Early Identification of Violent Criminal Gang Members

Elham Shaabani
Ashkan Aleali
Paulo Shakarian

Arizona State University
Tempe, AZ
shaabani@asu.edu

John Bertetto

Chicago Police Department
Chicago, IL

CySIS Tech '15
Cyber-Socio Intelligent Systems Laboratory
Annual Technical Symposium – August 19, 2015

Challenge

- Can we identify potential violent offenders ahead of time?
  - Not trying to create a crystal ball
  - Instead, try to better use police forces to avoid violence and reduce homicides
  - not to direct arrests, but to direct police presence in time and violent spikes

  - Given:
    - Co-arrestee social network structure
    - Meta-data from the arrest records
      - do not leverage features concerning the race, ethnicity or gender of individuals

Main results

- We leverage a combination of social network analysis and supervised learning
- Precision 0.89, recall 0.78 when the entire social network is known
- Improved precision and recall over currently used approach when the social network is learned over time – producing 4x more true positives

Overview of Network Data

Network Data

Name	Value
Number of records	64466
Violent offense	4450
Homicide	312
Criminal sexual assault	153
Robbery	1959
Aggravated assault	1441
Aggravated battery	896
Non violent offense	60016

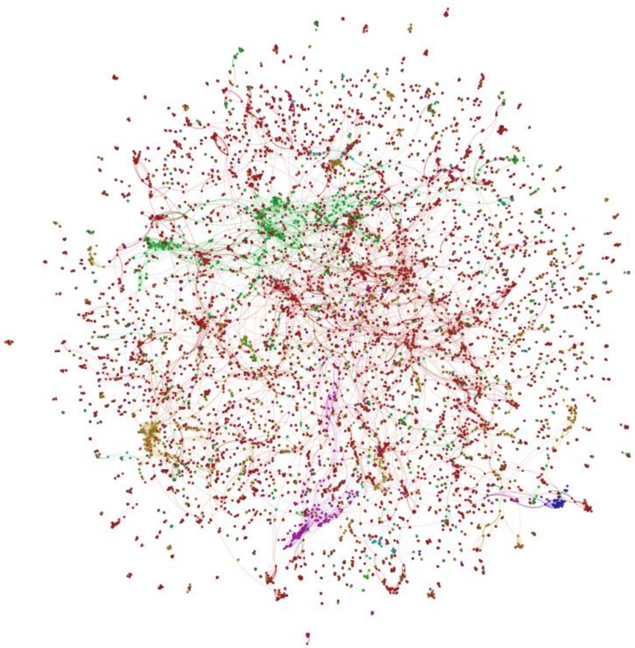
August 2011 – August 2014

In Chicago

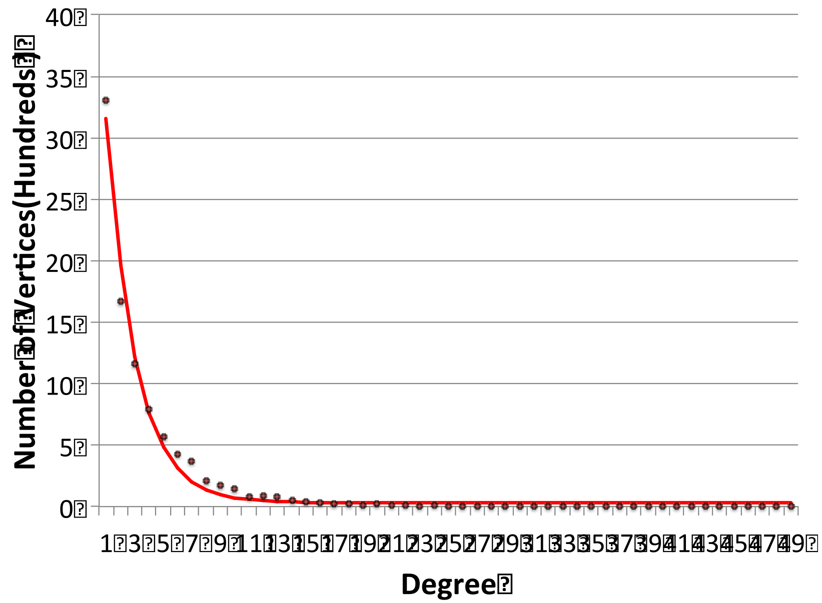
Highly imbalanced

Network Properties

CO-ARRESTEE NETWORK

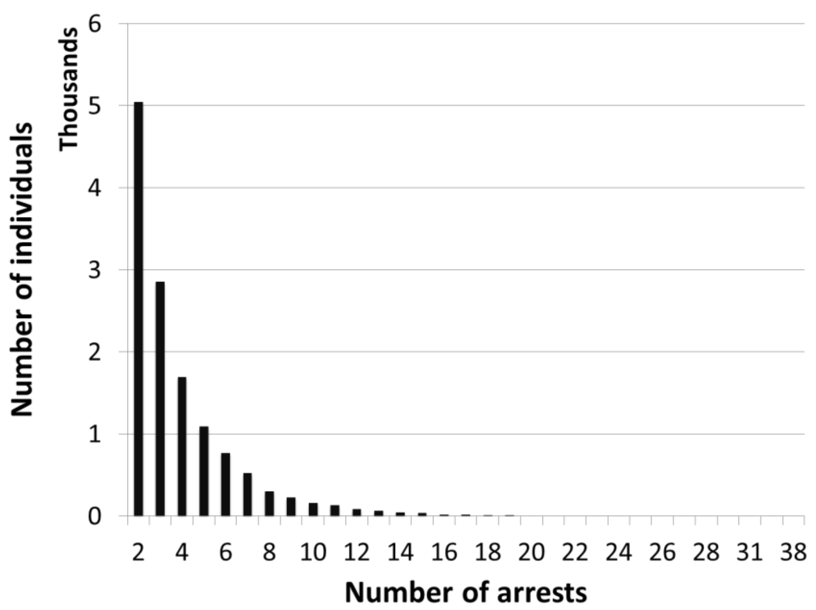


DEGREE DISTRIBUTION

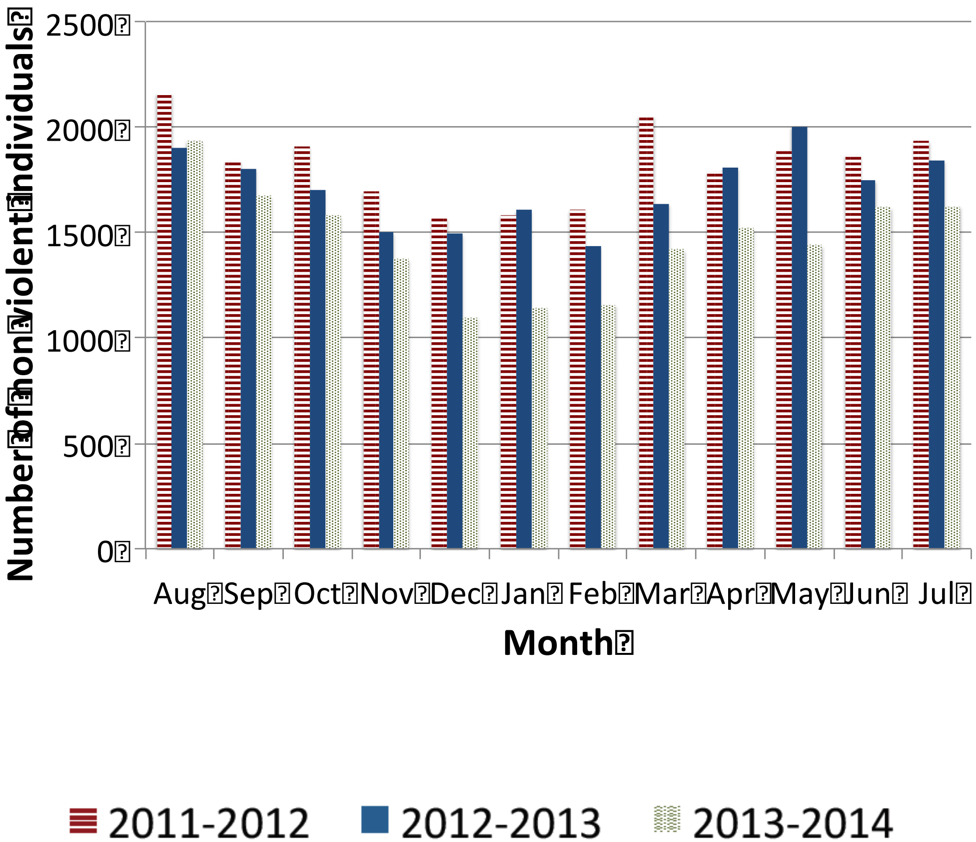


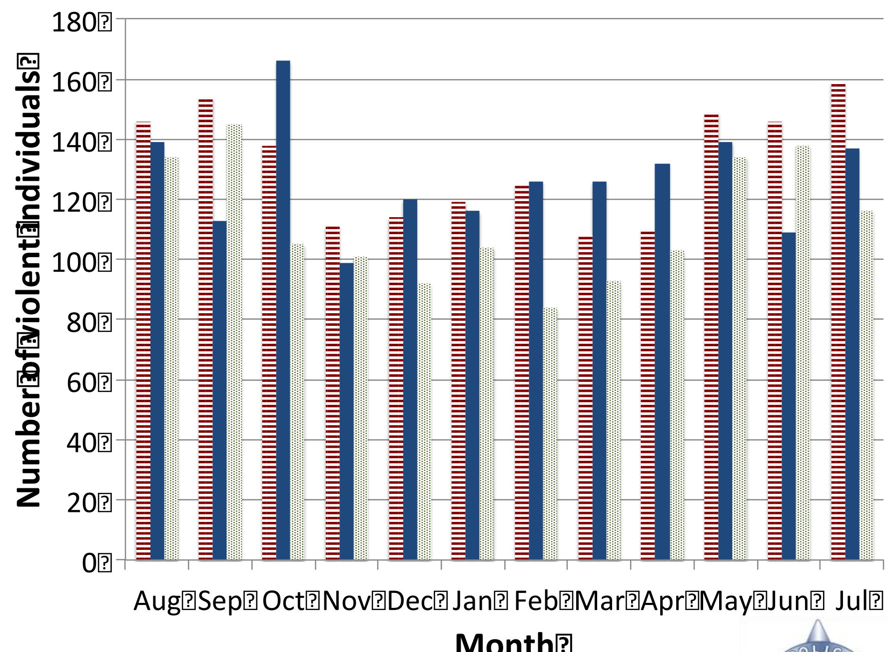
Many Future Violent Offenders are Known to Law Enforcement

