Information Integrity

Eric Graves
Advisor: Dr. Tan Wong
University of Florida

Center for Science of Information
NSF Science and Technology Center

Introduction

- Alice and Bob want to communicate through an untrusted and possibly Byzantine relay, Romeo. Thus our goal is the following:
  - Detect and if possible correct any malicious activity of Romeo
  - Minimize redundancy as not to incur any penalty if Romeo is not malicious
  - Design for all possible attacks by Romeo, making no assumptions about type of attack
- Additionally we feel it is important to investigate the following:
  - Relationship between secrecy and integrity
  - How successful modification of symbols by Romeo changes what Romeo can learn about the transmitted sequences
  - When can the channel alone provide integrity versus when coding is required

Manipulability, Information and Magic

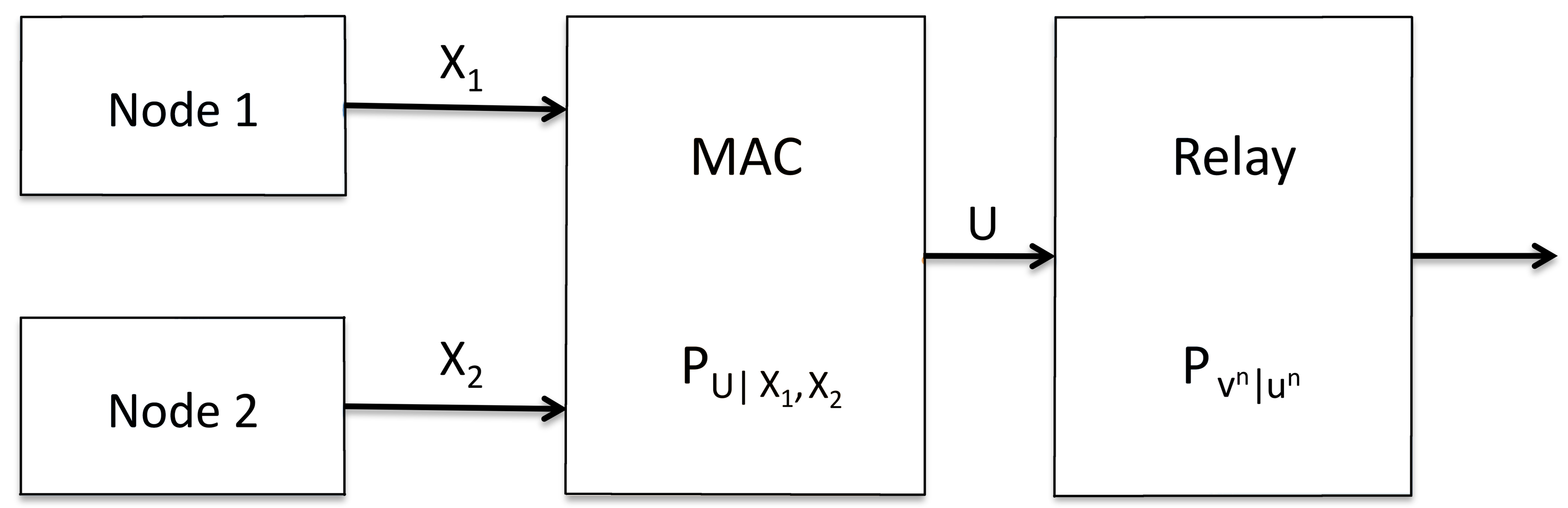
- A $N \times M$ stochastic matrix A is manipulable if there exists a $N \times N$ stochastic matrix Φ such that
$$\Phi A = A.$$
- Slap-Happy from [1, problem 2.7]
$$\left| T^n_{[X|U]}(u^n) \bigcap T^n_{[X|U]}(v^n) \right| \doteq \max_{\substack{P_{\tilde{X}|\tilde{U}} = P_{X|U} \\ P_{\tilde{X}|\tilde{V}} = P_{X|U} \\ P_{\tilde{U},\tilde{V}} = P_{u^n,v^n}}} 2^{nH(\tilde{X}|\tilde{U},\tilde{V})}$$
- For any x^n chosen uniformly random over $T^n_{[X|U]}(u^n)$, then
$$\Pr\left(x^n \in T^n_{[X|U]}(u^n) \bigcap T^n_{[X|U]}(v^n) \middle| x^n \in T^n_{[X|U]}(u^n) \right) \doteq \max 2^{-nI(\tilde{X};\tilde{V}|\tilde{U})}$$
- If the matrix defined by $A_{i,j} = P_{U|X}(\mathrm{u}_i|\mathrm{x}_j)$ is not manipulable, then
$$\min I(\tilde{X};\tilde{V}|\tilde{U}) = 0 \;\rightarrow\; v^n = u^n.$$

