Formal Methods in Security

Prakash Panangaden
School of Computer Science
McGill University

Summary

Summary

- The role of formal methods

Summary

- The role of formal methods
- Probabilistic reasoning

Summary

- The role of formal methods
- Probabilistic reasoning
- Channel capacity as a measure of anonymity

Summary

- The role of formal methods
- Probabilistic reasoning
- Channel capacity as a measure of anonymity
- Games, Capacities and Previsions

Summary

- The role of formal methods
- Probabilistic reasoning
- Channel capacity as a measure of anonymity
- Games, Capacities and Previsions
- Conclusions

The role of formal methods 1

The role of formal methods 1

- Cryptography is assumed unbreakable.

The role of formal methods 1

- Cryptography is assumed unbreakable.
- Attackers have access to every message and can synthesize messages.

The role of formal methods 1

- Cryptography is assumed unbreakable.
- Attackers have access to every message and can synthesize messages.
- They can perform statistical analysis of intercepted messages.

The role of formal methods 1

- Cryptography is assumed unbreakable.
- Attackers have access to every message and can synthesize messages.
- They can perform statistical analysis of intercepted messages.
- What can be done to preserve secrecy or anonymity?

Formal methods 2

Formal methods 2

- We model agents - including attackers - as processes in some formal system and

Formal methods 2

- We model agents - including attackers - as processes in some formal system and
- use tools like model checkers, bisimulation checkers to verify properties of the protocol.

Formal methods 2

- We model agents - including attackers - as processes in some formal system and
- use tools like model checkers, bisimulation checkers to verify properties of the protocol.
- The models may be probabilistic.

Formal methods 2

- We model agents - including attackers - as processes in some formal system and
- use tools like model checkers, bisimulation checkers to verify properties of the protocol.
- The models may be probabilistic.
- Legendary success: Gavin Lowe and the Needham-Schroeder protocol.

Probability

Probability

- Using non probabilistic models does not allow one to analyze situations where the attacker uses statistical techniques to extract information.

Probability

- Using non probabilistic models does not allow one to analyze situations where the attacker uses statistical techniques to extract information.
- Probabilistic process algebra and metrics were used by John Mitchell et al.

Probability

- Using non probabilistic models does not allow one to analyze situations where the attacker uses statistical techniques to extract information.
- Probabilistic process algebra and metrics were used by John Mitchell et al.
- Anonymity protocols analyzed by Palamidessi et al.

Probability

- Using non probabilistic models does not allow one to analyze situations where the attacker uses statistical techniques to extract information.
- Probabilistic process algebra and metrics were used by John Mitchell et al.
- Anonymity protocols analyzed by Palamidessi et al.
- Probabilistic model checking developed by Kwiatkowska et al.; the PRISM system.

What is anonymity?

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.
- Important in:

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.
- Important in:
  - Electronic elections

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.
- Important in:
  - Electronic elections
  - Posting to bulletin boards

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.
- Important in:
  - Electronic elections
  - Posting to bulletin boards
  - File sharing, refereeing (!), ...

What is anonymity?

- We want to ensure that the identity of an agent performing some actions remains secret; the action itself can be visible.
- Important in:
  - Electronic elections
  - Posting to bulletin boards
  - File sharing, refereeing (!), ...
- In some sense “dual” to secrecy.

Example Systems

Example Systems

- Crowds [Reiter and Rubin 1998]: initiator is anonymous

Example Systems

- Crowds [Reiter and Rubin 1998]: initiator is anonymous
- Onion Routing [Syverson, Goldschlag and Reed 1997]: anonymous communication

Example Systems

- Crowds [Reiter and Rubin 1998]: initiator is anonymous
- Onion Routing [Syverson, Goldschlag and Reed 1997]: anonymous communication
- Freenet [Clarke et. al. 2001]: anonymous information retreival

Nondeterministic or Probabilistic?

Nondeterministic or Probabilistic?

- Nondeterministic analysis can use the machinery of concurrency theory, but it does not allow one to reason about adversaries that make repeated observations and make statistical inferences

Nondeterministic or Probabilistic?

- Nondeterministic analysis can use the machinery of concurrency theory, but it does not allow one to reason about adversaries that make repeated observations and make statistical inferences
- The probabilistic approach is essential when the protocols themselves use randomization

Nondeterministic or Probabilistic?

- Nondeterministic analysis can use the machinery of concurrency theory, but it does not allow one to reason about adversaries that make repeated observations and make statistical inferences
- The probabilistic approach is essential when the protocols themselves use randomization
- However, usually both probability and nondeterminism is present.

Levels of Anonymity: Needs Probability

Levels of Anonymity: Needs Probability

- Beyond suspicion: to the observer, the culprit is not more likely than any other agent to be the culprit.

