An Indistinguishability-based Characterization of Anonymous Channels

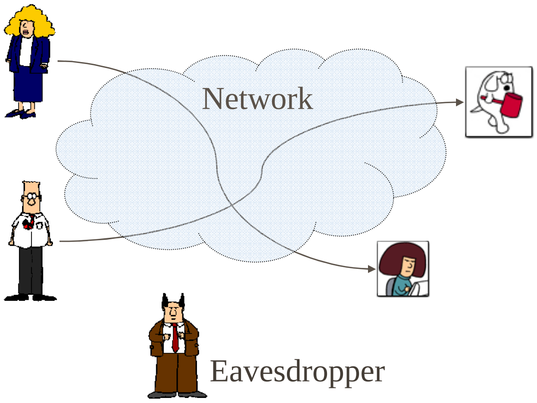
Alejandro Hevia
Dept. Computer Science,
U. of Chile, Chile

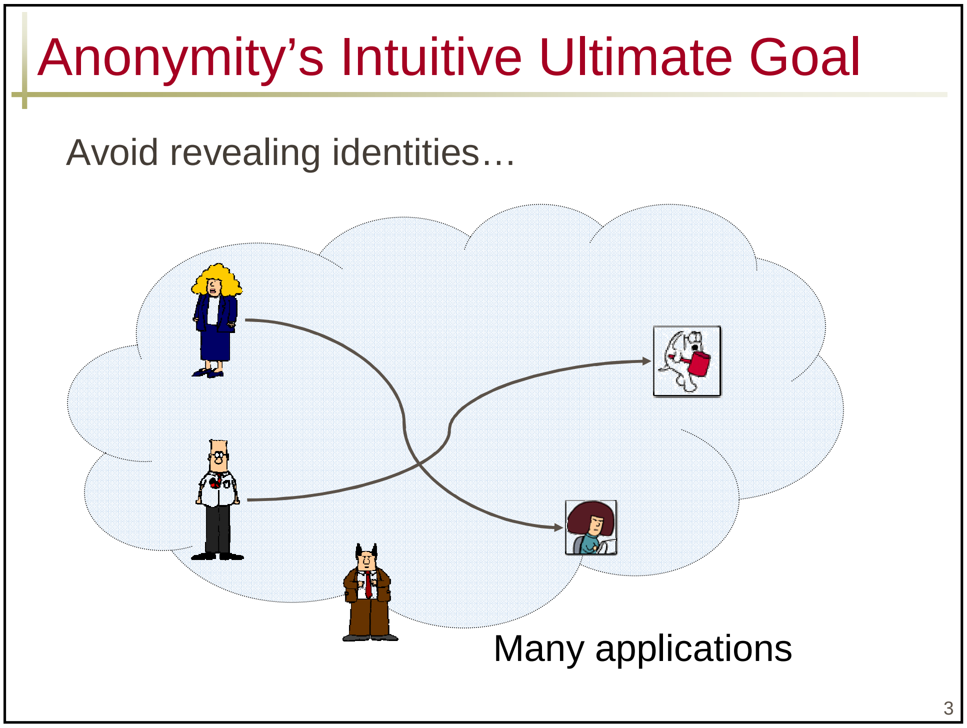
Daniele Micciancio
Dept. Computer Science &
Engineering, U. of California
at San Diego, USA

PETS'08, Leuven, Belgium, July 23, 2008

Anonymous Communication

The problem: Send/receive messages often reveals identities

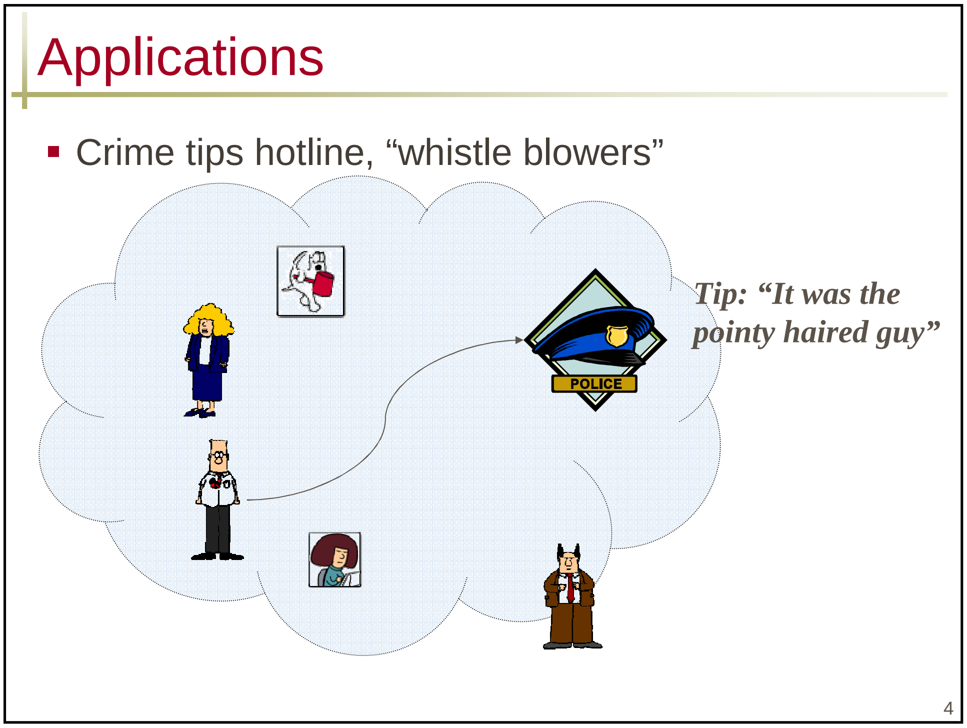




Anonymity's Intuitive Ultimate Goal

Avoid revealing identities…

Many applications



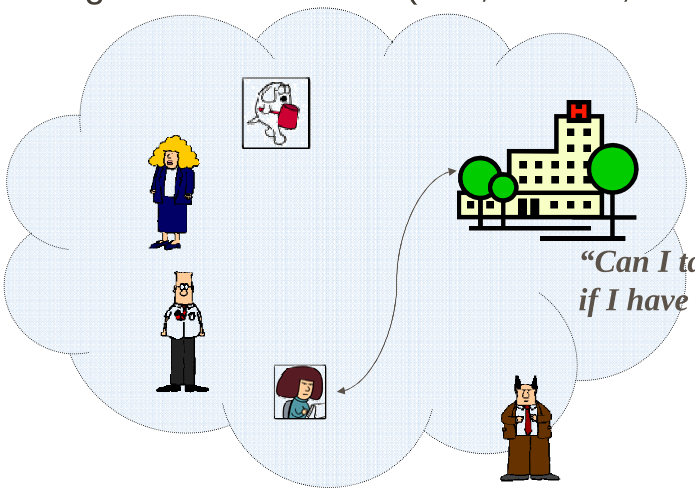
Applications

- Crime tips hotline, "whistle blowers"

