Privacy II. Network Anonymity and Secure Multiparty Computation

Bo Luo
Associate Professor, EECS
Director, Information Assurance Lab, ITTC
The University of Kansas, Lawrence, KS, USA
bluo@ku.edu; http://www.ittc.ku.edu/~bluo

Anonymity

- Data anonymity
  - Unidentifiability
  - Database and data mining
  - Privacy-preserving data publishing
- Network anonymity 网络匿名
  - Unobservability
  - Unlinkability
  - Sender anonymity
  - Receiver anonymity

Anonymous Network

- Chaum's MIX
- Onion Routing
- Crowds 不考

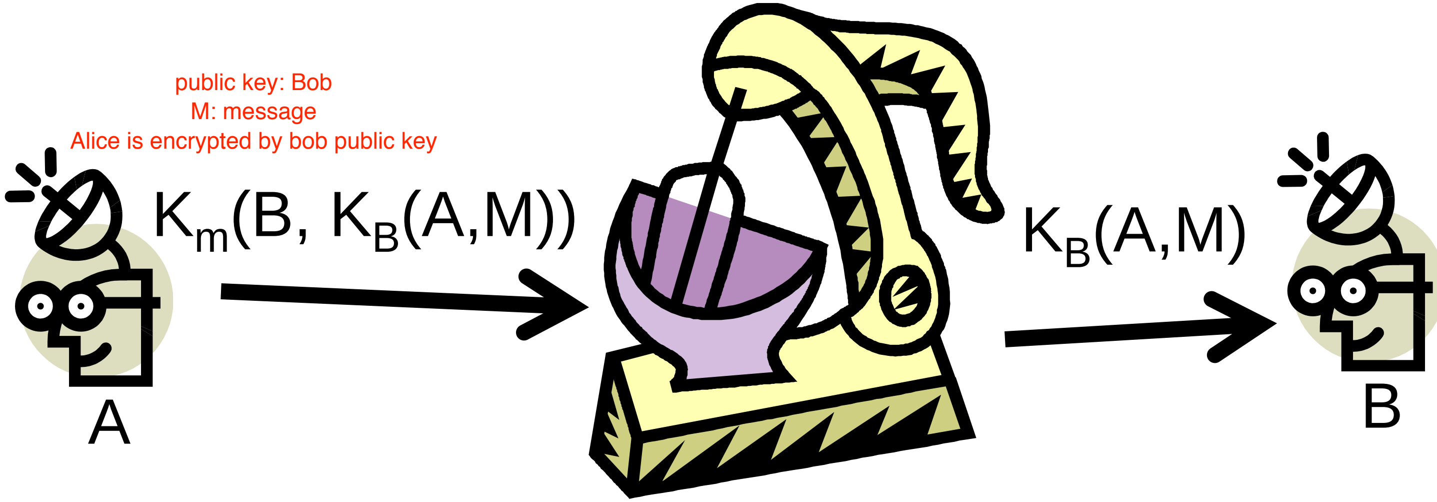
Chaum's MIX

- Presented first in 1981 by David Chaum
- Uses public key cryptography for anonymous e-mail
- Basic Idea:
  - E-mails would be sent to a "Mix" which would then forward them onto recipients
  - Unlinkability: The adversary knows all the senders and receivers but cannot link senders to receivers
- Key building block for anonymity systems

