ISI 2011

Leveraging social networks to detect of anomalous insider actions in collaborative environments

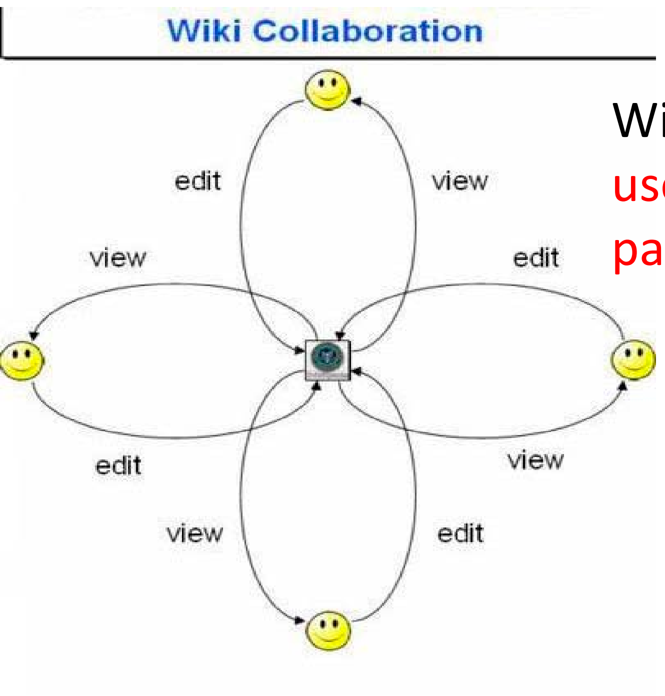
You Chen, Steve Nyemba, Wen Zhang, Bradley Malin

Health Information Privacy Lab, Vanderbilt University

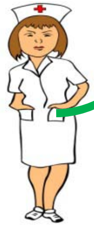
Outline

- Motivation
- Typical attacks
- Methods
- Experiments
- Limitations and future works

Motivation



Wiki Editing System: multiple users can view or edit common pages

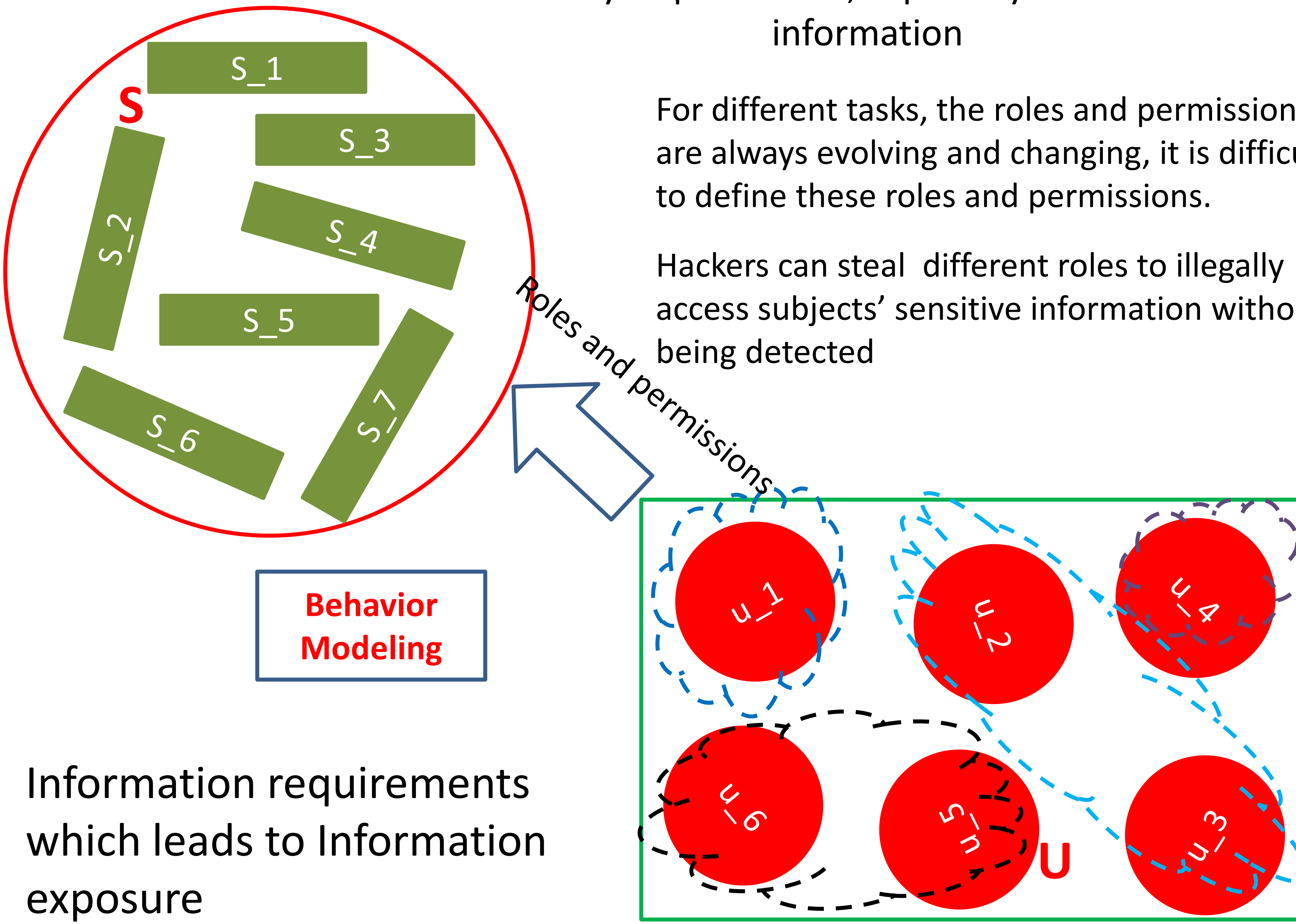


Electrical Health Records Systems: multiple users access common patients' records

Privacy requirements, especially for sensitive information

For different tasks, the roles and permissions are always evolving and changing, it is difficult to define these roles and permissions.

Hackers can steal different roles to illegally access subjects' sensitive information without being detected



Information requirements which leads to Information exposure

Outline

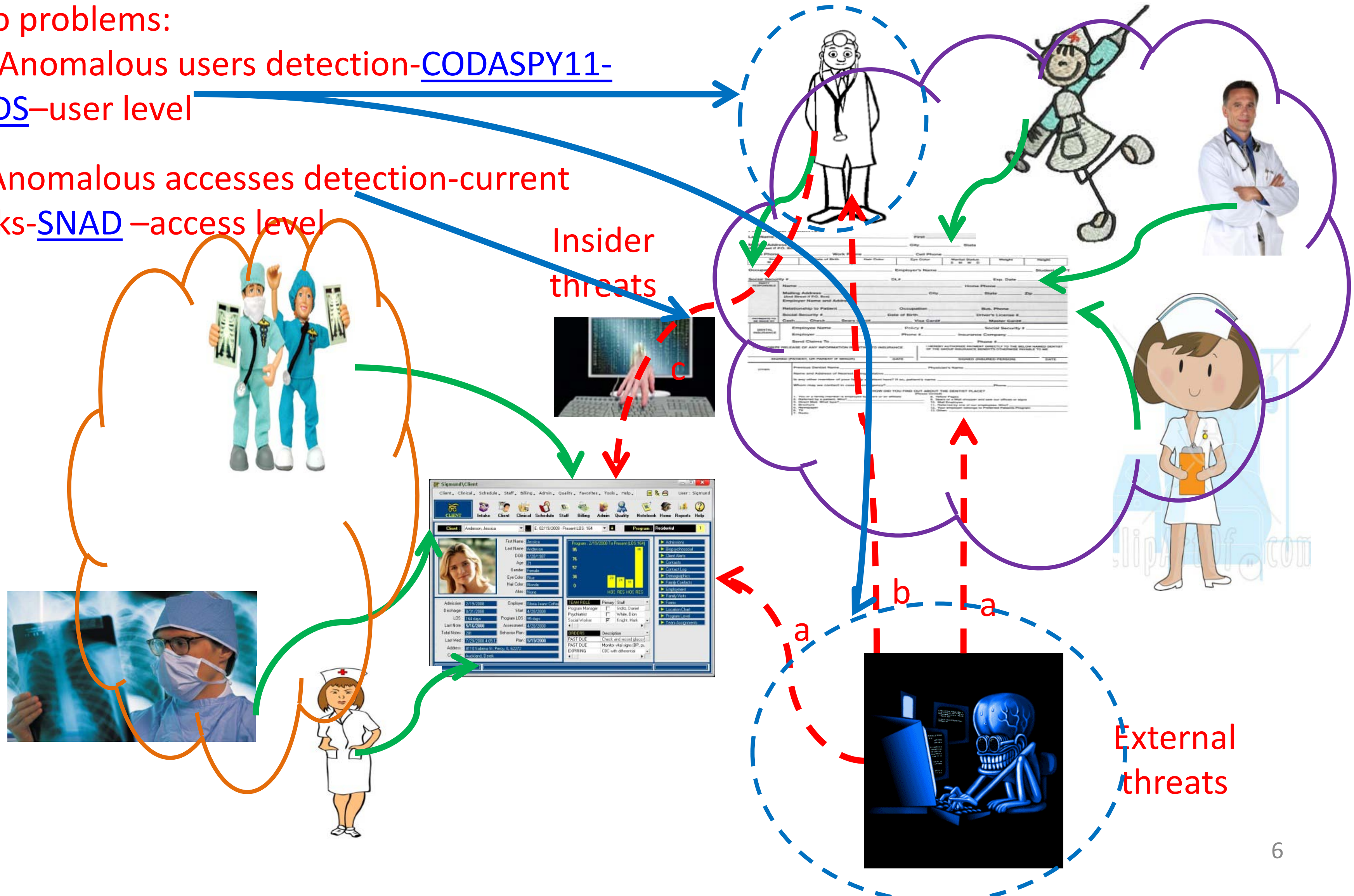
- Motivation
- Typical attacks
- Methods
- Experiments
- Limitations and future works

