Security of
Networks, Services, and Systems

Privacy

Ricardo Morla

FEUP – SSR/M.EEC, SR/M.EIC

Data privacy

- Ability of individuals to control their personal information
- What is personal information?
  - Information that could identify an individual
  - Information related to an identifiable person (more general)
- How can someone lose control of their personal information?

Threat models

1. Theaft
   - innapropriate security measures
2. Willful disclosure to service provider
   - in exchange for using the service
   - because it is required for the service
   - because it is the purpose of a (data computing) service
3. Eavesdropping
   - of data sent over the network
   - with some (or no) expectations of confidentiality

GDPR and legal obligations (wrt someone else’s data)

- If you’re keeping someone else’s data:
    - you should keep it from third parties
        - by not sharing it with 3rd parties
        - by doing reasonable efforts to prevent attackers from obtaining the data
    - you should only keep it only while the owner agrees
        - after which you should delete the data
- If you’re using someone else’s data:
    - you should only use it for the purposes which the data owner agreed
    - you should make it clear what you intend to use the data for, and ask for the permission of the owner

Beyond legal … what options are there?

- Don’t give away your data!
  - Limit the data you share with the service provider
  - Use an alternative service that doesn’t ask you for that data
  - Don’t use the service
- What if need to:
  1. Process my data in the provider?
  2. Send my data over some channel I don’t trust?

* Process my data in the provider? Secure Multiparty Computation

- Client encrypts data, sends it to provider
  - doesn’t send secret to provider
- Provider applies *special* computations on encrypted data
  - Depends on algorithm an on operation
  - Doesn’t require secret to compute operation
  - Results are naturally encrypted – no need for key to encrypt
- Example: compute product, textbook RSA
  - Client sends encrypted values of M_1, M_2 to provider as $C_1 = M_1^d; C_2 = M_2^d$
  - Provider computes the product $C_P = C_1 C_2 = M_1^d M_2^d$ – no need for secret *e*
  - Provider sends C_P back to client
  - Client decripts C_P with secret e : $\mathrm{M_P} = C_P^e = \left(M_1^d M_2^d\right)^e = (M_1 M_2)^{de} = M_1 M_2$
- Really easy example, other operations will be harder (sum? generic?)

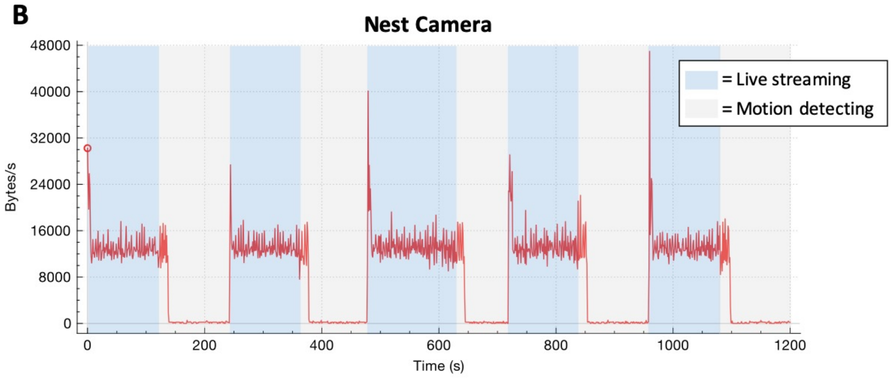
* Send my data over some channel I don’t trust? Encrypt data – and what else?

- You should use a secure mechanism that prevents an eavesdropper from accessing the data you send
  - Like a stream cipher
- But does this mean the attacker won’t be able to get any personal information from you?
- Depends
  - Think information theory – what information can you get from a constant signal?
  - Does on-off provide any personal information?
  - On – camera streaming ;; off – camera not streaming
  - Relevant if the attacker knows that you only turn the camera on when you’re not at home