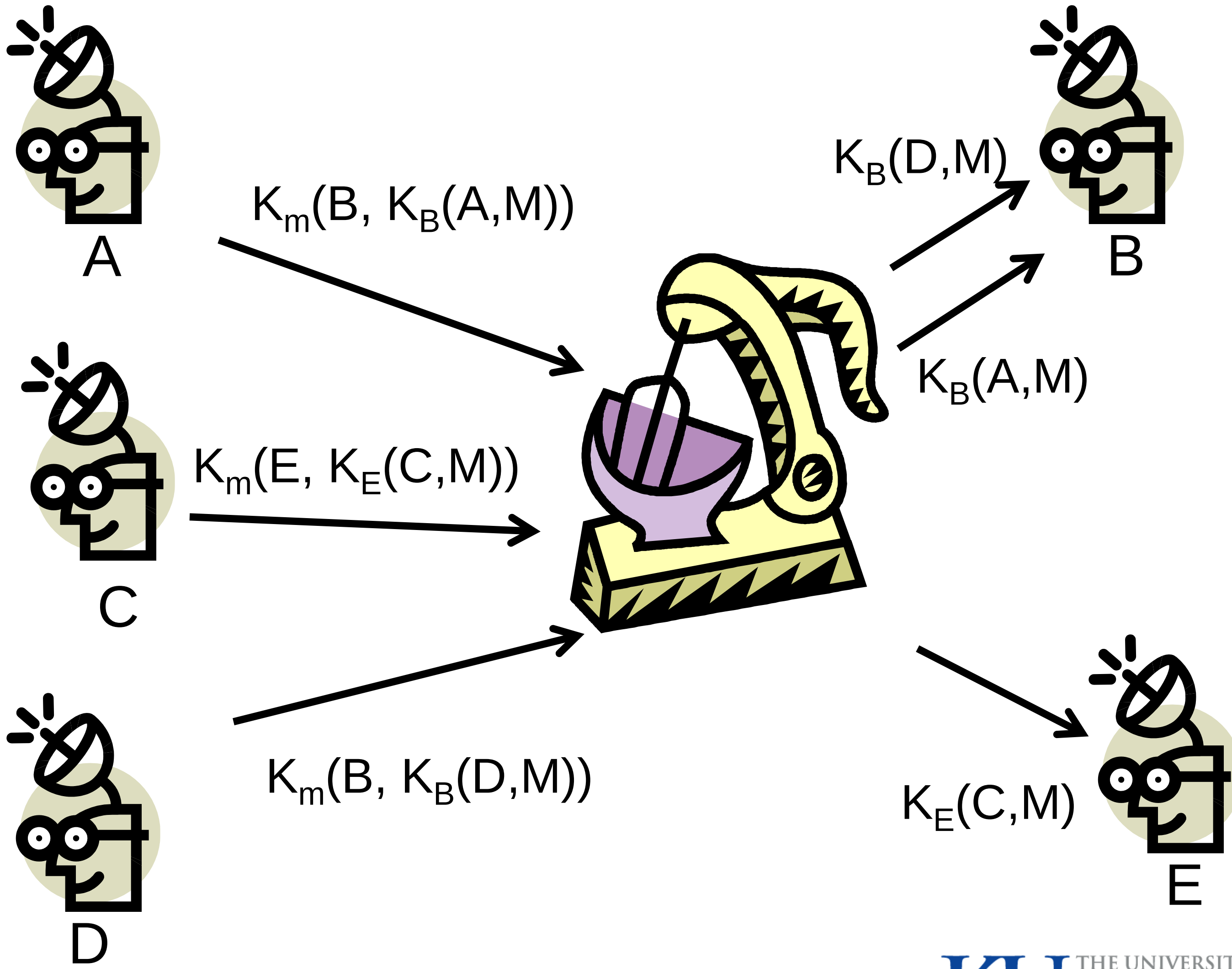
Chaum's MIX

Alice sent email to Bob

public key: Bob
M: message
Alice is encrypted by bob public key

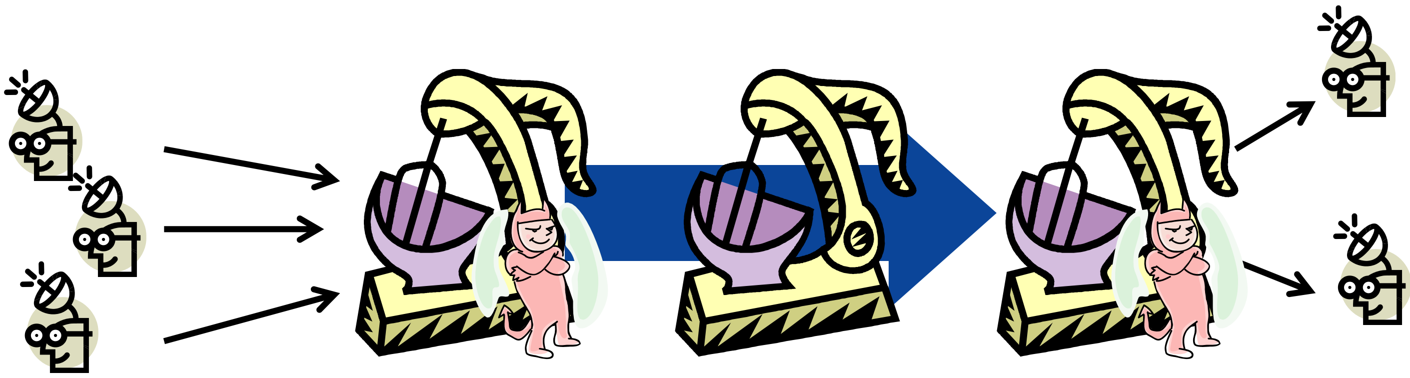


Chaum's MIX



MIX Cascade

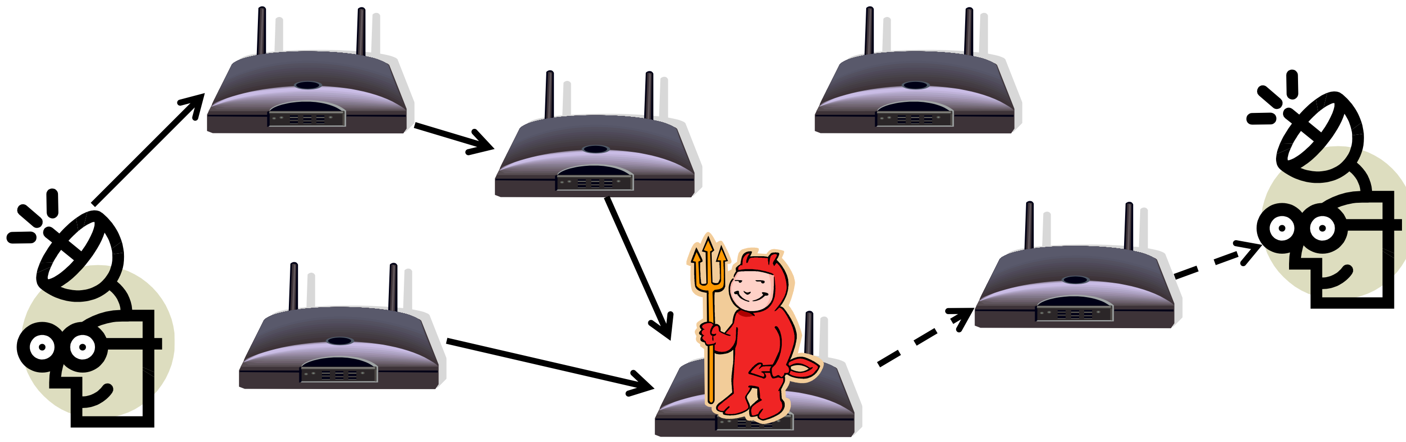
- What if some of the mixes are controlled by adversaries?
- A cascade of mixes can be used to handle compromised mixes



- How many adversaries can this withstand?
  - N-1

Anonymity via Random Routing

- Hide message source through random routing
