Auditing EMR System Usage

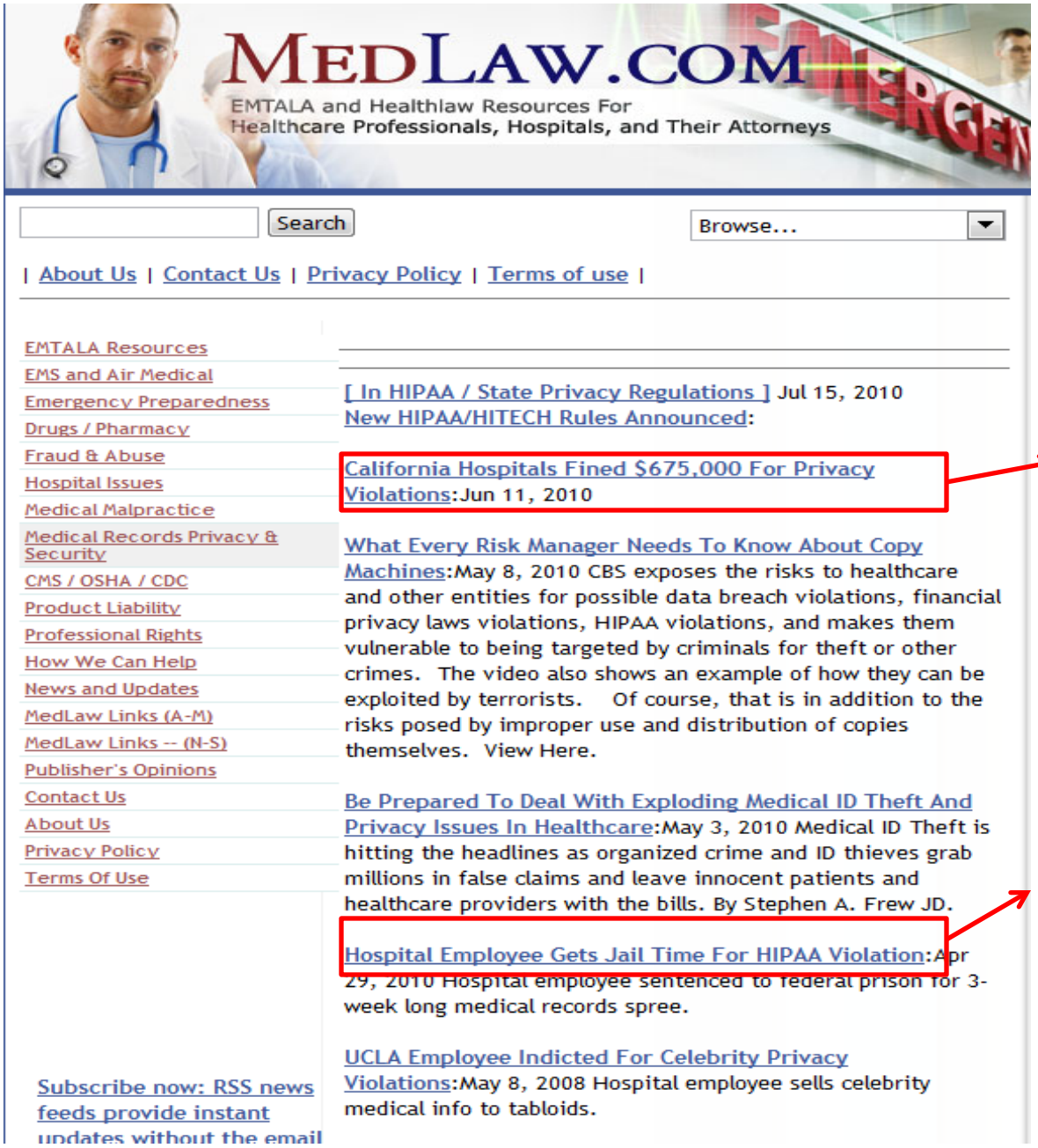
You Chen

Jan, 17, 2013

You.chen@vanderbilt.edu

Health data being accessed by hackers, lost with laptop computers, or simply read by curious employees

However, HIPAA's shortcomings and lack of clarity have fed the public's concern about the potential risks to privacy associated with having the most personal data imaginable stored in electronic format. Add to this, the nearly constant barrage of news stories about health data being accessed by hackers, lost with laptop computers, or simply read by curious employees, and it is little wonder consumers are concerned about privacy.



Current HIPAA Security Rules are not enough

$675,000 for Privacy Violation

Jail Time for **Malicious Accesses**

HIPAA Security Rules

- Administrative Safeguards
  - Administrative actions, and policies and procedures, to manage the selection, development, implementation, and maintenance of security measures to protect electronic protected health information (PHI)
- Physical Safeguards
  - Physical measures, policies and procedures to protect a covered entity's electronic information systems and related buildings and equipment, from natural and environmental hazards, and unauthorized intrusion
- Technical Safeguards
  - The technology and the policy and procedures for its use that protect electronic protected health information [PHI] and control access to it

Current State

Monitor VIPs (the Clooney effect- finding more attractive man)

Monitor employee-employee access

Follow-up on external suspicion

Spot checks

Technical Safeguards

- Access Control
- Audit controls: Implement systems to record and audit access to protected health information within information systems
  - Track & audit employees access to patient records
  - Store logs for $\geq$ 6 years

Access Control?

- "We have *-Based Access Control."
- "We have a mathematically rigorous access policy logic!"
- "We can specify {*context*, *team*, *temporal*} policies!" (Georgiadis et al, 2001; Park et al. 2001;)
- "We can control your access at a fine-grained level!"
- "Isn't that enough?"

C. Georgiadis, I. Mavridis, G. Pangalos, and R. Thomas. Flexible team-based access control using contexts. *Proceedings of ACM Symposium on Access Control Model and Technology*. 2001: 21-27.

J. Park, R. Sandhu, and G. Ahn. Role-based access control on the web. *ACM Transactions on Information and System Security*. 2001; 4(1): 37-71.

Why is the Problem So Hard?

- Hospital system is inherently dynamic

Roles

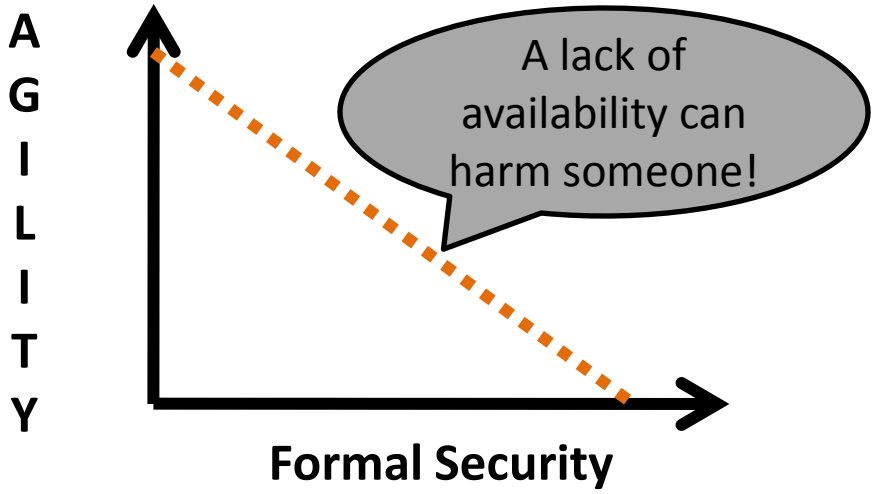
- Multiple responsibilities
- Fuzzy "
- Changing "

Teams

- Situation dependent
-Constrained by availability
"Can you work today?"

