CIC-IoMT-2024 Veri Seti Detaylı Analiz Raporu

Bu rapor, CIC-IoMT-2024 (Canadian Institute for Cybersecurity - Internet of Medical Things 2024) veri setinin kapsamlı analizini sunmaktadır. Veri seti, tıbbi cihazların siber güvenlik açıklarını ve saldırı türlerini incelemek için tasarlanmıştır.

# 📋 İçindekiler

1. 1. Genel Bakış
2. 2. Veri Seti Yapısı
3. 3. WiFi ve MQTT Verileri
4. 4. Bluetooth Verileri
5. 5. Özellik Mühendisliği
6. 6. Saldırı Kategorileri
7. 7. Veri İstatistikleri
8. 8. Öneriler ve Sonuçlar

# 1. 🌐 Genel Bakış

CIC-IoMT-2024 veri seti, Internet of Medical Things (IoMT) cihazlarının siber güvenlik açıklarını analiz etmek için geliştirilmiş kapsamlı bir veri setidir. Veri seti iki ana kategoriye ayrılmıştır:

* • WiFi ve MQTT Protokolleri: Akıllı tıbbi cihazların ağ trafiği
* • Bluetooth Protokolleri: Kablosuz tıbbi cihazların iletişimi

# 2. 📁 Veri Seti Yapısı

Veri seti aşağıdaki hiyerarşik yapıya sahiptir:

CICIoMT2024/  
├── WiFi\_and\_MQTT/  
│ ├── attacks/  
│ │ ├── csv/  
│ │ │ ├── train/ (51 CSV dosyası)  
│ │ │ └── test/ (21 CSV dosyası)  
│ │ └── pcap/  
│ │ ├── train/ (51 PCAP dosyası)  
│ │ └── test/ (21 PCAP dosyası)  
│ └── profiling/  
│ ├── CSV/ (37 profil CSV dosyası)  
│ └── PCAP/ (37 profil PCAP dosyası)  
└── Bluetooth/  
 ├── attacks/  
 │ ├── csv/ (CSV dosyaları mevcut değil)  
 │ └── pcap/  
 │ ├── train/ (2 PCAP dosyası)  
 │ └── test/ (2 PCAP dosyası)  
 └── profiling/  
 └── pcap/ (10 cihaz profil PCAP dosyası)

# 3. 📶 WiFi ve MQTT Verileri

## 3.1 Saldırı Verileri

WiFi ve MQTT saldırı verileri, farklı türde siber saldırıları simüle eden ağ trafiği kayıtlarını içerir. Toplam 72 dosya bulunmaktadır:

|  |  |  |
| --- | --- | --- |
| **Saldırı Kategorisi** | **Açıklama** | **Dosya Sayısı** |
| Benign | Normal, zararsız trafik | 2 (train+test) |
| MQTT Saldırıları | MQTT protokolüne yönelik saldırılar | 12 |
| TCP/IP DoS/DDoS | Hizmet reddi saldırıları | 32 |
| Keşif Saldırıları | Port tarama, ping sweep vb. | 8 |
| ARP Spoofing | ARP protokolü manipülasyonu | 2 |
| Malformed Data | Bozuk veri paketleri | 2 |
| Diğer | Çeşitli saldırı türleri | 14 |

## 3.2 Profil Verileri

Profil verileri, gerçek tıbbi cihazların normal çalışma durumlarını kaydeder. Bu veriler, normal trafik desenlerini anlamak için kullanılır.