Tip: "It was the pointy haired guy"

Querying disease databases

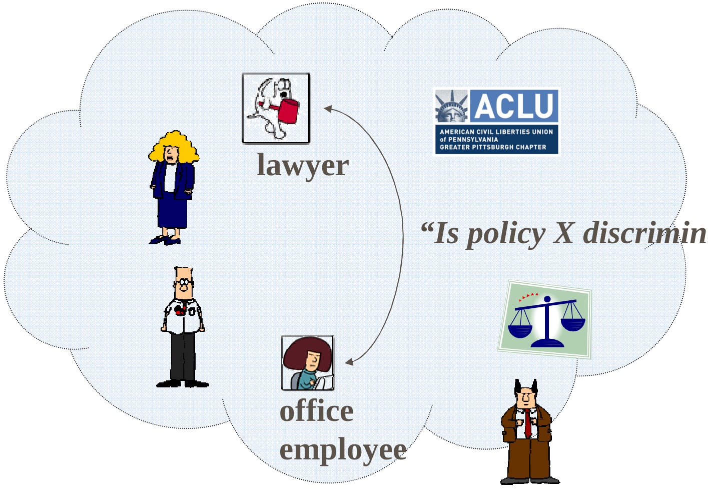
- Stigmatized diseases (HIV, Cancer, STDs)



“Can I take medicine X if I have disease Y?”

Political chat rooms

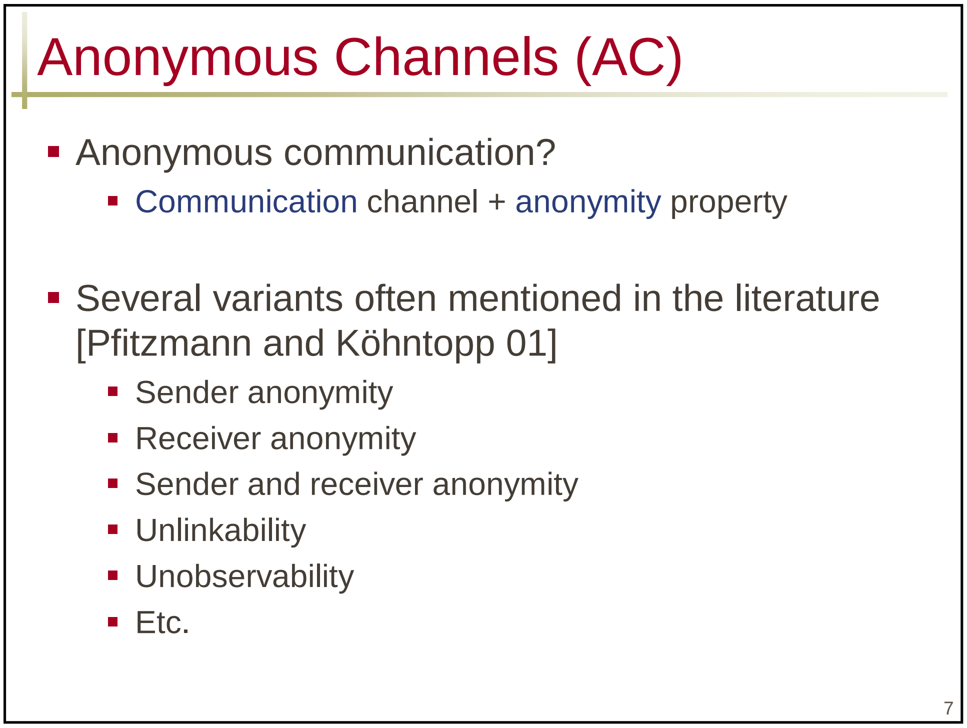
- Or “forum for unpopular/sensitive topics”



lawyer

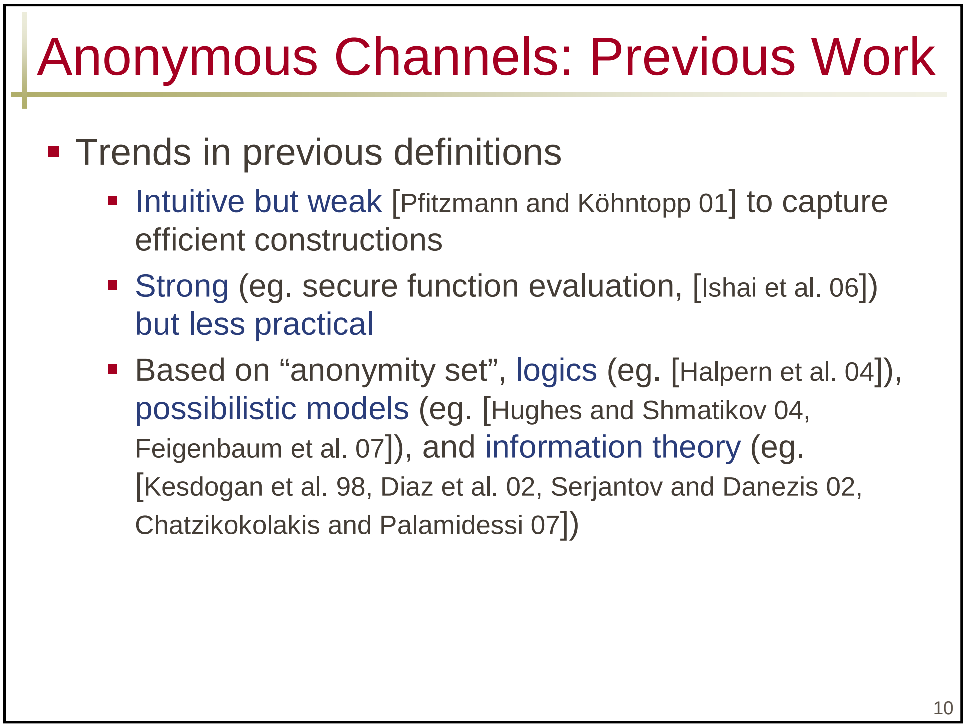
office employee

“Is policy X discriminatory?”