Reason

- Who defines the policies?
- Often Vague
- Cluttered by "multi-use" systems



But If You Let Them, They Will Come

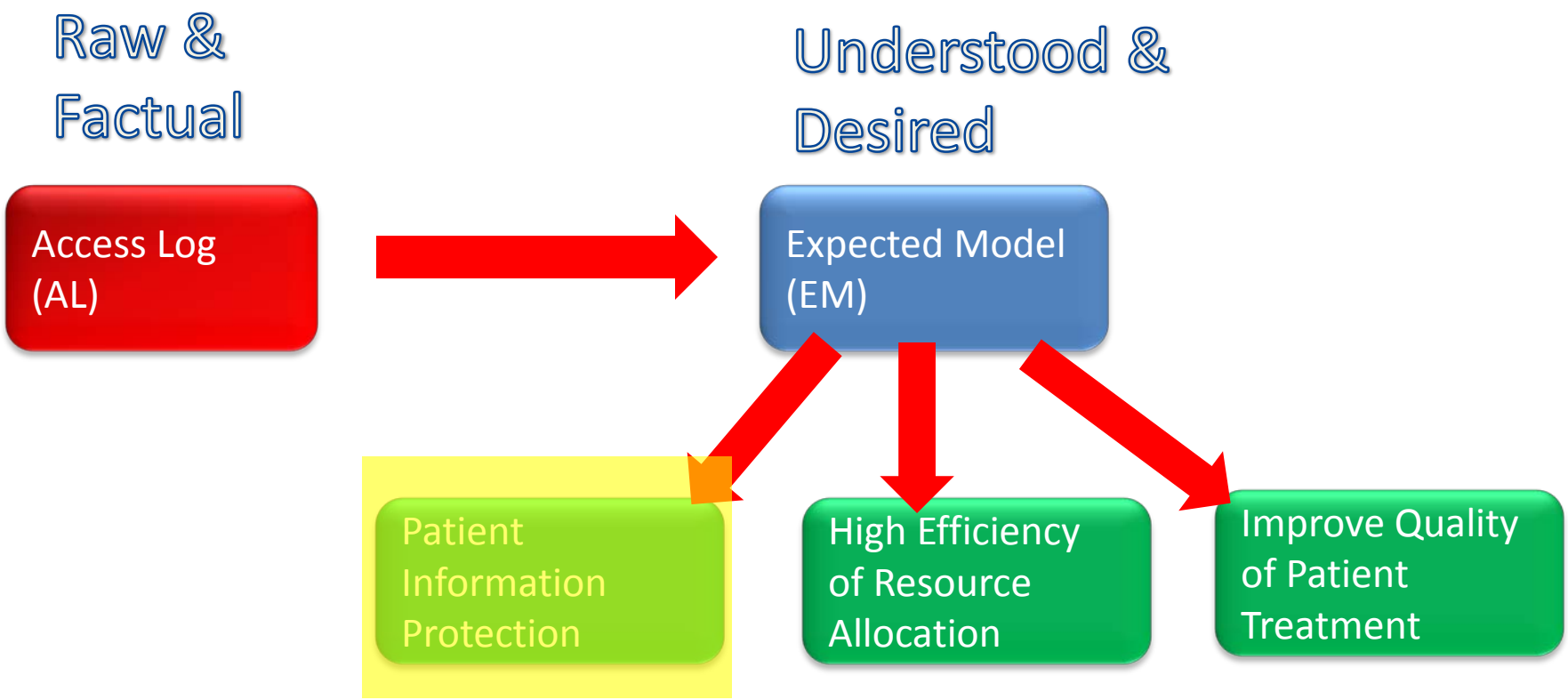
- In March 2006, researchers carried out an investigation on hospitals in the Central Norway Health Region
- Users were assigned to an initial set of privileges and could invoke actualization, temporarily escalating their rights as necessary
- Such an access control system is feasible when the number of actualizations is small

This case:
•53,650 of 99,352 patients actualized
•5,310 of 12,258 users invoked actualization
•Over 295,000 actualizations in one month

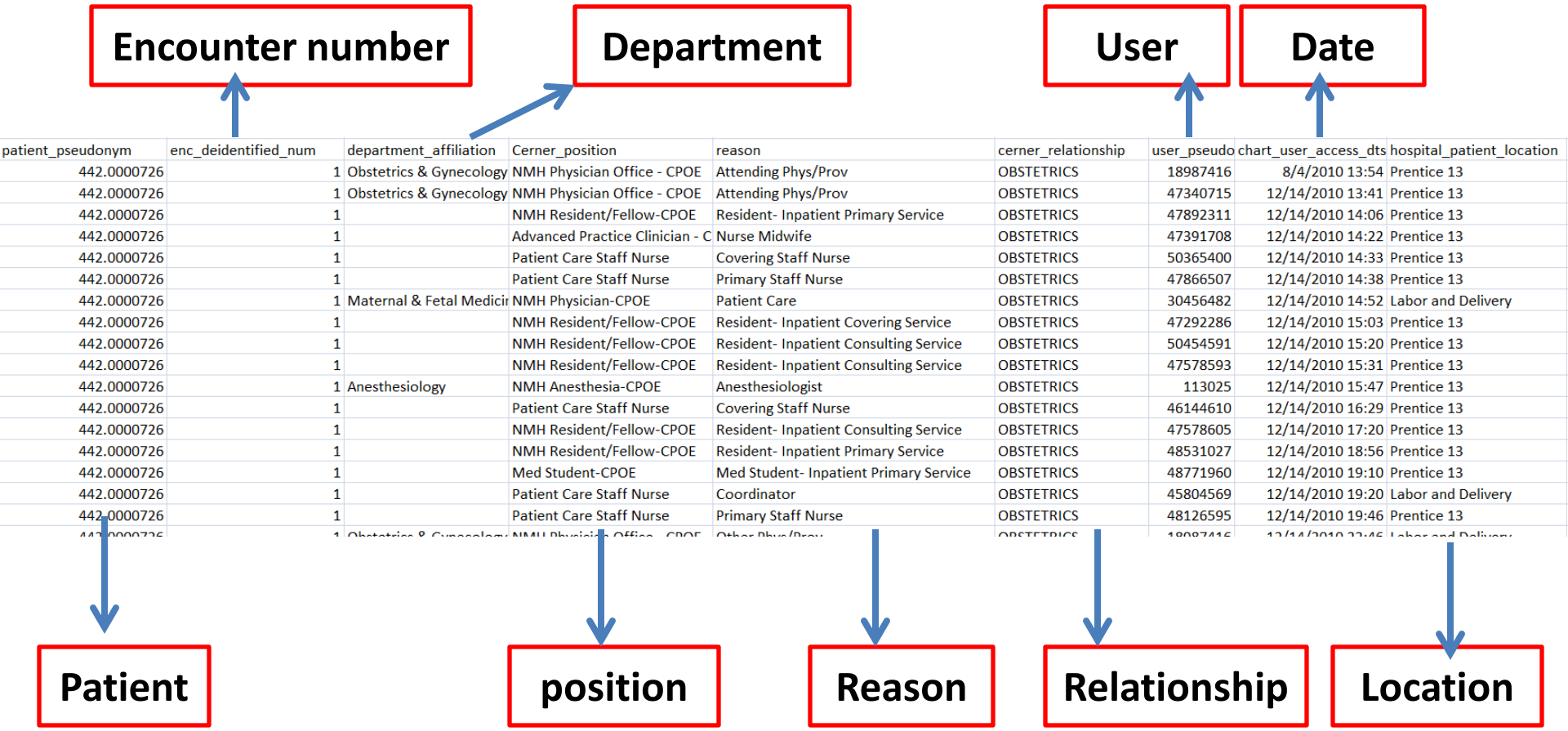
Role	Users	Invoked Actualization in Past Month
Nurse	5633	36%
Doctor	2927	52%
Health Secretary	1876	52%
Physiotherapist	382	56%
Psychologist	194	58%

L. Røstad and N. Øystein. Access control and integration of health care systems: an experience report and future challenges. *Proceedings of the 2nd International Conference on Availability, Reliability and Security (ARES)*. 2007: 871-878.