Levels of Anonymity: Needs Probability

- Beyond suspicion: to the observer, the culprit is not more likely than any other agent to be the culprit.
- Probable innocence: the culprit has less than 50% chance of being the culprit.

Levels of Anonymity: Needs Probability

- Beyond suspicion: to the observer, the culprit is not more likely than any other agent to be the culprit.
- Probable innocence: the culprit has less than 50% chance of being the culprit.
- Possible innocence: the culprit has less than 100% chance of being the culprit.

Dining Cryptographers: Chaum 1988

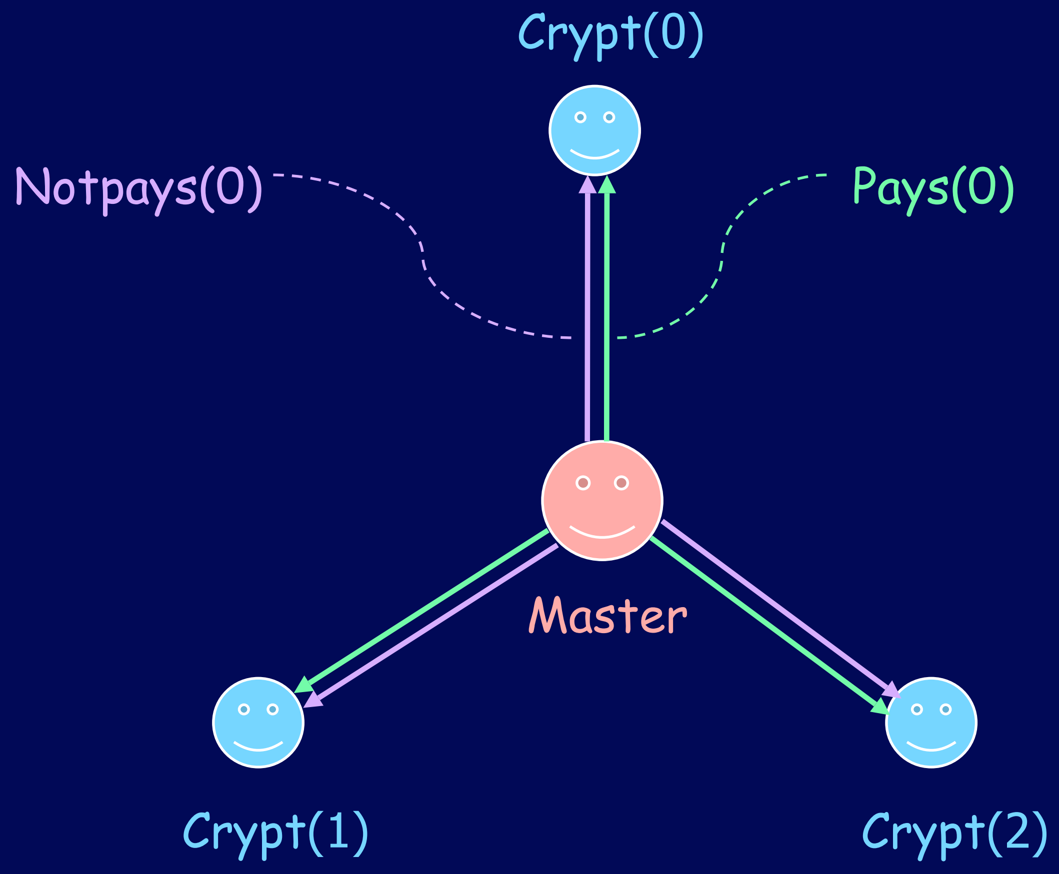
Dining Cryptographers: Chaum 1988

- The problem:
  - Three cryptographers share a meal
  - The meal is either paid by M or by one of the diners, M decides who will pay
  - M informs each one whether they will pay or not

Dining Cryptographers: Chaum 1988

- The problem:
  - Three cryptographers share a meal
  - The meal is either paid by M or by one of the diners, M decides who will pay
  - M informs each one whether they will pay or not
- The goal: the cryptographers want to find out if one of them is paying without knowing who.

