Hidden Communication in P2P Networks

Steganographic Handshake and Broadcast

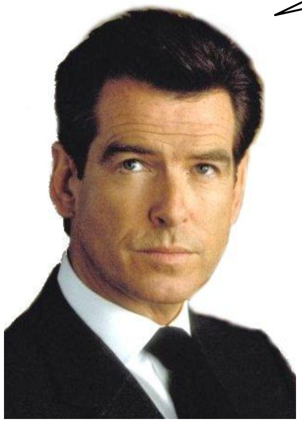
Raphael Eidenbenz, Thomas Locher, Roger Wattenhofer

INFOCOM 2011

ETH Zurich – Distributed Computing Group

Spy Rendezvous

In London, April's a spring month.



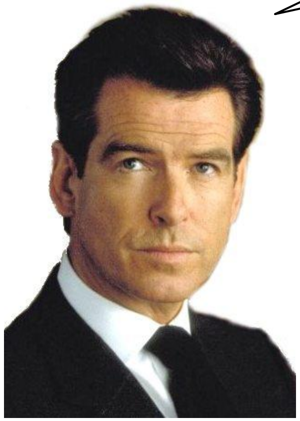
..whereas in St.Petersburg we're freezing our butts off.

MI6

CIA

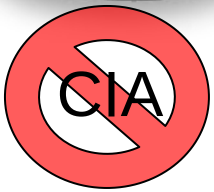
Spy Rendezvous

In London, April’s a spring month.



Yes, Russian winters are cold. Enjoy your stay, sir.

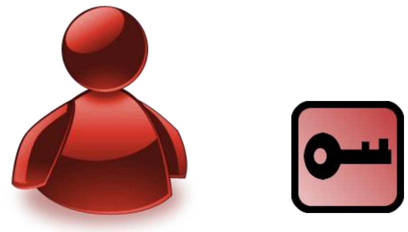
??