Aim of Access Logs Auditing



Examples of Accesses



patient_pseudonym	enc_deidentified_num	department_affiliation	Cerner_position	reason	cerner_relationship	user_pseudo	chart_user_access_dts	hospital_patient_location
442.0000726	1	Obstetrics & Gynecology	NMH Physician Office - CPOE	Attending Phys/Prov	OBSTETRICS	18987416	8/4/2010 13:54	Prentice 13
442.0000726	1	Obstetrics & Gynecology	NMH Physician Office - CPOE	Attending Phys/Prov	OBSTETRICS	47340715	12/14/2010 13:41	Prentice 13
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Primary Service	OBSTETRICS	47892311	12/14/2010 14:06	Prentice 13
442.0000726	1		Advanced Practice Clinician - C	Nurse Midwife	OBSTETRICS	47391708	12/14/2010 14:22	Prentice 13
442.0000726	1		Patient Care Staff Nurse	Covering Staff Nurse	OBSTETRICS	50365400	12/14/2010 14:33	Prentice 13
442.0000726	1		Patient Care Staff Nurse	Primary Staff Nurse	OBSTETRICS	47866507	12/14/2010 14:38	Prentice 13
442.0000726	1	Maternal & Fetal Medicir	NMH Physician-CPOE	Patient Care	OBSTETRICS	30456482	12/14/2010 14:52	Labor and Delivery
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Covering Service	OBSTETRICS	47292286	12/14/2010 15:03	Prentice 13
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Consulting Service	OBSTETRICS	50454591	12/14/2010 15:20	Prentice 13
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Consulting Service	OBSTETRICS	47578593	12/14/2010 15:31	Prentice 13
442.0000726	1	Anesthesiology	NMH Anesthesia-CPOE	Anesthesiologist	OBSTETRICS	113025	12/14/2010 15:47	Prentice 13
442.0000726	1		Patient Care Staff Nurse	Covering Staff Nurse	OBSTETRICS	46144610	12/14/2010 16:29	Prentice 13
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Consulting Service	OBSTETRICS	47578605	12/14/2010 17:20	Prentice 13
442.0000726	1		NMH Resident/Fellow-CPOE	Resident- Inpatient Primary Service	OBSTETRICS	48531027	12/14/2010 18:56	Prentice 13
442.0000726	1		Med Student-CPOE	Med Student- Inpatient Primary Service	OBSTETRICS	48771960	12/14/2010 19:10	Prentice 13
442.0000726	1		Patient Care Staff Nurse	Coordinator	OBSTETRICS	45804569	12/14/2010 19:20	Labor and Delivery
442.0000726	1		Patient Care Staff Nurse	Primary Staff Nurse	OBSTETRICS	48126595	12/14/2010 19:46	Prentice 13