RE-ARREST DISTRIBUTION



Seasonality of Crime





ARIZONA STATE UNIVERSITY

Identifying Violent Offenders

Existing methods

- Past Violent Activities (PVA)
  - If an offender has committed a violent crime in the past, we claim that he will commit a violent crime in the future.
- Two-Hop Heuristic (THH)
  - All neighbors one and two hops away from previous violent criminals

Supervised Learning Approach

Supervised learning approaches

- **Random Forest**
- Naïve Bayes
- Linear Regression
- Decision Tree
- Neural Network
- Support Vector Machine

Features

- Neighborhood Based
- Network Based
- Temporal
- Geographic

Neighborhood-Based Features

Description	Definition
Degree (w.r.t. C)	$\lvert\{u \mid u \in N_v^1 \cap V_C\}\rvert$
Fraction of 1-hop neighbors committing a crime in C	$\lvert\{u \mid u \in N_v^1 \cap V_C\}\rvert / \lvert N_v^1\rvert$
Fraction of 2-hop neighbors committing a crime in C	$\lvert\{u \mid u \in N_v^2 \cap V_C\}\rvert / \lvert N_v^2\rvert$
Majority of 1-hop and 2-hop neighbors committing a crime in C	$maj_v(C,1) \wedge maj_v(C,2)$
Minority of 1-hop and majority of 2-hop neighbors comitting a crime in C	$\neg maj_v(C,1) \wedge maj_v(C,2)$

- Each node and its first/second level neighbors
- *maj_v(C,i)* is TRUE if at least half of the nodes within a network distance of *i* from node *v* have committed a crime in *C* and FALSE otherwise.

Network-Based Features

Community Based

Description	Definition
Component size when v is removed	$\lvert \mathsf{C}(\mathsf{C}_v(G) \setminus \{v\}) \rvert$
Largest component size with a violent node after v is removed	$\max_{v' \in \mathsf{C}(\mathsf{C}_v(G)\{v\} \cap V_{\mathcal{V}}} \lvert X_{v'} \rvert$ where $X_{v'} = \mathsf{C}_{v'}(\mathsf{C}_v(G)\{v\})$
Group size	$\lvert \mathsf{P}_v(G_{gang_v}) \rvert$
Relationships within the group	$\lvert \{(u,v) \in E \ s.t. \ u,v \in \mathsf{P}_v(G_{gang_v})\} \rvert$
Number of violent members in the group	$\lvert \{v' \in \mathsf{P}_v(G_{gang_v}) \ s.t. \ \mathcal{V}_v \neq \emptyset\} \rvert$
Triangles in group	No. of triangles within subgraph $\mathsf{P}_v(G_{gang_v})$
Transitivity of group	$\frac{No.\ of\ triangles\ in\ \mathsf{P}_v(G_{gang_v})}{No.\ of\ \text{“}\vee\text{”'s}\ in\ \mathsf{P}_v(G_{gang_v})}$
Group-to-group connections	$\lvert \{u \in \mathsf{P}_v(G_{gang_v}) \ s.t. \ \exists(u,w) \in E \ where \ w \notin \mathsf{P}_v(G_{gang_v})\} \rvert$
Gang-to-gang connections	$\lvert \{u \in G_{gang_v} \ s.t. \ \exists(u,w) \in E \ where \ w \notin G_{gang_v}\} \rvert$

Path Based

Description	Definition
Betweenness (w.r.t. C)	$\sum_{u,w \in V_C} \frac{\sigma_v(u,w)}{\sigma(u,w)}$
Closeness (w.r.t. C)	$(\lvert V_C \rvert - 1) / \sum_{u \in V_C} d(u,v)$
Shell Number (w.r.t. C)	$shell_C(v)$ (see appendix for further details)
Propagation (w.r.t. C)	1 if $v \in \Gamma_\kappa(V_{\mathcal{V}})$, 0 otherwise. (see appendix for further details)

- Leveraged the intuitions from social network analysis and criminology to generate new and useful features