Steganographic Handshake in Networks

Regular peers

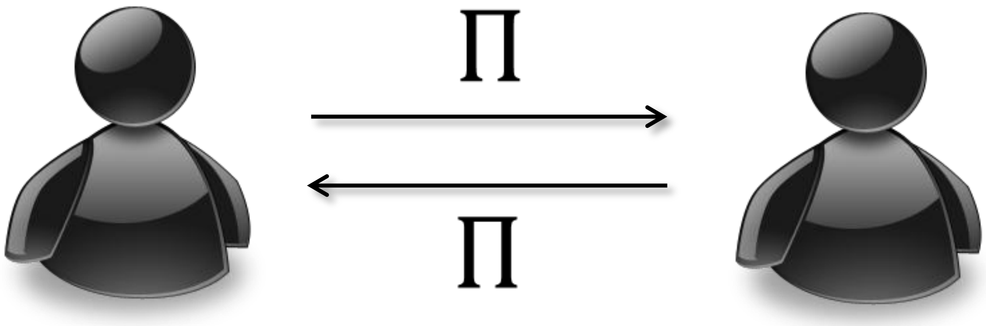
Conspirers



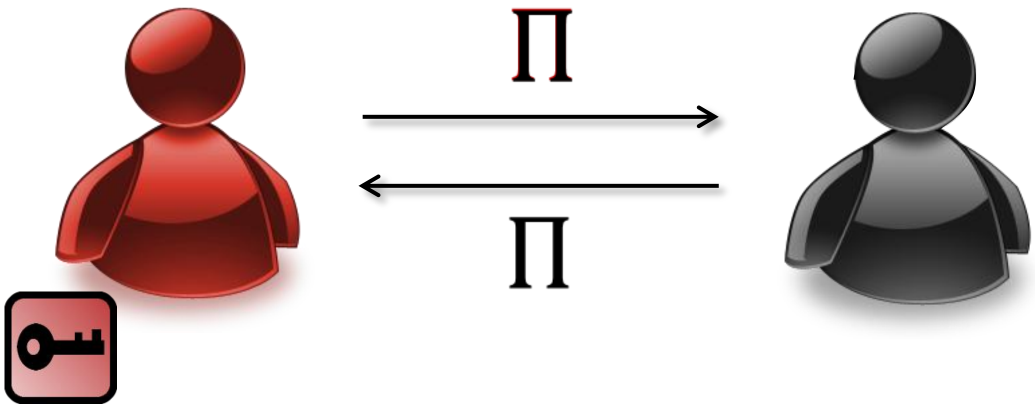
Share files

Talk to other conspirers
without raising suspicion

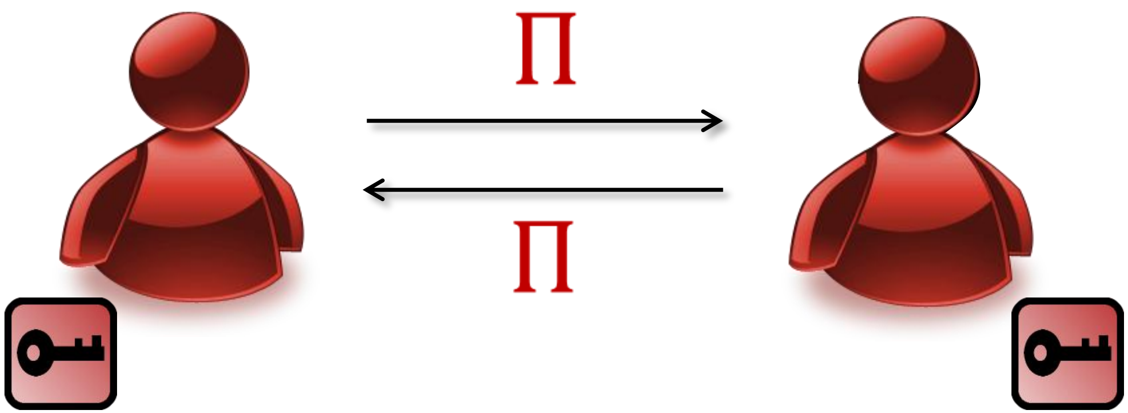