Examples of Patient Diagnose Codes

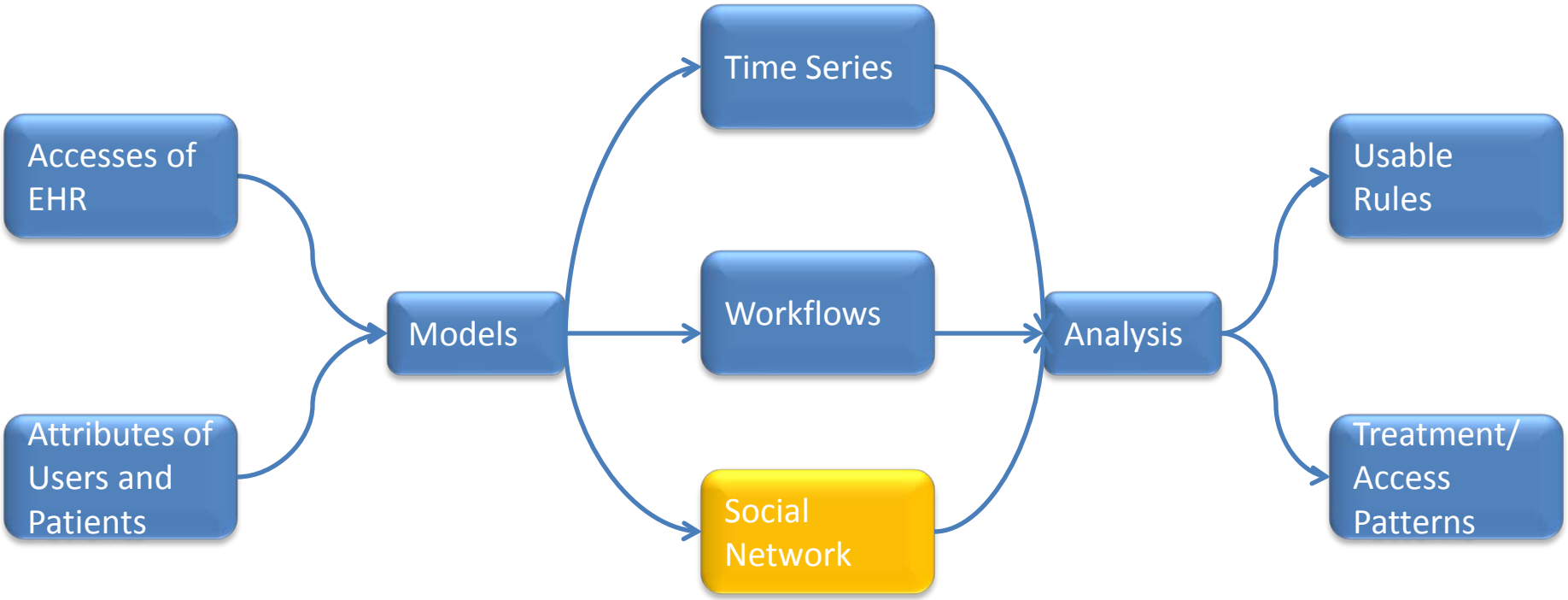
Encounter number

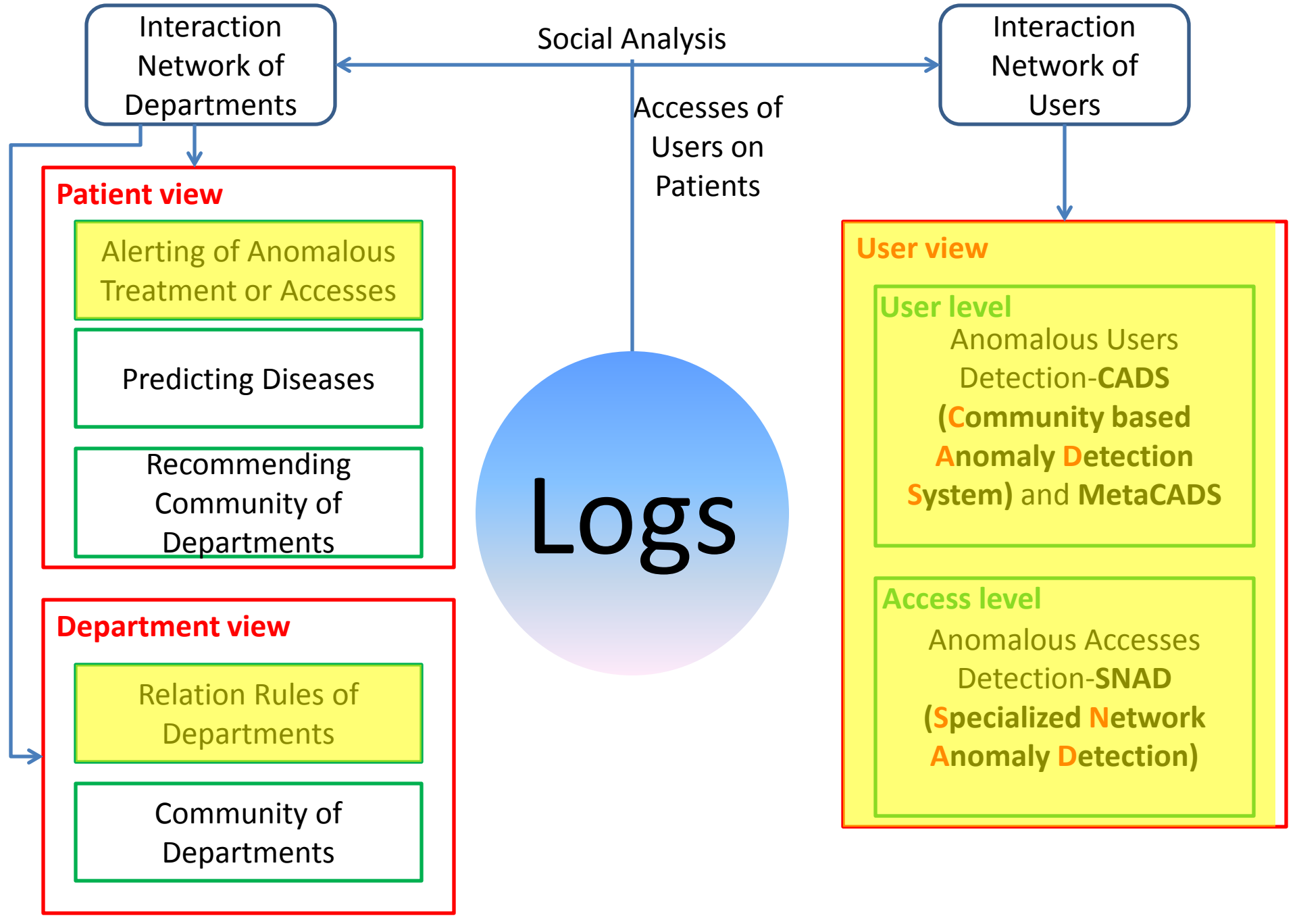
Diagnose codes

patient_study_id	enc_deiden	DX_codes
442.0000726	1	2165 , 65921, 65951, 66401, 66481, V270
442.0001714	1	V053 , V3000
442.0002396	1	4019 , 4111 , 41401, 4142 , 4739 , 49390
442.0002775	2	1122 , 20300, 25000, 27651, 40390, 5845 , 5859 , 591 , 5933 , V1005, V1046
442.0002775	1	1534 , 185 , 1962 , 1974 , 20300, 25000, 2809 , 40390, 56089, 5849 , 5859 , 59080, 591 , 78791, 7907
442.0003301	1	76408, 76529, V053 , V3100
442.0004873	1	V270 , V8535, 27800, 64911, 64971, 65841, 66401
442.0005024	1	4019 , 72252, V1582
442.0005968	1	5990 , 2724 , 311 , 4019 , 44022
442.0006352	1	65971, V270
442.0007008	1	25000, 6144 , 99859, V1042
442.0007371	1	V707 , 2859 , 33394, 4019 , 71690, 74190, V420
442.0007707	1	30000, 49121, 51889, 60000, 7850
442.0007707	2	78052, V1083, 30001, 496 , 0549 , 1120 , 2768 , 30000
442.0008016	1	V053 , V3001
442.0008405	1	2449 , 25080, 4019 , 41400, 42731, 4280 , 60000, V4581, V5861
442.0009617	1	V053 , V3000

Patient

Various Ways of Access Logs Auditing




Interaction Network of Departments
Social Analysis
Interaction Network of Users
Accesses of Users on Patients
Patient view
Alerting of Anomalous Treatment or Accesses
Predicting Diseases
Recommending Community of Departments
Department view
Relation Rules of Departments
Community of Departments
Logs
User view
User level
Anomalous Users Detection-CADS (Community based Anomaly Detection System) and MetaCADS
Access level
Anomalous Accesses Detection-SNAD (Specialized Network Anomaly Detection)

Uncovering Anomalous Usage of Medical Records via Social Network Analysis

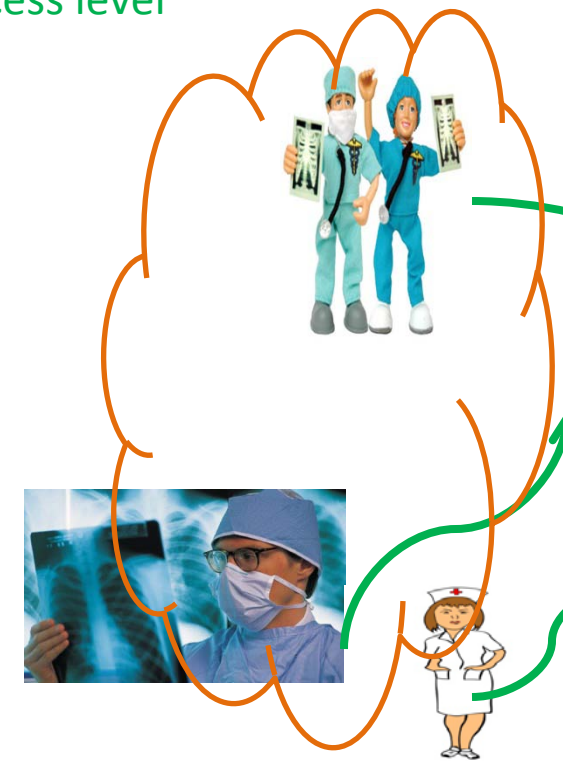
Two Typical Attacks

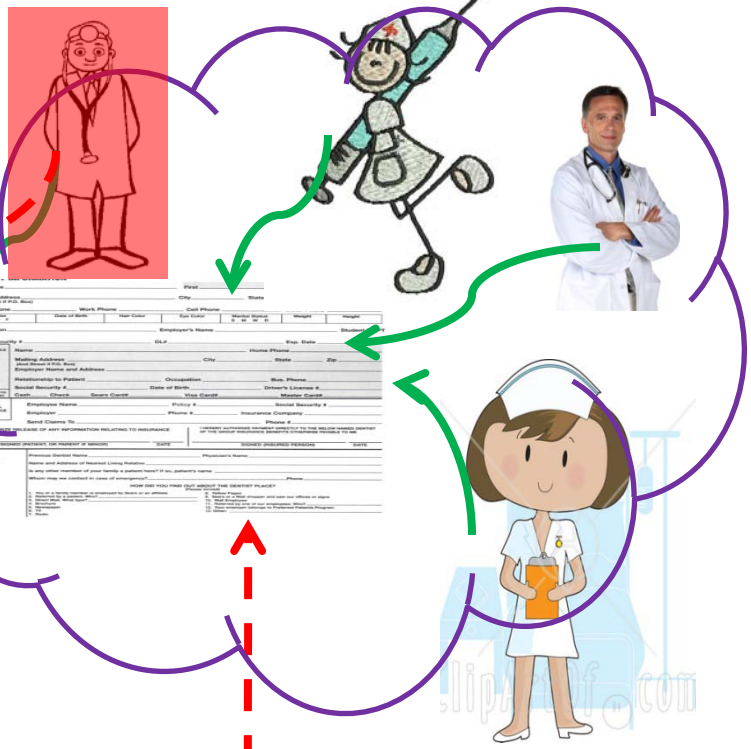
Intruders have complete knowledge of the system and its policies

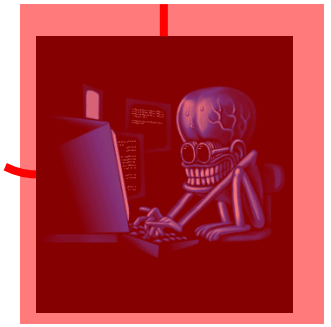
(1) Anomalous users detection –user level