- Routers don't know for sure who the source of the message is



Anonymity via Random Routing

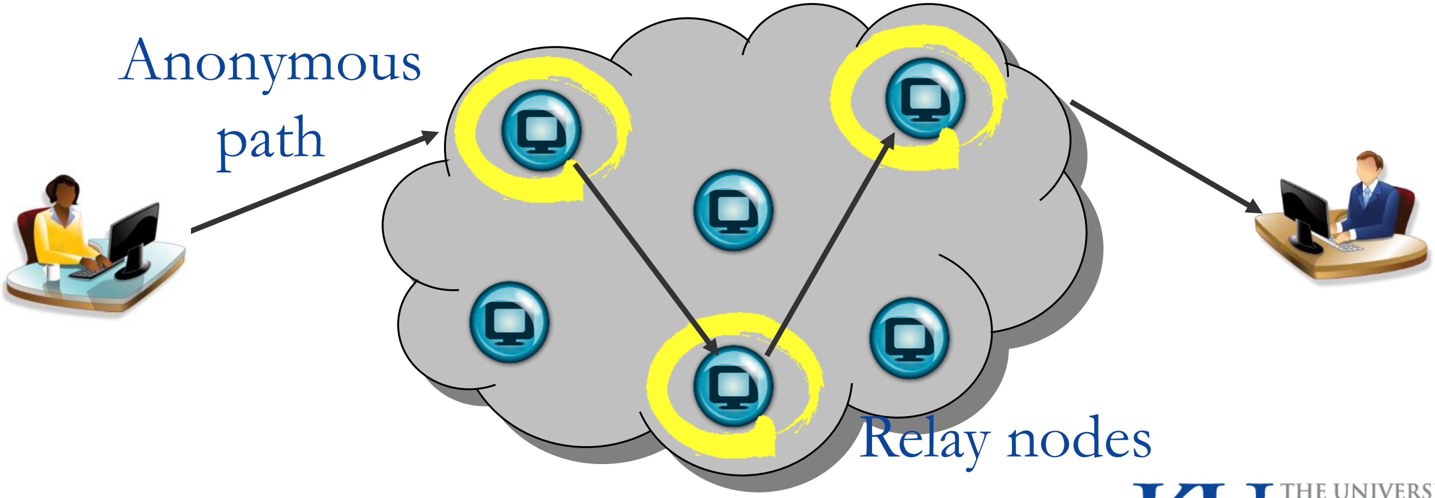
- Chaum's Mix (Chaum 1981)
  - Decryption and re-encryption, and reorder
- Onion routing (Syverson et al. 1997)
  - Layered encryption using pair-wise symmetric keys
- Crowds (Reiter et al. 1998)
  - Probabilistic random walk with pf
- P5 (Sherwood et al. 2001)
  - Dining cryptographer network
- Tarzan, MorphMix, Freedom, Tor, Cashmere, Salsa, …

