



A Lightweight Privacy Preserving SMS-based Recommendation System for Mobile Users (ACM RecSys 2010)

Elisa Baglioni, Luca Becchetti, Lorenzo
Bergamini, Ugo Colesanti, Luca Filipponi,
Giuseppe Persiano and Andrea Vitaletti



Motivation

Social matching

The image shows a Facebook profile for Daniel Hui. On the left sidebar, the 'Overview' section is visible, with a blue arrow pointing to the '36 Mutual Friends' link. The main content area shows wall posts from Daniel Hui and his friends. On the right, the 'Browse Friendships' section is highlighted with a large red oval, showing a list of mutual friends and a search bar. Below this, there are sections for 'Friends' Events' and 'Sponsored' ads.

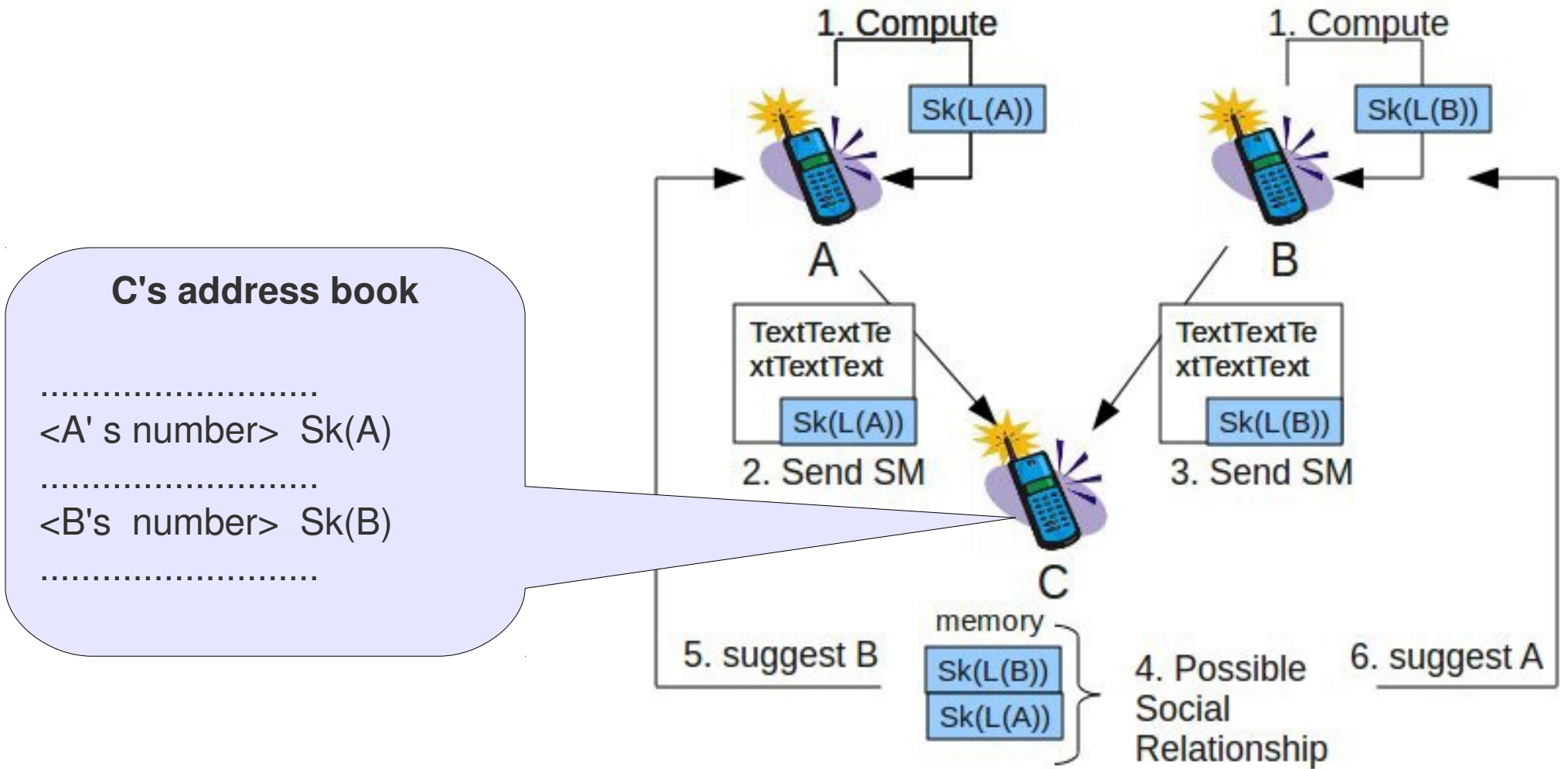
FRIENDSHIP is a KEY CONCEPT
RECOMMENDING FRIENDSHIP IS A KEY FEATURE