Geographic Features

Name	Definition
District Frequency	$\lvert\{(t,v')\ s.t.\ arr^t_{v'} = \mathsf{true} \wedge \exists t'\ s.t.\ dstr^t_{v'} = distr^{t'}_v\}\rvert$
Beat Frequency	$\lvert\{(t,v')\ s.t.\ arr^t_{v'} = \mathsf{true} \wedge \exists t'\ s.t.\ beat^t_{v'} = beat^{t'}_v\}\rvert$
Beat Violence	$\lvert\{(t,v')\ s.t.\ arr^t_{v'} = \mathsf{true} \wedge \mathcal{V}^t_{v'} \neq \emptyset \wedge \exists t'\ s.t.\ beat^t_{v'} = beat^{t'}_v\}\rvert$
District Violence	$\lvert\{(t,v')\ s.t.\ arr^t_{v'} = \mathsf{true} \wedge \mathcal{V}^t_{v'} \neq \emptyset \wedge \exists t'\ s.t.\ dstr^t_{v'} = distr^{t'}_v\}\rvert$

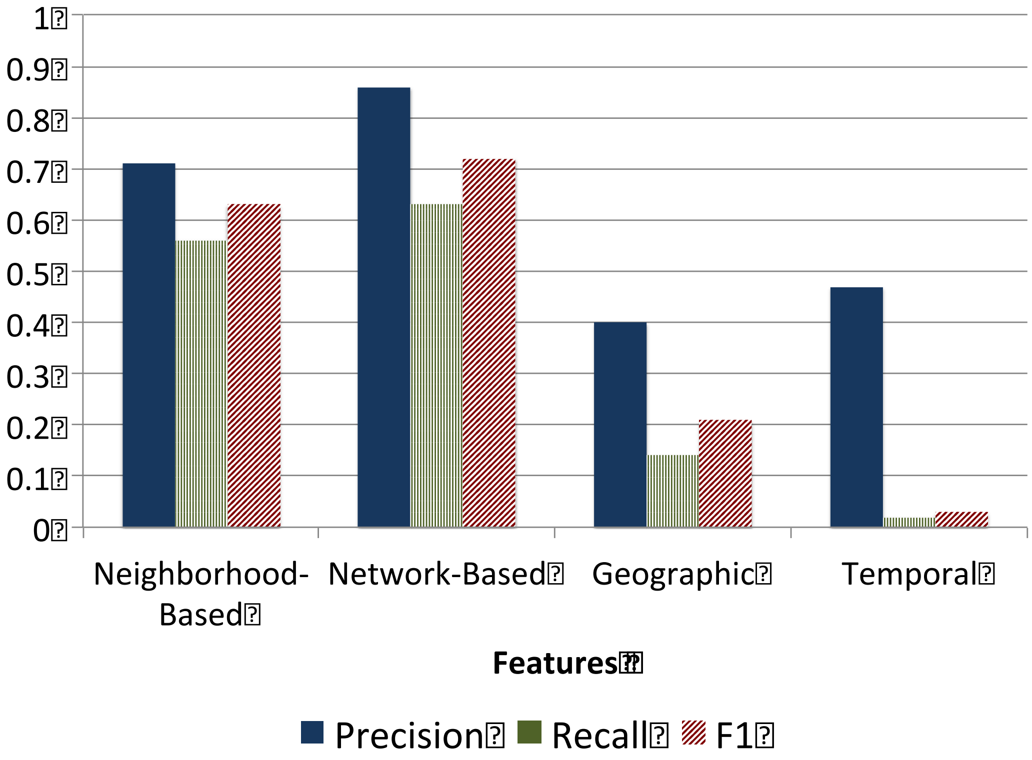
- Capture the information related to the location of a crime incident
- In accordance with well known literature in criminology

Temporal Features

Name	Definition
Average interval time (w.r.t. C)	$\sum_i \Delta_i^v(C)/\lvert t_C^v \rvert$
Number of violent groups	$\lvert\{t\ s.t.\ arr_v^t = \mathsf{true} \wedge \exists v'\ s.t.\ arr_{v'}^t = \mathsf{true} \wedge \mathcal{V}_v^t \neq \emptyset \wedge v' \in N_v^t\}\rvert$

Results: Known Co-Arrestee Network

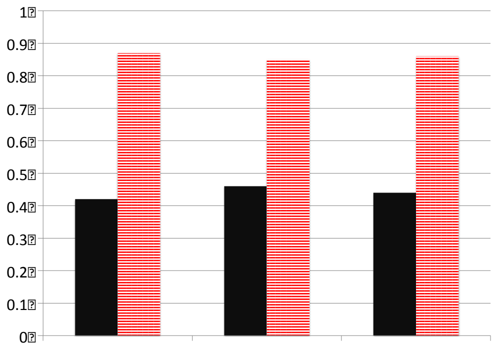
Classification using Single Feature Categories