Typical Attacks

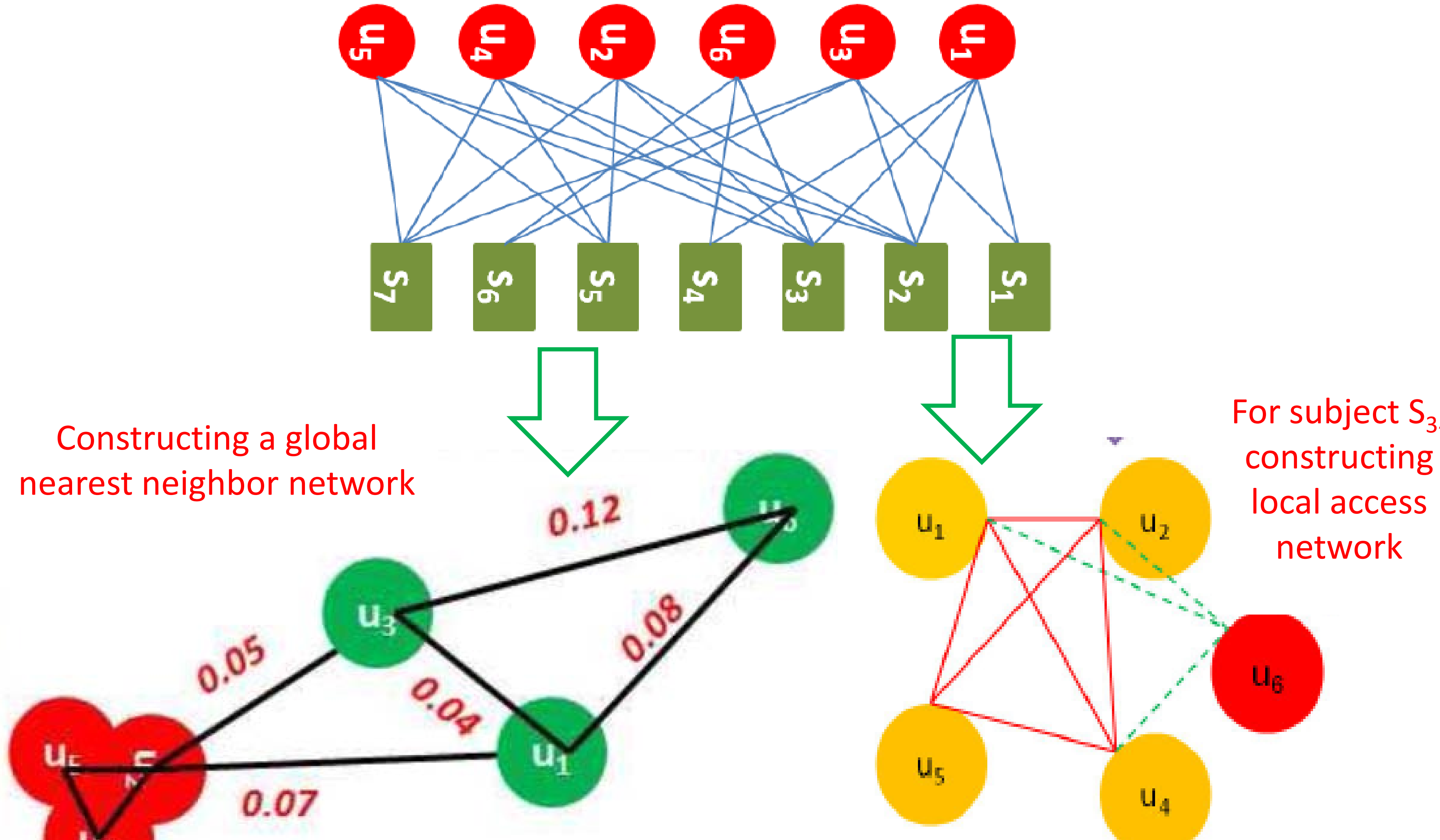
Two problems:

(1) Anomalous users detection-CODASPY11-CADS–user level

(2) Anomalous accesses detection-current works-SNAD –access level



Detecting anomalous users from a global view is a cool way to protect privacy of subjects information, and CADS is a typical model by using global social network



SNAD detects attacks from local view, and it detects more specific attacks such as anomalous accesses

Outline

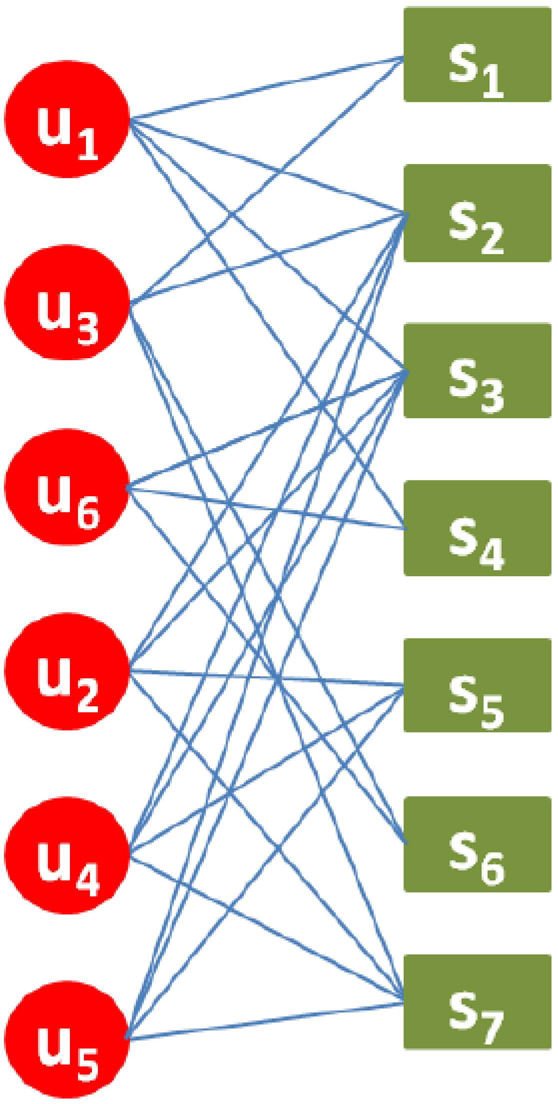
- Motivation
- Typical attacks
- Methods
- Experiments
- Limitations and Future works