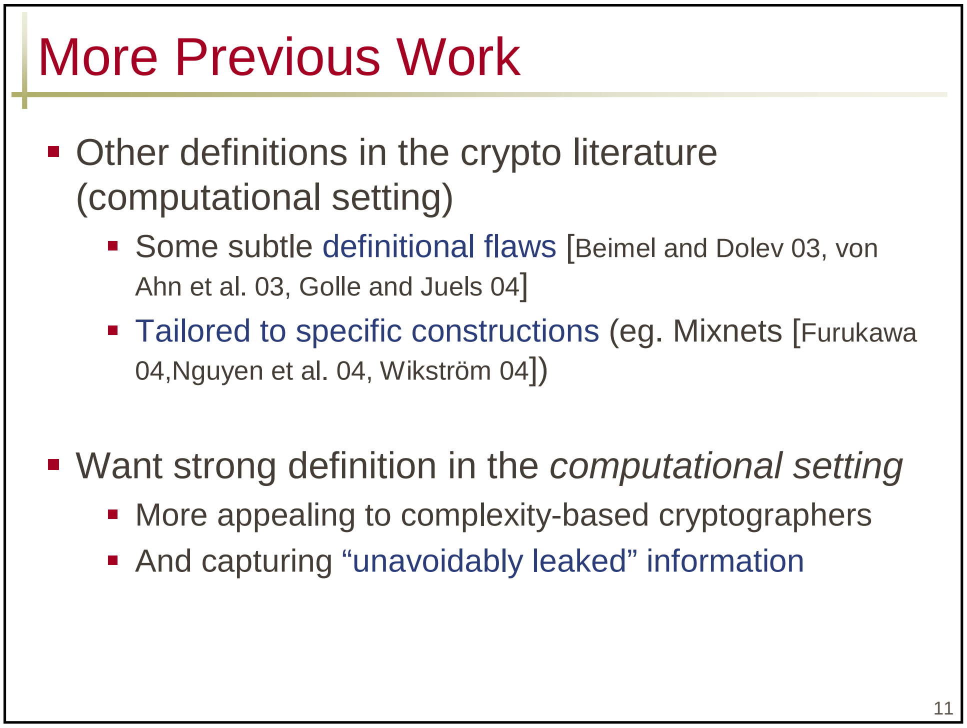
Anonymous Channels (AC)

- Anonymous communication?
  - Communication channel + anonymity property
- Several variants often mentioned in the literature [Pfitzmann and Köhntopp 01]
  - Sender anonymity
  - Receiver anonymity
  - Sender and receiver anonymity
  - Unlinkability
  - Unobservability
  - Etc.



Anonymous Channels: Previous Work

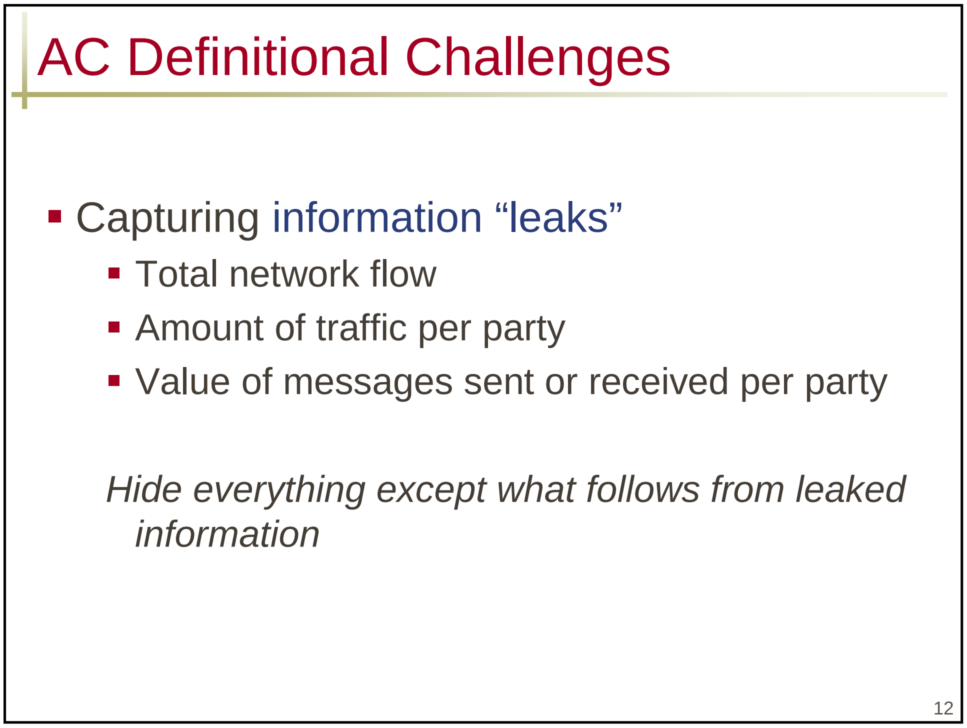
- Trends in previous definitions
  - Intuitive but weak [Pfitzmann and Köhntopp 01] to capture efficient constructions
  - Strong (eg. secure function evaluation, [Ishai et al. 06]) but less practical
  - Based on "anonymity set", logics (eg. [Halpern et al. 04]), possibilistic models (eg. [Hughes and Shmatikov 04, Feigenbaum et al. 07]), and information theory (eg. [Kesdogan et al. 98, Diaz et al. 02, Serjantov and Danezis 02, Chatzikokolakis and Palamidessi 07])



More Previous Work

- Other definitions in the crypto literature (computational setting)
  - Some subtle definitional flaws [Beimel and Dolev 03, von Ahn et al. 03, Golle and Juels 04]
  - Tailored to specific constructions (eg. Mixnets [Furukawa 04,Nguyen et al. 04, Wikström 04])

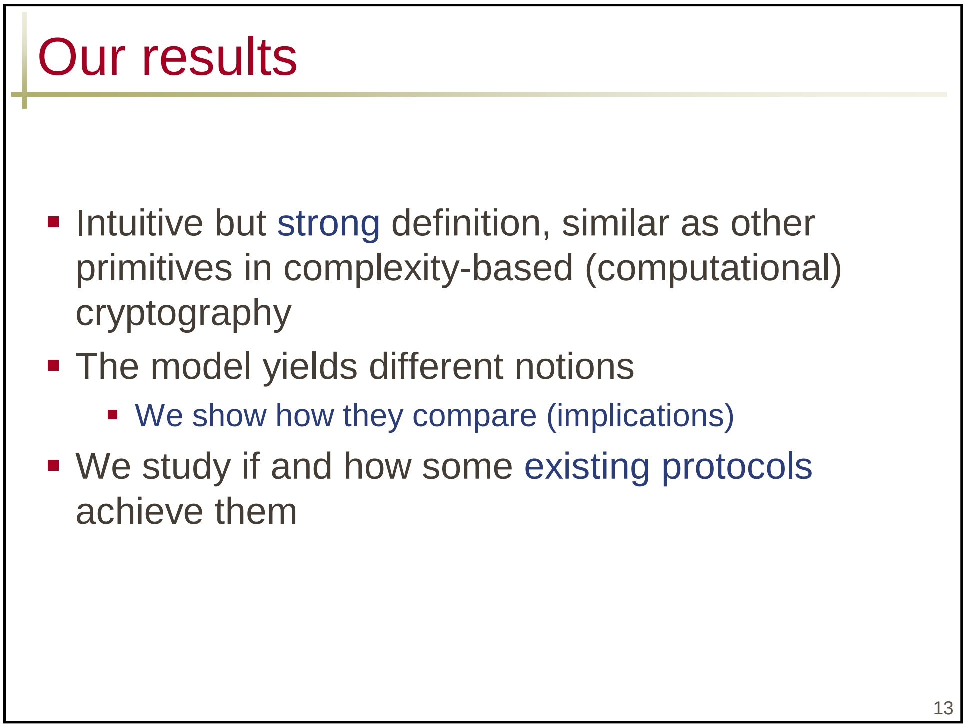
- Want strong definition in the *computational setting*
  - More appealing to complexity-based cryptographers
  - And capturing “unavoidably leaked” information



AC Definitional Challenges

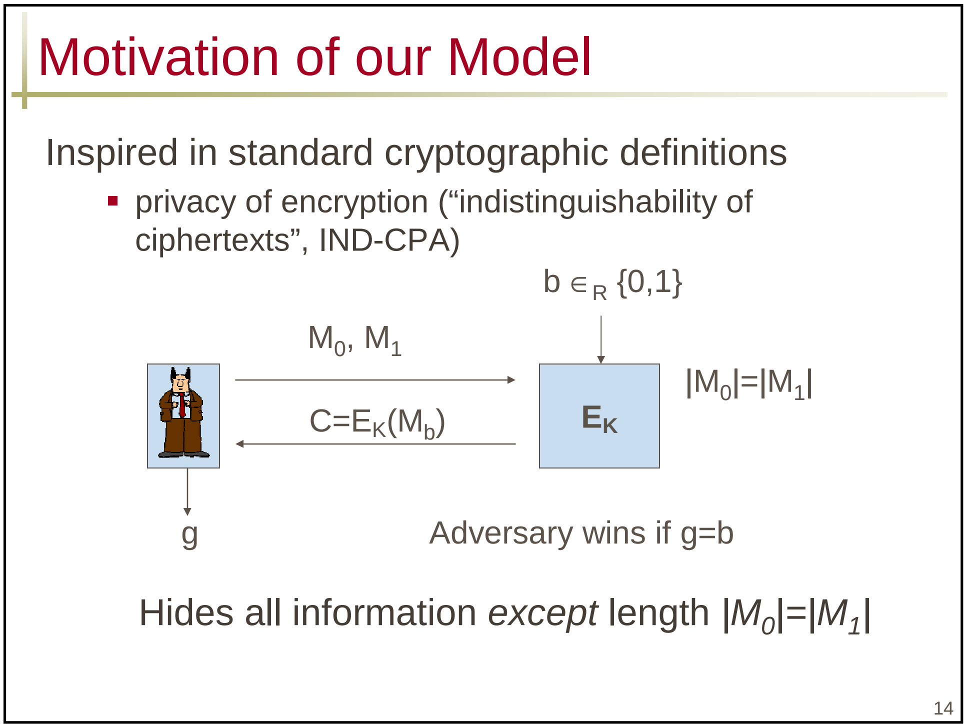
- Capturing information “leaks”
  - Total network flow
  - Amount of traffic per party
  - Value of messages sent or received per party

Hide everything except what follows from leaked information



Our results

- Intuitive but strong definition, similar as other primitives in complexity-based (computational) cryptography
- The model yields different notions
  - We show how they compare (implications)
- We study if and how some existing protocols achieve them



Motivation of our Model

Inspired in standard cryptographic definitions

- privacy of encryption (“indistinguishability of ciphertexts”, IND-CPA)

$b \in_R \{0,1\}$

M_0, M_1

$C=E_K(M_b)$

E_K

$|M_0|=|M_1|$

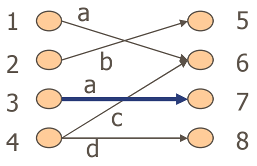
g

Adversary wins if g=b

Hides all information *except* length $|M_0|=|M_1|$

The Model – Basics

- Fixed number of parties
- Communication patterns as matrices

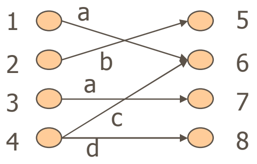


$$M = \begin{array}{c|cccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 1 & & & & & & a & & \\ 2 & & & & & b & & & \\ 3 & & & & & & & a & \\ 4 & & & & & & c & & d \\ 5 & & & & & & & & \\ 6 & & & & & & & & \\ 7 & & & & & & & & \\ 8 & & & & & & & & \end{array}$$

We model *logical messages* (not packets)

m_{ij} = sets of messages from party i to party j

The Model – Sent/Received Messages



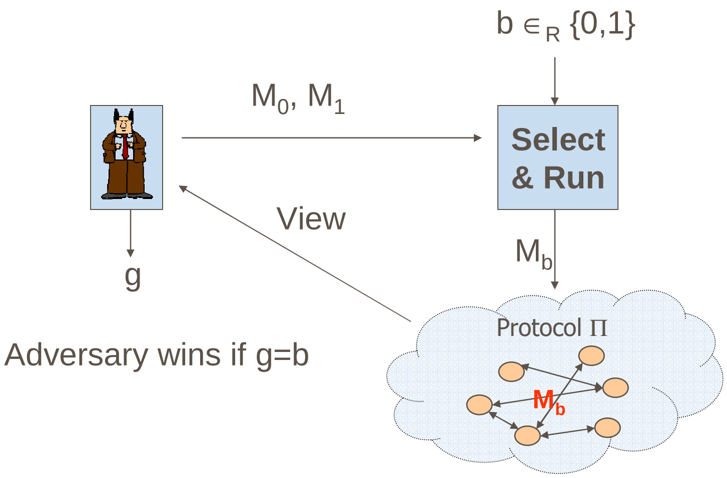
$$M = \begin{array}{c|cccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 1 & & & & & & a & & \\ 2 & & & & & b & & & \\ 3 & & & & & & & a & \\ 4 & & & & & & c & & d \\ 5 & & & & & & & & \\ 6 & & & & & & & & \\ 7 & & & & & & & & \\ 8 & & & & & & & & \end{array}$$

Send msgs = Row
Received msgs = Column

Sent by player 4 = { c , d }
Rcvd by player 6 = { a , c }

Towards a Definition

- Indistinguishability-based definition


$b \in_R \{0,1\}$

M_0, M_1

Select & Run

View

M_b

g

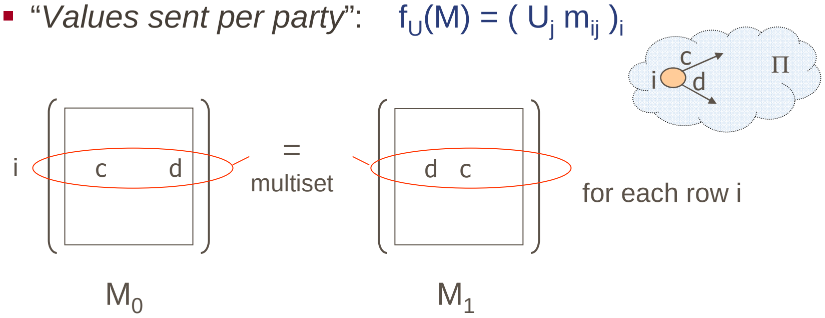
Protocol Π

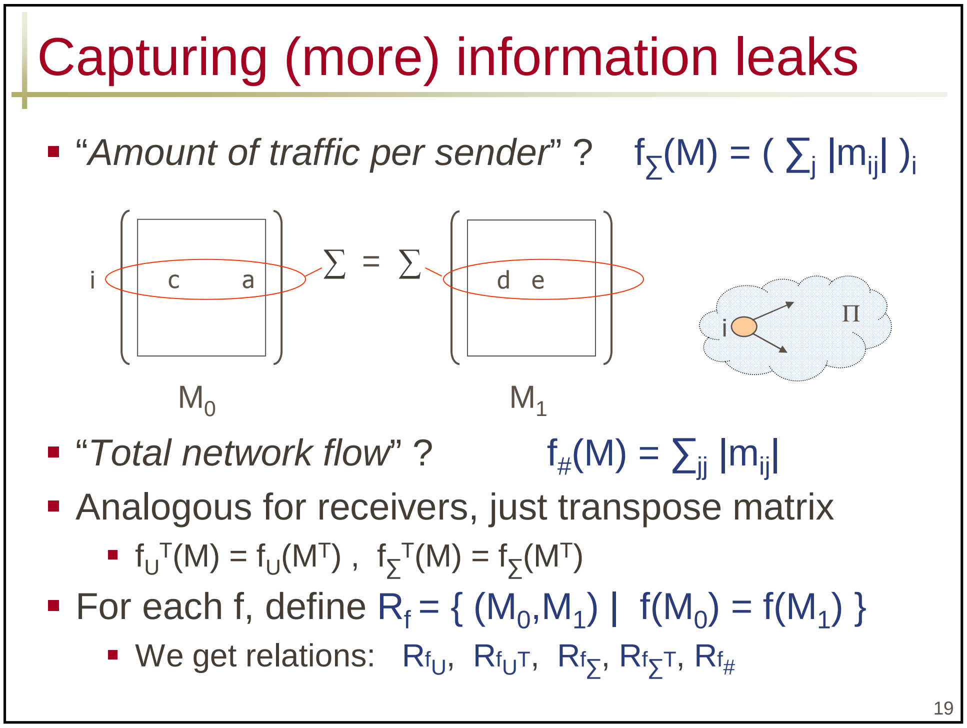
M_b

Adversary wins if g=b

Capturing information leaks

- By restricting the matrix pair M_0, M_1
  - Let f(M) be the information leaked
  - "Select & Run" requires $f(M_0) = f(M_1)$
- Example of leaked information:
  - "*Values sent per party*": $f_U(M) = (\cup_j m_{ij})_i$

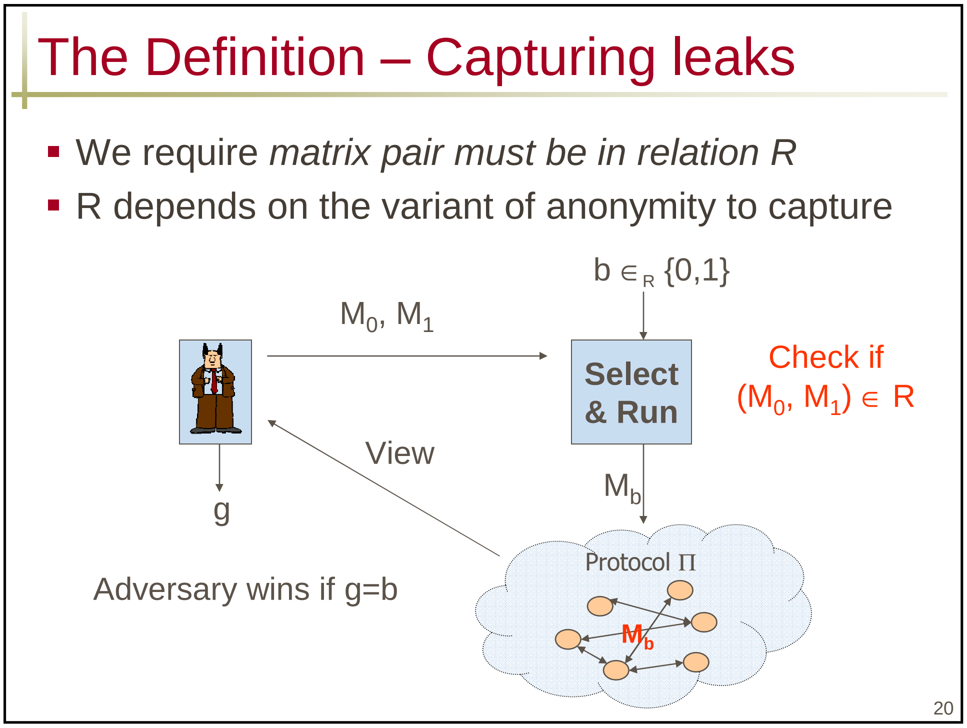




Capturing (more) information leaks

- "*Amount of traffic per sender*" ? $f_\Sigma(M) = (\sum_j |m_{ij}|)_i$

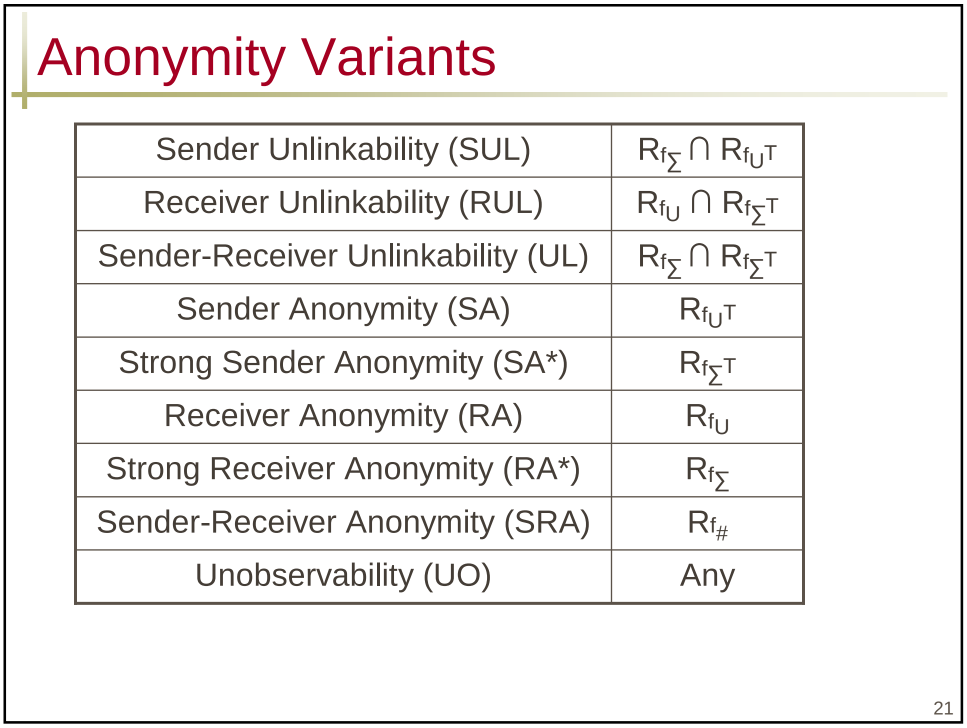
- "*Total network flow*" ? $f_\#(M) = \sum_{ij} |m_{ij}|$
- Analogous for receivers, just transpose matrix
  - $f_U^T(M) = f_U(M^T)$, $f_\Sigma^T(M) = f_\Sigma(M^T)$
- For each f, define $R_f = \{ (M_0,M_1) \mid f(M_0) = f(M_1) \}$
  - We get relations: R_{f_U}, $R_{f_U^T}$, R_{f_Σ}, $R_{f_\Sigma^T}$, $R_{f_\#}$



The Definition – Capturing leaks

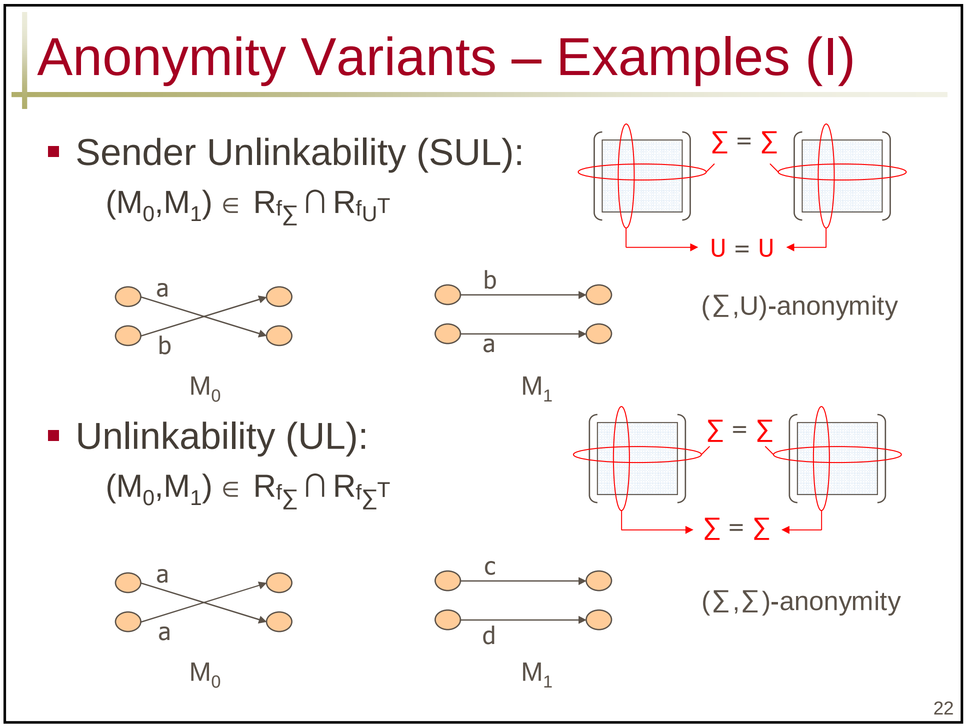
- We require *matrix pair must be in relation R*
- R depends on the variant of anonymity to capture

Adversary wins if g=b



Anonymity Variants

Sender Unlinkability (SUL)	$R_{f_\Sigma} \cap R_{f_U^T}$
Receiver Unlinkability (RUL)	$R_{f_U} \cap R_{f_\Sigma^T}$
Sender-Receiver Unlinkability (UL)	$R_{f_\Sigma} \cap R_{f_\Sigma^T}$
Sender Anonymity (SA)	$R_{f_U^T}$
Strong Sender Anonymity (SA*)	$R_{f_\Sigma^T}$
Receiver Anonymity (RA)	R_{f_U}
Strong Receiver Anonymity (RA*)	R_{f_Σ}
Sender-Receiver Anonymity (SRA)	$R_{f_\#}$
Unobservability (UO)	Any



Anonymity Variants – Examples (I)

- Sender Unlinkability (SUL):

  $(M_0,M_1) \in R_{f_\Sigma} \cap R_{f_U^T}$

(Σ,U)-anonymity

- Unlinkability (UL):

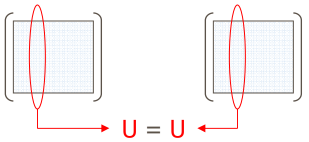
  $(M_0,M_1) \in R_{f_\Sigma} \cap R_{f_\Sigma^T}$

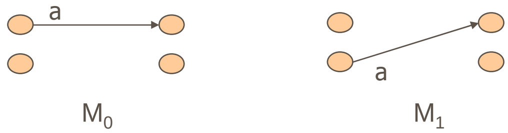
(Σ,Σ)-anonymity

Anonymity Variants – Examples (I)

- Sender Anonymity (SA):

  $(M_0,M_1) \in R_{f_{U^T}}$

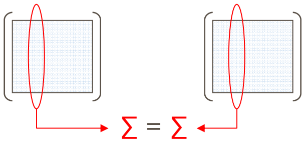


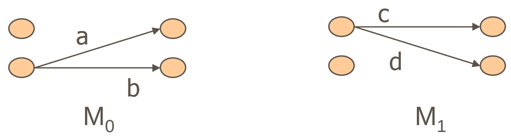


(?,U)-anonymity

- Strong Sender Anonymity (SA*):

  $(M_0,M_1) \in R_{f_{\Sigma^T}}$



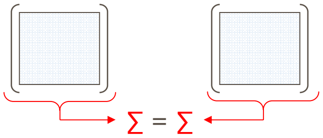


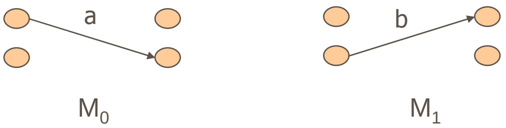
(?,∑)-anonymity

Anonymity Variants – Examples (II)

- Sender-Receiver Anon. (SRA):

  $(M_0,M_1) \in R_{f_\#}$

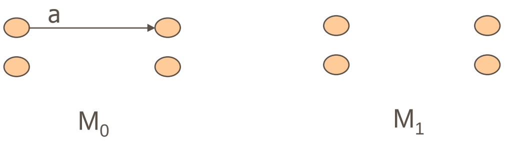




(?, ∑)-anonymity

- Unobservability (UO):

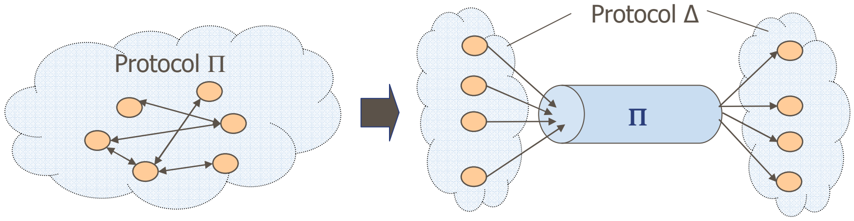
  Any (M_0,M_1)



(?, ?)-anonymity

Comparing the Notions

We say A → B if there exits Δ such that



Achieves notion A

Protocol Δ^{Π}
Achieves notion B

(Black box implications)

Implications under Computational Assumptions

Lemma 1: Under the PKI model, if semantically secure key-private encryption exists then

- SUL → UL
- RUL → UL
- SA → SA*
- RA → RA*

Proof Idea: (say N→N')

- Relax equality into computational indistinguishability in the notions (I-N anonymity) and prove I-N does not weaken the adversary.
- Use encryption to achieve N' from protocol achieving I-N (as black box). Prove the reduction works.

Implications using dummy messages

Lemma 2a: Assume the total traffic flow is upper bounded by known value. Then,

SUL → SA, UL → SA*, RA* → UO

Lemma 2b:

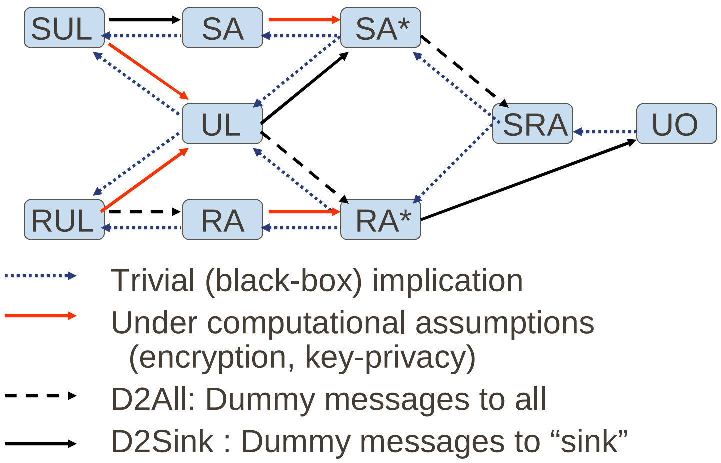
RUL → RA, UL → RA*, SA* → SRA

Proof Idea: Two simple strategies that work

- (2b) Each sender pads its traffic up to the known value.
- (2b') For each message to party i, sender sends a dummy to each other party j ≠ i

Lemma 3: Both strategies are optimal in number of dummy mesages

Summary of Relations between Notions



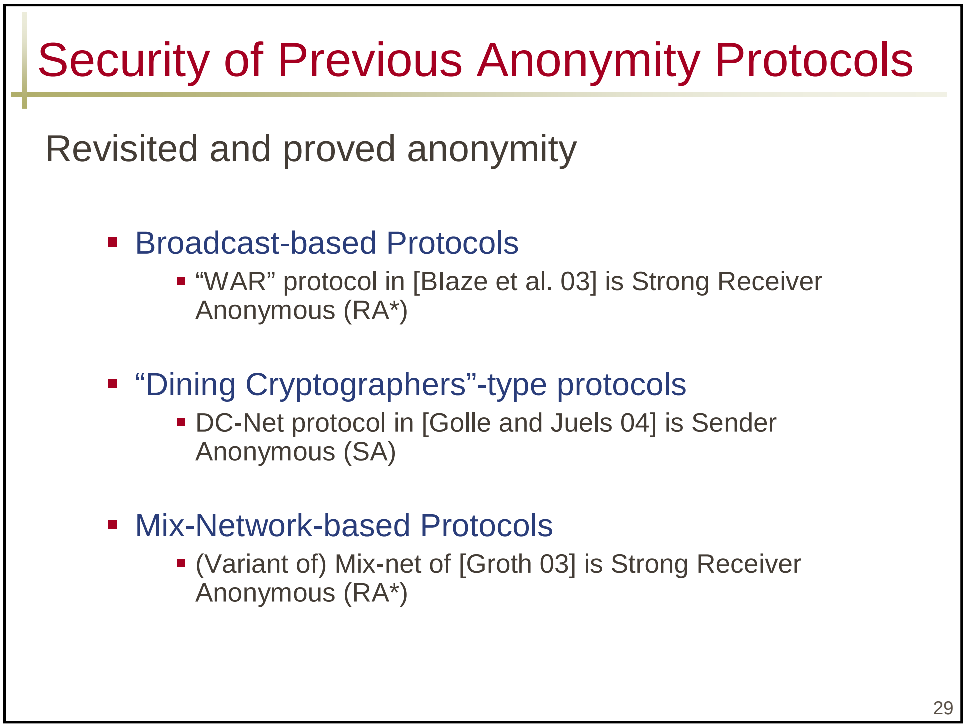
Trivial (black-box) implication

Under computational assumptions (encryption, key-privacy)

D2All: Dummy messages to all

D2Sink : Dummy messages to "sink"

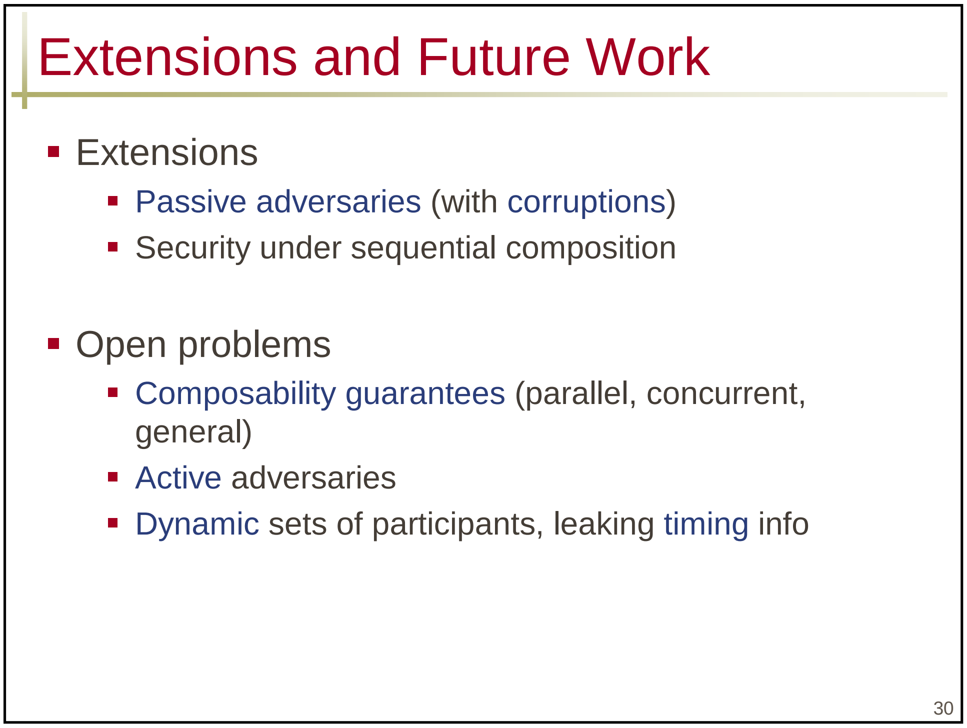
Amount of dummy traffic is provably optimal



Security of Previous Anonymity Protocols

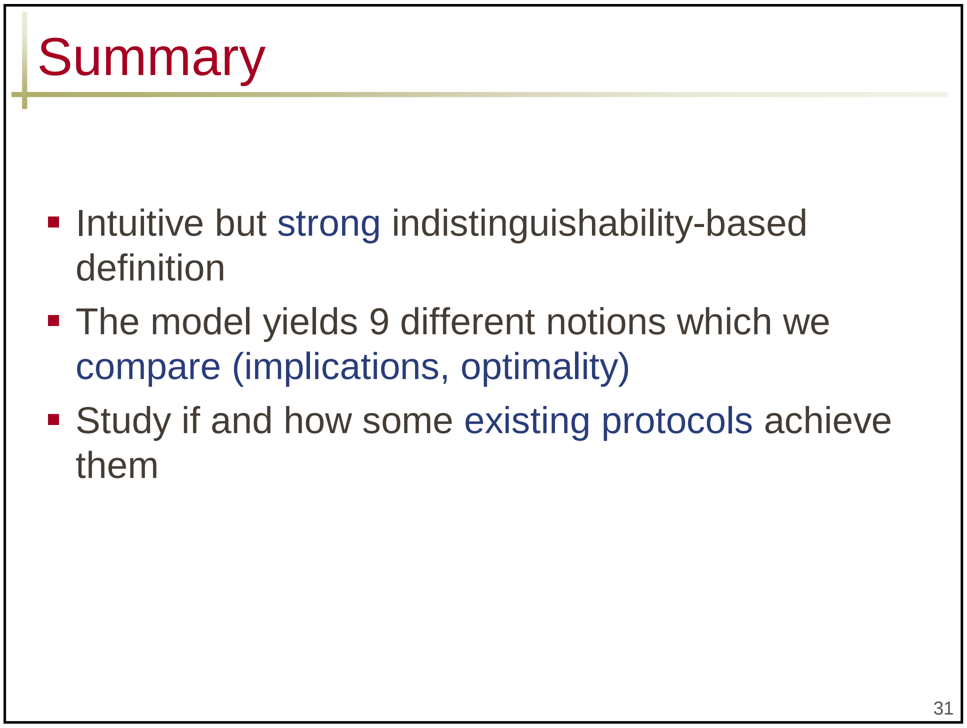
Revisited and proved anonymity

- Broadcast-based Protocols
  - “WAR” protocol in [Blaze et al. 03] is Strong Receiver Anonymous (RA*)
- “Dining Cryptographers”-type protocols
  - DC-Net protocol in [Golle and Juels 04] is Sender Anonymous (SA)
- Mix-Network-based Protocols
  - (Variant of) Mix-net of [Groth 03] is Strong Receiver Anonymous (RA*)



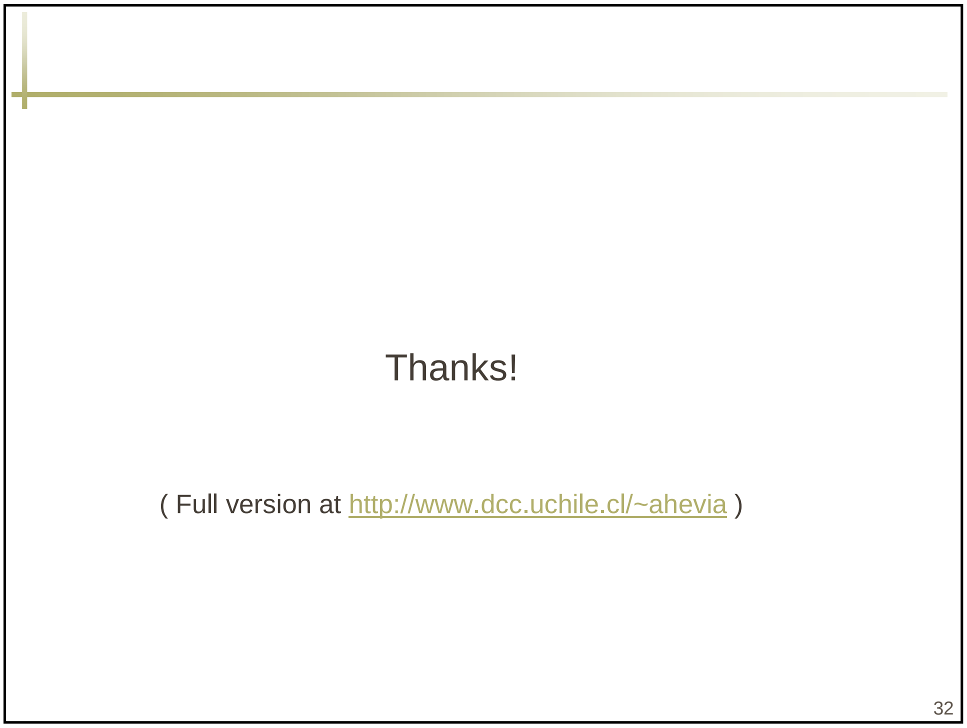
Extensions and Future Work

- Extensions
  - Passive adversaries (with corruptions)
  - Security under sequential composition
- Open problems
  - Composability guarantees (parallel, concurrent, general)
  - Active adversaries
  - Dynamic sets of participants, leaking timing info



Summary

- Intuitive but strong indistinguishability-based definition
- The model yields 9 different notions which we compare (implications, optimality)
- Study if and how some existing protocols achieve them



Thanks!

(Full version at http://www.dcc.uchile.cl/~ahevia)