Anonymizing network

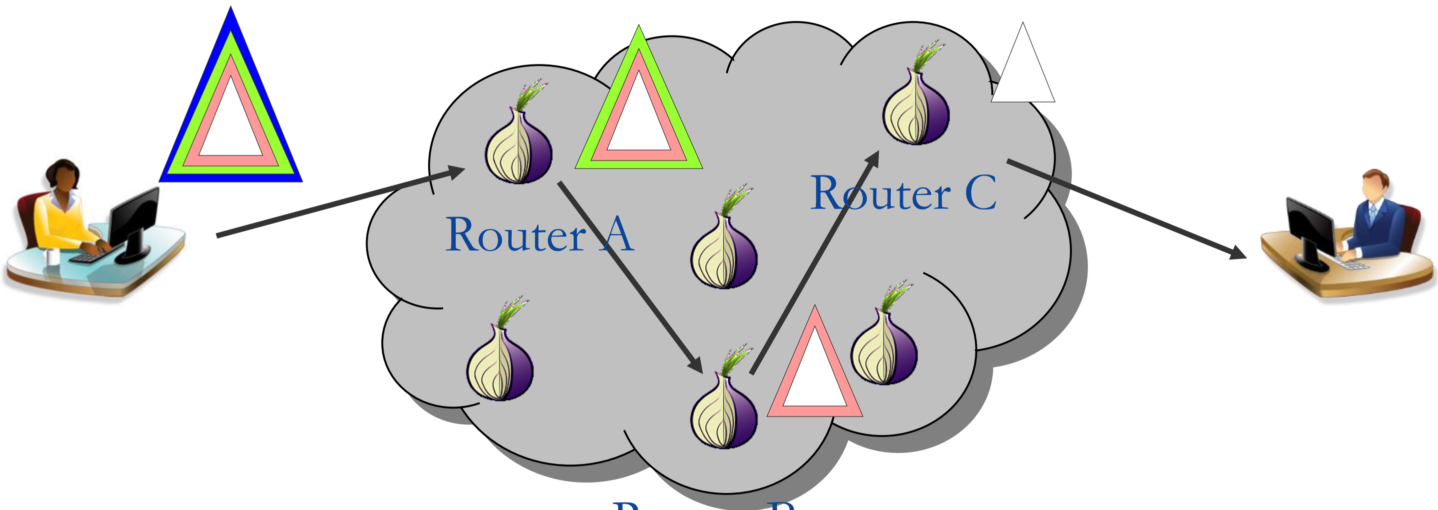
- Sender chooses a random sequence of routers
  - Some are honest, some aren't
  - Similar to mix cascade

Anonymizing network

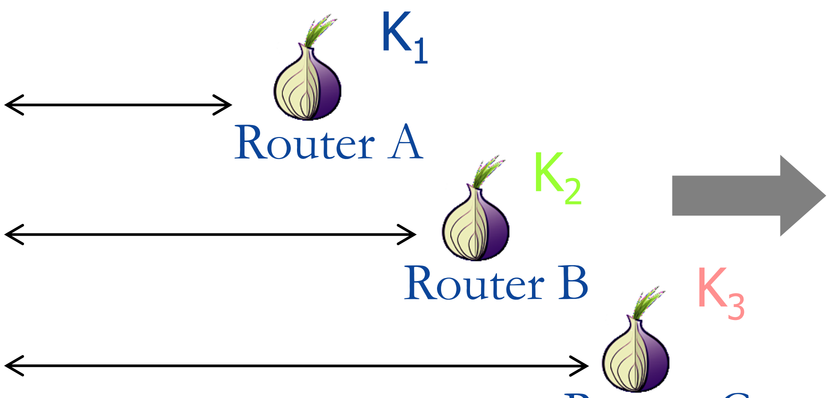
- An anonymizing network is an overlay with relay nodes
  - Server-based or peer-to-peer
- Selecting a set of nodes from available relays to construct a circuit to relay the packets
- Packets are encrypted along the anonymous path
  - Goal: Hostile routers shouldn't learn Alice is talking to Bob



