Security in the Internet of Things

Information Theoretic Insights

Vince Poor

(poor@princeton.edu)

Support by NSF CNS-1702808: "Secure Inference in the Internet of Things"

Disneyland
Hotel:
August 29, 1988

Importance of the IoT

The Internet of Things (IoT) makes possible Smart-X where

$$X \in \left\{ \begin{array}{c} city, factory, grid, \\ building, home, transportation, \\ healthcare, agriculture, metering \end{array} \right\}$$

IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*

IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*

WIRED

AUGUST 10, 2018 | LOUISE MATSAKIS

Hackable Touchscreens Could Spy on Hotel Rooms and Meetings

The technology company Crestron makes touchscreen panels and other equipment for places like conference rooms, which a researcher found can be turned into hidd...

IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*

AUGUST 10, 2018 | LOUISE MATSAKIS

Hackable Touchscreens Could Spy on Hotel Rooms and Meetings

The technology company Crestron makes touchscreen panels and other equipment for places like conference rooms, which a researcher found can be turned into hidd...

DECEMBER 20, 2017 | BRIAN BARRETT

Don't Get Your Kid an Internet-Connected Toy

They can be hacked. They're a privacy nightmare. This year, it's not too late to keep the IoT toys away from the tree.

IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*

AUGUST 10, 2018 | LOUISE MATSAKIS

Hackable Touchscreens Could Spy on Hotel Rooms and Meetings

The technology company Crestron makes touchscreen panels and other equipment for places like conference rooms, which a researcher found can be turned into hidd...

DECEMBER 20, 2017 | BRIAN BARRETT

Don't Get Your Kid an Internet-Connected Toy

They can be hacked. They're a privacy nightmare. This year, it's not too late to keep the IoT toys away from the tree.

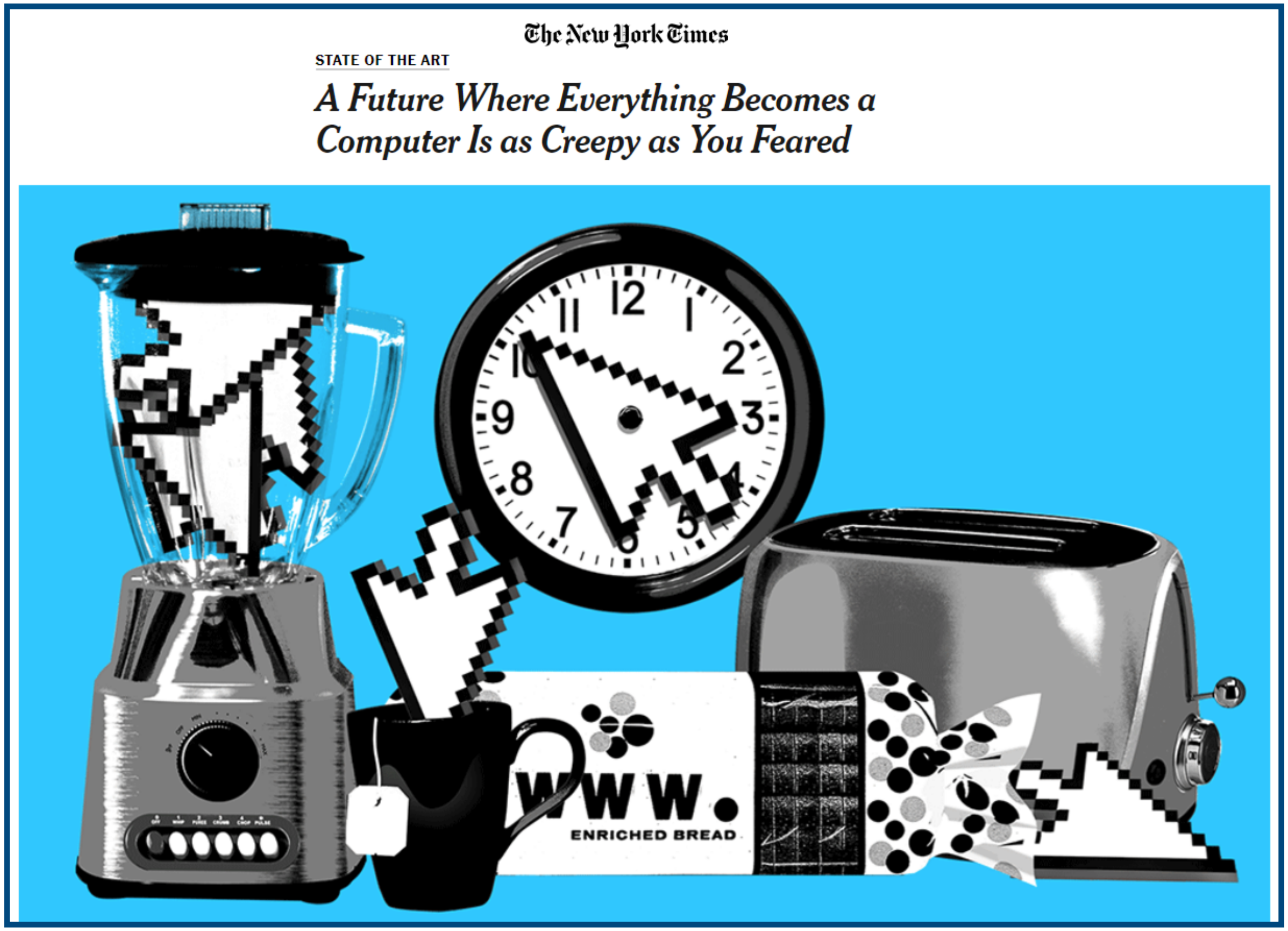
MARCH 2, 2017 | LILY HAY NEWMAN

Medical Devices Are the Next Security Nightmare

More internet-connected medical devices flood into healthcare industry every day, but we're not moving fast enough to defend them.

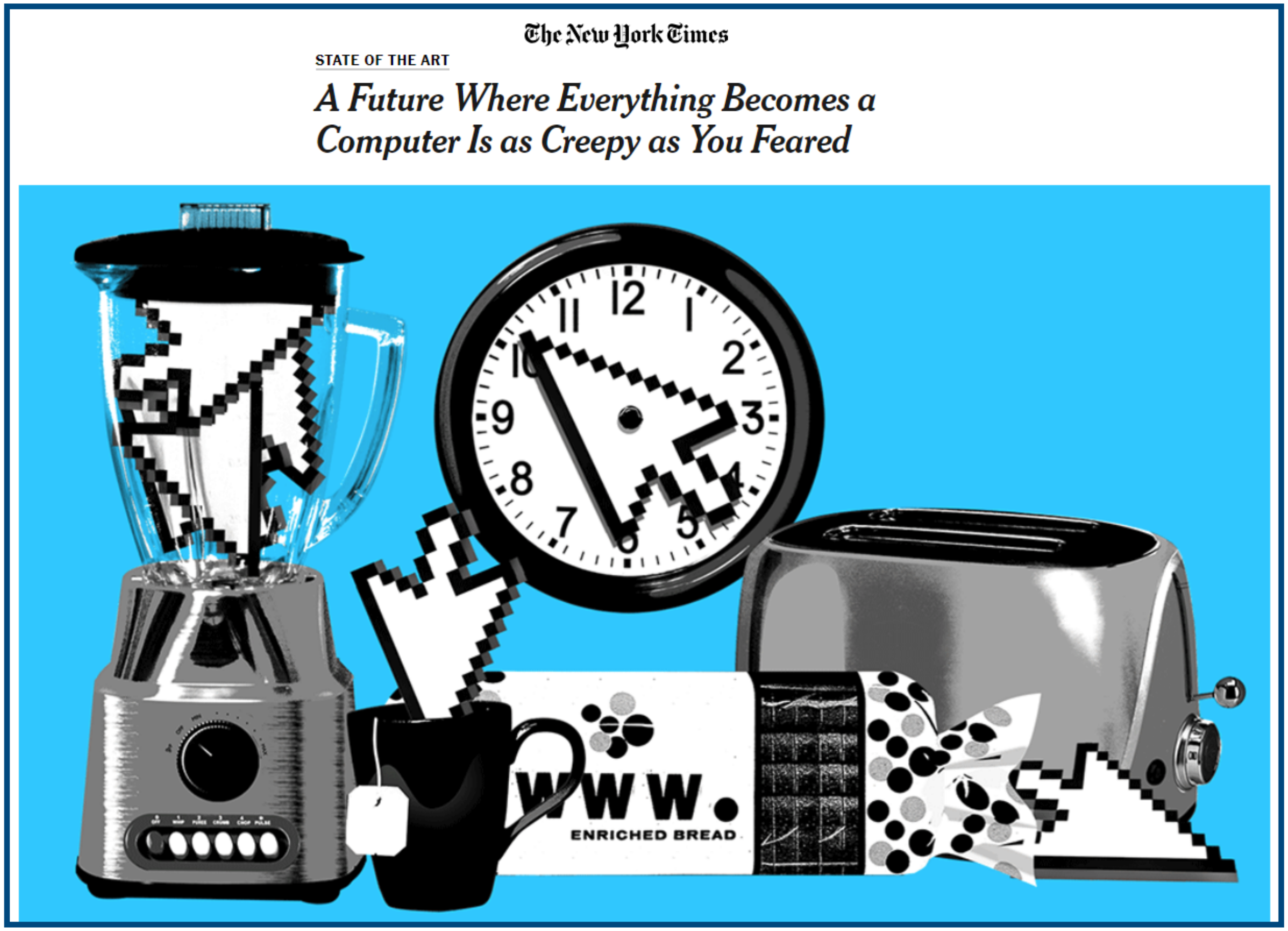
IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*



IoT Security – A Major Concern

- IoT vulnerabilities to cyber attacks → Mostly concern *personal privacy and security*

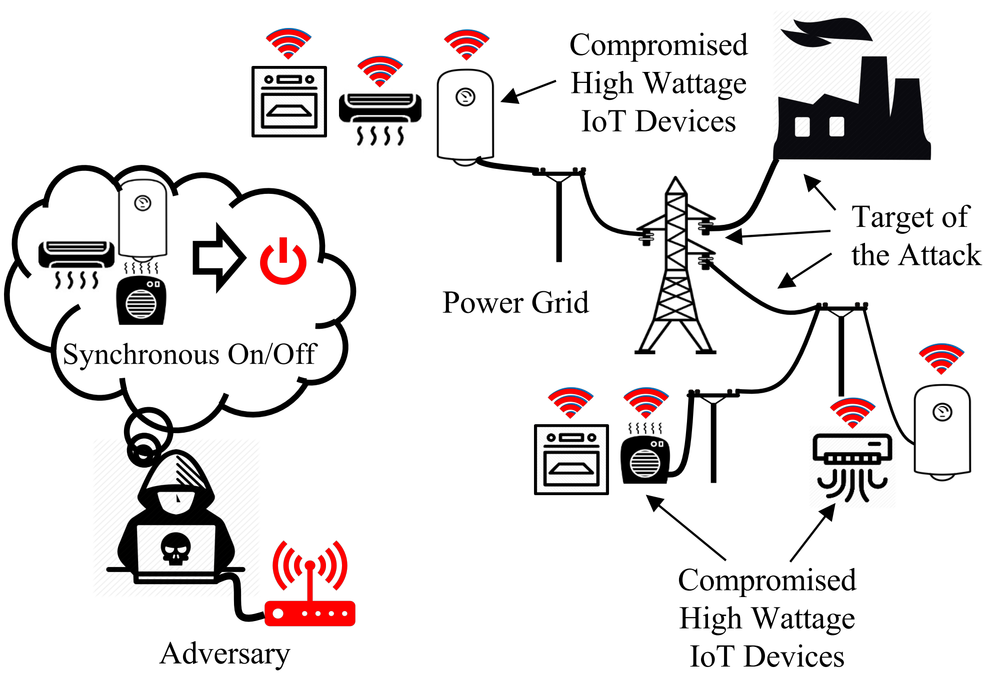


- "IoT Security: Let's forget all the lessons from traditional network security …," James Mickens