(2) Anomalous accesses detection –access level

Anomalous Access



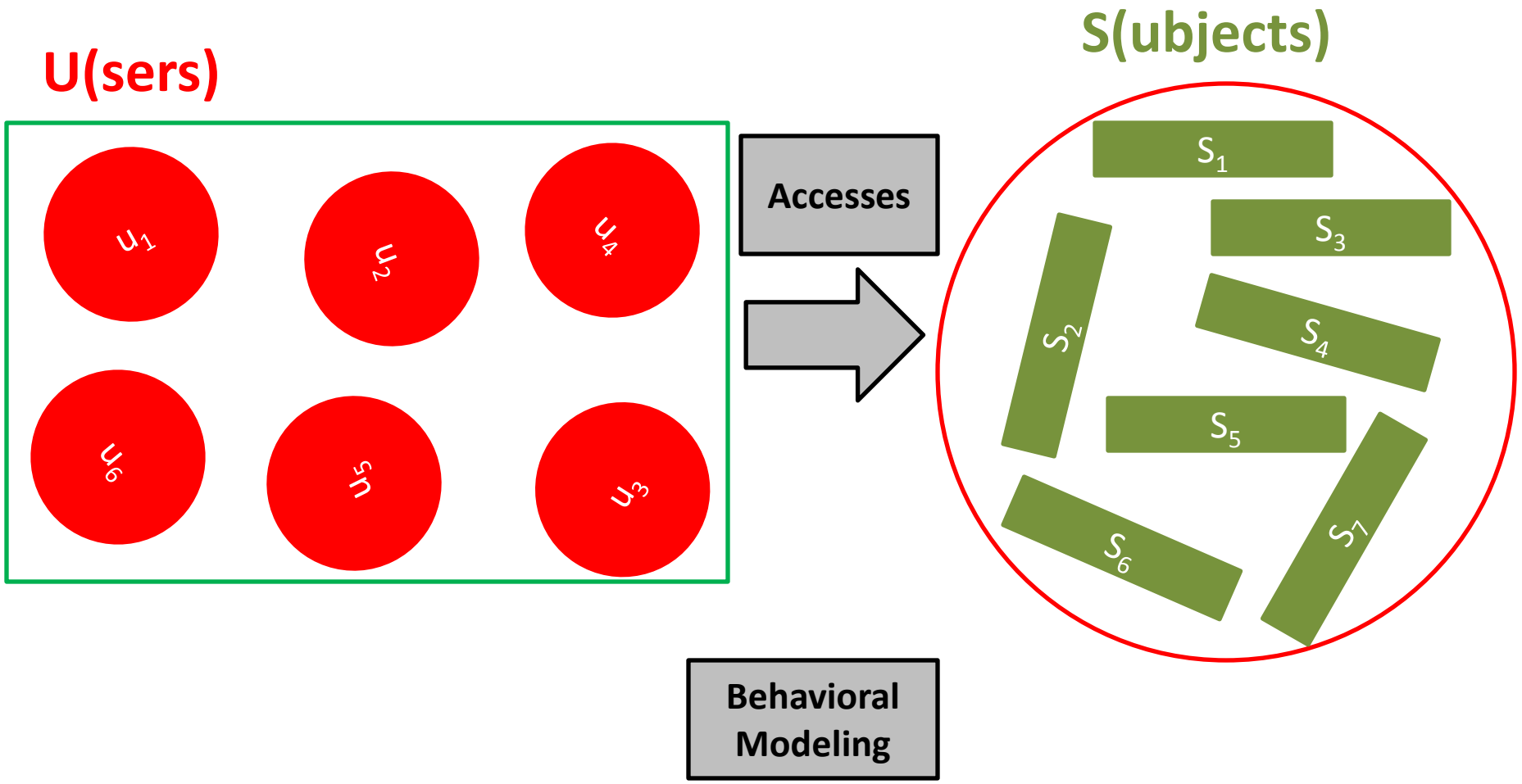




Anomalous users

Intruders have little knowledge of the system and the anticipated behavior

Two general objects of health information system



Where are We Going?

User Level Anomaly Detection

Community Anomaly Detection System (CADS) and Its Extension MetaCADS

(IEEE TDSC)

Access Level Anomaly Detection

Specialized Network Anomaly Detection (SNAD)

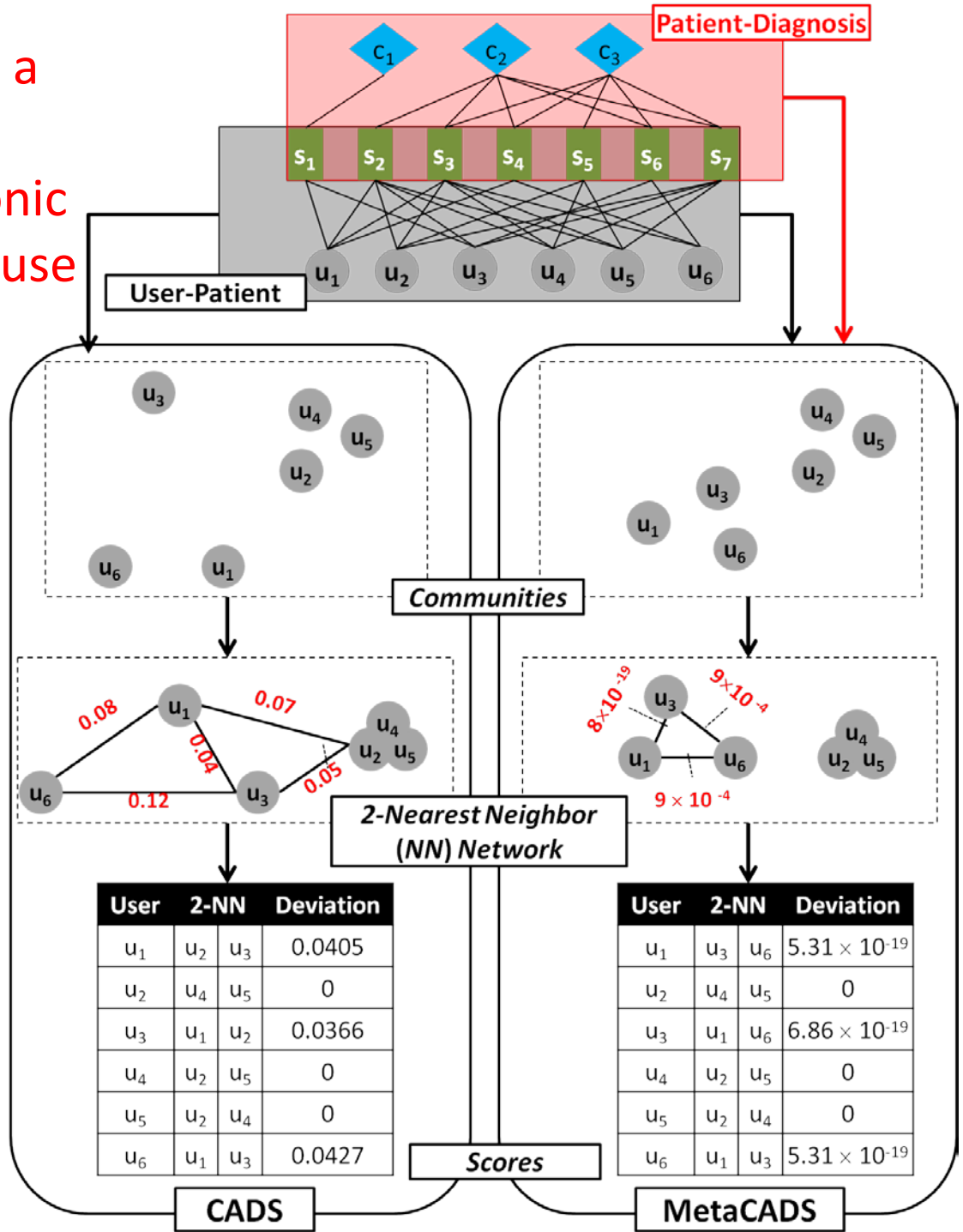
(Security Informatics)

You Chen, Steve Nyemba and Bradley Malin. Detecting Anomalous Insiders in Collaborative Information Systems. IEEE Transaction on Dependable and Secure Computing. Vol.9.No 3, p332-344.

You Chen and Bradley Malin. Detection of Anomalous Insiders in Collaborative Environments via Relational Analysis of Access Logs. Proceedings of ACM Conference on Data and Application Security and Privacy. 2011, p63-74

You Chen, Steve Nyemba, Wen Zhang and Bradley Malin. Specializing Network Analysis to Detect Anomalous Insider Actions. Security Informatics. **1**:5, 2012, p1-24.

