

Drac: An architecture for Anonymous Low-Volume Communications

Claudia Diaz

K.U.Leuven ESAT/COSIC

G. Danezis, C. Diaz, C. Troncoso, and B. Laurie
(under submission)

Outline

- Motivation
- Drac system
- Evaluation
- Discussion
- Conclusions

Motivation

- Anonymity against global passive adversary
- Limited bandwidth or regular traffic
 - VoIP, IM (not web traffic)
 - Padding
- Peer-to-peer architecture
 - Scalability, no need for centralized directory server

Properties

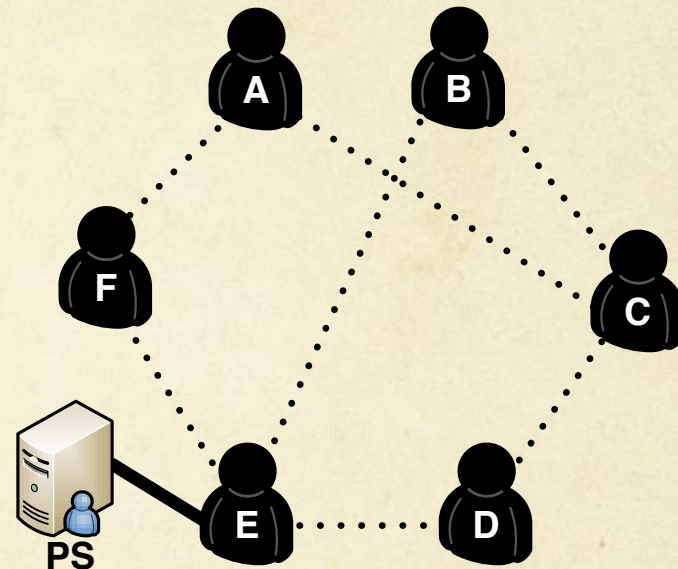
- Anonymity of certain relationships (not others)
- Difficulty concealing long-term, persistent relationships
 - Public information anyway
- Unobservability of communication events
- Use social network for routing
 - Sybil prevention
 - Trusting strangers?
 - Build incentives
 - Correlation between executions: robust anonymity sets (though smaller sets)

Drac System

- Social network, modeled as a graph, with N nodes (users):
 - edges connect *friends*:
 - public relationships, friends are trusted for routing, shared secret keys
 - privacy goal: conceal timing, duration, frequency of communications
 - *contacts*: other people Alice may wish to talk to (e.g., her doctor)
 - not trusted, know each other by pseudonyms, shared (long-term) secret key
 - relationship must be kept confidential by Drac
- Connections
 - Heartbeat (signaling)
 - Communication circuits
- Private Presence Server
- Epochs
 - Synchronous start and end of communications
 - Otherwise possible to correlate based on start and end of communications

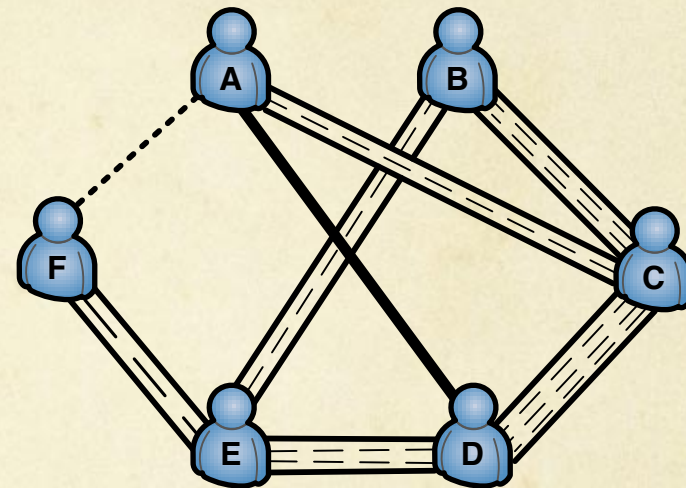
Heartbeat connections

- between each pair of friends
- very low bandwidth, bidirectional
- signal presence to friends
- signaling for creating communication circuits
- connections to presence server (depth D_p)
- observing heartbeat gives no additional info to adversary, since those relationships are considered “public”



