcı karşısına sadece üretilmesi gereken iz kayıtları ile birlikte saldırı detayları paylaşılır. Kullanıcı kendi sisteminde internete açık uygulama veya servislere göre modüller çalıştırmadır (örneğin FTP servisi varsa: kaba kuvvet, bilgi toplama, port tarama modülleri gibi). İnternet üzerinden gerçekleştirilen saldırı simülasyonu port tarama, kaba kuvvet, bilgi toplama, flood modülleri ile gerçekleştirilebilmektedir.

Zararlı son kullanıcı aktiviteleri ve zararlı iç ağ aktiviteleri aynı veri kaçırma veya veri alışverişi modülünde olduğu gibi test edilen sistemin son kullanıcı bilgisayarları (ağa dahil olan herhangi bir sistem) üzerinde çalışmalıdır. Yetki yükseltme, iç ağ port tarama, antivirüs atlatma, kalıcı olma, işletim sistemi hakkında bilgi toplama, kullanıcı ekleme, zararlı yazılım yükleme modülleri ile bu kategorideki saldırı simülasyonları gerçekleştirilebilir.

Saldırı simülasyonları sayesinde hem olası bir siber saldırı öncesi benzer bir saldırı yöntemi ile insan kaynağının eğitilmesi sağlanabilmekte hem de mevcut siber güvenlik altyapısına yönelik eksiklikler tespit edilebilmektedir. Düzenli olarak saldırı simülasyonunu yapılması sürekli olarak değişen büyük bilişim alt yapılarında hızlı tespit olanağı sağlayacağı değerlendirilmektedir.

VI. Sonuç ve Gelecek Çalışmalar

Geliştirilen yeni saldırı simülasyonu modelinde, bilişim sistemleri alt yapılarında siber güvenlik önlemleri olarak

kullanılan yazılım ve/veya donanımların daha etkin şekilde kullanılması ve sıkılaştırmalarının güncel siber saldırıları önleyecek şekilde yapılması ile siber güvenlik önlemlerini kullanan insan kaynağının etkin şekilde kullanılmasına yönelik önerilerde bulunulmuştur. Güncel siber saldırı senaryoları ile mevcut sistemlerin test edilmesi, gerçekleştirilen testlerin literatürde bulunan sıkılaştırma önerileri ile karşılaştırılması, ilgili sistemlere yeni sıkılaştırma önerileri ile birlikte katkıda bulunması gelecek çalışmaların konusudur.

TEŞEKKÜRLER

Makalenin hazırlanması aşamasında destek ve katkıları dolayısıyla Prof. Dr. Ali Aydın Selçuk'a teşekkür ederim.

KAYNAKÇA

[1] https://www.giac.org/paper/gsec/501/information-security-process-prevention-detection-response/101197 (Erişim Tarihi; 01.06.2018)

[2] A. Bendovschi, "Cyber-attacks–trends, patterns and security countermeasures", Procedia Economics and Finance, 28, ss. 24-31.

[3] A. Fielder, E. Panaousis, P. Malacaria, C. Hankin, & F. Smeraldi," Decision support approaches for cyber security investment", Decision Support Systems, 86, 13–23, 2016

[4] N. Khochare, & B. B. Meshram, "Tool to Detect and Prevent Web Attacks", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 1(4), pp-375, 2012.

[5] https://www.sans.org/reading-room/whitepapers/analyst/show-on-2017-incident-response-survey-37815 (Erişim Tarihi; 01.06.2018)

[6] https://www.gartner.com/doc/2254717/brand-firewall-best-practice-enterprises (Erişim Tarihi; 01.06.2018)

[7] https://blog.erratasec.com/2013/09/masscan-entire-internet-in-3-minutes.html (Erişim Tarihi; 01.06.2018)

[8] C. Kolias, G. Kambourakis, A. Stavrou, & J. Voas "DDoS in the IoT: Mirai and other botnets", Computer, 50(7), 80-84, 2017.

[9] T. Xu, J. Zhang, P. Huang, J. Zheng, T. Sheng, D. Yuan, S. Pasupathy, "Do not blame users for misconfigurations", In Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles (pp. 244-259), 2013.

[10] L. A. Barroso and U. H¨olzle, "The Datacenteras a Computer: An Introduction to the Design of Warehouse-Scale Machines", Morgan and Clay pool Publishers, 2009.

[11] N. S. Safa, & C. Maple,"Human errors in the information security realm–and how to fix them", Computer Fraud & Security, 17-20, 2016.

[12] E. Al-Shaer and H. Hamed, "Discovery of Policy Anomalies in Distributed Firewalls," IEEE INFOCOM '04, vol. 4, pp. 2605-2616, 2004

[13] J. Alfaro, N. Boulahia-Cuppens, and F. Cuppens, "Complete Analysis of Configuration Rules to Guarantee Reliable Network Security Policies," Int'l J. Information Security, vol. 7, no. 2, pp. 103- 122, 2008.

[14] F. Baboescu and G. Varghese, "Fast and Scalable Conflict Detection for Packet Classifiers," Computer Networks, vol. 42, no. 6, pp. 717-735, 2003.

[15] L. Yuan, H. Chen, J. Mai, C. Chuah, Z. Su, P. Mohapatra, and C. Davis, "Fireman: A Toolkit for Firewall Modeling and Analysis," Proc. IEEE Symp. Security and Privacy, p. 15, 2006.

[16] H. Hu, G. J. Ahn, & K. Kulkarni, "Detecting and resolving firewall policy anomalies" IEEE Transactions on dependable and secure computing, 9(3), 318-331, 2012.

[17] D. Appelt, A. Panichella, & L. Briand, "Automatically repairing web application firewalls based on successful SQL injection attacks", In Software Reliability Engineering (ISSRE), 2017 IEEE 28th International Symposium on (pp. 339-350). IEEE, 2017.

[18] J. J. Singh, H. Samuel, & P. Zavarsky," Impact of Paranoia Levels on the Effectiveness of the ModSecurity Web Application Firewall", In Data Intelligence and Security (ICDIS), 2018 1st International Conference on (pp. 141-144). IEEE, 2018.

[19] D. Appelt, C. D. Nguyen, A. Panichella, & L.C. Briand, "Machine-Learning-Driven Evolutionary Approach for Testing Web Application Firewalls", IEEE Transactions on Reliability, (99), 1-25, 2018.

[20] A. K. Sood, & S. Zeadally, "A taxonomy of domain-generation algorithms". IEEE Security & Privacy, 14(4), 46-53, 2016.

[21] C. Baylon, "Lessons from Stuxnet and the Realm of Cyber and Nuclear Security: Implications for Ethics in Cyber Warfare", In Ethics and Policies for Cyber Operations, 213-229, 2017.

[22] M. E. Kuhl, J. Kistner, K. Costantini, & M. Sudit, " Cyber attack modeling and simulation for network security analysis". In Proceedings of the 39th Conference on Winter Simulation: 40 years! The best is yet to come (pp. 1180-1188). IEEE Press, 2017.

[23] K. Pan, A. Teixeira, C. D. López & P. Palensky, "Co-simulation for cyber security analysis: Data attacks against energy management system", In Smart Grid Communications (SmartGridComm), IEEE International Conference on (pp. 253-258) , 2017.

[24] M. McDonald, J. Mulder, B. Richardson, R. Cassidy, A. Chavez, N., Pattengale, ... & R. Halbgewachs, " Modeling and simulation for cyber-physical system security research, development and applications". Sandia National Laboratories, Tech. Rep. Sandia Report SAND2010-0568, 2017.

DRDoS Yükselticileri Üzerine Ülke Çapında Bir Çalışma

A Nationwide Study of DRDoS Amplifiers

Emre Murat Ercan
Dept. of Computer Engineering
TOBB University of Economics and Technology
Ankara, Turkey

Barikat Internet Security, Ankara, Turkey
emremuratercan@yandex.com

Ali Aydın Selçuk
Dept. of Computer Engineering
TOBB University of Economics and Technology
Ankara, Turkey
aliaydinselcuk@gmail.com

***Abstract*—In recent years distributed reflective denial-of-service (DRDoS) attacks have given an important advantage for attackers. Because of the UDP protocols' nature and some insufficient hardenings attackers can achieve hundreds of Gb/s bandwidth easily. They can even hit more than a Tb/s as we have seen in the GitHub attack. Attack nature is simple. Attacker botnets spoof victims' IP address and request amplifiable services as him. Servers will send some gigantic amount of data as response. At that point, one key solution is to harden the servers running these amplifiable services.**

In this paper we focused on three UDP-based protocols, namely NTP, DNS, and Memcached. All these three protocols are already used by attackers in DRDoS attacks. We aimed to find servers in Turkey running these protocols. After that we researched whether these servers are available for being used as amplifiers in DRDoS attacks.

***Index Terms*— DDoS, DRDoS, Amplification Attack, NTP, DNS, Memcached.**

***Özet*—Son yıllarda Dağıtık Yansıtmalı Hizmet Engelleme Saldırıları saldırganlara önemli bir avantaj sağlamaktadır. UDP tabanlı protokollerin güvensizliği ve bazı eksik sıkılaştırmaların sonucu olarak saldırganlar saniyede yüzlerce gigabit bant genişliği elde edebilir. Hatta bu bant genişliği GitHub saldırısında saniyede 1 terabitin üstüne çıkmıştır. DRDoS saldırısında botnetler; kurbanların IP adreslerini taklit eder ve yükseltmeye el verişli servisleri onlar gibi talep eder. Talep edilen servislere cevap olarak çok büyük miktarda veri kurbana döner. Bu noktada temel çözümlerden bir tanesi yükseltici olarak çalıştırılabilen sunucuların sıkılaştırılmasıdır.**

Bu makalede daha önce saldırganlar tarafından DRDoS saldırılarında kullanılmış olan UDP tabanlı üç protokol (NTP, DNS, Memcached) üzerine odaklandık. Çalışmada Türkiye'de bu protokolleri çalıştıran sunucuları bulmayı hedefledik. Sunucular bulunduktan sonra DRDoS saldırılarında yükseltici olarak kullanılabilirliğini araştırdık.

Anahtar Kelimeler—DDoS, DRDoS, Yükseltme Saldırıları, NTP, DNS, Memcached

I. Introduction

Distributed denial-of-service (DDoS) attacks are one of the best-known and oldest attack types which aim to stop services. Attackers can aim every industry like universities, online services and even political sites. DDoS attacks have a simple logic and there are so many ways to perform it such as using malicious botnets and request thousands of get connections (GET FLOOD) [9]. Adversaries can aim to exhaust bandwidth and other resources [16].

Distributed reflective denial-of-service attack is a new generation of DDoS attack. It combines the same logic with amplification power. Relatively few sources can be dangerous as classical DDoS with an amplification power. The adversary directly targets bandwidth in DRDoS attacks with botnets. Botnets impersonate victim and request services which generally use UDP-based protocols. UDP-based protocols are one of the most important key components in DRDoS, which is also known as UDP-Based Amplification Attacks [8]. By its nature, UDP-based protocols are connectionless, and there is no handshake process that would help to mitigate IP spoofing. First discussion about spoofed IP usage from adversaries rose to the surface in 1989 [15]. In reflection attacks, while requests come from botnets, responses will be sent back to the victim. To make things worse, service responses could be as big as 51000 times the size of the request in worst cases [8]. In Figure 1, we present the nature of DRDoS attacks.

The potential of a DRDoS attack is determined by the "byte amplification factor (BAF)", which is defined as the rate of response byte to rate of request byte. Also another factor is the "packet amplification factor" which is defined as the ratio of response packet number to request packet number (PAF) [4]. BAF is more dangerous than the PAF, because BAF can reach gigantic numbers. For this reason, most of former studies defined "amplification factor" as BAF. In our study we also use the amplification factor as BAF.

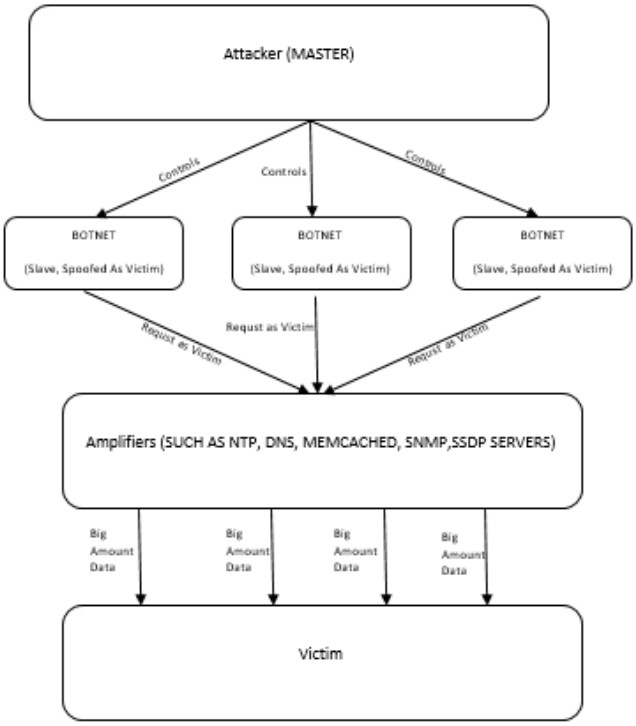
Some of the most vulnerable protocols for DRDoS attacks are NTP, SNMP v2, DNS, NetBios, SSDP, CharGen, QOUTD and Memcached because of their high amplification factors [4]. Some early studies showed that NTP and DNS could be extremely harmful for service providers. DNS amplification

factor could be up to 76x in worst cases. As an average amplification factor, open DNS resolvers' respond with more than 28x. NTP has a much more terrifying amplification factor as 556x to 4670x in some cases [4]. Also these studies showed us amplification attacks could be still useful with a 79x amplification factor with TCP based protocols [19], but in our study this fact is out of scope. Beyond these studies we have witnessed the greatest DDoS attack to GitHub. That attack was also DRDoS attack which was done using Memcached [17].

In this study we focused on DNS, NTP and Memcached servers in Turkey. The DNS protocol is used for translating domain names to IP addresses. This protocol uses both TCP and UDP port 53. NTP is used for time synchronization. Synchronization could be server to server or server to client. Memcached is a system to use distributed memory object caching which can run on both TCP and UDP port 11211. We found these three services in IPv4 domain and we requested some services which already attackers use to decide if these servers allow adversaries to attack victims on them. This study analyzed the state of the servers in Turkey running these protocols.

Organization of this paper is as follows: In Section 2, we give brief information about DNS, NTP and Memcached protocols and give some well-known occurred attacks' examples with these protocols. We explain our discovery methods in Section 3. In Section 4 we represent poorly managed DNS, NTP and Memcached servers and hardening methods for server administrators. After indicating hardening methods, we conclude the paper in Section 5.

Figure 1: Nature of DRDoS Attacks.



II. Related Work

This work is inspired by the amplification research by Rossow et al. [4]. Rossow et al. researched on 14 UDP-based protocols and discovered these protocols potential amplification factors and countermeasures. Their study is one of the most definitive papers about DRDoS terms. Another paper by the same group warned about TCP protocols as amplifiers and tackled with NTP servers [5]. Kuhrer et al. worked on potential TCP amplification factors and countermeasures. Their paper showed how TCP could be compromised for DRDoS attacks [19]. Also real life experiments showed us how UDP based protocols could be dangerous without early warnings [6,7,17].

Over the last few years DRDoS attacks have been become more and more dangerous. Adversaries started to generate massive traffic volume with amplification attacks. While the first warnings about usage of UDP based protocols such as DNS with spoofed IP for DDoS attacks in 1999 [12], first noticeable DRDoS attack was performed in 2012 [2] and then DRDoS attacks became popular in 2013. First spectacular DRDoS attack in 2013 was performed by the abuse of SNMP protocol. After that day, there were four more DRDoS attacks with a volume exceeding 100 Gb/s in 2013.

DNS is one of the most important protocols used by attackers. Not only its amplification factor is significant, but also there are so many servers running this protocol on the Internet. The protocol works both TCP and UDP at port 53. In DNS the largest amplification factor appears with the "ANY" query. This query demands all known information about a DNS zone in a single request [11]. Causing denial of service with DNS servers is identified as CVE-2006-0987 [22] and CVE-2006-0988 [23]. In 2015 Turkey suffered from a DNS amplification attack. That attack came to be known as the "nic.tr attack" [6]. There was also another well-known attack as the "Dyn attack" in 2016 [7]. This attack was similar to the nic.tr attack. They were both done by the same group with the similar tactics, techniques and procedures.

The NTP protocol is simply used for time synchronization. This protocol runs at UDP port 123. Time synchronization has a key role for exploring attacks. Most defense systems need NTP to understand if there is a real incident. NTP is also significant for DRDoS attackers. Its responses can be 4670 times larger than the request. That amplification factor occurs in "monlist" request [13]. This request normally is used for server's monitor data. As a response server sends back last 600 clients IP addresses and some more information such as these clients' NTP mode and version information. Causing denial of service with NTP servers is also identified as CVE-2013-5211 [14]. As a "Cloudflare" announcement their customers were attacked 400 Gb/s in February 2014 with an NTP amplification attack [1].

The Memcached protocol is designed to speed up dynamic database-driven websites. With uses of key/value pairs it reduces database loads [24]. The protocol works on both TCP and UDP at port 11211. Causing denial of service with Memcached servers is also identified as CVE-2018-1000115 [25]. According to CVE-2018-1000115 Memcached

amplification factor could be up 10,000x to 51000x [8,17]. Before the last day of February 2018, Memcached was not that much a hot topic. But the attack on GitHub which peaked at more than 1.3 Tb showed us this protocol could be so harmful for service providers [17].

III. Methodology

In this paper we focused on these three main protocols, NTP, DNS, and Memcached, and the servers in Turkey running these protocols. First we worked on finding country-based IPv4 addresses. There are quite a few different IPv4 databases but we choose "Ivan Erben's database because we decided that his script gives the most recent country base IPv4 addresses [10]. This study has been done according to the file dated 26.05.2018.

As the second phase of study we found these 3 protocols on IPv4 domain. For this purpose, we used "zmap" [3]. Zmap is a project which is written for fast internet scanning.

We started our study with discovering NTP servers. We scanned all Turkey IP addresses with specialized NTP module in zmap. After zmap exploring was executed, we used results as input of "nmap". Nmap is a well-known tool for internet scanning. We used one of the nmap scripts that was written to obtain information about stated IP addresses.

While we were working on Memcached, we executed two respective commands. The first command was executed for finding Memcached servers in Turkey. For this purpose, we scanned all stated IPs with zmap. In this scan we used a specialized module for Memcached. After that phase, we kept digging to conceive much more harmful servers. For this examination we used nmap script that gives all information about Memcached server.

Our third research was concentrated on DNS. First, we started discovering DNS servers. For this research we used zmap again. We have done an aggressive research and scanned all IPs without using any module. Only restriction was target port at that scan. As the second phase of our DNS study, we prepared a script to find servers that allows both "ANY" request and recursive query. We used "nslookup" to obtain all information records.

IV. Results and hardening methods

According to our research there are 52,139 NTP servers in Turkey. 1,336 of them are available for monlist request. 1,315 of them give information only about associated servers, public clients, and private clients' IP information. These are still amplifiers with a smaller amplification factors. On the other hand, 21 of these servers give information on all associations between the server and clients. This is the case where the amplification factor can be as high as 4670x.

In Turkey, 13,359 servers run as Memcached servers. Generally, Memcached servers do not have to be open for Internet. But it is impossible to be sure if these servers should be open for internet without communication with the server administrator. On the other hand, we observed 508 of them still uses UDP port 11211.

According to our research 91,552 DNS servers are available at port 53. Most of these servers refused our query. They will not be an amplifier of DRDoS with using "ANY" request. On the other hand, 13,730 of these servers responded and gave all recorded information about our requests.

These three protocols can be hardened by their administrators. NTP monlist requests can be disabled with two methods [20,21]. The first method is updating NTP servers. It is always recommended to use a protocol's most up-to-date version possible. Since NTP server version 4.2.7, monlist request is disabled by default. Also monlist request can be disabled by manual configuration [14]. In early versions of Memcached protocol both TCP and UDP ports were enabled. But after the GitHub attack, the newer version (1.5.6) was published and the amplification problem was fixed by disabling UDP port 11211 by default. Also if there is no mandatory need for internet communication, this server should be closed to internet access. For DNS servers, one of the most important prevention is to whitelist IP restriction for "ANY" queries [18]. With this restriction, only a few IP addresses can perform this query, and others requests will be refused and dropped. Also recursive queries should be restricted [11].

V. Conclusion and future works

As the internet usage increases, adversaries' knowledge for different attacks increases even faster. Our study should motivate service administrators to harden their servers. It is well-known that attacking is easier than defending. This is why network administrators should be wary of their systems. As we mentioned earlier, they should use up-to-date services as much as possible. If this option is not possible, they should harden their systems for not being a part of an attack.

Acknowledgments

We would like to thank Bahtiyar Bircan, Kamil Seyhan, and Sertaç Katal from Barikat Internet Security for helping us improve our discovery methods.

References

[1] M. Prince. Technical Details Behind a 400 Gbps NTP Amplification DDoS Attack. https://blog.cloudflare.com/technical-details-behind-a-400gbps-ntp-amplification-ddos-attack/ , February 2014.

[2] Prolexic Quarterly Global DDoS Attack Report Q2 2013, "Prolexic Stops Largest-Ever DNS Reflection DDoS Attack," May 2013. [Online]. https://sm.asisonline.org/ASIS%20SM%20Documents/Prolexic%20Quarterly%20Global%20DDoS%20Attack%20Report.pdf.

[3] Z. Durumeric, E. Wustrow, and J. A. Halderman. ZMap: Fast Internet-wide Scanning and Its Security Applications. In Proceedings of the 22nd USENIX Security Symposium, Washington, D.C., USA, August 2013.

[4] C. Rossow. Amplification Hell: Revisiting Network Protocols for DDoS Abuse. In Symposium on Network and Distributed System Security (NDSS) (2014).

[5] M. Kührer, T. Hupperich, C. Rossow, and T. Holz. Exit from Hell? Reducing the Impact of Amplification DDoS Attacks

Proceedings of the 23rd USENIX Security Symposium, San Diego, USA, August 2014.

[6] 14/12/2015 Tarihinde Başlayan DDoS Saldırısı Kamuoyu Duyurusu. https://www.nic.tr/2015-12-DDoS-Saldirisi-Kamuoyu-Duyurusu-20151221.pdf . 21 Dec 2015.

[7] S. Hilton. Dyn Analysis Summary Of Friday October 21 Attack https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/. 26 Oct 2016.

[8] CERT Advisory, "UDP-Based Amplification Attacks" https://www.us-cert.gov/ncas/alerts/TA14-017A

[9] A. Buscher, T. Holz. Tracking DDoS Attacks: Insights into the Business of Disrupting the Web. In Proceedings of the 5th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET), San Jose, CA, USA, April 2012.

[10] I. Erben. http://www.iwik.org/ipcountry/TR.cidr

[11] CERT Advisory, "DNS Amplification Attacks" https://www.us-cert.gov/ncas/alerts/TA13-088A

[12] A. AusCERT, "Domain Name System (DNS) Denial of Service (DoS) Attacks," August 1999. [Online]. Available: http://www.securityfocus.com/advisories/1727

[13] F. J. Ryba, M. Orlinski, M W¨ahlisch, C. Rossow, T. C. Schmidt. "Amplification and DRDoS Attack Defense – A Survey and New Perspectives". arXiv:1505.07892v3 [cs.NI] 17 May 2016.

[14] CERT Advisory, "NTP Amplification Attacks Using CVE-2013-5211" https://www.us-cert.gov/ncas/alerts/TA14-013A

[15] S. M. Bellovin, "Security Problems in the TCP/IP Protocol Suite," ACM SIGCOMM Comput. Commun. Rev., vol. 19, no. 2, pp. 32–48, Apr. 1989.

[16] S. M. Specht, R. B. Lee. Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures. In Proceedings of the International Conference on Parallel and Distributed Computing (and Communications) Systems (ISCA PDCS), San Francisco, CA, September 2004.

[17] S. Kottle, February 28th DDoS Incident Report, https://githubengineering.com/ddos-incident-report/ , March 2018.

[18] Use DNS Policy for Applying Filters on DNS Queries. https://docs.microsoft.com/en-us/windows-server/networking/dns/deploy/apply-filters-on-dns-queries, March 2018.

[19] M. Kuhrer, T. Hupperich, C. Rossow, T. Holz. Hell of a Handshake: Abusing TCP for Reflective Amplification DDoS Attacks. In Proceedings of the 8th USENIX Workshop on Offensive Technologies, San Diego, CA, August 2014.

[20] Team Cymru. Secure NTP Template. https://www.team-cymru.com/secure-ntp-template.html

[21] J. Graham-Cumming. Understanding and Mitigating NTP-Based DDoS Attacks. https://blog.cloudflare.com/understanding-and-mitigating-ntp-based-ddos-attacks/

[22] National Vulnerability Database, "CVE-2006-0987 Detail" https://nvd.nist.gov/vuln/detail/CVE-2006-0987#vulnCurrentDescriptionTitle.