Steganographic Handshake in Networks



Steganographic Handshake in Networks



Steganographic Handshake in Networks

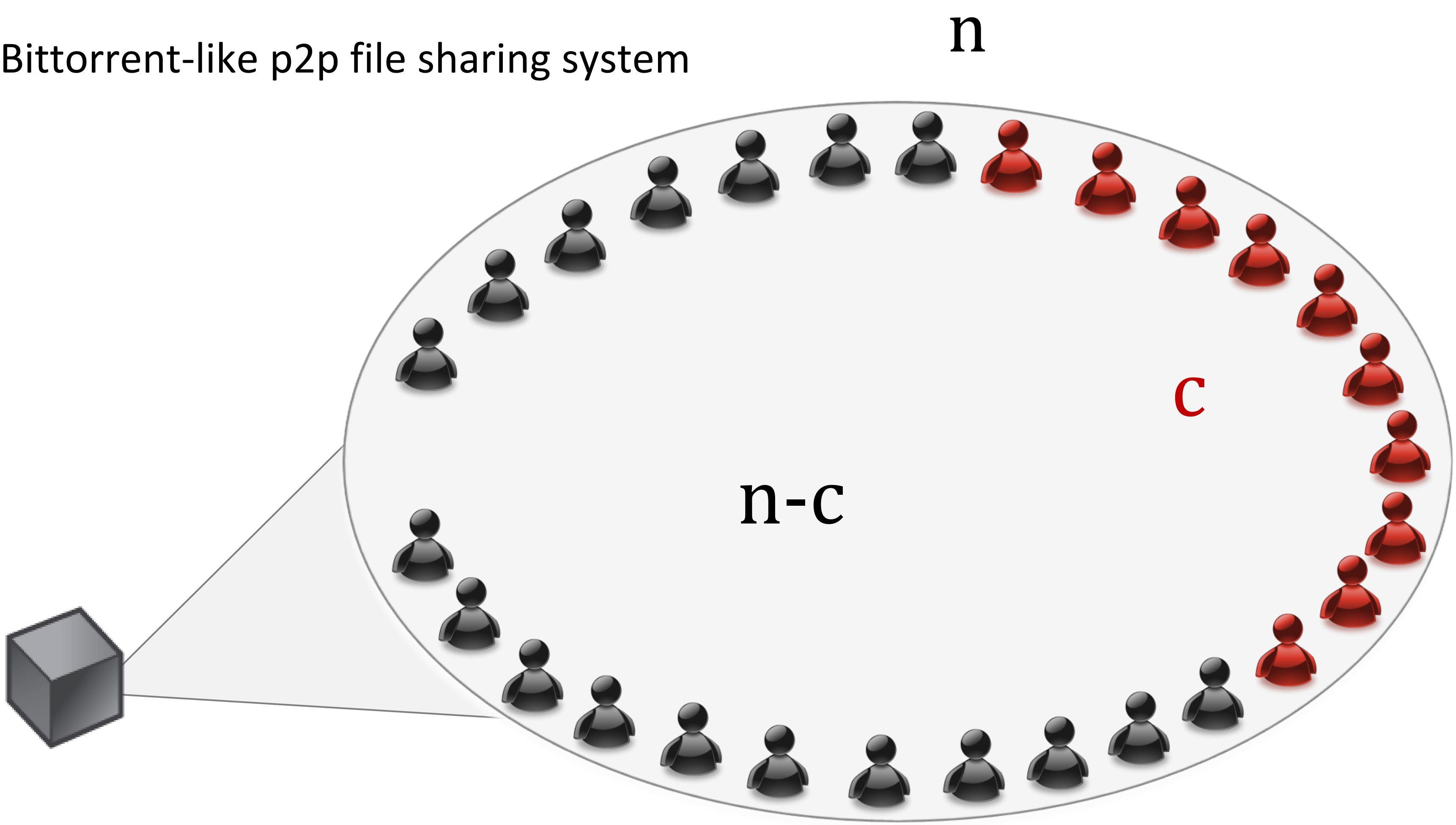


Steganographic Channels

- P2P File sharing
  - Block request sequence
  - Block subset selection
- Timing
- Bandwidth
- Ports