Communication connections

- circuit of depth D
- circuit creation algorithm
 - Communication with contacts: Entry point
 - Bridges
- friends: direct links, D hops anyway
- Adversary: can see nr of circuits per link, and some bridges



--- circuit

▬ link

— bridge

Onion encryption

$$u_X \rightarrow u_Y \rightarrow u_Z \Rightarrow u_U \rightarrow u_V \rightarrow u_W$$

$$u_X \rightarrow u_Y : E_{k_{XY}} (E_{k_{XZ}} (E_{k_{XW}} (M)))$$

$$u_Z \Rightarrow u_U : E_{k_{XW}} (M)$$

$$u_V \rightarrow u_W : E_{k_{VW}} (E_{k_{UW}} (E_{k_{XW}} (M)))$$

Private presence server

- Private Presence server: Honest but curious
- There could be several of them
- User u_A has long-term identifier ID_A (user may have several, one per circle of contacts, so they cannot find out they know the same user)
- Contacts A and B share a key K_{AB}

Presence

- unlinkability between time periods (epochs), avoid long-term pseudonymous profiling: “*id du jour*” IDJ
- T published by Presence server

$$IDJ_A = H(T, ID_A)$$

- B sends this message to the PS:

$$E_{PK_{PS}}(IDJ_A, E_{K_{AB}}(E_B, g^{r_B}))$$

- If A wants to talk to B, she sends g^{r_A} to E_B (next epoch)
- session key: $k_{AB} = g^{r_A r_B}$
- update long term key: $K'_{AB} = H(k_{AB}, K_{AB})$

Epochs

- why needed
 - avoid traffic analysis based on looking at start/end of communications
- during an epoch: preparation for next epoch
- tradeoffs (duration epoch)
 - delay in starting a communication
 - overhead creating circuit, switching circuits during conversation
 - long-term disclosure

Evaluation

- Experimental setup
- Anonymity towards the presence server
- Contact communication anonymity
- Unobservability

Experimental setup

- Simulator implemented in python
- Topologies: small world, scale free, random
 - f friends on average (selected according to topology)
 - f randomly selected contacts
- Single epoch per experiment (no multiple epoch analysis)
 - heartbeat connections: between friends, and between end of presence circuit and presence server
 - communication circuits and bridges; adversary can see nr of circuits per link and distinguish bridges
 - 10% of users communicating with contacts (randomly selected)
- One sample per experiment:
 - contact communication anonymity
 - presence anonymity
 - contact communication unobservability