[23] National Vulnerability Database, "CVE-2006-0988 Detail" https://nvd.nist.gov/vuln/detail/CVE-2006-0988

[24] What is Memcached? https://memcached.org/

[25] Vulnerability Database, "CVE- 2018-1000115 Detail" https://nvd.nist.gov/vuln/detail/CVE-2018-1000115

Büyük Genomik Verilerde Mahremiyet

Privacy on Big Genomic Data

Enes Canbaz
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
enescanbaz@gmail.com

M. Emir Çakıcı
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
m.emircakici@hotmail.com

Yılmaz Vural
Kişisel Verileri Koruma Kurumu
Ankara, Türkiye
yilmazvural@gmail.com

Yavuz Canbay
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
yavuzcanbay@gazi.edu.tr

***Özet*—Günümüzde genomik verilerin mahremiyeti birçok bilim dalı açısından çok önemli bir noktaya gelmiştir. Genomik verilerin mahremiyetinin önemli olmasının en büyük sebebi bu verilerin sadece biz değil akrabalımız ve atalarımız hakkında da bilgiyi içermesidir. Bu çalışmada büyük genomik verilerde mahremiyeti ele alan çalışmalar incelenmiş, büyük genomik veride hassas veri parçaları, bu hassas parçaların korunması ve saklanması üzerine bir literatür çalışması yapılmıştır. Bu kapsamda kısa tandem tekrarları (STR), hastalıklar ile ilgili genler ve genetik varyasyonlar hakkında literatür taraması yapılmıştır. Genomik verinin mahremiyeti, bu mahremiyetin teknolojinin ilerlemesi ile birlikte nasıl korunacağı gibi konular üzerinde çalışmalar incelenmiştir. Sonuç olarak insan genom verilerinin incelenmesinde mahremiyet konusu üzerinde bilgisayar bilimlerinin ve tıbbi bilimlerin ortak şekilde çalışması gerektiği ve verilerin mahremiyet sorununun önemli ölçüde ortadan kaldırılmadan teknolojik olarak ilerlemenin kişiler açısından büyük bir sorun olduğu sonucu elde edilmiştir.**

Anahtar Kelimeler—hassas genom verisi, kısa tandem tekrarları, mahremiyet

***Abstract*— Hereditary and non-inherited DNA properties of a living thing are called genomic data. Nowadays, the privacy of genomic data has become a very important point for many scientists. The greatest reason why the privacy of genomic data is important is that these data are not only about us but also about our relatives and our ancestors. In this study, studies dealing with privacy of big genomic data were examined, sensitive fragments of big genomic data and the protecting and storing of these fragments were examined. In this context, short tandem repeats (STR), genes related to diseases and genetic variations were investigated. The study of such issues as the privacy of genomic data, how this privacy will be preserved along with the progress of the technology has been studied. As a result, the study of human genomes has revealed that privacy and medical science must work in common on privacy.**

Keywords—sensitive genome data, short tandem repeats, privacy

I. GIRIŞ

Büyük veri, farklı kaynaklardan elde edilen ve klasik yöntemler ile işlenemeyen tüm verilerin anlamlı ve işlenebilir hale getirilmesi problemi olarak adlandırılabilir. Büyük verinin oluşumunda etkili 5 farklı bileşen vardır. Bu bileşenler hız (velocity), veri büyüklüğü (volume), çeşitlilik (variety), doğrulama (verification) ve değer (value) olarak adlandırılır. Büyük veri ilk olarak astronomi ve genetik alanında ortaya çıkmış bir kavramdır. Genetik alanda büyük veriler üzerinde yapılan çalışmalar genom verileri üzerinde yapılmaktadır. Her canlıya ait genom verileri üzerinde kendine özgü genleri ve genetik elementleri incelenmektedir. Büyük veri teknolojileri sayesinde çok büyük kapasitelere sahip olan genom verileri işlenebilmekte, canlıların genetik özellikleri ortaya çıkarılabilmektedir.

Günümüzde genom verisinin kişilerle paylaşılması ve korunması arasındaki dengeyi kurmak en önemli zorluklardan birisi haline gelmiştir. Mahremiyetin korunması için yeni teknolojilerin veri paylaşımını olumsuz etkilemeyecek şekilde oluşturulması gerekmektedir. Bu çalışmada, büyük genom verileri üzerinde hassas parçaların mahremiyeti üzerinde literatür taraması yapılmıştır.

İnsan genomunun incelenmesi günümüzde birçok alanın ilgisini ve dikkatini çekmektedir. Genomik veriler birçok hastalığın tespiti ve tedavisi için çok miktarda bilgi içermektedir. Bu verilerin incelenmesi ve analizi hastalıkların genetik açıdan aktarımı ve kişi üzerinde erken tespiti için önemli rol oynamaktadır. Bütün genom sekanslaması işlemleri (Whole Gene Sequencing-WGS) bu sebeple sağlık ve büyük veri alanında devrim yaratacak bir teknoloji olarak görülmektedir. İnsan genomik verilerinin paylaşılması tıbbi buluşları hızlandırmasına karşın genomik verilerin mahremiyetine yönelik riskleri de artırır. Mahremiyetin korunması ile açık veri paylaşımı arasında geçmişten gelen bir gerilim vardır. Kişi kendi verisini kendi isteğiyle paylaşsa dahi bu veri sadece ona değil ona ve onun yakın akrabalarına ait verilerdir. Bu sebeple kişi aslında kendi genomik verisini paylaşırken bile tam söz hakkına sahip değildir. Bir insan genomu, sahibini eşsiz şekilde tanımlayabilir ve geçmiş, gelecek nesiller için bile kendisi ve yakınları hakkındaki bilgileri açığa çıkarabilecek kadar çok veriye sahiptir [1].

Verilerin mahremiyeti tıbbi araştırmalardaki hızı kesmesine karşın kişilerin özelini korumak büyük bir zorunluluktur. Sayısallaştırılmış genom verilerinin nerede saklanacağı, kimin tarafından saklanacağı, saklanılan yer ve saklayan kişinin güvenilir olup olmadığı önemli bir sorundur.

Genom verileri üzerinde hassas parçaların korunması bütün genom verisinin korunmasından daha önemli bir husustur.

Bu çalışmanın ikinci bölümünde büyük genomik veri mahremiyeti ile ilgili literatür çalışmaları özetlenmiş, üçüncü bölümde hassas genomik veri mahremiyeti üzerinde durulmuş ve son bölümde de yapılan araştırma ve incelemeler üzerine sonuç ve değerlendirmelere yer verilmiştir.

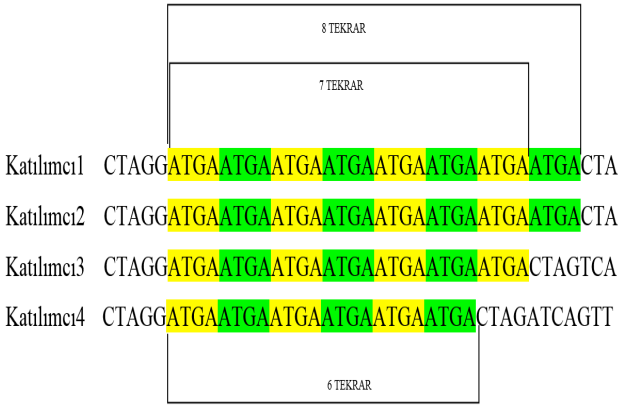
II. LITERATÜR TARAMASI

İnsanların genetik diziliminde bir insanı diğerinden ayıran genetik varyasyon %0.5 olduğu ve bu oranın korunmasının kalan %99.5'lik kısımdan daha önemli olduğunu bilmek ve bunu göz ardı etmemek zaruridir [2]. V.Cogo ve arkadaşları [1] büyük genomik veriler üzerinde hassas verilerin mahremiyetin korunması için bir metot önermiş ve bunun ile ilgili analizler yapmışlardır. Bir insan genomunun sahibini eşsiz şekilde tanımlayabildiği, kişinin geçmiş ve gelecek nesilleri hakkında bilgiler ortaya çıkardığı belirtilmiştir. Büyük genomik veriler üzerindeki hassas parçalar tüm genlere oranla %11.3 olarak tespit edilmiş ve bu hassas verilerin kısa tandem tekrarları, hastalıkla ilgili genler ve genomik varyasyonlar olarak tespit edilmiştir. Bu %11.3'lük verinin kalan %88.7'lik kısımdan ayrı saklanması ve farklı yöntemlerle korumayı öneren bir metot geliştirmişlerdir. Bu kapsamda, genel parçaları bulut gibi sistemlerde saklanması önerilirken hassas olan önemli verilerin homomorfik şifreleme yöntemleri ile daha güvenli şekilde saklanması gerektiğini belirtmişlerdir. Yazarlar, genom verilerinin bazı parçalarının kişinin isteği dâhilinde olmadan kişilik bilgilerini ortaya çıkarabilme potansiyeline sahip olduğunda bu verilerin mahremiyet duyarlı yani hassas veri olarak tanımlamışlardır. Yazarlar bu verilere saldırının üç şekilde olabileceğinden bahsetmiş ve bunların kısa tandem tekrarları, hastalıklarla ilgili genler ve genomik varyasyonlar olduğunu söylemişlerdir. V.Cogo ve arkadaşlarının önerdiği metot basitçe şu şekilde açıklanabilir; Önceden tanımlanmış DNA parçası verildiği zaman bu DNA segmentinin mahremiyet açısından bilinen bir bilgiye sahip olup olmadığını tespit edip daha önce yayınlanmış verilerden mahremiyet duyarlı olan nükleik ve amino asit sekanslarının seçilmesi ve bunların ayrı yerde saklanmasıdır. Fakat bu sistemde iki sorun olduğunu ve bunların doğruluk ve performans olduğunu belirtmişlerdir [1]

A. Kısa Tandem Tekrarları

Kısa tandem tekrarları (Short Tandem Repeats-STR) A,C,G,T karakterlerinden oluşan küçük tekrarlı DNA dizilimleridir. Genetik kimlik profilleri, adli kimlik, babalık testleri gibi konularda kullanılmaktadır. Son zamanlarda bazı saldırılar kişilerin bireysel genom verileri üzerinden soyadı ve soyağacı doğru tahmin edilmiştir. M.Gymnek ve arkadaşları [3], Y kromozomunda kısa tandem tekrarları oluşturarak ve rekreasyonel genetik veri tabanlarını sorgulayarak soyadlarının kişisel genomlardan alınacağından bahsetmiştir. Gymnek açıkladığı saldırıda her soyadı çıkarımının yaklaşık 16 kişinin mahremiyetini ihlal ettiğinden bahsetmiştir. Gymnek için %12'lik bir oran etkileyici olmamakla birlikte yürütülen 1000 Genomes projesine katılan 1092 katılımcının 131'inin mahremiyetini hayatı boyunca tekrar geri kazanamayacağını kabul etmiştir

H. Fan ve arkadaşları çalışmalarında [9] STR mutasyonu ile ilgili konulara odaklanmaktadırlar. STR, 100 nükleotid uzunluğa sahip seri oluşturan 1-6 bp'lik tekrarlayan bir bölüm içeren kısa süreli tekrarlanan DNA sekanslarıdır. Yazarlar STR'lerin insanlar dahil olmak üzere prokaryotlarda(çok hücreli) ve ökaryotlarda(tek hücreli) yaygın olarak bulunduğundan bahsetmişlerdir. Fan ve arkadaşları insan genomunun yaklaşık olarak tamamının eşit olarak dağıldığını, tüm genomun yaklaşık %3'ünü [10] STR'lerin oluşturduğundan bahsetmişlerdir. Yazarlar ortalama olarak insan genomunda 2,000 bp başına bir STR oluştuğundan, yüksek mutasyon oranları sebebiyle STR'lerin biyolojik araştırmalarda yaygın olarak kullanıldığından bahsetmişlerdir. Yayın kapsamında STR mutasyon modellerinde sonsuz aleller ve kademeli mutasyon modelinden bahsedilmektedir. Yazarlar tekrar sayısı, tekrar ünitesi, tekrar yapısı, tekrar biriminin temel bileşeni, rekombinasyon, STR'deki kesintiler, cinsiyet ve yaş STR mutasyonunu etkileyen faktörler olarak belirtmişlerdir. Fan ve arkadaşları STR'lerin daha fazla biyolojik işlevi keşfedilse dahi bunların çoğunun bilinmediğinden ve STR'lerin mutasyonu ve değişkenliğinin daha fazla araştırılması gerektiğinden bahsedilmektedir.



Şekil 1. Kısa Tandem Tekrarları (STR)

V.Cogo ve arkadaşları tarafından yapılan çalışmada incelenmesi gereken en önemli noktalardan birinin STR olduğu görülmektedir. Gymnek tarafından açıklanan saldırının devamının geleceği öngörülebilmektedir. V.Cogo ve arkadaşları bu tür saldırılar için geliştirdiği metotta bilinen tüm Y-STR(Y kromozom kısa tandem tekrarları)'leri veritabanına kaydederek ve bunları hassas olmayan çıktı akışından ayırarak koruyabileceklerinden bahsetmişlerdir [1].

Homer ve arkadaşları bireyin genomik DNA'sının çok sayıda kişiye ait genomik veriler içerisinde mevcut olup olmadığını çözümlemenin çok sayıda alanın ilgisini çekeceğinden bahsetmektedirler. Yazarlar genom verileri karışımlarını incelemekte kullanılan çok sayıda yöntem

olduğundan ve en yaygın olan yöntemin kısa tandem tekrarları (STR) olduğunu belirtmişlerdir. Y kromozomundaki STR'ler, sıklıkla DNA karışımının erkek bileşenlerini çözerken yararlıdır [18]. Homer ve arkadaşları, hiper değişken dizilemesine dayanan mitokondriyal DNA (mtDNA)'ın yüksek kopya sayısı nedeniyle bozulmuş DNA analiz ederken daha yararlı olduğunu belirtmişlerdir. Fakat mtDNA, tek değerli kalıtım ve düşük ayrımcılık gücü gibi zayıf yönlere sahiptir [17].

Mountain ve arkadaşları, yeni bir kombinasyon polimorfizmini, yani SNP (single nucleotid polimorphism) STR'leri geliştirdi, bu da nüfus tarihine dair bilgiler sağladı. Şu anda, STR bölgeleri, farklı bölgelerdeki nüfusların ilişkisini ve antik halkların göç yolunu ortaya çıkarmak için kullanılmaktadır [11].

Ruitberg ve arkadaşları çalışmalarında kısa tandem tekrarları hakkında bilgiler vermişlerdir. Yazarlar STR'lerin adli laboratuvarlarda artık popüler hale geldiğinden ve bunun sebebinin de düşük miktarda DNA'nın bozulmuş bir formda bile başarılı bir şekilde yazılabildiğinden bahsetmişlerdir. CODIS olarak adlandırılan FBI Birleşik DNA İndeks Sistemi hakkında bilgiler vermişler ve 13 adet STR işaretli çekirdek kümesi kullanıldığından bahsetmişlerdir. Bu sistem sayesinde suçlulardan ve suç mahalli kanıtlarından DNA profillerine bağlamada başarılı olunduğundan, aynı zamanda babalık testleri vakalarında yardımcı olmak için kullanıldığından bahsedilmiştir [12].

Gelfand ve arkadaşları, kısa tandem tekrarlarının nedensel olarak hastalıklar ile ilişkili olduğunun yaygın bir şekilde kabul edildiğinden bahsetmişlerdir. Hatta duygusal bozukluklar ve bağımlılık davranışlarının da kısa tandem tekrarları ile ilgili olduğu söylenmektedir. Mikrosatellitler bugün adli tıp alanında kullanılan DNA parmak izlerinin temelini oluşturmaktadır. Daha fazla sayıda olan SNP(single nucleotid polimorphism)'ler ile yapılan karşılaştırmalara rağmen mikrosatellitleri içeren polimorfik tandem tekrarları genetik test ve bağlantı analizinde önemli bir araçtır [13].

Butler ve arkadaşları, kişi genomundaki STR lokusunun fiziksel lokasyonu tanımladığını ve insan popülasyonlarında gözlemlenen alel aralıkları ve varyantları ebeveyn testinden gözlemlenmekte olan mutasyon oranları olarak özetlemektedirler. Aynı zamanda mevcut çekirdek lokuslara gelecekte eklenebilecek olan STR lokusları için istenen özellikler çalışma kapsamında tartışılmıştır. Yazarlar bu çekirdek STR lokusunun, gelecekte adli bilimlerde de önemli noktalara geleceğinden bahsetmişlerdir. Butler ve arkadaşları, kullanılan çekirdek lokusların ceza ve ebeveyn testinin kararlarında yararlı olduklarını göstermişlerdir [15].

B. Genetik Varyasyonlar ve Hastalıkla İlgili Genler

Bilim adamlarının insan genomlarından topladığı referans genom toplam insan genomunun %99.5'ini oluşturur. Kalan %0,5'lik kısım ise insanı eşsiz bir şekilde tanımlar. Bu %0,5'lik kısım toplam 3,2 milyar baz çiftinin arasından birkaç milyon nükleotide karşılık gelmektedir. Genetik varyasyon SNP gibi çeşitli biçimlerde ifade edilebilir. Ayday ve arkadaşları SNP'yi bir genom dizisinde nükleotidlerin pozisyon değiştirmesi olarak tanımlamışlardır. Araştırmacılar insanların 50 milyon benzersiz SNP'ye [4] sahip olduklarını ve bu SNP'lerin bireylerin bozuklukları ve hastalıklara yatkınlığını belirlemede yardımcı olduğu belirlemişlerdir. Bir SNP nükleotidinin diğer SNP nükleotidlerinin içeriğini ortaya çıkarmasının mümkün olabileceğini ve aynı zamanda bu durumun genom verilerinin mahremiyetinin korunmasını zorlaştıracağını vurgulamışlardır. Yapılan deney sonuçlarında genetik mutasyonların ilaç metabolizmasını değiştirdiğin ve bazı genomik testlerin hastanın hangi ilaçlara ne gibi tepki verdiğini tahmin etmede yardımcı olabileceğini göstermişlerdir [5].

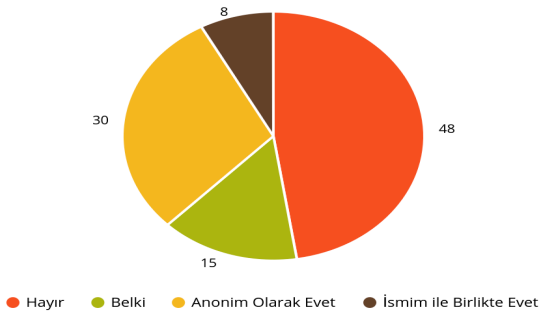
King ve arkadaşları çalışmalarında soyadı ve DNA arasındaki ilişkiyi incelediler. Kalıtsal soyadının babadan oğula geçtiği için soyadı ve Y-kromozomal haploitler arasındaki ilişkilere odaklandılar. Her ne kadar DNA her iki ebeveynden miras alınsa da Y kromozomunun rekombinan olmayan bir bölgesi olduğunu belirtmişlerdir. Bundan dolayı soyadının, baba atalarından miras kalmış olan Y kromozomu ile ilişkili olması beklenmektedir. Genetikte soyadların ilk uygulamasının nüfusun akrabalık derecesini tahmin etme üzerine kullanıldığını belirtmiş, çok çeşitli popülasyonların ve onların soyadlarının örneklenmesi sosyal demografik tarihe yeni ve ilginç bir bakış açısı getireceğini ifade etmişlerdir [14].

Genom çapında ilişkilendirme çalışması tarafından hâlihazırda tanımlanmış olan hastalık lokuslarında, lokus atfedilebilir risk genellikle şu an tahmin edilenden daha yüksek olmaktadır. Bunun nedeni, Genom çapında ilişkilendirme çalışmalarında kullanılan SNP'lerin, tipik olarak, asistanlık sinyaline yol açan gerçek nedensel mutasyon için kusurlu proksiler olacağıdır. Nedensel gen, sıklıkla hem yaygın hem de nadir görülen başlangıç işaretleyici SNP'leri tarafından etiketlenmemiş ek mutasyonlar içerecektir. Her genin katkısını belirlemek, her lokusta varyantların yoğun çalışmalarını gerektirecektir. Genom çapında ilişkilendirme çalışması ile daha birçok hastalık lokusu saptanmaya devam etmektedir. Yukarıda belirtildiği gibi, bugüne kadar Genom çapında ilişki çalışmaları istatistiksel olarak düşük bir güce sahip olmuştur ve bu nedenle de benzer ve daha küçük etkilerin ortak varyantlarıyla birçok lokusu kaçırmıştır. İlk çalışmalarda ortak yapısal varyantlar için proksiler yoktu ve düşük frekanslı ortak varyantları yakalayamadılar (% 0.5 ila% 5). Ayrıca, çalışmaların büyük çoğunluğu sadece Avrupa soylarının örneklerinde gerçekleştirilmiştir. Daha geniş, daha kapsamlı ve daha çeşitli GWAS'lar daha birçok yerel yeri ortaya çıkaracaktır [19].

C. Mahremiyet ile İlgili Çalışmalar

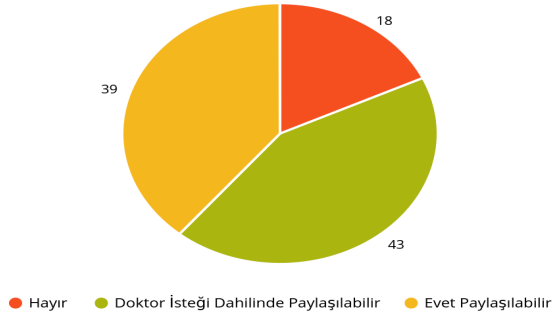
M.Naveed ve arkadaşları günümüzde genom dizileme teknolojisinin gelişmekte olduğunu ve artık detaylı genotipler üretmenin mümkün olduğunu vurgulamışlar, genetik verilerin toplanması ve analiz edilmesi ile kişi hakkında belirli hastalıklar ile ilgili bir ilişki ve aile ilişkilerinin açığa

çıkartılması gibi durumlarda kişinin mahremiyetine büyük etkilerde bulunulduğunu belirtmişlerdir. Yazarlar genomik verilerin mahremiyeti sorununun bilgisayar bilimi, tıp ve kamu politikasının ortak bir noktası olduğuna işaret etmişlerdir. Özellikle bilgisayar bilimcilerin genomik verilerin mahremiyeti ile daha çok ilgilendiklerinden bahsedilmiştir. Günümüzde genomik verilerin sağlık hizmetleri, biyomedikal araştırmalar, hastalık riski tespitleri vb. gibi daha fazla alanda kullanıldığına ve bu durumun güvenlik ve mahremiyet açısından birçok problem ortaya çıkardığına değinmişlerdir. Bu mahremiyet sorunun aşılamamasındaki en büyük sebeplerden birinin multidisipliner alanlarda çalışmanın zorluğu ve bu alanların belirli bir soruna ortak paydada buluşulamaması olduğunu söylemişlerdir. Bu başlık altında Naveed ve arkadaşları tarafından yapılan, katılımcıların %36'sının genom konusunda, %3'güvenlik konusunda uzman olduğu, tüm katılımcıların genom konusu hakkında bilgisi olduğu, katılımcıların %82'sinin güvenlik konusunda bilgisinin olduğu anketler yapılmıştır. Yapılan ilk ankete göre katılımcıların %20'si genomik mahremiyetin korunmasının imkânsız olduğuna inanmaktadır. Katılımcıların neredeyse yarısı kişinin genomik verilerinin diğer sağlık verilerinde farklı olmadığını düşünmektedir. Katılımcıların %7'si genomik mahremiyetin korunmasında konuyu biyoinformatik alanında uzman kişilere bırakılması gerektiğini düşünmektedir. Katılımcıların %20'si genom mahremiyetinin hukuki ve yasal yönlerle tamamen korunabileceğine inanmaktadır. Katılımcıların %7'si mahremiyet arttırıcı teknolojilerin genetik durumunda bir sıkıntı olduğunu düşünmektedir. Katılımcıların %10'una göre genomik verinin mahremiyetinin gereksiz olduğunu çünkü kişinin genomik verilerinden başka bir kişiyi tanımlamanın zor olduğunu düşünmektedir. Katılımcıların %30'u sağlık hizmetlerinde genomik verilerin incelenmesinden gelecek olan avantajların mahremiyet sorunlarının neden olacağı zarardan daha yüksekte kalacağı böylece mahremiyet sorununun göz ardı edilebileceğini düşünmektedir [7].



Şekil 2. Naveed ve arkadaşları tarafından yapılan *"Genom verilerinizi internette paylaşır mısınız?"*[7] anketi cevap oranları

Diğer bir ankette kişilerin genomik verilerinin webde herkese açık olarak paylaşıp paylaşmayacağı sorulmuştur. Şekil 2.'de de görüldüğü gibi katılımcıların cevapları ise aslında insanların durum ile ilgili bir endişesinin olduğunu ortaya çıkartmıştır. Katılımcıların sadece %8'lik bir kısmı kendi genomik verisini web üzerinde kendi ismi dahil olarak herkese açık bir şekilde paylaşabileceğini ve bunun bir sorun teşkil etmediğin söylemiştir. Katılımcıların yaklaşık %50'lik bir kısmı ise hiçbir şekilde verilerinin internet ortamında bulunmasına izin vermeyeceğini belirtmiştir. Katılımcıların %15'i kararsız yaklaşırken kalan %30'luk kısım ise anonimize edilmiş şekilde verilerinin kullanılabileceğini eğer anonim kalınacak ise bir sorun teşkil etmediğini belirtmiştir [7].



