Topics in Information-Theoretic Cryptography

Course Introduction

Yanina Shkel, September 30, 2021

About the Instructor

At EPFL: since December 2019

BS: mathematics & computer science, UW-Madison

MS/PhD: Information Theory, UW-Madison

Postdoc: Princeton University, UIUC

Research Interests: theory of data compression, mathematical models for privacy and secrecy, fundamental limits of privacy-aware information processing systems, information-theoretic methods in cryptography

Dr. Yanina Shkel

Course Logistics

Time and Location: Thursdays 8:15-10am, BC03 and on Zoom

Course Webpage: Moodle

Course Format: Based on paper reading and presentation during lecture

Grading: The final grade will be based 20% on course participation and 80% on the final project

Enrolment: Current enrolment on ISAcademia is low :(
We will decide on Monday/Tuesday, October 4/5 wether to go on with the course

Format

Paper Reading

- There will be 1-2 papers assigned each lecture
- We will discuss the papers during lecture
- You will get more from the course if you read the papers
- You are not expected to read papers labeled as “Further Reading”
- Final project assignments includes extension of existing results, implementation tasks, critical summary of a paper, etc.
- You may use a paper from the provided reading list or suggest their own paper.
- You should communicate your final project topic to the instructor a month before the final due date

Tentative Course Schedule:

Lecture 1 – September 30 – Course Introductions

Reading Assignment

- The Princeton Companion to Applied Mathematics. Princeton University Press, 2015, ch. IV.36 Information Theory by Sergio Verdú

Further Reading

- C. E. Shannon, “A mathematical theory of communication,” The Bell System Technical Journal,, vol. 27, pp. 379–423 and 623–656, July/October 1948.

Lecture 2 – October 7 – Réniy Entropy and Axiomatic Definitions of Entropy

Reading Assignment

- Rényi, “On measures of entropy and information,” in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. University of California Press, 1961, pp. 547–561.

Further Reading

- L. Campbell, “A coding theorem and Rényi’s entropy,” Information and Control, vol. 8, no. 4, pp. 423 – 429, 1965.
- S. Verdú, “Error exponents and alpha-mutual information,” Entropy, vol. 23, no. 2, 2021.
- Y. Shkel and S. Verdú, “A coding theorem for f-separable distortion measures,” Entropy, vol. 20, no. 2, 2018.

Lecture 3 – October 14 – TBD

Reading Assignment

Further Reading

…

Lecture 7 – November 11 – TBD

Reading Assignment

Further Reading

November 18 – Final Project Proposal Due

Lecture 8 – November 18 – TBD

Reading Assignment

Further Reading

…

Lecture 12 – December 14 – TBD

Reading Assignment

Further Reading

December 21 – Final Project Due

Format
Paper Reading

- As you read each paper, consider:
  - What is the “broad strokes” problem to be addressed?
  - Why is it important?
  - How does the paper advance the state-of-the-art?
  - What is the history and the previous state-of-the-art?
  - What is the main result?
  - What are the main technical tools/ideas/insights?
  - What is the impact (for older papers)? What is possible future work (for newer papers)?
- It is OK to not understand everything in the paper
  - These are deep and technical works

Tentative Course Schedule:

Lecture 1 – September 30 – Course Introductions
Reading Assignment
- The Princeton Companion to Applied Mathematics. Princeton University Press, 2015, ch. IV.36 Information Theory by Sergio Verdú

Further Reading
- C. E. Shannon, “A mathematical theory of communication,” The Bell System Technical Journal,, vol. 27, pp. 379–423 and 623–656, July/October 1948.

Lecture 2 – October 7 – Réniy Entropy and Axiomatic Definitions of Entropy
Reading Assignment
- Rényi, “On measures of entropy and information,” in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. University of California Press, 1961, pp. 547–561.