Onion routing



encrypted by C,B,A order!
decrypted by A B C



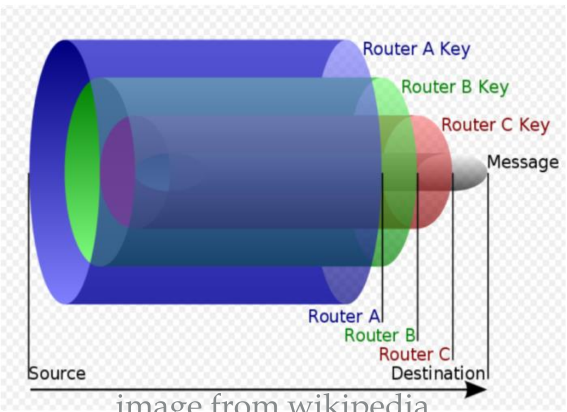


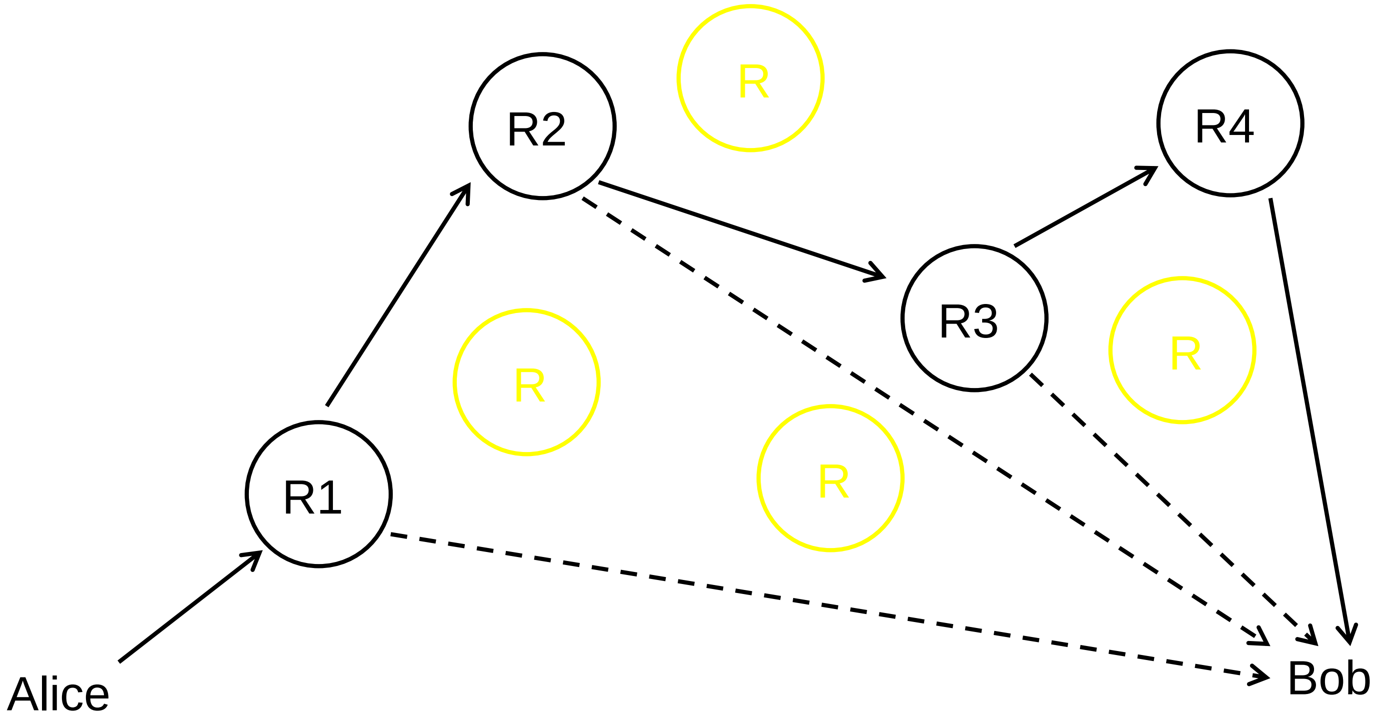
image from wikipedia

Crowds

不考

- Routers form a random path
  - Different than onion routing because the routers choose path, not sender
- After receiving a message router flips a biased coin
  - With probability p, the router forwards the message to another router
  - With probability 1-p, the router forwards the message to the recipient