Şekil 3. Naveed ve arkadaşları tarafından yapılan *"Genom verilerinizin akrabalarınızla ilgili özel bilgi sızdırdığını varsayarsak, genom verilerinizi paylaşma hakkına sahip olduğunuzu düşünüyor musunuz?"* [7] anketi cevap oranları

Naveed ve arkadaşlarının yaptığı diğer bir ankette ise kişilerin genomik verilerinin akrabaları hakkında büyük miktarda veri sızdırdığı düşünülürse bir kişi kendi genomik verisini vermeye hakkının var olup olmadığının cevabı aranmıştır. Ankete göre Şekil 3'te de görüldüğü gibi kullanıcıların %39'u genetik verilerinin herkese açık bir şekilde paylaşılabileceğini belirtmiştir. %43'lük kısmı ise eğer doktorunun tıbbi bir amaç için ihtiyacı var ise paylaşabileceğini belirtmiştir. Kalan %18'lik katılımcılar ise durumun yanlış olduğunu ve verilerini asla paylaşmayacaklarını belirtmişlerdir. Diğer bir ankette ise genomik verinin mahremiyetini arttırmak için ne gibi önlemler alınabileceği ve bu konuda nasıl önerilerin olduğu sorularına cevap aranmıştır. Katılımcıların %67'lik kısmı bu konu hakkında daha fazla yatırım yapılarak genomik verilerin mahremiyetinin arttırılabileceğini belirtmişlerdir. Katılımcılardan %59'u daha fazla test ve denemelerin yapılması gerektiğini belirtmişlerdir. Katılımcılardan %69'u genomik verilerin incelenmesi ile sağlık alanında daha fazla fayda sağlanabileceğini belirtmişlerdir. Katılımcılardan %31'i ise mahremiyeti artırmak için hiçbir şey yapılmaması gerektiğini belirtmişlerdir [7].

III. Hassas Genomik Verilerde Mahremiyet

Genom verilerinde mahremiyet diğer verilere göre daha önemli ve korunması zaruri olan bir durumdur. Çünkü genom verisi üzerinden oluşturulacak bir ifşa durumu sosyal medya ya da kişisel sağlık kayıtları üzerinden ortaya çıkabilecek ifşa durumundan ayrıldığı en önemli nokta genom verilerinin bireyin sadece kendisi değil aynı zamanda ataları ile ilgili de zengin bilgiler içermesidir. Bir insan genom verisini kaybettiği zaman anonimliğini tamamen ve sadece kendi anonimliğini değil aynı zamanda akrabaları, kendinden önceki ve sonraki nesillerin anonimliğini de tamamen kaybetmektedir. Bu

sebeple çalışmada, Tam Genom Dizilemesi (Whole Genome Sequencing-WGS), mahremiyet ifşası ve mahremiyet ile ilgili verilerden bahsedilmektedir.

A. Tüm Genom Dizilemesi

Ayday ve arkadaşları uygun fiyatlı hazır WGS'nin biyoinformatik alanına teşviki artıracağına ancak genomik veriler üzerindeki mahremiyet endişesinin artacağını belirtmişlerdir. Yazarlar bu durumda bilgisayar bilimcilerin, hukukçuların, sağlık hizmeti sağlayıcıların vb. farklı dallarda çalışan bilim insanlarının iş birliğinin gerektiğini vurgulamış ve mahremiyet sorunlarını ele almak için bir iş birliği çağrısı olduğundan ve WGS'nin tıbbi olarak en üst seviyede kullanılmasının bireyler ve toplum için faydalı olacağından bahsetmişlerdir. WGS'in yakın zamanda bütün insanlar için daha uygun hale geleceğini fakat mahremiyet ve etik konularının bu popülerleşmeyi kısıtlayıp potansiyel tıbbi ilerlemeyi yavaşlatabileceğini öngörmüşlerdir. Her genomik sekansın hastalığın anlaşılmasında ve tedavi edilmesinde önemli ilerlemeler sağlayan çok miktarda bilgi içerdiğinden bahsetmişlerdir. Teknoloji ilerledikçe insanlar için genom verisi analizi yani WGS işlemleri uygun maliyetli hale geldiğinde kişilerin genom verisine erişimi için hukuki güvencelere ihtiyaç duyacaklardır [2]. Burada en önemli mahremiyet sorunu genom verilenini kimsenin erişimi olmadan ya da izinsiz şekilde sızıntı yapılamayacak şekilde nasıl korunacağı ve kontrol edileceği olmuştur.

B. Mahremiyetten Doğan Sorunlar

Ayday ve arkadaşlarına göre başlangıçta sosyal medya ve kişisel sağlık kayıtlarının ortaya çıkardığı veri maruziyeti meselesi DNA dizilimlerinin oluşturulması ile büyük ölçüde artmaktadır. Genomik verilerin kişinin çevresi ve yaşam tarzı hakkındaki veriler ile birleştirerek saldırgan tarafça kişinin hastalıklara yatkınlığı da dâhil olmak üzere kişinin tüm fenotipini ortaya çıkartabilir. DNA sekansını iptal etmek veya değiştirmek imkânsızdır. Yazarların düşüncesi ise araştırmacıların bu problemleri ciddi anlamda ele alıp çözüm sunana dek kişiselleştirilmiş ilaçların oluşturulmasının ertelenmesi kişiler için daha doğru ve mahremiyetleri için daha güvenli olacağıdır [2]. Kişinin saç teli, deri veya tükürük gibi biyolojik malzemelerin kişinin genomik mahremiyetini günler sonra toplansa bile ortaya çıkarmakta kullanılabilecek kanıtlar olduğunu göstermişlerdir. Bu tür mahremiyeti ihlal edecek saldırılar sadece topluluklar, veritabanları ve çok sayıda sayısallaştırılmış genom için değil kişiye veya küçük bir gruba yönelik de olabilir. Yazarlar, adlar ve sosyal güvenlik numaralarının silinmesi ve maskelenmesinin kişiyi anonimize etmesi konusunda yetersiz kaldığını çünkü genomun nihai yani ebedi tanımlayıcı olduğunu düşünmüş ve bunu belirtmişlerdir. Bir kişi kendi genomik verisinin sızdırılması ile sadece kendisi değil, yakın akrabalı hakkında da genomik bilgileri sızdırmış olmaktadır. Yazarlar, bu konu ile ilgili olarak bu verilerin verilmesinde gönüllü tesadüfi veya kötü niyetli olup olmadığına bakılmaksızın bunun toplumsal olarak büyük bir sorun olduğunu çalışmalarında belirtmişlerdir. Başkalarının kimliğinin açığa verilmesi genomik veri sızıntısını benzersiz ve üzerinde uzun mesailer harcanması gereken bir sorun haline getirmektedir çünkü bu durum kişinin akraba sayısına bağlı olarak büyük bir grubu da etkileyebilmektedir [6]. Gymnek'in açıkladığı 131 kişinin verisinin sızdığı saldırıda, 131 kişi ve onların 2100'e yakın akrabalarının etkilendiği belirtilmiştir [3].

Proteinler, birçok insan hastalığında çok önemli bir role sahip olan biyolojik bir sistemin fonksiyonel unsurlarıdır. Spesifik gen mutasyonları kişilerde belirli hastalıklara yatkınlığına dair bir anlatımdır. Yazarlar genlerin maskelenmesinin bireylerin genomik verilerinde mahremiyetini korumak için uygulanabilir bir yaklaşım olduğundan bahsetmişlerdir. Veri maskelemesi ile oluşturulan yöntemin daha önce de bahsedilen Gymnek tarafından açıklanan saldırılarda olduğu gibi kişilerin hassas verileri ile ilgili bilgi edinmeyi amaçlayan saldırıları engellemek için yetkisiz açıklamalarla sızan DNA dizileri olduğu zaman bireyin mahremiyetini koruyacağı V.Cogo ve arkadaşları tarafından düşünülmüştür [1]. Bu saldırılar kişiler üzerinde bir tanı oluşturmaz fakat yanlış ellerde kişiler için dezavantajlı şekilde kullanılabilir.

Bilgisayar bilimi açısından duruma yaklaşıldığı zaman genomik verilerin işlenmesinde her adımda oluşacak sorunların güvenlik ve mahremiyet kapsamında değerlendirilmesi ve önlem alınmasına ihtiyaç vardır. Diğer bir yandan tıbbi bilimler açısından bakıldığı zaman güvenlik ve mahremiyet kavramlarının kapladığı alan genomik veriler üzerinde geliştirilebilecek teknolojilerin hızına ket vurup durumu ve teknolojinin keşfini yavaşlattığı düşünülmektedir. Bu iki disiplin arasında ortak bir paydada buluşmak zaruridir.

C. Mahremiyet ile İlgili Çıkarımlar

Naveed ve arkadaşları hem mahremiyet eksikliğinin hem de mahremiyet fazlalığının genomik verilerin incelenmesi açısından beklenen faydaları azaltma potansiyelinin ortaya çıkabileceği hakkında bilgiler vermektedirler. Bununla beraber yazarlar Hipokrat Yemini'nde hastanın mahremiyeti hakkında bulunan ilgili beyandan ve hassas genomik verilerin sızıntısında oluşabilecek endişe ve problemlerden bahsetmişlerdir. Aynı zamanda mahremiyetle ilgili problemlerin fazlalığı yapılacak olan genomik araştırma ve incelemeleri engelleyebileceği görüşündedirler. Buna çözüm olarak mahremiyetin korunması için eğitimler verilmesi gerektiği ve biyoinformatik derslerinin müfredatında verilecek bilgilerin en üst düzeyde olması gerektiği yazarlar tarafından vurgulanmıştır. Yazarlar, genomik verilerin toplanması, depolanması ve işlenmesi için gerekli olan donanımların ucuz olması ve yüksek verimli sekanslama teknolojilerinin de birleşimi ile giderek daha kolay hale geldiğini düşünmektedirler. Fakat genetik verilerin mahrmeiyet ve güvenlik sorunlarının henüz yeterince ele alınmadığını, aynı zamanda bu genetik verilerin gelecekte nasıl kullanılacağı hakkında somut bir bilgi olmadığını belirtmişlerdir [7].

Zhen Lin, Art B. Owen ve Russ B. Altman, ilk olarak genetik varyasyonların kalıtımsal hastalıkları nasıl etkilediğini anlama ve tıbbi tedavilere yanıt verme konusundaki ilginin

yoğun olduğundan bahsetmektedirler. İnsan genomik verilerinin özel ve hassas olduğu vurgulanmaktadır. Yazarlar bu genomik verilerin kamuya açılmasının bilinçsiz mahremiyet sorunlarını ortaya çıkaracağını belirtmişlerdir. Yazarların ortaya çıkacağını düşündükleri mahremiyet sorunu; bireyin genetik verilerinden halka açık SNP verisi ile eşleşmeler olduğunda kişinin başarılı bir şekilde eşleştirilmesine yol açabileceğidir. Mahremiyeti koruma arzusu ile bilimsel verilere erişimi sağlama ihtiyacı arasındaki gerginliğin yeni teknolojilere yönelik bir arayışa yol açtığından bahsedilmektedir. Genomik verilerin mahremiyetinin ve güvenliğinin sağlanması hakkında yapılması gereken strateji ve yöntemlerden bahsetmişlerdir. Fakat yazarların yaptığı hesaplamalar sonucunda bu tür stratejilerin mahremiyeti korumadığını göstermektedir. Yazarlar genomik verilerin mahremiyeti ve güvenliği hakkında teknolojik yenilikler yeterli seviyeye gelene kadar yasalar ve düzenlemeler ile çözümler bulunması gerektiğini savunmaktadırlar. Bireysel genotip verileri ve ilişkili fenotip bilgileri içeren Farmakogenetik bilgi bankası inşa ettiklerinden bahsetmektedirler. Böylece endüstriyel veya hükümet araştırma birimi ile ilişkili olduğunu göstermedikçe hiçbir genetik veri kimseye sağlanmayacağından ve genomik veriyi kötüye kullanımı önlemese de kullanımı izlemenin bir yolunu sağladığından bahsedilmektedir [8].

Greenbaun ve arkadaşları çalışmalarında, açık kaynak ve verilerin geçmişte biyoinformatikte destekleyici nedenler olduğunu, bununla beraber mahremiyetle ilgili endişelerin de gelecekte kişisel genom verileri gibi veri setlerine erişimi sınırlayabileceğine değinmişlerdir. Yazarlar biyoinformatik alanında her zaman bir gerilim olduğunu ve bu gerilimin hasta mahremiyeti üzerine olduğunu belirtmişlerdir. Yazarlar kişiye ait genom diziliminin birey hakkında daha çok bilgi ortaya koyduğundan ve tıbbi kayıtlar ve insan genomu arasında ayrım yapmanın zor olacağından bahsetmektedirler [16].

IV. SONUÇLAR

Bu çalışmada büyük genomik verilerin üzerinde mahremiyet çalışmaları, büyük genomik veride hassas veri parçaları, korunması ve saklanması üzerine incelemeler ve araştırmalar yapılmıştır. Literatür taramaları sonucunda görüldüğü üzere genomik verilerin mahremiyeti hem bilgisayar bilimcileri için hem de tıbbi bilimlerde çalışanlar için büyük bir sorundur. Tıbbi keşifler mahremiyet sebebi ile yavaşlarken bilgisayar bilimcileri kişilerin mahremiyetini tıbbi keşiflerin önüne koymayı tercih etmektedir.

Büyük genomik verilerde hassas parçaların bulunması ise tamamen ayrı bir şekilde üzerinde durulması gereken bir konudur. Araştırmalar sonucunda elde edilen bilgilere göre büyük genomik verilerde hassas parçaların korunması, saklanması ve kullanılması insan genomunun kalan kısmının korunması ve saklanmasından çok daha önemlidir. STR(kısa tandem tekrarları), Y-STR(Y kromozom kısa tandem tekrarları), hastalık bağlantılı genler ve genomik varyasyonlar insanlar için genomik veri mahremiyetinde korunması gereken noktalardır.

Genomik verilerin hassas parçalarının sızıntısı sadece kişiye değil kişinin yakın akrabalarına da sorun oluşturmaktadır. Hassas genomik verinin istek dâhilinde verilmesi bile bir kişiye bırakılamayacak kadar önemli ve ciddi bir konudur. Kişi kendi hassas verisini anonim olarak paylaşmak istese dahi paylaştığı verinin sadece kendisine değil aynı zamanda yakın akrabalarına da ait olduğunun bilincinde olmalı ve buna göre hareket etmelidir. Bu konu hakkında eğitimler verilmeli ve toplum hassas verisini koruma konusunda eğitilmelidir.

Genomik verilerin ciddi boyutlarda olduğunu ve işlenmesi için büyük veri analitiği yöntemlerinin kullanılması gerektiği ortadadır. Biyoinformatik mühendisleri ve büyük veri uzmanları bu konu üzerinde ortak ilerleyerek hem hassas gen parçalarının analizinde hem anonimleştirilmesinde ortak çalışmalıdır. Teknolojini gelişmesi ile birlikte WGS yani bütün genom sekanslaması işlemlerinin fiyatları düşeceğinden ötürü bu işlemin herkes tarafından yapılacağı bir dönem oluşacaktır. Bu dönem gelmeden önce genomik verilerin anonimliğinin ve korunur hale getirilmesinin hem hukuki hem de teknolojik olarak sağlanması gerekir. Genomik verinin asla tam anlamıyla anonim hale getirilemediğini çünkü genom insan için ana tanımlayıcı olduğunu görmekteyiz. Bu sebeple bütün veririnin anonimleştirilmesindense hassas verinin korunması daha önemli bir noktaya gelmiştir. Aslına bakılırsa bu sorun tıbbi alanlarda sadece günümüzde değil tıbbın babası olan Hipokrat zamanında bile var olan bir sorundur. Hasta mahremiyeti günümüzde genom bazına kadar düşmüş olsa da temelde sorun her zaman aynı kalmıştır. Bizim yapmamız gereken ise zaman ve teknoloji ile birlikte evrilen sorunu anlamak ve yeni dönemin getirdiği şartlar ile birlikte ele almaktır.

KAYNAKÇA

[1] V. V. Cogo, A. Bessani, F. M. Couto, P. Verissimo, "A High-Throughput Method to Detect Privacy-Sensitive Human Genomic Data", *Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society - WPES '15,* 2015.

[2] E. Ayday, E. De Cristofaro, J.-P. Hubaux, and G. Tsudik. Whole genome sequencing: Revolutionary medicine or privacy nightmare? *Computer*, (2):58{66, 2015.

[3] M. Gymrek, A. McGuire, D. Golan, E. Halperin, Y. Erlich, "Identifying Personal Genomes by Surname Inference", *Science*, 2013

[4] (06.08.2018). *Nat'l Center for Biotechnology Information*, İnternet Syfası: www.ncbi.nlm.nih.gov/projects/SNP.

[5] (07.08.2018). G. Naik, *Gene Maps Are No Cure-All*, İnternet Sayfası: www.wsj.com/articles/SB10001424052702304023504577319604245325644.

[6] M. Humbert, E. Ayday, A. Telenti, "Addressing the Concerns of the Lacks Family: Quantification of Kin Genomic Privacy," *Proc. 20th ACM Conf. Computer and Communications Security* (CCS 13), 2013, pp. 1141–1152.

[7] M. Naveed et al., "Privacy in the genomic era", *ACM Comput. Surv*, 48(1), 2015.

[8] Z. Lin, A. B. Owen, R. B. Altman, "Genomic research and human subject privacy", *Science*, 305(5681):183, 2004.

[9] H. Fan, J.Y. Chu, "A brief review of short tandem repeat mutation", *Genomics, Proteomics & Bioinformatics*, 5(1):7{14, 2007.

[10] J. Koreth *et al*., "Microsatellites and PCR genomic analysis", *Journal of Pathology.* 178: 239-248.

[11] J.L Mountain, "SNPSTRs: empirically derived, rapidly typed, autosomal haplotypes for inference of population history and mutational processes", *Genome Res.* 12: 1766-1772.

[12] C. M. Ruitberg, D. J. Reeder, J. M. Butler, "STRBase: a short tandem repeat DNA database for the human identity testing community", *Nucleic Acids Res.*, 29(1):320-322, 2001.

[13] Y. Gelfand, A. Rodriguez, G. Benson, "TRDB: the tandem repeats database", *Nucleic Acids Res.*, 35, 2007.

[14] T. E. King, M. A. Jobling. "What's in a name? Y chromosomes, surnames and the genetic genealogy revolution", *Trends Genet*., 25(8):351-360, 2009.

[15] J. M. Butler, "Genetics and genomics of core short tandem repeat loci used in human identity testing", *J. Forensic Sci*., 51(2):253-265, 2006.

[16] D. Greenbaum, A. Sboner, X. J. Mu, M. Gerstein, "Genomics and privacy: Implications of the new reality of closed data for the_field", *PLoS Computational Biology*, 7(12), 2011.

[17] N. Homer et al. "Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays", *PloS Genet.*, 4(8), 2008.

[18] M.A. Jobling, P. Gill, "Encoded evidence: DNA in forensic analysis", *Nat Rev Genet*. 5: 739–751.

[19] D. Altshuler, M.J. Daly, E.S. Lander, "Genetic Mapping in Human Disease", *Science*, 2008

[20] (05.07.2018). A. D. Long, P. Beldade, P. Brakefield, A. Monteiro, *Why sequence a butterfly?*, İnternet Sayfası: https://jgi.doe.gov/why-sequence-a-butterfly/

[21] (16.05.2018). *Büyük Veri - Big Data nedir?,* İnternet Sayfası: https://www.bilginc.com/tr/egitim-haber/buyuk-veri-big-data-nedir

Büyük Genomik Verisinin Mahremiyetinin Korunarak İşlenmesi

Privacy Preserving Processing of Big Genomic Data

Arian Ajdari
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
arianajdari23@gmail.com

Zeynep Fenerci
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
zeynepfenerci9@gmail.com

Yılmaz Vural
Kişisel Verileri Koruma Kurumu
Ankara, Türkiye
yilmazvural@gmail.com

Yavuz Canbay
Bilgisayar Mühendisliği
Gazi Üniversitesi
Ankara, Türkiye
yavuzcanbay@gazi.edu.tr

***Özet* – DNA sekansları, teknolojinin gelişimiyle birlikte çeşitli hastalıklara ve soy ilişkilerine yatkınlığı ortaya koyan çok sayıda bilgi zenginliği sunmaktadır. DNA sekanslamanın getirdiği bu bilgi zenginliği, mahremiyet kavramının da bu alanda ele alınması gerekliliğini ortaya çıkarmıştır. Son zamanlarda, genomik veriler üzerindeki mahremiyet sorunları ve genomik verilere yetkisiz erişim ile ilgili olası çözümler tartışılmaktadır. Ancak güvenlik ve mahremiyet kavramları üzerine yapılan yanlış ve karmaşık tanımlamalar dolayısıyla, önerilen çözümlerin gerçek yaşam problemlerine uygulaması zor olmaktadır. Bu çalışmada, lgenom verisi mahremiyeti hakkında iteratürdeki çeşitli teknikler incelenip özetlenmiş, teknolojinin bu konudaki karmaşıklığı ve olgunlaşma süreci değerlendirilmiş ve son olarak da genetik mahremiyetin ihlal edilmesine sebep olan durumlar gözden geçirilmiştir.**

***Anahtar kelimeler*—mahremiyet, büyük genom verisi, DNA analizi**

Abstract – DNA sequencing coupled with technological advancements offers enrichment of information technology in accordance with different genetic diseases and ancestry. Enrichment of information technology has also pawed way for privacy concerns. Privacy concerns arising from genetic data, unauthorized access and their respective solutions have been debated lately. It is agreed that proposed solutions in the area of security and privacy, while some being wrong and complicated, are hard to be implemented properly in real-world scenarios. In this work, current privacy preserving techniques were researched, technological complications, maturity and real-world examples of genetic privacy breaches were examined.

***Keywords*—privacy, big genomic data, DNA analysis**

I. GİRİŞ

Biyoinformatik, dünya genelinde bilim için önemli bir araştırma konusu olan ve ülkemizde son yıllarda bilinirliği artan disiplinler arası bir alandır. Bu bilim dalının temelini moleküler biyoloji oluşturmaktadır. Fakat aynı zamanda bilgisayar ve istatistik alanlarından da yararlanarak olası problemlere çözüm yolu üretmeye çalışmaktadır. Birey varoluşu gereği, benliğinin bilincine sahip olarak kendi kararlarını verebilmekte ve mahremiyet konusunda hassasiyete sahip olmaktadır.

Tüm bu özellikler mahremiyet kavramının temellendirilmesini sağlamaktadır. Böylece mahremiyet kişinin özerk olması konusunda temel nitelikler arasında sayılmaktadır. Mahremiyet kavramı sadece özel alana işaret eden bir kavram değildir. Mahremiyet, hem özgürlük bilinci gereği kendi kendini belirleme, hem de kendisi dışındaki çevreyi kendi ilgisine göre kontrol etme anlamına gelmektedir [1]. Mahremiyet kavramı düşünce özgürlüğü, müdahaleden bağımsız olma, belirli konuları başkalarından gizleme, kişisel bilgileri kontrol etme, itibarını koruma gibi pek çok farklı alanı kapsayan bir kavramdır. Mahremiyet, bir insan hakkıdır. İnsanlar kendileri hakkındaki bilgilerin adil bir şekilde kullanılacağı konusunda emin olmalıdırlar. Kişisel veriler ya da farklı bir ifadeyle mahrem veriler, kamu yararına toplanır, saklanır ve işlenir. Elektronik veri yönetimi ile birlikte veri toplama, işleme ve saklama işleri daha da artmıştır [2].

Genetik mahremiyeti, DNA sekanslama esnasında elde edilen verilerden oluşmaktadır. DNA verileri nesilden nesile aktarıldığı ve büyük miktarda enformasyon içerdiği için verilerin mahremiyeti büyük önem taşımaktadır. Bu bağlamda genetik mahremiyette genetik materyalin mahremiyeti ve bilginin mahremiyeti konuları ele alınır.

Bu makalede, mahremiyet ihlali stratejileri ilk önce kategorize edilip, temel teknik kavramlarını açıklanıp, performanslarını ve sınırlamaları değerlendirilip, verilerin analizi sırasında kullanılan ve gizliliği koruyan teknolojilerden bahsedildi.