An Example of What Can Go Wrong

[Soltan, et al. USENIX'18]

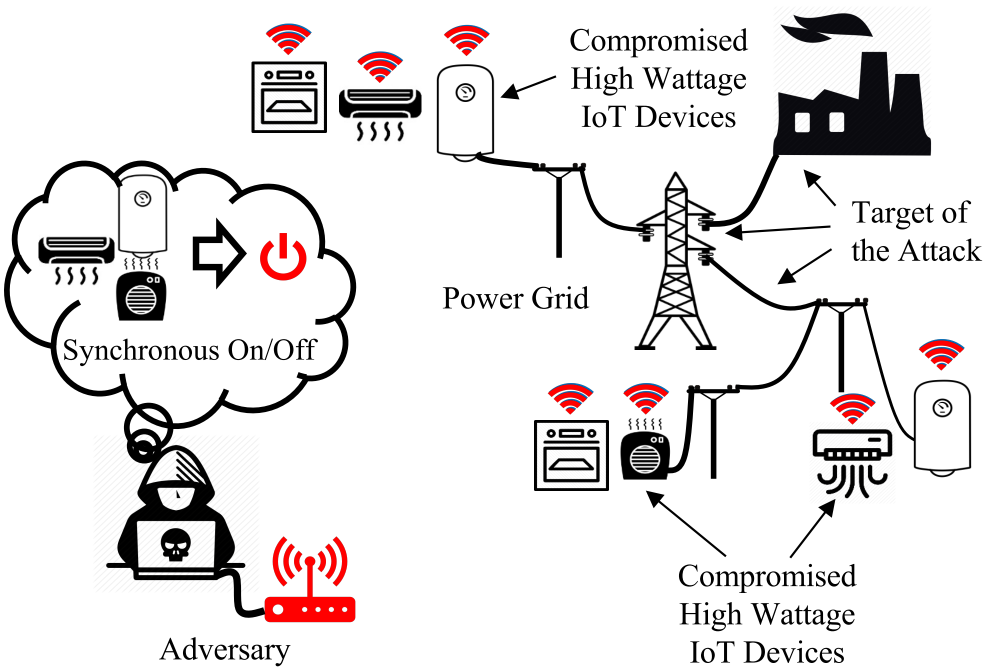
- Manipulation of demand via IoT: Botnets controlling high-wattage IoT devices (air conditioners, refrigerators, etc.) can disrupt the power grid.



An Example of What Can Go Wrong

[Soltan, et al. USENIX'18]

- Manipulation of demand via IoT: Botnets controlling high-wattage IoT devices (air conditioners, refrigerators, etc.) can disrupt the power grid.
- A Mirai-sized (600,000 bots) botnet of water heaters can change the demand instantly by 3GW – similar to having access to the largest currently deployed nuclear plant!



IoT - Characteristics

- Some salient characteristics:
  - Very large numbers of (possibly) low-complexity terminals
  - Low-latency, short-packet communications (e.g., for automation)
  - Possibly light or no infrastructure (e.g., ad hoc networking)
  - Used primarily for data gathering, inference & control

IoT - Characteristics

- Some salient characteristics:
  - Very large numbers of (possibly) low-complexity terminals
  - Low-latency, short-packet communications (e.g., for automation)
  - Possibly light or no infrastructure (e.g., ad hoc networking)
  - Used primarily for data gathering, inference & control
- These characteristics shape the issues of security and privacy, and introduce new regimes to consider for these issues

Overview of Today's Talk

The theme:

- A role for information theory in this area

Overview of Today's Talk

The theme:

- A role for information theory in this area

Begin with two main topics motivated by the characteristics of IoT :

- Security in wireless data transmission: physical layer security
- Privacy in sensing systems: privacy-utility tradeoffs

Overview of Today's Talk

The theme:

- A role for information theory in this area

Begin with two main topics motivated by the characteristics of IoT :

- Security in wireless data transmission: physical layer security
- Privacy in sensing systems: privacy-utility tradeoffs

Other issues – some new, some older (briefly):

- Authentication, security in MANETs, data injection attacks on electricity grids, attacks on sensor networks

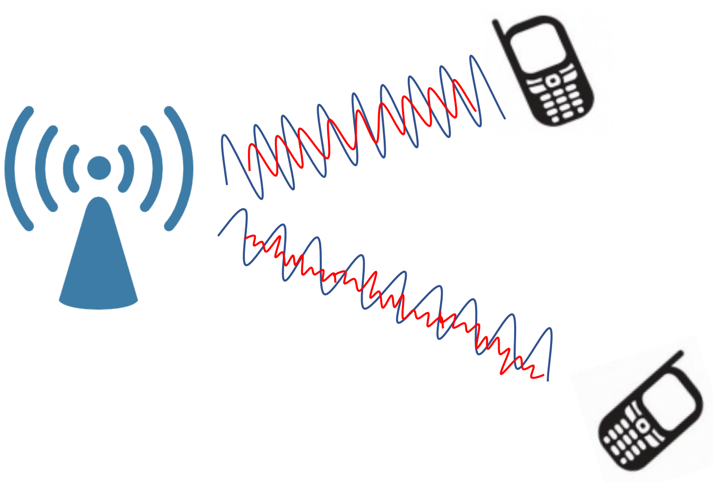
Physical Layer Security

in

Wireless Networks

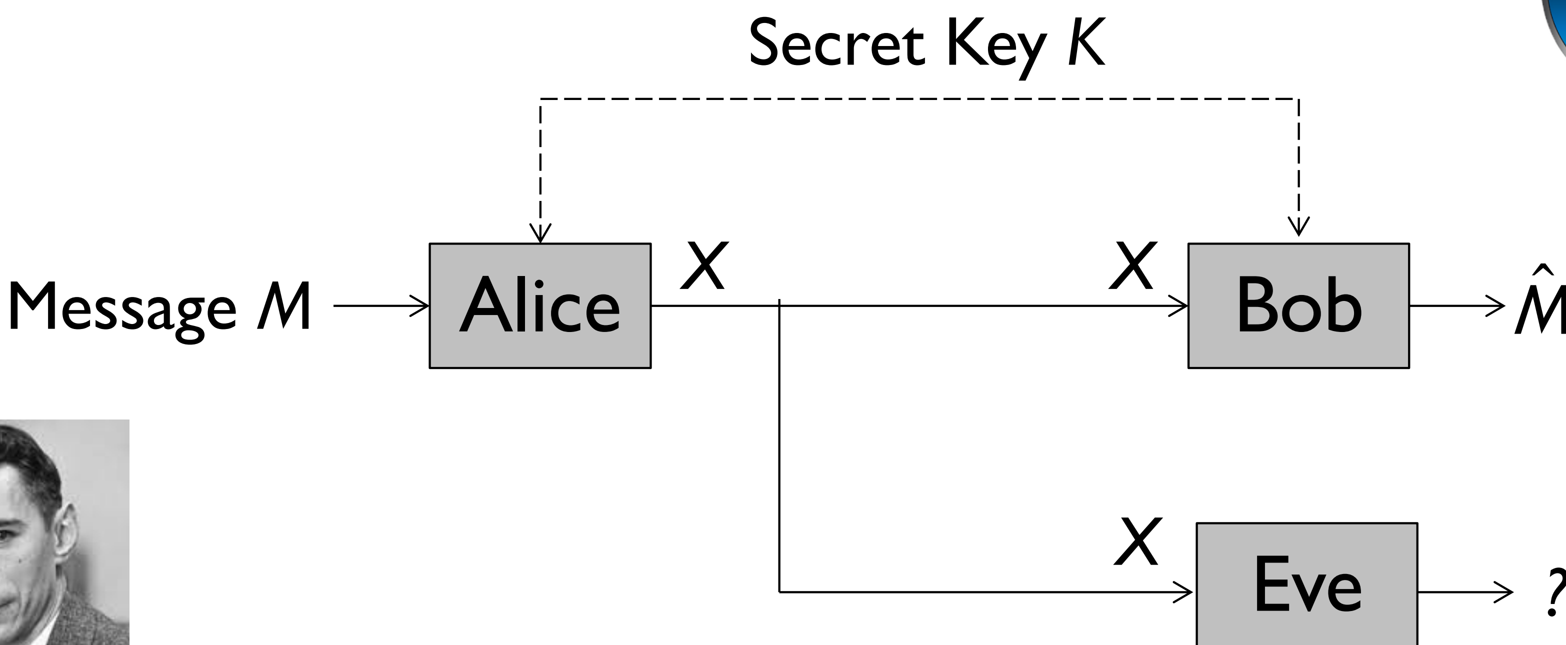
Rethinking Security Design

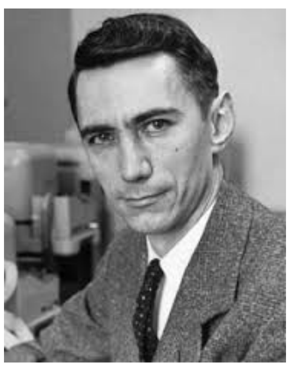
- Conventionally a higher layer issue: encryption, key distributions, …
- Difficult with massive number of devices (esp. with no infrastructure), low cost, low latency.



- Physical layer security provides security by exploiting imperfections in physical channels: noise, fading, …
- Joint encoding for reliability and security.

Information Theoretic Security:
Shannon's Model



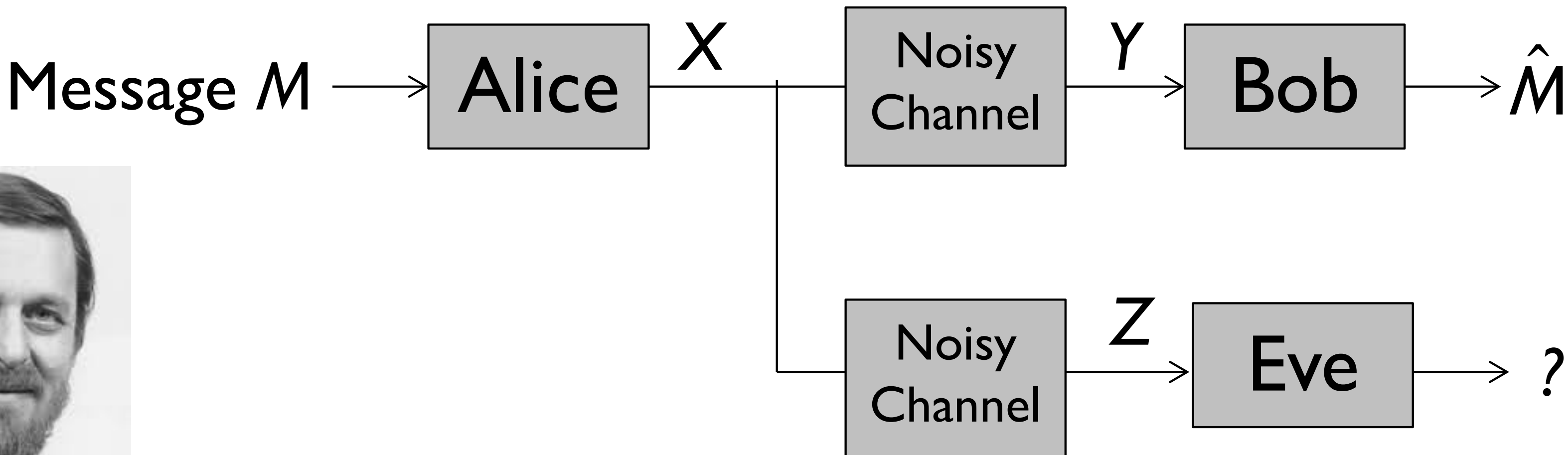


Shannon (1949): For cipher, perfect secrecy requires a one-time pad.

[I.e., the entropy of the key must be at least the entropy of the source: $H(K) \geq H(M)$]

Information Theoretic Security: Wyner's Model

"The Wiretap Channel"



- Tradeoff: reliable rate R to Bob vs. the equivocation $H(M|Z)$ at Eve
- Secrecy capacity = maximum R such that $R = H(M|Z)$
- Wyner (1975): Secrecy capacity > 0 iff. Z is degraded relative to Y

Physical Layer Security

- There has been a resurgence of interest in these ideas.

- In general, the legitimate receiver needs an advantage over the eavesdropper – either a secret shared with the transmitter, or a better channel.

Physical Layer Security

- There has been a resurgence of interest in these ideas.

  - In general, the legitimate receiver needs an advantage over the eavesdropper – either a secret shared with the transmitter, or a better channel.

- The physical properties of radio propagation (diffusion & superposition) provide opportunities for this, via
  - fading: provides natural degradedness over time
  - interference: allows active countermeasures to eavesdropping
  - spatial diversity (MIMO, relays): creates "secrecy degrees of freedom"
  - random channels: sources of common randomness for key generation

Physical Layer Security

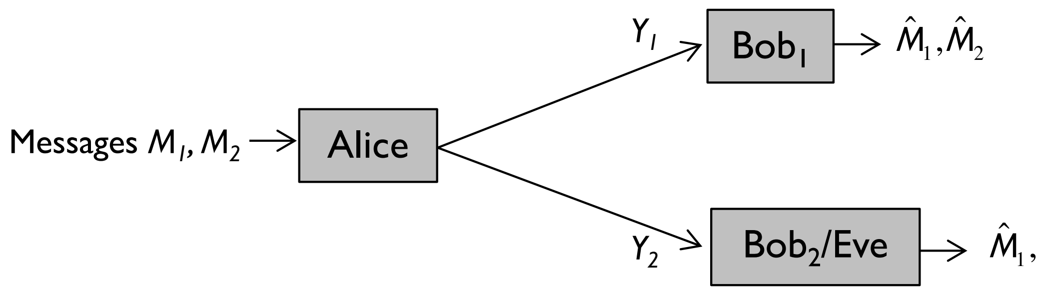
- There has been a resurgence of interest in these ideas.

  - In general, the legitimate receiver needs an advantage over the eavesdropper – either a secret shared with the transmitter, or a better channel.

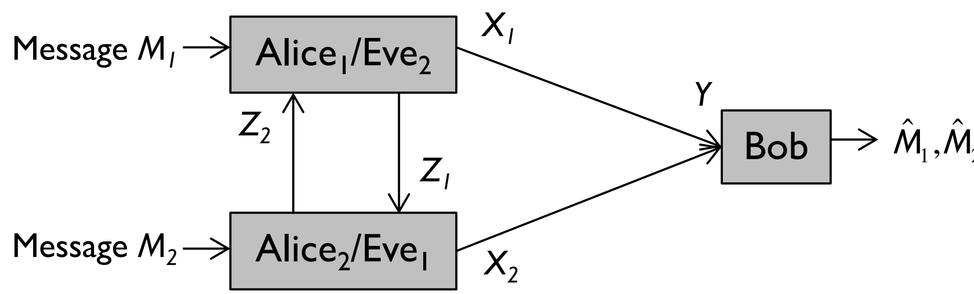
- The physical properties of radio propagation (diffusion & superposition) provide opportunities for this, via
  - fading: provides natural degradedness over time
  - interference: allows active countermeasures to eavesdropping
  - spatial diversity (MIMO, relays): creates "secrecy degrees of freedom"
  - random channels: sources of common randomness for key generation
- The first three of these phenomena lead to rich secrecy capacity regions for the fundamental channel models used to understand wireless networks.

Secrecy in Fundamental Channel Models

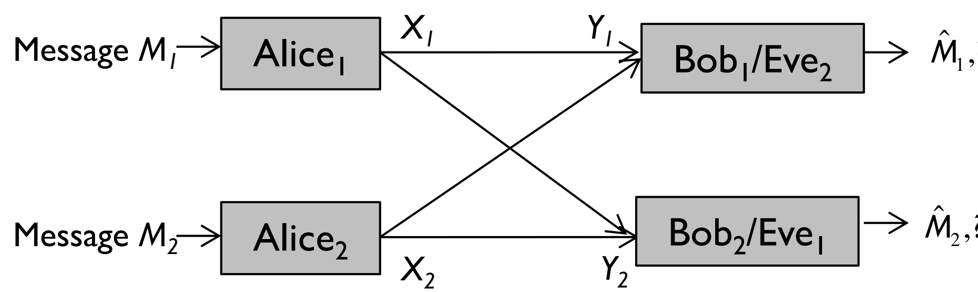
- <u>Broadcast Channels</u>:



- <u>Multiple-Access Channels</u>:



- <u>Interference Channels</u>:



- <u>Relay Channels, MIMO Channels, etc.</u>

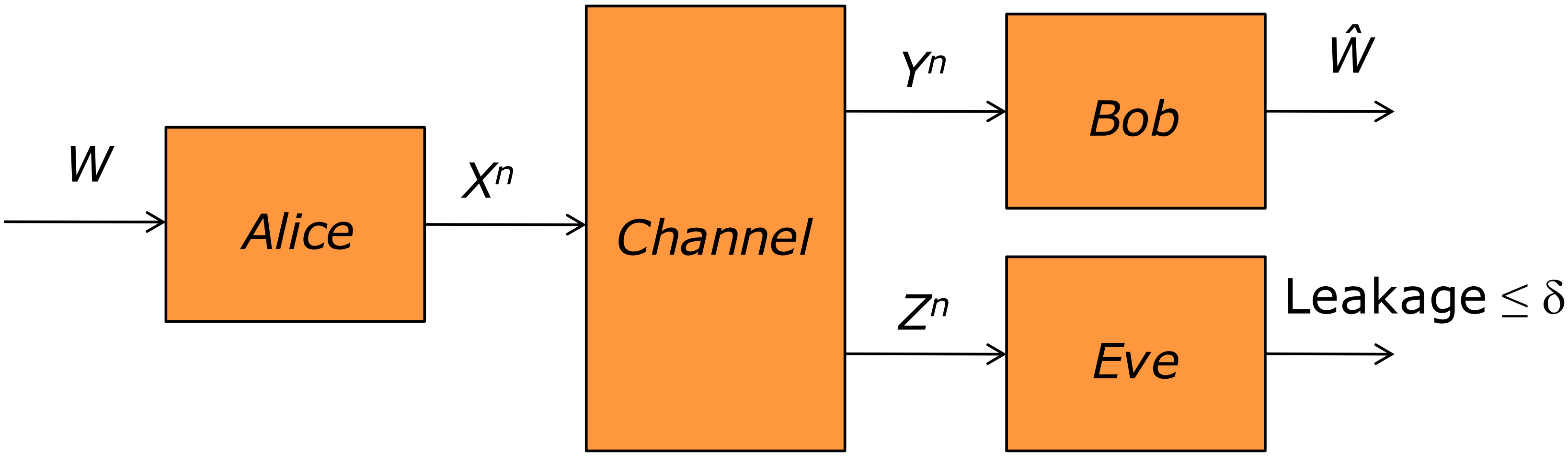
Poor, Schaefer (2017) Wireless Physical Layer Security PNAS

Key Generation from Common Randomness

- Passive Eavesdropper:
  - Public discussion
  - Channel reciprocity: joint source-channel model
  - Relay assisted: trusted or oblivious
- Active Eavesdropper:
  - Channel reciprocity: joint source-channel model

Lai, Liang, Du, Poor (2015)
Key Generation from Random Channels
in *Physical Layer Security in Wireless Communications* (CRC)

Wiretap Channel and Secrecy Capacity

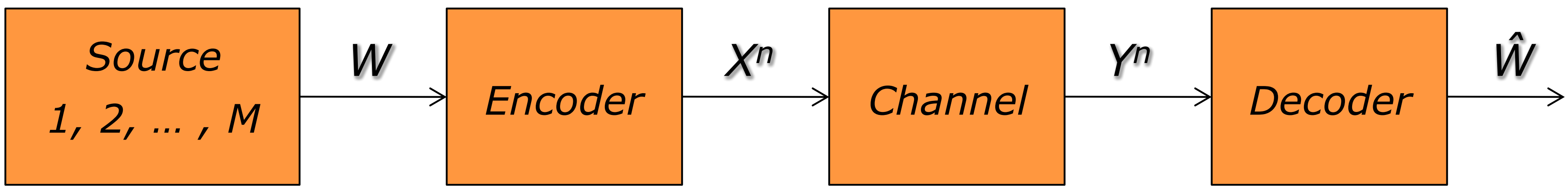


- Secrecy capacity: largest rate in the asymptotic regime of
  - Blocklength $n \to \infty$
  - Probability of error $\mathbb{P}\left(W \neq \hat{W}\right) \to 0$
  - Information leakage $\delta \to 0$

$$C_s = \max_{P_X} \{I(X;Y) - I(X;Z)\}$$

- Limitation: not suitable for low-latency applications as in IoT.