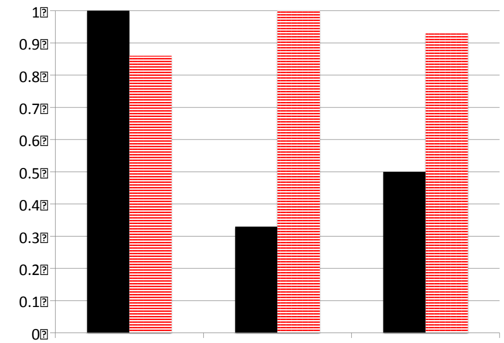
Network based is highly correlated to violent behaviour

Performance of One Feature

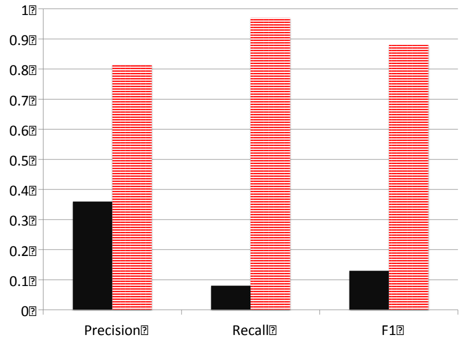
Minority of 1-hop and majority of 2-hop neighbors committing a violent crime
(Neighborhood-Based)



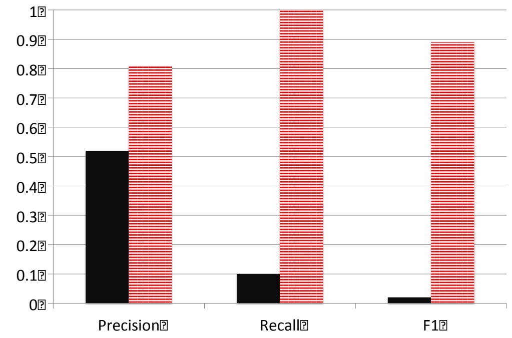
Closeness
(Network-Based)



Beat violent
(Geographic)



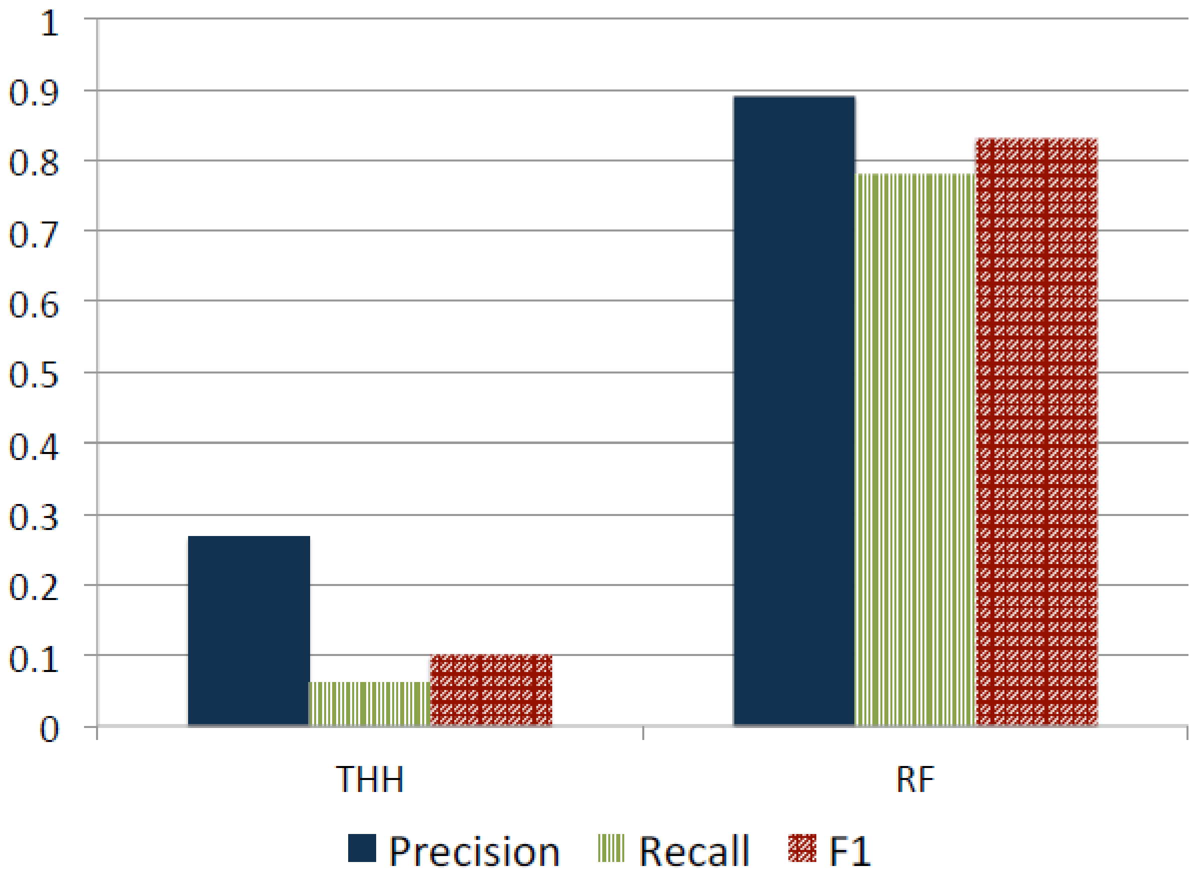
Average interval month
(Temporal)