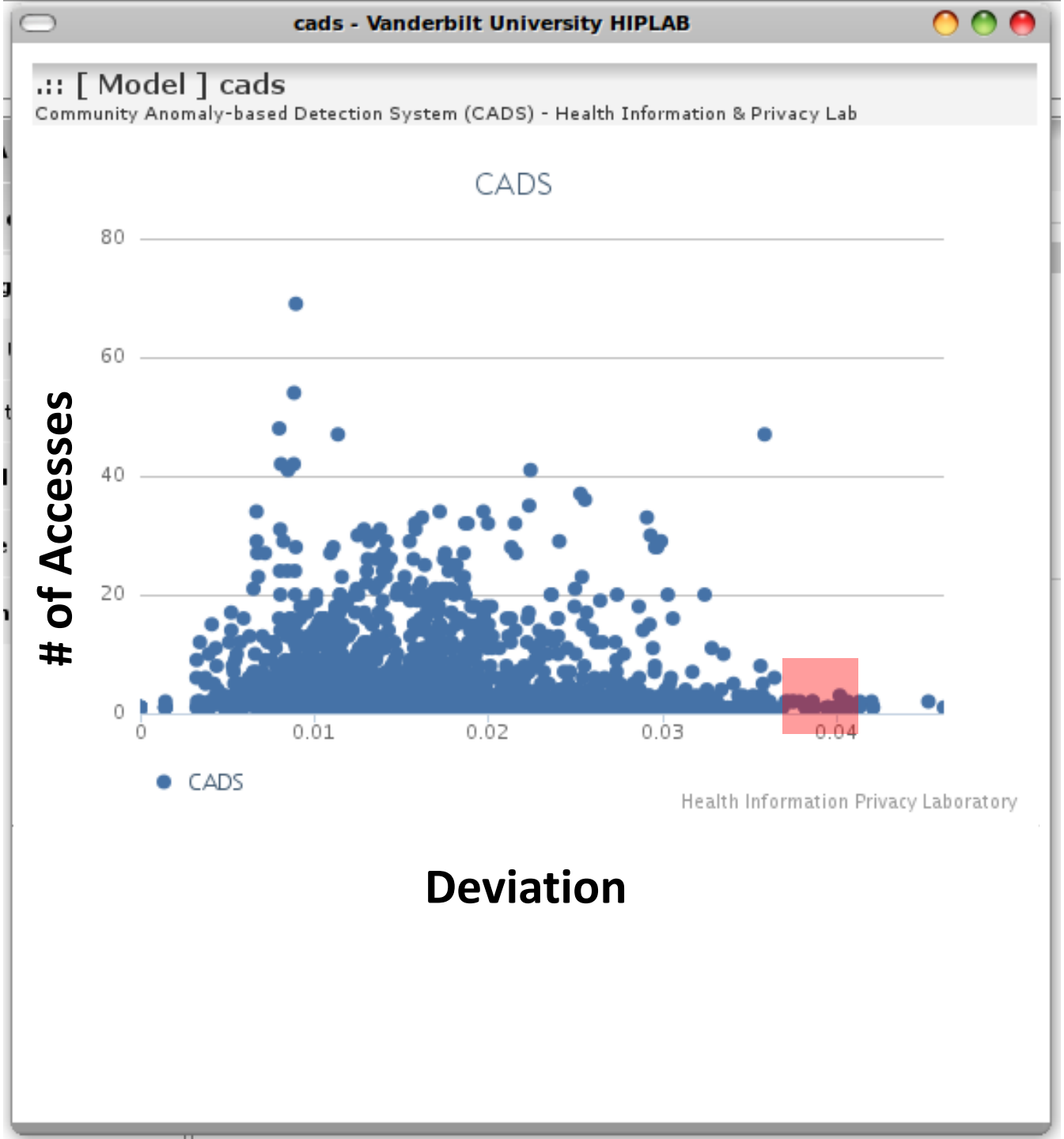
Social Networks are a Novel Approach to Discovery of Electronic Medical Record Misuse



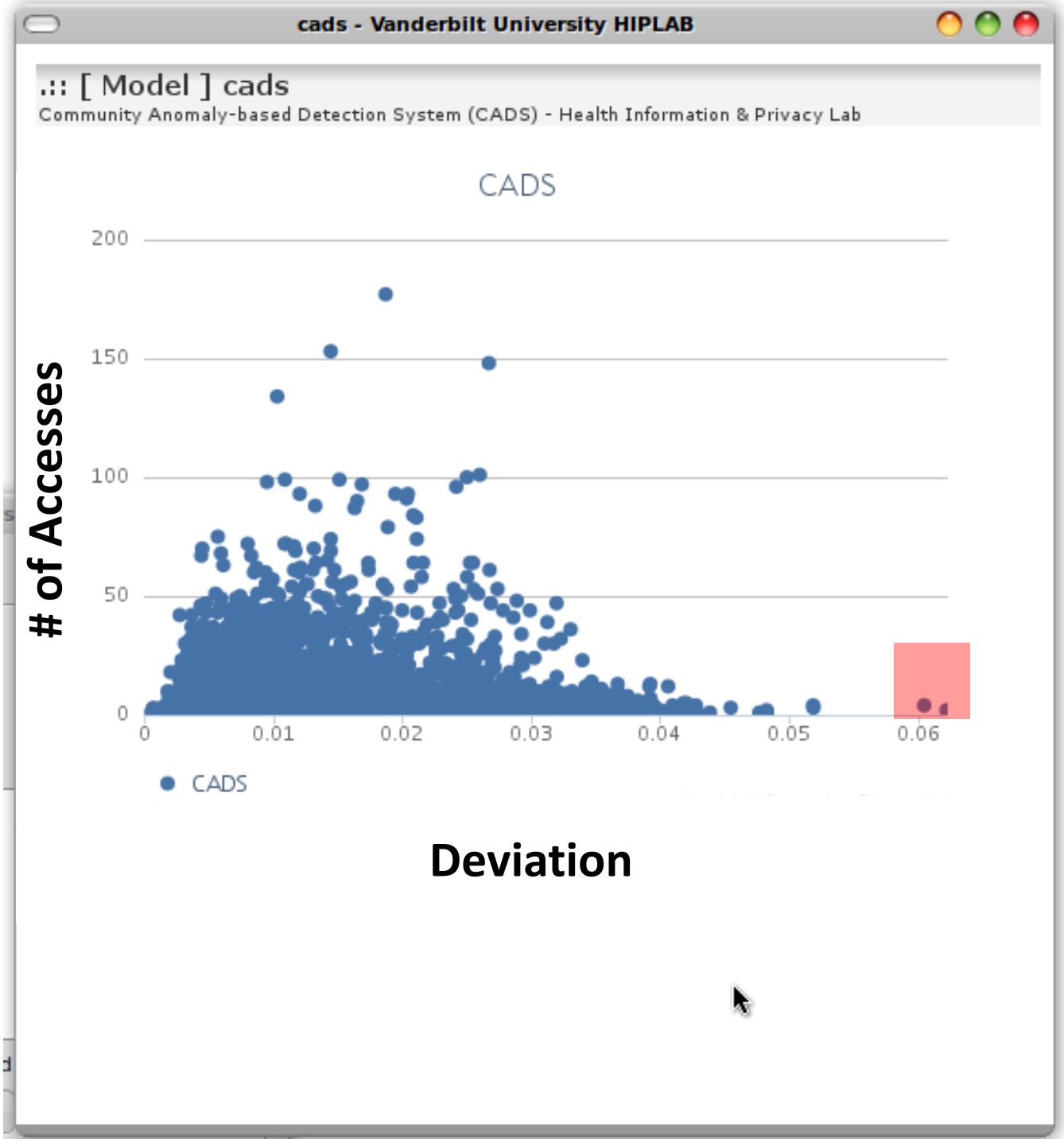
User	2-NN		Deviation
u_1	u_2	u_3	0.0405
u_2	u_4	u_5	0
u_3	u_1	u_2	0.0366
u_4	u_2	u_5	0
u_5	u_2	u_4	0
u_6	u_1	u_3	0.0427

User	2-NN		Deviation
u_1	u_3	u_6	5.31×10^{-19}
u_2	u_4	u_5	0
u_3	u_1	u_6	6.86×10^{-19}
u_4	u_2	u_5	0
u_5	u_2	u_4	0
u_6	u_1	u_3	5.31×10^{-19}

CADS on Vanderbilt Dataset



CADS on Northwestern Dataset



Example Environments

Electronic Health Records (EHR)

- Vanderbilt University Medical Center “StarPanel” Logs
- 3 months in 2010
- Arbitrary Day
  - ≈ 4,208 users
  - ≈ 1,006 patients
  - ≈ 1,482 diagnoses
  - ≈ 22,014 accesses of subjects
  - ≈ 4,609 assignments of diagnoses