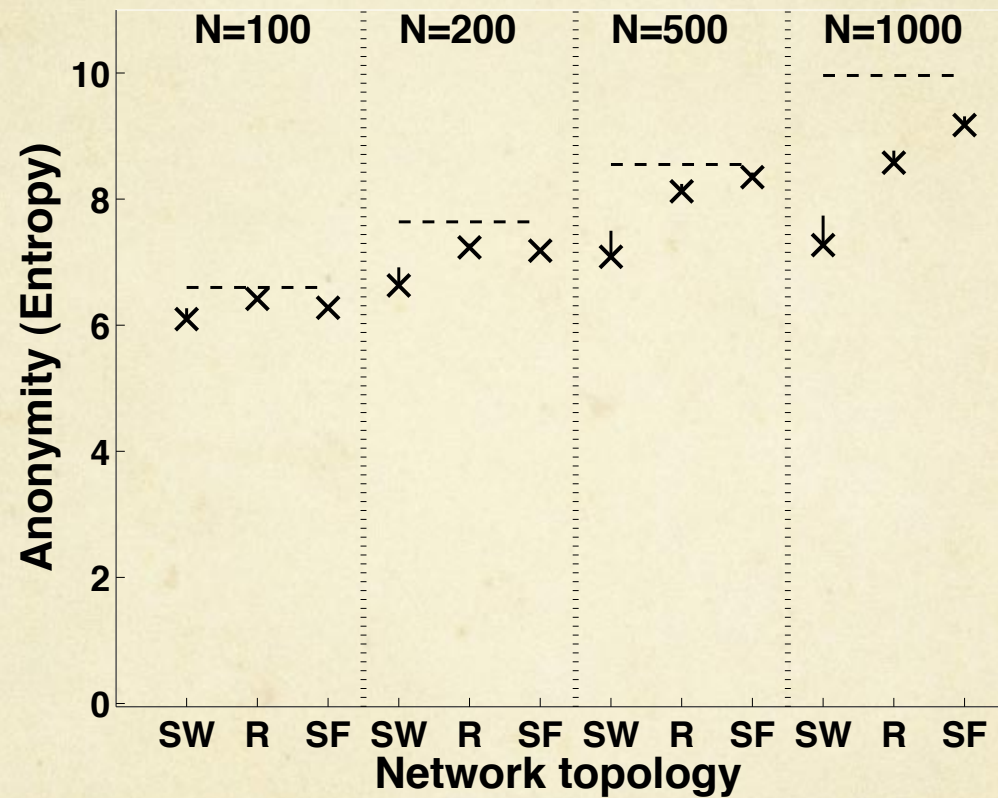
Anonymity towards the presence server

- start from connection to Presence Server (end of circuit)
- check all paths that lead to each of the initiators