Finite Blocklength Information Theory



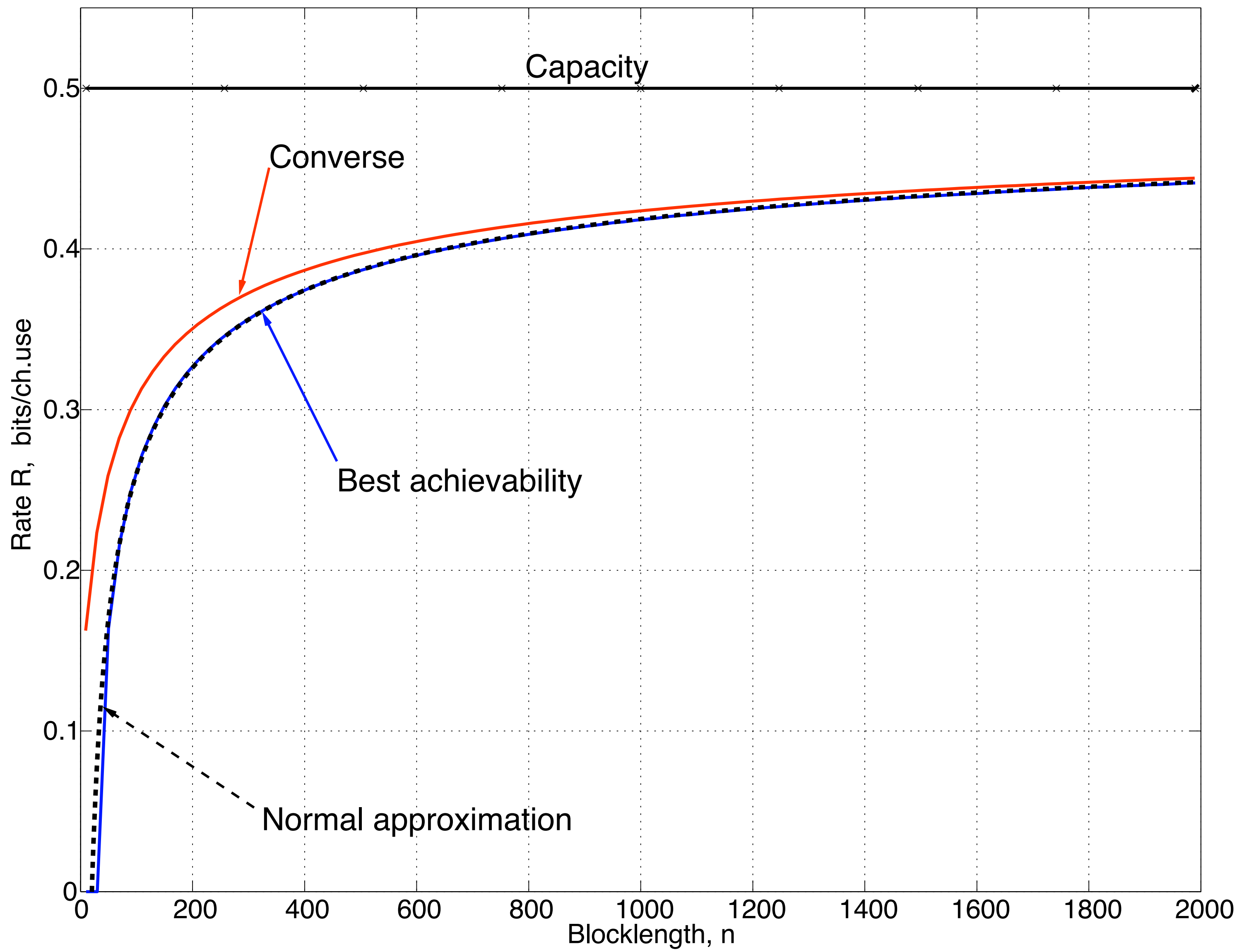
- <u>(n,M,ε) code</u>: $P(W \neq \hat{W}) \leq \varepsilon$
- <u>Fundamental limit</u>: $M^*(n,\varepsilon) = \max\{M: \exists \text{ an } (n,M,\varepsilon) \text{ code}\}$

$$\log M^*(n,\varepsilon) = n\,C - \sqrt{nV}\,Q^{-1}(\varepsilon) + O(\log n)$$

$C = E[i(X^*,Y^*)]$ (Shannon's capacity); $V = \mathrm{Var}[i(X^*,Y^*)]$ ("dispersion")

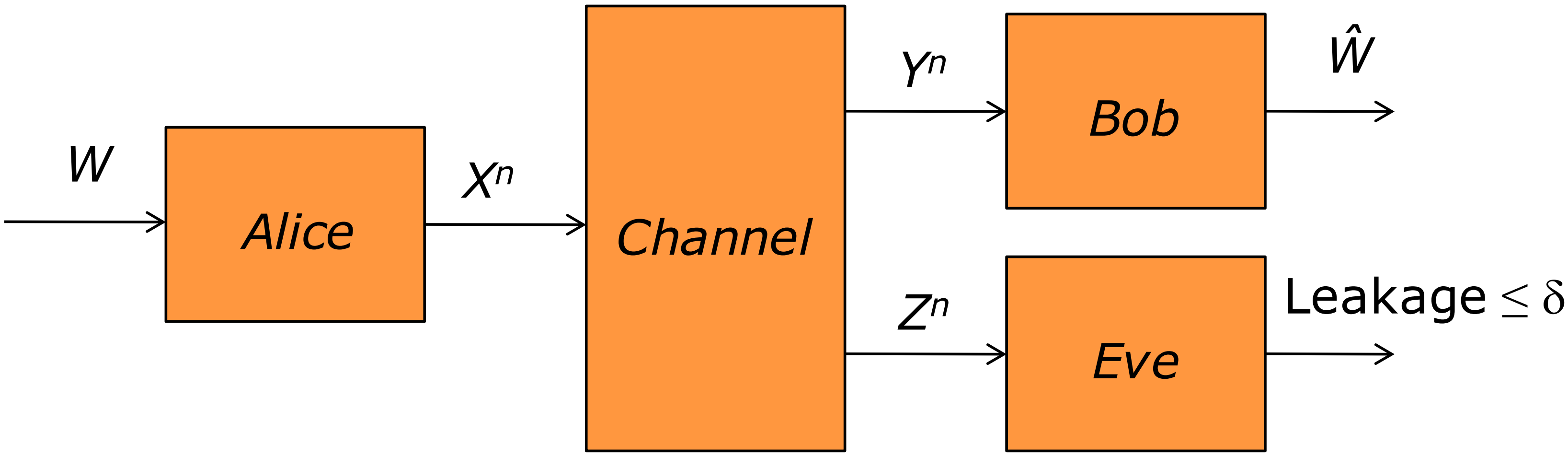
[Polyanskiy, et al. (2010), etc.]

Example: AWGN (SNR = 0 dB; $\varepsilon = 10^{-3}$)



[Polyanskiy, et al. (2010)]

PHY Layer Security: Finite Blocklength



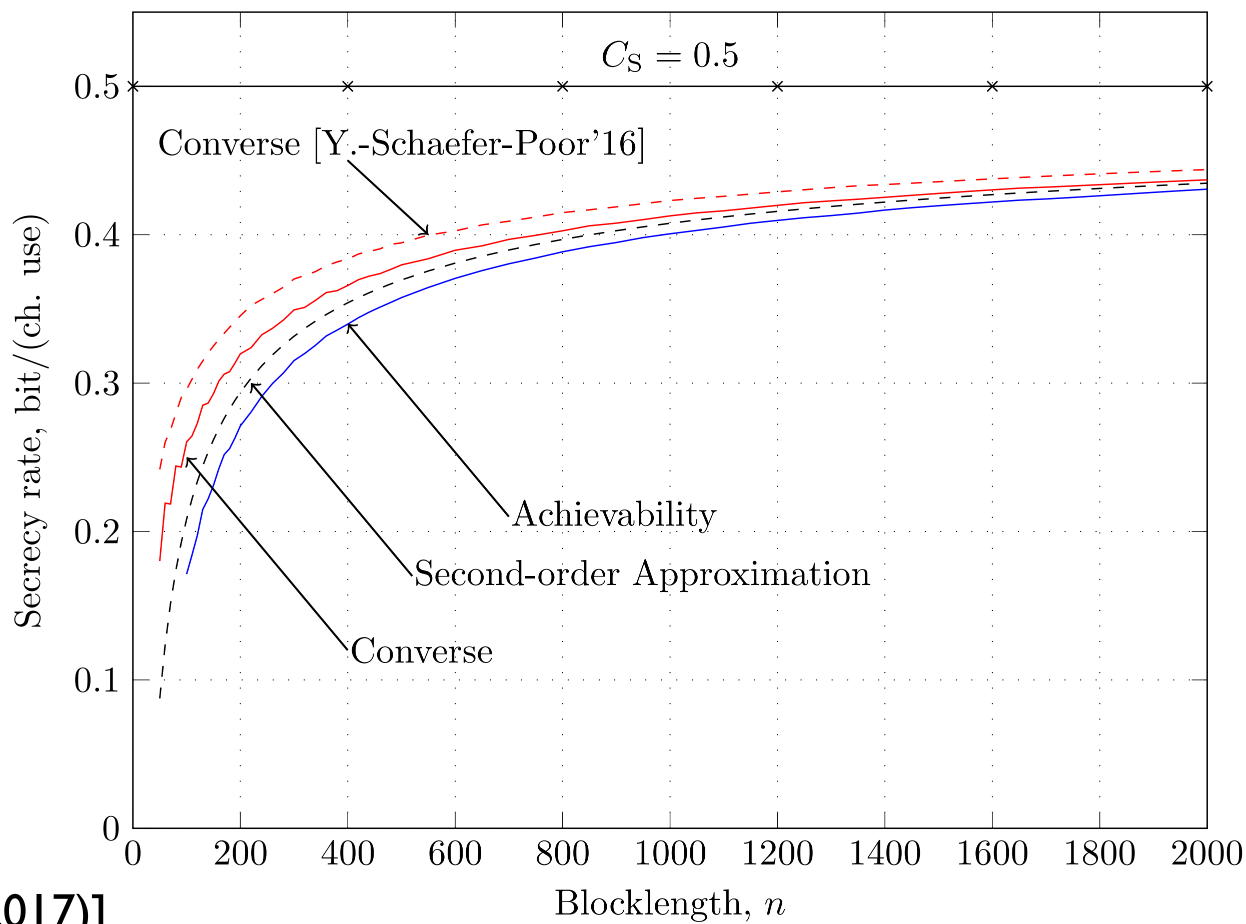
- (M, ϵ, δ) secrecy code:
  - Message $W \in \{1, \ldots, M\}$
  - Encoder $P_{X|W} : \{1, \ldots, M\} \to \mathcal{A}$; decoder $g : \mathcal{B} \to \{1, \ldots, M\}$
  - Average error probability: $\mathbb{P}\left(W \neq \hat{W}\right) \leq \epsilon$
  - Secrecy constraint: information leakage $\leq \delta$

- $R^*(n, \epsilon, \delta)$: maximum secret rate at a given blocklength.

Semi-deterministic Wiretap Channel (BSC): $\delta = \epsilon = 10^{-3}$

- Legitimate channel is deterministic, eavesdropper channel is BSC:

$$R^*(n,\epsilon,\delta) = C_s - \sqrt{\frac{V}{n}} Q^{-1}\left(\frac{\delta}{1-\epsilon}\right) + \mathcal{O}\left(\frac{\log n}{n}\right)$$



[Yang, et al. (2017)]

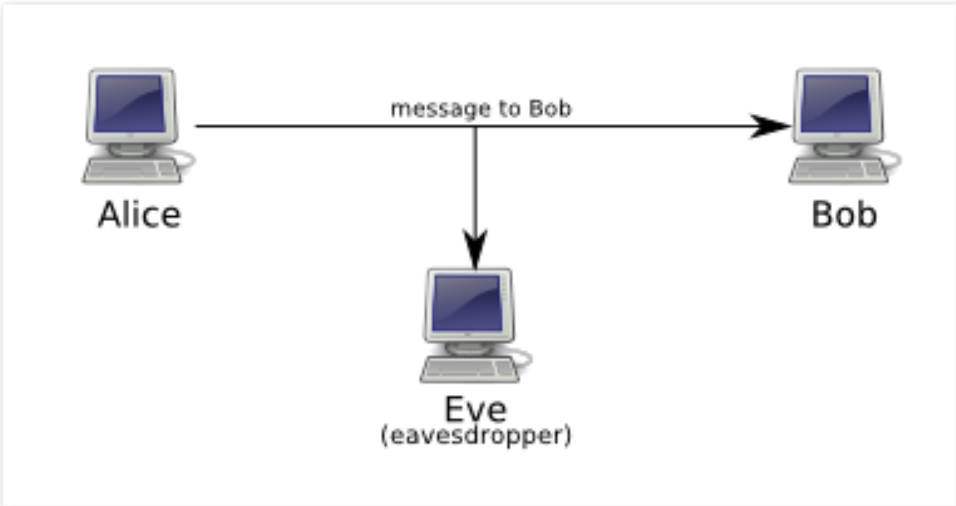
Privacy-Utility Tradeoffs

in

Sensing Systems

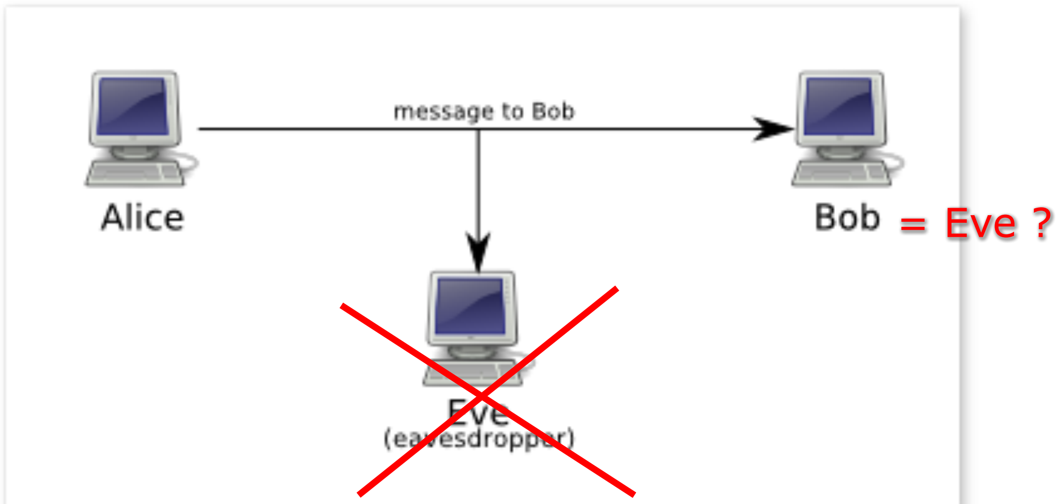
Privacy vs. Secrecy

- Privacy is not secrecy:



Privacy vs. Secrecy

- Privacy is not secrecy:

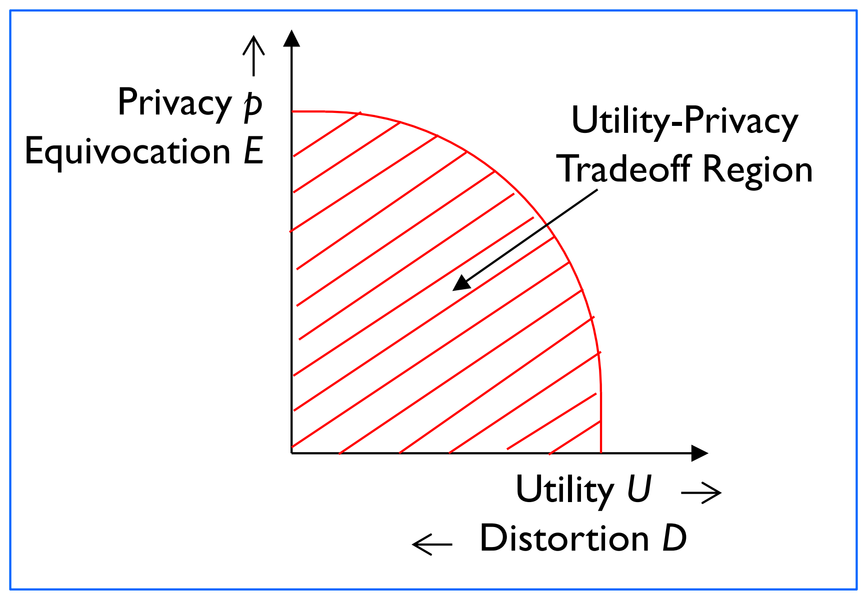


- Denial of access (secrecy) makes a data source useless.

Privacy-Utility Tradeoff

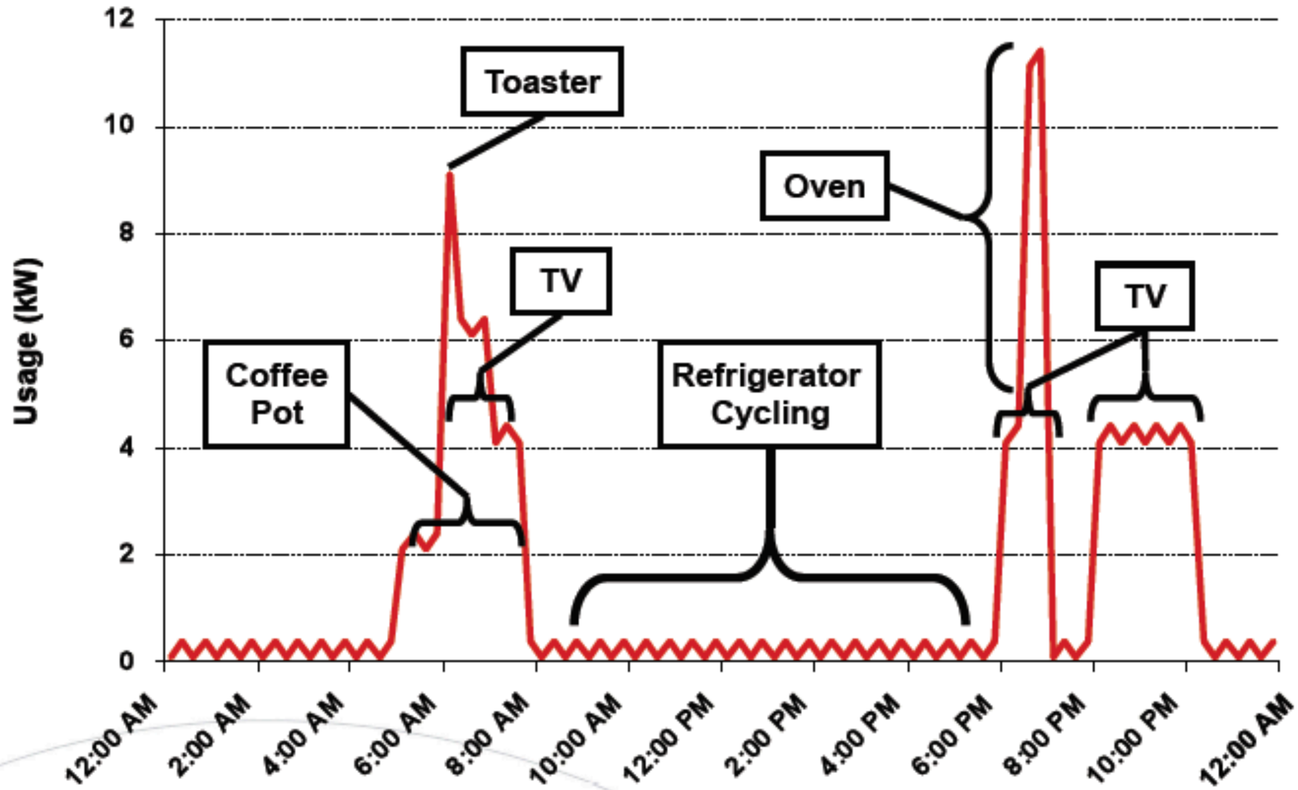
- Sensing systems generate considerable electronic data:

- Data's utility depends on its accessibility.
- Accessibility endangers privacy.
- This fundamental tradeoff can be characterized via information theory.



Example: Smart Meter Privacy

- Smart meter data is useful for price-aware usage, load balancing
- But, it leaks information about in-home activity

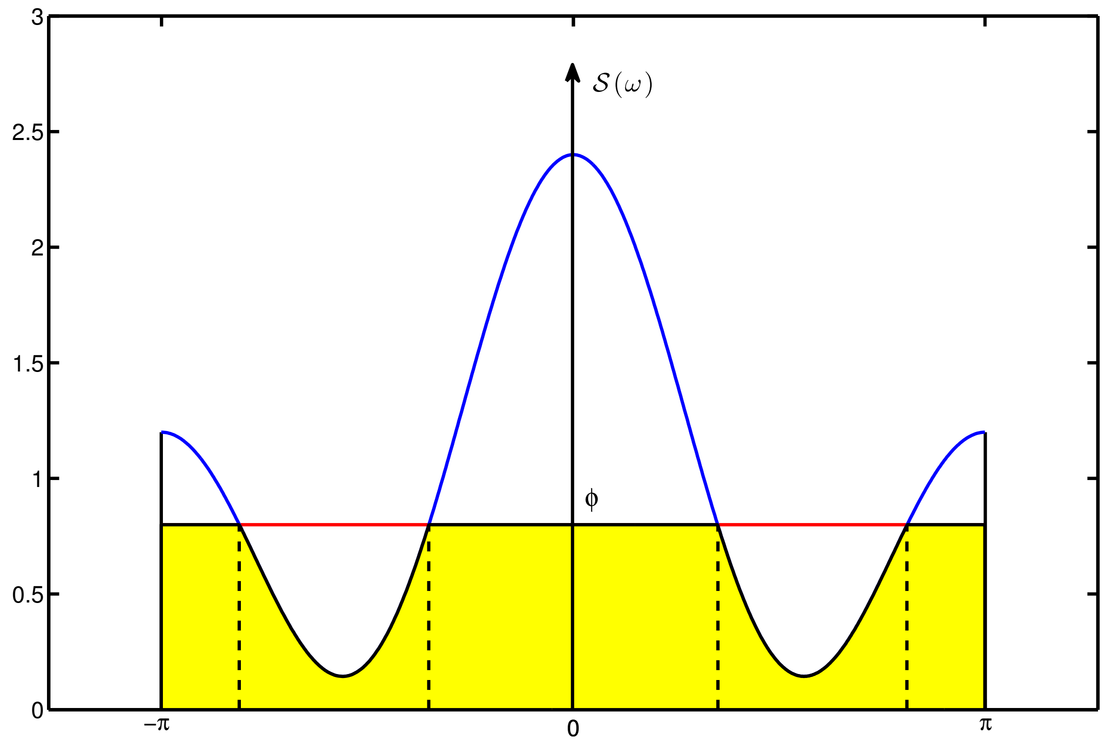


Poor (2017)
Privacy in the Smart Grid: Information, Control & Games
In *Information Theoretic Security and Privacy of Information Systems* (Cambridge)

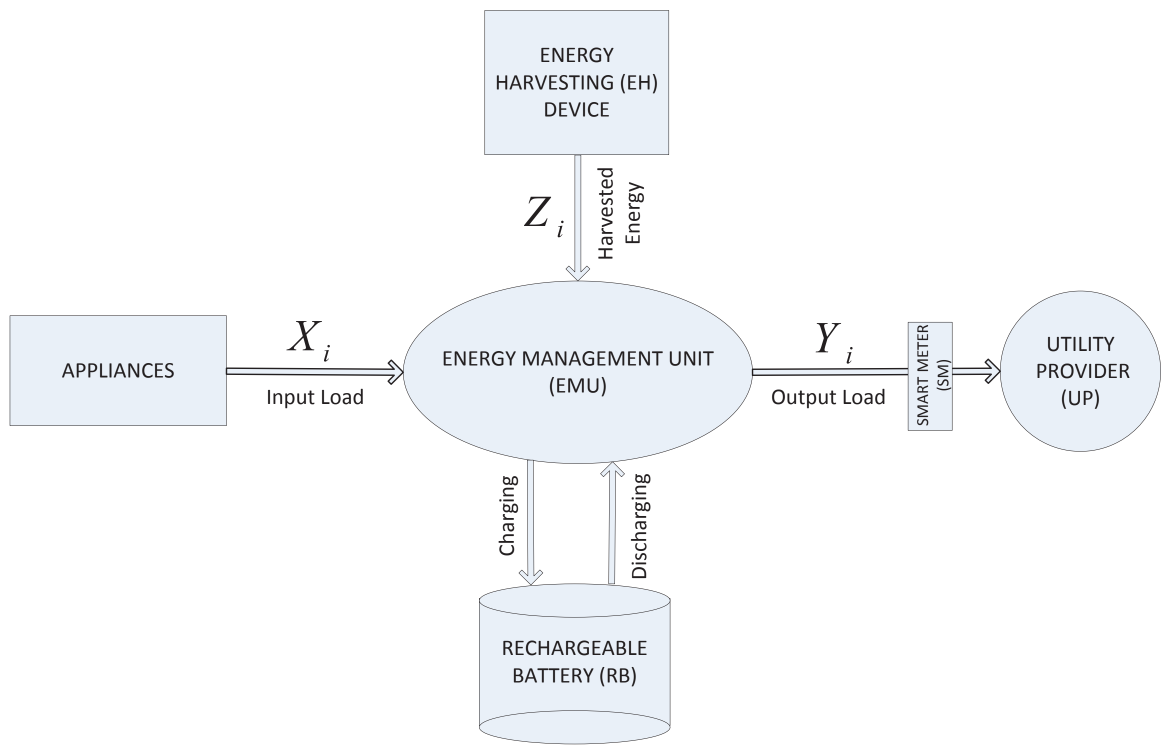
Source Coding Solution:

Hidden Gauss-Markov Model (protection of the hidden intermittency state)

P-U tradeoff leads to a spectral ‘reverse water-filling’ solution

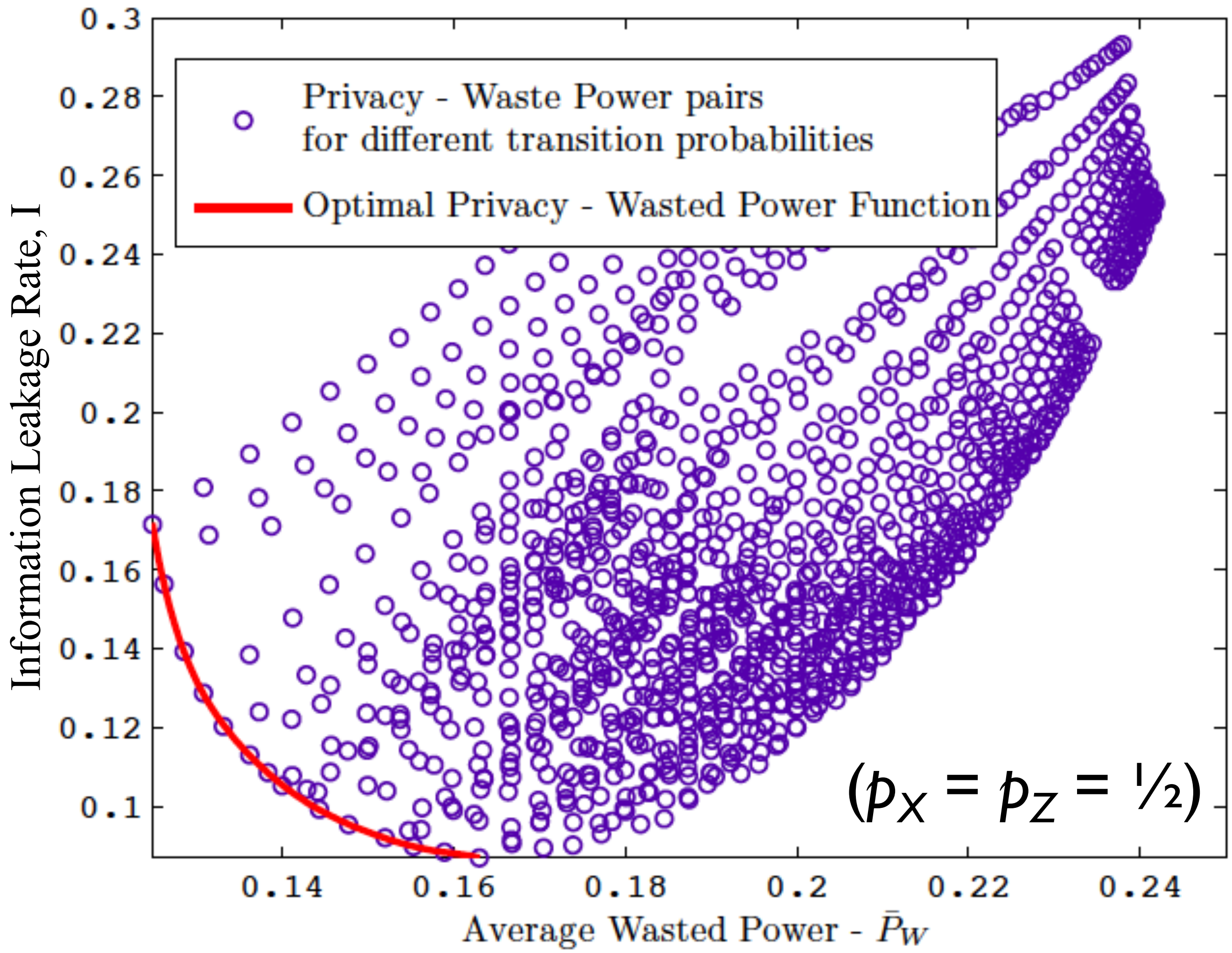


A Control Approach:
Energy Harvesting and Storage



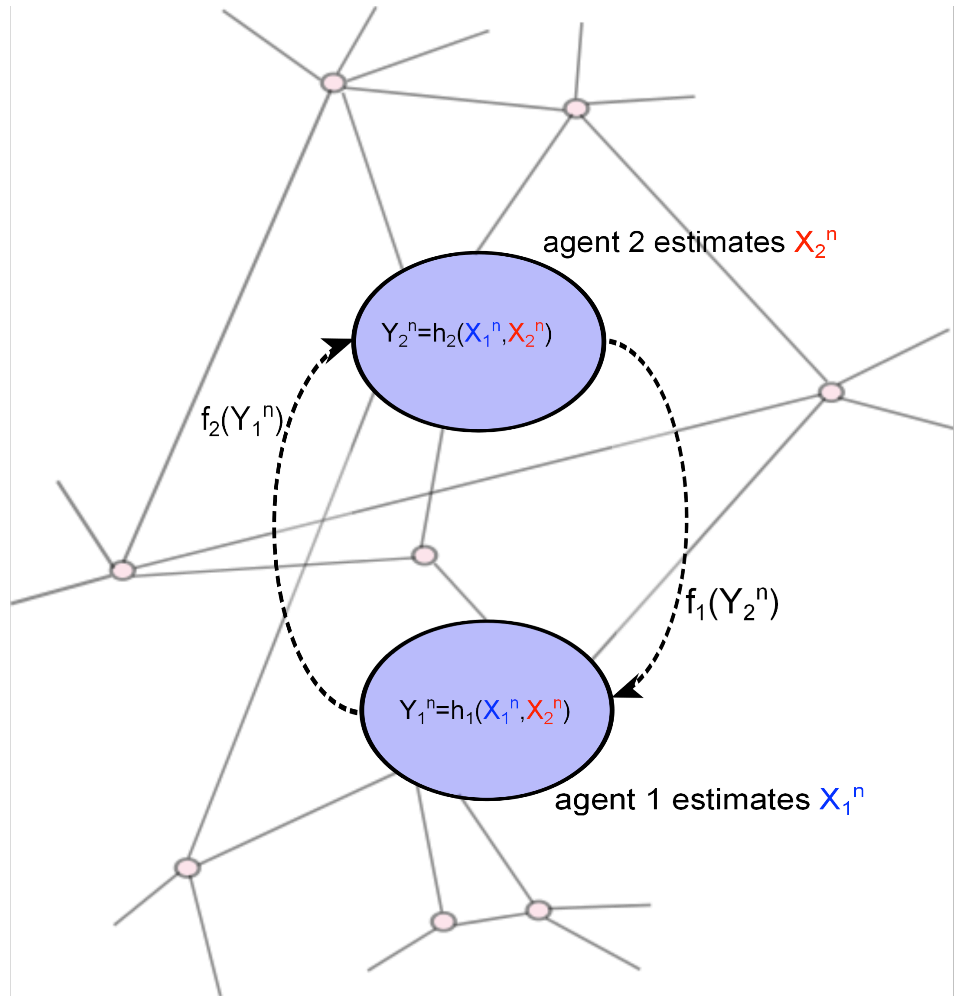
Tradeoff: wasted energy versus information leakage

Privacy-Utility Tradeoff:
Binary Variables

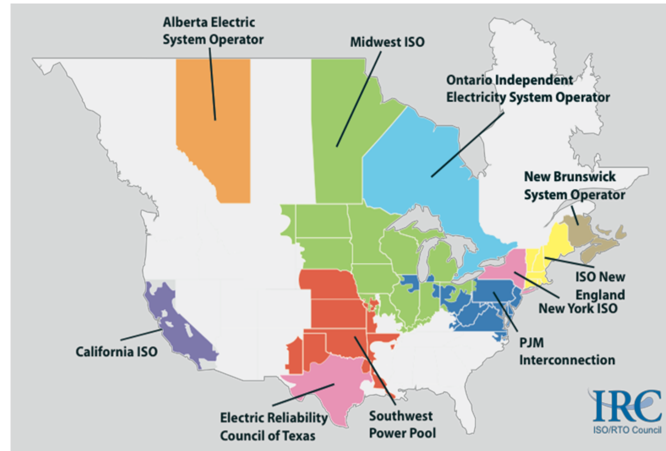


Competitive Privacy: Privacy-Utility Tradeoffs for Interacting Agents

- Multiple interacting, but competing, agents (or groups of agents) with coupled measurements.
- Each wants to estimate its own parameters, or state.
- They can help each other by sharing data, but wish to preserve privacy.
- Each has a privacy-utility tradeoff, but they are competitive ones.
- How should they interact?

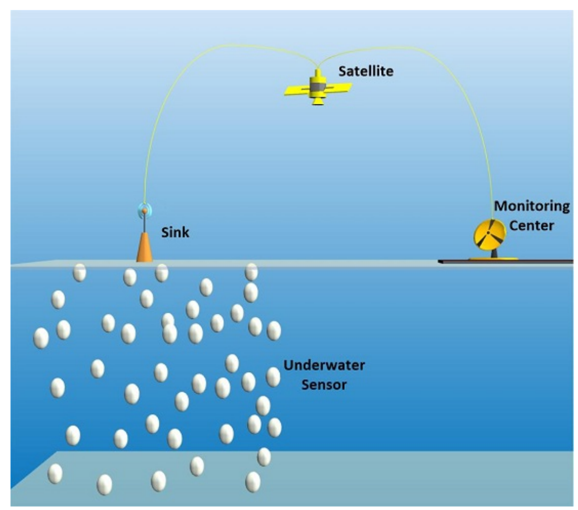


Motivating Examples



Electricity Grids: grid management

Radar: untrustworthy allies



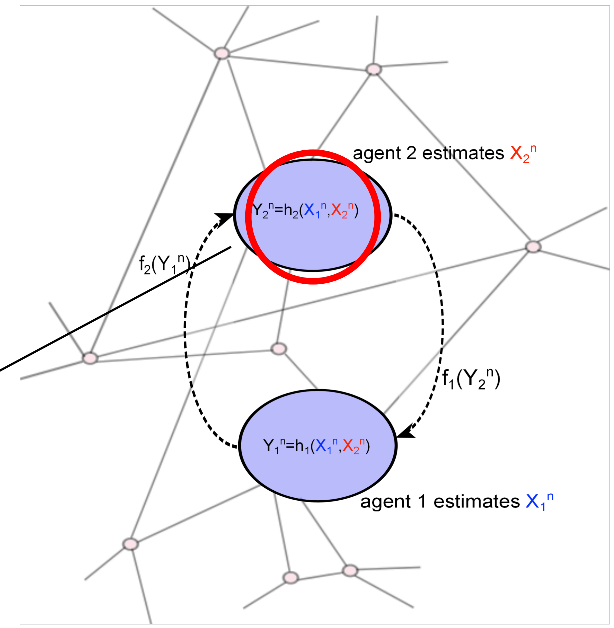
Sensor Networks: resource localization

Linear Measurement Model

- Noisy measurements at agent k:

$$Y_k = \sum_{m=1}^{M} H_{k,m} X_m + Z_k,\; k = 1,2,\ldots,M$$

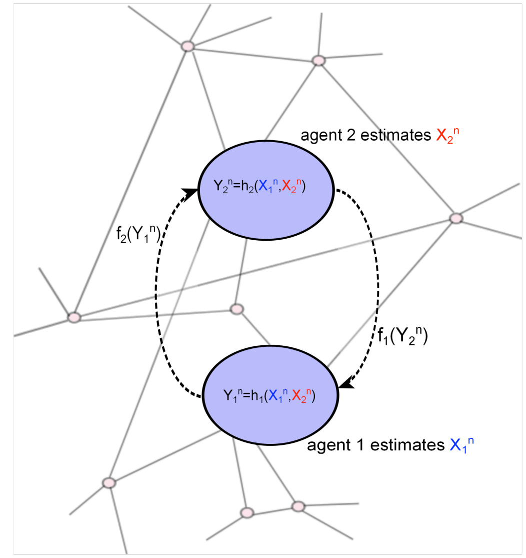
m^{th} system state



- Utility for agent k: mean-square error for its own state X_k

- Privacy for agent k: leakage of information about X_k to other agents

How Should Agents Exchange Data?