Further Reading
- L. Campbell, “A coding theorem and Rényi’s entropy,” Information and Control, vol. 8, no. 4, pp. 423 – 429, 1965.
- S. Verdú, “Error exponents and alpha-mutual information,” Entropy, vol. 23, no. 2, 2021.
- Y. Shkel and S. Verdú, “A coding theorem for f-separable distortion measures,” Entropy, vol. 20, no. 2, 2018.

Lecture 3 – October 14 – TBD
Reading Assignment
Further Reading

…

Lecture 7 – November 11 – TBD
Reading Assignment
Further Reading

November 18 – Final Project Proposal Due

Lecture 8 – November 18 – TBD
Reading Assignment
Further Reading

…
Lecture 12 – December 14 – TBD
Reading Assignment
Further Reading

December 21 – Final Project Due

Tentative Reading List:

- C. E. Shannon, “A mathematical theory of communication,” The Bell System Technical Journal,, vol. 27, pp. 379–423 and 623–656, July/October 1948.
- N. Higham, M. Dennis, P. Glendinning, P. Martin, F. Santosa, and J. Tanner, Eds., The Princeton Companion to Applied Mathematics. Princeton University Press, 2015, ch. IV.36 Information Theory by Sergio Verdú.
- A. Rényi, “On measures of entropy and information,” in Proceedings of the Fourth Berkeley Sympo- sium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. University of California Press, 1961, pp. 547–561.
- L. Campbell, “A coding theorem and Rényi’s entropy,” Information and Control, vol. 8, no. 4, pp. 423 – 429, 1965.
- S. Verdú, “Error exponents and alpha-mutual information,” Entropy, vol. 23, no. 2, 2021. [Online]. Available: https://www.mdpi.com/1099-4300/23/2/199
- Y. Shkel and S. Verdú, “A coding theorem for f-separable distortion measures,” Entropy, vol. 20, no. 2, 2018. [Online]. Available: https://www.mdpi.com/1099-4300/20/2/111
- I. Issa, A. B. Wagner, and S. Kamath, “An operational approach to information leakage,” IEEE Trans- actions on Information Theory, vol. 66, no. 3, pp. 1625–1657, 2020.
- J. Liao, O. Kosut, L. Sankar, and F. du Pin Calmon, “Tunable measures for information leakage and applications to privacy-utility tradeoffs,” IEEE Transactions on Information Theory, vol. 65, no. 12, pp. 8043–8066, 2019.
- C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, “Our data, ourselves: Privacy via distributed noise generation,” in Advances in Cryptology - EUROCRYPT 2006, S. Vaudenay, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 486–503.
- C. Dwork, F. McSherry, K. Nissim, and A. Smith, “Calibrating noise to sensitivity in private data analysis,” in Proceedings of the Third Conference on Theory of Cryptography, ser. TCC’06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 265–284. [Online]. Available: http://dx.doi.org/10.1007/11681878 14
- C. Dwork, “Differential privacy,” in Automata, Languages and Programming, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1–12.
- G. Barthe and F. Olmedo, “Beyond differential privacy: Composition theorems and relational logic for f-divergences between probabilistic programs,” in Automata, Languages, and Programming, F. V. Fomin, R. Freivalds, M. Kwiatkowska, and D. Peleg, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 49–60.
- P. Kairouz, S. Oh, and P. Viswanath, “The composition theorem for differential privacy,” IEEE Transactions on Information Theory, vol. 63, no. 6, pp. 4037–4049, 2017.
- F. d. P. Calmon, A. Makhdoumi, M. M´edard, M. Varia, M. Christiansen, and K. R. Duffy, “Principal inertia components and applications,” IEEE Transactions on Information Theory, vol. 63, no. 8, pp. 5011–5038, Aug 2017.
- Y. Y. Shkel, R. S. Blum, and H. V. Poor, “Secrecy by design with applications to privacy and compres- sion,” IEEE Transactions on Information Theory, vol. 67, no. 2, pp. 824–843, 2021.
- K. Chatzikokolakis, G. Cherubin, C. Palamidessi, and C. Troncoso, “The bayes security measure,” 2020.
- E. Balsa, C. Troncoso, and C. Diaz, “A metric to evaluate interaction obfuscation in online social networks,” International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 20, 12 2012.
- R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, “Quantifying location privacy,” in 2011 IEEE Symposium on Security and Privacy, 2011, pp. 247–262.