II. BÜYÜK GENOM VERİSİNDE MAHREMİYET SORUNLARI VE KORUMA YÖNTEMLERİ

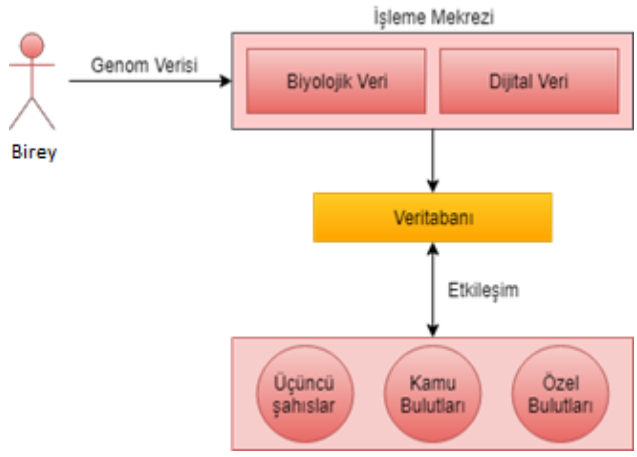
İnsan Genom Projesi biyoloji alanında, bugüne kadar yapılmış en karmaşık projelerden bir tanesidir. Bu projenin ana amacı, bir insanın genom verisinin tümünü anlaşılabilir şekilde kodlanmasıdır. İstenilen sonuca varabilmek için, bilim adamları 13 sene içerisinde yeni teknikler geliştirip ilk genom dizilimini elde etmiştir [3-5]. Sonuç olarak bilim adamlarının araştırmaları

sonucunda 3 milyar DNA çifti çıkmakla beraber bilim adamlarının bulanık bir döneme sürükleyişi başlamıştır. 3 milyar DNA çiftlerin arasında anlaşılmayan bir ilişkiyi kurabilmeleri için çeşitli istatistik yöntemlerine başvuruldu ve kısa zamanda bilgisayarların hesaplama gücünden faydalanabileceği öğrenilmiştir [6]. Son yüzyılda istatistik alanında yapılan geliştirmeler sayesinde büyük miktardaki verilerin analizi kolay hale getirilmiştir [7]. Bilgisayarlar hesaplama gücünü her geçen gün iki kat arttırdığına göre büyük verinin analizinin karmaşıklığı düzeltilmiştir [8]. Büyük bir veri havuzunun içerisinde çok kısa sürede ilişkilerin kurulması mahremiyet ihlaline neden olabilir [9].

Büyük genom verisi işlenirken, bireylerin mahrem bilgilerinin korunması için çeşitli teknikler kullanılmaktadır. Anonimleştirme ve kimliksizleştirme yöntemleri bu kapsamda önerilen tekniklerden bazılarıdır. Bahsedilen bu yöntemler genom verinin bütünlüğünü ve işlevselliğini korumaktadır. Bu sayede verinin mahremiyeti de korunmuş olur [9].

A. Büyük genom verisinde mahremiyet ihlalleri ve koruma yöntemleri

Genom verisinin 'büyük veri'nin alt ürünü olarak başarılı ve hızlı bir şekilde işlenebilmesi için güçlü bilgisayarlara ve dağıtık sisteme ihtiyaç duyulmaktadır. Her ne kadar insandan insana genom verisi sadece %0,5 farklı olsa da dağıtık sistemlerde genom verisi işleme esnasında büyük mahremiyet ihlalleri ortaya çıkabilmektedir [1]. Ayrıca üçüncü taraflar verileri işleme ve saklama esnasında kullanılan makine öğrenmesi algoritmaları ile kimlik ifşası gerçekleştirebilmektedir. Bu konuda çıkabilecek mahremiyet ihlaline örnek olarak, yapılan bir çalışmada genom verisinden göğüs kanseri tespit edilmesindeki kolaylık örnek olarak gösterilebilir. Sadece yukarıda bahsedilen %0,5'teki dizilimin SNP'lere (tek nükleotid polimorfizm) bakarak şahsın mahrem olarak nitelendirilen kanser bilgisi bilgisayarlar tarafından birkaç dakikada tespit edilebilir [4, 11, 12].



Şekil 1- Birey, işleme merkezi ve üçüncü taraflar arası etkileşim [9]

Şekil 1'de bireyin, işleme merkezi ve veritabanına erişebilecek olanların arasındaki etkileşimini göstermektedir. Birey, kendi genetik veresini işleme merkezine vermektedir. İşleme merkezi, biyolojik veriyi sayısallaştırıp dijital veriye dönüştürüp veritabanında kaydetmektedir. Veritabanı, üçüncü şahıslar, kamu bulutları ve özel bulutları arasında etkileşim gerçekleşme esnasında çeşitli mahremiyet sorunları ortaya çıkmaktadır.

- *Kamu ve özel bulutlarda mahremiyet ihlalleri ve koruma yöntemleri*

DNA'daki dizilim işlemi pahalı bir işlem olduğuna göre, dizilimin sıralanması üçüncü taraflarca yapılmaktadır. Burada iki tane mahremiyet sorunu söz konusudur. İlk sorun, veri kamu ağlar üzerinde iletildiğinde yetkisiz taraflar tarafından ele geçirilebilir ve kötü amaçlı kullanılabilir. İkinci sorun ise, veriler kamu bulutlarda ve işleme merkezlerinde bulunduğunda kamu bulutlarında bulunan taraflarca kötü amaçlı kullanılmasıdır. Bu sorunlardan ilkine çözüm olarak ise, çeşitli şifreleme teknikleri önerilmektedir. Verilerin şifrelemesi yerel sunucuda yapılarak iletim esnasında verilerin doğruluğu ve mahremiyeti korunabilmektedir [11]. Şifreleme yerel sunucular tarafından hızlı ve güvenli şekilde yapılmalı ve ayrıca bu veriler üzerinde kolay bir şekilde okuma, yazma ve silme işlemlerini desteklemelidir. Bununla ilgili Chen bir öneride bulunmuştur. Chen'e göre mahremiyeti korumak için kamu ve özel bulut yaklaşımı kullanılmalıdır. Özel bulutta ya da güvenirliği yüksek olan bulutta mahrem veri geçici olarak saklanıp işlenebilir. Ağır teknik işlemleri gerçekleştirilmesini kamu bulutlar kullanılabilir. Kamu bulutların özel bulutlarda bulunan mahrem verilere erişememesi için Chen tarafından hashleme teknikleri önerilmiştir. Hash anahtarı kamu bulutlarda hesapladığında, özel buluta geçen veri şifrelenmiş halde gitmektedir. Bu şekilde özel bulutlarda şifrelenmiş veri üzerinde işlem yapıldığında esas veri korunmaktadır. İşlem bitince şifrelenmiş veri özel bulutlarda öz hale getirebilmektedir. Bunun sonucunda çift bulut yaklaşımı ile mahremiyet kabul edebilir seviyede korunmasını sağlamaktadır [24].

- *Üçüncü şahısların etkileşimden kaynaklanan mahremiyet ihlalleri ve koruma yöntemleri*

Tüm tarafların mahremiyet konusundaki hassasiyetinden dolayı çeşitli koruma yöntemleri önerilmektedir. Örneğin; geliştirilen yeni saldırı türüne göre art arda yapılan bir sorgu, belli bir veri seti üzerinde, çıkan sonuca göre kimlik ifşasına neden olabilir. Bu probleme yönelik ilk çözüm Atallah tarafından ortaya atılmıştır. Bu çözüm ise, değiştirilmiş mesafe protokolünü içermektedir [25]. Atallah'ın çalışmalarından faydalanarak Troncoso-Pastoriza güvenli bir şekilde genom dizilimini doğrulayacak otomatı geliştirmiştir. Bu şekilde ortaya çıkan düzenli ifade otomat tarafından kullanılır ve işlenmiş verinin üzerinde işlem gerçekleşebilir. Otomat veriden habersiz olduğuna göre, tarafların hangi işlemde bulunduğu bilinmediği için, mahremiyet tamamen korunabilir [9]. Bu tekniklerin yanında genomik veri setine erişim sırasında mahremiyeti koruma yöntemleri de eklenebilir. En çok kullanılan yöntem anonimleştirmedir. Bu yöntem ile veritabanındaki her satır için tanımlanmayacak şekilde birer kayıt üretilmektedir.

Anonimleştirmenin prensiplerini kullanarak aynı değer ile ifade edilen hassas verinin mahremiyeti korunabilir [13].

- *Açık veritabanı etkileşiminden kaynaklanan mahremiyet ihlalleri ve koruma yöntemleri*

İşlenen veriler sayesinde insan genom analizi kolaylaşmaktadır. Veriler arasında analizlerin yapılabilmesi ve işlenmesi için verinin belirli bir yerde toplanması ve saklanması gerekmektedir. Saklama yeri ve şekline göre çeşitli mahremiyet sorunları oluşabilmektedir. Bir problem senaryosunda, bir hasta kendi genom verisine ulaşıp bu verileri üzerinde çeşitli testler yapma istediğinde bulunabilmektedir. Kendi genom verisinin nerede saklandığını bilmediği durumlarda, hasta servis sağlayıcılara güvenmemektedir. Bu sebepten, hasta kendisi hakkında oluşabilecek problemleri 3. taraflara paylaşmak istememektedir [10]. İşleme esnasında çeşitli şifreleme teknikleri kullanıldığı gibi saklama esnasında da veriler şifrelenmiş halde kaydedilmektedir. Homomorfik şifreleme teknikleri kullanıcıların mahremiyet seviyesini arttırmaktadır. Homomorfik tekniği ve esas verilere yalnız süzgeçten geçirilmiş şirketlerin tarafından ulaşım sağlanabildiğinden dolayı kullanıcıların mahremiyet koruma seviyesi kabul edilebilir hale getirilmektedir. Homomorfik şifreleme şifrelenmiş verilerin işlem yapmasını sağlamaktadır. Homomorfik işlemden sonra çıkan sonuç dizisi doğru sonucu tutmaktadır. Açık bir veritabanı kamu bulutlar, özel bulutlar ve 3. şahıslar tarafından erişildiği için asıl veriye ulaşması ciddi mahremiyet nedenine olabilir ancak homomorfik şifreleme sayesinde asıl veri yanlız veritabanında kaydedilmiş olmaktadır [9]. Genomik veriler kullanıcıların kişisel depolama cihazlarında depolanmayacak kadar hassastır. Bu sebeple kullanıcıların kendi genomik verilerini kendi ellerinde bırakmak bile yüksek oranda risk içerir. Tüm bunlar sonucunda genomik verileri depolamak, işlemek ve mahremiyet sorunundan korumak için Depolama ve İşleme Birimi (SPU) kullanılmıştır [10].

B. Büyük genom verisi işlemede ortaya çıkabileck mahremiyet sorunları

Son yıllarda, DNA işleme yöntemlerinde büyük adımlar atılmıştır. Bu kısmın amacı, genomik veri işleme esnasında kullanılan algoritmalardan kaynaklanan mahremiyet sorunları ele alınıp çözüm yöntemleri önerilecektir. Makine öğrenmesi, Yapa Zekâ ve Veri Madenciliği konuları üzerinde mahremiyetin nasıl korunabileceğine dair bilgi verilecektir. Genom verisi işlenme esnasında kullanılan algoritmalar üçüncü şahıs tarafların bilgi sızdırmalarına engel olmalı ve şahsıların mahremiyetini korumalıdır. Algoritmaların yanında mahremiyeti koruyacak şekilde çeşitli protokoller tasarlanmıştır. Protokollerin verimli şekilde tasarlanması için deterministik otomatlar kullanılmıştır. Bu şekilde protokolün verimli ve esnek şekilde çalışması garanti edilmiş olur [11, 14].

- *Makine Öğrenmesi algoritmalarından kaynaklanan sorunlar*

Makine öğrenmesi algoritmaları, zamanla ya da kazandırılmış 'tecrübe' ile daha iyi sonuç veren algoritma kümesine denilmektedir. Bilim adamları, makine öğrenmesi algoritmalarıyla büyük bir genom veri setinin içerisinde insan tarafından anlaşılmayacağı ya da zor bulunacağı desenler ve özelliklerini bulmayı amaçlamaktadır. Makine öğrenmesi algoritmaları genel olarak 3 aşamada geliştirilmektedir [10]. Birinci aşamada algoritma istenilen sonuca götürecek şekilde geliştirilmektedir. Bu aşamadan sonra algoritma bilgisayar tarafından anlaşılabilir olmuştur. İkinci aşamada bu algoritma üzerinde daha önce belirlenmiş veriler çalıştırılır ve algoritma bulduğu özelliklerle beraber 'öğrenir'. Öğrendiği noktaları etiketler üzerinden sınıflandırma yapar. Sınıflandırma yaptıktan sonra algoritmanın öğrenme kabiliyetini ölçmesi için algoritmanın daha önce görülmemiş test verileri üzerinde yoğunlaşması istenmektedir. Bu yaklaşımla makine öğrenmesi algoritmaları bir genom verisi üzerinde yüksek başarıyla genlerin başlangıcını, genlerin ifade edilişini, hastalıkları ve fonksiyonu tanımlayabilir [14, 16, 17].

Linear Regression sıklıkla kullanılan bir makine öğrenmesi algoritmasıdır. Bu algoritma sınıflandırma yapabilmesi için bir tane doğru çizer ve gelen veri seti üzerinde sınıf kararını yapar. Genom veri seti, daha önce açıklandığı gibi, büyük bir veri olduğu için Linear Regression tarafından üretildiği ağırlıklar mahremiyeti bozacak şekilde bilgi içirebilir. Linear Regression modelinde ağırlıklar modelin nasıl öğrendiğini ve ne şeklinde sınıflandırmayı yapma konusunda bilgi içerdiği için, ters mühendislikle genom verisi hakkında bilgi edinebilir. Bu problemi önlemek için ağırlıkların şifrelenmesi gerekmektedir. Ağırlıkları şifrelemekle ve şifreleme anahtarı yalnız bilgisayar tarafından bilindiğinde, insan faktörü ortadan kaldırılmış olmaktadır. Bu şekilde, makine öğrenmesi algoritmaları tarafından yapılan çıkarımlar yalnız bilgisayar tarafından bilinmektedir [16].

- *Yapay Zekâ algoritmalarından kaynaklanan sorunlar*

Yapay Zekâ algoritmalarından kaynaklanan sorunları ele alabilmek için ilk olarak Yapay Zekâ'yı toplu bir bilim dalı olarak açıklamak gerekmektedir. Yapay Zekâ'nın amacı, düşünebilen ve karar verebilen makineler veya sistemler tasarlamaktadır. Böyle bir tasarımı gerçekleştirilebilmek için, sistemin çok fazla miktarda veriye ihtiyacı vardır. Büyük veriden gereken çıkarımların yapılabilmesi için daha önce makine öğrenmesi algoritmalarıylandan eğitilmiş veri kümelerinden faydalanılması gerekmektedir. Yeni girdiler gelince yapay zekâya sahip olan sistem derin öğrenmeyi kullanarak yeni bilgiler edinebilir. Yapay zekâ birden fazla teknik bilimi bir araya getirerek düşünen ve karar verebilen sistemleri geliştirmeyi amaçlar [13].

Makine öğrenmesi uygulamalarında koruyucu yöntemler kullanılmalıdır. Bugünün teknolojisinde yapay zekâ, derin öğrenmeyi kullanarak oldukça geliştirilmiştir. Derin öğrenme tekniklerin başarılı oranı yüksek olduğu için, pek çok şirket tarafından sık sık kullanılmaktadır. Derin öğrenme teknikleri genom verisi üzerinde çıkarımlarla beraber mahremiyet sorununu da getirmektedir. Bir derin öğrenme algoritmanın başarabilmesi, orantılı olarak verinin miktarına bağlıdır. Derin öğrenme gerçekleşme esnasında mahremiyeti korunabilmesi için stokastik süreçlerin uygulanması önerilmektedir [14, 15].

Stokastik süreç, olasılıktan etkileyen ve ona göre sonuçları çıkartan süreci kapsamaktadır. Stokastik süreci verilen model ile ilgili tüm olasılıklarını tutmakla beraber, modellerini sadece öğrenilmiş katsayılarını ve verisetinin küçük bir parçasını paylaşarak mahremiyet kabul edilebilir seviyede korunabilir [15].

- *Deterministik otomatların doğruladığı protokollerin mahremiyeti*

Deterministik Otomatlar, durum değişikliğini takip ettiği için protokollerin tasarımını basitleştirip doğru bir şekilde işlenmesinden sorumludur. Genom verisinde deterministik otomatların diziliminin verilen özelliklerine sahip olup olmadığını tespit etmek için kullanılabilir [14].

Genom verisi farklı taraflarca işlendiğinde kullanılan tekniklerden dolayı verinin alıcısı veya göndericisi istenilmeyen bilgileri elde edebilir. Genom verisi üzerinde basit işlemler gerçekleştiğinde, verinin mahremiyetini korumak için çeşitli yöntemler mevcuttur. Bu yöntemler, bulunan makine üzerinde kullanılan algoritmalardan dolayı mahremiyetin bozulmaması için algoritmalar tek tek incelenip hakkındaki koruma yöntemleri açıklanır. Algoritmaların oluşturduğu kural kümesi deterministik otomatlardan doğrulanabildiğine göre işlem esnasında (özellikle karşılaştırma işlemlerinde), hem alıcıdan hem de göndericiden istenilmeyen bilgi sızabilir. Mahremiyeti koruması için, algoritmaların güvenli bir şekilde işlenmesinin yanında, bilgilerin sızıntılara da engel olması gerektiği bilinmelidir. Bu konuda ciddi çalışmalar yapılmaktadır. Bir çözüm olarak habersiz deterministik otomatlar önerilmektedir. Koruma yöntemleri üzerinde, çeşitli kamu ve özel bulutların haberleşme sırasında habersiz deterministik otomatlar asıl gitmesi gereken şifrelenmiş verinin dışında başka bir verinin bulunmamasını garanti eder. Böyle bir teknik 3 kriterle değerlendirilebilir: Doğrululuk, verimlilik ve güvenlik. Bu 3 kriter, habersiz deterministik otomatları tasarlarken kullanılmaktadır [18].

- *DNA eşleşme esnasında oluşan mahremiyet sorunları*

Teknolojinin ve algoritmaların gelişmesiyle beraber genom veritabanı oluşturma işlemi en kolay şekle indirgenmiştir. İlk olarak, çeşitli güvenlik teşkilatları kendi işlemleri için bulunduğu ülkelerin nüfusuna göre genom veritabanı oluşturmuştur. Amaç, suç işleme sırasında kimlik tespiti yapılmasıdır. Teknolojinin gelişmesiyle bu tür teknikler kamuya da sunulmuştur. Düşük bir ücretle, birey kendi genom verisi üzerinde soyağacını oluşturabilmektedir. Günümüzde bu tür testler ve uygulamalar internet üzerinde sunulmakta ve herkes tarafından kullanılabilmektedir. Soyağacı oluşturulduğu veya suçlu olanı tespit edilmekteyken kullanılan genom verisinin kısmına STR (Short Tandem Repeat) denilmektedir.

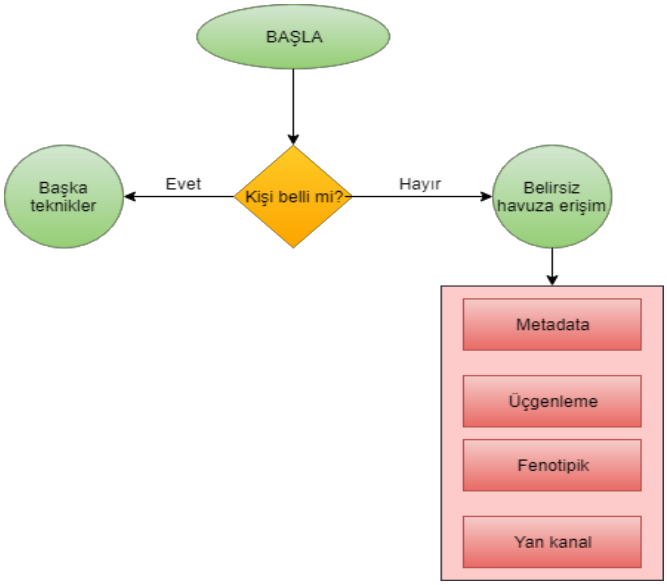
İlk olarak, STR genom verisinin işlevsiz kısmını ifade etmektedir. STR kısmı, genom verisinin bir kısmının kopyasını ya da belli bir kuralla göre tekrarlanmasını ifade ettiği için mahrem bilgi içermemektedir. Yalnız STR kısmın etrafında genetik hastalığa kaynak oluşturacak bilgi bulunduğu için, özel bir yaklaşımla şahsı hakkında mahrem bilgi elde edinebilir. Genom hakkında bilinen bilgi ile STR eşleşme esnasında kolay bir şekilde bireyin hakkında mahrem bilgi sızabilir ve bu bilgiler kötü amaçlı kişiler tarafından kullanılabilir. Sorumsuz şekilde kullanım bireyi tam anlamıyla kimlik ifşasına kadar götürebilir.

Bulanık dizi arama yöntemin kullanım kapsamı homomorfik şifrelemeden daha kısıtlıdır. Eğer eşleşmeyi yapan tarafı yalnız STR kısmından eşleşmeyi yapacaksa bulanık dizi arama yöntemi uygundur. Veritabanında her genom verisi belli bir anahtarla şifrelenir. Eşleşme yapıldığında kamu anahtar hesaplanır ve tüm veritabanı üzerinde çalıştırılır. Anahtar, veritabanındaki herhangi bir genom verisini açtığı takdirde eşleşmeyi bitermiş olur. Bu şekilde verilen genom verisini yalnız kim ait olduğu açık kalır ve hakkındaki bilgiler korunmuş olmaktadır [11].

C. Büyük genom verisi üzerinde kimlik ifşası

Son zamanların en çok konuşulan konuları arasında mahremiyet tehditleri ve genomik verilere erişimle ilgili çözümler yer almaktadır. Ancak güvenlik tanımlarının karmaşık doğası nedeniyle önerilen çözümler gerçek yaşam problemlerine uygulamak zor olmaktadır. Dijital genomik veriler, kimlik ifşası riskine neden çok sayıda biyoinformatik sürecin oluşmasına neden olur. Özel genomik verilerin analizi, sorgulanması, genomik veritabanının araştırılması ve yeterli önlemlerin alınmaması sonucu kişisel mahremiyet ihlaline neden olan biyoinformatik süreçlerin oluşumu başlar [18].

Kimlik ifşası ve mahremiyetin ihlaline neden olan teknikler genom verisi üzerinde yapılan soyağacı oluşturma tekniklerinden kaynaklanmaktadır. Milyonlarca kişi kendi soyağacını oluşturmasıyla beraber kimlik ifşası tekniklerinin artmasını tetiklemiştir. Bahsedilecek teknikler 4 tane madde altında toplanmıştır ve arasındaki ilişki aşağıdaki şekilde gösterilmiştir [2].



Şekil 2- Kimlik ifşasına yönelik kullanılabilecek yöntemler [19]

Şekil 2'de, verilen bir genom seti üzerindeki kimlik ifşasına yönelik kullanılabilir yöntemleri göstermektedir. Enformasyon teoreminden faydalanıp, bir genom verisi üzerinde kimlik tespiti için gereken enformasyon miktarı hesaplanabilmektedir. Bu

kısımda, metadata, üçgenleme, fenotipik ve yan kanal enformasyon miktarı üzerinde odaklanıp bilimsel bir şekilde mahremiyet ihlali açıklanabilir ve koruma yöntemleri önerilebilir.

- *Üçgenleme tekniği ile kimlik ifşası*

Bu teknik ile üçgenleme yaklaşımını kullanılarak bir genom verisinin kime ait olduğunu bulmaya çalışmaktadır. Üçgenleme tekniği, belirsiz bir havuzdan başlayıp, sıkıştırma yöntemiyle kimlik tespiti yapmaya çalışmaktadır. Örneğin, Amerika'da soyadı ve metadata ile kimlik ilişkisi kullanılacaktır. Amerika'nın nüfusu 300 million civarında olduğuna göre, verilen genom verisinin herhangi bir kişinin olma olasılığı miktar olarak aşağıdaki gibi ifade edilebilir [20]:

$$Entropy = \log_2 \frac{1}{\frac{1}{300.000.000}} = \log_2 300.000.000 \approx 28bit$$

Enformasyon miktarı 28 bit olduğuna göre, belirsizlik oldukça yüksektir. Üçgenleme tekniğinde ilk olarak genom verisi üzerinde cinsiyet tespiti yapılmaktadır. Bununla birlikte yarısı elenir ve enformasyon sayısı 27 bit'e düşürülmüş olur. Veritabanında saklanmış metadata üzerinde il ve yaş tespiti yapılıp enformasyon miktarı 16 bite düşer. Soyadı bilgisi ile enformasyon miktarı 3 bite düşer. Son 3 bite, kamu arama motorları yardımıyla çözülüp kimlik ifşası gerçekleştirilebilir. Bunu önlemek için, herhangi bir genom verisi üzerinde metadata saklanmamalı ve yapılan işlemlerin bilgi sızıntısına izin verilmemelidir. Önerilen yaklaşıma göre, kişiler kendi genom verisini kamu kuruşlarınca paylaştığında soyadını hiçbir şekilde açık tutmamalıdırlar [19].

- *Fenotik tahmin ile kimlik ifşası*

İkinci yaklaşım ise, genom verisi üzerinde fenotopik özelliklerden kimlik tespiti yapılmasıdır. Fenotopik bilgiler eğer doğru bir şekilde belirtilirse 6 bite kadar enformasyon içerebilir. Fenotipik bilgisinde, boy, BMI (Vücut Kütle Endeksi), yüz şekli, göz rengi vb. yer almaktadır. Yapılan araştırmalara göre, boy bilgisi santimetre boyutunda 5 bite kadar bilgi içerebilir ve üçgenleme tekniğine büyük bir katkı sağlayabilmektedir. Ancak fenotipik bilgileri günümüzde olsa bile yeterli bir doğruluğa sahip olmadığı için çıkacak enformasyon miktarı düşük olup kullanım alanı kısıtlıdır [20]. Daha önce açıkladığımız gibi, enformasyon teoremi verilen genom verisinin üzerinde kimlik tespiti için başarılı bir şekilde kullanılabilir. Fenotipik özellikleri bol miktarda enformasyon içerdiği için kimlik tespitini kolaylaştırmaktadır. Fakat boy bilgisi, yüz ifade bilgisi yeterince bir doğruluğa sahip olmadığı cihazlardan ölçtüğü için enformasyon miktarında düşüş görülmektedir. Günümüzde, iris izi gelişmiş cihazlar tarafından ölçebildiği için, genom verisinden edinen enformasyon miktarı çoğu zaman yüksekte tutulmaktadır. Demografik özelliklerini, genom verileri ve cihaz tarafından yansıtılmış iris izi, güçlü bir enformasyon miktarına sahip ve kimlik tespiti için kullanılabilen bir araçtır. Bu şekilde mahremiyet ihlali gerçekleştirilebilir. Koruma yöntemi olarak yukarıdaki detaylı açıklanan yöntemler uygulanabilir. Daha fazlası, hukuki çerçevesinde mahremiyeti koruyacak kanunlar kamunun lehine oluşturulabilir [3].