Steganographic Broadcast

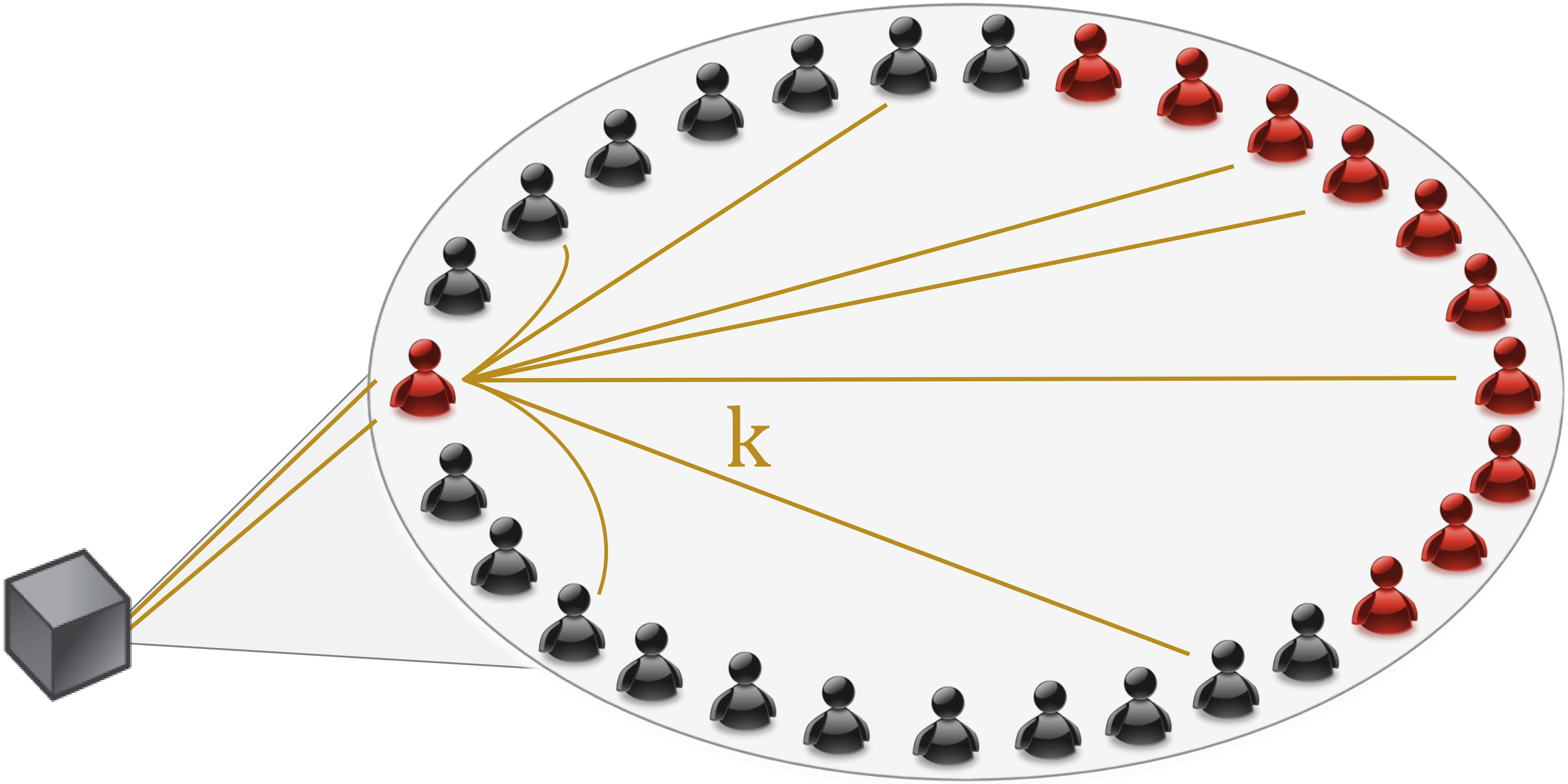
- Send a message to all conspirers
- Bittorrent-like p2p file sharing system