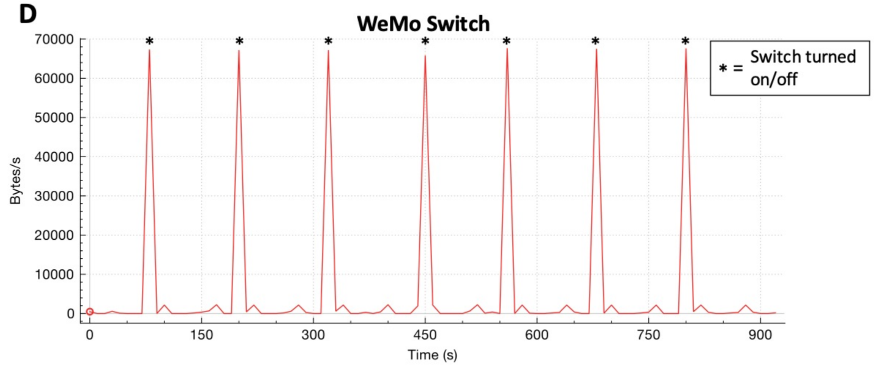
Encrypted Traffic

Smart home traffic analysis

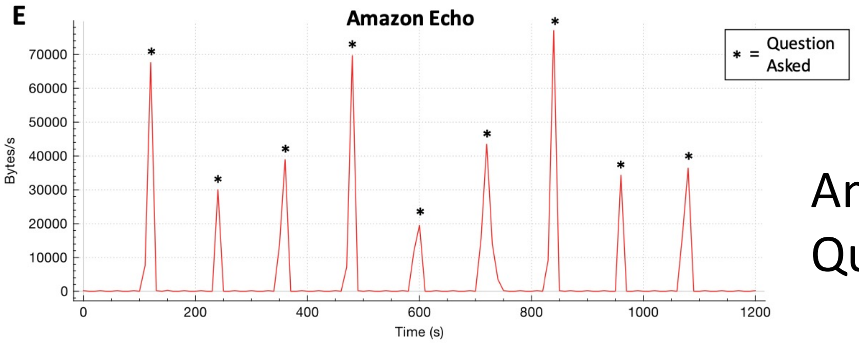
Apthorpe, Noah, Dillon Reisman, and Nick Feamster. "A smart home is no castle: Privacy vulnerabilities of encrypted IoT traffic." *arXiv preprint arXiv:1705.06805* (2017).



Motion detected



Switch on/off



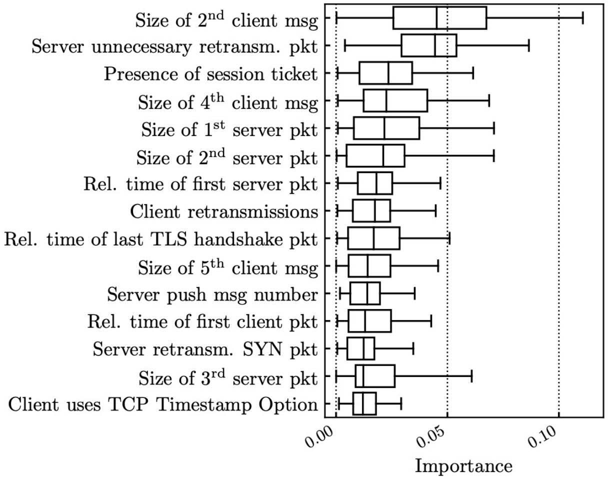
Amazon Echo Question asked

Domain name attacks to privacy

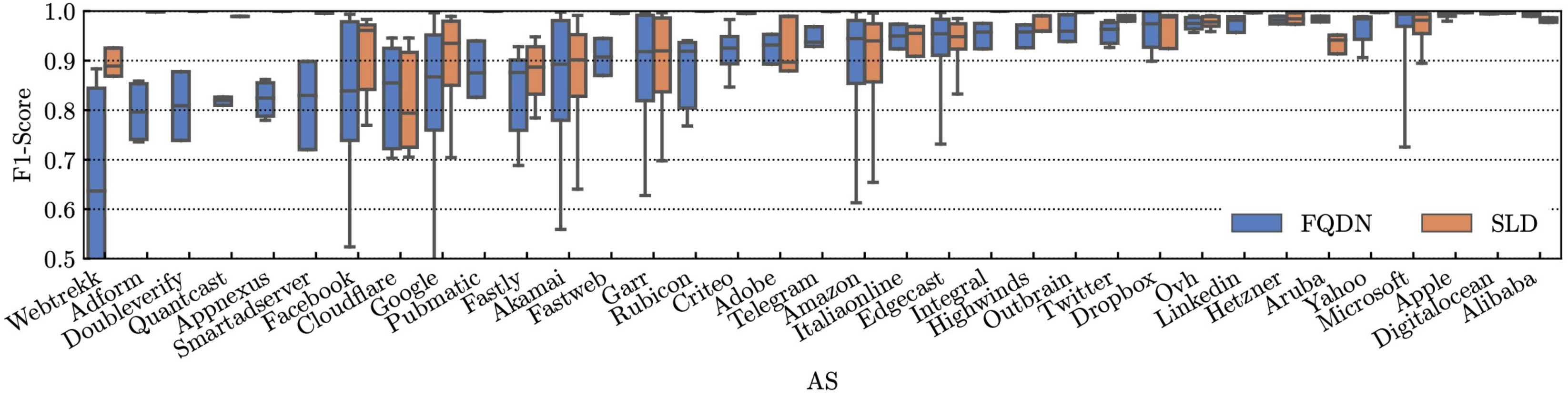
- Domain names are personal information for someone that is browsing the web
  - Profiling of services, political, sexual, other preferences
- For each web site that is visited we can get domain names from:
  - DNS request
  - HTTP header with servername
  - Server name indication in TLS
- Although the traffic is encrypted (e.g. with TLS), domain names are sent in clear. Solutions:
  - TLS hides HTTP servername header
  - Encrypting the DNS transport with DoH, DoT, etc.
  - Encrypting the Server Name Indicator in TLS (eSNI, draft-ietf-tls-esni)

Traffic analysis to determine domain names

Trevisan, Martino, et al. "Does domain name encryption increase users' privacy?." *ACM SIGCOMM Computer Communication Review* 50.3 (2020): 16-22.



- Infer domain names from traffic patterns
- 80% flows with F1 score larger than 0.8
- Size and timing of first packets important



Security of Networks, Services, and Systems

Privacy

Ricardo Morla

FEUP – SSR/M.EEC, SR/M.EIC