- *Yan kanal sızıntılar ile kimlik izleme*

Kimlik ifşası yan kanallar üzerinden de gerçekleştirilebilir. Genom verileri veritabanında kaydedilğinde şahsi bilgiler de kayıt altına alınabilir. Bu şekilde bir genom verisi eşleşme sırasında ve kayıt bulunduğu durumlarda mahrem bilgiler de bulunabilir. Ters mühendislik ile kamu arama motordan veri indirildiğinde dosya üzerinden isimden kullanıcı bilgileri elde edilebilir. Bu konuda koruma yöntemi olarak, şifreleme önerilmektedir [9].

III. MAHRMİYET SORUNLARI İÇİN GENEL ÇÖZÜM ÖNERİLERİ

Makalede açıklanan problem doğasından dolayı mahremiyeti koruyacak teknikleri açıklamadan önce büyük genom verisi işleme esnasında kaynaklanan mahremiyet sorunları yolunda ilerlenmiştir. Bu şekilde büyük genom verisi ile mahremiyet tanımı ve arasındaki ilişki kurulmuştur. İşleme esnasında mahremiyet ihlaline neden olabilecek sorunlar ispat edilmiştir ve korumaya dair bilgi verilmiştir. Mahremiyeti genel bir şekilde koruması için, 4 halka testi yapılmalıdır.

Birinci halka kanuni çerçevesidir. Mahremiyet kanuni bir şekilde tanımlandığı takdirde mahremiyeti koruması için gereken ilk adımlar alınmış olmaktadır. Bu yaklaşım doğru bir şekilde yapıldığında, mahremiyetin herhangi bir boyutu kabul edilebilir şekilde korunabilir [21].

İkinci halka, mahremiyetin nerede korunacağına dair bilgiler içermektedir. Genom verisi bir dizilim oluşumu olarak açıklandığına göre dijital dünyada kullanılan yöntemler açıklanmıştır. Genom verisi oluşturma esnasında (işleme merkezinde) mahremiyeti koruması için hashleme tekniği önerilmiştir. Daha sonra veritabanında yerleştiğinde, asıl verinin yanında gürültü olarak ekleme veya çarpma yoluyla veri yerleştirilmiştir. Son olarak açık bir veritabanın üzerinde yapılan işlemlerden dolayı, asıl veriye ulaşmadan işlem yapılabildiği doğrulandı. Koruma yöntemi olarak homomorfik şifreleme önerilmiştir. Devasa bir güce sahip olan bilgisayarlar genom verisi hakkında bir sürü çıkarım yapabilmek durumda olduğuna göre, genom verinin üzerinde çalışan algoritmalardan dolayı çeşitli sızıntılara neden olabilmektedir. Bu şekilde istenmeyen bilgi sızıntıdan dolayı mahremiyet ihlaline neden olabilmektedir. Kullanılan makine öğrenmesi algoritmaların parametreleri ters mühendislikle bilgi sızdırabilir. Koruma yöntemi olarak parametrelerin şifrelenmesi önerilmektedir [24]. Derin Öğrenme ve Yapay Zekâ'daki derin öğrenmeden kaynaklanan mahremiyet sorunları habersiz otomatlardan çözülebilir. Mahremiyete bilimsel boyutu kazandırmak için ve aynı anda mahremiyet ihlalinin miktarı ölçülmesi için Shannon'un enformasyon teoremi kullanılabilir ve kimlik ifşasında uygulanabilir [22].

Üçüncü halka gelecekte kullanılabilecek şifrelenmiş donanımı içermektedir. Özel tasarlanmış donanımlar şifreleme işlemini daha basit bir hale getirip, mahremiyetin korunmasından kaynaklanan maaliyet minimuma indirgenmiş

olmaktadır. Böyle bir systemin özellikleri çeşitli standardlara göre tasarlanmış. P1357 spesifikasyonu kapsamlı ve kullanışlı şifrelenmiş donanımını dijital biçimde anlatmaktadır. P1357 amacı her hangi bir sistemde, hem donanım hem de yazılımsal açıdan şifrelemeyi gerçekleşmektedir. Bu spesifikasyon yardımıyla dağıtık mimaride donanım açısından gereken önlemler alınabilir [19].

Son halka, insan faktörünü minimuma indirgenmesini amaçlamaktadır. Bu yalnız ilk üç halkanın başarılı bir şekilde gerçeklendiğinde yapılabilir. Herhangi bir sistemde insan faktörü ihlal edilemez, yalnız minimuma indirgenir [23].

IV. SONUÇ

Sonuç olarak, büyük genom verisi ve bilgisayarın hesaplama kapasitesi birbirine bağlıdır. Birinin gelişmesi diğerini tetiklediğine göre mahremiyet de bu faktörlerden etkilenmektedir. Yukarıdaki çalışmalar ve çözümler genom verisini ve bilgisayarları tek bir varlık olarak görmekle mahremiyeti korumaya dair bilgi verilmiştir. Makalede yer alan mahremiyet ihlali ile ilgili örnekler ve önerilen çözümlerin sonucunda mahremiyetin tam korunamaması açıklanmıştır. Fakat önerilen çözümler ile teknikleri tıbba ve sağlık profesyonellerinin yansıtması için bilgilerin korunmasına yönelik direktifler oluşturmuştur.

Bu konuda, sağlık hizmetlerinden faydalanması devlet tarafından sağlandığı için şahısların mahremiyetini korumak için gereken tıbbi ve etik kurallar çerçevesinde önlemler alınmalıdır. Günümüz teknolojinin sağladığı veri akışındaki özgürlük ile bilgi kaynağını sağladığı yarar mahremiyeti koruyacak şeklinde örtüşebilmelidir.

Genetik veri sayısallaştırdığında ve matematiksel bir matrise yerleştirildiğinde, genetik veri korumasız bir hale gelmektedir. Mahremiyeti kabul edilebilir bir şekilde korunması makalede geçen teknikler, zamana ve ihtiyaca göre uygulanmalıdır.

KAYNAKÇA

[1] (10.07.2018). National Human Genome Research Institute, *Human Genome Project produces many benefits,* İnternet Sayfası: www.genome.gov/27549135

[2] D. D. Tataroğlu, "Mahremiyet Sorunlarının Önlenmesinde Mahremiyet Etki Değerlendirmesi", *Yönetim ve Ekonomi,* 20(1).

[3] M. Sariyar, S. Suhr, I. Schlünder, "How Sensitive Is Genetic Data?", *Bioprservation and Biobanking,* 15(6), 494-501, 2017.

[4] Y. Erlich, A. Narayann, "Routes for breaching and protecting genetic privacy", *Nature Reviews Genetics*, 409–421.

[5] (10.07.2018). National Human Genome Research Institute, *What is the Human Genome Project?*, İnternet Sayfası: www.genome.gov/11511417

[6] R. E. Stevenson, "Mapping and Sequencing the Human Genome", 6th chapter on the Collection, Analysis, and Distribution of Information and Materials.

[7] M. S. Kaiser, "Advanced Statistical Methods", *Statistics 601*.

[8] T. N. Theis, H.S. P. Wong, "The End of Moore's Law: A New Beginning for Information Technology", *Computing in Science & Enginnering,* 19(2), 41-50, 2017.

[9] M. Akgün, A.O. Bayrak, B. Özer, M.Ş. Sağıroğlu, "Privacy preserving processing of genomic data: A survey", *Journal of Biomedical Informatics*, 103-111.

[10] F. Bruekers, S. Katzenbeisser, K. Kursawe, P. Tuyls, "Privacy-Preserving Matching of DNA Profiles", *IACR Cryptology ePrint Archive,* 16–17, 2008.

[11] M. Blanton, M. Aliasgari, "Secure Outsourcing of DNA Searching via Finite Automata", *IFIP Annual Conference on Data and Applications Security and Privacy*, 49-64, 2010.

[12] D.M. Roden et al., "Development of a large-scale de-identifie DNA biobank to enable personalized medicine", Clinical Pharmacology and Therapeutics, 84 (3), 362–369, 2008.

[13] M. Langheinrich, "Privacy in ubiquitous computing", Ubiquitous Computing Fundamentals, Boca Raton, FL, USA, CRC Press, 95-159, 2009.

[14] D. Szajda et al., "Toward A Practical Data Privacy Scheme for A Distributed Implementation of the Smith-Waterman Genome Sequence Comparison Algorithm", *Network and Distributed System Security Symposium,* 2006.

[15] P. Mohassel, Y. Zhang., "SecureML: A System for Scalable Privacy-Preserving Machine Learning". *IEEE Symposium on Security and Privacy,* 19-38, 2017.

[16] M. W. Libbrecht, W. Stafford, "Machine learning in genetics and genomics", *Nature Rviews Genetics.*

[17] J. Ramón, P. Troncoso, K. Stefan, M.S. Celik, "Privacy Preserving Error Resilient DNA Searching through Oblivious Automata", ACM Conference on Computer and Communications Security, 519-528, 2017.

[18] M. C. İzgi, "Mahremiyet kavramı bağlamında kişisel sağlık verileri", *Türkiye Biyoetik Dergisi,* 1, 2014.

[19] M. Kantarcıoğlu, W. Jiang, Y. Liu, B. Malin, "A cryptographic approach to securely share and query genomic sequences", Transactions on Information Technology in Biomedicine, 12 (5), 606-617, 2008.

[20] D. C. Barnlund, "A Transactional Model of Communication", *Communication Theory,* 47-57.

[21] P. J. Palsbøll, M. Allendorf, W. Fred, "Identification of management units using population genetic data", *Trends in Ecology & Evolution*, 22(1), 11-16, 2007.

[22] T. Gürsel, "İnsanlar Üzerinde Yapılan Biyomedikal Araştırmalarda Etik Değerlendirme", *Gazi Medical Journal*, 19(3), 2008.

[23] K.S. Leung et al., "Data mining on DNA sequences of hepatitis B virüs", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8(2), 428-440, 2012.

[24] Y. Chen, B. Peng, X. Wang, H. Tang, "Large-scale privacy-preserving mapping of human genomic sequences on hybrid clouds", *Network and Distributed System Security Symposium,* 2012.

[25] M.J. Atallah, F. Kerschbaum, W. Du, "Secure and private sequence comparisons", *Proceedings of the 2003 ACM Workshop on Privacy in the Electronic Society*, ACM, New York, NY, USA (2003), pp. 39-44,

Bilgi Güvenliği Bağlamında Yeni Teknolojik Devrim: Kuantum Teknolojiler

New Technological Revolution in the context of Information Security: Quantum Technologies

İhsan YILMAZ
Bilgisayar Müh.Böl., Müh.Fak., Çanakkale Onsekiz Mart Üniversitesi
Çanakkale, TÜRKİYE
iyilmaz@comu.edu.tr

Özet

Bu çalışmada, bilgi güvenliği bağlamında kuantum teknolojilerin dünyadaki gelişmeleri kısaca özetlenmiştir. Ayrıca bu yeni teknolojik devrimin kaçırılmaması bağlamında öneriler sunulmuştur.

Anahtar Kelimer: Kuantum teknolojiler, Bilgi güvenliği, süperpozisyon, Dolanıklık

Abstract

In this study, the developments in the world on quantum technologies in the context of information security are summarized briefly. In addition, recommendations were made in the context of not missing the new emerging technological revolution.

key words: **Quantum Technologies, Information Security, Super Position, Entanglement.**

I. KUANTUM TEKNOLOJİLER

Bilindiği gibi sanayi devrimi, bilişim devrimi, nükleer devrimi ülke olarak çok geç yakaladık. Bu zamana kadar insanoğlu işlerini elektriği kullanarak yaptırmaktaydı. Yaklaşık 30 yıldır insanoğlu yeni bir teknoloji devrimi üzerine çalışmaktadır. Bu devrimde kuantum teknoloji devrimidir. Bu devrimi kaçırmamak, ülkemizin geleceği bağlamında çok önemlidir. Bu yarışta geri kalmamak için bizimde kuantum teknolojilerine yönelip diğer ülkeleri yakalamamız gerekmektedir. Yoksa üreten değil tüketen bir toplum olmaya mahkûm oluruz.

Kuantum mekaniği; madde ve ışığın, atom ve atom altı seviyelerdeki davranışlarını inceleyen bir bilim dalıdır. Kuantum mekaniği; moleküllerin, atomların ve bunları meydana getiren elektron, proton, nötron, kuark, gluon gibi parçacıkların özelliklerini açıklamaya çalışır.

Günümüzde üretilen birçok cihaz elektronik devreleri kullanmaktadır. Kullanılan bu mikroçipler küçültülebileceği en son noktaya gelmiştir. Teknolojiyi geliştirenler yeni arayışlar içine girmiş ve kuantum teknolojisi üzerine yoğunlaşmıştır. Yaklaşık olarak son 30 yıldır bu teknoloji üzerine çalışılmakta ve birçok alanda kullanılmaktadır. İnsanlık var olduğundan beri diğerlerinden üstün olma çabası içerisindedir. Bunu da ancak diğerlerinin sahip olmadığı bilgiye sahip olarak sağlayabilir.

Temeli çok daha önceleri atılan fakat 1985'te Oxford Üniversitesinden David Deutsch tarafından "kuantum mantık kapıları" fikri ileri sürülerek, kuantum fiziği alanı bilgisayarlar içine dahil edildi. Deutsch'un bilimsel makaleleri herhangi bir fiziksel hesaplamanın kuantum bilgisayarlarına uygulanabileceğini gösterdi.

Kuantum teknoloji devriminin dayandığı temelleri aşağıdaki şekilde maddeler halinde özetleyebiliriz;

- Süperpozisyon (SuperPosition) durumu
- Dolanıklık (Entanglement)

A. Süperpozisyon(SuperPosition)

Kuantum veri, klasik bitlerin {0;1} kuantum mekaniği özelliklerini sağlayan sistemlerce ifade edilmesidir. Aslında yapılan 1930 lu yıllardan beri bilim insanlar tarafından bilinen ve deneysel olarak da ispatlanan bilgilerin teknolojik olarak insanlık yararına uygulanmasıdır.

Klasik bilgisayarlar sadece “0” ya da “1” değerlerini alabilen “bit” denilen bilgi parçacıkları üzerinde hesaplamaları gerçekleştirirler. Kuantum bilgisayarlar hesaplamalar için; “0” ve “1” değerlerinin her ikisini ve tüm olası durumlarını aynı anda barındıran “kuantum bit” ya da “q-bit” denilen bilgi birimini kullanmaktadır. 0 ya da 1 olduğu belirsizdir. Her iki

durumu aynı anda barındırır. Bu durum "Kuantum Süper pozisyon" olarak adlandırılmaktadır [1].

Kuantum durumları, elektrodinamik durumları gibi durumlardan yararlanılarak oluşturulabilir.

Örneğin;

Lazer Işığındaki Foton Polarizasyonu

0: $|H\rangle$ yatay polarizasyon, 1: $|V\rangle$ dikey polarizasyon

aynı klasik veriler farklı ve daha güvenli şekilde süper pozisyon olarak aşağıdaki gibi de ifade edilebilir [1].

$$0: \frac{|H\rangle + |V\rangle}{\sqrt{2}} \qquad 1: \frac{|H\rangle - |V\rangle}{\sqrt{2}}$$

Bu bağlamda, kuantum bilgisayarlardaki temel bazlar 0 ve 1'lerin her birinin durum (state) olarak ele alındığı durumlar olacaktır.

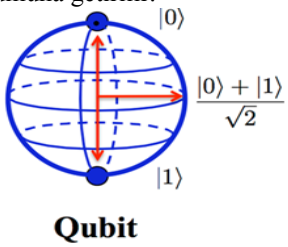
$$|0\rangle = \begin{vmatrix}1\\0\end{vmatrix} \quad |1\rangle = \begin{vmatrix}0\\1\end{vmatrix}$$

Bu durumda bir q-bitlik bilgi aşağıdaki şekilde ifade edilecektir [1].

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

$|\Psi\rangle$'nin ikili (binary) uzaydaki $|0\rangle$ bazında bulunma olasılığı $\alpha^2 = \alpha * \alpha$ ve $|1\rangle$ bazında bulunma olasılığı $\beta^2 = \beta * \beta$ şeklindedir.

Kuantum bilgisayarlarda süperpozisyon durumunda olmayan herhangi bir durum Hadamard kapısı ile süperpozisyon durumuna getirilir.



Şek. 1: Blok Küre üzerinde q-bit gösterimi [1].

B. Dolanıklık (Entanglement)

Dolanıklık kuantum mekaniğine özgü bir olgudur. Örneğin $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ şeklinde yazılabilen Hilbert uzayı içinde $|\psi\rangle$ gibi durumlar düşünüldüğünde; eğer $|\psi\rangle$ durumu $|a\rangle \in \mathcal{H}_A, |b\rangle \in \mathcal{H}_B$ iki durumun tensörel çarpımı şeklinde yazılamıyorsa $|\psi\rangle$ durumunun dolanık bir hali temsil ettiği ifade edilir. Dolanık durumdaki bir sistemde birinde yapılan ölçüm anlık olarak diğerini de etkilemektedir. Bu etkileme aralarındaki mesafe çok uzak olsa dahi geçerlidir [1].

Kuantum bilgisayarlarda dolanık olmayan sistemler Hadamard ve CNOT kapısı uygulanarak dolanık hale getirilir.

Dolanıklık hızının ışık hızından fazla olduğu ile ilgili deneyler yapılmıştır ve ışık hızından daha hızlı bir hıza sahip olduğu gösterilmiştir.

- Testing spooky action at a distance [2]
- Bounding the speed of `spooky action at a distance [3]
- A strong loophole-free test of local realism (NIST) [4]

II. KUANTUM TEKNOLOJİLERİN ÜSTÜNLÜKLERİ

Kuantum teknolojilerin klasik teknolojilerde olmayan süper pozisyon ve dolanıklık gibi üstün özelliklerinin yanı sıra, bu iki özelliğe bağlı olarak aşağıdaki özelliklerin de gerçekleştirilmesi kuantum teknolojilerinin üstünlüğünü sağlamaktadır.

- Teleportasyon (Teleportation)
- Süperyoğun Kodlama (Superdense Coding)
- Dolanıklık Transferi (Entanglement Swapping)
- Terslenebilirlik (Reversible)
- Kopyalanamama Teoremi (Nocloning Theorem)

A. Teleportasyon(Teleportation)

Teleportasyon; madde ya da enerjinin veya herhangi bir durumun aralarındaki fiziksel alanı dolaşmadan, bir noktadan diğer bir noktaya anlık transferidir. Kuantum teleportasyon; bir konumdan diğer bir konuma mesafeye bağlı olmaksızın kuantum bilginin anlık olarak (ışık hızında), gönderen ve alıcının daha önceden paylaşmış olduğu kuantum dolanıklık ve klasik iletişim ile iletilmesi sürecidir.

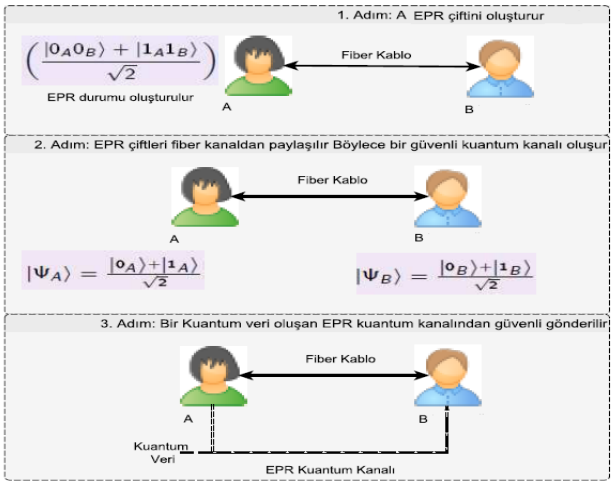
Diğer bir deyişle teleportasyon; bir kuantum verinin herhangi bir iletişim kanalı olmadan istenilen mesafeye aktarmaktır. Bu bize verinin dinlenememesini sağlar. Çünkü dinlenilecek kanal yok! Bunun için EPR durumları (dolanıkl) kullanılır. Yani iki katılımcı aşağıdaki EPR durumlarından birini paylaşarak dolanık hale gelmiş olur [1].

$$|\psi_{AB}\rangle = \left\{ \begin{matrix} |\phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, |\phi^-\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}} \\ |\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}, |\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}} \end{matrix} \right\}$$

Dolanıklık (örneğin EPR durumu) sadece kuantum düzeyde vardır ve iki kişinin paylaştığı verilerin korelasyona sahip olmasını ve birindeki değişiklerin diğerinde de anlık olarak mesafeden bağımsız olarak ortaya çıkmasına olanak sağlamaktadır.

EPR durumunun oluşumu için iki kuantum veri üzerinde belirli işlemler gerçekleştirilir. Daha sonra bu iki kuantum veri iki katılımcı arasında paylaşılır.

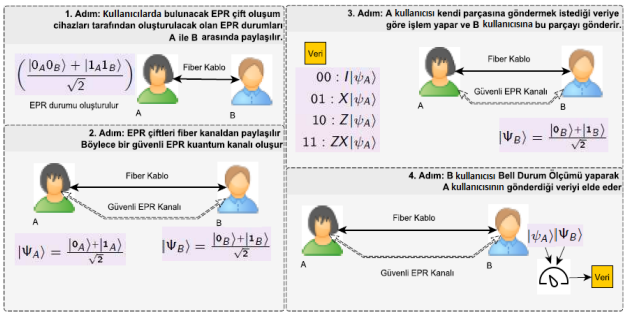
Bu paylaşım sonrası bir taraf diğerine herhangi bir iletim hattı olmadan bir kuantum veriyi aktarabilir. Bu kavrama teleportasyon(verinin ışınlanması) denilmektedir.



Şek.2. İki kişi arasında teleportasyon.

B. Süper-yoğun Kodlama (Superdense Coding)

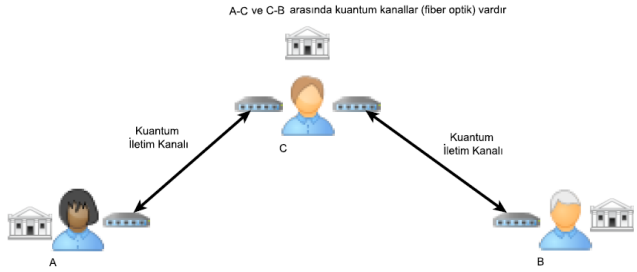
Kuantum bilgi teorisinde; süper-yoğun kodlama 2 klasik bit bilgisinin tek bir q-bit kullanılarak anlık (ışık hızında) gönderilmesidir [1].



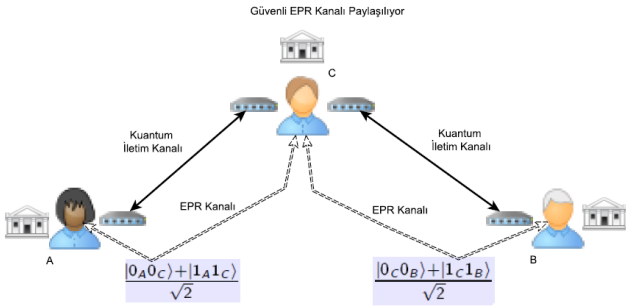
Şek. 3. İki kişi arasında süper yoğun kodlama

C. Dolanıklık Transferi (Entanglement Swapping)

Katılımcılar arasında $\frac{|00\rangle+|11\rangle}{\sqrt{2}}, \frac{|00\rangle-|11\rangle}{\sqrt{2}}, \frac{|01\rangle+|10\rangle}{\sqrt{2}}, \frac{|01\rangle-|10\rangle}{\sqrt{2}}$ dolanık durumlarından (Bell durumları) biri paylaşılır [1].



Şek. 4. Üç kişi arasındaki dolanıklık

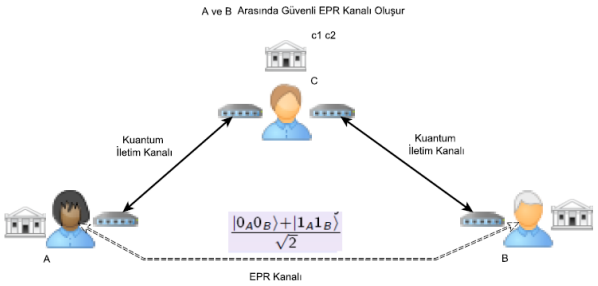


Şek. 5. C kişisinin dolanıklık ölçümü.

C kullanıcısı, A kullanıcısı ile olan EPR kanal parçasını Kuantum Teleportasyon ile B kullanıcısına aktarır. Bunun için Bell-Durum Ölçümü olarak adlandırılan bir fiziksel işlem gerçekleştirir.