|  |  |  |
| --- | --- | --- |
| **Cihaz Kategorisi** | **Cihazlar** | **Açıklama** |
| Kamera Cihazları | Blink, Ecobee, M1T, Owltron | Güvenlik kameraları |
| Sağlık Monitörleri | Checkme, COOSPO, Lookee | Kalp ritmi, oksijen monitörleri |
| Bebek Monitörleri | SenseU, SenseUBaby | Bebek takip cihazları |
| Çağrı Sistemleri | Multifunctional Pager, Singcall | Hastane çağrı sistemleri |
| Diğer Cihazlar | Rhythm+, SleepU, Wellue | Çeşitli tıbbi cihazlar |
| Ağ Durumları | Active, Idle, Broker | Ağ altyapısı durumları |
| Etkileşim Türleri | LAN, WAN, Power | Farklı bağlantı türleri |
| Fonksiyonlar | MIC, WATCH, PHOTO, RECORDING | Cihaz fonksiyonları |

# 4. 📱 Bluetooth Verileri

Bluetooth verileri, kablosuz tıbbi cihazların Bluetooth protokolü üzerinden yaptığı iletişimi kaydeder. Bu veriler daha sınırlıdır:

* • Saldırı Verileri: Sadece 2 dosya (Benign ve DoS)
* • Profil Verileri: 10 farklı tıbbi cihaz
* • Veri Formatı: Sadece PCAP formatında
* • CSV Dönüşümü: Mevcut değil

# 5. 🔧 Özellik Mühendisliği

CSV dosyaları 45 farklı özellik içerir. Bu özellikler, ağ trafiği analizi için mühendislik uygulanmış metriklerdir:

## 5.1 Temel Ağ Özellikleri

* • Header\_Length: Paket başlık uzunluğu
* • Protocol Type: Protokol türü (sayısal kod)
* • Duration: Bağlantı süresi
* • Rate: Veri hızı
* • Srate: Kaynak hızı
* • Drate: Hedef hızı

## 5.2 TCP Bayrakları

* • fin\_flag\_number: FIN bayrak sayısı
* • syn\_flag\_number: SYN bayrak sayısı
* • rst\_flag\_number: RST bayrak sayısı
* • psh\_flag\_number: PSH bayrak sayısı
* • ack\_flag\_number: ACK bayrak sayısı
* • ece\_flag\_number: ECE bayrak sayısı
* • cwr\_flag\_number: CWR bayrak sayısı

## 5.3 Protokol Sayıları

* • HTTP, HTTPS: Web protokolleri
* • DNS: Alan adı sistemi
* • Telnet, SMTP, SSH, IRC: Uygulama protokolleri
* • TCP, UDP: Taşıma katmanı protokolleri
* • DHCP, ARP, ICMP, IGMP: Ağ katmanı protokolleri
* • IPv, LLC: Düşük seviye protokolleri

## 5.4 İstatistiksel Özellikler

* • Tot sum: Toplam değer
* • Min, Max, AVG: Minimum, maksimum, ortalama
* • Std: Standart sapma
* • Tot size: Toplam boyut
* • IAT: Inter-Arrival Time
* • Number: Sayısal değer
* • Magnitude, Radius: Geometrik özellikler
* • Covariance, Variance: İstatistiksel özellikler
* • Weight: Ağırlık değeri

# 6. ⚔️ Saldırı Kategorileri Detayı

## 6.1 MQTT Saldırıları

* • MQTT-DDoS-Connect\_Flood: Dağıtık hizmet reddi - Connect flood saldırısı
* • MQTT-DDoS-Publish\_Flood: Dağıtık hizmet reddi - Publish flood saldırısı
* • MQTT-DoS-Connect\_Flood: Hizmet reddi - Connect flood saldırısı
* • MQTT-DoS-Publish\_Flood: Hizmet reddi - Publish flood saldırısı
* • MQTT-Malformed\_Data: Bozuk MQTT veri paketleri

## 6.2 TCP/IP Saldırıları

* • TCP\_IP-DoS/DDoS-SYN: SYN flood saldırıları (tek/dağıtık)
* • TCP\_IP-DoS/DDoS-ICMP: ICMP flood saldırıları (tek/dağıtık)
* • TCP\_IP-DoS/DDoS-TCP: TCP flood saldırıları (tek/dağıtık)
* • TCP\_IP-DoS/DDoS-UDP: UDP flood saldırıları (tek/dağıtık)