Steganographic Broadcast

- Send a message to all conspirers

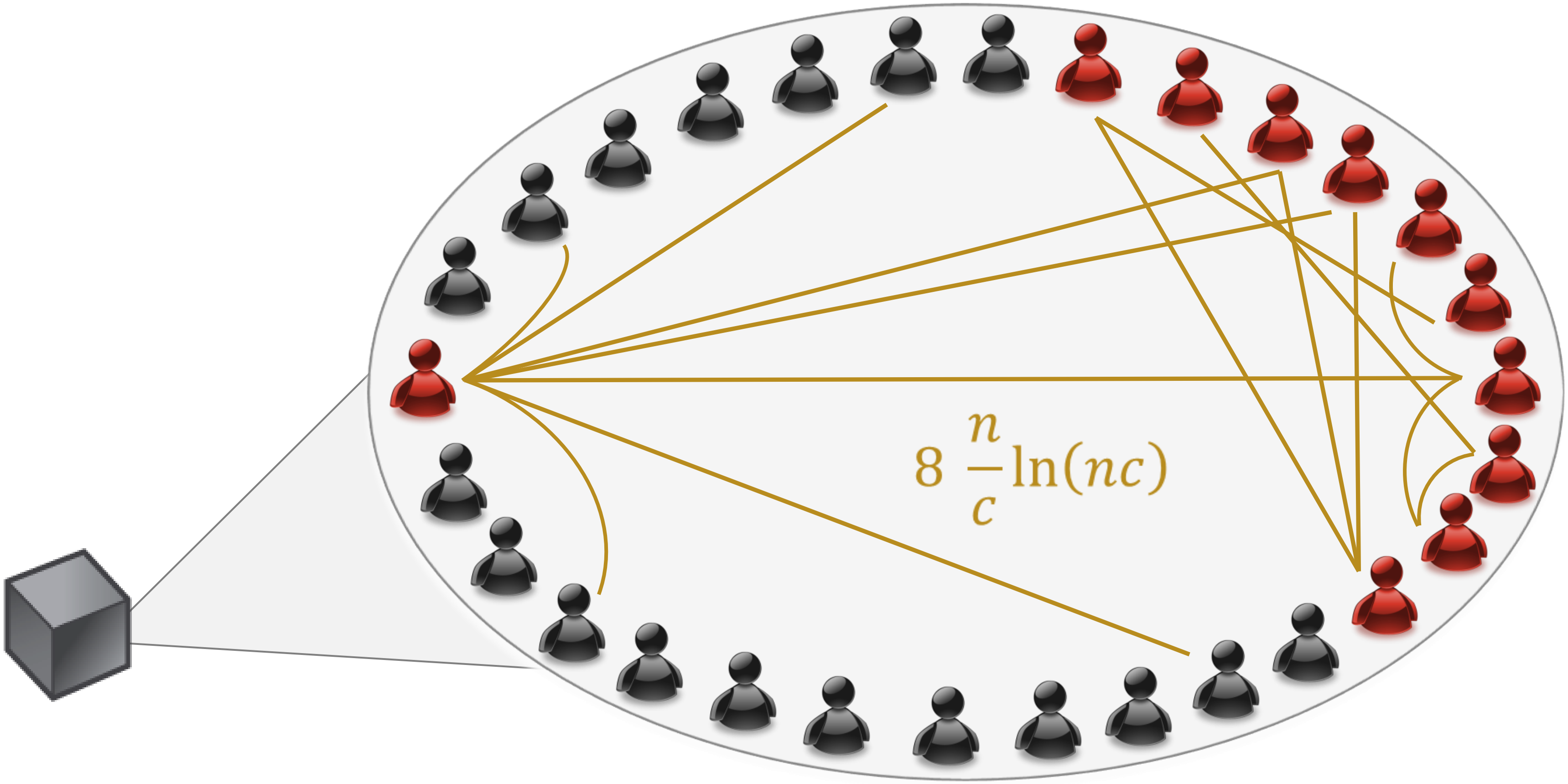
- Bittorrent-like p2p file sharing system



Efficient Broadcast

Lemma

If each conspirer randomly connects to $8\ \frac{n}{c}\ln(nc)$ peers, then the subnetwork induced by the c conspirers is connected w.h.p.



Efficient Broadcast

Lemma

If each conspirer randomly connects to $8\,\frac{n}{c}\ln(nc)$ peers, then the subnetwork induced by the c conspirers is connected w.h.p.

Algorithm

Get $8\,\frac{n}{c}\ln(nc)$ peer addresses
Acquire $6\log n$ blocks
Reveal types of connected peers
Broadcast message M in conspirer subnetwork

- Space complexity $O\left(\frac{n}{c}\log n + |M|\right)$
- Communication complexity $O\left(\frac{n}{c}\log n + \log^2 n + |M|\log n\right)$ w.h.p.