Bell durum ölçümü sonucunda $c^1c^2 = \{00,01,10,11\}$ değerlerini elde eder. Bu ölçüm ile birlikte A kullanıcısı ve B kullanıcısı arasında daha önce var olmayan bir güvenli EPR kanal oluşmuş olur.

$$\frac{|0_A0_B\rangle + |1_A1_B\rangle}{\sqrt{2}}$$



Şek.. 6. C kişisinin Aile olan dolanıklığını B kişisine aktarması

D. Terslenebilirlik (Reversible)

Terslenebilirlik; herhangi bir duruma uygulanan bir olayın ilk durumu değiştirmesinden sonra, yeni duruma aynı olayın tersi tekrar uygulanmasında ilk durumun elde edilmesidir. Kuantum bilgisayarlarda ölçme işlemi çökmeye sebep olduğundan, ölçme operatörü dışında, q-bitler üzerinde uygulanan kapılar (işlemciler, operatörler) birimsel ($(U^*)^T = U^{-1}$) ve terslenebilir olmalıdır.

$$U|\psi\rangle = |\psi'\rangle \rightarrow U^{-1}|\psi'\rangle = |\psi\rangle$$

Örneğin 2 q-bitlik bir duruma CNOT kapısı uygulayalım daha sonra yeniden CNOT($CNOT^{-1}$=CNOT) uygulayıp ilk durumu elde edelim [1].

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$
$$|\psi'\rangle = CNOT|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|11\rangle + \delta|10\rangle$$
$$|\psi\rangle = CNOT|\psi'\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

E. Kopyalanamama Teoremi (Nocloning Theorem)

Veri iletiminden kuantum teknolojilerinin güvenli olmasının nedeni verinin kopyalanamamasıdır. Buna «no-cloning» denir. İletilen bir q-bitin 0 veya 1 olduğu iletim esnasında belirsizdir. Kopyalamak isteyen kişinin ölçme yapması gerekir. Buda belirli bir duruma çökmesine neden olur. Çöken durumun gönderilmek istenen durum olduğu kesin değildir.

Bilindiği gibi kuantum bilgisayarlarda uygulanan operatörlerin hemen hemen hepsi (ölçme operatörleri hariç) birimseldir. Operatörün birimsel olması demek $(U^*)^T = U^{-1}$ şartını sağlaması demektir. Buna göre herhangi bir U operatörü bir $|\psi\rangle$ durumunu kopyalayan birim dönüşüm operatörü olduğunu düşünelim. Bu durumda U operatörünün kuantum bilgisayarların temel bazları olan $|0\rangle, |1\rangle$ ve herhangi bir $|\psi\rangle$ durumuna etkisi;

$$U|0\rangle \rightarrow |00\rangle, \qquad U|1\rangle \rightarrow |11\rangle, \qquad U|\psi\rangle \rightarrow |\psi\psi\rangle$$

şeklinde olmalıdır. Ancak $U|\psi\rangle$ bu durum aşağıdaki gibi gerçekleşmekte ve bilginin kopyasını oluşturamamaktadır.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$
$$U|\psi\rangle = \alpha U|0\rangle + \beta U|1\rangle = \alpha|00\rangle + \beta|11\rangle$$

şeklinde olur.

Halbuki $|\psi\rangle$ durumunun kopyası $|\psi\psi\rangle$ durumunu oluşturabilseydi;

$$|\psi\psi\rangle = (\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle)$$
$$= \alpha^2|00\rangle + \alpha\beta|01\rangle + \beta\alpha|10\rangle + \beta^2|11\rangle$$

elde edilirdi [1].

Yukarıdaki ana dayanaklar kapsamında şu anda uygulanan ve geliştirilen teknolojileri kısaca aşağıdaki şekilde özetleyebiliriz.

III. KUANTUM BİLGİ VE İLETİŞİM

A. Kuantum Bilgisayarlar ve Çipler

Dünyanın ilk ticari kuantum bilişim şirketi D-Wave Systems [5]'dir. Lockheed Martin, Google, NASA, ABD Ulusal Güvenlik Ajansı, Los Alamos Ulusal Laboratuvarı gibi bir çok kuruluş D-Wave firmasının satmış olduğu bilgisayarı almışlardır. D-Wave Systems en son olarak «D-Wave 2000Q» adını verdikleri yeni kuantum bilgisayarı üretmişlerdir. Bu bilgisayarlar üretilirken işlemcinin dış ortamdan etkilenmesini önlemek için sıfır Kelvine yakın yeni bir soğutma teknolojisi geliştirmişlerdir. Bunlara ilave olarak işlemcinin korunduğu ortam Dünya manyetik alanından 50000 kat daha korumalıdır. Aynı zamanda yüksek vakumda basınç, atmosferik basınca göre 10 milyar kat daha düşüktür. Soğutucu ve sunucular sadece 25kW güç tüketmektedir [6,7].

Bununla birlikte IBM, uzun yıllardır üzerinde çalıştığı ve laboratuvarlarında geliştirmiş oldukları kuantum bilgisayarını herkesin kullanıma açmıştır. Bulut üzerinden çalışan bu bilgisayar, dünyanın her yerindeki öğrenciler, akademisyenler ve bilim adamları için kullanılabilmektedir.

IBM son yıllarda mevcut süper klasik bilgisayarlardan çok daha hızlı çalışabilen 50 q-bitlik, kuantum bilgisayarı geliştirdiğini duyurmuştur [8].

Bununla birlikte Google 72 q-bitlik kuantum bilgisayar çipini geliştirdiğini duyurmuştur [9].

Bristol, Tokyo, Southampton Üniversiteleri ve NTT Device Technology Laboratuvarlarında uluslararası bilim insanları ekibi tarafından, kuantum dolanıklığı oluşturan ve tespit edebilen kuantum teleportasyonun çekirdek devreleri fotonik çip haline getirilmiştir. Bu sonuçlar ultra hızlı kuantum bilgisayarların geliştirilmesine ve iletişim güvenliğinin güçlenmesine yol açmaktadır [10].

Referans [11] te kuantum bilgisayarların kullanım alanları ile birlikte güvenlik anlamında nelerin olabileceği ayrıntılı bir şekilde vurgulanmıştır.

NASA ve Google, 2013 yılında birlikte Kuantum Yapay Zeka Laboratuvarını (QuAIL) kurmuşlardır.

B. Kuantum Programlama Dilleri

Kuantum programlama mantığı ile programlama yapabileceğimiz bazı simülasyonlar şu şekildedir;

Quantum Pseudocode: E. Knill tarafından tanıtılan ve kuantum algoritmaların gerçekleştirilmesi için tasarlanan ilk dil olma özelliğine sahiptir. QRAM (Quantum Random Access Machine) adı verilen bir kuantum makine ile bağlantılıdır.

Q Language: Geliştirilen ikinci kuantum programlama dilidir. C++ programlama dilinin bir uzantısı şeklinde tasarlanmıştır.

QCL (Quantum Computing Language): ilk geliştirilen kuantum programlama dillerinden biridir. Yapısı C programlama diline çok benzemektedir. Sözdizimi, C programlama dili sözdizimini andırır ve klasik veri türleri, C'deki ilkel veri türlerine benzer. Klasik kod ile kuantum kod aynı programda birleştirebilir.

QCL'deki temel yerleşik kuantum veri türü qureg (quantum register) 'dir. Bir dizi q-bit (kuantum biti) olarak yorumlanabilir.

Quipper: 2013 yılında Haskell tabanlı çalışan gömülü bir dil olarak ortaya çıkmıştır. Bu nedenle Quiper ile yazılan kuantum programlar Haskell tarafından sağlanan kütüphaneleri kullanır.

LIQUi| (liquid): Microsoft bünyesindeki Quantum Architectures and Computation Group (QuArC) tarafından geliştirilmiştir. F# programlama dili üzerinde bir kuantum simülasyonudur. Liquid, fiziksel kuantum bilgisayarı olmadan teorisyenlerin kuantum algoritma tasarımlarını denemelerine imkan sağlar. Bu programlama dili, optimizasyon ve zamanlama algoritmaları ve kuantum simülatörleri içerir. Liquid, üst düzey bir program biçiminde yazılan bir kuantum algoritmayı bir kuantum cihaz için düşük seviyeli makine talimatlarına çevirmek için kullanılabilir.

QML: Altenkirch ve Grattage tarafından geliştirilmiş Haskell benzeri bir kuantum programlama dilidir.

Diğer kuantum programlama dilleri: Quantum Lambda Calculi, QFC, QPL, qGCL.

C. Kuantum işletim sistemleri ve derleyiciler

Kuantum derleyiciler hakkında genel düşünce; giriş olarak keyfi birimsel U operatörünü alan ve onun ayrıştırılmış halini döndüren bir bilgisayar programıdır. Çoğunlukla U operatörleri, Temel İşlemler Dizisi (SEO – Sequence of Elemantary Operations) olarak ifade edilmektedir. Temel işlemler kümesi

genellikle tek bir q-biti döndürmekte veya CNOT uygulamaktadır. SEO kuantum bilgisayarların makine dilidir. Keyfi birimsel matris demek, bilgisayar kullanıcısının daha önceden bilgisinin olmadığı bir yapıya sahip olmayan birimsel matris demektir. Özel amaçlı bir kuantum derleyici, bilinen bir yapıya sahip bir giriş birimsel matrisi U'yu bir SEO'ya parçalayan bilgisayar programıdır. Örneğin; önceden Kesikli Fourier Dönüşüm matrisi olarak bilinen bir U matrisini onun SEO'suna parçalayabilir [12].

D. Kuantum Kriptografi

Kuantum bilgi iletişiminde bilginin iletilmesi her ne kadar güvenli olsa da kuantum kriptografi ile ilgili çalışmalar da devam etmektedir. Günümüzde bilgi iletişimi fiber optik kablolar ile sağlandığından bu alandaki çalışmaların çoğu ışığın polarizasyonu kullanan kuantum anahtar üretim ve dağıtım cihazlarının geliştirilmesine odaklanmıştır. E bölümünde bu cihazlar kısaca özetlenmiştir. Bu cihazların çoğunda üretilen anahtarlar kesiklidir. Bugünlerde sürekli anahtar üreten cihazlar üzerine çalışılmaktadır. Bununla birlikte kriptoloji algoritmalrı üzerine de birçok çalışma yapılmaktadır[13,14].

Örnek çalışmalar;

- Multiparty quantum key agreement protocol secure against collusion attacks [15]
- Practical Attacks on Decoy State Quantum Key Distribution Systems with Detector Efficiency Mismatch [16]
- An Update On Quantum Cryptography [17]
- Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices [18]
- Quantum key distribution over multicore fiber [19]
- Optimizing Decoy State Enabled Quantum Key Distribution Systems to Maximize Quantum Throughput and Detect Photon Number Splitting Attacks with High Confidence [20]
- Quantum key distribution over 122 km of standard telecom fiber [21]
- New security notions and feasibility results for authentication of quantum data [22]
- Quantum Cryptography Beyond Quantum Key Distribution [23]
- Practical challenges in quantum key distribution [24]

E. Kuantum Bilgi İletişim Cihazları

Mevcut kuantum anahtar dağıtımı cihazları:

ID Quantique: http://www.idquantique.com/

Clavis3 Quantum Key Distribution platformudur. Manuel ve otomatik işlemlere izin verir ve akademik araştırmalar için geliştirilmiştir. 100 km ye kadar güvenli anahtar değişimi sağlar. Anahtar üretimi hızlıdır. Coherent One-Way protokol kullanır.

SECURENET: http://www.sequrenet.com/

Telecom Paris Tech firmasının çalışmalarıdır (http://www.secoqc.net/). Cygnus, geliştirdikleri QKD modülüdür. 25 - 80 km arasında çalışmaktadır.

ETSI (European Telecommunications Standards Institute):

http://www.etsi.org/technologiesclusters/technologies/quantum-key-distribution

Bu chaz FP6 Avrupa Birliği projeleri kapsamında kuantum kriptografi kullanılarak geliştirilen SECOQC(Secure Communication based on Quantum Cryptography) projesi ETSI standardıdır.

BATTLE: http://www.battelle.org/our-work/nationalsecurity/cyber-innovations/quantum-key-distribution

Bir Amerikan firması olup, ID Quantique firmasının ortağıdır ve onun cihazlarını kullanmaktadır. 2013'den beri QKD cihazları satmaktadır.

Magiqtech: http://www.magiqtech.com/Products.html

Q-BOX isimli QKD cihazı bulunmaktadır. BB84 temelli simetrik anahtarlama ile çalışmaktadır.

Toshiba: http://www.toshiba.eu/eu/Cambridge-Research-Laboratory/Quantum-Information-Group/Quantum-Key-Distribution/Toshiba-QKD-system/

Single Photon Detector geliştirmiştir ve 100km'ye kadar dolanık temelli QKD yapabilmektedir. Key Rate'i 1 MB/sn dir. BB84 ün değiştirilmiş versiyonu olan T12 protokolünü (pX > pZ bilginin X baz ile ifade edilme olasılığı Z baz ile ifade edilme olasılığından büyüktür) kullanmaktadır.

F. Kuantum Algoritmalar

Kuantum bilgi iletişiminde güvelik bağlamında BB84, B92,SARG gibi birçok algoritmanın cihazlar ile gerçekleştirilimi yapılmıştır. Bununla birlikte asal çarpanları üstel bir hızla bulan shor algoritması geliştirilmiştir. Ayrıca rastgele dizili bir veri grubunda istenileni çok hızlı bulabilen genlik yükseltme tabanlı Grover algoritması geliştirilmiştir.

G. Kuantum İnternet

Son yıllarda kuantum internet ve güvenliği üzerine birçok çalışma bulunmaktadır [25].

Bununla birlikte Çin 9 Ağustos 2016 tarihinde Gobi çölünden dünyanın ilk kuantum internet uydusunu fırlatmıştır. Bu test amaçlı uydu Çin'in kuantum internette uyduya çıkmasını sağlayacaktır [26].

NASA radyo sinyalleri yerine lazer ışınları kullanarak süper hızlı haberleşen ve süper ucuza üretilen "mini küp uyduları" test etmektedir. Amerika tarafından DARPA da gerçekleştirilen bir proje kapsamında 10 bağlantılı bir optik temelli kuantum internet gerçekleştirilerek kullanılmaktadır .

FP6 Avrupa birliği projeleri kapsamında kuantum kriptografi kullanılarak bir proje geliştirilmiştir. Bu bağlamda Avusturya'nın Viyana kentinde bir kuantum internet ağı kurulmuştur.

Benzer bir kuantum internet ağıda Tokyo da kurulmuştur. Bu ağ kullanılarak çok güvenli telekonferans görüşmesi gerçekleştirilmiştir.

Ayrıca çok kısa zamanda bu teknoloji telefonlara da uygulanabilecektir.

Örnek çalışmalar;

- Free space quantum communication with quantum memory [27]
- Covert Optical Communication [28]
- Key Pre Distribution Using Quantum Key Channel – A Survey [29]
- Field test of quantum key distribution in the Tokyo QKD Network [30]
- Quantum internet using code division multiple Access [31]

Ayrıca çok kısa zamanda bu teknoloji telefonlara da uygulanabilecektir.

IV. KUANTUM DETEKTÖRLER

A. Kuantum Radar

Kuantum radarı, kuantum dolanıklığına dayalı bir uzaktan algılama yöntemidir.

Kuantum radar prototipi mevcut teknolojiler ile gerçekleştirilebilir ve gizli nesnelerin uzaktan algılanmasından elektrik devrelerinin çevresel taramalarına kadar çeşitli potansiyel uygulamalara uygundur.

Kuvvetlendirilmiş hassaslığı sayesinde bu cihaz ayrıca protein spektroskopisi ve biyomedikal görüntüleme için teknikler sağlayabilir.

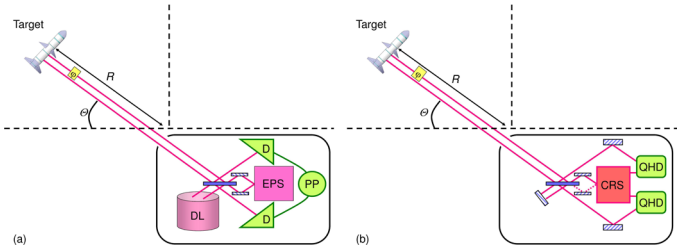
Klasik radarın sağlayabileceğinden daha iyi bir çözünürlük ve daha yüksek detay sağlayan bir radar sistemi yaratmak için alternatif yöntemler savunma müteahhidi Martin Lockheed tarafından da düşünülmektedir.

Çin'in önde gelen askeri elektroniği şirketi China Electronics Technology Group Corporation (CETC), Ağustos 2016 tarihinde, bilim adamlarının 100 km'lik bir kuantum radarı test ettiğini açıklamıştır. Bir kuantum radarı teorik olarak uzun menzilli gizli uçakları tespit edebileceği için bu önemli bir iddiadır.

Geleneksel radarlar hedefleri yansıtacak şekilde radyo dalgaları gönderirken, bir kuantum radarı; fiber bağlayıcılar, kuantum noktaları veya diğer yöntemlerle dolanık fotonları kullanmaktadır.

Dolanık fotonlar hedeflenen nesneden, geri dönüş süresine göre, nesnenin konumunu, radar kesitini, hızını, yönünü ve diğer özelliklerini gözlemleyebilen kuantum radarına geri dönerler.

Ayrıca, dolanık fotonları değiştirmeye veya çoğaltmaya yönelik herhangi bir girişim radar tarafından hemen tespit edilecektir.



Şek. 7. (a) Dolanık foton kaynağı(EPS) ve Foton sayısı çözümleyici detektörler(D) kullanmaktadır. (b) Tutarlı radar kaynağı(CRS) ve kuantum homodin detektör(QHD) kullanmaktadır [32].

Örnek çalışmalar;

- Design Considerations for Quantum Radar Implementation [33]
- Super-Resolving Quantum Radar: Coherent-State Sources with Homodyne Detection Suffice to Beat the Diffraction Limit [34]
- Range detection using entangled optical photons [35]

B. Kuantum Görüntüleme

Viyana Üniversitesi Kuantum Optik ve Bilgi Merkezi'nden bilim insanları, dolanıklıktan yararlanarak nesneye çarpıp geçen fotondan değil nesneye hiç çarpmayan fotondan görüntü elde etmeyi başardılar [36]. Bununla birlikte klasik ve kuantum görüntüleri analiz eden bir çok algoritma geliştirilmiştir (son çalışmaların da özetlendiği çalışma [37]).

C. Kuantum Sensörler

Ref.[38] daki bir grup bilim insanı tüm radyo sinyallerini ışık sinyallerine çeviren bir dedektör geliştirmişlerdir [39,40]. Bu aygıt maliyetli kriyojenik soğutma (-273 santigrata kadar soğutma) gerektirmemektedir ve radyo anstronomiden manyetik rezonans görüntülemeye kadar bir çok uygulamada pratik kullanım alanı bulabilmektedir. Araştırmacılar ayrıca bu teknolojinin geleceğin kuantum internetinin önemli bir yapı taşı olacağını düşünmektedirler

Kuantum Noktalar kullanılarak kuantum sensörler geliştirilmektedir.

Benzer şekilde Biyo-ajanlar eklenerek farklı tümör hücrelerini bulması için ayarlanmış kuantum noktaların UV ışığı altında görüntülenmesi de yapılmıştır. Kuantum noktaların hastalık teşhisinde çığır açması beklenmektedir.

Kuantum Noktalarının hassas olarak algılanması güvenlik uygulamalarından biyolojik uygulamalara kadar birçok farklı alanda ihtiyaç duyulmaktadır..

Benzer olarak sensör teknolojilerinde Karbon nanotüpler ve manyetik moleküller gelecekteki nanoelektrik sistemlerinin yapıtaşları olarak kabul edilmektedirler. Böyle bir sistem kuantum bilgisayarında q-bit olarak kullanılmayı mümkün kılmaktadır.

Örnek çalışmalar;

- Quantum Sensors: Improved Optical Measurement via Specialized Quantum States [39]
- Optimal and Secure Measurement Protocols for Quantum Sensor Networks [40]

- Quantum Sensing [41]

V. UZAY ARAŞTIRMALARI

Yukarıdaki bilgi ve teknolojiler kullanılarak; karadelik, karanlık madde, karanlık enerji gibi gizemli konulara cevaplar aranmaktadır. Böylece uzayın bilinmeyenleri hakkında cevaplar açıklığa kavuşturulabilecektir.

VI. ÖNERİLER

Yukarıdaki bilgilerden anlaşılacağı gibi kuantum teknolojiler hem hız hem de güvenlik anlamında çok büyük avantajlar sağlamaktadır. Bu nedenle kuantum teknoloji devriminin kaçırılmaması ve bilgi güvenliği bağlamında aşağıdakiler yapılabilir.

- Bilgisayar Mühendisliği bölümlerindeki Fizik dersleri yerine Modern fizik dersleri okutulabilir. Bu bölümlerde kuantum bilgisayar dersleri de okutulabilir.
- Nano teknoloji ve kuantum alanlarında farklı disiplinlerde (fizik, kimya, bilgisayar mühendisliği, matematik, biyoloji, tıp, uzay) çalışanlardan oluşan bir araştırma merkezi kurulmalıdır.
- Bu araştırma merkezi tarafından kuantum bilgisayar, kuantum çip, kuantum anahtar dağıtım sistemleri, kuantum görüntüleme ve dinleme sistemleri acilen üretilmelidir.
- Kuantum sistemler çevresel etkileşimlerden çok etkilendiği için atomik ve moleküler düzeyde kuantum bilgisayar ve kuantum çip oluşturulurken çevresel etkileşimlerden az etkilenen sekiz fosfat iyonundan oluşan Posner molekülleri ve çevresel etkileşimlerden hiç etkilenmeyen nötrinolar düşünülmelidir.
- Milli bir kuantum işletim sistemi geç kalmadan geliştirilmelidir.
- Milli bir kuantum programlama dili geliştirilmelidir.
- Kuantum kriptoloji alanında algoritmalar geliştirilmelidir.
- Bu çalışmaların tüm alanlara uygulanması acilen başlatılmalıdır.

REFERENCES

[1] D. McMahon, Quantum Computing Explained, John Wiley &Sons, Inc. Publication, 2008.

[2] D. Salart, A. Baas, C. Branciard, N. Gisin, H. Zbinden, Testing spooky action at a distance, *Nature* **volume454**, pages861–864, 2008.

[3] J. Yin, Y. Cao, H. Yong, J. Ren, H. Liang, S. Liao et.al., Bounding the speed of spooky action at a distance, Phys. Rev. Lett. 110, 26040 , 2013.

[4] L.K. Shalm, E. Meyer-Scott, B.G. Christensen, P. Bierhorst, M.A. Wayne, M.J. Stevens et.al.. A strong loophole-free test of local realism, **Phys.Rew.Lett**,115, 250402, 2015.

[5] Link:http://www.dwavesys.com/

[6] Link:http://www.dwavesys.com/d-wave-two-system

[7] Link: http://www.dwavesys.com/news/media-coverage

[8] link:https://futurism.com/ibm-announced-50-qubit-quantum-computer

[9] Link:https://quantumcomputingreport.com/news/google-announces-a-72-qubit-superconducting-quantum-chip

[10] Link:https://www.sciencedaily.com/releases/2015/04/150401114519.htm

[11] B.Arslan, M.Ulker, S. Akleylek and Ş. Sagiroglu, A Study on the Use of Quantum Computers, Risk Assessment and Security Problems" IEEE *Xplore*, 2018.

[12] Link::https://qbnets.wordpress.com/2009/04/17/brief-introduction-to-quantum-compilers/

[13] Link:https://www.slideshare.net/converse2006/bb84-0718

[14] Link:http://www.cki.au.dk/experiment/qrypto/doc/QuCrypt/b92coding.html

[15] S. Zhiwei, S. Xiaoqiang, W. Ping, Multiparty quantum key agreement protocol secure against collusion attacks, Quantum Inf Process,s.11128-017-1621,2017.

[16] F. Yangyang, G. Ming, W. Weilong, L. Chaobo, M. Zhi, Practical attacks on decoy state quantum key distribution systems with detector efficiency mismatch, Pys.Rew A,91, 052305, 2015.

[17] C.H. Bennett, G. Brassard, An update on quantum cryptography, G.R. Blakley and D. Chaum (Eds.): Advances in Cryptology - CRYPTO '84, LNCS 196, pp. 475-480, 1985.

[18] C.A. Miller, Y. Shi, Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices, Journal of the ACM, Vol. 63, Issue 4, Article No. 33, 2016.

[19] J.F. Dynes, S.J. Kindness, S.W.B. Tam, A. Plews, A. W. Sharpe, M. Lucamarini et.al, Quantum key distribution over multicore fiber, Optics Expres, vol.24,issue 8,pp.8081-8087,2016.

[20] L.O. Mailloux, M.R. Grimaila, D.D. Hodson, R. Engle, C. McLaughlin, G. Baumgartner, Optimizing decoy state enabled quantum key distribution systems to maximize quantum throughput and detect photon number splitting attacks with high confidence,link:https://arxiv.org/ftp/arxiv/papers/1606/1606.07313.pdf, 2016.

[21] C. Gobby, Z. L. Yuan, A. J. Shields, Quantum key distribution over 122 km of standard telecom fiber, Appl. Phys. Lett. **84**, 3762,2004.