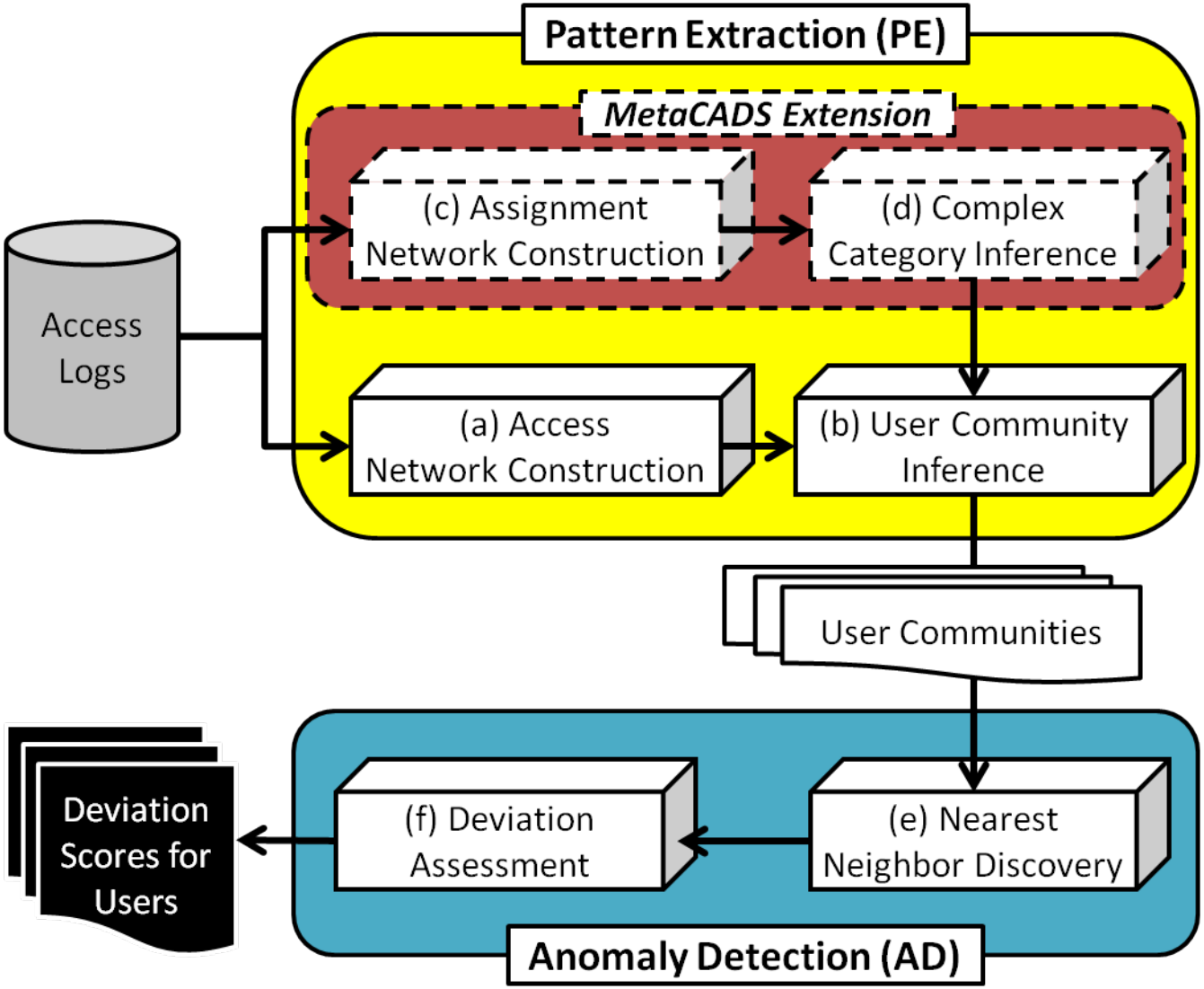
Where are We Going?

- User Level: CADS and MetaCADS
  - **Framework of CADS and MetaCADS**
  - An Example of CADS
  - Experimental Evaluation
  - Limitation

- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)

CADS and MetaCADS

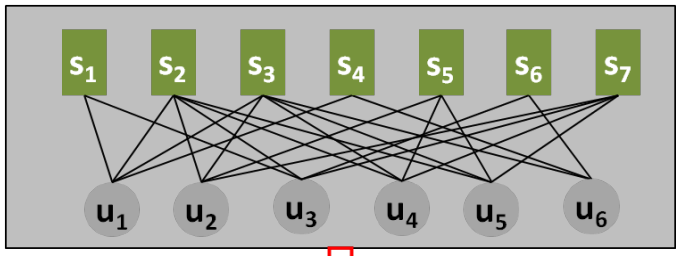


Where are We Going?

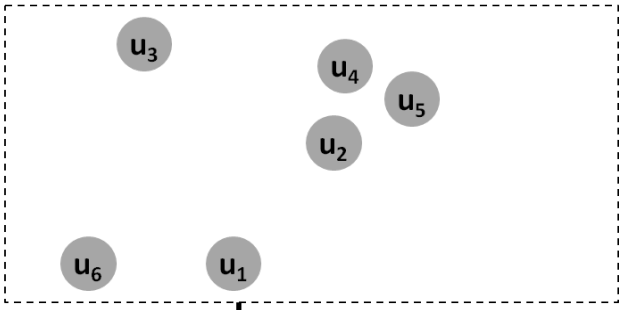
- User Level: CADS and MetaCADS
  - Framework of CADS and MetaCADS
  - **An Example of CADS**
  - Experimental Evaluation
  - Limitation

- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)

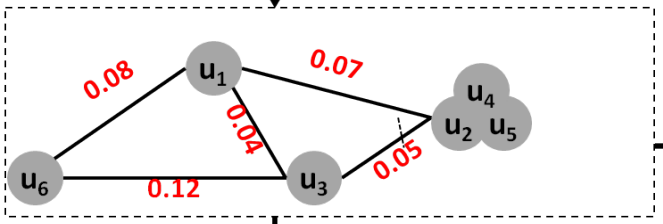


Bipartite Graph->Access Network of Users



Communities via Singular Value Decomposition

Distance via Weighted Euclidean Distance



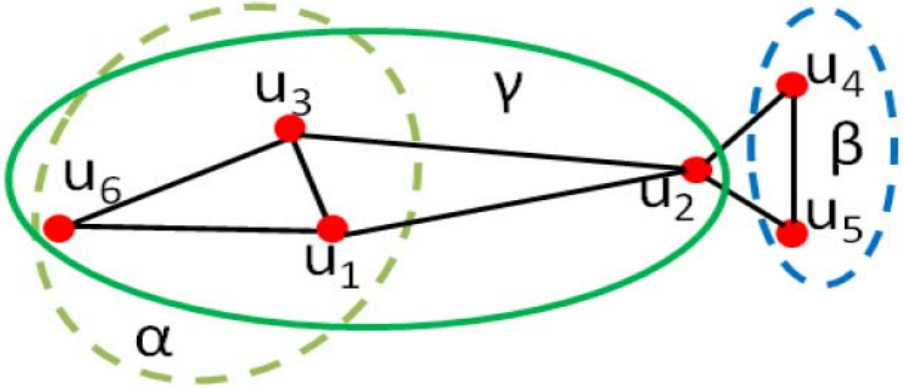
Nearest Neighbor Network

User	2-NN		Deviation
u_1	u_2	u_3	0.0405
u_2	u_4	u_5	0
u_3	u_1	u_2	0.0366
u_4	u_2	u_5	0
u_5	u_2	u_4	0
u_6	u_1	u_3	0.0427

Deviation Scores Calculation

How Do We Set “k”-NN?