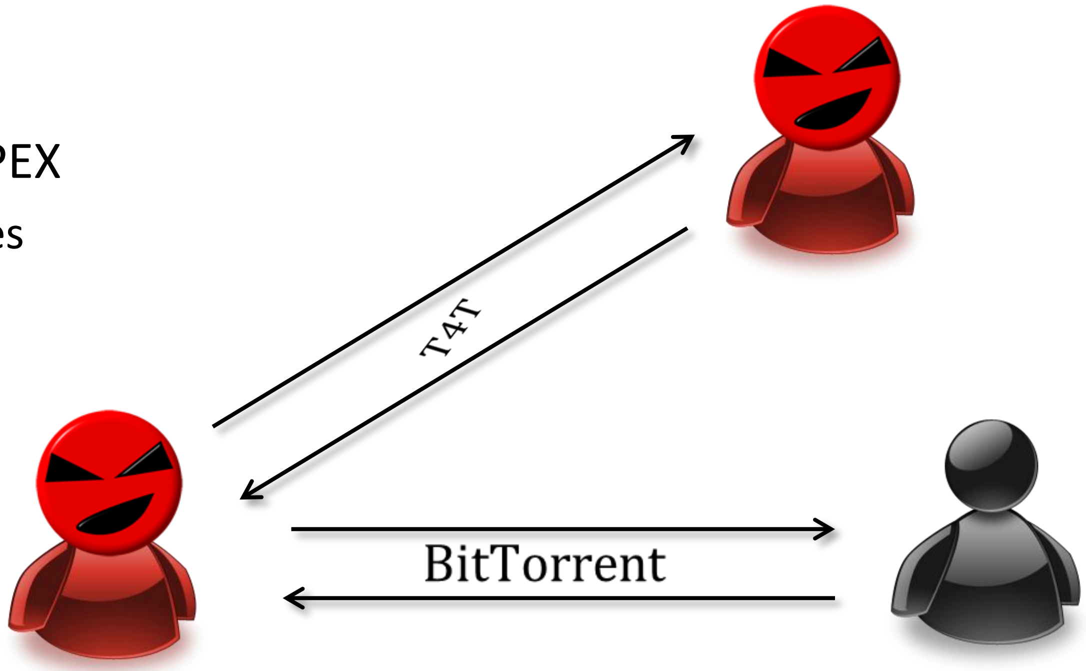
Stronger Authority Models

- Individual Monitoring
  - Authority monitors individual communication links, no correlation
  - $|M| \in \Theta(m \log m)$ where m is the # of blocks
- Complete Monitoring
  - Authority monitors complete network
  - $|M| \in \Theta(\sqrt{m} \log^2 m)$
- Stochastic Monitoring
  - Trade-off: Hidden communication vs. False positives

Steganographic Handshake in BitThief

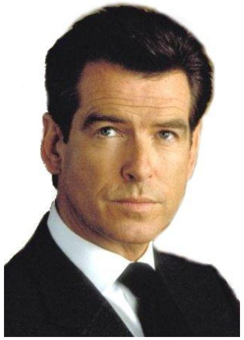
- BitThief is a BitTorrent client that
  - Free rides with BitTorrent clients [1], and
  - Trades tit-for-tat (T4T) with other BitThiefs [2]

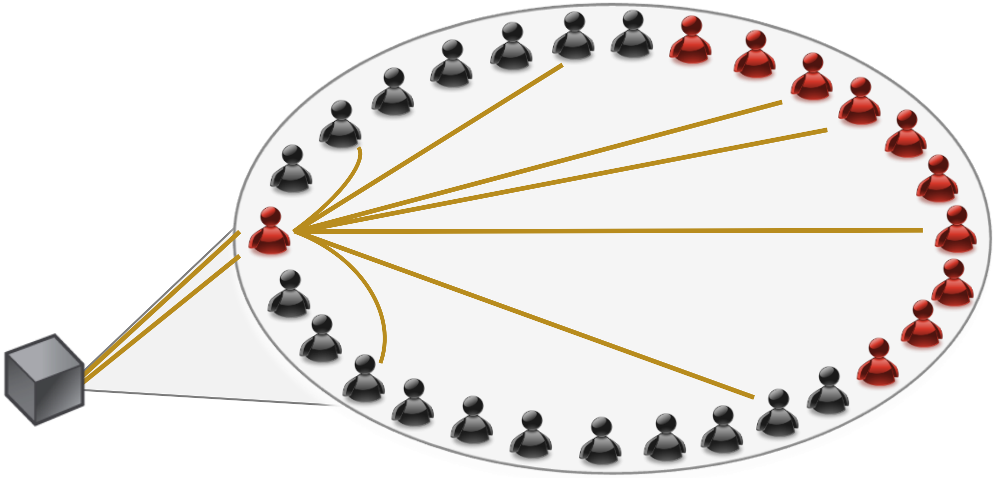
- ~~Block request sequence~~
- Hybrid approach using PEX
  - Order of peer addresses
  - Forged peer address



[1] Locher et al., *Free Riding in Bittorrent is Cheap*, HotNets 2006
[2] Locher et al., *Rescuing Tit-for-Tat with Source Coding*, P2P 2007

Reprise





Thank You!