Violent Non violent

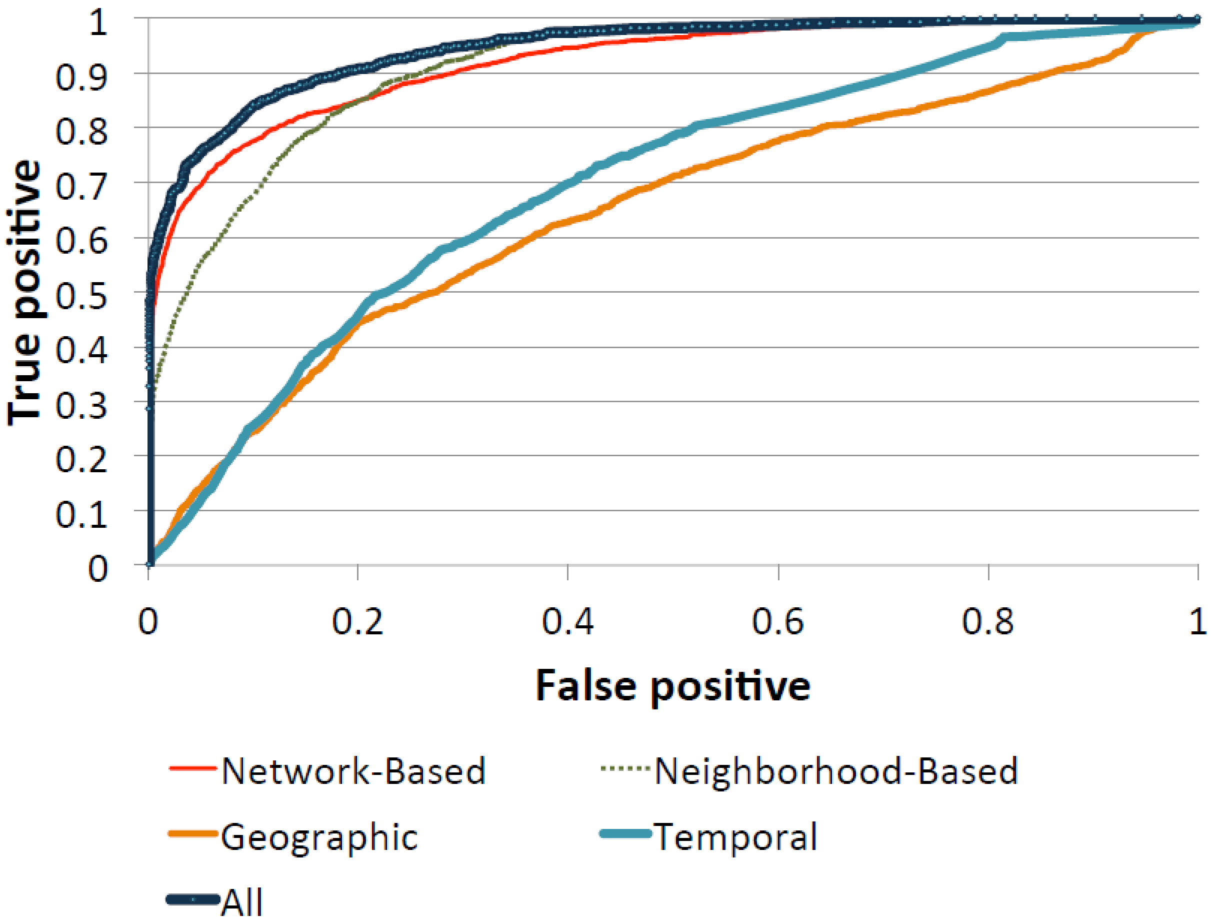
Arizona State University

Results: social network known



- Significant improvement in performance over currently-used method

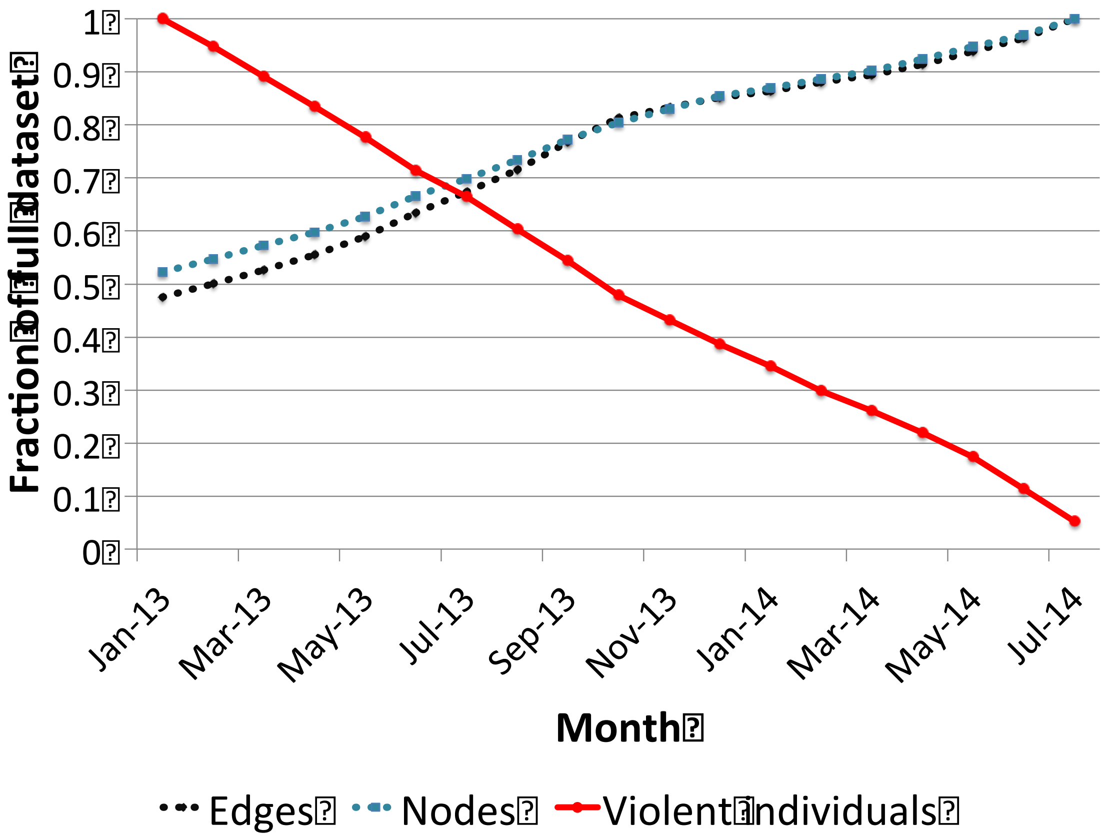