Methods

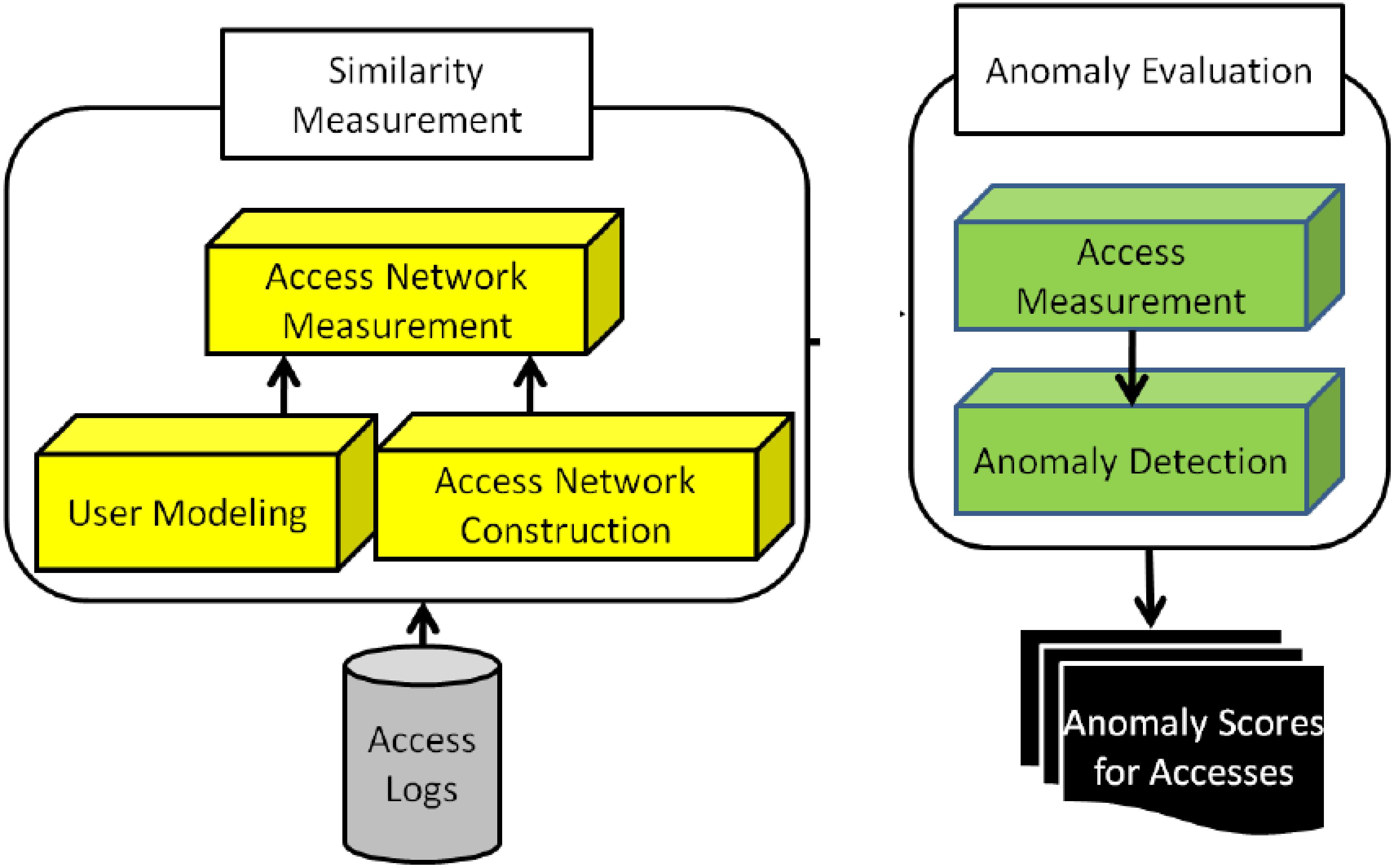
The methods were both based on access patterns which can be retrieved from a bipartite graph of users and subjects

SNAD: A specialized network anomaly detection model

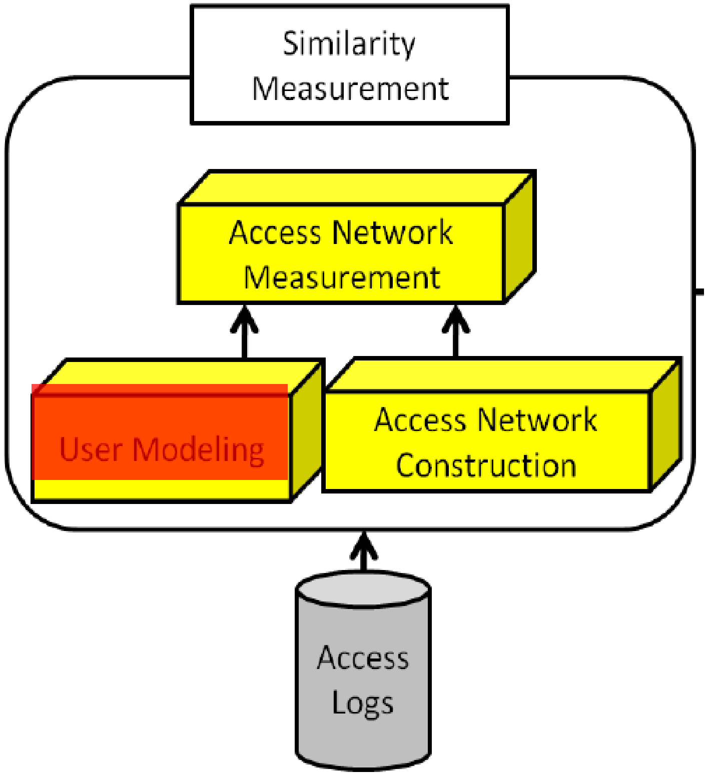
SNAD hypothesizes that if a user in the local network is anomalous, the similarity between this user to the network will be lower than the remaining users