Crowds



Problems

- Static paths suffer from node failures
  - Node failure → Path failure
  - Detection of a node failure is slow
  - Reconstructing an anonymous path is expensive
  - Frequent path reformations increase the vulnerability to the predecessor attack
  - The problem gets worse in P2P anonymizing networks

Secure Multiparty Computation

- Participants: $p_1, p_2, \ldots, p_N$,
- Private inputs, $d_1, d_2, \ldots, d_N$
- Objective: compute the value of a public function

$$F(d_1, d_2, \ldots, d_N)$$

while keeping the private inputs secret.

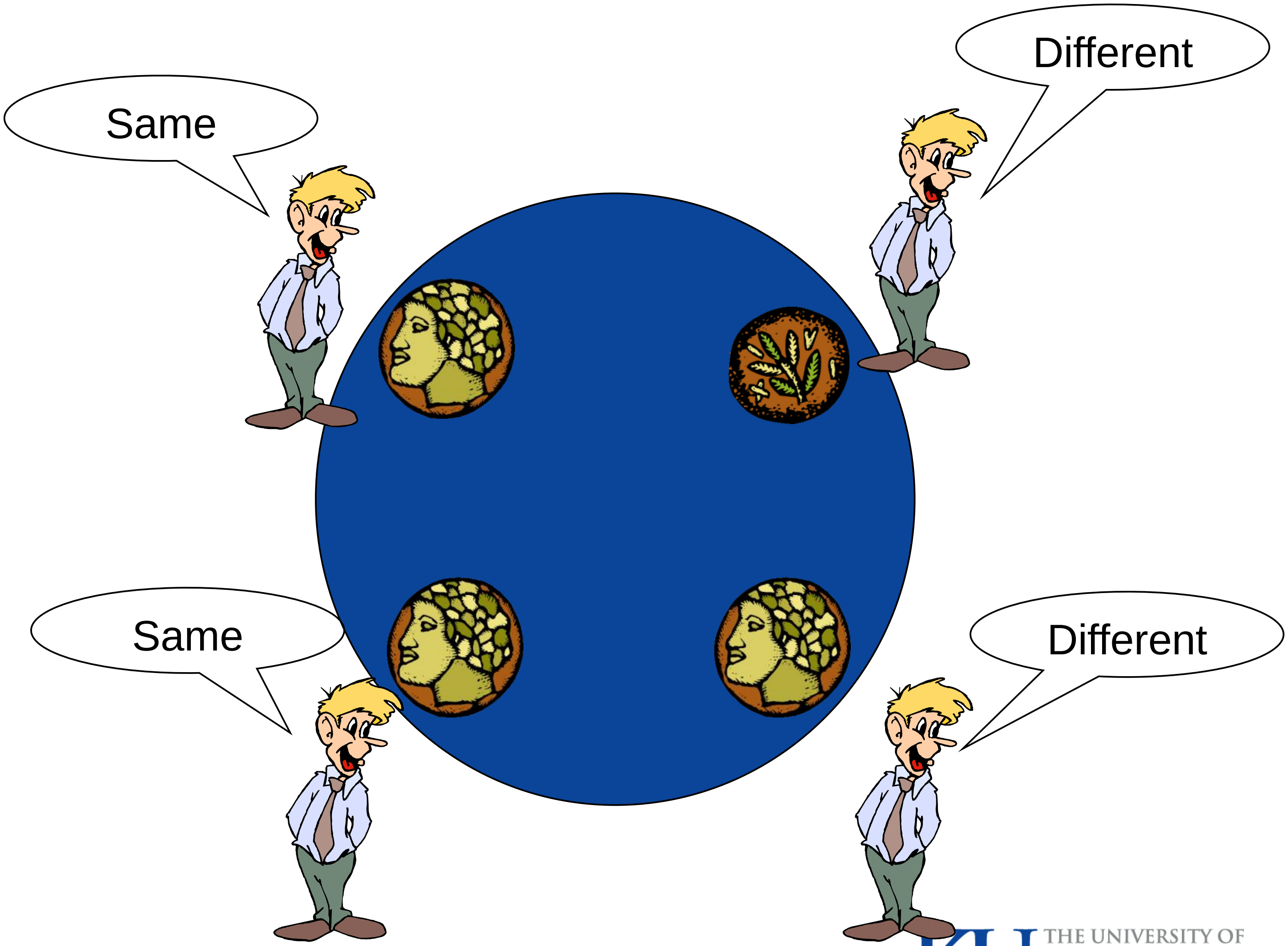
Dining Cryptographers

- Introduced by Chaum
- To release a public message in a perfectly untraceable manner
  - N cryptographers are having dinner
  - Waiter tells them that the dinner has been paid for but they want to know whether it was one of them that paid or the NSA agent in the corner