Integrity Types

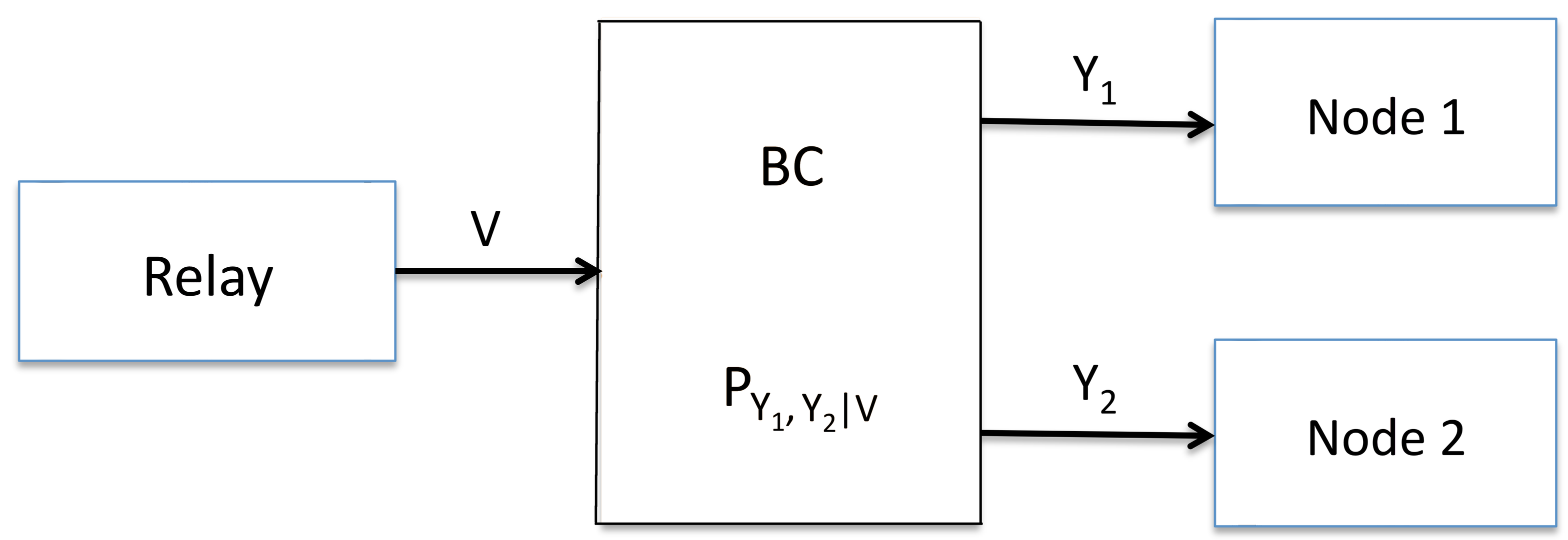
- **Strong Integrity**
  - Channel provides integrity
  - Decode *and* Detect
  - For any v^n such that $\|v^n - u^n\| > \epsilon$
$$\Pr\left(v^n \in T^n_{[U|X]}(x^n) | u^n \in T^n_{[U|X]}(x^n) \right) < 2^{-n\epsilon}.$$
- **Weak Integrity**
  - Coding provides integrity
  - Decode *or* Detect
  - For any v^n such that $\|v^n - u^n\| > \epsilon$
$$\Pr\left(\exists x_2^n \in \mathbf{c}_2^n : v^n \in T^n_{[U|X_1,X_2]}(x_1^n, x_2^n) | \exists \hat{x}_2^n \in \mathbf{c}_2^n : u^n \in T^n_{[U|X_1,X_2]}(x_1^n, \hat{x}_2^n) \right) < 2^{-n\epsilon}.$$

