Visualizing Privacy Implications of Access Control Policies in Social Networks

Mohd Anwar*, Philip W. L. Fong*, Xue-Dong Yang†, Howard Hamilton†

* University of Calgary, Alberta, Canada
† University of Regina, Saskatchewan, Canada

DPM 2009

Motivation

- In social networks, privacy settings allow users to choose access control policies

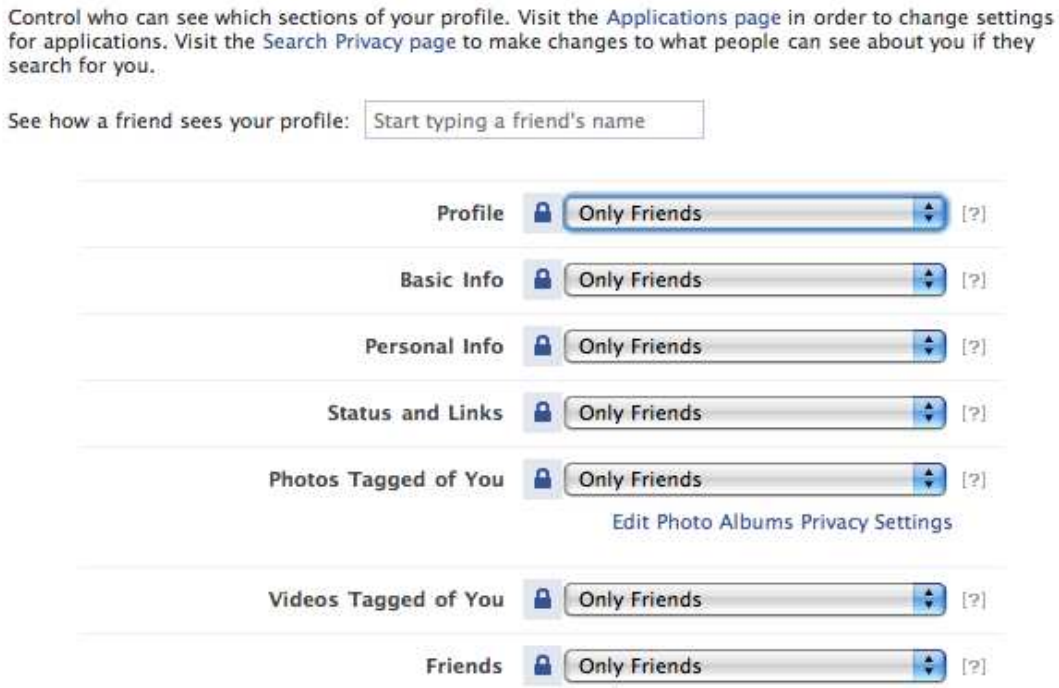


Figure: Privacy Setting in Facebook

- What are the privacy implications of these policies?
- How do we help users assess topology-based policies?

Related Work

- Privacy for Impression Management: Goffman 1961, Patil & Kobsa 2003
- Privacy Preservation Model for Social Networks: Fong, Anwar, & Zhao 2009
- Generating Social Graph: Chakrabarti *et al.* 2007
- Visualization (Social Graph/Security Policies): Freeman 2000, Heer & boyd 2005, Reeder *et al.* 2008

Outline

- Privacy in Social Networks
- User Specified Policies in Facebook-style Social Network Systems (FSNS)
- Topology-based Policies
- **Reflective Policy Assessment (RPA)**
- **Tool Support for RPA**
- Issues & Discussions
- Work in Progress

Privacy in Social Networks

What is Privacy?

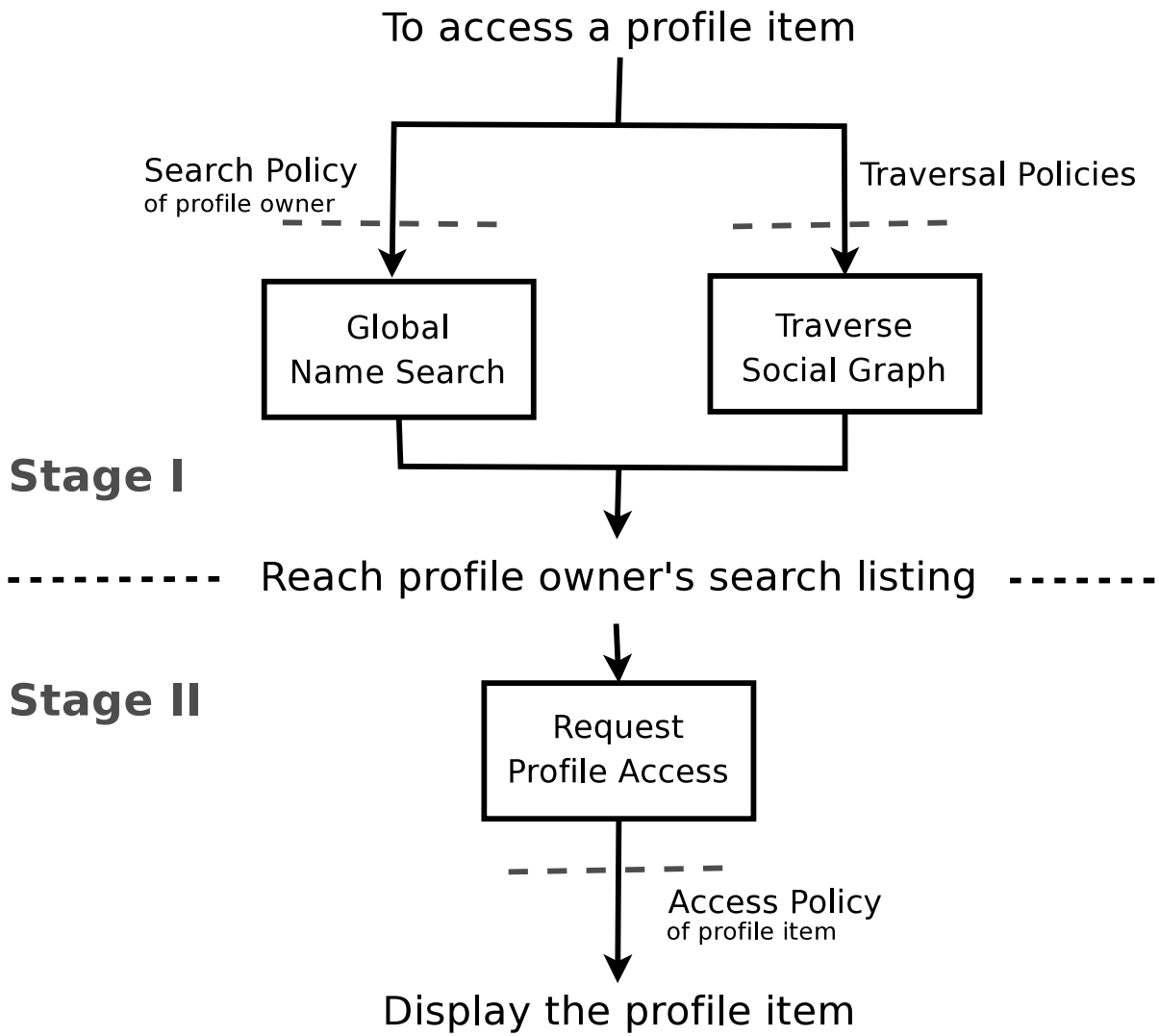
- Purpose of privacy is impression management
  - Control the impression that other people form
- Control over what impression one wants to convey to whom
  - What profile items to present to whom?
  - e.g. disclose the sorority photos to only friends, but siteseeing photos to everybody

Privacy and Access Control Policies

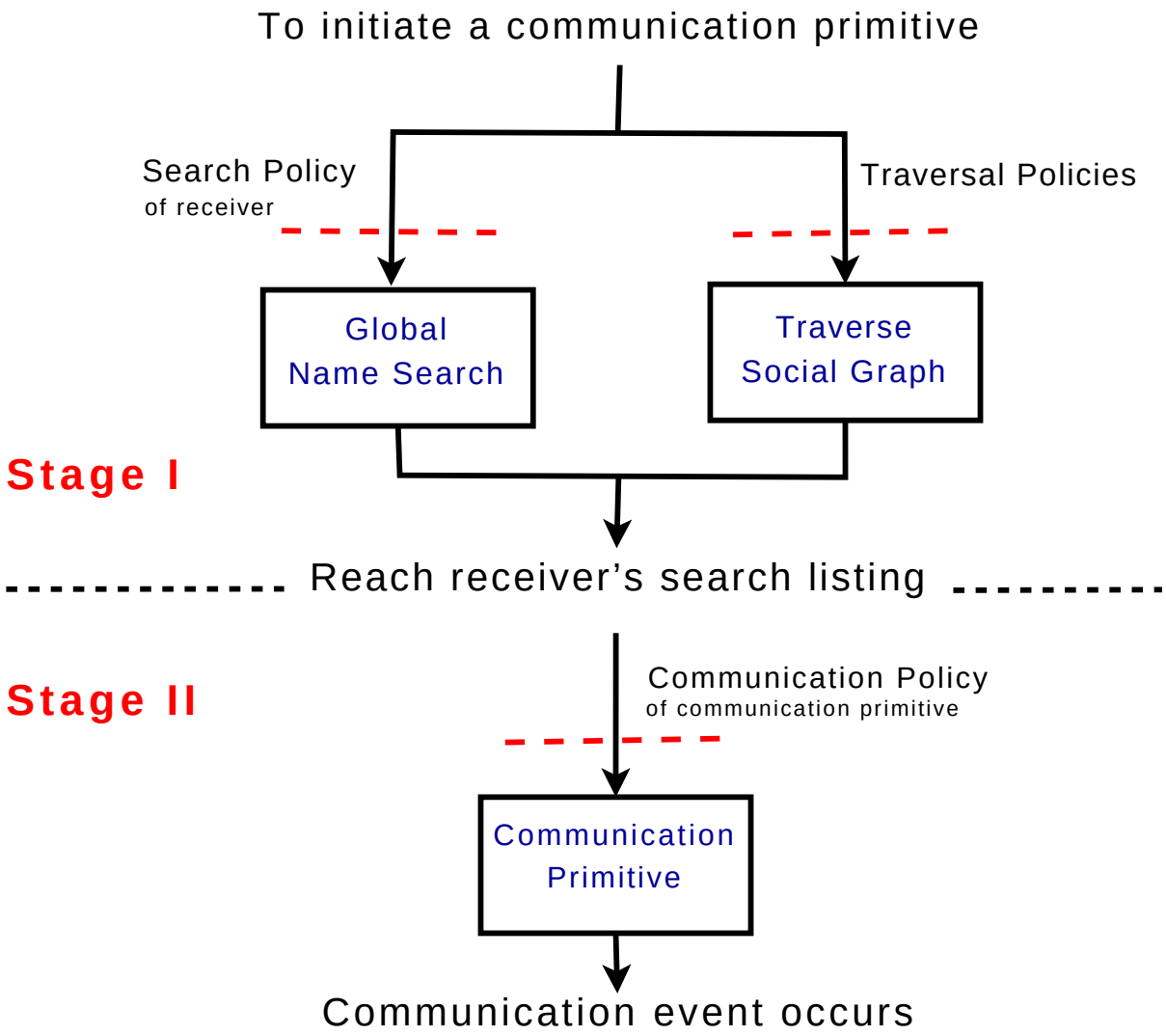
- Impression is conveyed according to relationship
- Relationship can be encoded into the topology of a graph (e.g. social graph)
- Therefore, topological access control policies help users control impression

User Specified Policies in FSNS

Search, Traversal, and Access Policies



Communication Policies



Topology-based Policies

- Facebook offers more general topology-based policies: " only friends" and "friends of friends"
- Richer form of acquantance relationships can be represented:

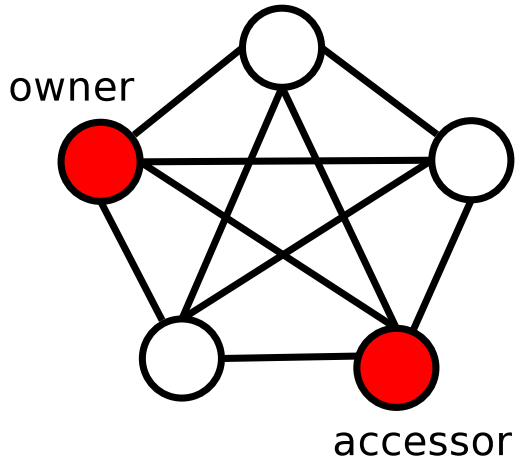


Figure: 5-clique

Topology-based Policies

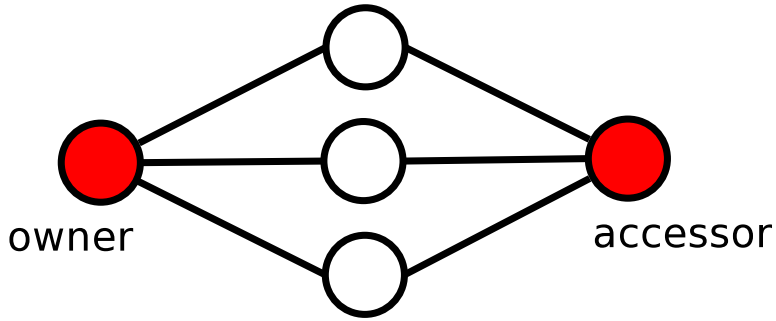


Figure: 3 common-friends

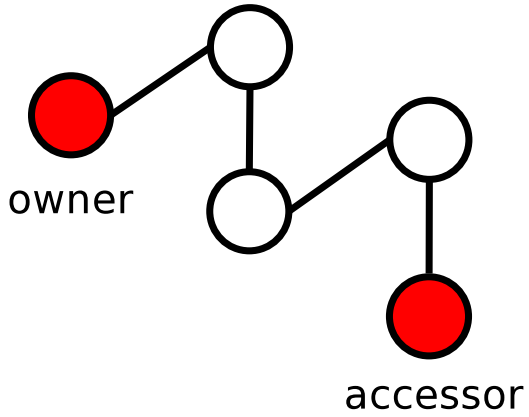


Figure: $distance_4$

Anti-monotonic Policies

- Under an anti-monotonic policy, access becomes more difficult as the social graph becomes denser
- Disclosure of information only to those who do not know you well
  - e.g. stranger ($\neg distance_k$)

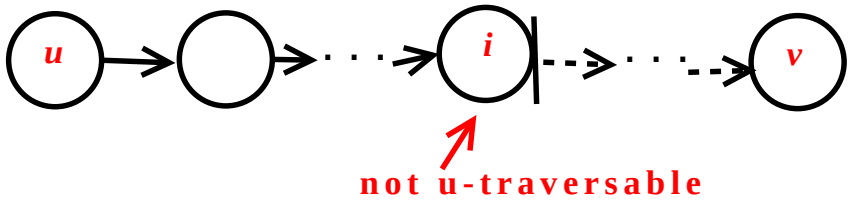
Reflective Policy Assessment (RPA)

Idea of RPA

- A mirror allows us to see what others see when they look at us
- To create a desired impression, we repeatedly look into the mirror and adjust our getup
- The process of formulating access control policies is similar to what it takes to create a desired look
- A user needs to repeatedly assess and adjust their policies
- We propose that a profile owner inspect her profile from the view point of a potential accessor

Privacy Dilemma with RPA

- A user must begin with identifying a potential accessor who is of interest to her.
- A potential accessor may not want her identity to be disclosed to the user conducting the policy assessment.
- This dilemma is rooted in the asymmetric nature of trust.



- To address this dilemma, we propose approximating the extended neighbourhood of a user.

Tool Support for RPA

Policy assessment is nontrivial

- Authorization depends on the existing topology of social graph
- Social Graph constantly changes, so do privacy needs
- It is nontrivial to comprehend the privacy consequence of adjusting privacy settings

Proposed Tool

To facilitate RPA, we devise a tool that

- visually depicts the extended neighbourhood
- allows the profile owner to point to any user in the extended neighbourhood as a potential accessor
- The tool displays a succinct representation of the profile, as seen from the eyes of the potential accessor

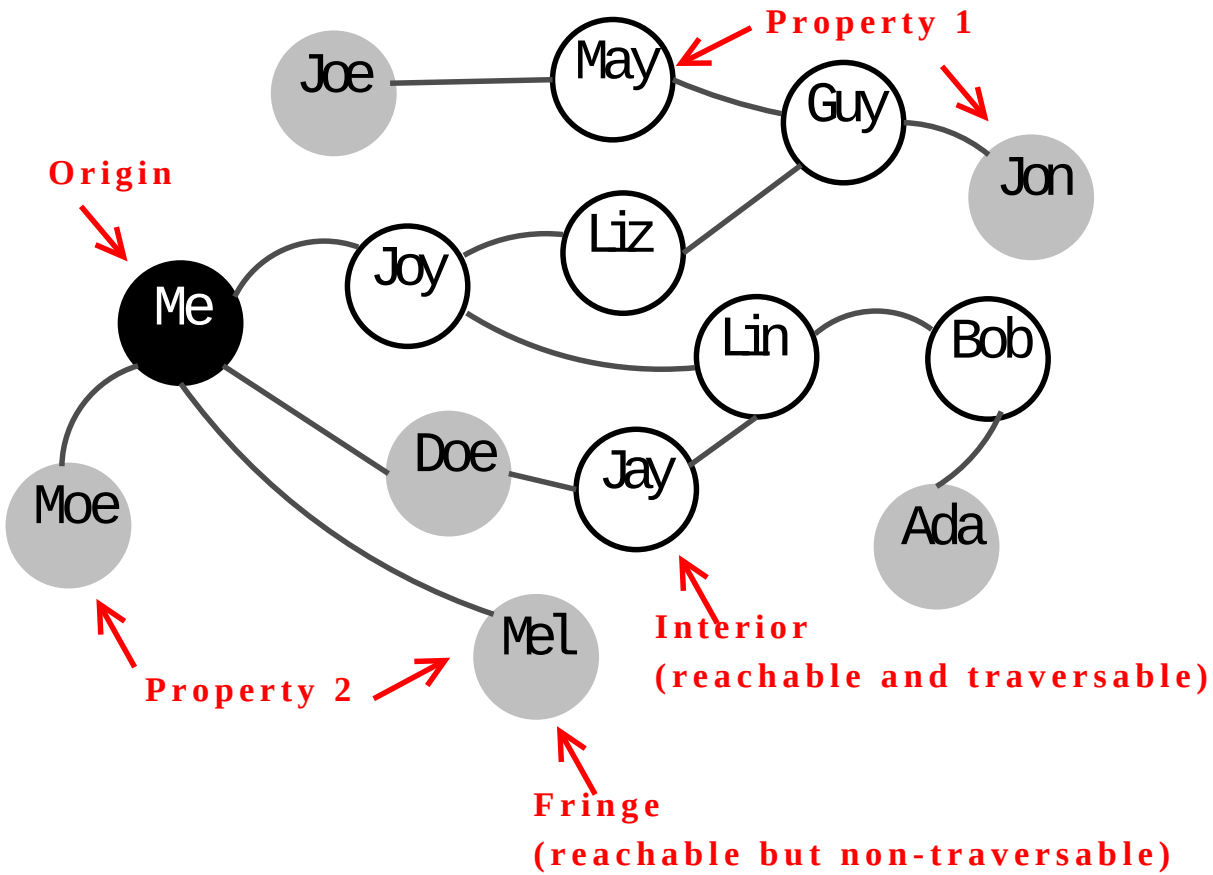
Properties of Social Graph

We use the following properties to establish the correctness of algorithm for generating social graph:

- *Property 1.* Given an origin, **every neighbour of an interior node is reachable**, and thus, no hidden edge can have an interior node as an end.
- *Property 2.* Suppose an origin is given. By definition, at least one end of each visible edge is an interior node. Therefore, **no visible edge can join two fringe nodes**.

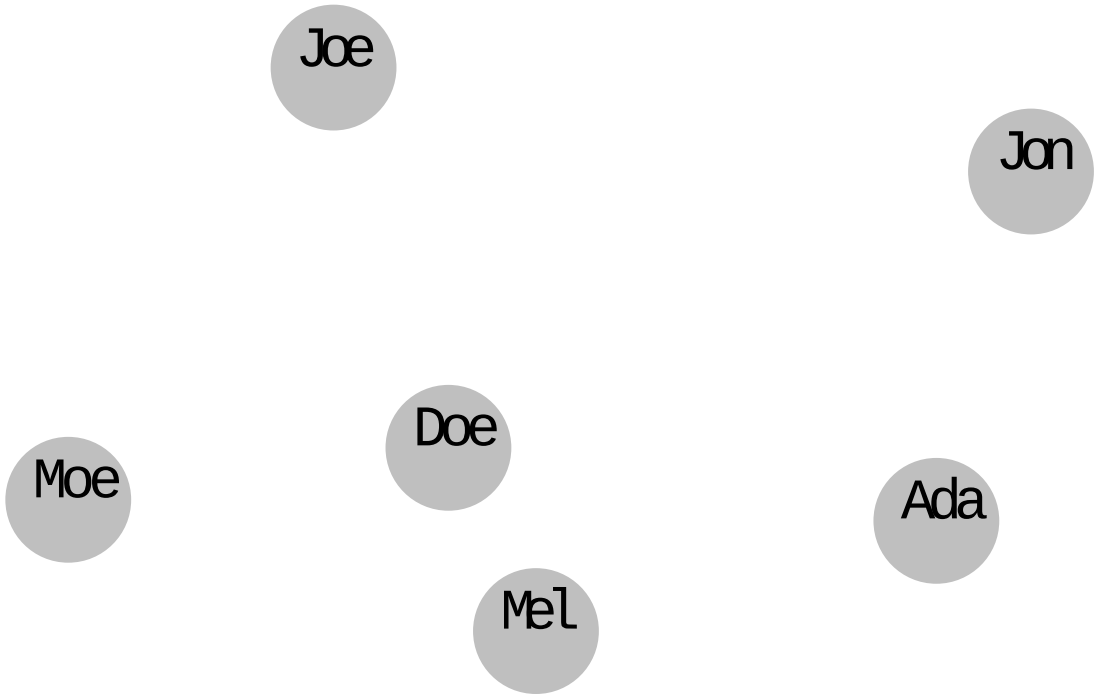
Graph Generation Algorithm

1. Construct a graph consisting of all reachable nodes and visible edges



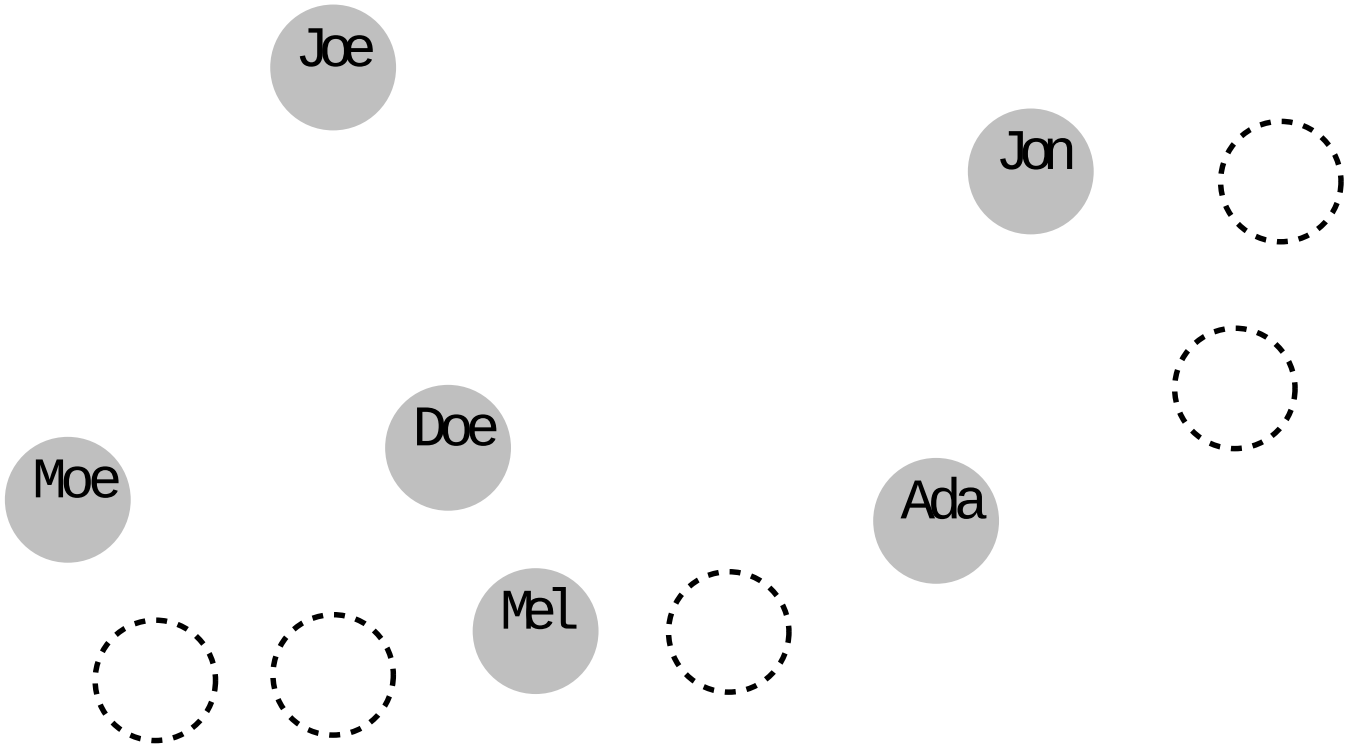
Graph Generation Algorithm

2. Temporarily remove all interior nodes and visible edges.



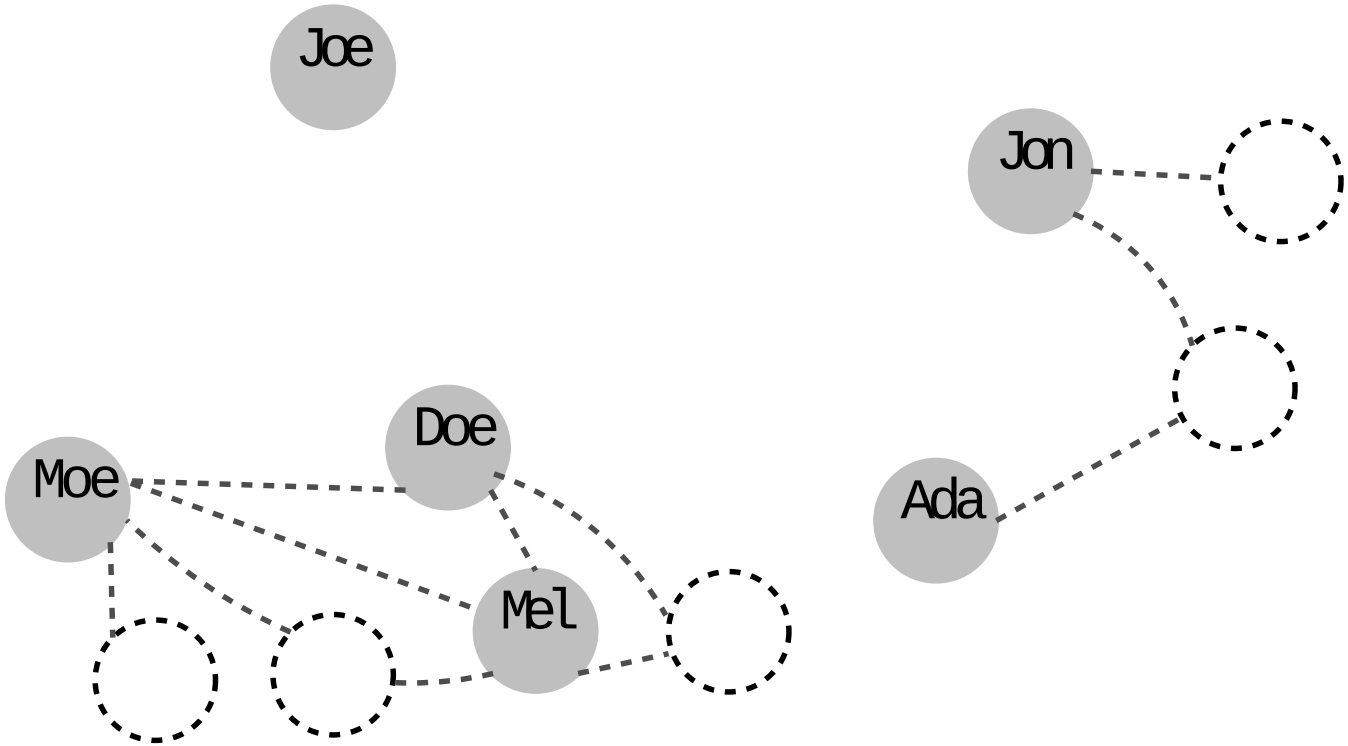
Graph Generation Algorithm

3. Add n "synthetic nodes" in the social graph.



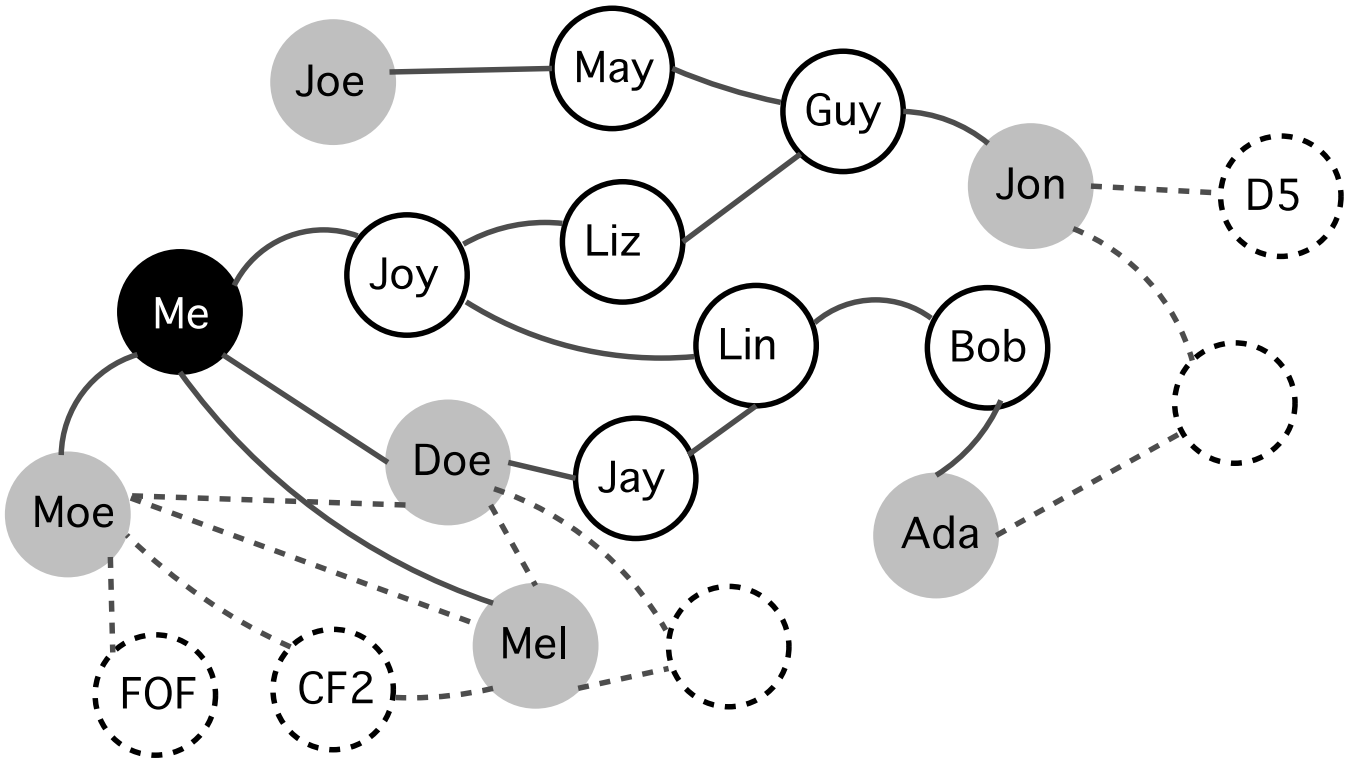
Graph Generation Algorithm

4. Use **R-MAT** (*Chakrabarti et al. 2007*) to randomly generate m "synthetic edges"



Graph Generation Algorithm

5. Add back the interior nodes and visible edges removed in step 2, and return the resulting graph.



Prototypical visualization tool

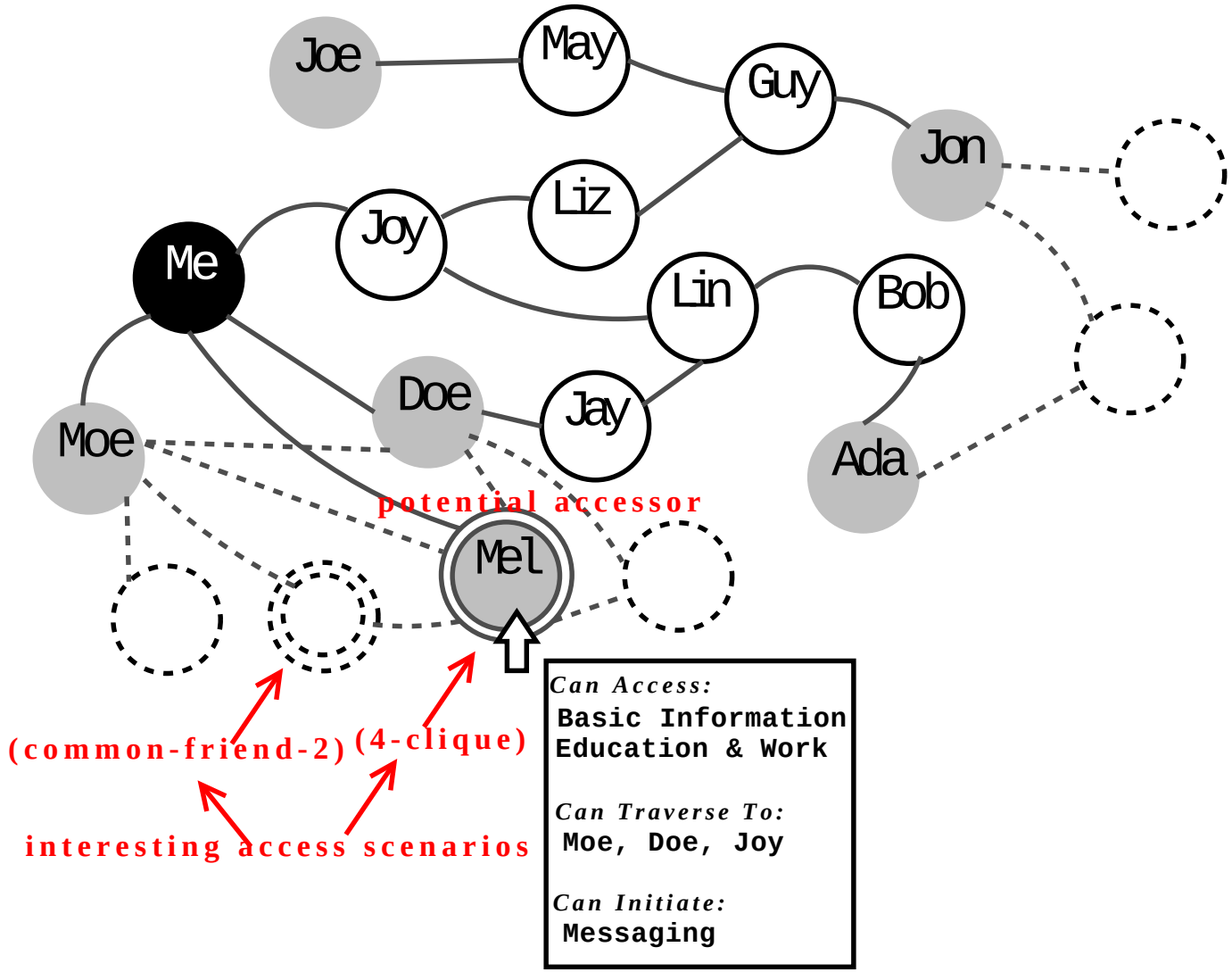


Figure: The black node is the profile owner, the double-circled node depicts a potential accessor representing an interesting access scenario.

Issues & Discussion

No Information Leakage by RPA

- Visible edges are already accessible by the profile owner.
- Hidden edges do not take part in the policy assessment.
- Topological information revealed by RPA is either **already available** (visible edges) or **anonymized** (synthetic edges).

RPA Recommends Access Scenarios

- Our visualization tool recommends nodes (potential accessors) that represent interesting access scenarios
- Based on the various profile appearances, partition the nodes into equivalence classes.
  - Two nodes that (both satisfies and violates the same policy predicates) produce the same profile appearance belong to the same access scenario.
- Each equivalent class represents a distinct access scenario.
- The tool will selectively highlight a node if it corresponds to a novel access scenario.

Work in Progress

We are in the process of addressing the following set of open questions:

(a) How effective is our tool?
- A user study is in order.

(b) How many graphs does one need to generate in order to gain enough confidence on the policies under assessment?
- A probabilistic analysis needs to be done.

(c) How well does our tool perform in a very large extended neighbourhood? –
- The profile owner needs not conduct assessment on every node (just one per equivalent class).
- Apply *focus + context* technique on a hyperbolic plane to effectively render a large neighbourhhod

Questions & Comments

Thanks!

Mohd Anwar
Post-Doctoral Fellow
Computer Science Department
University of Calgary
manwar@ucalgary.ca