ABOUT THE COURSE

Overview

Fall 2020

Computational security: relies on computational infeasibility of breaking the system.

Unconditional (information-theoretic) security: relies on theoretical impossibility of breaking the system (even given a computationally unbounded adversary).

This course: compare and contrast the two notions of security, look at methods that have information-theoretic guarantees.

Topics

Fall 2020

Secret Communication Problem: Perfect Secrecy vs Computational approach (one-way functions, secret key agreement, semantic security)

Secret Key Generation Problem: Randomness extraction, privacy amplification, secret key capacity

Differential Privacy: Definition, motivation, properties, information-theoretic perspective

Emerging Trends: Other measures of leakage, Maximal leakage, Perfect privacy/secrecy by design

Overview

Fall 2021

Information Measures: How do we model information mathematically? How do we measure information?

Example: Communication Information measures arise as answers to specific engineering questions.

Example: Privacy Information measures are postulated from first principles and used to measure privacy loss.

Course Takeaway: The ‘best’ way to measure information often depends on application.

Topics

Fall 2021

Information theory: entropy, mutual information, and relative entropy; their extensions to Renyi entropy, Renyi divergence, alpha-mutual information; various notions of common information;

Differential Privacy: definition, motivation, extensions, properties, information-theoretic perspective

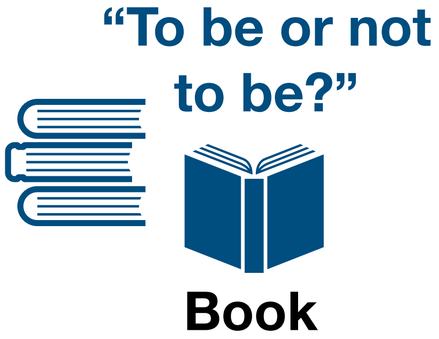
Maximal Leakage: definition, motivation, extensions, properties

Other Topics: secrecy by design, location privacy, fairness, etc.

COURSE CONTENT

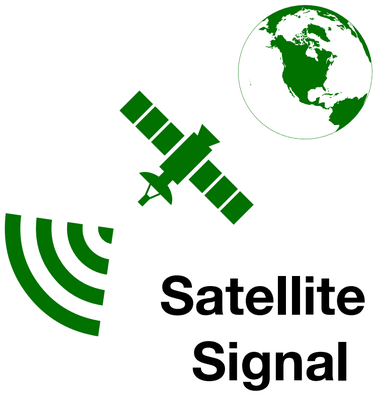
Single Information Source

How much information is there?



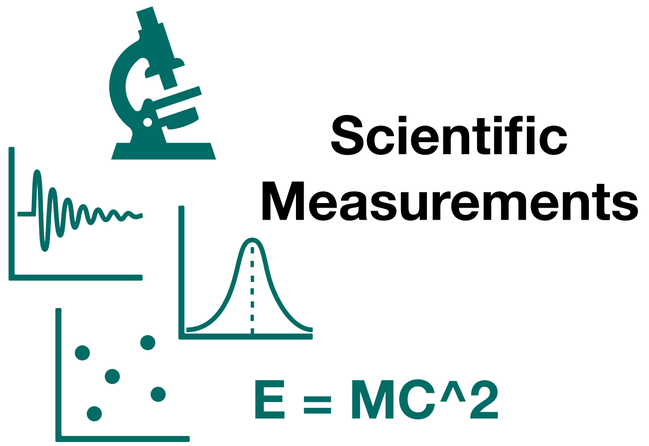
Single Information Source

How much information is there?