Questions & Comments?

References

- P. Erdös and A. Rényi, *On Random Graphs*, Publicationes Mathematicae, 1959.
- R. Van der Hofstad, *Random Graphs and Complex Networks*, 2007.
- *BitThief – A Free Riding BitTorrent Client*. http://bitthief.ethz.ch
- Locher et al., *Free Riding in Bittorrent is Cheap*, HotNets 2006
- Locher et al., *Rescuing Tit-for-Tat with Source Coding*, P2P 2007

Encoding Bits Into a Permutation

- Encode a message M in a permutation
  - Represent M as a factorial number
  - $M = 10001_2 = 2210_!$ because $0 \cdot 0! + 1 \cdot 1! + 2 \cdot 2! + 2 \cdot 3! = 17 = M$
  - M is encoded into $\Pi = (3,4,2,1)$ as the Lehmer Code of Π is 2210.

- Lehmer Code
  - Counts the # swaps to get to Π
  - (1,2,3,4) ..2 swaps..
  - (3,1,2,4) ..2 swaps..
  - (3,4,1,2) ..1 swap..
  - (3,4,2,1) ..0 swaps..
  - (3,4,2,1)

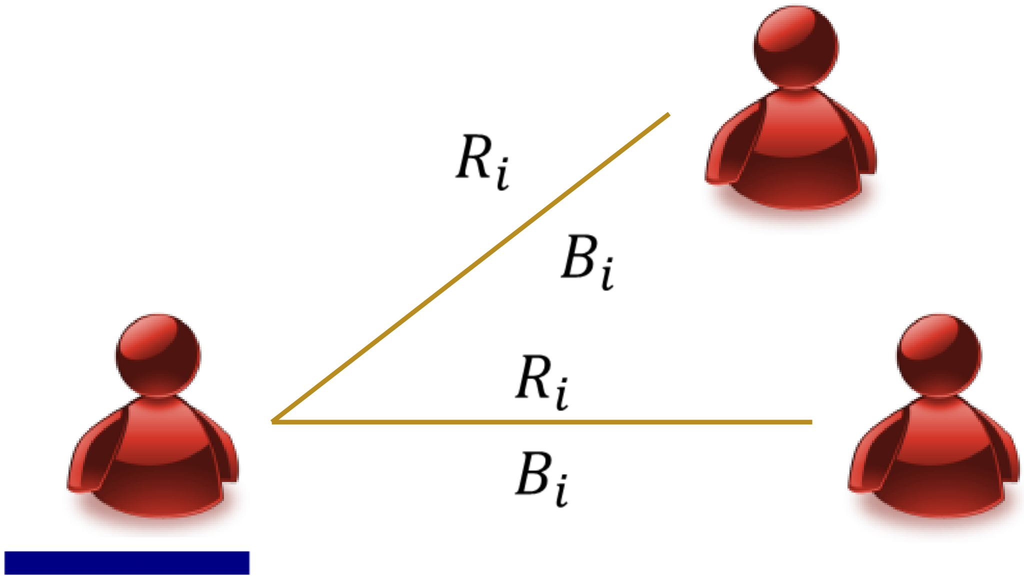
Proof of Lemma 3.2

- If each conspirer randomly connects to $8\,\frac{n}{c}\ln(nc)$ peers, then the subnetwork induced by the c conspirers is connected w.h.p.