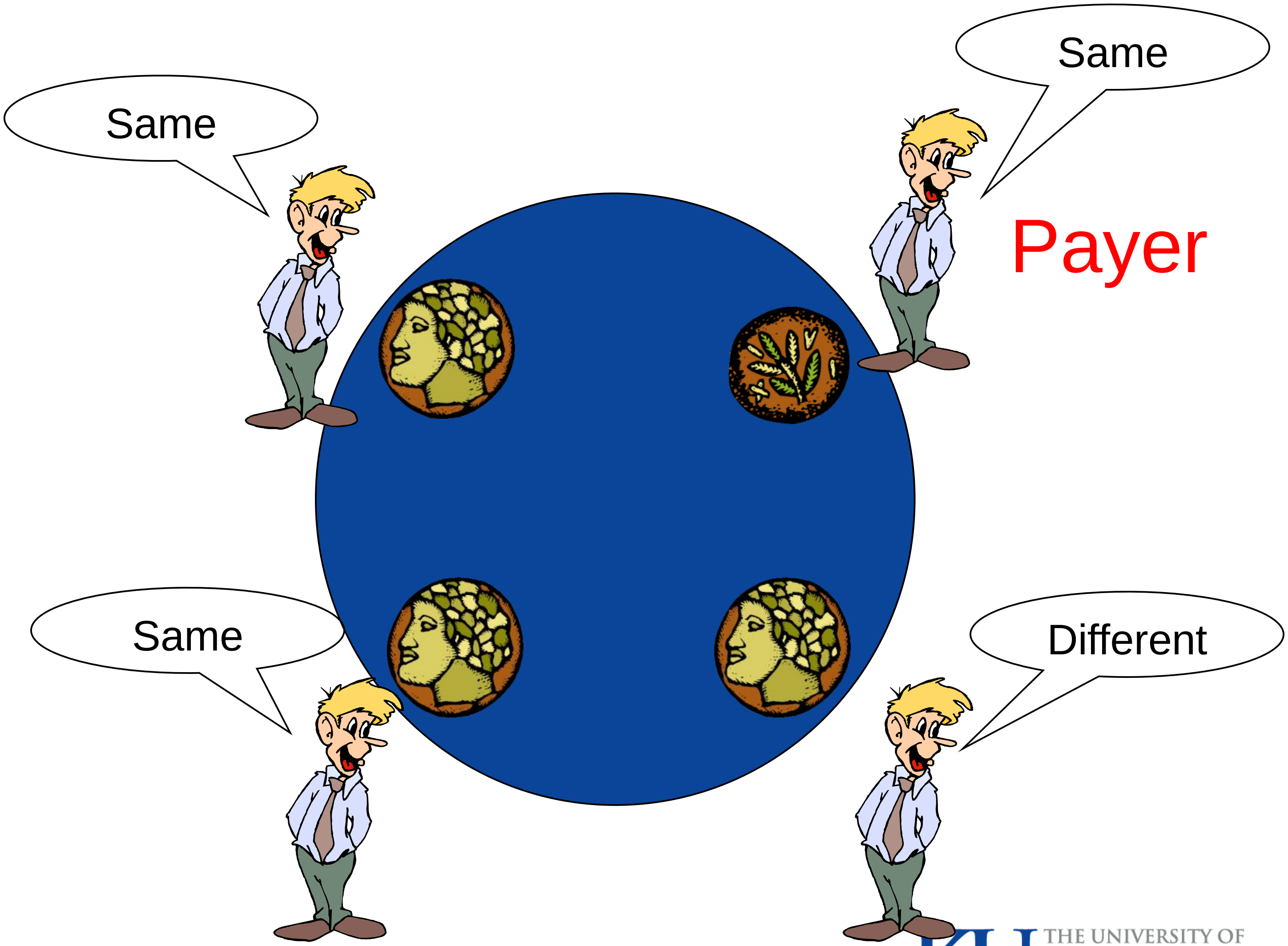
Dining Cryptographers

- The Protocol
  - Each diner flips a coin and shows it to his left neighbor
  - Each diner announces whether he and his neighbor’s coin flips are the same or different. The payer lies.
  - Even number of “different” => no one lied => NSA paid
    Odd number of “different” => one the diners paid

Dining Cryptographers



Dining Cryptographers



Problems with DC

- Very Impractical
  - Only one bit sent at a time
  - Each party has to have pairwise secure channels
  - Massive communication overhead
    - For N ‘diners’
    - N messages sent to share coins
    - N broadcast messages to share
    - All this for 1 bit

Secure two-party computation

- Yao's Millionaires' problem: two millionaires are interested in knowing which of them is richer without revealing their actual wealth.