Social-Network Based Features



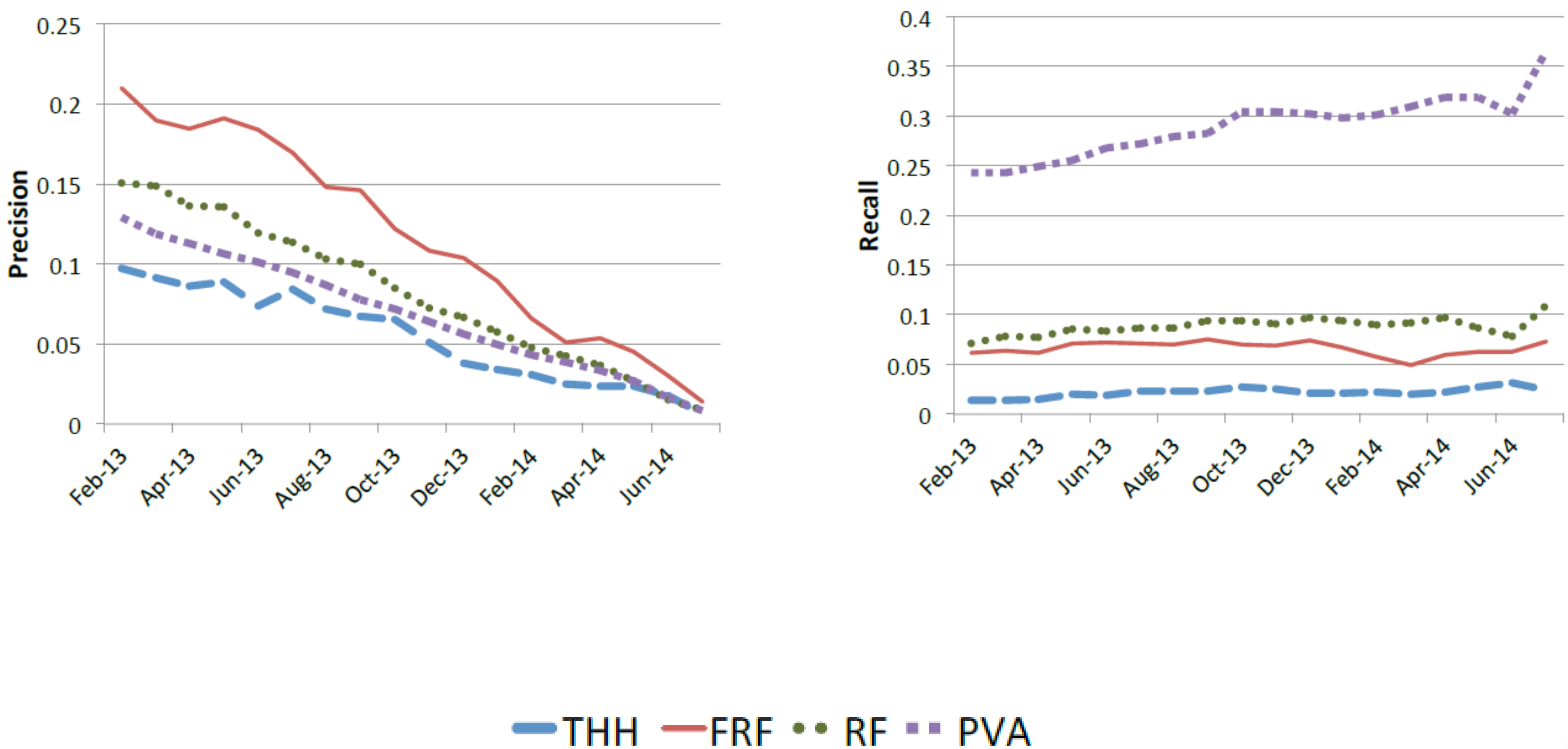
- Previous work in criminology focuses primarily on temporal and geographic features. We found network-based features to be more powerful