[22] S. Garg, H. Yuen, M. Zhandry. New security notions and feasibility results for authentication of quantum data, Advances in Cryptology – CRYPTO 2017, pp 342-37, 2017.

[23] A. Broadbent, C. Schaffner. Quantum cryptography beyond quantum key distribution, Designs, Codes and Cryptography, Volume 78, Issue 1, pp 351–382,2016.

[24] E. Diamanti, H.K. Lo, B. Qi, Z. Yuan, Practical challenges in quantum key distribution, *npj Quantum Information* volume2, Article number: 16025, 2016.

[25] V. Makarov, J.P. Bourgoin, P. Chaiwongkhot, M. Gagne, T. Jennewein, S. Kaiser et.al., creation of backdoors in quantum communications via Laser damage, Phys.RevA 94, 030302,2016.

[26] S.K.Liao et all. Satellite-to-ground quantum key distribution, doi:10.1038/nature23655,nature,2017.

[27] M. Namazi, G. Vallone, B. Jordaan, C. Goham, R. Shahrokhshahi, P. Villoresi et.al., Free space quantum communication with quantum memory, Phys.Rew.Aplied 8,064013, 2016.

[28] B.A. Bash, A.H. Gheorghe, M. Patel, J.L. Habif, D. Goeckel, D. Towsley et.al., Covert optical communication, Nature Communications 6: 8626, 2014.

[29] MD. Sarwar Pasha, A. Bala Ram, Key pre distribution using quantum key channel, International Journal of Computer Science and Mobile Applications, Vol.2 Issue. 3, 2014.

[30] M. Sasaki, Field test of quantum key distribution in the tokyo QKD network, Optical Express, vol.19.,issue 11,pp.10387 2011.

[31] J. Zhang, Y. Liu, Ş.K. Özdemir, R. BingWu, F. Gao, X. Wang et.al., Quantum internet using code division multiple access, *Scientific Reports* **volume3**, Article number: 2211 , 2013.

[32] K. Jiang, H. Lee, C. Gerry, J.P. Dowlin, Super-resolving quantum radar: coherent-state sources with homodyne detection suffice to beat the diffraction limit, Journal of Applied Physics **114,2013**.

[33] M.J. Brandsema, R.M. Narayanan, M. Lanzagorta, Design considerations for quantum radar implementation, Proceedings Volume 9077, Radar Sensor Technology XVIII; 90770, 2014.

[34] K. Jiang, H. Lee, C. Gerry, J.P. Dowling, Super-resolving quantum radar: coherent-state sources with homodyne detection Suffice to beat the diffraction limit, Journal of Applied Physics **114**, 193102, 2013.

[35] M.J. Brandsemaa, R.M. Narayanana, M. Lanzagorta, Range detection using entangled optical photons, Radar Sensor Technology XIX; and Actve and Passve Sgnatures VI, 946111, 2015.

[36] G.B. Lemos, V. Borish, G.D. Cole, S. Ramelow, R. Lapkiewicz, A. Zeilinger, "Quantum imaging with undetected photons", Nature, 512 (7515), 2014.

[37] E. Şahin & İ. Yılmaz, "Qrmw: Quantum Representation Of Multi Wavelength Images", Turkısh Journal Of Electrical Engineering & Computer Sciences, 1300-0632, 26, 2, 768-779, 2018.

[38] T.Bağcı et all. " Optical detectation of Radio waves through a nanomecanical transcuder" **507**, pages81–85,2014.

[39] D.S. Simon, Quantum sensors: improved optical measurement via specialized quantum state, Journal of Sensors, article ID 6051286, 2015.

[40] Z. Eldredge, M. Foss-Feig, S.L. Rolston, A.V. Gorshkov, Optimal and secure measurement protocols for quantum sensor networks, Phys. Rev. A 97, 042337, 2018.

[41] C. L. Degen, F. Reinhard, P. Cappellaro, Quantum sensing, Rev.Mod.Phys. 89,035002, 2017.

Latis Tabanlı Kriptografi Algoritmalarının Hız Testi

Performance Comparison of Lattice Based Cryptosystems

Damla Acar
Hacettepe Üniversitesi,
Matematik Bölümü
Ankara, Türkiye
damla.acar@hacettepe.edu.tr

Oğuz Yayla
Hacettepe Üniversitesi,
Matematik Bölümü
Ankara, Türkiye
oguz.yayla@hacettepe.edu.tr

***Özet*— Bu çalışmada latis tabanlı kriptografi algoritmaları incelenmiştir. Latis tabanlı kriptografinin temelinde Ajtai tarafından bulunan latis problemlerinin olağan durum ile en kötü durumdaki zorlukları arasında yer alan ilişki vardır. Bu çalışmada ele alınacak latis tabanlı algoritmalar sırasıyla [1] ve [2]'de tanıtılan SIS (Small Integer Solution Problem) ve LWE (Learnin With Errors Problem)'ye dayanmaktadır. Latis tabanlı kriptografinin gündeme gelmesinin sebepleri arasında hesaplamaların oldukça basit olması ve genelde sadece modüler çarpma işlemine dayanması vardır. Bu ise düşük maliyetli cihazlarda şifreleme için avantajlı olabilir. Bir diğer sebep ise RSA [3] gibi geleneksel sayı teorisi tabanlı kriptografinin çok fazla alternatifi olmamasıdır. Günümüzde kullanılan şifreleme algoritmalarının dayandığı problemler kuantum bilgisayarlar tarafından kısa sürede çözülebileceği için bu alternatif algoritmalar önem kazanmaktadır. Ayrıca zor latis problemlerini çözebilecek polinom zamanlı kuantum algoritmalar henüz mevcut değildir. Bu çalışmada NIST'in 2017 yılında yaptığı çağrıya gönderilen ve ilk aşamayı geçen latis tabanlı kriptografi algoritmaları zamanlamaları açısından kıyaslanmıştır**

***Anahtar Kelimeler*—Latis-tabanlı kriptografi algortimaları, şifreleme, kuantum sonrası kriptografi.**

***Abstract*— In this paper we study lattice based cryptographic algorithms. These algorithms are based on the hardness of the worst-case and average-case of lattice problems. The lattice-based algorithms to be discussed in this study are based on SIS introduced in [1] and LWE (Learning With Errors Problem) introduced in [2]. Calculations in lattice-based cryptography are quite simple and generally based on modular multiplication. This can be advantageous for encryption on low-cost devices. Therefore lattice-based cryptography is discussed. In addition, there is not so much alternative to traditional number theory-based cryptography such as RSA [3]. The problems of encryption algorithms used today will be solved by quantum computers in a short time. Therefore these alternative algortihms have become very important. Also polynomial time quantum algorithms that can be solve difficult lattice problems are not yet available. In this paper we study lattice-based cryptographic algorithms, which passed the first the first step nd which were send to the call of NIST in 2017 were compared in terms of their timing.**

***Index Terms*—Lattice-based cryptographic algorithms, encryption, post quantum cryptography.**

I. GİRİŞ

Kuantum bilgisayarların gelmesiyle günümüzdeki kriptografik sistemlerin çoğunun kırılacağı bilinmektedir. Bu yüzden kuantum bilgisayarlara dayanıklı yeni algoritmalar üretilmesi gerekmektedir. Bununla birlikte, kuantum bilgisayarlardaki ilerlemeler, şu anda kullanılan genel anahtar şifreleme algoritmalarının dayandığı güvenlik varsayımlarını zayıflatmakta ve tehdit etmektedir [7]. Bu yüzden dünya genelinde kuantum sonrası(post quantum) için en etkili ve en güvenli algoritmayı geliştirme yarışı vardır [7]. Bu algoritmalar genelde latis tabanlı, kod tabanlı ve çok değişkenli polinom tabanlı matematiksel yapılara dayanmaktadır [5]. Latis tabanlı algoritmalar ile üretilen açık anahtarlı kriptografik sistemler, kuantum sonrası kriptografik algoritmalar için en güçlü seçeneklerden biri olarak ortaya çıkmıştır. İnternet üzerinden işlem yapmak için yaygın olarak açık anahtar algoritmaları kullanılır [6].

Bu çalışmada NIST'in 2017 yılında yaptığı çağrıya [4] gönderilen ve ilk aşamayı geçen anahtar kapsülleme (kriptografik anahtar üretimi), şifreleme ve imzalama algoritmaları zamanlamaları açısından kıyaslanmıştır.

II. LATİS TABANLI KRİPTOGRAFİ

Kuantum bilgisayarlara dayanıklı bir algoritma geliştirmek için NP-tam ya da NP-zor olan problemlere ihtiyaç vardır. Kuantum bilgisayarlar günümüzde kullanılan açık anahtar kriptografi algoritmalarını (RSA, ECDSA, DSA, DH vb) kırabilecektir. Ayrıca, simetrik anahtar kriptografisi (AES, TDES vb.) daha uzun anahtara ihtiyaç duyacaktır. Özet fonksiyonlar (SHA-1, SHA-2, SHA-3 vb.) ise daha uzun çıktı vermeleri gerekecektir. Aşağıdaki tabloda [5] bu karşılaştırmaları görebilirsiniz.

Latis tabanlı kriptografik yapılar güvenlik açısından ikiye ayrılır. Bunlar, uygulama açısından oldukça pratik ve etkili olup güvenlik açısından zayıf olan yapılar ile latis problemlerinin en kötü durum zorluğuna dayanıklı olan kriptografik yapılardır. Ancak bu yapılardan sadece birkaç tanesi pratikte uygulanabilirdir. İkinci grupta yer alan yapılar güçlü bir güvenlik garantisi sunarlar. Çünkü kriptografik bir yapının kırılması altta yatan herhangi bir latis problemini çözen bir algoritma gerektirir. Bu ise latis tabanlı kriptografinin ayırt edici özelliklerindendir. Diğer kriptografik yapıların hemen hemen hepsi orta durum zorluğuna dayanır.

TABLO I. GÜVENLİK SEVİYELERİ

Algoritma	Fonksiyon	Kuantum Öncesi	Kuantum Sonrası
Simetrik			
AES-128	Şifreleme	128	64(Grover)
AES-256	Şifreleme	256	128(Grover)
Salsa20	Şifreleme	256	128(Grover)
GMAC	MDK	128	128
Poly1305	MDK	128	128
SHA-256	Özet	256	128(Grover)
SHA3-256	Özet	256	128(Grover)
Açık Anahtarlı			
RSA-3072	Şifreleme	128	Kırıldı(Shor)
RSA-3072	İmzalama	128	Kırıldı(Shor)
DH-3072	Anahtar Değişimi	128	Kırıldı(Shor)
DSA-3072	İmzalama	128	Kırıldı(Shor)
256-bit ECDH	Anahtar Değişimi	128	Kırıldı(Shor)
256-bit ECDSA	İmzalama	128	Kırıldı(Shor)

Bu çalışmada kuantum bilgisayarlara dayanıklı latis tabanlı problemlere dayanan kriptosistemler incelenmiş ve hız karşılaştırmaları yapılmıştır. Bu algoritmaları yapabildikleri işleve göre üç gruba ayırabiliriz: anahtar kapsülleme, şifreleme, imzalama.

Hız karşılaştırması yapılan test ortamı CentOS 7, 1 GB RAM, 25 GB SSD ve 1vCPU(2.4Ghz) özellikleri olan bir sanal bilgisayardır.

A. Anahtar Kapsülleme Algoritmaları

Anahtar kapsülleme, asimetrik (public-key) algoritmalar kullanılarak; simetrik kriptografide kullanılmak üzere tasarlanmış bir şifreleme tekniğidir. Pratikte, açık anahtar sistemleri uzun mesajların iletilmesinde kullanışsızdır. Simetrik anahtarlar mesajı şifrelemek için kullanılır; ancak bu anahtarların paylaşımı zor olduğu için, tam güvenlik sağlamada yetersiz kalırlar. Açık anahtar sistemi bu noktada devreye girer ve simetrik anahtarların değişimi için kullanılırlar [5]. Kapsülleme yapabilmek için önce iki taraf da açık anahtar ve gizli anahtar ikililerini üretmiş olması gereklidir. Gönderen kişinin gizli anahtarı ve alacak kişinin açık anahtarı kapsülleme algoritmasına girer ve karşı tarafa kapsüllenmiş veri (simetrik anahtar) gönderilir. Karşı taraf ise bu kapsüllenmiş veriyi kendi gizli anahtarı ve karşı tarafın açık anahtarı ve dekapsülleme algoritmasının yardımıyla öğrenir.

Kuantum sonrası kriptografik anahtar değişimi için latis tabanlı bazı anahtar kapsülleme yöntemleri bulunmaktadır. Bu çalışmada NIST tarafından açılan yarışmaya katılan latis tabanlı birçok anahtar kapsülleme algoritmaları incelenmiş; ayrıca algoritmaların hızları birbiriyle kıyaslanmıştır. İncelenen algoritmalar, NIST tarafından açılan yarışmada birinci turu geçen algoritmalardan, KINDI, LIMA, TITANIUM, NTRUencrypt, Round2, LAC, OKCN/AKCN/CNKE, NewHope, NTRU-HRSS-KEM, Odd Manhattan, SABER, Three Bears, HILA 5 algoritmalarıdır. Aşağıda bu algoritmaların açık/gizli anahtar çiftinin üretimi, kapsülleme ve dekapsülleme algoritmalarının hız kaşılaştırmalarını mikrosaniye türünden sunuyoruz. Bu hız karşılaştırmasında herbir algoritma ile 100 farklı anahtar üretilmiş ve süreleri ölçülmüştür. Bu ölçümlere göre herbir algoritmanın 100 adet anahtar üretiminde geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo 2'de verilmiştir.

TABLO II. ANAHTAR ÜRETME

Algoritma	Maksimum	Minimum	Ortalama
NTRU443	3009	2565	2620
NewHope	355	267	281
NTRUHRSS	116198	89765	91600
Odd Manhattan	167879	137294	142549
AKCN-MLWE	282	189	196
Round2	4005	2731	2892
SABER	1350	897	923
Three Bears	1447	896	935
Titanium	99892	89255	90583
Lima	28507	4975	5352
Lac	264	149	163
Kindi	439	138	175
Hila5	13921	11691	11858

Tablo II'de ortaya çıkadan değerlere göre anahtar üretme işleminde ortalama değerlere göre en hızlı olan LAC algoritmasıdır. Diğer taraftan Kindi algoritması ise Minimum değerlere bakıldığında en kısa sürede anahtar üretilmiştir. Ancak aynı algoritmanın Maksimum değerine bakıldığında Minimum değerinden iki katından daha uzun sürdüğü görülmektedir. Bu analiz de Kindi'nin Anahtar üretme sürelerinin farklı girdilere göre oldukça farklı süreler tuttuğunu söyleyebiliriz. Anahtar üretme işleminde en yavaş olan algoritma ise Odd Manhattan'dır. Algoritmanın hızlandırılması yönünde yapılan optimizasyon işlemlerinden yarı Mersenne asallarının kullanılması ve hesaplama paylaşma yönteminin yeterince etkili olamadığı söylenebilir.

Latis tabanlı anahtar kapsülleme algoritmaları ile bir anahtar herbir algoritma için 100 farklı kapsülleme işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo III'te verilmiştir.

Tablo III'te yer alan verilerden kapsülleme işleminde Maksimum değerlerde en hızlı olan algoritmanın AKCN-MLWE olduğu söylenebilir. Ancak Minimum ve Ortalama değerlerde ise Kindi'nin oldukça hızlı olduğu görülmektedir.

Dolayısıyla KINDI algoritması farklı değerler için farklı sonuçlar üretmektedir. En yavaş algoritmalar ise Odd Manhattan'dır.

TABLO III. KAPSÜLLEME

Algoritma	Maksimum	Minimum	Ortalama
NTRU443	594	403	421
NewHope	355	267	281
NTRUHRSS	3236	1849	1903
Odd Manhattan	79232	58992	63242
AKCN-MLWE	272	225	233
Round2	7678	5407	5649
SABER	1617	1169	1198
Three Bears	1813	1181	1219
Lima	12625	10847	11113
Titanium	70144	62490	63118
Lac	1141	673	726
Kindi	453	179	223
Hila5	23454	14639	15004

Latis tabanlı anahtar kapsülleme algoritmaları ile kapsülleme işlemi uygulanan bir anahtar herbir algoritma için 100 farklı dekapsülleme işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo IV'te verilmiştir.

Tablo IV incelendiğinde en hızlı algoritmanın tüm değerlerde AKCN-MLWE olduğu görülmektedir. Titanium algoritması ise NTT (Number Theoretic Transform) olmasına rağmen en yavaş algoritmadır.

B. Şifreleme Algoritmaları

NIST'in sayfasında [4] da yer aldığı üzere latis tabanlı şifreleme algoritmalarının bazıları Compact LWE, Giophantus, KINDI, LAC, LIMA, NTRUEncrypt, OKCN/AKCN/CNKE, Round2 ve Titanium'dur. Şifreleme yapabilmek için açık/gizli anahtar çiftinin karşılıklı üretilmiş olması gerekmektedir. Yukarıda adı geçen algoritmaların aşağıda anahtar üretimi ve şifreleme hız kaşılaştırmalarını mikrosaniye biriminde sunuyoruz. Bu hız karşılaştırmasında herbir algoritma ile 100 farklı anahtar üretilmiş ve sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo V'te verilmiştir.

Tablo V'te yer alan verilere göre maksimum değerlerde en hızlı algoritma LAC'tır ancak minimum ve ortalama değerlerde en hızlı KINDI algoritmasıdır. KINDI'nin maksimum değeri ile minimum değeri arasındaki fark oldukça fazladır. Buna karşın LAC algoritmasında bu fark daha azdır. Dolayısıyla anahtar üretimi için LAC algoritmasının daha etkili olacağı söylenebilir. En yavaş algoritma ise Titanium'dur.

TABLO IV. DEKAPSÜLLEME

Algoritma	Maksimum	Minimum	Ortalama
NTRU443	688	577	592
NewHope	615	456	476
NTRUHRSS	7035	5406	5529
odd manhattan	96630	66804	71010
AKCN-MLWE	71	49	51
Round2	3635	2647	2753
SABER	1923	1419	1456
Three Bears	3768	2447	2522
Lima	16658	12371	12667
Lac	2017	1200	1291
Kindi	546	224	277
Hila5	4499	2875	2957

TABLO V. ANAHTAR ÜRETME

Algoritma	Maksimum	Minimum	Ortalama
Compact LWE	2687	1528	1649
KINDI	312	149	156
LAC	272	162	168
LIMA	30720	4934	5646
NTRUEncrypt	3412	2809	2883
OKCN/AKCN	496	291	308
Round2	4347	3285	3457
Titanium	114702	101453	102868
Giophantus	34842	32629	33222

Latis tabanlı şifreleme algoritmaları kullanılarak ele alınan bir mesaj herbir algoritma ile 100 farklı şifreleme işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo VI'da verilmiştir.

Tablo VI'daki değerlerden en hızlı algoritmanın tüm değerler için KINDI olduğu elde edilir. LAC ise anahtar üretiminde hızlı olmasına rağmen şifreleme işleminde daha yavaştır. En yavaş algoritma ise Titanium'dur. Titanium kendisinden önceki en yavaş algoritmadan üç kat daha yavaştır. Buradan algoritmayı hızlandırmak için yapılan optimizasyon işlemlerinin yeterli olmadığı sonucu elde edilir.

Latis tabanlı şifreleme algoritmaları kullanılarak şifreli bir mesaj herbir algoritma ile 100 farklı deşifreleme işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo VII'de verilmiştir.

TABLO VI. ŞİFRELEME

Algoritma	Maksimum	Minimum	Ortalama
Compact LWE	986	371	531
KINDI	241	181	187
LAC	1205	726	753
LIMA	13511	10763	11485
NTRUEncrypt	433	340	365
OKCN/AKCN	589	347	373
Round2	8054	6574	6816
Titanium	76412	70714	71592
Giophantus	6548	70686	71998

Tablo VII ile şifre çözme işleminde Maksimum, Minimum ve Ortalama değerlerin tümünde en hızlı olan algoritma Compact-LWE olduğu elde edilir. Oldukça hızlı olmasına rağmen Maksimum ve Minimum değerleri arasındaki fark fazladır. KINDI ise Compact-LWE'den iki kat daha yavaş olmasına rağmen Maksimum ve Minimum değerler açısından daha tutarlıdır. Bu yönü ile KINDI şifre çözme işlemi için daha uygun olduğu söylenebilir. En yavaş algoritma ise Giophantus'tur.

TABLO VII. DEŞİFRELEME

Algoritma	Maksimum	Minimum	Ortalama
Compact LWE	151	61	89
KINDI	362	227	238
LAC	930	566	582
LIMA	1788	1519	1633
NTRUEncrypt	720	619	634
OKCN/AKCN	678	396	423
Round2	11412	9803	10118
Giophantus	152596	133833	135359

C. İmzalama Algoritmaları

NIST'in sayfasında [4] da yer aldığı üzere latis tabanlı imzalama algoritmaları pqNTRUsign, DRS, Falcon, qTesla'dır. İmzalama yapabilmek için açık/gizli anahtar çiftinin karşılıklı üretilmiş olması gerekmektedir.

Aşağıda bu algoritmaların açık/gizli anahtar çiftinin üretimi, imzalama ve imza çözme algoritmalarının hız kaşılaştırmalarını mikrosaniye türünden sunuyoruz. Bu hız karşılaştırmasında herbir algoritma ile 100 farklı anahtar üretilmiş ve süreleri ölçülmüştür. Bu ölçümlere göre herbir algoritmanın 100 adet anahtar üretiminde geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo VIII'de verilmiştir.

Tablo VIII'deki değerlerden anahtar üretme işlemi için tüm değerlerde en hızlı algoritmanın qTesla olduğu söylenebilir. Tüm değerlerde en yavaş algoritma ise pqNTRUsign algoritmasıdır.

TABLO VIII. ANAHTAR ÜRETME

Algoritma	Maksimum	Minimum	Ortalama
PqNTRUsign Uniform	93398	75483	79776
PqNTRUsign Gaussian	103511	75674	80755
DRS	603470	476656	496395
Falcon	114126	31915	49056
qTESLA	12789	2591	4239

Latis tabanlı imzalama algoritmaları kullanılarak ele alınan bir mesaj herbir algoritma ile 100 farklı imzalama işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo IX'da verilmiştir.

Tablo IX'da yer alan verilerden imzalama işleminde Maksimum değerlerde en hızlı algoritmanın Falcon oldduğu sonucuna ulaşılır. Ancak Minimum ve Ortalama değerlerde en hızlı algoritma qTesla'dır. Falcon'un maksimum ve minimum değeri arasındaki az olmasına rağmen bu fark qTesla'da oldukça fazladır. Dolayısıyla Falcon imzamala algoritmaları arasında en uygun seçenek olarak ortaya çıkmaktadır

TABLO IX. İMZALAMA

Algoritma	Maksimum	Minimum	Ortalama
PqNTRUsign Uniform	493308	6438	122064
PqNTRUsign Gaussian	1005870	6390	204379
DRS	62928	31679	37906
Falcon	6673	4781	5134
qTESLA	19814	1052	3198

Latis tabanlı imzalama algoritmaları kullanılarak imzalı bir mesaj herbir algoritma ile 100 farklı imza doğrulama işleminden geçirilmiştir. Bu işlem sırasında geçen sürelerin maksimum, minimum ve ortalama değerleri mikrosaniye türünden Tablo X'da verilmiştir.

Tablo X'da yer alan değerlerden imza doğrulama işleminde en hızlı olan algortimanın tüm değerlerde Falcon olduğu elde edilir. En yavaş algoritma ise DRS'dir.

TABLO X. İMZA DOĞRULAMA

Algoritma	Maksimum	Minimum	Ortalama
PqNTRUsign Uniform	2398	1546	1619
PqNTRUsign Gaussian	2364	1543	1636
DRS	351066	233338	247198
Falcon	600	285	371
qTESLA	1211	717	765

Genel olarak tüm tablolar incelendiğinde hızlı olan algoritmaların halka-LWE (Ring-LWE) problemine dayalı kriptosistemler olduğu görülmektedir.

III. SONUÇ

Post-kuantum kriptografisinde araştırma günden güne artmaktadır. NIST tarafından oluşturulan rekabet - daha fazla tasarım, daha fazla optimizasyon ve uygulama ve daha fazla saldırı olarak dönmektedir. Bu noktada önemli olan hangi sistemler saldırılara karşı savunmasız kestirebilmektir. Yaygın olarak kullanılabilen ve aynı zamanda güvenilebilen post-kuantum sistemleri oluşturmak için çalışmalara, testlere, analizlere, ataklara ihtiyaç vardır.

TEŞEKKÜR

Birinci yazar YÖK 100/2000 bursu kapsamında desteklenmektedir. İkinci yazar TUBİTAK 117E636 nolu proje tarafından desteklenmektedir.

KAYNAKLAR

[1] M. Ajtai, "Generating hard instances of lattice problems" In Proceedings of the twenty-eighth annual ACM symposium on theory of computing, pp. 99-108, ACM, July, 1996.

[2] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography", Journal of the ACM (JACM), 56(6), 34, 2009.

[3] R. Rivest, R. L.,A. Shamir, L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems" Communications of the ACM, 21(2), 120-126, 1978.

[4] Post Quantum Cryptography, Round 1 Submission, https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-1-Submissions, Nisan, 2018.