## 6.3 Keşif Saldırıları

* • Recon-Port\_Scan: Port tarama saldırıları
* • Recon-Ping\_Sweep: Ping sweep saldırıları
* • Recon-OS\_Scan: İşletim sistemi tarama
* • Recon-VulScan: Güvenlik açığı tarama

# 7. 📊 Veri İstatistikleri

## 7.1 Dosya Boyutları

|  |  |  |
| --- | --- | --- |
| **Veri Türü** | **Dosya Sayısı** | **Toplam Boyut (Yaklaşık)** |
| WiFi/MQTT Train CSV | 51 | ~15 GB |
| WiFi/MQTT Test CSV | 21 | ~5 GB |
| WiFi/MQTT Train PCAP | 51 | ~8 GB |
| WiFi/MQTT Test PCAP | 21 | ~2 GB |
| Bluetooth PCAP | 4 | ~52 MB |

## 7.2 Örnek Sayıları

Bazı önemli dosyaların örnek sayıları:

* Benign\_train.pcap.csv: 192,732 örnek
* MQTT-DDoS-Connect\_Flood\_train.pcap.csv: 173,037 örnek
* MQTT-DDoS-Publish\_Flood\_train.pcap.csv: 27,624 örnek
* Recon-Port\_Scan\_train.pcap.csv: 83,982 örnek
* ARP\_Spoofing\_train.pcap.csv: 16,048 örnek
* Recon-Ping\_Sweep\_train.pcap.csv: 741 örnek

## 7.3 Veri Kalitesi

* ✅ Tüm özellikler float64 formatında
* ✅ Eksik değer yok (NaN değerleri temizlenmiş)
* ✅ Standartlaştırılmış özellik adları
* ✅ Tutarlı veri formatı tüm dosyalarda
* ✅ Zaman serisi verisi (IAT özelliği ile)

# 8. 🚀 Öneriler ve Sonuçlar

## 8.1 Veri Seti Güçlü Yanları

* ✅ Kapsamlı saldırı türü çeşitliliği
* ✅ Gerçek tıbbi cihaz profil verileri
* ✅ Hem normal hem anormal trafik
* ✅ Çok boyutlu özellik seti (45 özellik)
* ✅ Train/test ayrımı mevcut
* ✅ Hem CSV hem PCAP formatları

## 8.2 Potansiyel İyileştirmeler

* 🔧 Bluetooth verileri için CSV formatı eklenebilir
* 🔧 Daha fazla tıbbi cihaz türü eklenebilir
* 🔧 Gerçek zamanlı veri akışı simülasyonu
* 🔧 Daha detaylı etiketleme (alt kategoriler)
* 🔧 Veri seti dengesizliği giderilebilir

## 8.3 Kullanım Önerileri

* 💡 Makine öğrenmesi için: Özellik seçimi ve normalizasyon gerekli
* 💡 Derin öğrenme için: Zaman serisi analizi uygun
* 💡 Güvenlik analizi için: Anomali tespiti algoritmaları
* 💡 Araştırma için: Kapsamlı karşılaştırmalı analiz
* 💡 Üretim için: Gerçek zamanlı izleme sistemleri

## 8.4 Sonuç

CIC-IoMT-2024 veri seti, tıbbi cihazların siber güvenlik analizi için kapsamlı ve değerli bir kaynak sunmaktadır. 45 özellik ve çeşitli saldırı türleri ile makine öğrenmesi ve derin öğrenme araştırmaları için uygundur. Veri seti, IoMT güvenliği alanında önemli araştırma fırsatları sunmaktadır.

Bu rapor, CIC-IoMT-2024 veri setinin kapsamlı analizini sunmaktadır. Veri seti, Canadian Institute for Cybersecurity tarafından geliştirilmiştir.