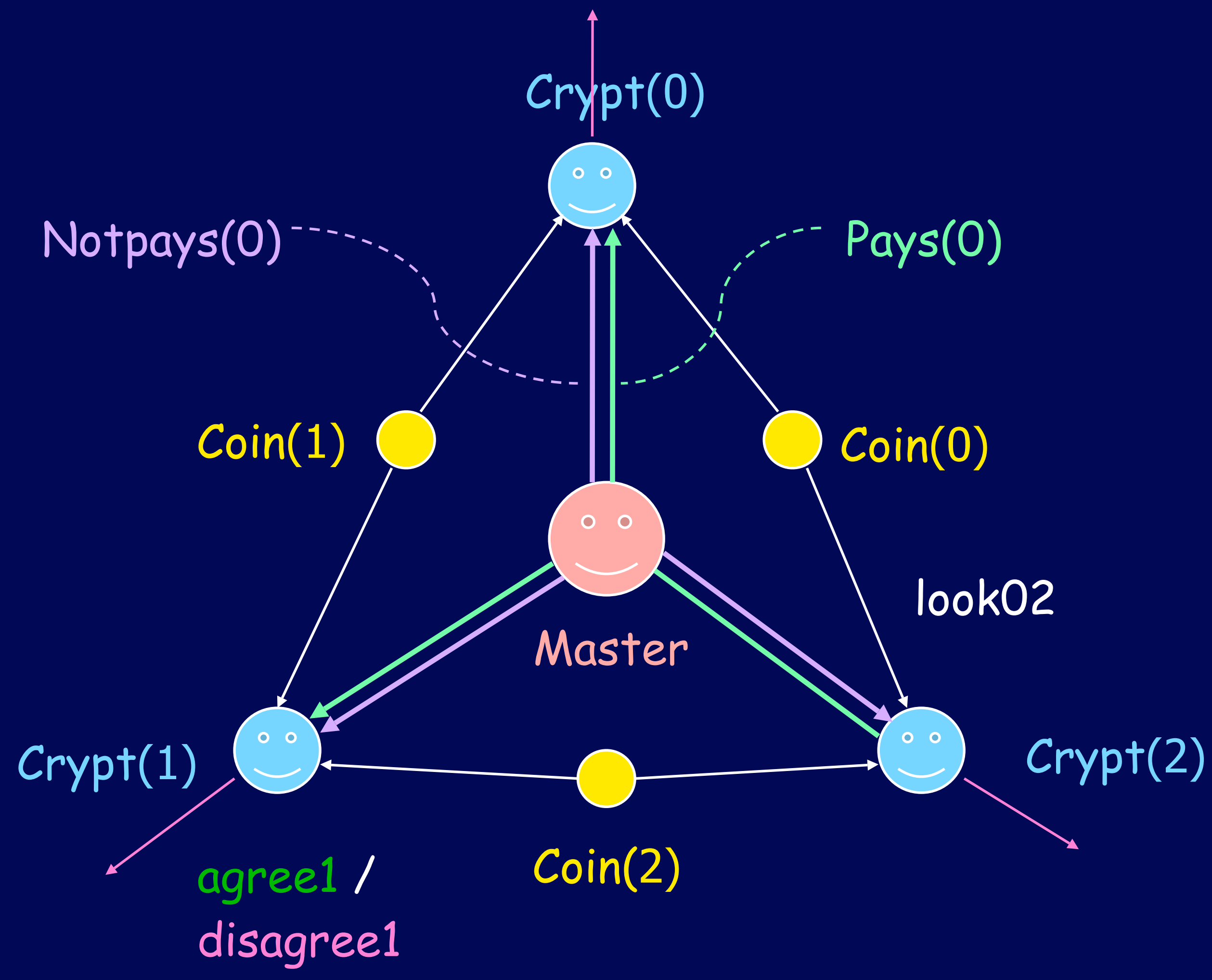
The dining cryptographers



Solution

- We insert a coin between each pair of cryptographers and toss it
- The result of each coin toss is visible only to the adjacent cryptographers
- Each cryptographer examines the two adjacent coins and says "agree" or "disagree"
- The one who pays (if any) will say the opposite of the truth.

The dining cryptographers



Does it work?

Does it work?

- The number saying “disagree” is even if and only if M is paying. (This works for arbitrary graphs.)

Does it work?

- The number saying "disagree" is even if and only if M is paying. (This works for arbitrary graphs.)
- If the coins are fair then an external observer and the non-paying cryptographers will not be able to deduce who is paying.

Does it work?

- The number saying "disagree" is even if and only if M is paying. (This works for arbitrary graphs.)
- If the coins are fair then an external observer and the non-paying cryptographers will not be able to deduce who is paying.
- In fact they will not even be able to increase their probabilistic estimates.

Biased Coins?

Biased Coins?

- In extreme cases it is easy to see that a statistical analysis of the outcomes will allow one to guess which way the coins are biased and thus who is paying.

Biased Coins?

- In extreme cases it is easy to see that a statistical analysis of the outcomes will allow one to guess which way the coins are biased and thus who is paying.
- This is not detected by the purely nondeterministic approaches.

Biased Coins?

- In extreme cases it is easy to see that a statistical analysis of the outcomes will allow one to guess which way the coins are biased and thus who is paying.
- This is not detected by the purely nondeterministic approaches.
- In less extreme cases of bias the situation is harder to analyze but clearly some information can leak out.

Coin 12 and Coin 13 are H, Coin 23 is T

M chooses the payer uniformly at random.

	1 pays	2 pays	3 pays
1 says	d	a	a
2 says	d	a	d
3 says	d	d	a

We never see 1 saying d while 2 and 3 say a.