$$\Pr_i[E_{PA}] = \frac{P_i}{\sum_{j=1}^N P_j}, 1 \leq i \leq N$$

$$H_A = - \sum_{i=1}^N \Pr_i[E_{PA}] \log_2 \Pr_i[E_{PA}]$$

Results: Topology

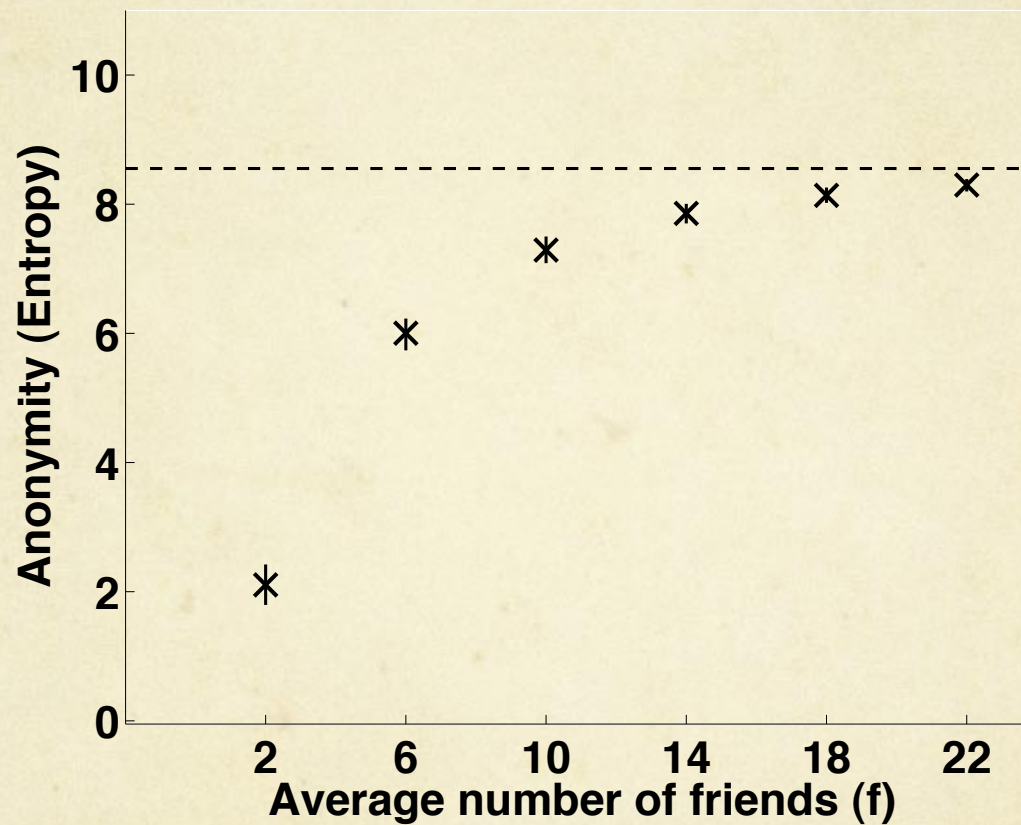


Parameters: 10 friends, $D_p = 3$

Grenoble, March 8, 2010

15

Results: number of friends

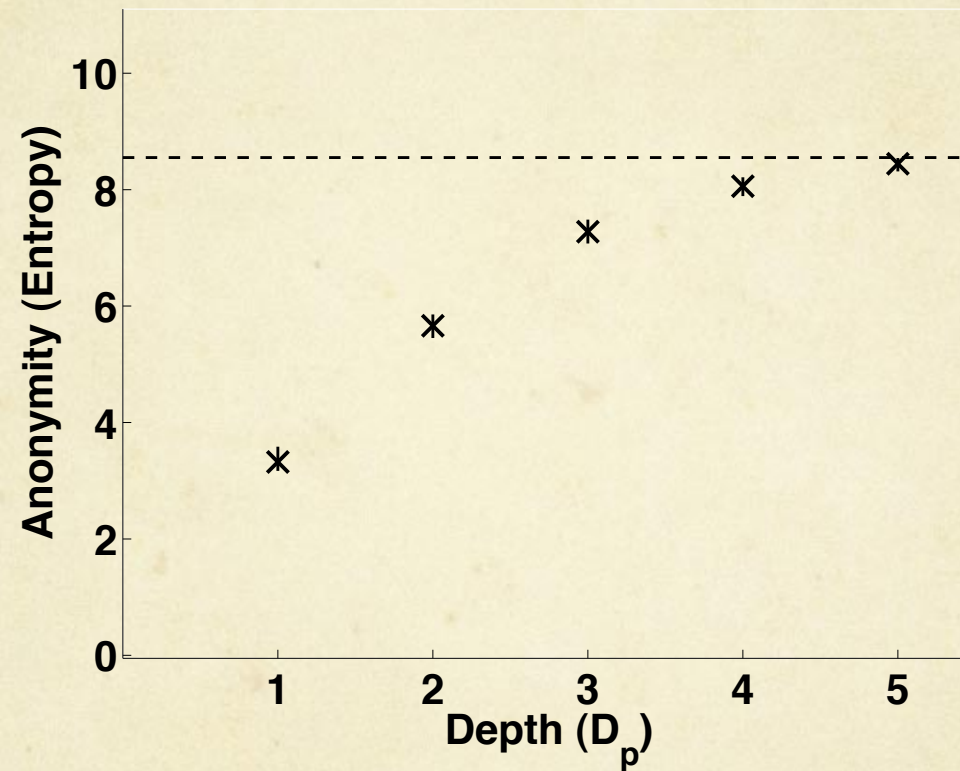


Parameters: SW net, $N = 500$, $D_p = 3$

Grenoble, March 8, 2010

16

Results: circuit depth

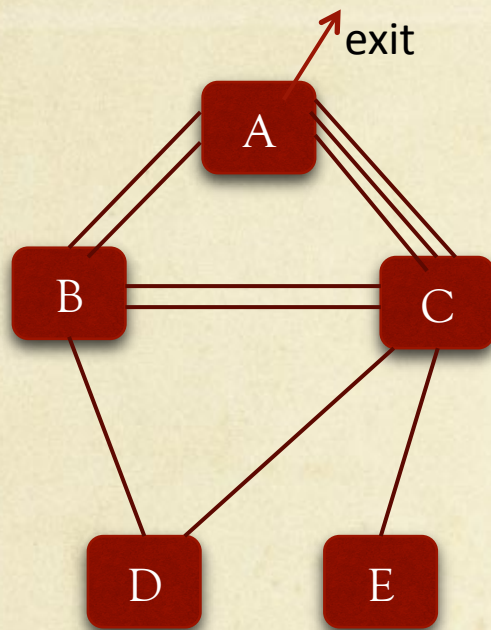


Parameters: SW net, $N = 500$, 10 friends

Contact communication anonymity