Goal

- Social matching of mobile phone users
- Users profiled according to their communication patterns (+ private profile)
- Not relying on a social networking app (e.g., Facebook)
- Fully decentralized
 - Profile info only exchanged between users
 - No central coordination (a novelty)
 - Profile info piggybacked within SMS body
 - Strong constraints on space
 - Viable in developing countries



In a nutshell...



Sk(X) = a sketch, i.e., a compact representation of user X's profile



Approach and Results

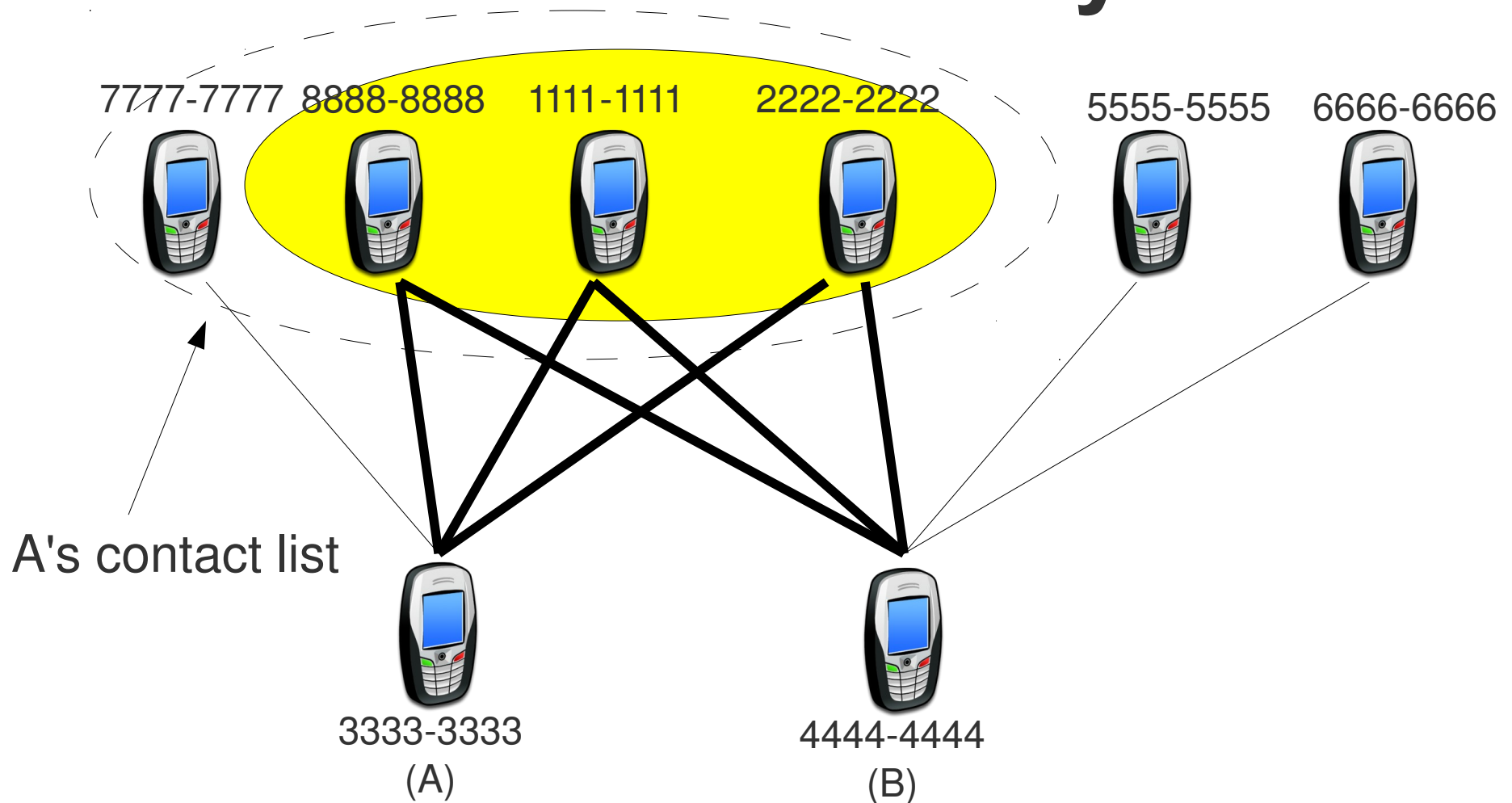
November 22nd – 23th

Perada Workshop

Issues

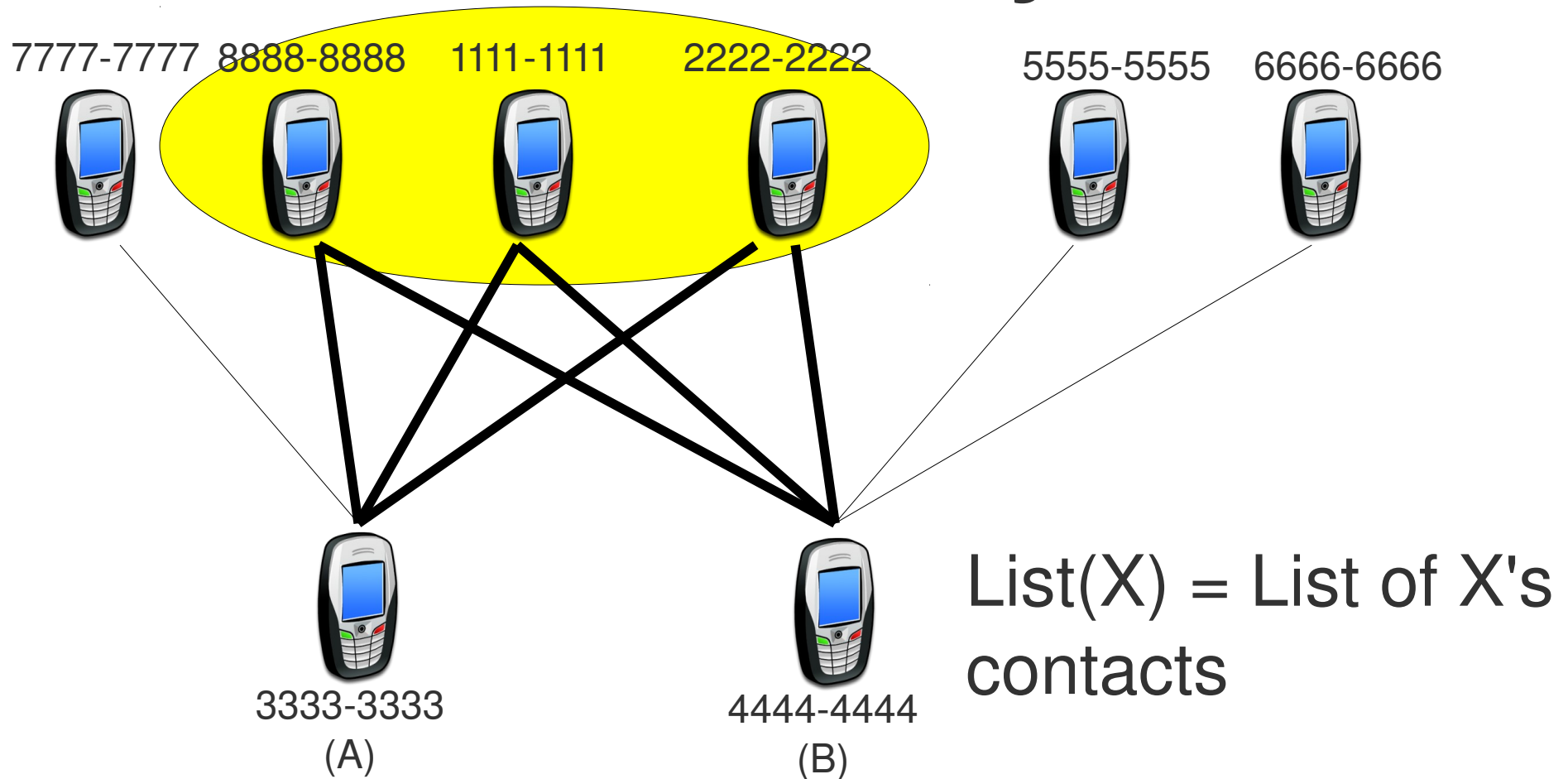
- **Profiles and user similarity**
 - How are profile sketches computed?
 - How is similarity between users defined?
- **Efficiency**
 - Space required (on user terminal)
 - Bandwidth (SMS body)
- **Privacy**
 - Call log/contact list contain sensitive info

User “similarity”



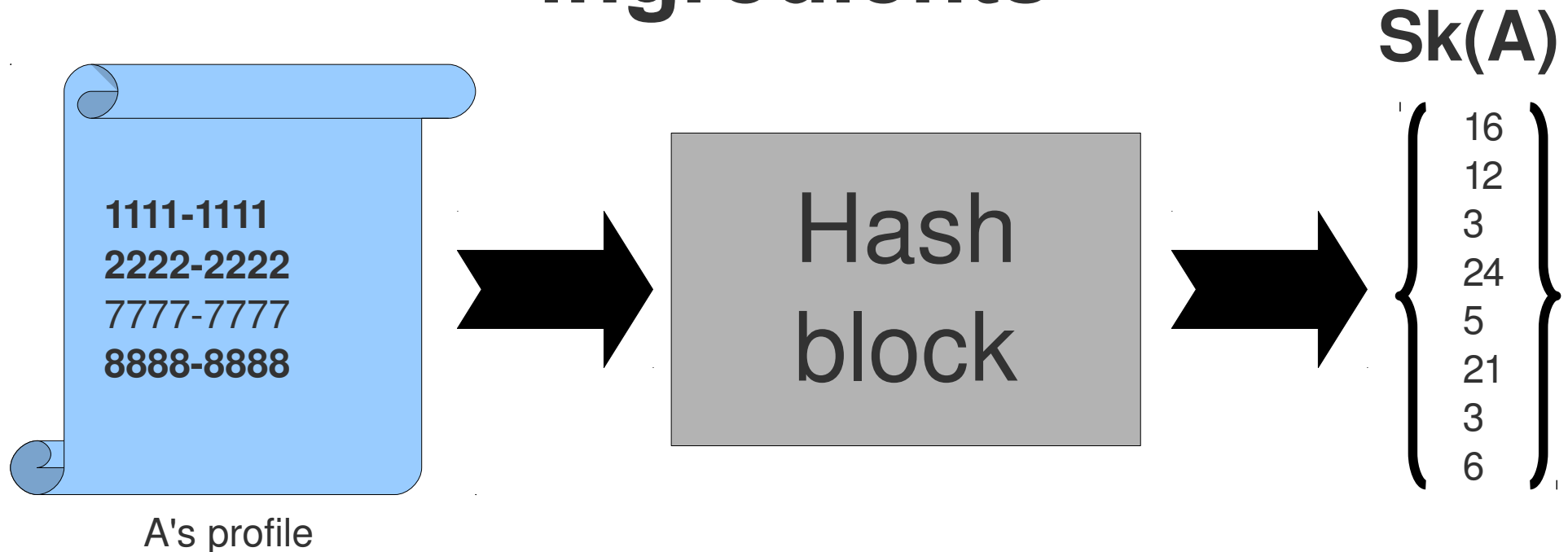
- Intuitively, 3333-3333 and 4444-4444 are similar in the social network... “mutual friendship”

User “similarity”



- We use Jaccard coefficient:
- $J(A, B) = \frac{|List(A) \cap List(B)|}{|List(A) \cup List(B)|}$

Computing profile summaries - ingredients






- Profiles viewed as feature sets
- Use of hash functions from special families [Broder et al. 00] to map feature sets into compact summaries $Sk(\cdot)$ so that:
 - Given $Sk(A)$ and $Sk(B)$ estimating $J(L(A), L(B))$ is inexpensive
 - Every $Sk(\cdot)$ is a (small) set of integers (e.g., 40/80 bytes)

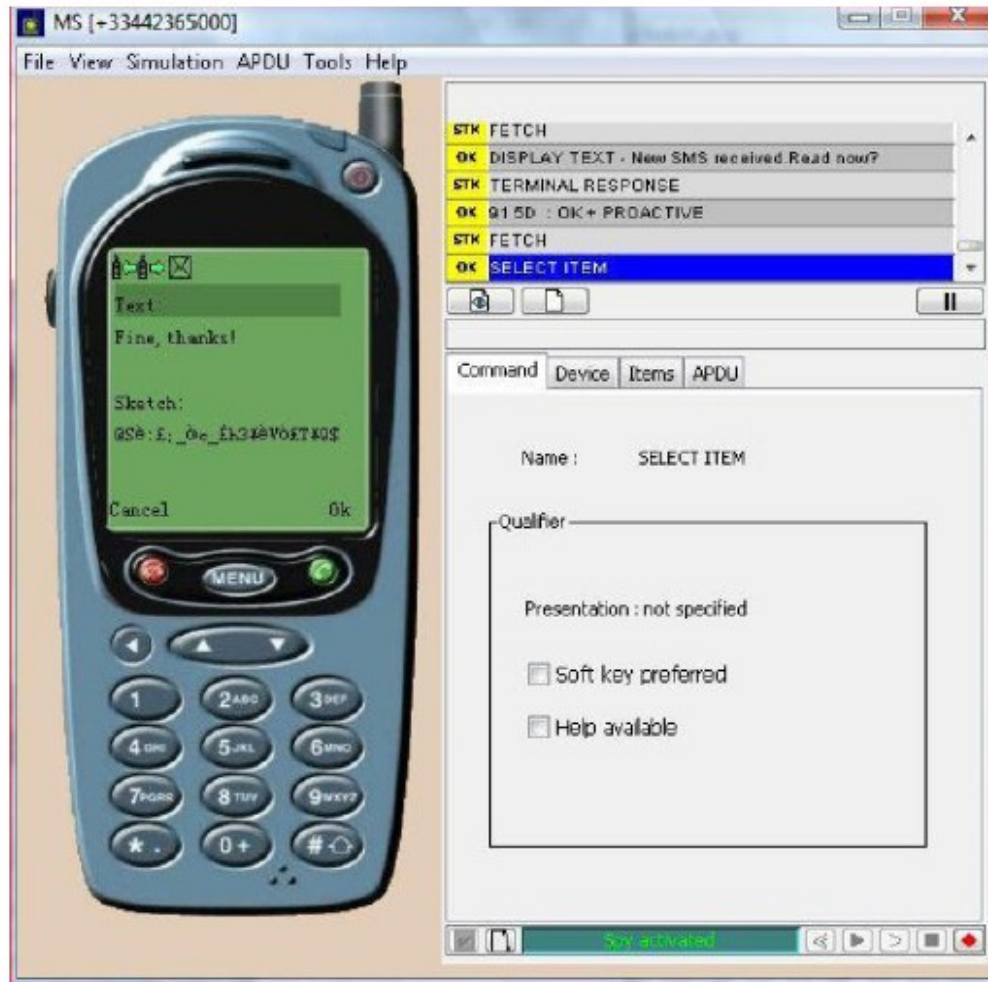
Preserving privacy

- User profiles are never explicitly exchanged
 - Only their sketches are
 - Basic technique of [Broder, Mitzenmacher 00]
may leak info
- Possible to make the protocol secure by public key encryption scheme
 - Homomorphic encryption scheme
 - Users exchange encrypted versions of their $Sk(\cdot)$'s
 - Clearly, number of SMS messages needed for A to send her sketch to B increases
 - Less than 1.4 messages in expectation w/o encryption

Experimental Analysis

- Reality Mining Project (MIT) [Eagle et al. 06]
 - largest mobile phone experiment in academia
 - thousands hours of continuous data on daily human behavior
 - About 100 users monitored
 - information on call logs
- Goals of the experiments
 - Is the Jaccard coefficient effective to detect social ties in the scenario we consider? 
 - Are techniques we use to estimate the Jaccard coefficient feasible wrt stringent space constraints? 
 - Are our suggestions of potential interest? 

Implementation on STK





Conclusions

Possible scenario

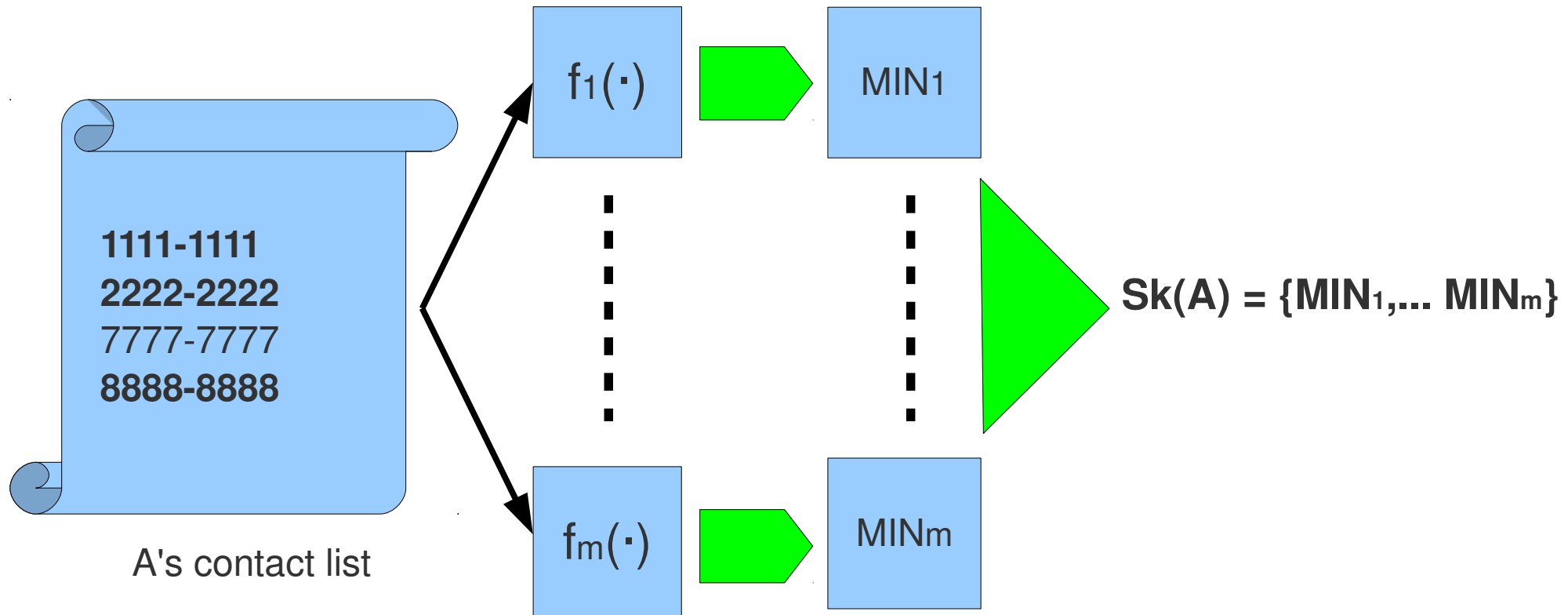
- Goal: deploy a cheap (for both operator and customers) social networking platform
- When/where: lack of network infrastructure but mobile phone network coverage
 - For example, many rural areas in Africa
- Mobile telephony plays key social role
 - E.g.: positive correlation with GDP increase in developing countries
 - SMS messaging key tool of social interaction [Gartner Cons. 2008]

Further work

A lot to do... A few examples:

- We only used topology information
 - Statistical indicators (e.g., contact frequency between user pairs) might improve results
- Dealing with ordered sets (e.g., preference lists)
 - Need more sophisticated techniques
- Keep temporal evolution into account
 - e.g., give more importance to recent calls
- Integrate with existing social networking apps

... may leak info



If $f_i(x) < MIN_j$ then x not in A's contact list