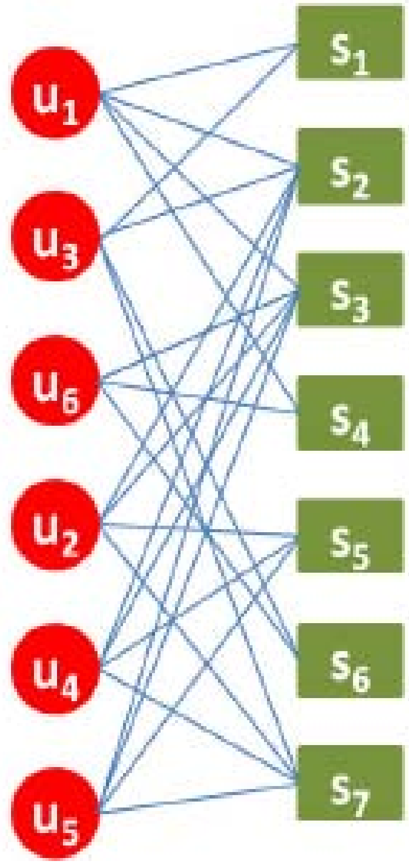


- Framework of SNAD



User Modeling



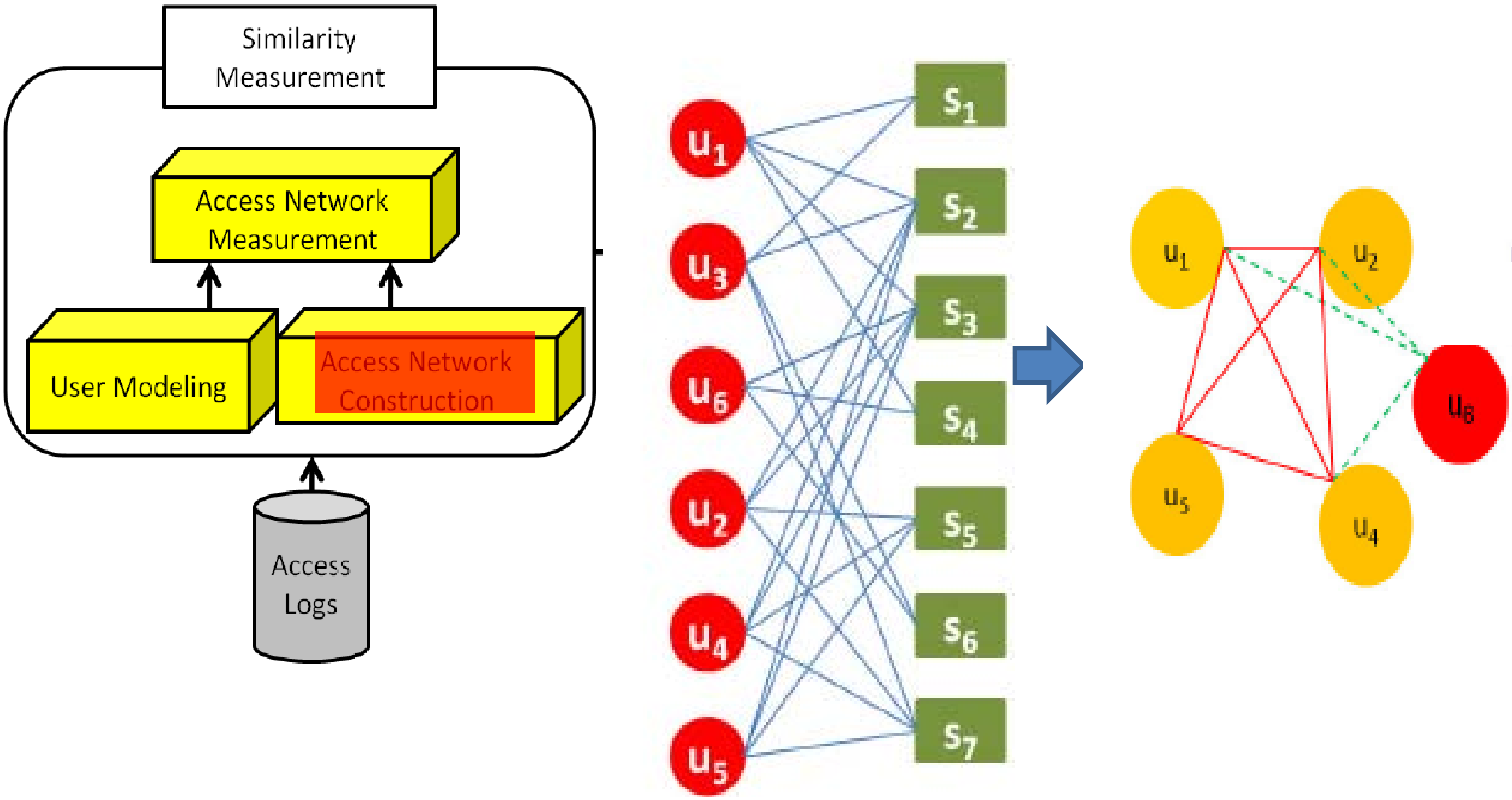


	u_1	u_2	u_3	u_4	u_5	u_6
s_1	1	0	1	0	0	0
s_2	1	1	1	1	1	0
s_3	1	1	0	1	1	1
s_4	1	0	0	0	0	1
s_5	0	1	0	1	1	0
s_6	0	0	1	0	0	1
s_7	0	1	1	1	1	0

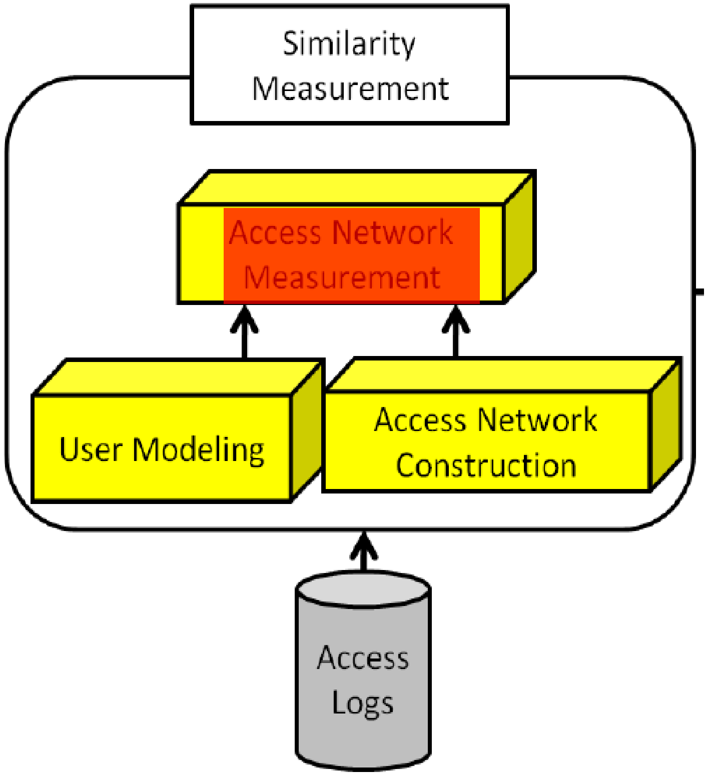
	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

$$IDF(u_i) = log\frac{|S|}{1 + |\{s_j, \quad where \quad SU(j,i) > 0\}|}$$

Access Network Construction



Access Network Measurement



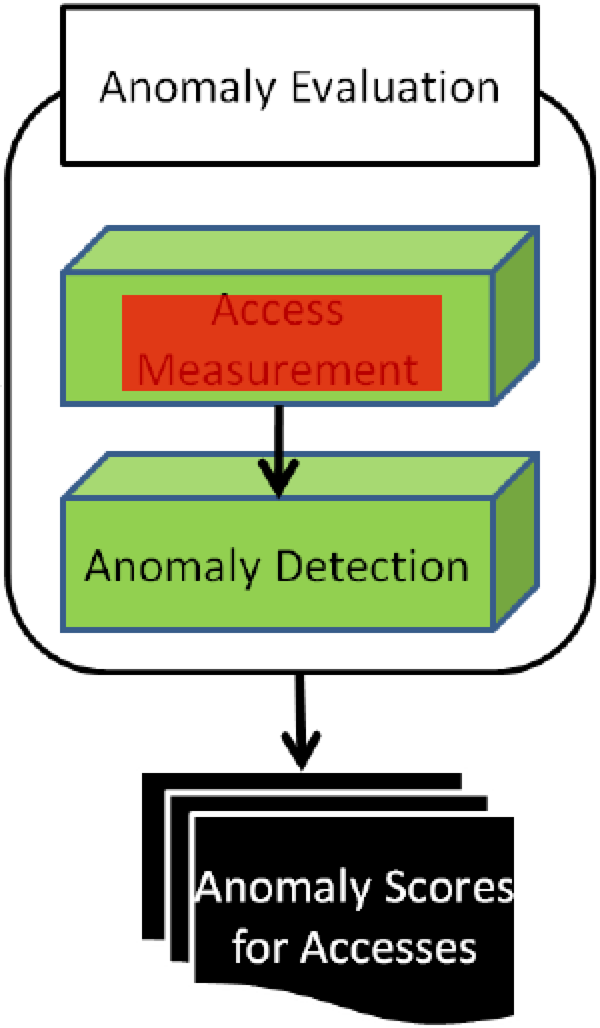
	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

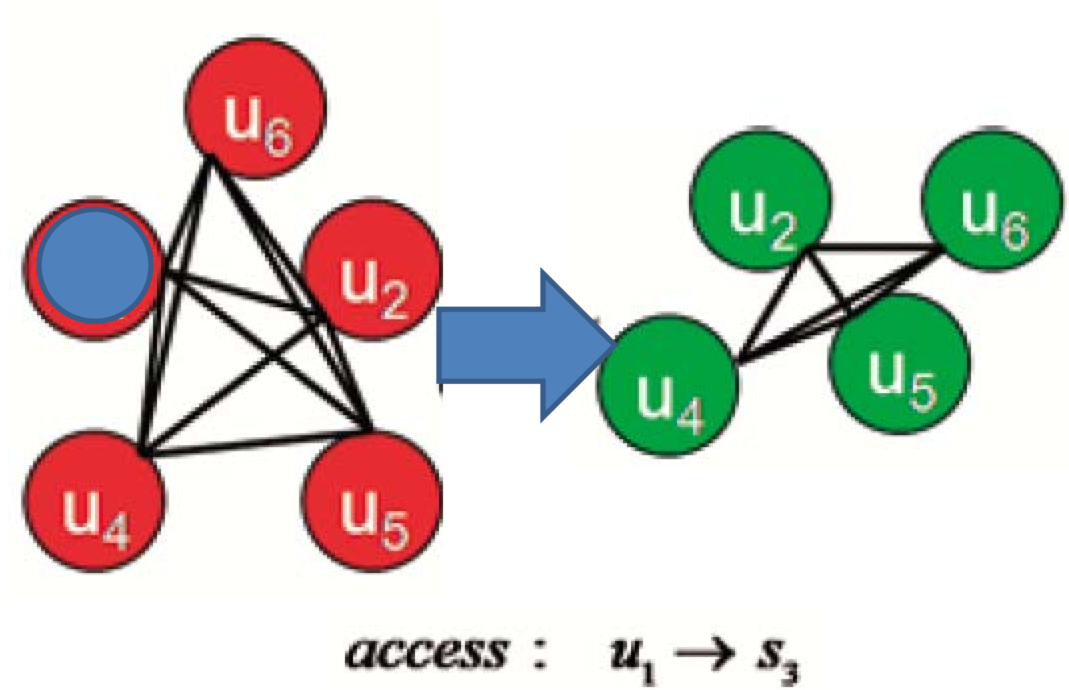
	u_1	u_2	u_4	u_5	u_6
u_1	1.00				
u_2	0.50	1.00			
u_4	0.50	1.00	1.00		
u_5	0.50	1.00	1.00	1.00	
u_6	0.58	0.29	0.29	0.29	1.00

$$Sim(u_i, u_j) = \frac{\mathbf{U_i} \cdot \mathbf{U_j}}{||\mathbf{U_i}|| \times ||\mathbf{U_j}||}$$

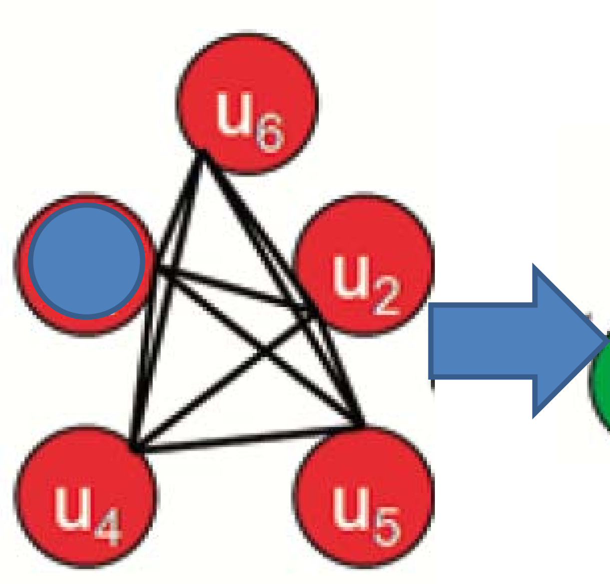
$$SIM(Net_{s_i}) = \frac{\sum_{i=1}^{|U_{s_i}|-1} \sum_{j=i+1}^{|U_{s_i}|} Sim(u_i, u_j)}{\frac{|U_{s_i}| \times (|U_{s_i}|-1)}{2}}$$

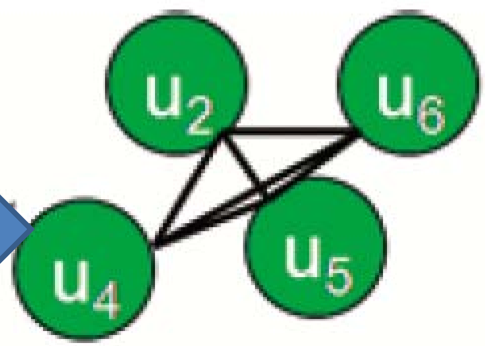
- Anomaly Evaluation - Access Measurement





- Anomaly Evaluation – Anomaly detection



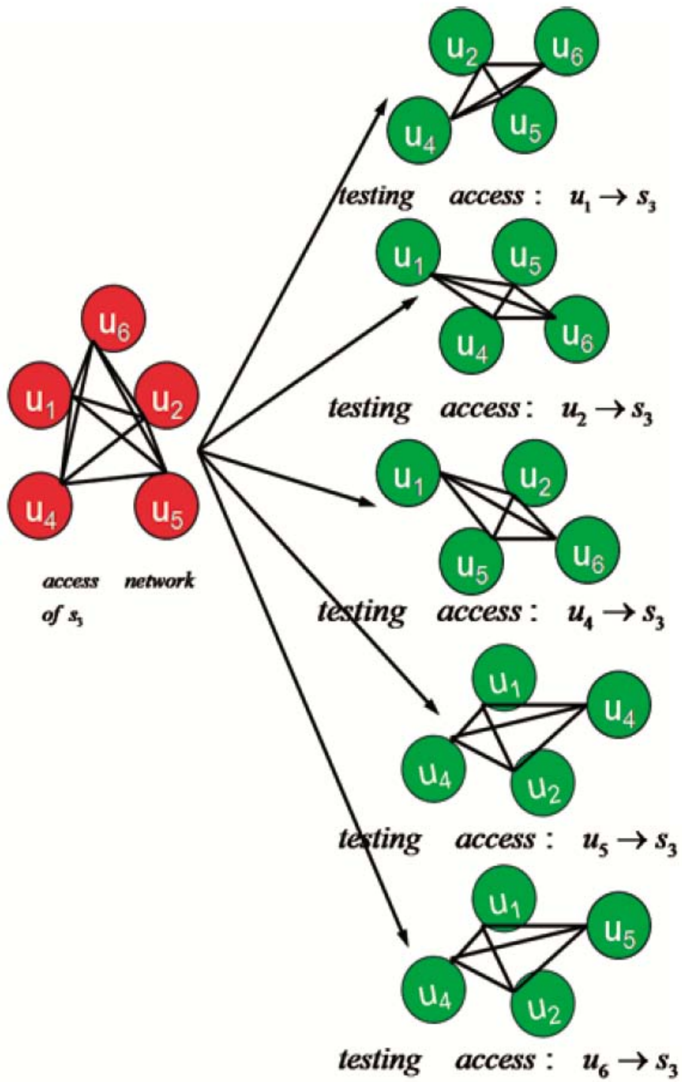


$access: \; u_1 \rightarrow s_3$

Network	Similarity	Size
u_1,u_2,u_4,u_5,u_6	0.59	5
u_2,u_4,u_5,u_6	0.64	4
u_1, u_4,u_5,u_6	0.52	4
u_1,u_2, u_5,u_6	0.52	4
u_1,u_2,u_4, u_6	0.52	4
u_1,u_2,u_4,u_5	0.75	4

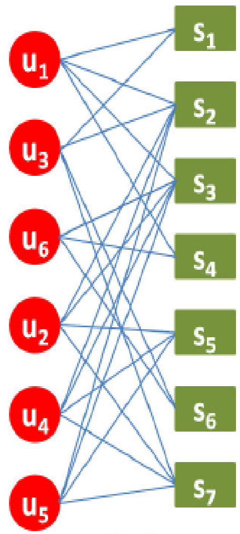
Access	Score	Size
u1-s3	0.05	4
u2-s3	-0.07	4
u4-s3	-0.07	4
u5-s3	-0.07	4
u6-s3	0.16	4

- Anomaly Evaluation



$$Score(u_j \to s_3) = SIM(Net_{s_{3j}}) - SIM(Net_{s_3})$$

- Compared Models: Spectral Anomaly Detection Model



(a) Bipartite graph of users and subjects

	u_1	u_2	u_3	u_4	u_5	u_6
s_1	1	0	1	0	0	0
s_2	1	1	1	1	1	0
s_3	1	1	0	1	1	1
s_4	1	0	0	0	0	1
s_5	0	1	0	1	1	0
s_6	0	0	1	0	0	1
s_7	0	1	1	1	1	0

(b) Binary matrix of subjects and users

Spectral decomposition binary matrix

Spectral analysis

Spectral decomposition on IDF matrix

	u_1	u_2	u_4	u_5	u_6
u_1	1.00				
u_2	0.50	1.00			
u_4	0.50	1.00	1.00		
u_5	0.50	1.00	1.00	1.00	
u_6	0.58	0.29	0.29	0.29	1.00

(d) Similarity matrix of pairs of users

	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

(c) IDF matrix of subjects and users

Spectral decomposition is a popular and general method to measure communities of users, so we here compare our model with this method

Experiments

- Motivation
- Typical attacks
- Methods
- Experiments
- Limitations and future works

Data Sets

Vanderbilt Medical Center. From a very large EHR system-patients' records are considered as subjects

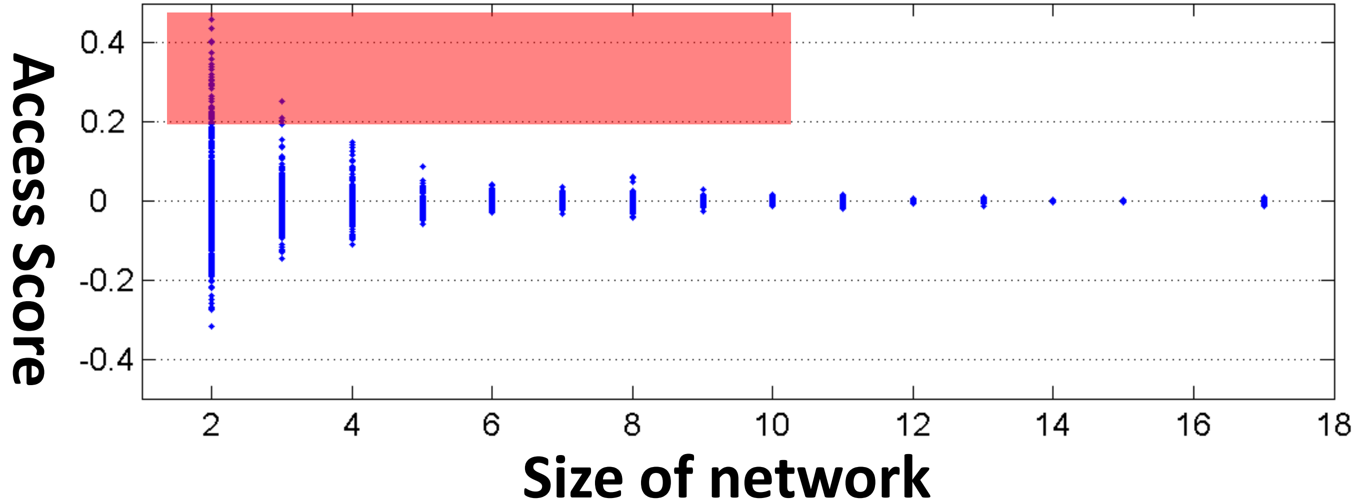
-Private dataset

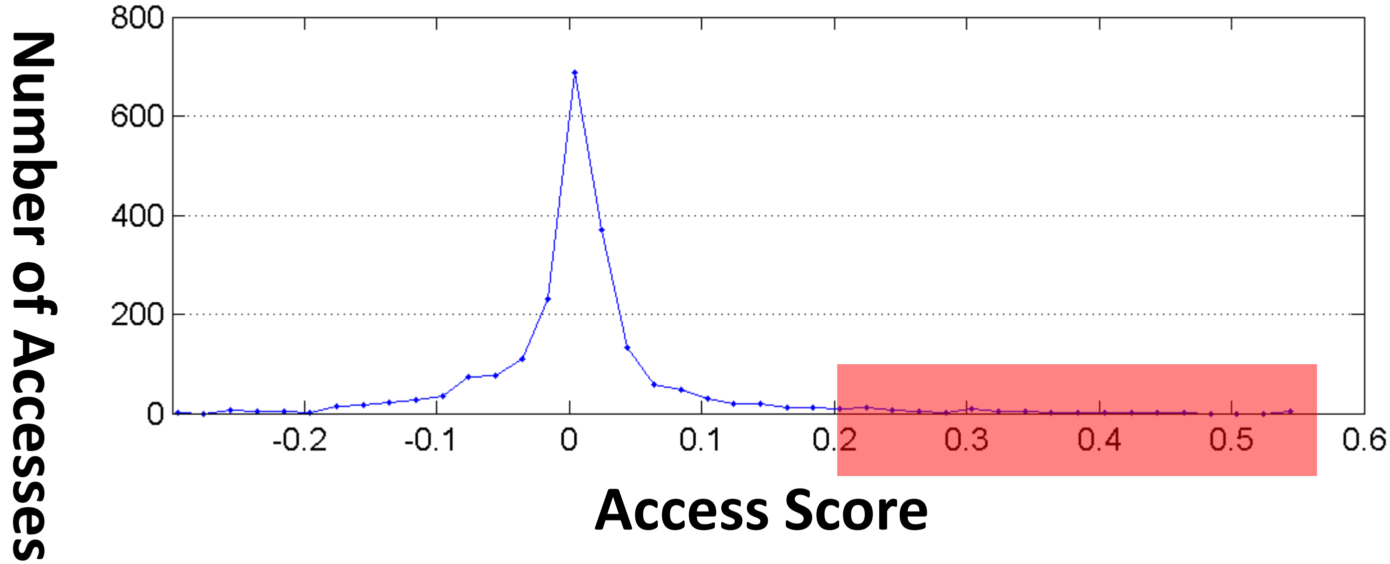
Wiki editing system-pages are considered as subjects

-Public dataset and can be replicated SNAD on this dataset

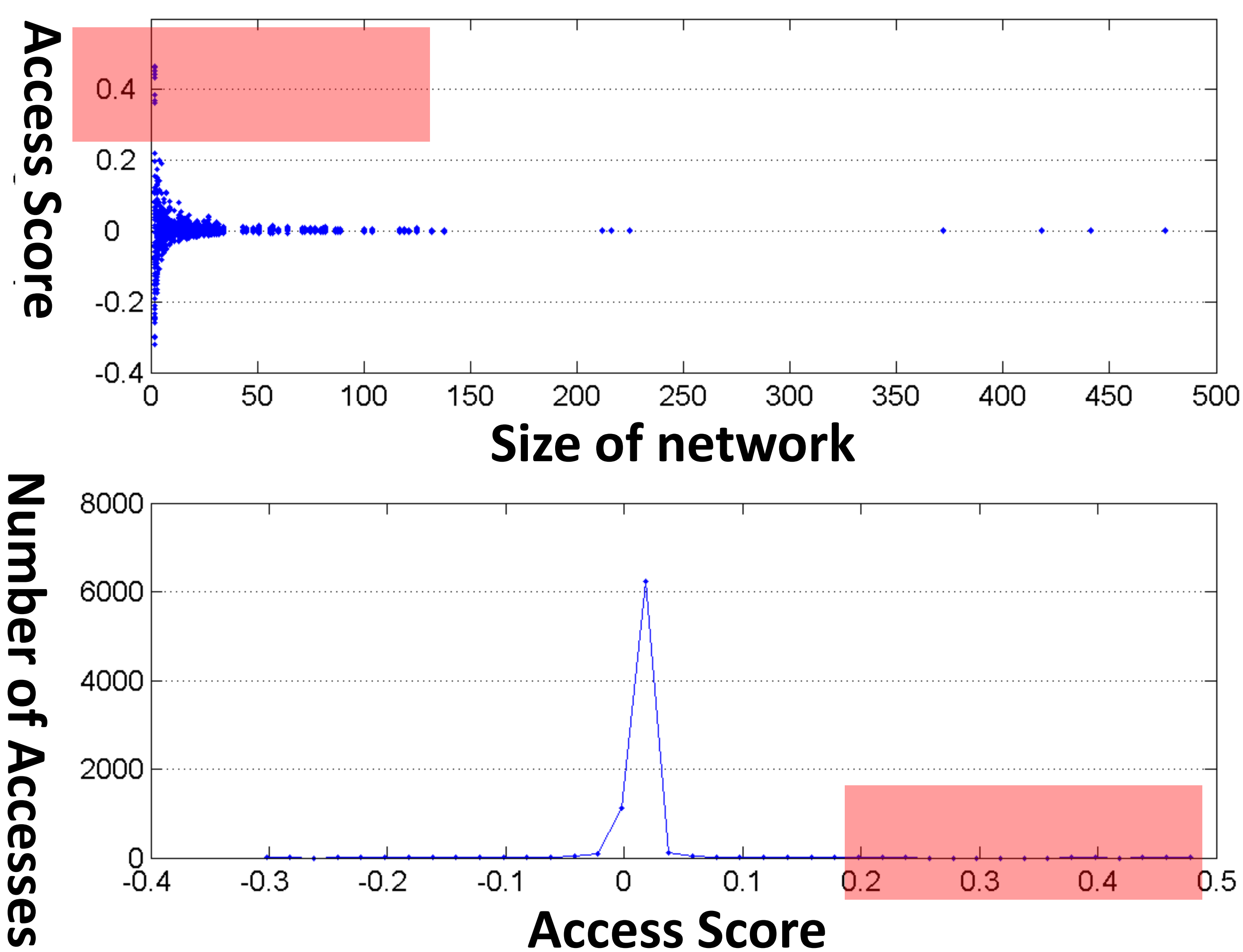
Dataset	Weeks	Users/week	Subjects/week	Accesses/week
EHR	30	2,281	13,148	44,250
Wiki	50	3,952	240	28,186