Results: Co-Arrestee Network Learned Over Time

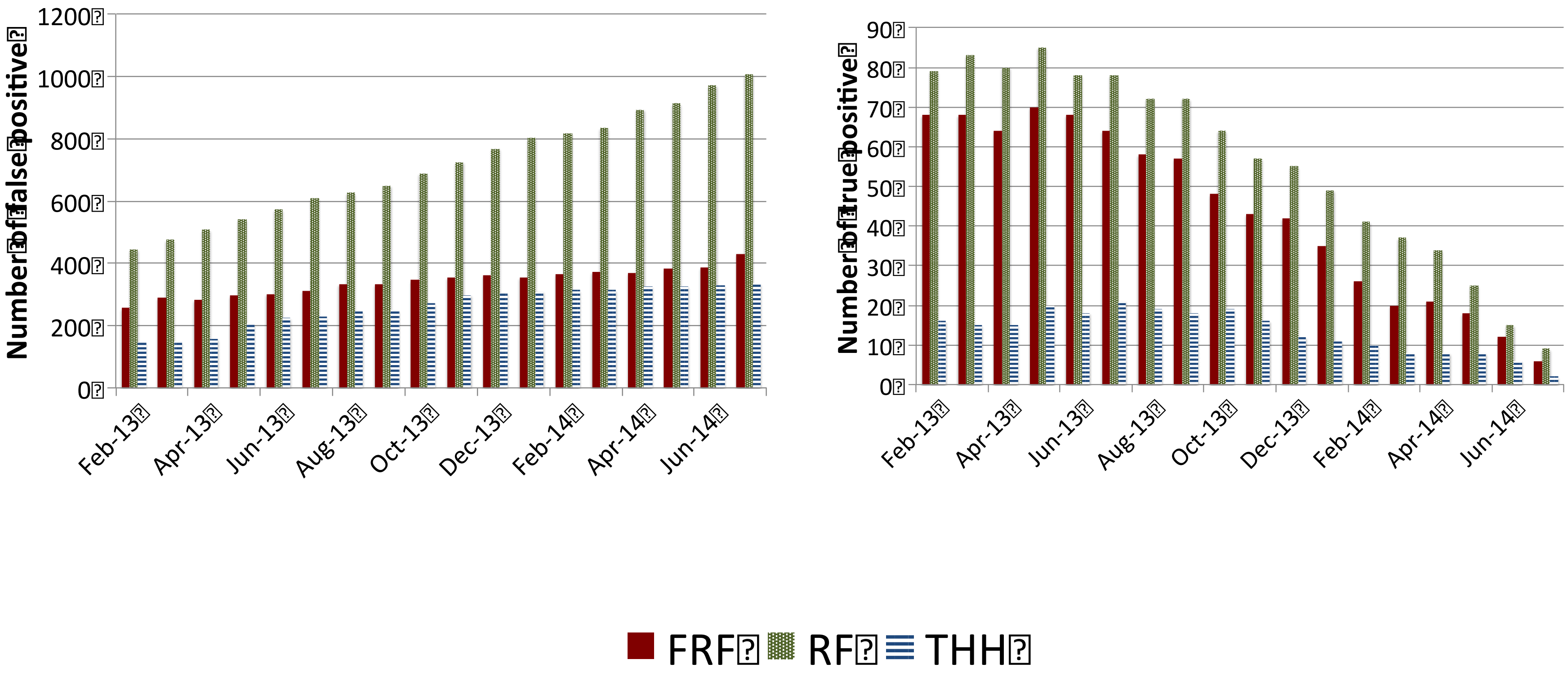
Network Properties



Results: social network learned over time



Results: social network learned over time



ARIZONA STATE UNIVERSITY

Ongoing works

- Now we are working with the Chicago Police Department to deploy this work in an operational setting.

- A provisional patent has also accepted

Conclusion

- Strong relationship between network-based features and violent crimes
- F1 score of 0.83 for the known social network
- Producing 4X more true positive if the network is discovered over time

Thank You