- This is a classical problem in information theory – the Wyner-Ziv problem (optimal distributed source coding) – which tells how to exchange information.
- But, doesn't say how much information to exchange.

- Because of the competitive nature, game theory or prospect theory can illuminate this.
- Leads to a number of interesting solutions:
  - a basic problem is a prisoners' dilemma
  - with pricing, cooperation or multi-play games, more meaningful solutions arise

Poor (2018)
Privacy in Networks of Interacting Agents
in *Emerging Applications of Control and System Theory* (Springer)

Other Issues

Other Issues

- Authentication
  - Information theoretic bounds on the probabilities of successful impersonation and substitution attacks [Lai, et al. IT-09]
  - Privacy-security tradeoffs in biometric authentication systems [Lai et al. IFS-11]

Other Issues

- Authentication
  - Information theoretic bounds on the probabilities of successful impersonation and substitution attacks [Lai, et al. IT-09]
  - Privacy-security tradeoffs in biometric authentication systems [Lai et al. IFS-11]
- Attacks on MANETs
  - Information theoretic guidance on how many malicious nodes can be tolerated [Liang, et al. IT-11]

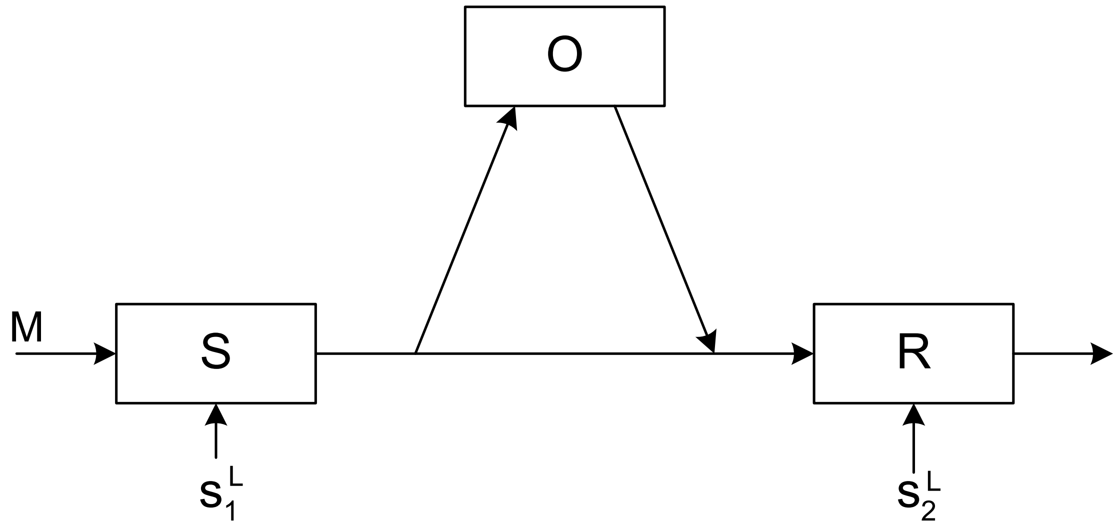
Other Issues

- Authentication
  - Information theoretic bounds on the probabilities of successful impersonation and substitution attacks [Lai, et al. IT-09]
  - Privacy-security tradeoffs in biometric authentication systems [Lai et al. IFS-11]
- Attacks on MANETs
  - Information theoretic guidance on how many malicious nodes can be tolerated [Liang, et al. IT-11]
- Data Injection Attacks on Smart Grids
  - Information theoretic guidance on protection against stealth attacks [Sun, et al. SmartGridComm'17]

Other Issues

- Authentication
  - Information theoretic bounds on the probabilities of successful impersonation and substitution attacks [Lai, et al. IT-09]
  - Privacy-security tradeoffs in biometric authentication systems [Lai et al. IFS-11]
- Attacks on MANETs
  - Information theoretic guidance on how many malicious nodes can be tolerated [Liang, et al. IT-11]
- Data Injection Attacks on Smart Grids
  - Information theoretic guidance on protection against stealth attacks [Sun, et al. SmartGridComm'17]
- Man-in-the-Middle and Spoofing Attacks on Sensor Nets
  - Effects on CRLB in parameter estimation [Zhang, et al. SPM'18]

Authentication with Correlated Sequences



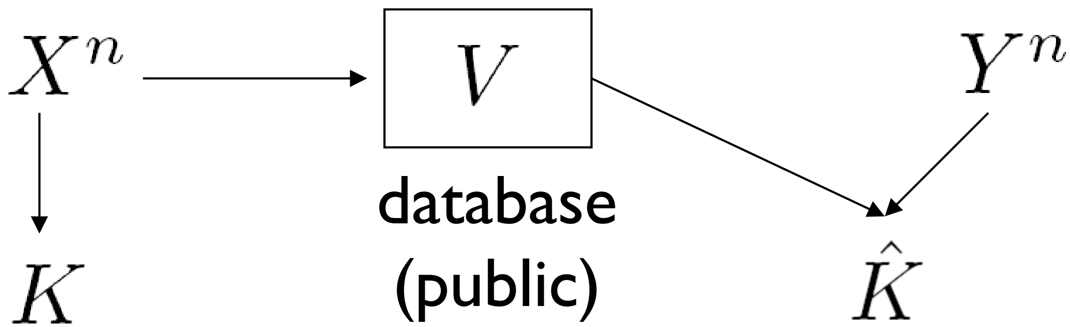
Impersonation attack: O transmits a message before S

Substitution attack: O replaces S's message with its own

Theorem [Lai, et al. IT-09]: If the S-O channel is not less noisy than the S-R channel, then

$$P_I = P_S = 2^{-LI(S_1;S_2)}$$

Biometric Authentication



Two performance metrics:

Utility = key rate: $R = n^{-1}H(K)$

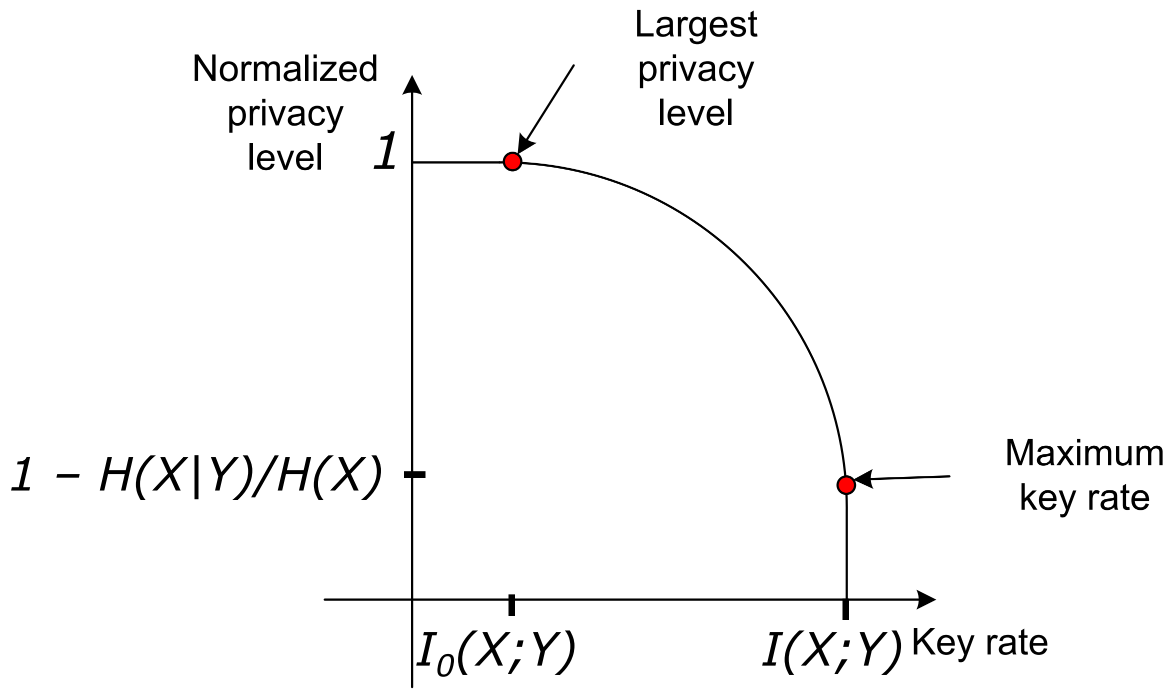
Authentication: the number of attacker's guesses

Privacy level: $\Delta_P = H(X^n|V)/H(X^n)$

Normalized privacy level of the biometric measurements.

What's the tradeoff between these two?

