

Securing Privacy in MQTT-Based Systems

—

By: Hussain Zainal

Project Overview

Objective:

- Design/Analyze a secure MQTT-based communication system in the context of disaster relief

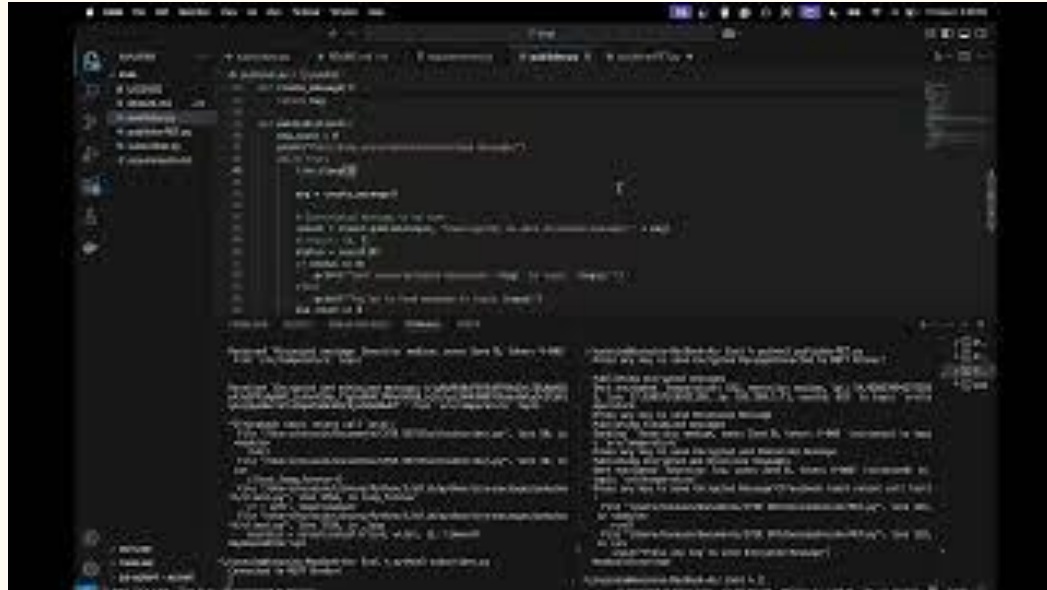
Challenge:

- Plain MQTT messages expose personal data (GPS, IPS, etc) leading to major privacy threats

Solutions:

- Apply Privacy-Enhancing Techniques (PETs) to protect users

Current Configuration



LINDDUN Privacy Threat Analysis 1

- Linkability

- Threat: An attacker can link multiple MQTT messages to the same user
- Justification: Messages share the

- Identifiability

- Threat: An attacker can identify someone based on the message content
- Justification: Messages could contain unique tokens or even account names

- Non-repudiation

- Threat: Users cannot deny sending a message
- MQTT logs do not contain any signatures/hashes

LINDDUN Privacy Threat Analysis 2

- Detectability

- Threat: An attacker can detect that someone is communicating, even without knowing what messages are being sent.
- Justification: MQTT topics are visible unless configured differently.

- Disclosure of Information

- Threat: Sensitive content can be exposed during MQTT transmissions
- Justification: MQTT messages are not TLS encrypted by default

LINDDUN Privacy Threat Analysis 3

- Unawareness

- Threat: Victims may not know what is being collected or who receives it. Might be automatic collection
- Justification: No interface is present to warn/confirm what is being shared

- Non-compliance

- Threat: System may be violating privacy laws
- Justification: Due to it being un-encrypted by default, the engineers are not complying to privacy policies

PETs to be implemented

- Encrypted Payloads
 - How it's implemented
 - Uses fernet symmetric encryption (from the cryptography library)
 - Publisher encrypts the payload before sending
 - Subscriber decrypts using the same secret key
 - Why
 - Prevents Disclosure of Information
- Pseudonymization
 - How it's implemented
 - Victim IDs become “VIC-XXX” instead of username.
 - GPS coordinates are assigned to zones
 - Why it matters
 - Prevents identifiability and Linkability

PET configurations



Improved Config with PETs

Before:

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
hussain@Hussains-MacBook-Air Eval % python3 subscriber.py
Connected to MQTT Broker!