In EHR System-one week





In Wiki-one week



Evaluation

For a random user, verifying how number of simulated access injected into this user influence the performances of SNAD

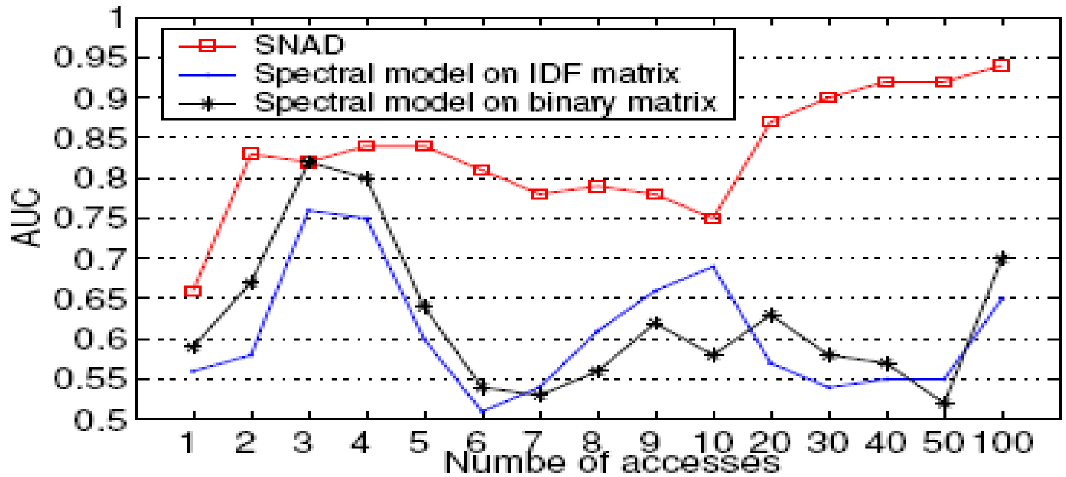
For a fixed number of simulated accesses, verifying how number of intruded users influence the performances of SNAD

The number of simulated accesses and intruded users are both diverse

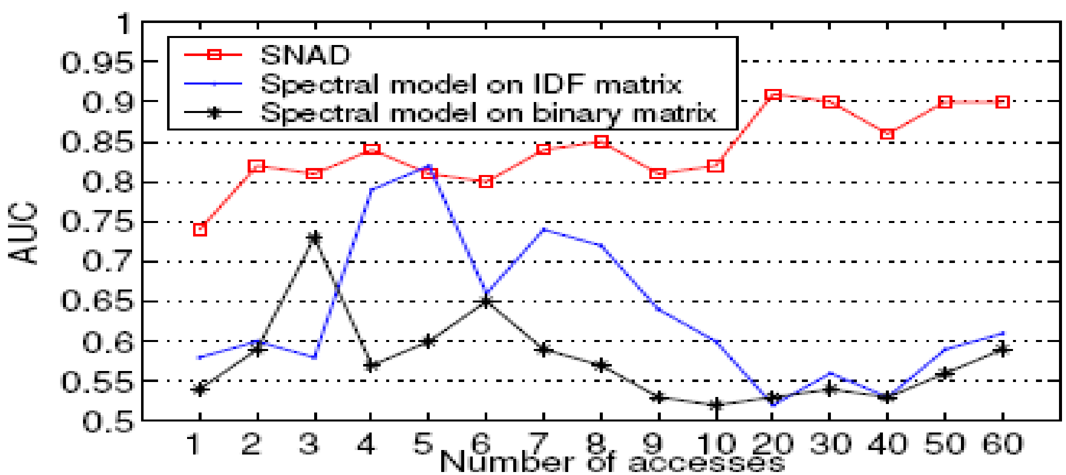
Model Evaluation-setting 1

For a random user, injecting simulated accesses

s_1	s_2	s_3	…	s_i	…	s_n	
0	1	0	…	0	…	0	
0	1	1	…	0	…	0	1
1	1	1	…	0	…	0	2
⋮							
1	1	1	…	1	…	1	100


SNAD
Spectral model on IDF matrix
Spectral model on binary matrix
AUC
Numbe of accesses

(a) EHR

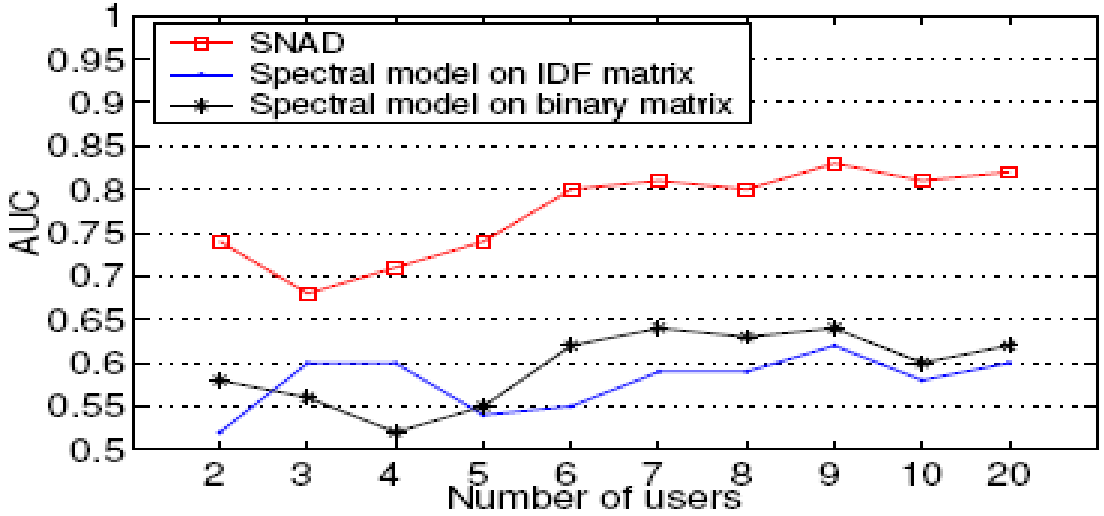

SNAD
Spectral model on IDF matrix
Spectral model on binary matrix
AUC
Number of accesses

(b) Wiki

Model Evaluation-setting 2

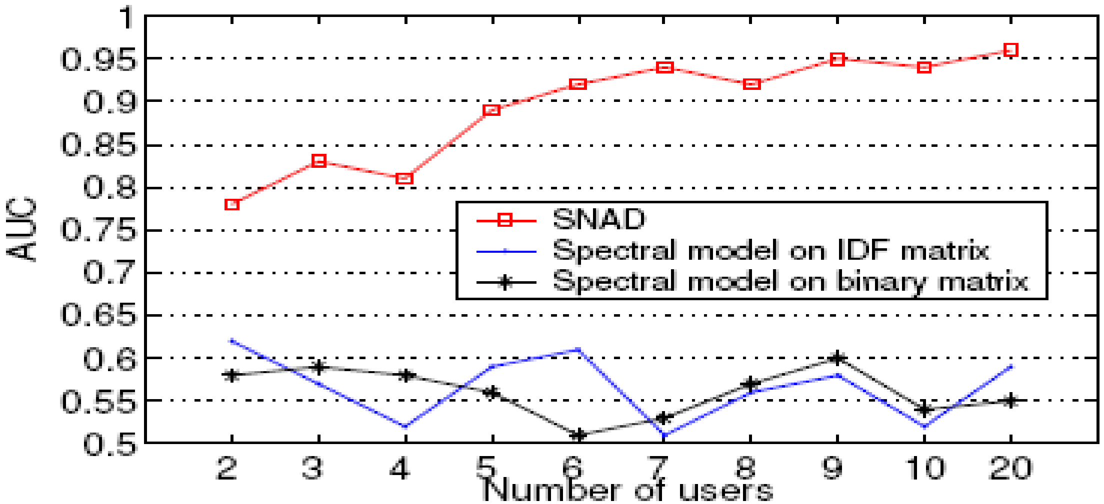
Fixing number of simulated accesses, number of intruders is random

s_1	s_2	s_3	…	s_i	…	s_n	
1	1	1	…	0	…	1	Intruder_1
0	1	1	…	1	…	1	Intruder_2
.							
.							
.							
1	1	1	…	1	…	0	Intruder_k