- Proof:
- For each conspirer u, it holds that $E[|N_u^c|] = 8\ln(nc)$.
- $P[|N_u^c| < 4\ln(nc)] = P\left[|N_u^c| < \frac{E[|N_u^c|]}{2}\right] \leq e^{-\frac{E[|N_u^c|]}{2^2 2}} = \frac{1}{nc}$ (Chernoff)
- $P[\forall u \in C: |N_u^c| \geq 4\ln(nc)] > 1 - \frac{1}{n}$
- If each edge of a graph G with c nodes is present with probability $\ln(nc)/c$ then G is connected with probability $> 1 - \frac{1}{n}$ (Corollary from [Hofstad 2007])
- In such a graph G, all nodes have less than $4\ln(nc)$ neighbors w.h.p.
- Each conspirer implicitly chooses $4\ln(nc)$ random neighbors in the conspirer subgraph. ■

Broadcast under Individual Monitoring

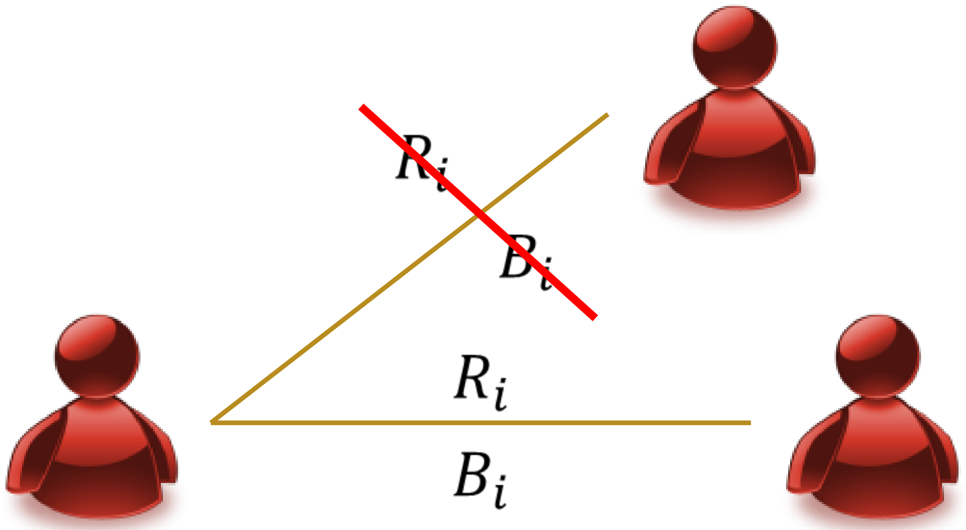
- Authority monitors individual communication links



- $|M| \in \Theta(m \log m)$ where m is the # of blocks

Broadcast under Complete Monitoring

- Authority monitors all connections, and correlates data
  - No under-reporting
  - No re-requesting



- Acquire $8\sqrt{n}\ln(nc)$ random blocks
- $|M| \in \Theta(\sqrt{m}\log^2 m)$

Broadcast under Stochastic Monitoring

- Regular peers choose their request order permutation according to a distribution C

- Authority classifies a peer as a conspirer if it uses a request order permutation Π with $p(\Pi) < \epsilon$

- Trade-off in the choice of threshold ϵ

  – Amount of hidden communication vs. False positives

Broadcast under Stochastic Monitoring

- Regular peers choose their request order permutation according to a distribution C
- Authority classifies a peer as a conspirer if it uses a request order permutation Π with $p(\Pi) < \epsilon$

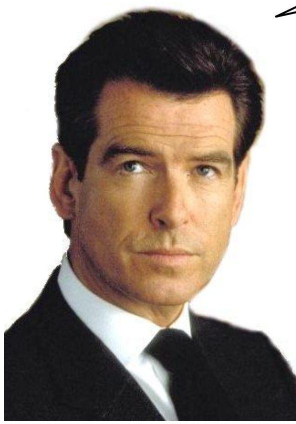
Algorithm 4 $ENC_{stochastic}$

```
i := 0;
repeat
    Π := ENC_order(M ⊕ K(i)||i);
    i++;
until p(Π) > ε
return Π;
```

- $\mathcal{K}$ is a deterministic PRG

Spy Rendezvous

In London, April’s a spring month.



..whereas in St.Petersburg we’re freezing our butts off.

MI6

CIA