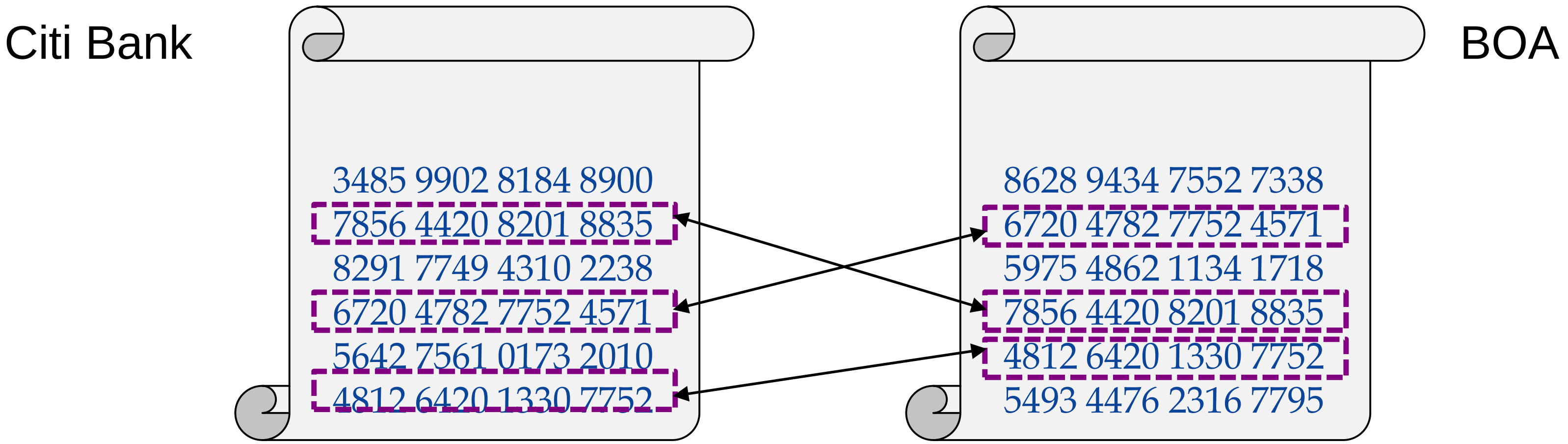
- 2-party Secure Function Evaluation (SFE)
  - Alice has $\{x_1, x_2, \ldots x_n\}$
  - Bob has $\{y_1, y_2, \ldots y_n\}$
  - They want to learn $f(x,y)$ without revealing their own values.

FairPlay

- Yao’s construction is about 20 years old. There were no known implementations (?).
- FairPlay - a full fledged secure two-party computation system, implementing Yao’s “garbled circuit” protocol.
- Nisan,Malkhi,Pinkas,Sella USENIX Security 2004

Record Linkage

- Record linkage is to identify related records associated with the same entity from multiple databases



Privacy-Preserving Record Linkage

- Privacy becomes an issue when data is sensitive.
  - I will only share with you on the "linked records"
  - I will not give you the plain text of my primary keys.

- Secure multi-party set intersection problem
  - Solutions based on commutative encryption
  - Solutions based on homomorphic encryption

THE UNIVERSITY OF KANSAS

Privacy-Preserving Record Linkage

- A Naïve Solution
  - Citi hashes its records
  - BOA hashes its records
  - They exchange the hashes
  - Identical hash → shared record

  - What is wrong here?

Agrawal's method

- Commutative encryption: using the same set of commutative keys, the encrypted content can be recovered in any arbitrary order.

$$f(g(v)) = g(f(v))$$

Agrawal's method

- Protocol
  - Hashing
  - Encryption
  - Exchange
  - Encryption
  - Compare
  - Decryption

Commutative Encryption

- Commutative Encryption: using the same set of commutative keys, the encrypted content can be recovered in any arbitrary order.

- AES Protocol [Agrawa et. al., SIGMOD 2003]: