

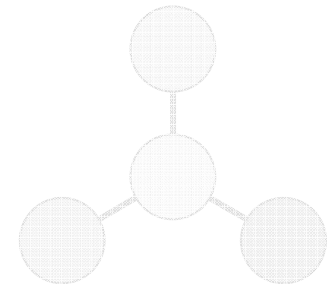


Implicit Affinity Networks (IAN)

Matthew Smith, Thesis Defense
Department of Computer Science
Brigham Young University

December 2006

smitty@byu.edu
<http://dml.cs.byu.edu/~smitty/>





Outline

- **Introduction**
- **Related Work**
- **Implicit Affinity Network (IAN)**
 - Building
 - Filtering
 - Social Capital
- **Implementation**
- **Experimental Results**
- **Conclusion and Future Work**



Introduction

- Individuals are complex and dynamic

- Complex

- No two individuals are exactly the same

- Dynamic

- Due to changing circumstances of life (e.g., marriage, retirement), geographical location, age, social interactions, hobbies, friends, family, etc.





Motivation

- Homophily
 - “Similarity begets friendship” - Plato
 - “Birds of a feather flock together”
- Belonging
 - Claimed by psychologists, such as Maslow, to be a basic need of humans
- Small World Phenomenon
 - People tend to be connected to each other by short chains of social acquaintances



Thesis Statement

- Although they clearly exist, affinities among individuals are not all easily identified. Yet, they offer unique opportunities to discover new social networks, strengthen ties among individuals, and provide recommendations.
- We propose the idea of Implicit Affinity Networks (IAN) to build, visualize, and track affinities among groups of individuals. IANs are simple, interactive graphical representations that users may navigate to uncover interesting patterns.
- This thesis describes a system supporting the construction of IANs and evaluates it in the context of online communities and family history.



Related Work

- Match-making “communities”
 - Examples
 - Match.com, LDSSingles.com, etc.
 - Static set of attributes to match on
- Circle of Friends communities
 - Examples
 - LinkedIn.com, Facebook.com, MySpace.com, etc.
 - Dynamic friends (explicitly linked)
 - Static set of attributes
 - Community evolution limited



More Related Work

- Network Modeling
- Social Network Analysis Methods
- Social Influence / Viral Marketing
- Social Capital
- Social Bookmarking by Tagging
- Collaboration and Communities
- Blogging / Web 2.0
- Targeted Advertising



Outline

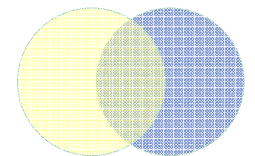
- Introduction
- Related Work
- **Implicit Affinity Network (IAN)**
 - Building
 - Filtering
 - Social Capital
- Implementation
- Experimental Results
- Conclusion and Future Work



Implicit Affinity

Let a community be comprised of an

- *Arbitrary number* of individuals $\{1, \dots, n\}$
 - With an *arbitrary number* of attributes for each individual $\{1, \dots, m_i\}$
 - Having an *arbitrary number* of values for each attribute $\{1, \dots, p_{ij}\}$
- We define an *affinity* between two individuals as the overlapping of attribute-values for any common attribute





Affinity Scoring

$$score_{A_j}(I_X, I_Y) = \frac{|V_j^X \cap V_j^Y|}{|V_j^X \cup V_j^Y|}$$

$$scorew_{A_j}(I_X, I_Y) = score_{A_j}(I_X, I_Y) \cdot \Delta_{A_j}^{X,Y} \cdot w_{A_j}$$

combined user weight

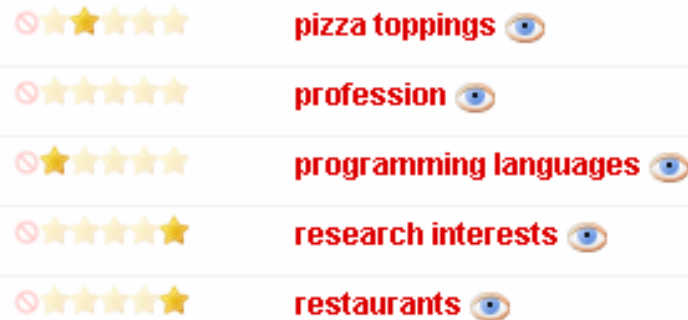
community weight

$$overall_score(I_X, I_Y) = \frac{1}{|A^X \cap A^Y|} \sum_{A_j \in A^X \cap A^Y} scorew_{A_j}(I_X, I_Y)$$



Combined User Weight

- Specified by each user ($\delta_{A_j}^X$)



Combined User Weight

$$\Delta_{A_j}^{X,Y} = \frac{\delta_{A_j}^X + \delta_{A_j}^Y}{2}$$

- Somewhat obtrusive, however, they allow the IANs to be tailored to the user's indicated interests (e.g., identifying clusters of individuals along well-defined characteristics)

$$score_{w_{A_j}}(I_X, I_Y) = score_{A_j}(I_X, I_Y) \cdot \Delta_{A_j}^{X,Y} \cdot w_{A_j}$$



Community Weights

- Dynamically learned for each attribute favoring *frequently* used attributes

$$w_{A_j} = \frac{|\{I_i \in I : V_j^i \neq \emptyset\}|}{|I|}$$

- Change as the community evolves

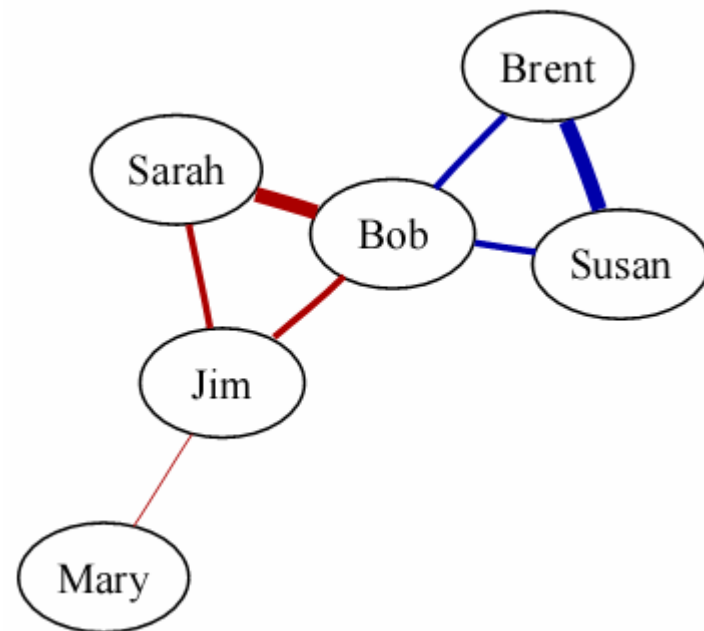
$$score_{w_{A_j}}(I_X, I_Y) = score_{A_j}(I_X, I_Y) \cdot \Delta_{A_j}^{X,Y} \cdot w_{A_j}$$



Affinity Network Building

Sample of Individuals and their Attributes

Individual	Attributes
Jim	A: $\{a_1, a_2, a_3\}$
Sarah	A: $\{a_1, a_2\}$
Mary	A: $\{a_3\}$
Bob	A: $\{a_1, a_2\}$ B: $\{b_1, b_2\}$
Susan	B: $\{b_1, b_2, b_3\}$
Brent	B: $\{b_1, b_2, b_3\}$



IAN

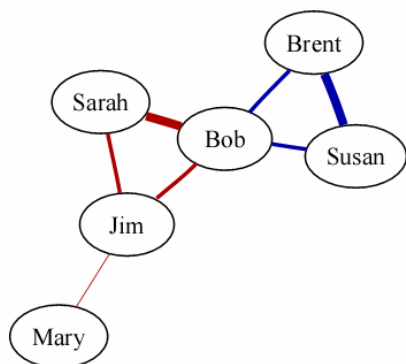
Sarah		Bob		Jim		Mary		Susan		Brent	
1	0	2	0	1	0	0	0	0	0	1	0
2	0	2	0	1	0	0	0	0	0	0	0
3	0	3	0	3	0	0	0	0	0	0	0
A	B	A	B	A	B	A	B	A	B	A	B



Affinity Network Filtering

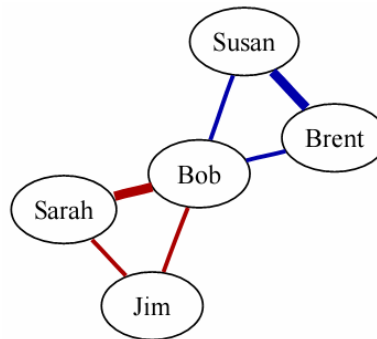
- Affinity Strength Threshold
- N-Affinities (strongest)
- User-Centralized n -Clique

Full Network



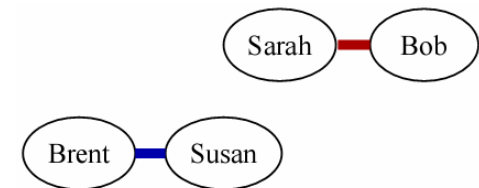
2-clique from Sarah

Threshold=0.5



2-clique from Susan

2-Affinities
(strongest)





Affinity Network Social Capital

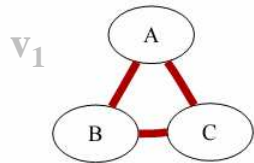
- Social Capital: The advantage available through connections between individuals within a particular network
- Bonding and Bridging Metrics

$$bonding_strength = \frac{\sum_{I_X \in I} \sum_{I_Y \in I} overall_score(I_X, I_Y)}{|I| \cdot \frac{(|I|-1)}{2}}$$

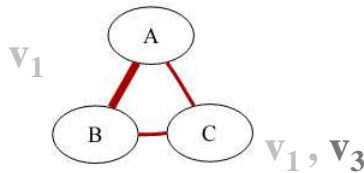
$$bridging_strength = 1 - bonding_strength$$



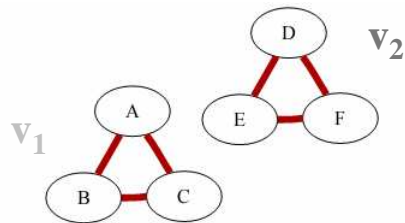
IAN Social Capital Evolution



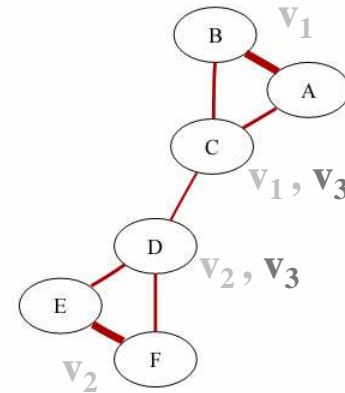
Bonding Strength: 1.00, Bridging Strength: 0.00 (t=1)



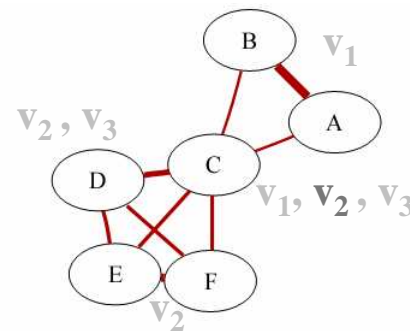
Bonding Strength: 0.67 ↓ - Bridging Strength: 0.33 ↑ (t=2)



Bonding Strength: 0.60 ↓ - Bridging Strength: 0.40 ↑ (t=3)



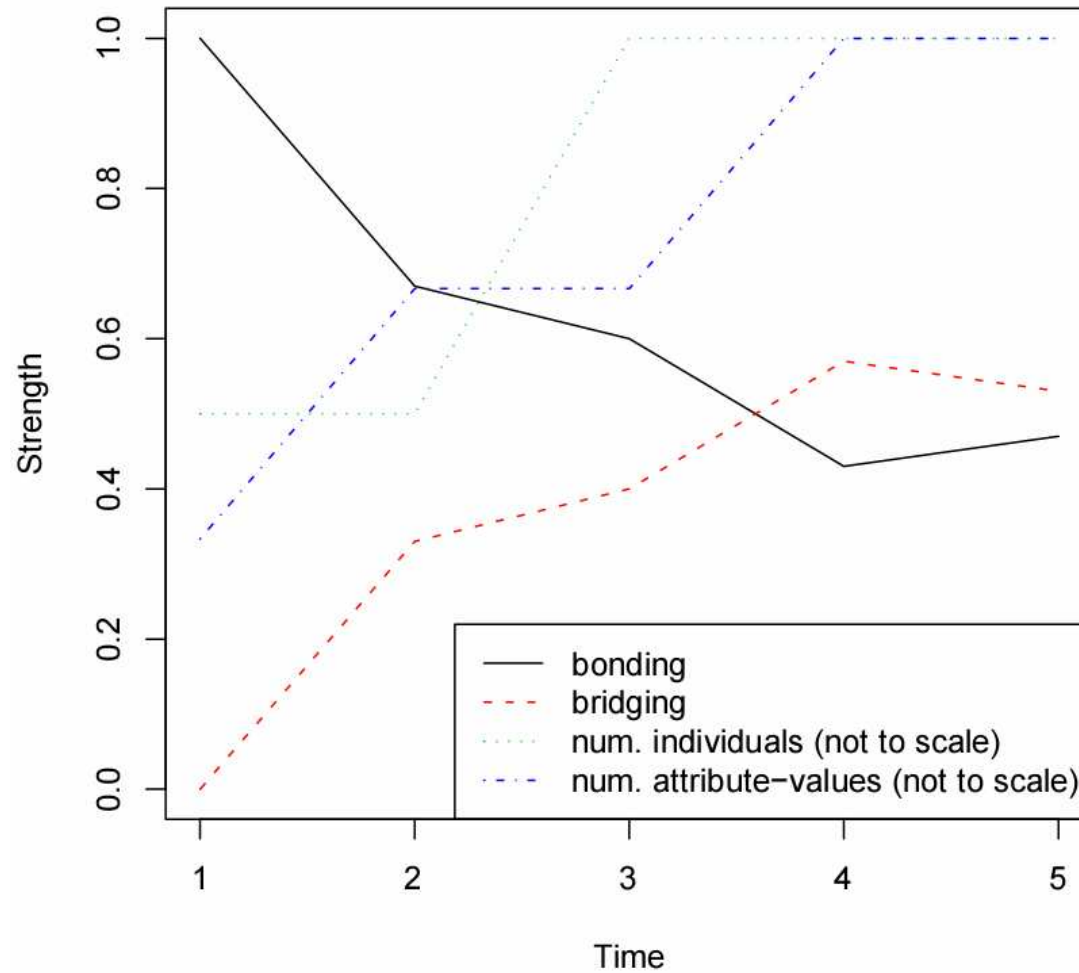
Bonding Strength: 0.43 ↓ - Bridging Strength: 0.57 ↑ (t=4)



Bonding Strength: 0.47 ↑ - Bridging Strength: 0.53 ↓ (t=5)



Network Strength Evolution





Outline

- Introduction
- Related Work
- **Implicit Affinity Network (IAN)**
 - Building
 - Filtering
 - Social Capital
- **Implementation**
- Experimental Results
- Conclusion and Future Work



IAN Community

- Online Community
 - Driven by members rather than managers
 - Members can describe themselves however they wish
 - Any change that a member makes affects all other members to some degree
- Online at:
 - <http://dml.cs.byu.edu/IAN>



Outline

- Introduction
- Related Work
- **Implicit Affinity Network (IAN)**
 - Building
 - Filtering
 - Social Capital
- Implementation
- **Experimental Results**
- Conclusion and Future Work

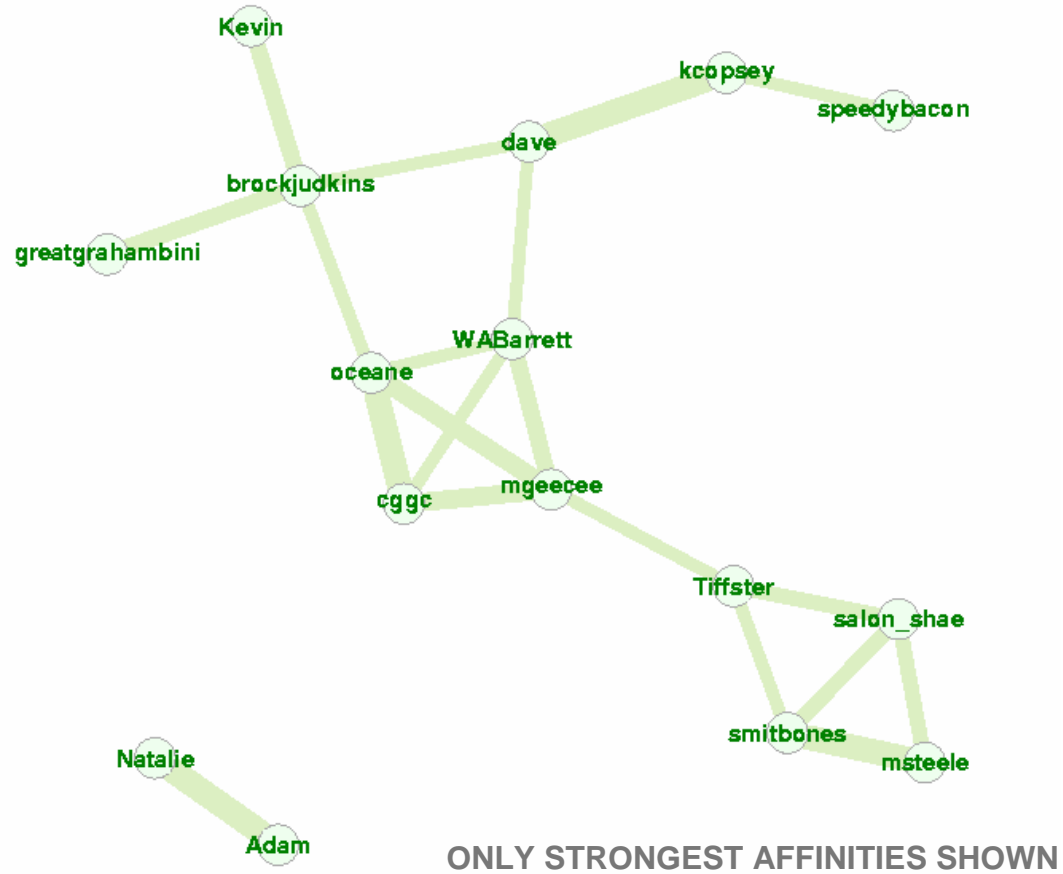
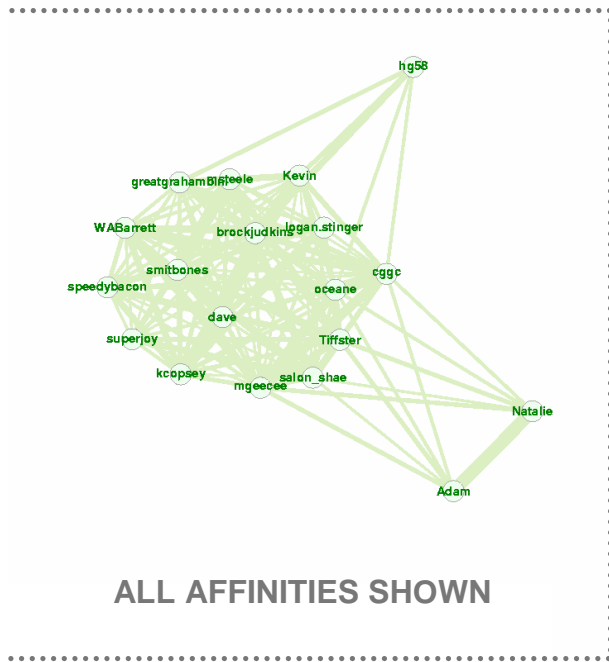


IAN Experiment

- 221 days
 - Since May 10
- 69 users participated in the experiment
- During first 183 days, on average,
 - a user was active for 39 days
 - a user visited every 8 days
 - a user added 2.38 attribute-values to their profiles per visit
 - a user had 93 attribute-values across 21 attributes



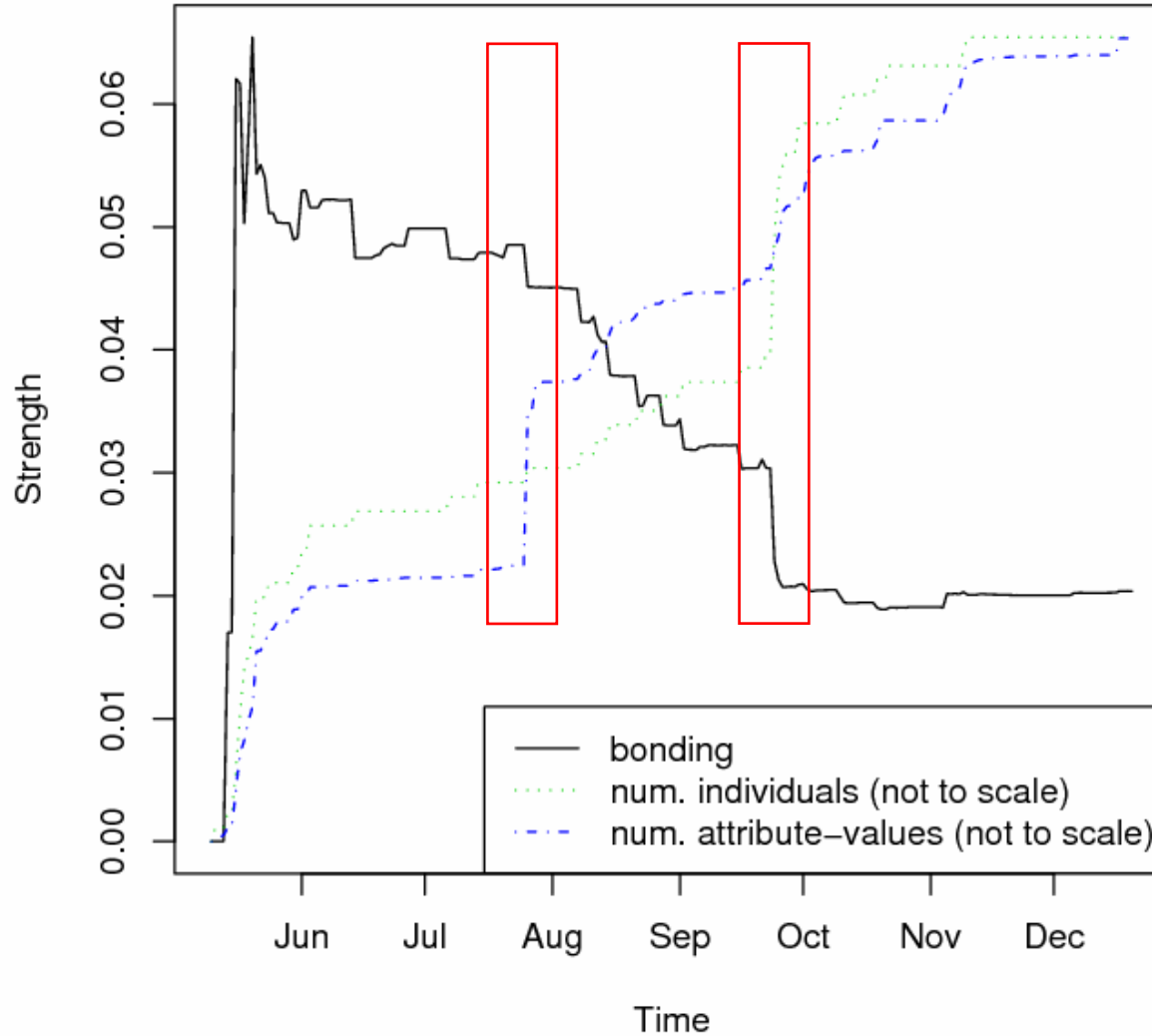
Languages Spoken



(IAN Snapshots taken: December 7, 2006)

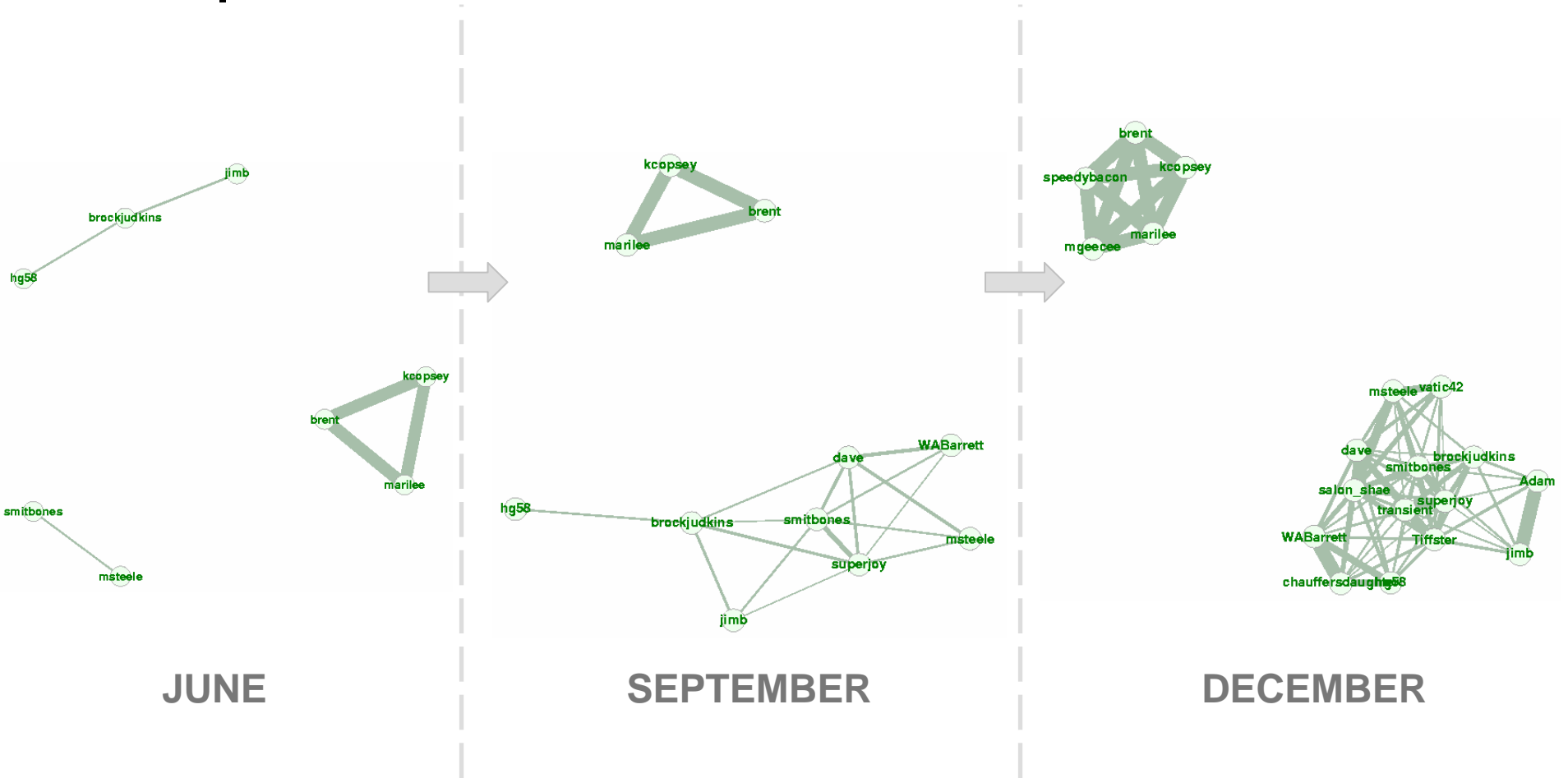


IAN Community Evolution



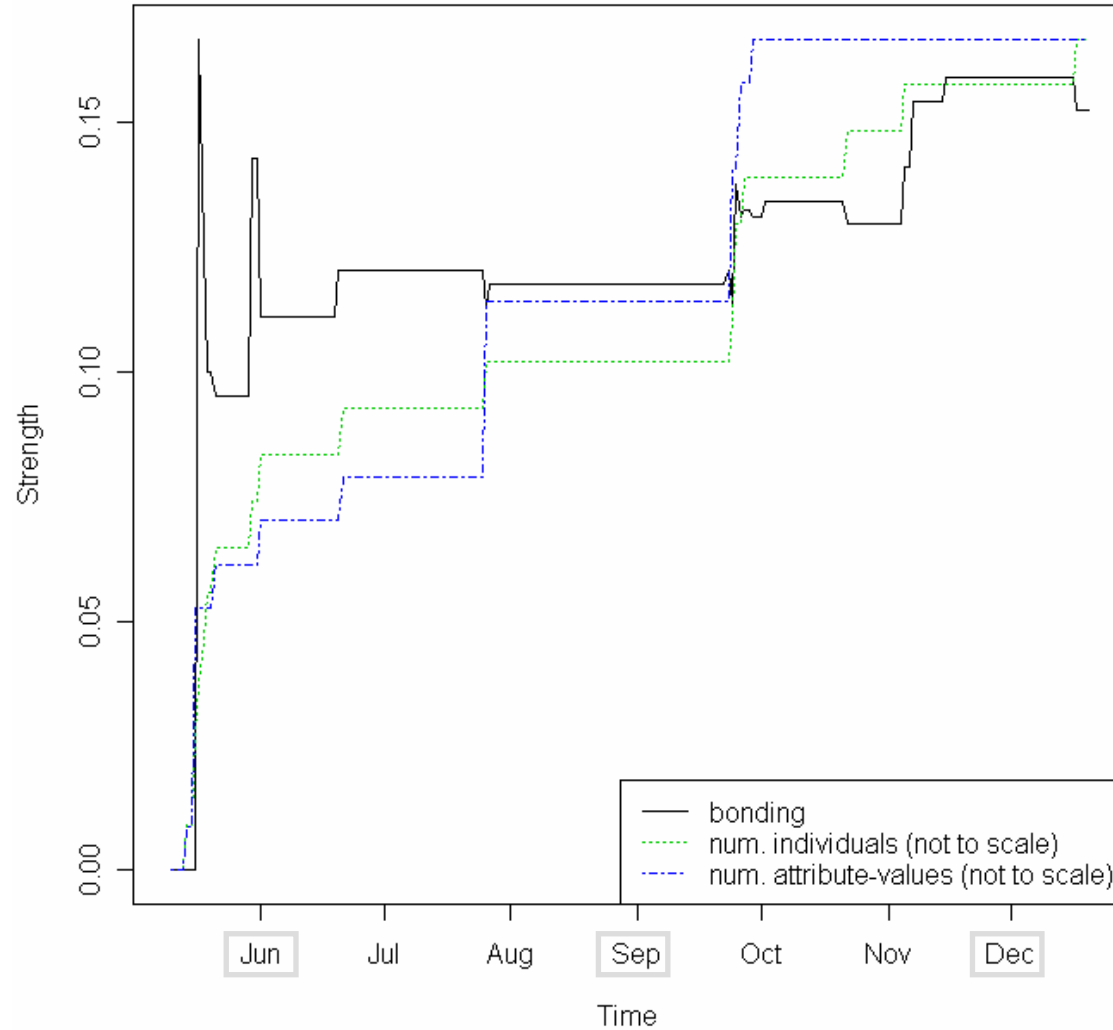


Musical Talents Evolution



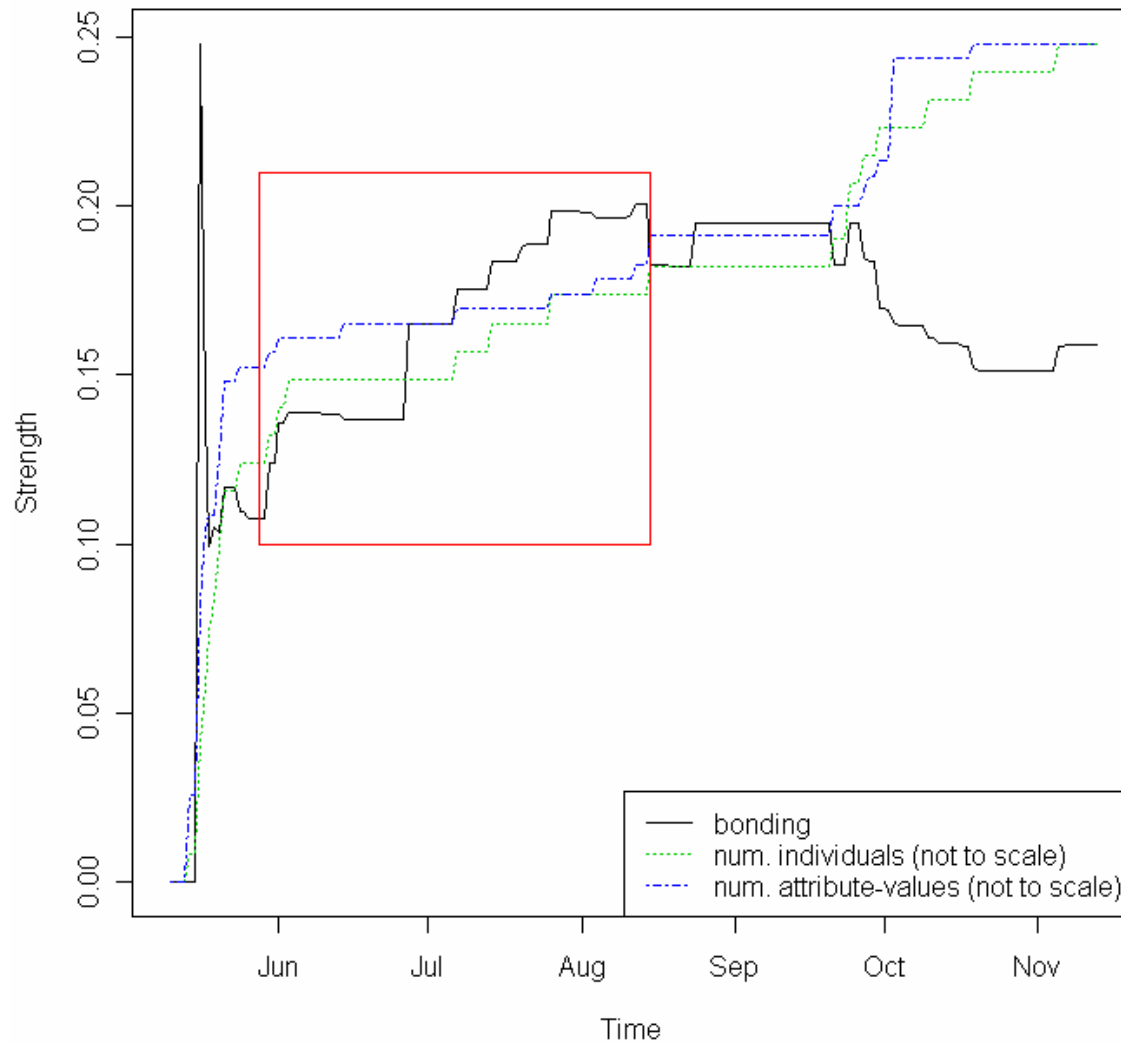


Musical Talents Evolution





Food Evolution





Qualitative Assessment – Survey Snippets

Q: What did you find most interesting about IAN?

- It was great to see the different things I have in common with the community
- It was interesting to see who was most related to me, and why
- It is cool that I can add my own interests

Q: What did you find most uninteresting about IAN?

- It was sometimes arduous to sort through attributes that I cared about
- Not knowing who some of the people in the community were made the affinities less interesting



Qualitative Assessment – Survey Snippets

Q: What new discoveries did the affinity networks help to highlight?

- It was interesting to find that I actually had stronger affinities with different people than I would have expected
- I am part of a musically talentless group
- There is good, solid, English ancestral origin
- I'm surprised by the handedness affinity network. I would have thought it would have a hard split, but hasn't (most people in that sub-community are right-handed)

Q: What, if any, thoughts or suggestions do you have?

- It would be interesting to do a study with businesses to see if grouping people based on their attribute affinities correlates to productivity in the workplace
- I misspelled a word and did not see a way I could fix it
- The graphical depiction of the network was interesting to see, but I didn't understand the way the spatial relationships were plotted.
- I think this concept could be used as a social tool for networking and friendship building



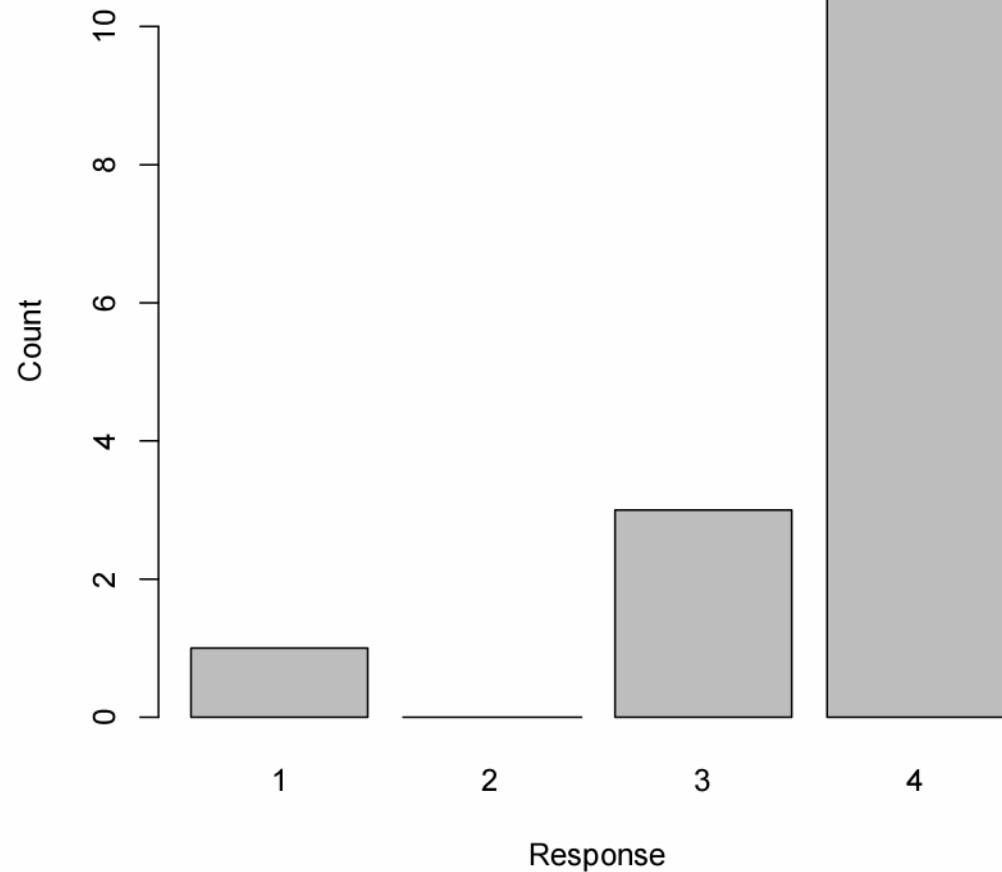
Qualitative Assessment - Survey

Please choose one answer to complete the following sentence:

IAN _____

- 1. did nothing for me
- 2. showed me things I already knew
- 3. highlighted things I suspected but was unsure about
- 4. helped me discover new things

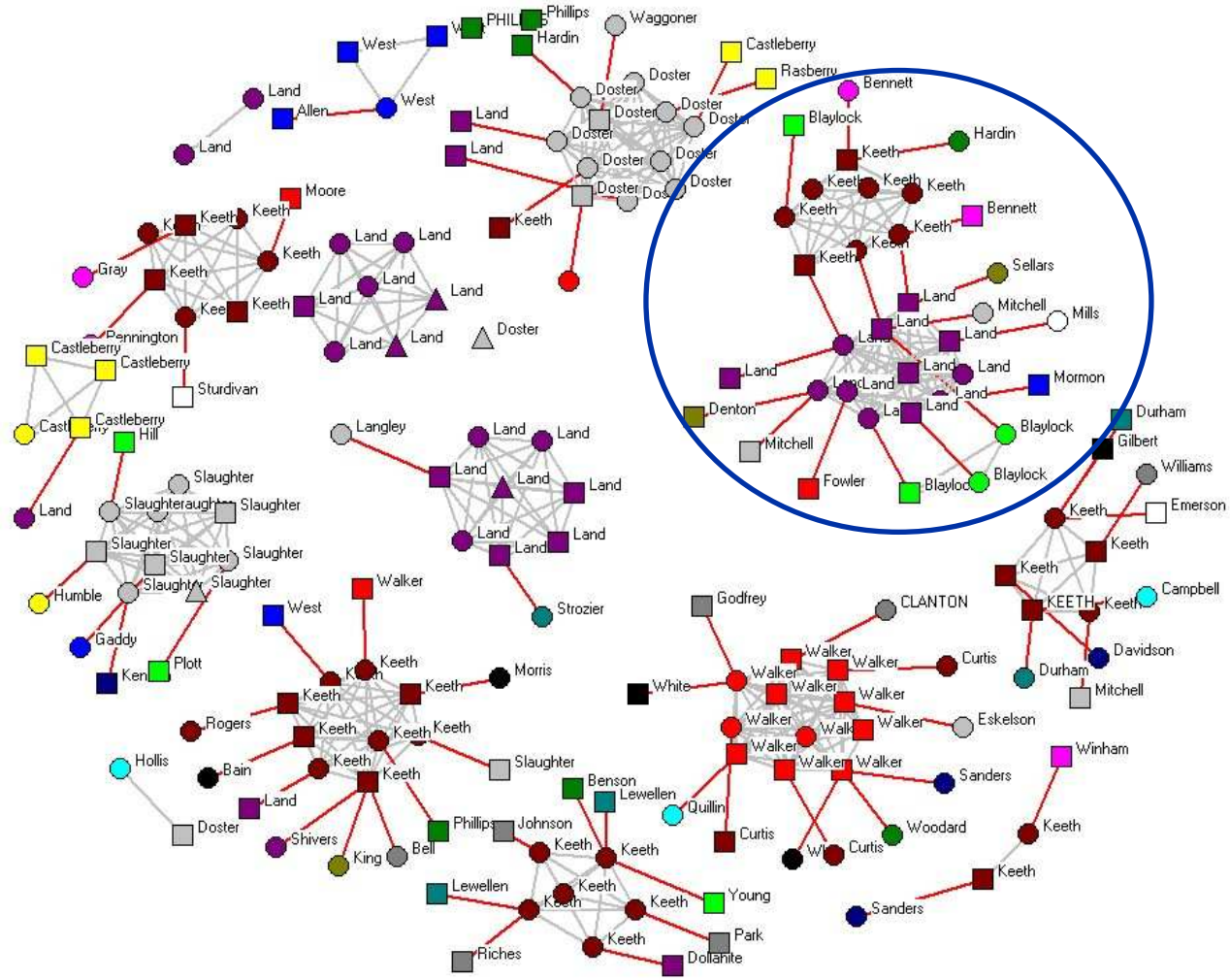
~23% responded (15 of 65)



GIAN



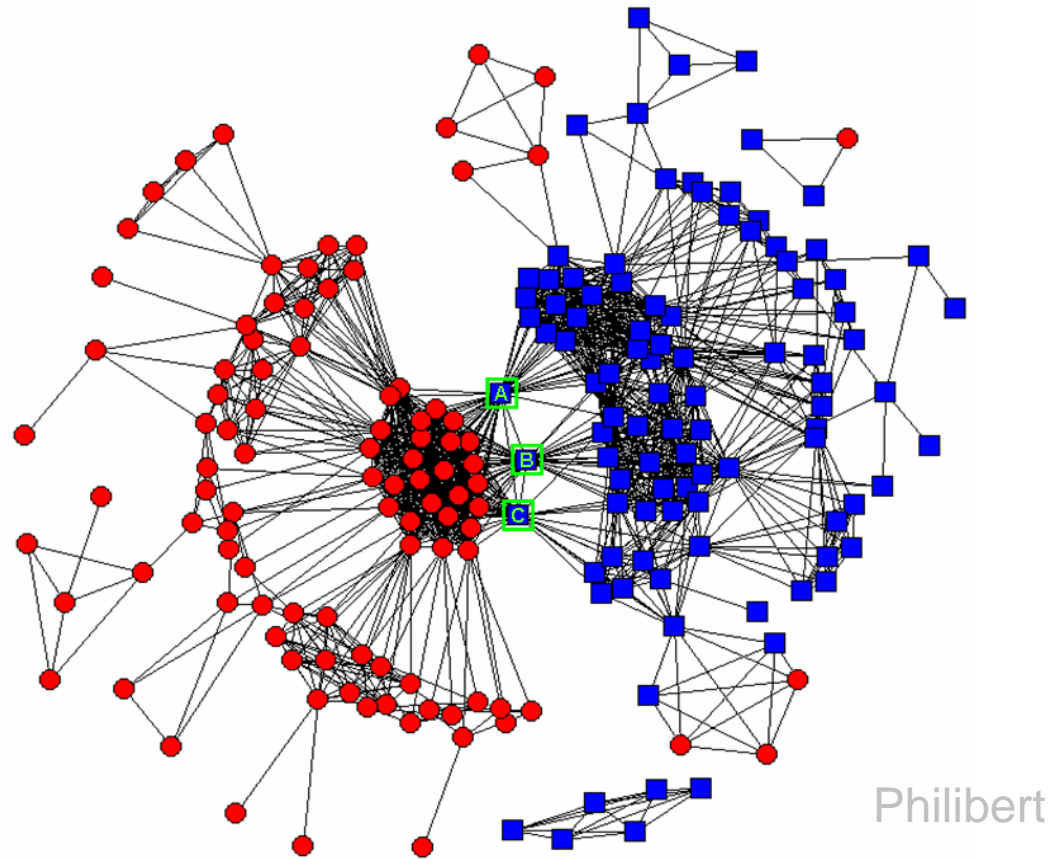
Spouse-Sibling IAN



GIAN



Gender Naming Patterns



A Joseph Marie
B Francois Marie
C Jean Marie

Philibert



Outline

- Introduction
- Related Work
- **Implicit Affinity Network (IAN)**
 - Building
 - Filtering
 - Social Capital
- Implementation
- Experimental Results
- **Conclusion and Future Work**



Conclusion and Future Work

- Conclusion
 - We presented an inventive methodology of creating and using IANs for discovery
 - IANs can be used to better understand how complex entities are interconnected and how they evolve
 - We have begun to show how network strength metrics of bonding and bridging can be used to measure social capital and track community evolution
- Future Work
 - Apply IANs to more domains
 - Medical patient communities
 - Search communities
 - Extend IAN interactivity
 - Reproduce existing / create new communities



Questions?

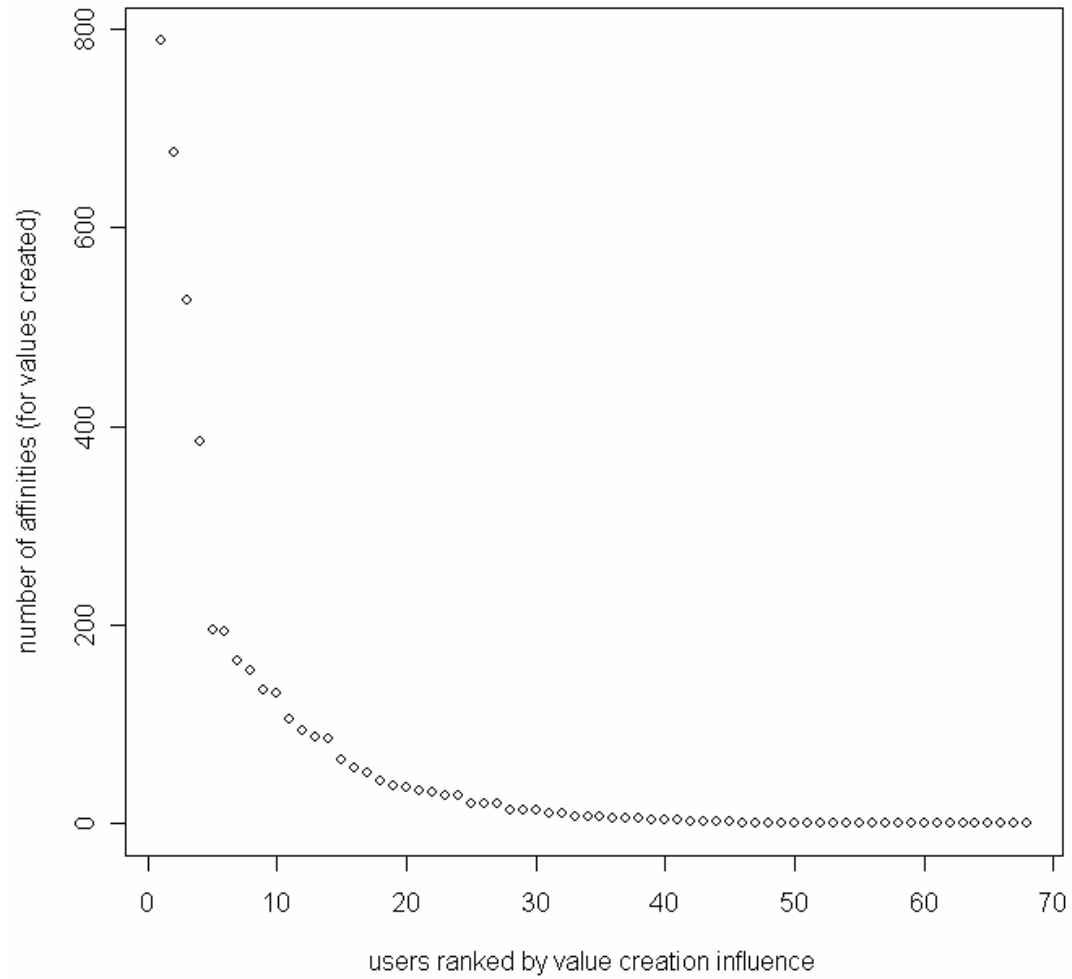


Extra Slides

- Influencers
- Search Phrase Community
- Research Interests
- Connecting Attributes
- IAN Snapshots
- GIAN Examples
- Jon Kleinberg Quote

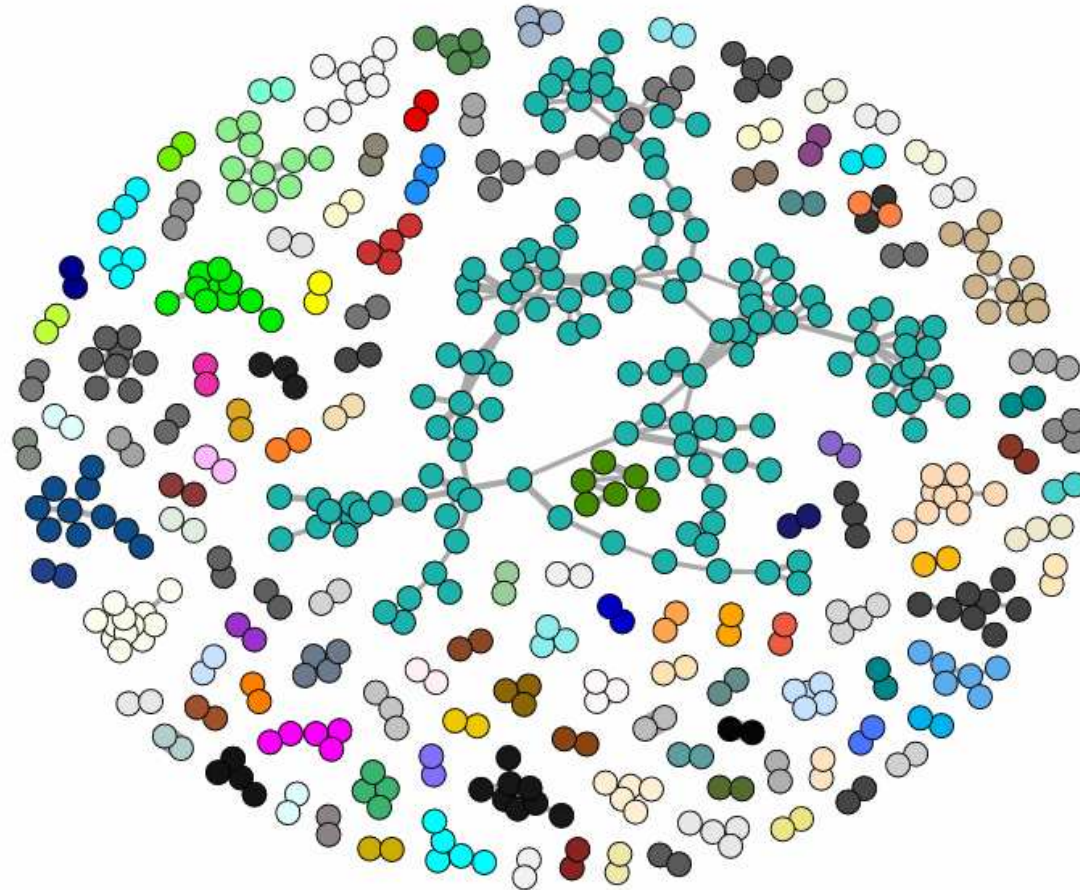


Influencers



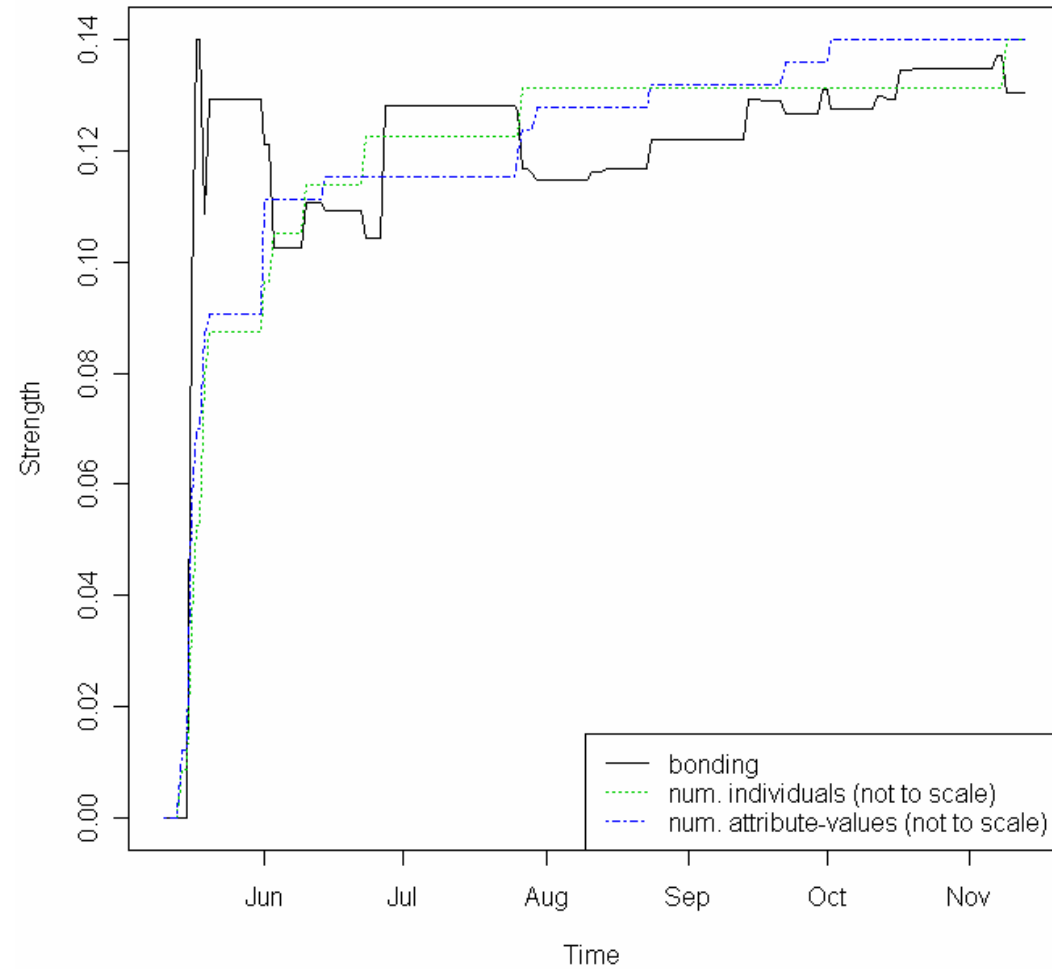


Search Phrase Community



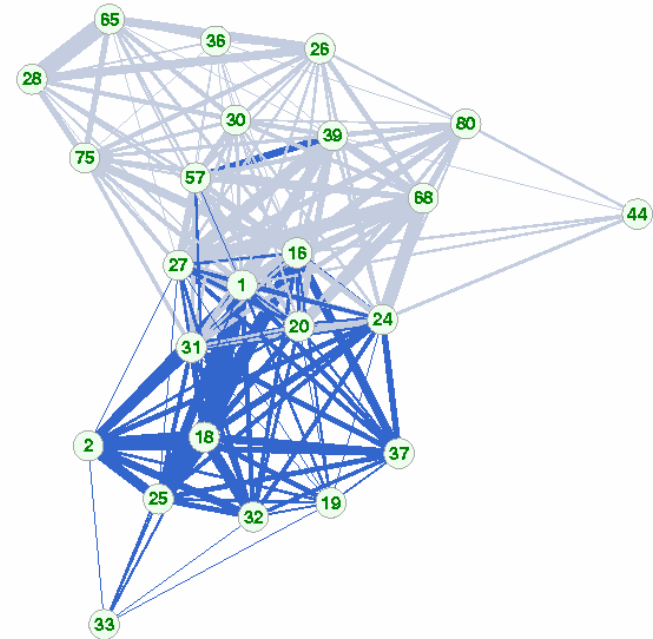
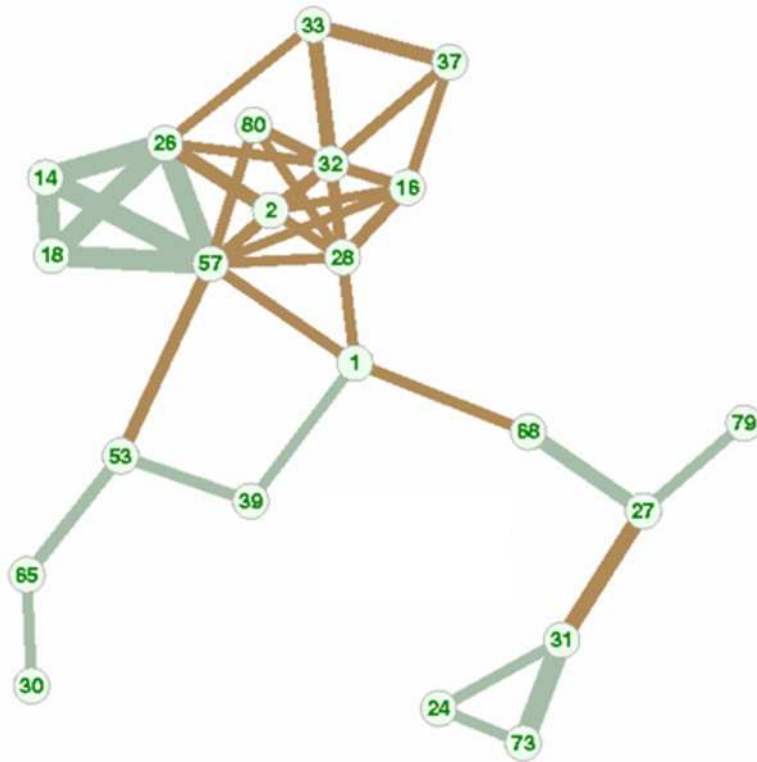


Research Interests Evolution



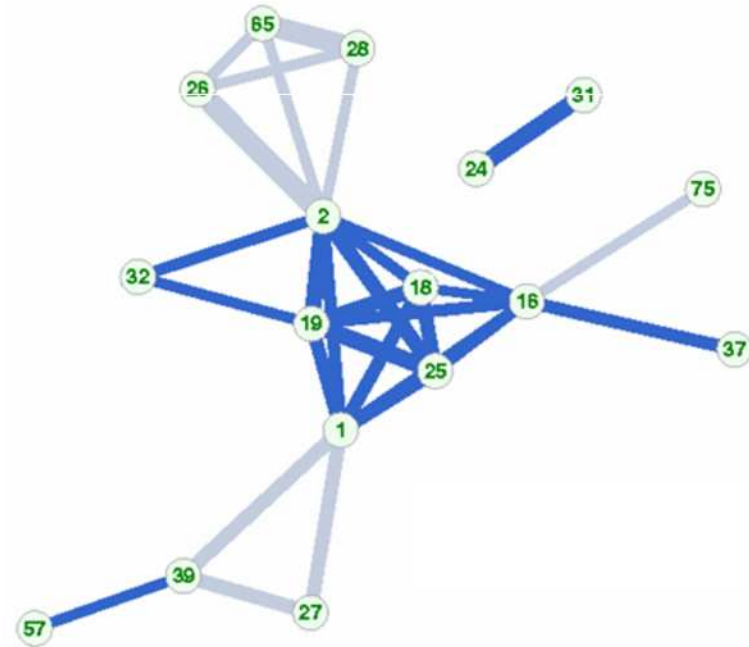
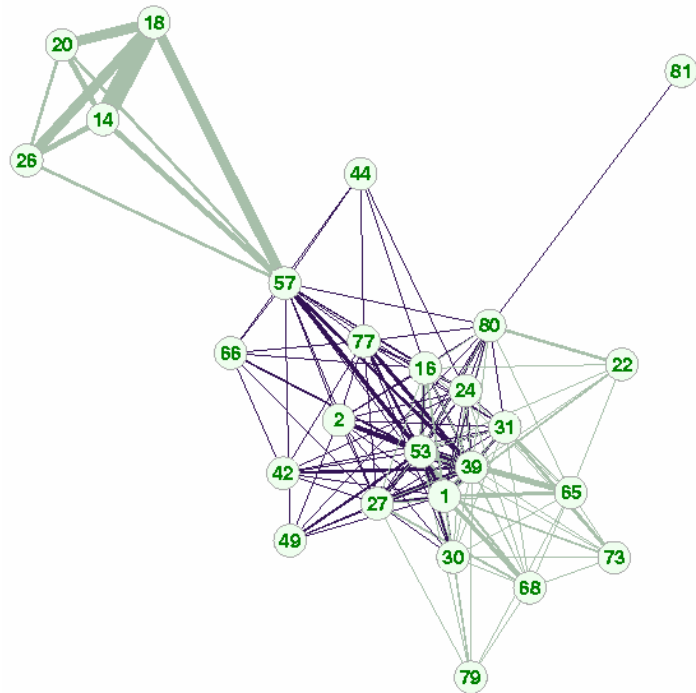


Connecting Through Attributes





Connecting Through Attributes





IAN Screenshots

IAN - Home - Mozilla Firefox

Search Attributes

Signed In as **smilbones** [settings] Community | My Attributes | My Affinities | Sign out

Popular Attributes

- food (40 affinities)
- hobbies (35 affinities)
- sports (32 affinities)
- programming languages (30 affinities)
- board games (29 affinities)

New Affinities

- ian (26 affinities) 1.2 hours ago
- what's on Tom Cruise's ipod? (2 affinities) 1.5 hours ago
- food (40 affinities) 3.1 days ago
- sports (32 affinities) 3.1 days ago
- board games (29 affinities) 3.1 days ago

New Attributes [\[view all\]](#)

favorite season, birth year, what's on the president's ipod?, what's on Tom Cruise's ipod?, what's on Greg Ostertag's ipod?, Phone service, never done, handed?, guilty pleasure, Political Affiliation, can't live without, Favorite Utah Spots, Pet Peeves, kids?, electronics for entertainment purpose, cartoon fever, best sci-fi movie characters, lucky number(s), favorite websites, Toys, pizza toppings, Energy Drinks, The 80's, shoes, tv, more...

Discussion

- sports** 2 minutes ago
We will be playing basketball this Saturday. If you are interested in join... - smilbones
- what's on Tom Cruise's ipod?** 1.9 hours ago
I agree, but I don't think people necessarily get it... superjoly
- what's on Tom Cruise's ipod?** 2 days ago
This is a creative topic... - smilbones
- Phone service** 5.3 days ago
Although the commercials claim the reverse, Cingular seems to be the network... - smilbones
- hobbies** 1 week ago
Although we are not all connected directly through our hobbies, there seems... - smilbones
- States Visited** 1 week ago
WABarrett and msteale are probably both very familiar with the USA... - smilbones
- musical talents** 1 week ago
This is a very interesting graph. Not only are there two distinct clusters... - smilbones

Community Members

smilbones, superjoly, natescottsdavis, test, anonymous, msteale, iun, brockjudkins, brent, hg58, greatgrahamini, fid1976, dave, hblinc, Katie, jhmsady, logan stinger, WABarrett, manlie, carriele, bplaton, Isabelle, jimb, cggc, oceane, mgecece, sive, kcoppsey, xokclothing

IAN | About | Terms of Service | Privacy Policy | Contact Us

© 2006 BYU Data Mining Lab - Matt Smith

IAN - User Attributes - Mozilla Firefox

Search Attributes

Signed In as **smilbones** [settings] Community | My Attributes | My Affinities | Sign out

My Attributes

Total: 68
[Add New Attribute](#)

Importance	Label	Options
★★★★★	academic major	edit delete
★★★★★	achievements	edit delete
★★★★★	bands	edit delete
★★★★★	best sci-fi movie characters	edit delete
★★★★★	beverage	edit delete
★★★★★	birth month	edit delete
★★★★★	birth year	edit delete
★★★★★	board games	edit delete
★★★★★	books	edit delete
★★★★★	can't live without	edit delete
★★★★★	cartoon fever	edit delete
★★★★★	cold cereal	edit delete
★★★★★	computer science classes taken	edit delete
★★★★★	countries lived in	edit delete
★★★★★	countries visited	edit delete
★★★★★	databases	edit delete
★★★★★	degrees obtained	edit delete
★★★★★	electronics for entertainment purpose	edit delete
★★★★★	eye color	edit delete
★★★★★	fan	edit delete
★★★★★	favorite season	edit delete
★★★★★	favorite sites	edit delete

IAN - Attribute View - Mozilla Firefox

Attribute: **hobbies** [\[Add a new value\]](#)

board games, Camping, cooking, Cooking new food, Crafts, Crossword Puzzles, disappearing, doing hair, drawing, drums, eating, Fishing, flying, gardening, graphic arts, guitar, Hiking, lawn games, making jewelry, playing Baduk, Playing soccer, Playing Tennis, programming, Quilting, Reading, Scrapbooking, shopping, sleeping, Stamp Collecting, star wars trivia, thinking, Video Games, watching tv, writing

Comments/Discussion [\[Add\]](#)

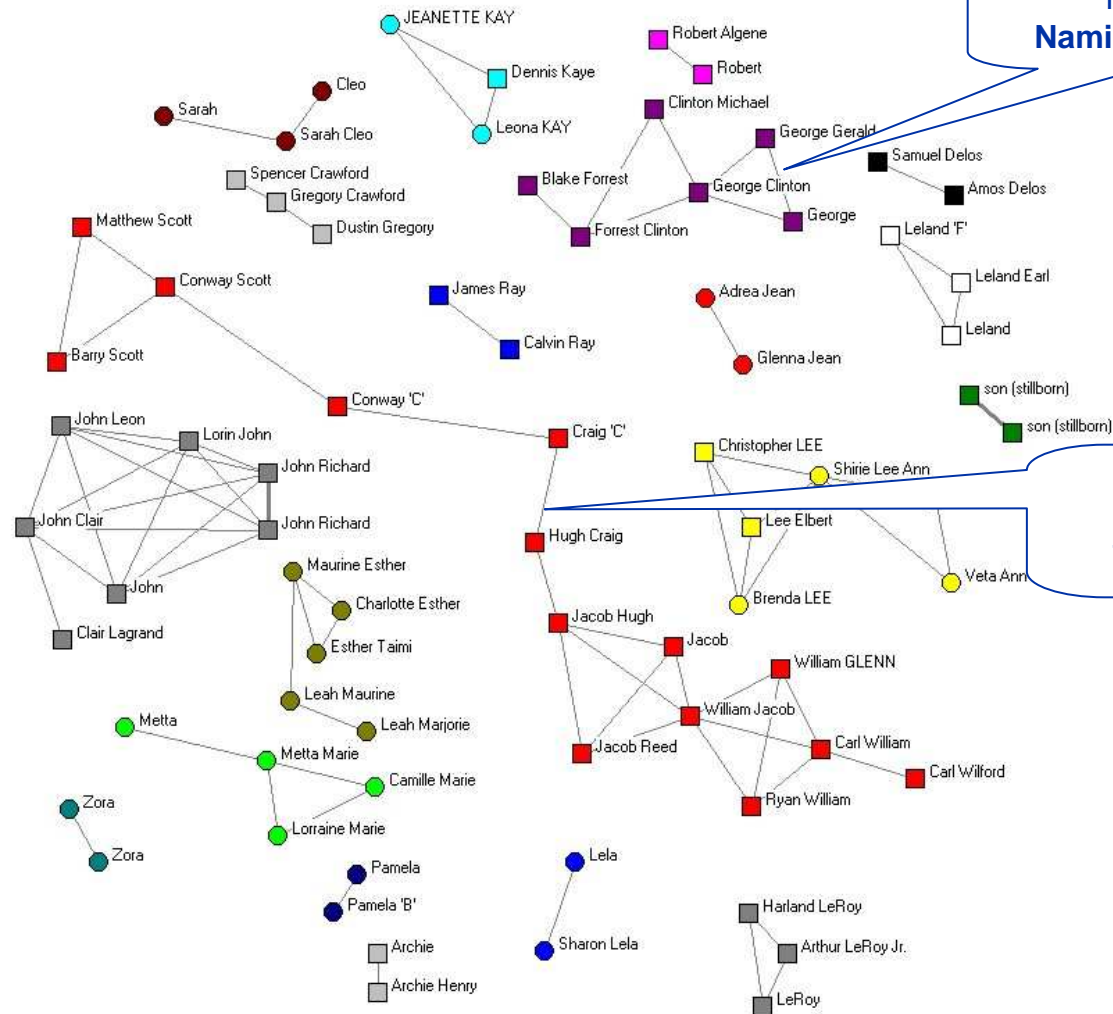
smilbones said: (1.1 weeks ago)
Although we are not all connected directly through our hobbies, there seems to be overlapping hobbies which tie us all together. It is interesting how that happens.

Network graph showing connections between users based on shared hobbies. Nodes represent users and edges represent shared attributes.



Given Name IAN

(One or more affinities --- Isolates removed)



More neat Naming Patterns...

Interesting Naming Pattern Through generations



Jon Kleinberg

- In a recent article in the New York Times, renowned computer scientist Jon Kleinberg had this to say about social networking research:
 - ``We're really witnessing a revolution in measurement...This is the introduction of computing and algorithmic processes into the social sciences in a big way, and we're just at the beginning''
- We are encouraged by this trend. IANs offer a way to build and analyze social networks in a way that may be useful to sociologists.