Diagram

- $t = 1, \ldots, n$



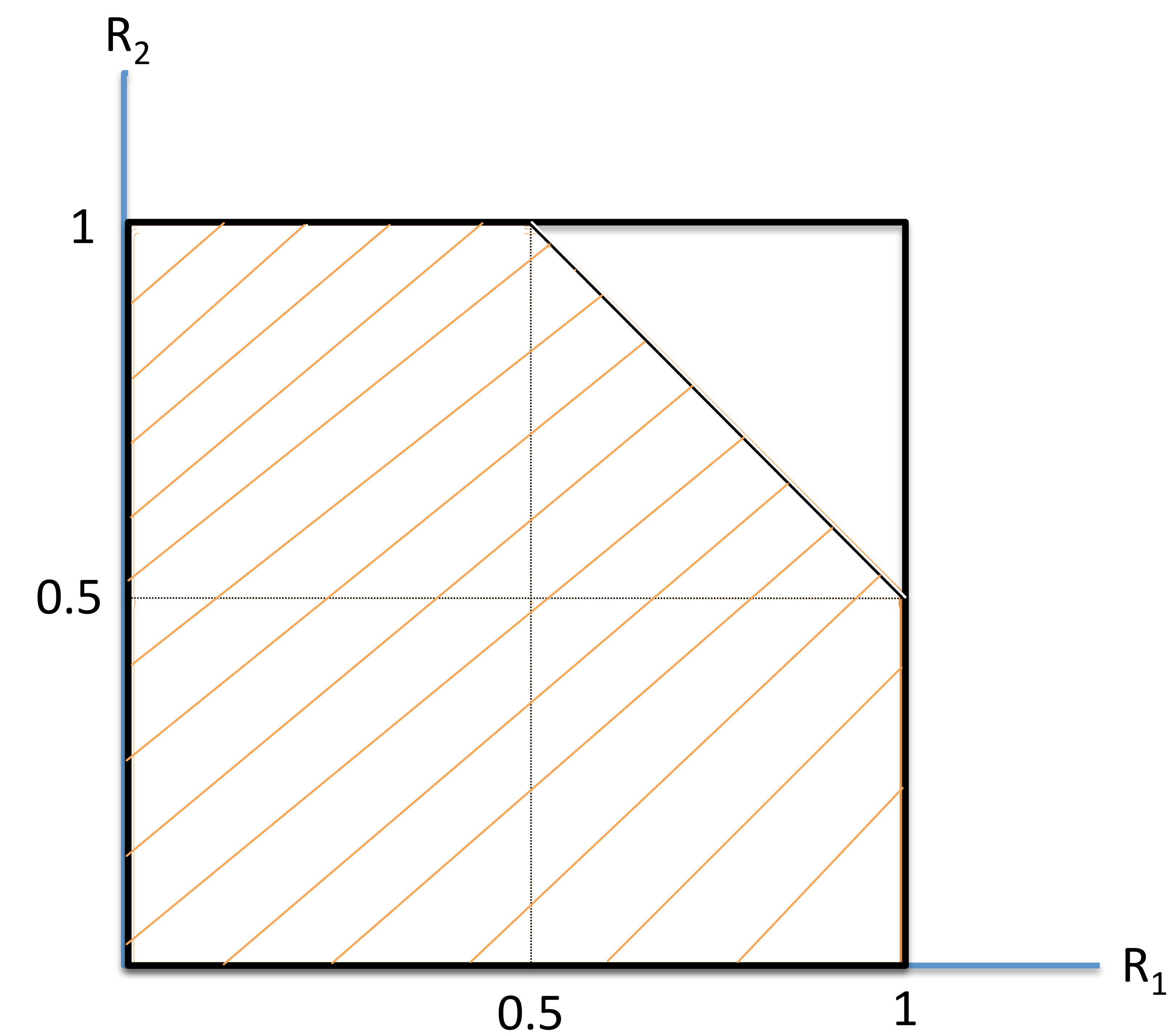
- $t = n+1, \ldots, 2n$



Extensions

- Can be extended to different channels, such as:
  - One-Way Relay
  - Multiple (possibly Byzantine, but uncooperative) Relays
- Secrecy
  - Integrity is not obtained through secrecy but relies upon same basic assumptions of secrecy, i.e., their exist multiple possible transmitted sequences (x^n) for every received sequence (u^n)
  - The random coding practices are easily extended to cover secrecy
  - For secret key exchange, the techniques easily show that a large portion of the keys have the same probability of occurrence, which is strictest definition of secrecy possible
  - Does not use hash functions
- Authentication
  - This work on integrity can be viewed as the analog to authentication first introduced by Maurer [2]
  - Same techniques that can be used to prove integrity can be used on authentication, but allow it to be done with random coding as opposed to hash functions
- Continuous Distributions
  - For finite, or polish, alphabets the method to prove integrity easily extends to continuous distributions

Binary Erasure MAC Example



The capacity region with guaranteed information integrity is the closure of the convex hull of all (R_1, R_2) satisfying
$$R_1 < I(X_1; U|X_2)$$
$$R_2 < I(X_2; U|X_1)$$
for some $P_{U,X_1,X_2}(\mathrm{u}, \mathrm{x}_1, \mathrm{x}_2) = P_{U|X_1,X_2}(\mathrm{u}|\mathrm{x}_1, \mathrm{x}_2) P_{X_1}(\mathrm{x}_1) P_{X_2}(\mathrm{x}_2)$.

Bullet-points!

- Can either detect or correct any manipulation
- Capacity region (when relay is not being malicious) is equal to that of a trusted relay
- Does not require any special code, random coding alone sufficient for proof
- Can be generalized to fit more complicated channel models

References

- Csiszár, I., Körner, J.: Information Theory: Coding Theorems for Discrete Memoryless Systems, 2nd edn. Cambridge University Press (2011)
- Maurer, U.: Information-theoretically secure secret-key agreement by not authenticated public discussion. In: W. Fumy (ed.) Advances in Cryptology — EUROCRYPT '97, *Lecture Notes in Computer Science*, vol. 1233, pp. 209–225. Springer-Verlag (1997)