Single Information Source

How much information is there?



Single Information Source

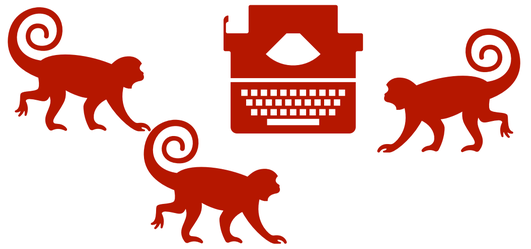
How much information is there?

Medical Data

Single Information Source

How much information is there?

#H*kehtfw20e;sp0s0gsl

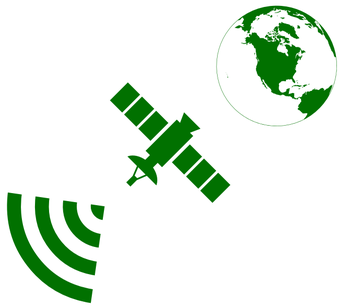


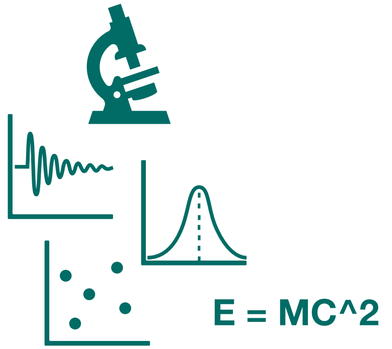
Random Noise

Single Information Source

How much information is there?

- Use tools from probability theory and statistics to model information
- A ‘source of information’ is a random variable
- The ‘amount’ of information measured depends…

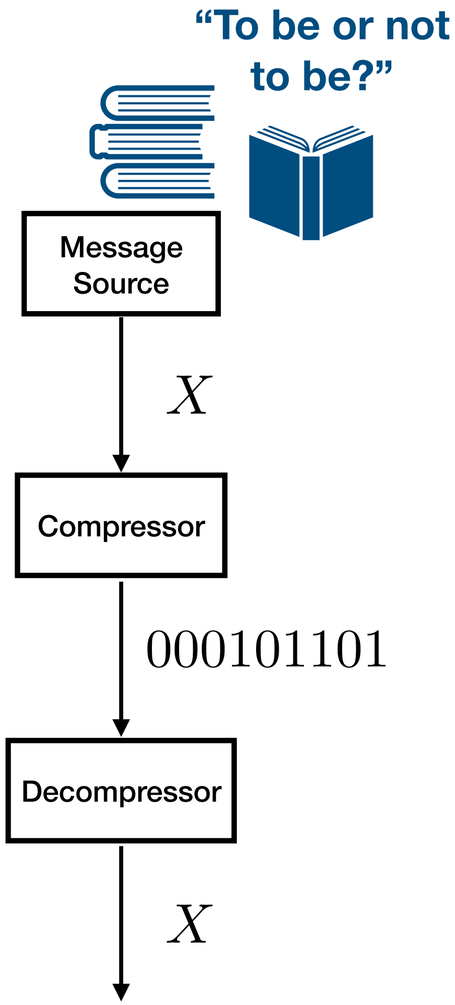




Single Information Source

Example: Lossless Data Compression

- Q: What is the smallest number of bits we need to represent the source?
- A: There is a fundamental lower bound called “entropy” (or Shannon entropy)
- Many practical algorithms exist that come very close to achieving this lower bound



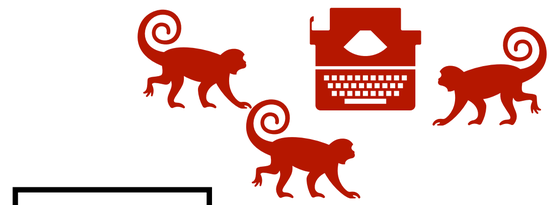
$$H(X) = \sum_{x \in \mathcal{X}} P_X(x) \log \frac{1}{P_X(x)}$$

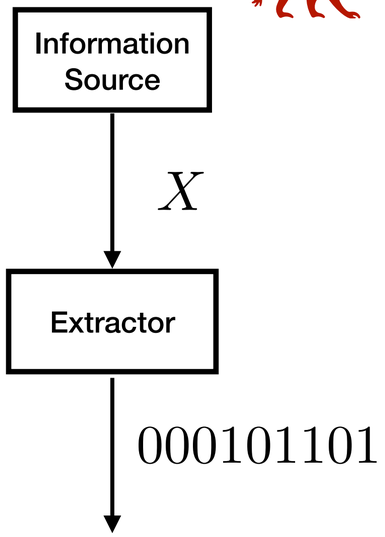
Single Information Source

Example: Randomness Extraction

- Q: What is the largest number of uniform bits we could get out of this information source?
- A: There is a fundamental upper bound called “min-entropy”
- Note: Min-entropy is a special case of Réniy entropy

#H*kehtfw20e;sp0s0gsl

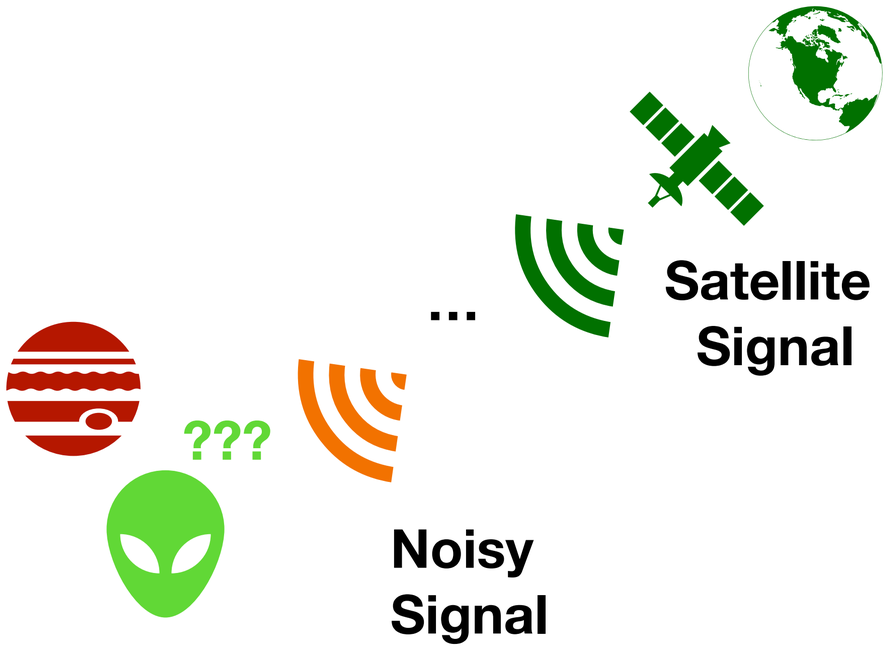