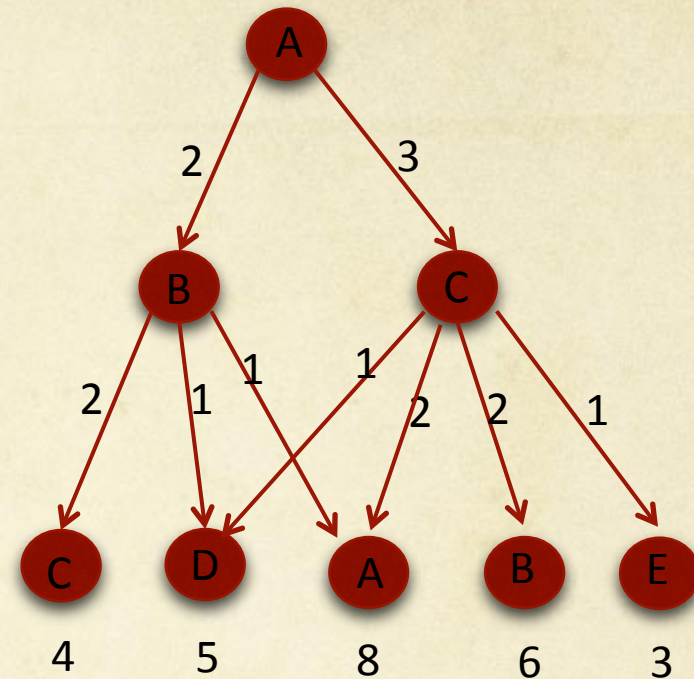
- Assume all bridges observable
- Evaluate anonymity of each half of circuit separately, starting from bridge (no end-to-end anonymity)
 - no certainty A is communicating: not straightforward to compute anonymity
 - anonymity: who is Alice talking to (assuming I know Alice is talking)
 - unobservability: is Alice talking?
- difference with presence: nr of circuits per link visible

Example



true paths:

- A-C-B
- B-C-A
- C-A-B
- D-B-A
- E-C-D



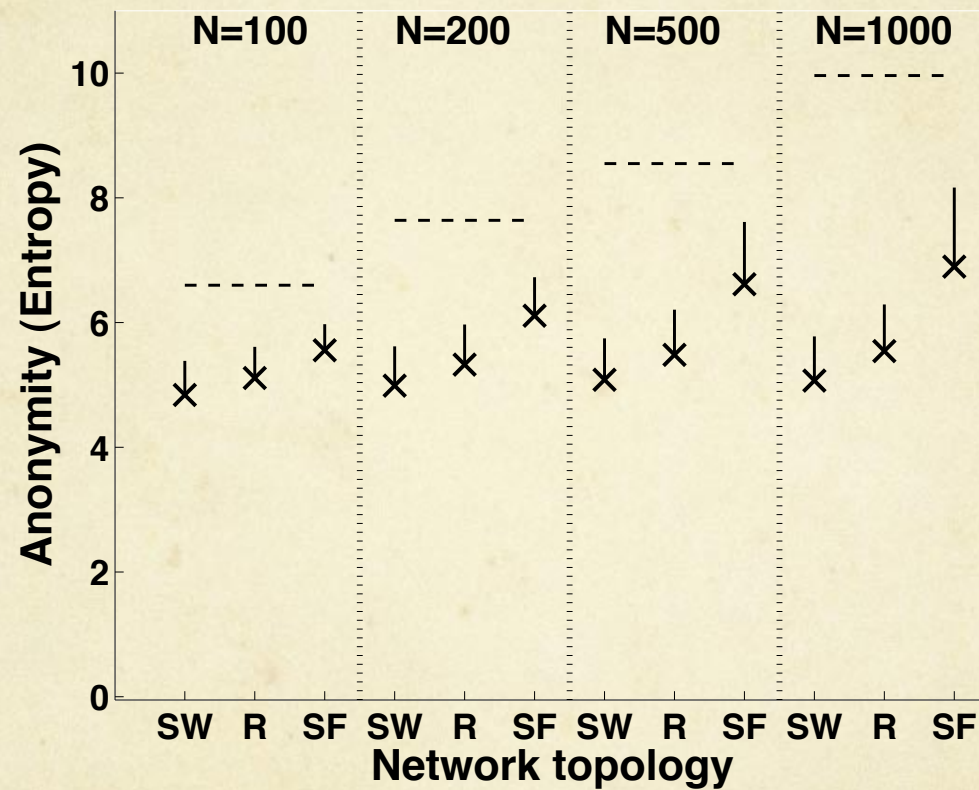
possible paths:

- C-B-A (x4)
- D-B-A (x2)
- A-B-A (x2)
- D-C-A (x3)
- A-C-A (x6)
- B-C-A (x6)
- E-C-A (x3)

Prob (caller, exit A):

- $\Pr(A) = 8/26 = 0,3$
- $\Pr(B) = 6/26 = 0,23$
- $\Pr(C) = 4/26 = 0,15$
- $\Pr(D) = 5/26 = 0,19$
- $\Pr(E) = 3/26 = 0,12$

Results: topology

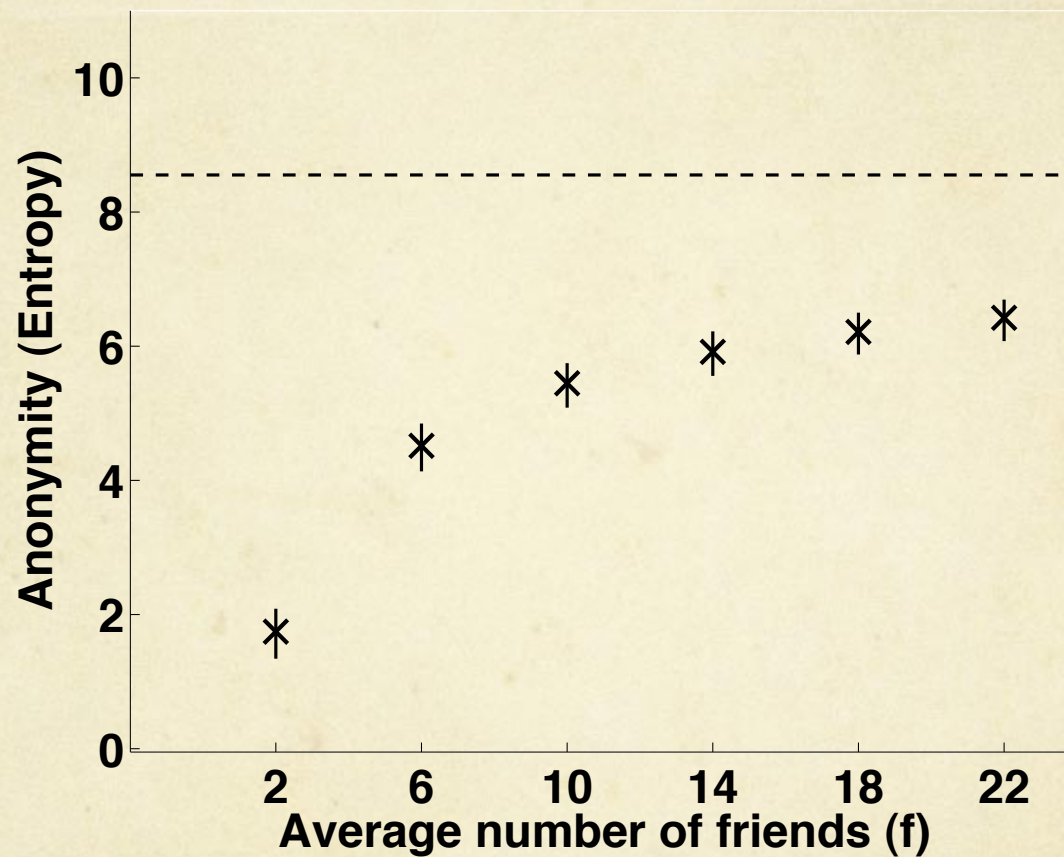


Parameters: 10 friends, $D = 3$

Grenoble, March 8, 2010

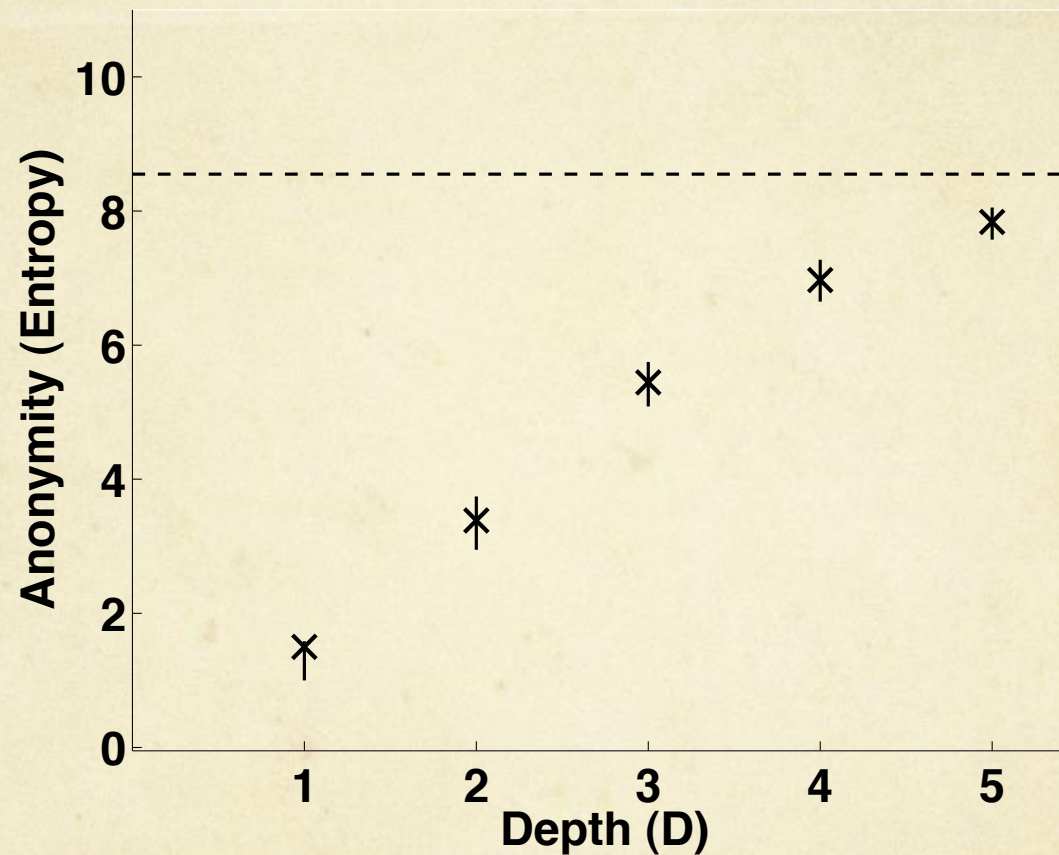
20

Results: number of friends



Parameters: SW net, $N = 500$, $D = 3$

Results: circuit depth



Parameters: SW net, $N = 500$, 10 friends

Unobservability

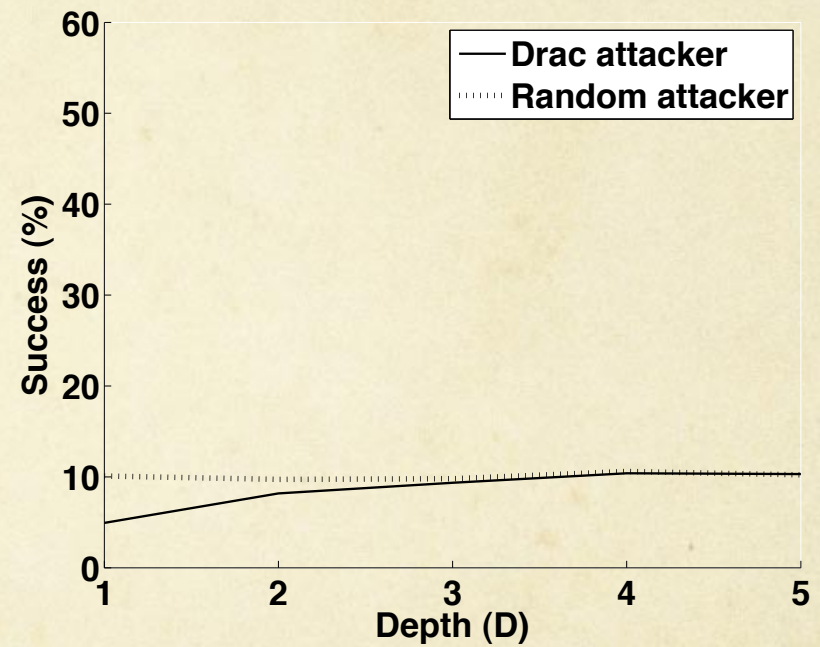
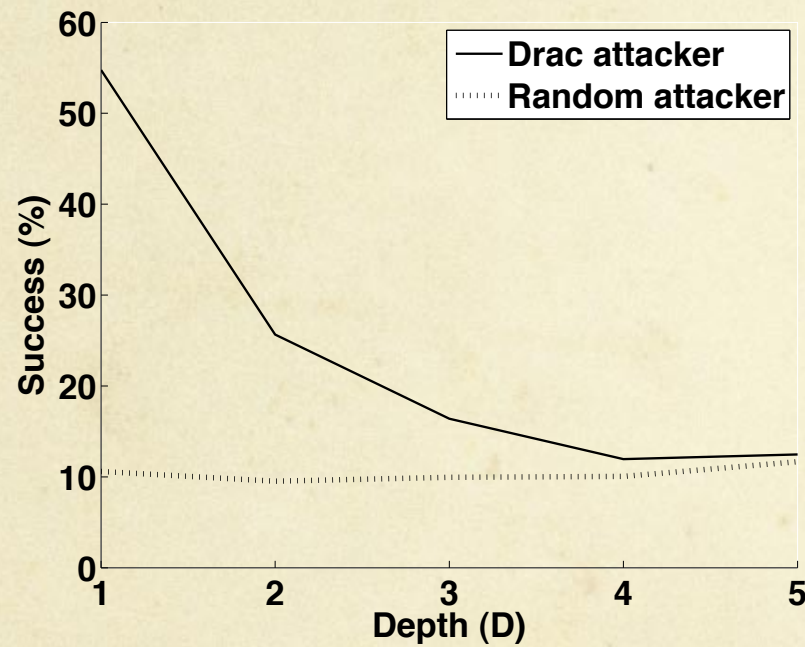
- communications with friends: fully unobservable
- communications with contacts: bridges observable
 - C : total nr of contact communications (assume known by adversary)
 - \mathcal{E} : set of active entries (assume identifiable by adversary)
 - if no coincidence of entries, then $|\mathcal{E}| = 2C$
 - $\Pr_i[E_j]$: probability that u_i is the user communicating through E_j
 - $\Pr[u_A]$: prob u_A is communicating through one of the active entries

$$\Pr[u_A] = \frac{\sum_{j=1}^{|\mathcal{E}|} (\Pr_A[E_j] \prod_{k=1, k \neq j}^{|\mathcal{E}|} (1 - \Pr_A[E_k]))}{\sum_{j=1}^{|\mathcal{E}|} (\Pr_A[E_j] \prod_{k=1, k \neq j}^{|\mathcal{E}|} (1 - \Pr_A[E_k])) + \prod_{k=1}^{|\mathcal{E}|} (1 - \Pr_A[E_k])}$$

Unobservability

- Adversary constructs set S with top 2C users (highest $\Pr[u_A]$)
- Random adversary: constructs set R with 2C random users
- Select user u_A who *is* communicating with a contact
 - Test if u_A in S and if u_A in R
- Select user u_Z who *is not* communicating with a contact
 - Test if u_Z in S and if u_Z in R

Results



Parameters: SW net, $N = 500$, 10 friends, $C = 25$

Grenoble, March 8, 2010

25

Discussion

- Regular, or low-bandwidth, traffic: VoIP, IM, etc. (not web!)
- depth of connections: security parameter, tradeoffs
- trust model: DoS, sybil, no centralized directory, scalability, avoids network discovery and random sampling attacks
- incentives to stay online, route traffic
- stable anonymity sets
- anon depends on mixing properties of social graph
- long-term relationships: no way to conceal them
 - actual comm events are unobservable

Conclusions

- Mix of properties
 - design seems promising
- Open questions
 - correlations between many epochs
 - MCMC for proper computation of probability distributions
 - unobservability metrics, deniability?
 - resistance to corrupted nodes