If we say "almost never" then the nondeterministic approach will say this is fine!

Information Theory Summarized

X, Y are random variables and x, y represent possible values.

Entropy: $H(X) = -\sum_x p(x) \log p(x)$
Uncertainty in X.

Conditional Entropy: $H(X|Y) = -\sum_y p(y)[\sum_x p(x|y) \log p(x|y)]$
Uncertainty in X when Y is known.

Mutual Information: $I(X;Y) = H(X) - H(X|Y)$
What Y reveals about X and vice versa.

Channel Capacity

A channel is just a triple

$$(\mathcal{X}, \mathcal{Y}, p(\cdot|\cdot))$$

where $\mathcal{X}$ is the set of input symbols, $\mathcal{Y}$ is the set of output symbols and $p(y|x)$ is the probability of observing y if x is input.

Given an input distribution $p(x)$ we can define random variables X and Y.

The **channel capacity** is given by

$$C = \max_{p(x)} I(X;Y).$$

What use is Channel Capacity?

What use is Channel Capacity?

- Channel capacity measures the propensity of a system to leak information.

What use is Channel Capacity?

- Channel capacity measures the propensity of a system to leak information.
- Usually we try to increase the channel capacity, but here

What use is Channel Capacity?

- Channel capacity measures the propensity of a system to leak information.
- Usually we try to increase the channel capacity, but here
- we want the channel capacity to be as low as possible.

Capacity of What?

Capacity of What?

- Ira Moskowitz et. al. studied the capacity of a covert channel to measure how much information could be leaked out of a system by an agent with access to a covert channel.

Capacity of What?

- Ira Moskowitz et. al. studied the capacity of a covert channel to measure how much information could be leaked out of a system by an agent with access to a covert channel.
- We are viewing the protocol itself as an abstract channel and thus adopting channel capacity as a quantitative measure of anonymity.

Sanity Check

Sanity Check

- To what does capacity 0 correspond?

Sanity Check

- To what does capacity 0 correspond?
- It corresponds precisely to strong anonymity, i.e. to the statement that A and O are independent.

Other Things

Other Things

- Palamidessi's group has modelled the DC protocol in the PRISM language and shown how to compute the capacity.

Other Things

- Palamidessi's group has modelled the DC protocol in the PRISM language and shown how to compute the capacity.
- One can consider the theory of hypothesis testing and analyze attacks made using Bayesian decision rules. We have bounds on the probability of error. This has been greatly extended in a new paper which uses some ideas from convexity theory to give new bounds.

Games, capacities and previsions 1

Games, capacities and previsions 1

- The right way to understand the interactions of adversaries is to model them as games.

Games, capacities and previsions 1

- The right way to understand the interactions of adversaries is to model them as games.
- This causes an interaction between probability and nondeterministic choices.

Games, capacities and previsions 1

- The right way to understand the interactions of adversaries is to model them as games.
- This causes an interaction between probability and nondeterministic choices.
- One has capacities rather than measures. Used in economics and in concurrency theory by Gupta, Jagadeesan, Desharnais and Panangaden.

Games 2

Games 2

- Far reaching generalization and development of these ideas by Jean Goubault-Larrecq

Games 2

- Far reaching generalization and development of these ideas by Jean Goubault-Larrecq
- He has a 641 page document (in French)!!

Games 2

- Far reaching generalization and development of these ideas by Jean Goubault-Larrecq
- He has a 641 page document (in French)!!
- Related work by Mislove, Keimel, Plotkin and Tix.

Games 2

- Far reaching generalization and development of these ideas by Jean Goubault-Larrecq
- He has a 641 page document (in French)!!
- Related work by Mislove, Keimel, Plotkin and Tix.
- The theory is ready to be used.

Conclusions

Conclusions

- Information theory is a rich and powerful way to analyze probabilistic protocols.

Conclusions

- Information theory is a rich and powerful way to analyze probabilistic protocols.
- The theory of games and capacities needs to be combined with information theory.

Conclusions

- Information theory is a rich and powerful way to analyze probabilistic protocols.
- The theory of games and capacities needs to be combined with information theory.
- All kinds of beautiful mathematics: convexity theory, domain theory in addition to traditional information theory.

Existing Collaborations

Existing Collaborations

- I am designated an Équipe étranger of INRIA Futur and work closely with Catuscia Palamidessi. Her part of the collaboration is supported by INRIA and mine by McGill university and to a small extent by FQRNT.

Existing Collaborations

- I am designated an Équipe étranger of INRIA Futur and work closely with Catuscia Palamidessi. Her part of the collaboration is supported by INRIA and mine by McGill university and to a small extent by FQRNT.
- Josée Desharnais and François Laviolette (U. Laval) collaborate with Jean Goubault-Larrecq. Looser ties with me, Vincent Danos and others.