[5] D. J. Bernstein, "Introduction to post-quantum cryptography", In post-quantum cryptography, pp. 1-14, Springer, Berlin, Heidelberg, 2009.

[6] C. Peikert, "Lattice cryptography for the internet", In International workshop on post-quantum cryptography, pp. 197-219, Cham, 2014.

[7] L. Chen, L. Chen, S. Jordan, YK Liu, D. Moody, R. Parelta, et al. D. Smith-Tone, "Report on post-quantum cryptography", US Department of Commerce, National Institute of Standards and Technology, 2016.

Popüler 500 Web Sitesinin Sertifika Şeffaflığı Uyum Değerlendirmesi

Certificate Transparency Conformity Assessment of Top 500 Websites

Erhan TURAN
WISE Lab.,
Hacettepe University,
Ankara, Turkey
erhan.turan@hacettepe.edu.tr

Tamer ERGUN
e-Signature Technologies, Kamu SM, BİLGEM,
TÜBİTAK,
Ankara, Turkey
tamer.ergun@tubitak.gov.tr

Sevil SEN
WISE Lab.,
Hacettepe University,
Ankara, Turkey
ssen@cs.hacettepe.edu.tr

***Abstract*— Secure Socket Layer (SSL) certificates are electronic files used to encrypt data flow between clients and servers and to verify the identity of websites. SSL certificates are published by Certificate Authorities (CA) that are considered to be completely trustworthy. However, it is necessary to check whether or not a certificate has been accidentally issued by a CA without the user's consent. The Certificate Transparency (CT) Project, developed by Google, aims to satisfy this need within the SSL certificate validation system and offers an open framework to monitor and audit SSL certificates. Chrome requires that all TLS server certificates issued after April 30, 2018 must be compliant with the Chromium CT Policy. In this current study, we investigate whether or not websites and CAs are following this policy. The most popular 500 websites were therefore checked for their CT compliance with the methods that they use.**

Keywords—Certificate Transparency, SSL, Public Key Infrastructure

Özet*— *Güvenli Yuva Katmanı (SSL) sertifikaları, istemciler ve sunucular arasındaki veri akışını şifrelemek ve web sayfalarının kimliğini doğrulamak için kullanılan elektronik dosyalardır. SSL sertifikaları, güvenilir Elektronik Sertifika Hizmet Sağlayıcıları (ESHS) tarafından üretilirler. ESHS'ler ne kadar güvenilir olsalar da, bir sertifikanın ESHS tarafından kullanıcının izni olmaksızın (kazara) üretilip üretilmediğinin de kontrol edilmesi gerekmektedir. Google tarafından geliştirilen Sertifika Şeffaflığı (CT) Projesi, SSL doğrulama mekanizmasında bu ihtiyacı karşılamak ve SSL sertifikalarının takibini/ denetlenebilirliğini sağlamak için açık bir sistem sunmayı amaçlamaktadır. Chrome, 30 Nisan 2018'den sonra yayınlanan tüm SSL/TLS sunucu sertifikalarının Chromium CT Politikası ile uyumlu olmasını zorlamaktadır. Bu çalışmada, web sitelerinin ve ESHS'lerin bu politikaya ne kadar uyduğu araştırılmıştır. Bu doğrultuda, en popüler 500 web sitesinin uyumluluğu, kullandıkları CT yöntemleri de ele alınarak kontrol edilmiştir.

Anahtar Kelimeler—Sertifika Şefaflığı, SSL, Açık Anahtar Altyapısı

I. INTRODUCTION

SSL certificates provide trust-based web security by establishing secure connections between clients and servers. These certificates need to be validated before their use. The steps of the certificate validation system that web browsers use to verify websites' SSL/TLS certificate chain is specified in RFC 5280 [1]. By using this validation system, browsers can detect erroneous certificates such as those that have expired, where they have been signed by a fake authority, or if they have been revoked [2]. There are also problems associated with certification authorities [3]. However, identifying violations of trusted CAs is difficult. In some cases, such fraudulent attempts cannot be detected for weeks or even months.

There have been some faults regarding SSL certifications in recent years. For example, a Dutch certification authority (DigiNotar) was compromised and hackers used the cryptographic system of the certificate authority to generate fake SSL certificates [5]. The Internet sites used for spying in Iran have been presented to users as popular websites such as Gmail and Facebook. Following this event, the certificates issued by DigiNotar were revoked and the certificate authority has since been closed down. In another example, a Malaysian sub-root certificate authority (DigiCert Sdn. Bhd. Sub-root of the Entrust certificate authority) issued certificate revocation information and 22 weak signing certificates without the Extended Key Usage field [6]. Two of these certificates were used to sign malicious software that was employed in phishing attacks against an Asian certification authority. As a result of this, browsers deployed updates and all certificates issued by this CA were removed from their trusted root lists.

Certificate Transparency (CT) focuses on fraudulent attempts that are hard to detect using the existing certificate validation system; making it possible to detect certificates issued in error or by malicious intent, and to identify the issuing certification authority [4]. It is important for audit purposes to detect incompatibilities and vulnerabilities that can occur on the part of Certification Authorities, which is considered a major deficiency for SSL. Certificate Transparency is an open framework that monitors and inspects SSL/TLS certificates and does not disrupt the existing SSL/TLS certificate validation system that browsers have been using. The system is not an alternative or a substitute to the existing validation system of browsers. Instead, it adds new functions to the validation system and expands the certification chain verification steps in order to provide support for inspection of all SSL/TLS certificates.

Google announced that certificates issued after April 30, 2018 must be compatible with CT. Before the announcement, Nykvist et al. [7] studied the server-side adoption of CT. In their work, they examined the compatibility of websites and characterized the overheads and the potential performance impact of the Signed Certificate Timestamp (SCT) delivery methods. Since there was no obligation before the announcement, it is important to now assess the current process. For this purpose, this current study analyzed the status of the top 500 websites and their certificates issuers [8]. In addition, the compliance of web browsers was also checked.

II. CERTIFICATE TRANSPARENCY COMPONENTS

Certificate Transparency focuses on the problems of the existing SSL system that are difficult to detect. These issues are briefly described as follows.

Malicious certification authorities and Internet sites can take steps to trick users such as issuing fraudulent SSL certificates by certification authorities including the domains of well-known Internet sites, and the deception of users using these certificates on Internet sites within a 'man in the middle' attack.

Even in the absence of malicious intent, it is possible for certification authorities to make a mistake when producing SSL certificates. Many mistakes have been made by certification authorities in the past. These mistakes may not be detected for weeks or even months, with users having been victimized as a result. Certificate Transparency is proposed as a solution to such problems, and has three main objectives:

1. To make it impossible for certificate authorities to issue SSL certificates for a domain without the domain owner's knowledge.
2. To support an open audit and monitoring system that allows domain owners or the certification authority to check whether or not certificates have been produced in error or through malicious intent.
3. It is intended to protect users from certificates produced in error or through malicious intent.

Certificate Transparency aims to achieve these objectives through three main components: certificate logs, monitors, and auditors.

A. Certificate Logs

The most important component of the Certificate Transparency Project is the certificate log servers. A certificate log server is a simple network service that holds and protects hash values of SSL certificates. Certificate log servers have three main features:

- A certificate can only be appended to the log server (append-only) and the certificate record cannot be deleted, modified or retrospectively added.
- In the certificate log servers, a special cryptographic mechanism known as the Merkle Hash Tree is used to prevent subsequent modifications to the records which are cryptographically protected.
- Certificate log servers can be audited explicitly; anyone can query a log server and verify that an SSL certificate has been properly added to the log server.

Certificates are logged to the log servers and maintained securely. Log servers return a Signed Certificate Timestamp which is proof of logging.

B. Monitors

Companies which have websites need to know if any certificates are issued for their websites. Taking into account all of these logs, it is possible to check for the issuance of certificates.

Monitors are servers that periodically connect to the log servers, continuously check for suspicious certificates, and work explicitly. The monitoring function is similar to the credit reporting service, which notifies individuals when a fake credit card is issued on their behalf.

Monitoring tools are progressively developing. Facebook developed a monitoring tool for users and users can check the certificate issuance of their domains [9].

C. Auditors

Auditors are software components that typically perform two functions. It can be used to check whether or not an SSL certificate to be authenticated is in the log server. Auditors can verify that SSL certificates have been correctly added to the log server and are cryptographically consistent.

If SSL certificates to be authenticated are not included in the log server, they are marked as suspicious and subsequently, the TLS client may refuse connection to sites with suspicious certificates.

III. CERTIFICATE TRANSPARENCY LOG AND CONTROL

Certificate Transparency can be achieved via three methods according to the logging and control architecture. These methods are described in the following sections.

A. X509V3 Extension Method

X.509 is a standard that defines the format of certificates [1]. SSL Certificates have numerous fields conforming to the Certificate Authority/Browser (CA/B) Baseline Requirements [10]. In the X509v3 Extension method, as shown in Fig. 1, firstly a pre-certificate is created by the CA. A pre-certificate has a "poison extension" and thereby cannot be used as an actual SSL Certificate. Secondly, the pre-certificate is logged to the log server and gathers a log response which is known as a Signed Certificate Timestamp (SCT). The SCT is placed as an extension to the certificate and then the certificate is signed. The SCT is shared in the process of the SSL/TLS handshake within the certificate.

B. TLS Extension Method

In the SSL/TLS Extension method, as shown in Fig. 2, the certificate is logged by the domain owner to the log servers and the SCT is serviced by the web server in the process of the SSL handshake. With this method, the website admin needs to log the certificate and deploy the SCT to the server.

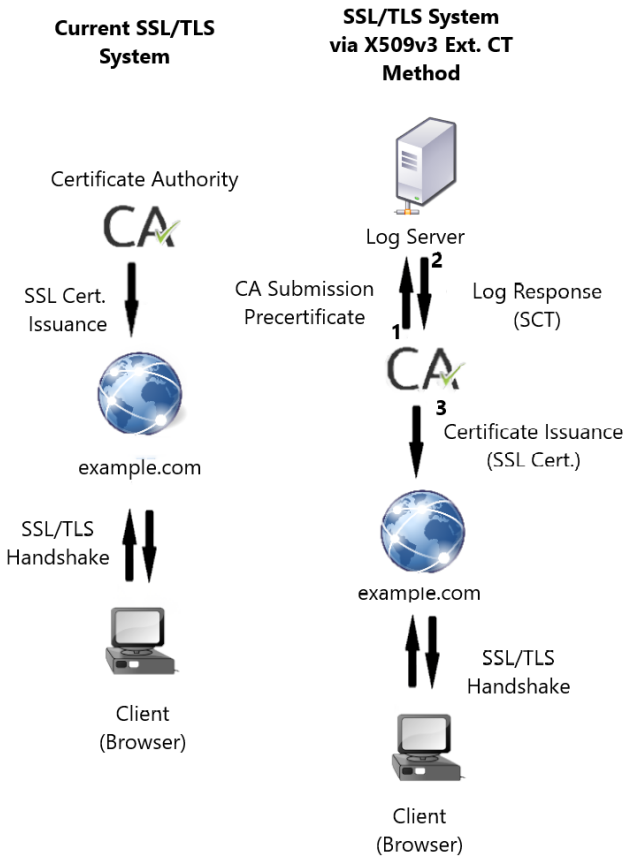


Fig. 1. X509v3 Extension Method

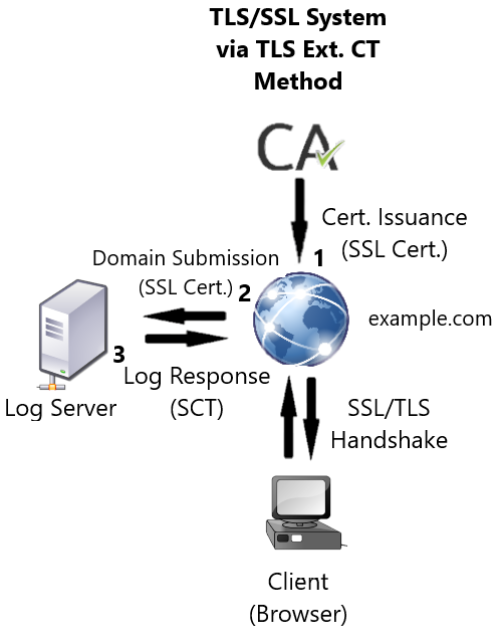


Fig. 2. SSL/TLS Extension Method

C. OCSP Stapling Method

Online Certificate Status Protocol (OCSP) is a protocol used for establishing the revocation status of a certificate [11]. In OCSP, a client sends an OCSP request to the OCSP server and the server creates and signs the OCSP response for the related request. OCSP stapling is a method for boosting the efficiency of the OCSP request and response process. In the OCSP stapling method, the server of the website sends a request for itself and gathers a response and serves this response to its clients. In the OCSP stapling method, as shown in Fig. 3, the certificate is logged to the log servers by the CA, and then the CA gathers the SCT from the log server and service inside of the OCSP response. Website servers obtain the OCSP response and serve it with its clients through OCSP Stapling.

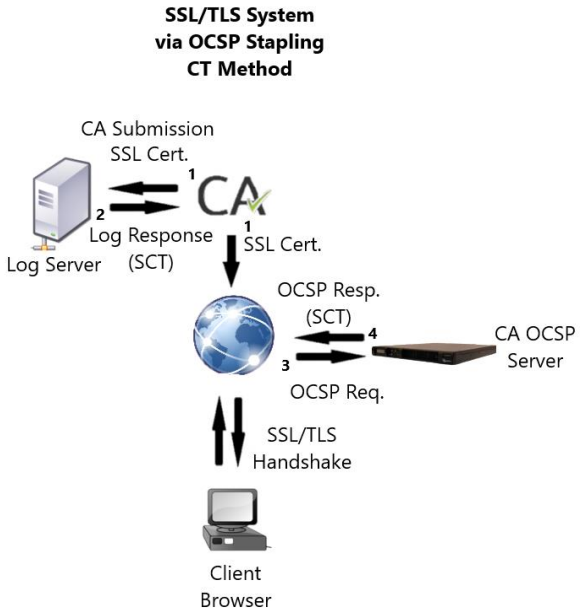


Fig. 3. OCSP Stapling Method

IV. The Process of Compliance With Certificate Transparency

The main components of Certificate Transparency (CT) in the light of Google's notifications are log servers and monitors, which can be operated by Google, certification authorities or third parties. Log servers are run by certification authorities such as DigiCert, WoSign, and StartCom. Auditors can also be run by browsers and clients who implement TLS.

Firstly, we examined browsers for their CT compliance as presented in Table I.

Google Chrome supports CT in versions released after January 2015.

Mozilla Firefox supports CT, and published its time schedule for CT on June 9, 2015 [12]. It is enabled within "about:config" security.pki.certificate_transparency.mode value=1 setting.

Safari announced that certificates issued after October 15, 2018, must meet their CT policy in order to be evaluated as trusted on Apple platforms [13]. However, the current version of Safari (v11) does not show any notification with regards to CT.

Yandex does not support CT. There is no information about CT on their website.

Internet Explorer does not support CT, but Microsoft developed a new extension to the Active Directory Certificate Services to support CT [14].

TABLE I. CT COMPLIANCE OF BROWSERS

	CT Compatibility of Browsers				
	Google Chrome v67.0	***Firefox v61.0.1***	***Safari v11***	***Yandex v18.6.1***	***Internet Explorer v11.165***
Compliance	✓	✓	✗	✗	✗

V. CERTIFICATE TRANSPARENCY CONFORMITY ASSESSMENT OF TOP 500 WEBSITES

Google announced that Chrome required all TLS server certificates issued after April 30, 2018 must be compliant with the Chromium CT Policy. After this date, when Chrome connects to a website serving a trusted certificate that is non-compliant to the Chromium CT Policy, Chrome will show a full-page warning that the connection is non-CT-compliant. CAs are strongly encouraged to work with their clients in order to ensure that their TLS certificates are compliant with the Chromium CT Policy through at least one of three methods mentioned in Section 3 before the end of March 2018 so as to ensure that any issues with deploying CT support are resolved in advance of the enforcement deadline. These changes were first rolled out to Desktop platforms, including macOS, Windows, Linux, and Chrome OS [15].

Experiments were conducted in this study in order to check the status of certificates for popular websites and CAs after Google's CT announcement. First, the names of the top 500 websites were obtained from the MoZ Top 500 on May 25, 2018. Then, a Certificate Transparency Control program was implemented in order to achieve the design needs. The program was developed on the Java platform using the Google Certificate Transparency API for handling SCT [16,17]. Where a certificate has a SCT extension, the browser can use it for checking. Windows OS shows a certificate transparency extension on their certificate viewer as shown in Fig. 4. OpenSSL is used for triple handshake packet capture in the SSL/TLS method.

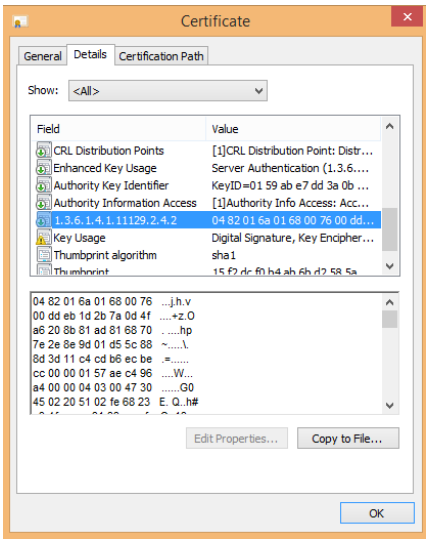


Fig. 4. Certificate with SCT Windows View

In our approach, we first fetch the certificate from the server and then parse the certificate in order to examine the SCT extension. If an SCT extension is found, other methods, OCSP Stapling and SSL/TLS handshake, are then checked. For the second method, OCSP stapling, a module was developed for checking the presence of the SCT in the OCSP responses. For the third method, TLS extension, OpenSSL is employed. The TLS responses are intervened and parsed in order to check the existence of SCT.

VI. RESULTS

The top 500 websites were analyzed. Details of the top 10 websites are given in Table II. It was determined that the dominant method (80%) used in the top 10 websites is the X509v3 extension method. The overall analysis of the top 500 websites is shown in Fig. 5. As shown in Fig. 5., while only half of the websites (54%) are CT-enabled, 16% of the websites do not even use SSL directly on their pages. As shown in Table III, the X509v3 extension method is widely used. Since the whole process could be achieved by only CAs without contribution from the domain owner with this method, it is therefore deemed easier to deploy. The OCSP stapling extension method was not used by the top 10 websites. The reason why it might not be the preferred method by CAs is that the OCSP stapling case domain owner has some responsibilities to perform. It can be hard to deal with domain-based problems during integration. The TLS extension method was found to be rarely used (15% of CT methods). We believe that the TLS extension method is only applied by domain owners who are CT-aware, and whose certificates do not include SCT. We found that 43 CA chains support CT methods and 17 CAs do not support any methods, as shown in Fig. 6. We found that some popular CAs are incompatible with CT. Microsoft and Yandex do not support CT as a CA (Microsoft IT TLS CA v5, Microsoft IT TLS CA v2, Yandex CA).

TABLE II. REPORT OF TOP 10 WEBSITES

	Top 10 Websites				
ID	***Site URL***	***X509v3 Ext. Method***	***OCSP S. Ext. Method***	***TLS Ext.***	***Stat***
1	https://facebook.com	✓	✗	✗	✓
2	https://twitter.com	✓	✗	✗	✓
3	https://google.com	✗	✗	✓	✓
4	https://youtube.com	✗	✗	✓	✓
5	https://instagram.com	✓	✗	✗	✓
6	https://linkedin.com	✓	✗	✗	✓
7	https://wordpress.org	✗	✗	✗	✗
8	https://pinterest.com	✓	✗	✗	✓
9	https://wikipedia.org	✓	✗	✗	✓
10	https://wordpress.com	✓	✗	✗	✓

These results show that Certificate Transparency is not implemented completely and that CAs commonly use the X509V3 method.

TABLE III. CT METHODS USAGE RATES

	CT Methods Usage Numbers		
ID	***X509v3 Ext. Method***	***OCSP Stapling Method***	***TLS Extension Method***
1	224	0	41

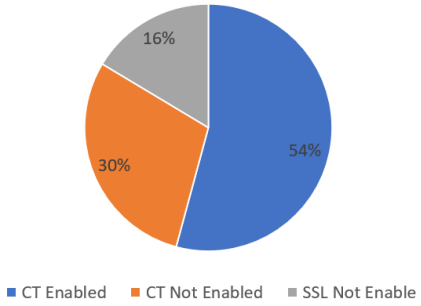


Fig. 5. CT Compliance of websites

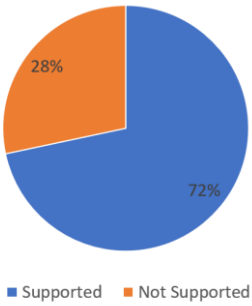


Fig. 6. CT Compliance of CAs

VII. CONCLUSION

In this study, the Certificate Transparency conformity assessment of top 500 websites and their certificate issuers were analyzed. It was observed that CT usage is not sufficiently widespread and that there are many CAs and websites which do not use CT or even SSL. The usage rates of CT methods were also explored. Although all three methods are deemed to be usable, user-friendly methods are preferred due to their ease of use for website admins. The OCSP stapling method is not used by the top websites. In this method, both the CA and the domain must work together for integration. It is believed that the CA chooses to implement the X509v3 extension method rather than the OCSP stapling method since logging and SCT deployment processes are challenging for website admins. This study is considered the first analysis of websites and browsers after Google's announcement on CT usage. There have already been some improvements seen on CT [18] and it is believed that studies in this area will increase in the near future. Hence, this current study makes a contribution to the literature by presenting the current status of websites and browsers.

VIII. REFERENCES

[1] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, W. Polk, "IETF, RFC5280 - Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", May 2008.

[2] D. Akhawe, B. Amann, M. Valentin, and R. Sommer, "Here's my cert, so trust me, maybe? Understanding TLS errors on the web", In Proceedings of the 22nd international conference on World Wide Web, pp. 59-70, May 2013.

[3] Z. Durumeric, J. Kasten, M. Bailey, J. A. Halderman, "Analysis of the HTTPS Certificate Ecosystem", IMC'13, pp. 23–25, October 2013,

[4] B. Laurie, A. Langley, E. Kasper, "IETF, RFC6962 – Certificate Transparency", June 2013.

[5] A. Johanna, G. Oliver, S. Quirin, B. Lexi, G. Carle, R. Holz, "Mission Accomplished? HTTPS Security after DigiNotar", Proceedings Of The 2017 Internet Measurement Conference IMC '17 pp. 325-340, November 2017.

[6] Mozilla Announcement for Digicert Revocation Process, https://blog.mozilla.org/security/2011/11/03/revoking-trust-in-digicert-sdn-bhd-intermediate-certificate-authority/".

[7] C. Nykvist, L. Sjöström, J. Gustafsson, N. Carlsson, "Server Side Adoption of Certificate Transparency", PAM 2018: Passive and Active Measurement, pp. 186-199, March 2018.

[8] Moz Top 500 Websites, https://moz.com/top500

[9] Facebook Certificate Transparency Monitor Tool, https://www.facebook.com/notes/protect-the-graph/introducing-our-certificate-transparency-monitoring-tool/1811919779048165/

[10] CA/Browser Forum. Baseline Requirements for the Issuance and Management of Publicly-Trusted Certificates, v.1.5.9, June. 2018. https://cabforum.org/wp-content/uploads/CA-Browser-Forum-BR-1.5.9.pdf

[11] M. Myers, R. Ankney, A. Malpani, S. Galperin, C. Adams, "IETF, RFC2560 - X.509 Internet Public Key Infrastructure Online Certificate Status Protocol – OCSP", June 1999.

[12] Mozilla Time Schedule about Certificate Transparency, https://wiki.mozilla.org/SecurityEngineering/Certificate_Transpareny

[13] Apple Announcement about CT compliance and obligations, https://support.apple.com/en-us/HT205280

[14] Microsoft Announcement about CT compliance for ACCS https://support.microsoft.com/en-us/help/4093260/introduction-of-ad-cs-certificate-transparency

[15] CT Certificate Transparency Enforcement in Google Chrome Announcement, https://groups.google.com/a/chromium.org/forum/#!topic/ct-policy/wHILiYf31DE

[16] Google's Certificate Transparency Code Archive and Wiki, https://code.google.com/p/certificate-transparency/

[17] Google's Certificate Transparency Java API and Sample Codes, https://github.com/google/certificate-transparency

[18] R. Ellgren, T. Löfgren, "Distributed Client Driven Certificate Transparency Log", Linköping University, Department of Computer and Information Science Bachelor thesis, 2018

ISC TURKEY 2018

11. ULUSLARARASI
BİLGİ GÜVENLİĞİ
ve **KRİPTOLOJİ**
KONFERANSI

11th INTERNATIONAL CONFERENCE
ON INFORMATION
SECURITY &
CRYPTOLOGY

17 - 18 Ekim - *October* 2018• ANKARA BTK MERKEZ BİNASI • *ICTA HEADQUARTER*

BİLGİGÜVENLİĞİ
DERNEĞİ

Maltepe Mahallesi Tuncer Sokak
No: 2/8 06570 Çankaya-ANKARA
0 (312) 231 18 10
bilgi@bilgiguvenligi.org.tr

ISBN: **978-605-86904-8-6**