Received `Unencrypted, no data minimized message: temperature: 196, severity: medium, lat: 57.44075521168091, lon: 94.23873945896395, ip: 192.168.1.125, userId: 991` from `srv/temperature` topic

Received `Unencrypted, no data minimized message: temperature: 420, severity: low, lat: 55.483430544510924, lon: 42.70805121565104, ip: 192.168.1.77, userId: 899` from `srv/temperature` topic

hussain@Hussains-MacBook-Air Eval % python3 publisher.py
Publishing unencrypted/nonminimized messages
Connected to MQTT Broker!
Sent unencrypted/no minimized `temperature: 196, severity: medium, lat: 57.44075521168091, lon: 94.23873945896395, ip: 192.168.1.125, userId: 991` to topic `srv/temperature`
Sent unencrypted/no minimized `temperature: 420, severity: low, lat: 55.483430544510924, lon: 42.70805121565104, ip: 192.168.1.77, userId: 899` to topic `srv/temperature`
```

After:

```
AWbzRbYGDddRw3bT4rllLcPqtb41qw0vkBjNZev57wRWJ6D2rPqe9SwAYX0_xqh6UqGQruqY8P
b2-e6ztmFV5yL2wzdK7RrPxQM18TPuST26d1ll5-fXAbRx-PkEcBPXiF7YM0tgJP9GcyFNcq0C
PoK8B5YH4q4isXEF7rhuSBRFJv6eTOPwpV2AE9S89VX1mSiChAZS4gyk=` from `srv/temperature` topic

Received `Minimized message: Severity: high, zone: Zone A, token: V-001` from `srv/temperature` topic

Received `Encrypted and minimized message: b'gAAAABoFR-u3_Zy38ESZiv0-UUVIwFLdGlbTvWFBlozFkcftgECK0cJrrPIy6g1CNZZ0xJhY-hgFx_pak2h80JwbaeLuQ4i-TTzHcaYjyPISmYSHDlw10NUqZOUAKVji3yeFF6h9xfo` from `srv/temperature` topic

Press any key to send Encrypted Message
Connected to MQTT Broker!

Publishing encrypted messages
Sent encrypted `temperature: 256, severity: low, lat: -41.85315880913369, lon: -118.80014219493853, ip: 192.168.1.247, userId: 970` to topic `srv/temperature`
Press any key to send Minimized Message
Publishing minimized messages
Sending `Severity: high, zone: Zone A, token: V-001` (minimized) to topic `srv/temperature`
Press any key to send Encrypted and Minimized Message
Publishing encrypted and minimized messages
Sent encrypted `Severity: low, zone: Zone C, token: V-002` (minimized) to topic `srv/temperature`
Press any key to send Encrypted Message
```

Privacy Metric

- Before PETs: 5 fields exposed (Victim ID, GPS, IP, Temperature, Severity)
 - After PETs: 1 field exposed
 - Other fields got tokenized
 - 80% reduction in exposed sensitive data
-
- To access messages without encryption:
 - Requires 3.4×10^{38} possible keys
 - Brute force requirements: 1.7×10^{14} years (too long!)

Effectiveness, Overhead, Trade offs

- Effectiveness:
 - Tokenization hides real identities and exact locations
 - Encryption protects the entire message payload from eavesdroppers
- Overhead:
 - Increased message size due to encryption
 - Have to convert zones into locations
 - Key maintenance
- Trade offs:
 - More secure but harder debugging for encryption
 - Increased processing power requirement for encrypting/decrypting
 - Rotation of keys

Privacy Impact on System Behaviour

- Encrypted messages are no longer human readable which hinders eavesdropping
- Randomized victim tokens prevent cross-session tracking
- System design uses pseudonyms
- Business objectives not affected
- Aligns system with privacy regulations