$$H_{\min}(X) = \min_{x \in \mathcal{X}} \log \frac{1}{P_X(x)}$$

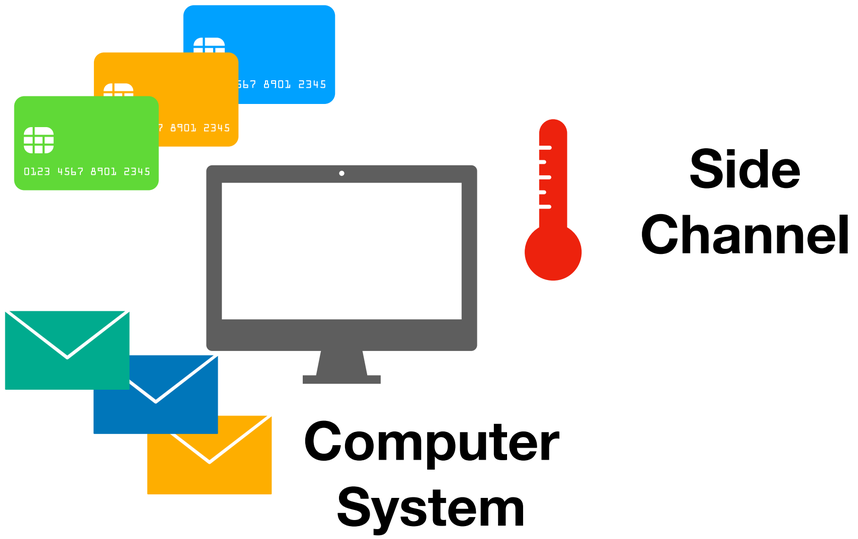
Two Information Sources

How much relevant information is there in a related observation?



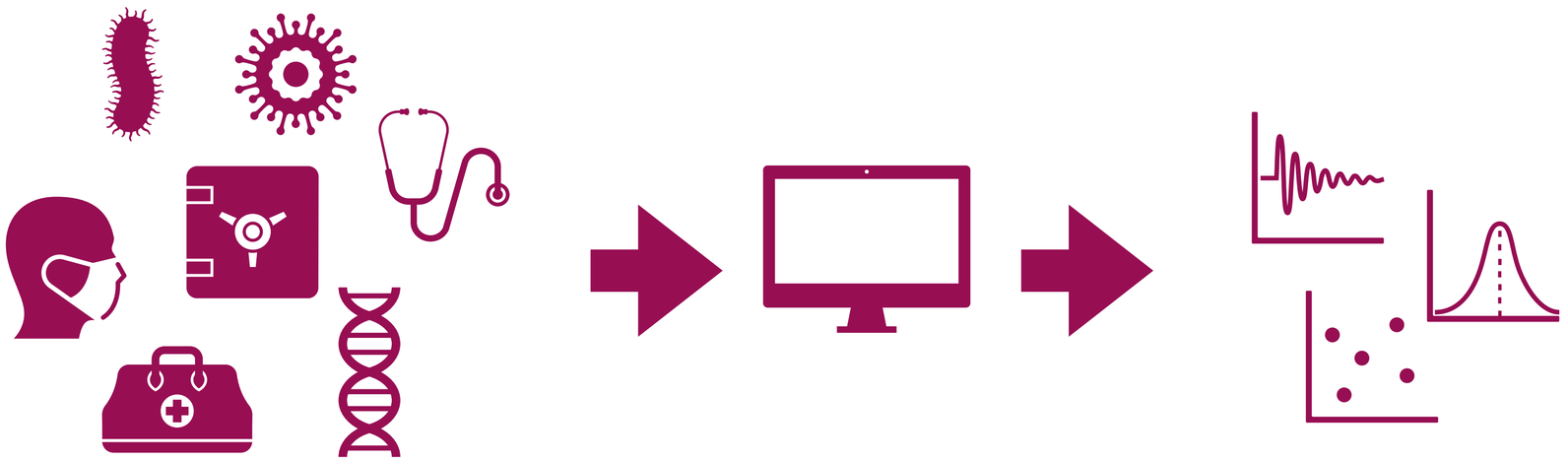
Two Information Sources

How much relevant information is there in a related observation?



Two Information Sources

How much relevant information is there in a related observation?

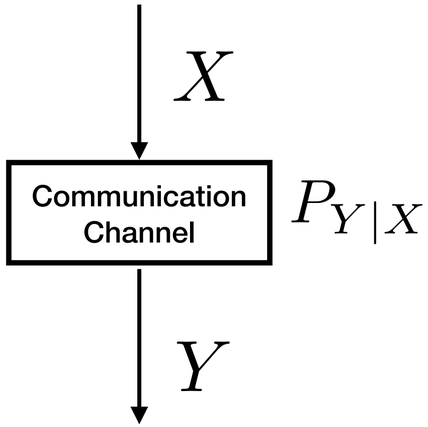


Medical Data **Data Statistics**

Two Information Sources

Example: Data Transmission

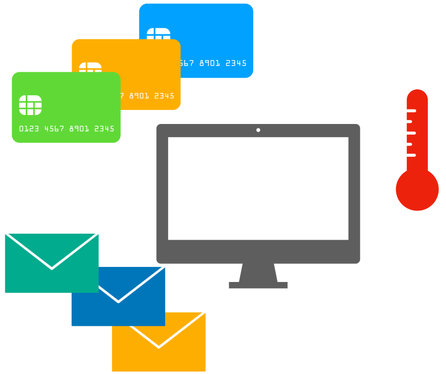
- Q: What is the largest number of messages we could reliably transmit through this communication channel
- A: The rate of transmission is the maximal mutual information across the channel
- Practical error correcting codes (e.g. LDPC, Polar codes) that achieve this



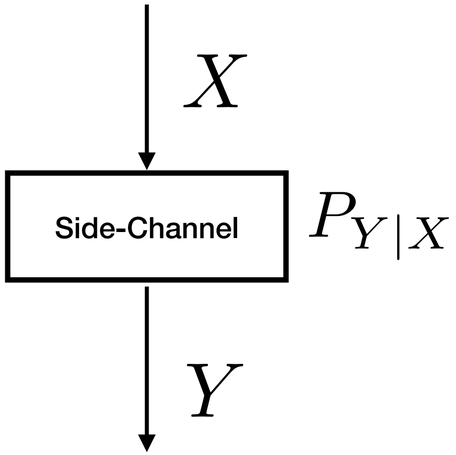
$$C = \max_{P_X} I(X;Y)$$

Two Information Sources

Example: Side-Channel Leakage



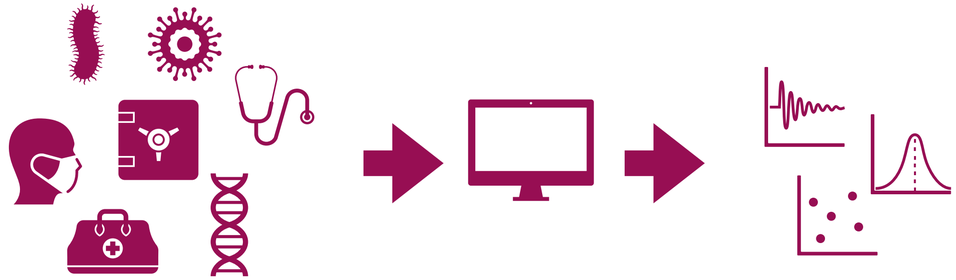
- Q: How do we measure how much information is leaked by a side channel?
- A: Measure how much better the adversary can compute functions of data
- This is the approach taken by maximal leakage



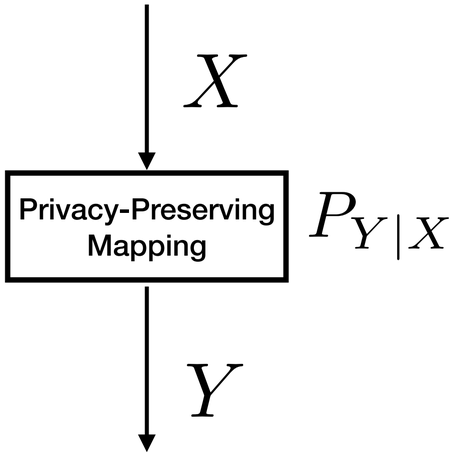
$$\mathcal{L}(X \to Y) = \sup_{U - X - Y - \hat{U}} \log \frac{\mathbb{P}[U = \hat{U}]}{\max_{u \in \mathcal{U}} P_U(u)}$$

Two Information Sources

Example: Database Privacy



- Q: How do we measure how much privacy is compromised (e.g. private information leaked) in a database query?
- A: Measure how much individual record changes the result of the query
- This is the approach taken by differential privacy



$$\mathcal{DP}(X \to Y) = \sup_{x_1, x_2 \in \mathcal{X}:\ d(x_1,x_2)=1, \mathcal{T} \subset \mathcal{Y}} \log \frac{\mathbb{P}[Y \in \mathcal{T} | X = x_1]}{\mathbb{P}[Y \in \mathcal{T} | X = x_2]}$$

Two Information Sources

Common Information

- There are many other perfectly reasonable ways to define the notion of ‘relevant’ information between two sources
- Example: Gács-Körner-Witsenhausen common information
- Example: Wyner common information

$$K(X;Y) = \sup_{W\,:\,W=g_1(X)=g_2(Y)} H(W)$$

$$0 \leq K(X;Y) \leq I(X;Y) \leq J(X;Y) \leq H(X,Y)$$

$$J(X;Y) = \inf_{W\,:\,X-W-Y} I(X,Y;W)$$

Information Sources

Rényi Entropy, alpha-Mutual Information

- There are also many perfectly reasonable ways to generalise Shannon entropy and Mutual Information
- Rényi entropy is one well-known one
- It has nice properties and axiomatic justifications (stay tuned for next lecture)

ON MEASURES OF ENTROPY AND INFORMATION

ALFRÉD RÉNYI
MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES

1. Characterization of Shannon's measure of entropy

Let $\mathcal{P} = (p_1, p_2, \cdots, p_n)$ be a finite discrete probability distribution, that is, suppose $p_k \geqq 0 (k = 1, 2, \cdots, n)$ and $\sum_{k=1}^{n} p_k = 1$. The amount of uncertainty of the distribution $\mathcal{P}$, that is, the amount of uncertainty concerning the outcome of an experiment, the possible results of which have the probabilities $p_1, p_2, \cdots, p_n$, is called the *entropy* of the distribution $\mathcal{P}$ and is usually measured by the quantity $H[\mathcal{P}] = H(p_1, p_2, \cdots, p_n)$, introduced by Shannon [1] and defined by

$$H(p_1, p_2, \cdots, p_n) = \sum_{k=1}^{n} p_k \log_2 \frac{1}{p_k}. \tag{1.1}$$

Different sets of postulates have been given, which characterize the quantity (1.1). The simplest such set of postulates is that given by Fadeev [2] (see also Feinstein [3]). Fadeev's postulates are as follows.

(a) $H(p_1, p_2, \cdots, p_n)$ *is a symmetric function of its variables for* $n = 2, 3, \cdots$.
(b) $H(p, 1-p)$ *is a continuous function of* p *for* $0 \leqq p \leqq 1$.
(c) $H(1/2, 1/2) = 1$.
(d) $H[tp_1, (1-t)p_1, p_2, \cdots, p_n] = H(p_1, p_2, \cdots, p_n) + p_1 H(t, 1-t)$

for any distribution $\mathcal{P} = (p_1, p_2, \cdots, p_n)$ *and for* $0 \leqq t \leqq 1$.

The proof that the postulates (a), (b), (c), and (d) characterize the quantity (1.1) uniquely is easy except for the following lemma, whose proofs up to now are rather intricate.

LEMMA. *Let* $f(n)$ *be an additive number-theoretical function, that is, let* $f(n)$ *be defined for* $n = 1, 2, \cdots$ *and suppose*

$$f(nm) = f(n) + f(m), \qquad n, m = 1, 2, \cdots. \tag{1.2}$$

Let us suppose further that

$$\lim_{n \to +\infty} [f(n+1) - f(n)] = 0. \tag{1.3}$$

Then we have

$$f(n) = c \log n, \tag{1.4}$$

where c *is a constant.*