SNAD
Spectral model on IDF matrix
Spectral model on binary matrix
AUC
Number of users

(a) EHR



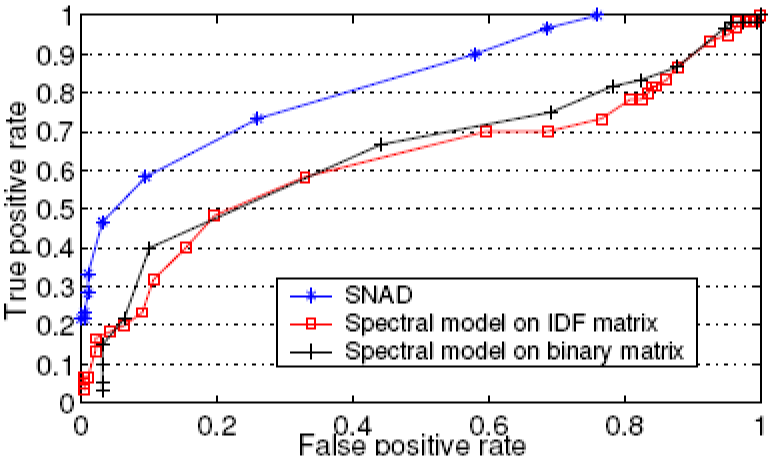
SNAD
Spectral model on IDF matrix
Spectral model on binary matrix
AUC
Number of users

(b) Wiki

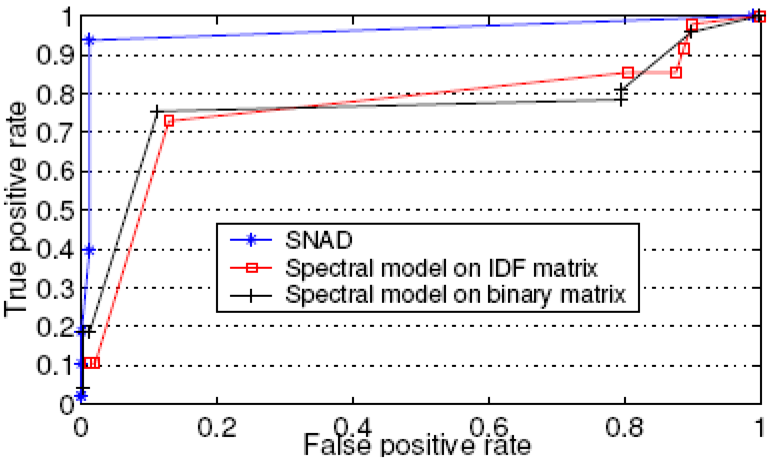
Model Evaluation-setting 3

Fixing number of simulated accesses, number of intruders is random

s_1	s_2	s_3	…	s_i	…	s_n	
1	1	0	…	0	…	0	Intruder_1
0	1	1	…	1	…	1	Intruder_2
			⋮				
0	1	1	…	1	…	0	Intruder_k



(a) EHR



(b) Wiki

Dataset	SNAD	Spectral IDF	Spectral Binary
EHR	0.83±0.03	0.74±0.06	0.69±0.05
Wiki	0.91±0.02	0.76±0.04	0.64±0.04

Experiments

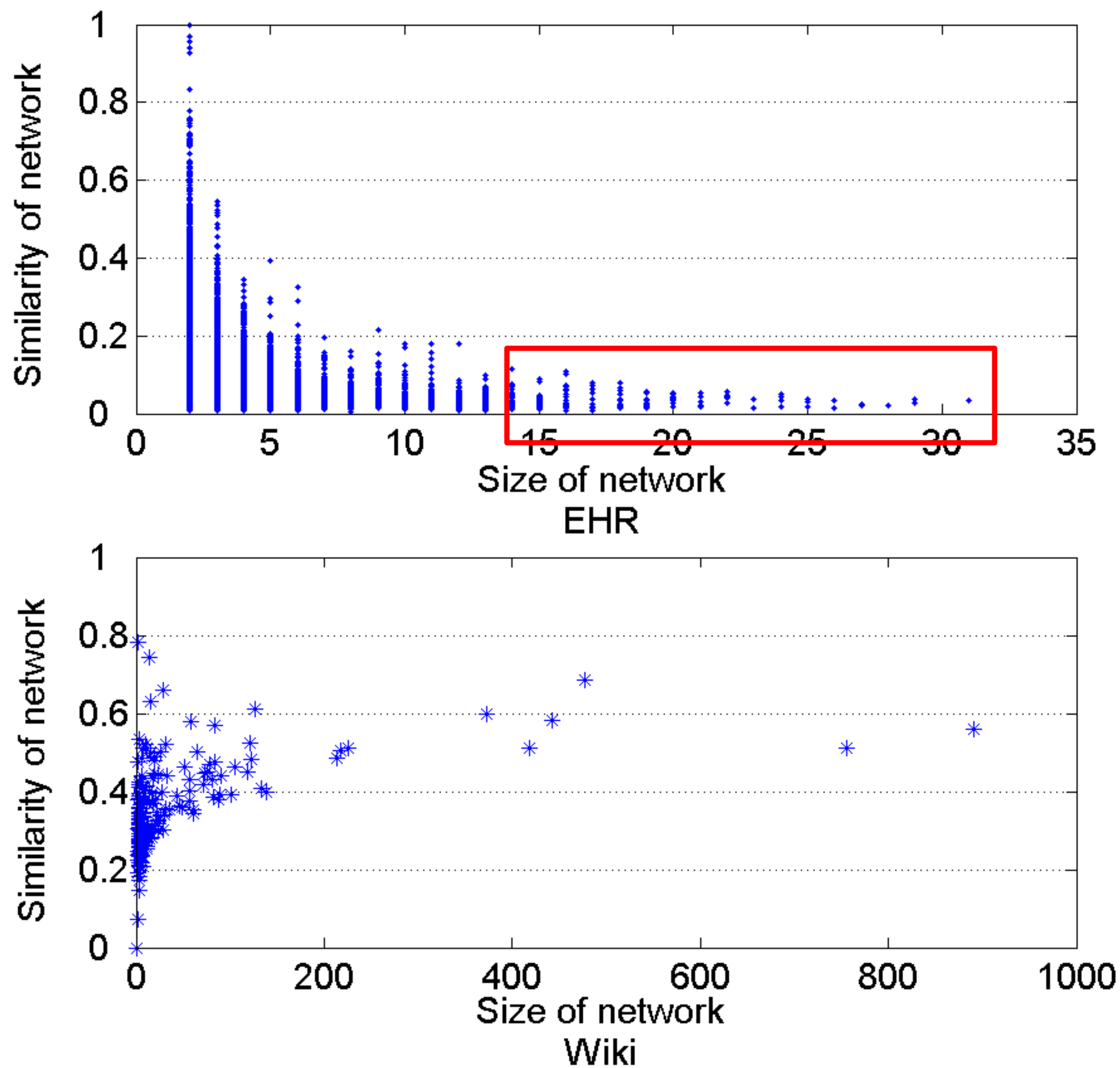
- Motivation
- Typical attacks
- Methods
- Experiments
- Limitations and future works

Limitations and Future Works

Only mining access patterns on users and subjects will lead to high false positive rates and low true positive rates

When mining normal patterns, we can include semantic information of users and subjects to decrease false positive rates and increase true positive rates

SNAD is not appropriate on large access network with low network similarity. In this case, the suppression of a user has little influence on the similarity of access network. In HER system and Wiki system, SNAD has high performance, since these systems are collaborative and access networks have high network similarity



1
0.8
0.6
0.4
0.2
0
Similarity of network
0
5
10
15
20
25
30
35
Size of network
EHR
1
0.8
0.6
0.4
0.2
0
Similarity of network
0
200
400
600
800
1000
Size of network
Wiki

Questions?

You.chen@vanderbilt.edu

This research was sponsored by grants CCF-0424422 and CNS-0964063 from the National Science Foundation and 1R01LM010207 from the National Institutes of Health.