Biometric Authentication: The Tradeoff



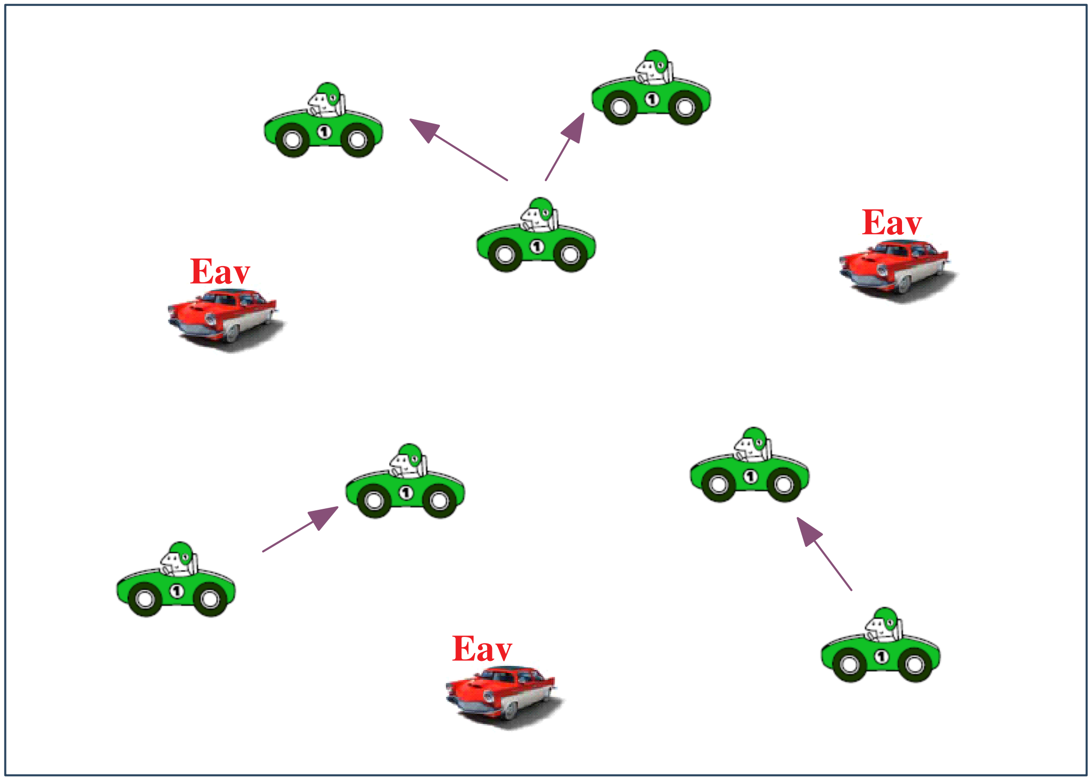
Theorem [Lai, et al. IFS-11]:

(Δ_P, R) is achievable, if and only if there exists $U \to X \to Y$ such that

$$\Delta_P \leq 1 - \frac{I(U;X) - I(U;Y)}{H(X)}$$

$$R \leq I(U;Y),$$

MANETs with Malicious Nodes



- n legitimate mobile nodes
- Each legitimate node is both a source and a destination.
- m malicious nodes

Secrecy Capacity Scaling

[Liang, et al. IT'11]

- Case I: $m = o(\sqrt{nD})$
  - # of malicious nodes is small
  - Type II packets (two-hop scheme) dominate
  - $C_s = \Theta\left(\sqrt{\frac{D}{n}}\right)$
  - Presence of malicious nodes has negligible impact
- Case II: $m = \Omega(\sqrt{nD}poly(n))$
  - # of malicious nodes is large
  - Type I packets (one-hop scheme) dominate
  - $C_s = \Theta\left(\frac{1}{m}\right)$
  - Secrecy throughput is determined by # of malicious nodes

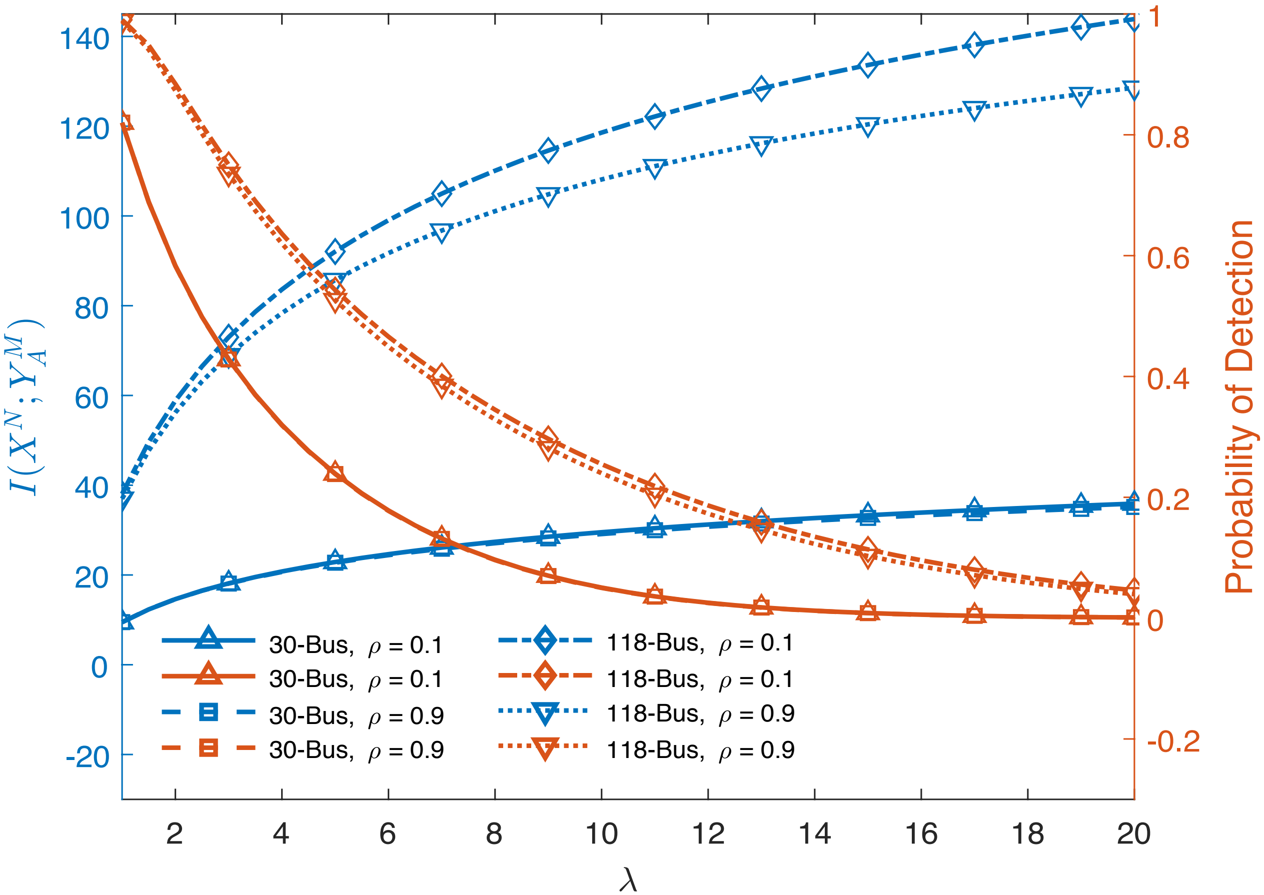
Stealth Attacks on Smart Grids

[Sun, et al., SG - under review]

Stealth attacks seek to trade off:

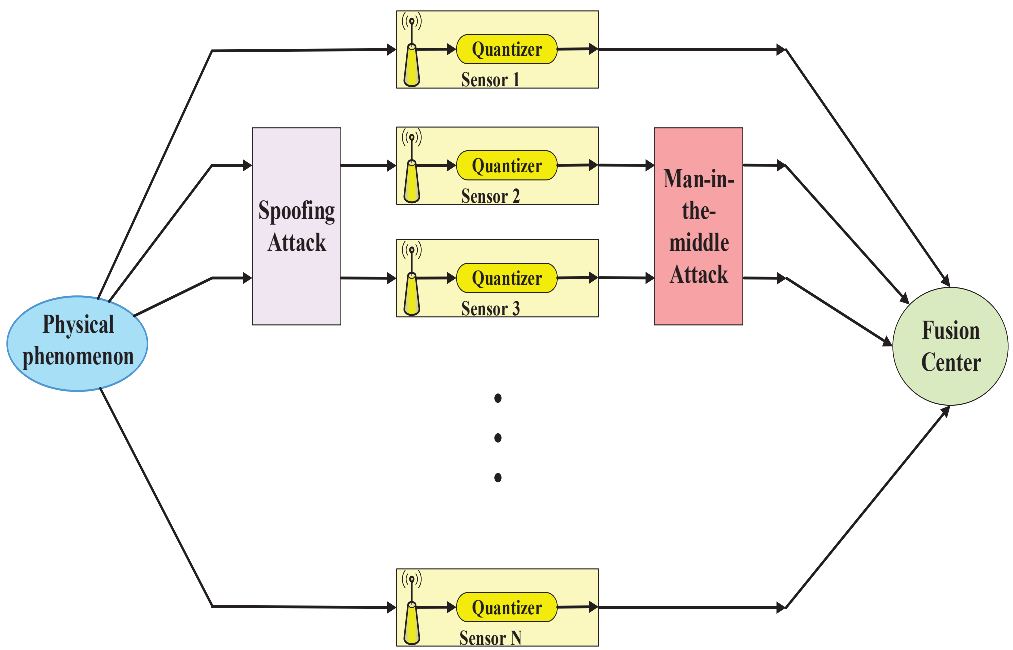
- mutual information between the grid state and operator's observations

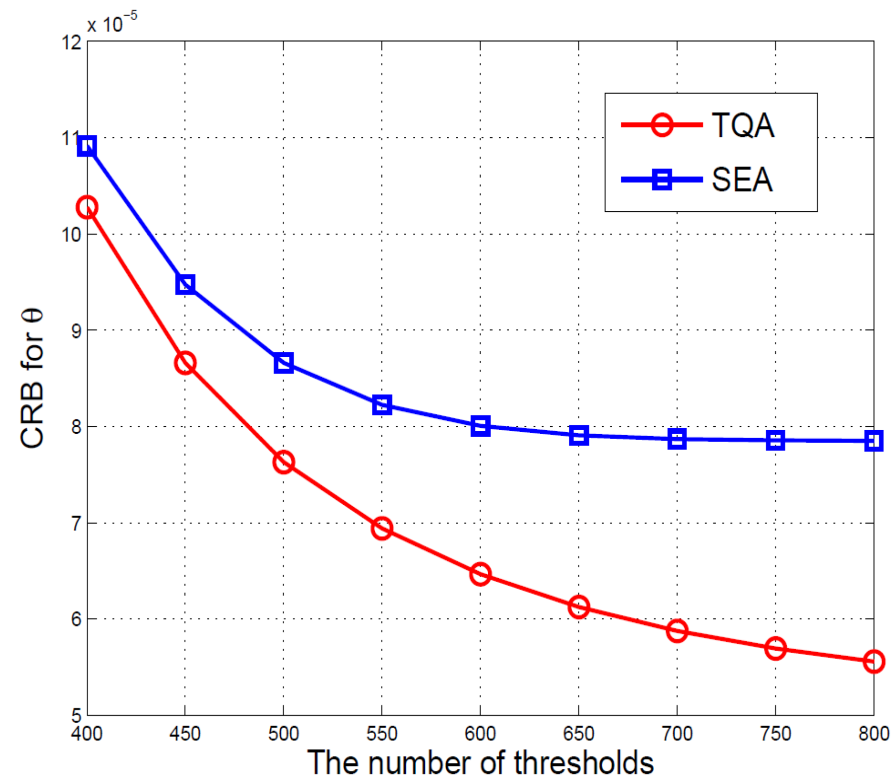
- probability of the attack being detected



Attacks on Sensor Nets

[Zhang, et al. SPM'18]





Man-in-the-middle attack:

- TQA uses attacked data
- SEA ignores attacked data

Summary

- Information theory can help understand some fundamental limits of security and privacy in IoT

- These are theoretical constructs; although they sometimes point to potential practical solutions, there are many needs to connect this kind of analysis to real networks, e.g.
  - more finite-blocklength analysis
  - scaling laws for large networks
  - practical coding schemes to achieve fundamental limits
  - other security primitives (signatures, certificates, etc.)

Some Basic References

Lai, Liang, Du, Poor (2015) Key Generation from Random Channels, in *Physical Layer Security in Wireless Communications* (CRC)

Schaefer, Boche, Khisti, Poor (2017) *Information Theoretic Security and Privacy of Information Systems* (Cambridge)

Poor, Schaefer (2017) Wireless Physical Layer Security, PNAS.

Poor (2017) Privacy in the Smart Grid: Information, Control & Games, in *Information Theoretic Security and Privacy of Information Systems* (Cambridge)

Poor (2018) Privacy in Networks of Interacting Agents, in *Emerging Applications of Control and System Theory* (Springer)

Thank You!