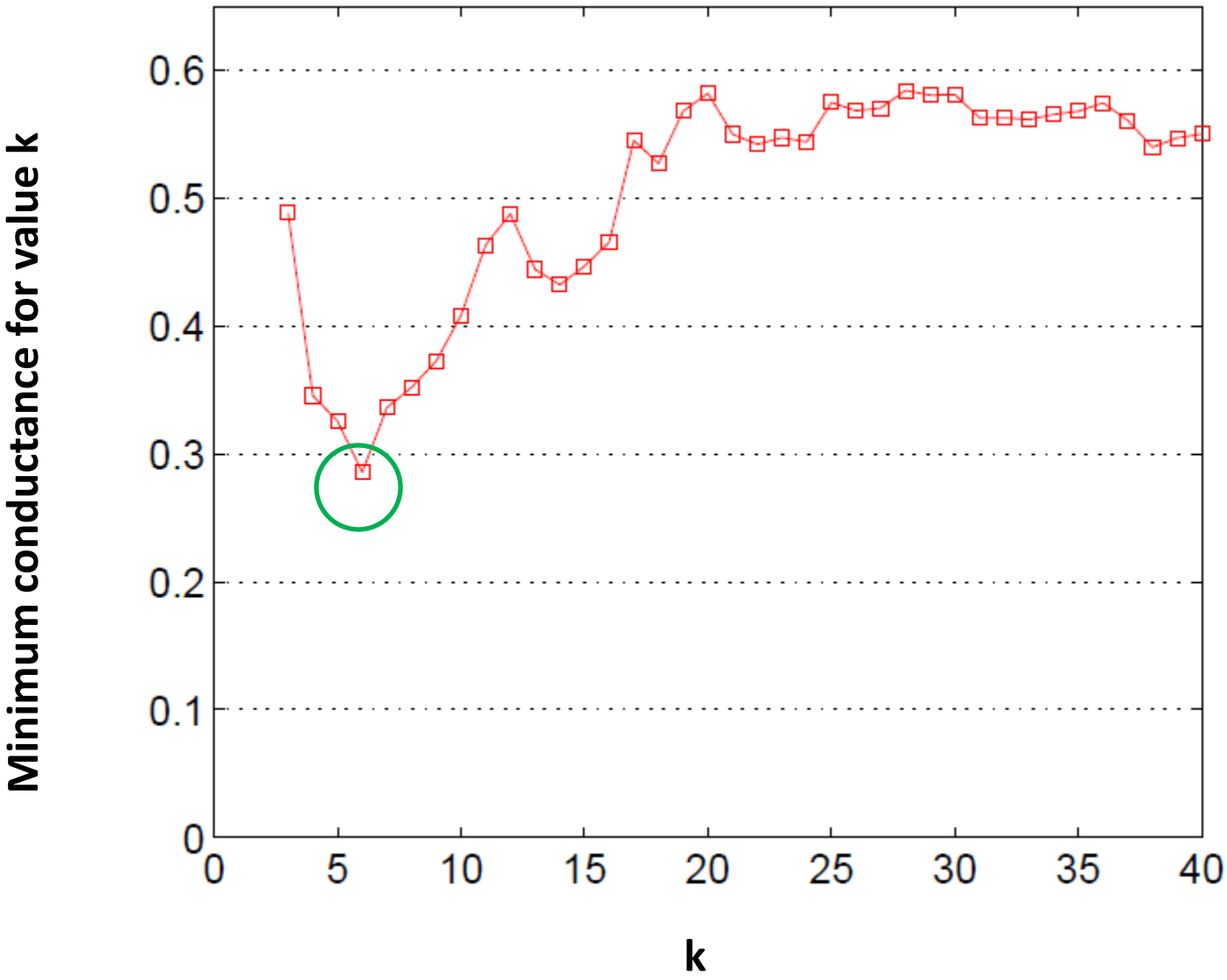
- Conductance- a measure of community quality (Kannan et al)



$$\psi(\beta) = \frac{2}{4}, \psi(\alpha) = \frac{2}{8}, \psi(\gamma) = \frac{2}{min\{4,12\}}$$

$$\psi(\alpha) < \psi(\beta) = \psi(\gamma)$$

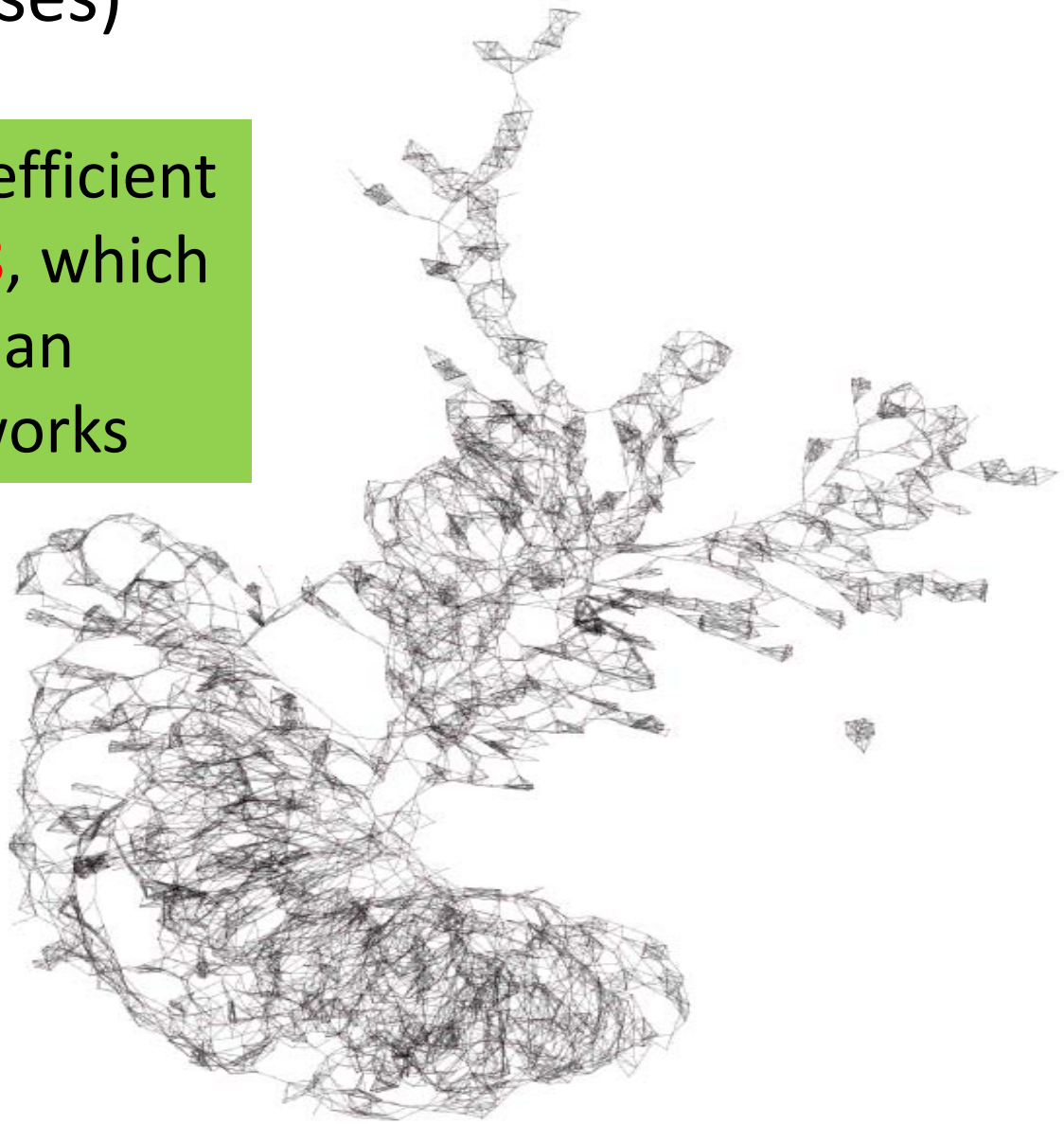
Minimum conductance at k=6



Example 6-Nearest Neighbor Network (1 day of accesses)

The average cluster coefficient for this network is 0.48, which is significantly larger than 0.001 for random networks

Users exhibit collaborative behavior in the health information system



Measuring Deviation from k-NN

- Every user is assigned a radius *r*:
  - the distance to his kth nearest neighbor
- Smaller the radius → higher density in user's network

$$Dev(u_i) = \sqrt{\frac{\sum_{u_j \in knni} (r_j - \bar{r})^2}{k}}$$

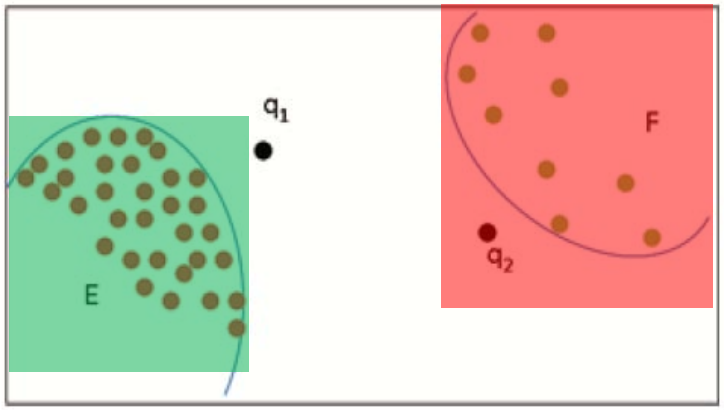
$$\bar{r} = \frac{\sum_{u_j \in knni} r_j}{k}$$

Radius for these points are nearly 2, and for q_1 is 3

$$\bar{r} = \frac{2+2+2+2+3}{5} = 2.2$$

$$Dev(q_1) = \sqrt{\frac{(2-2.2)^2 \times 4 + (3-2.2)^2}{5}} = 0.42$$

5 nearest



Radius for these points are larger than 10, and every r values significantly different

If we set threshold of radius as 10, than q_1 is a normal user, who in fact is anomalous

Where are We Going?

- User Level: CADS and MetaCADS
  - Framework of CADS and MetaCADS
  - An Example of CADS
  - **Experimental Evaluation**
  - Limitation

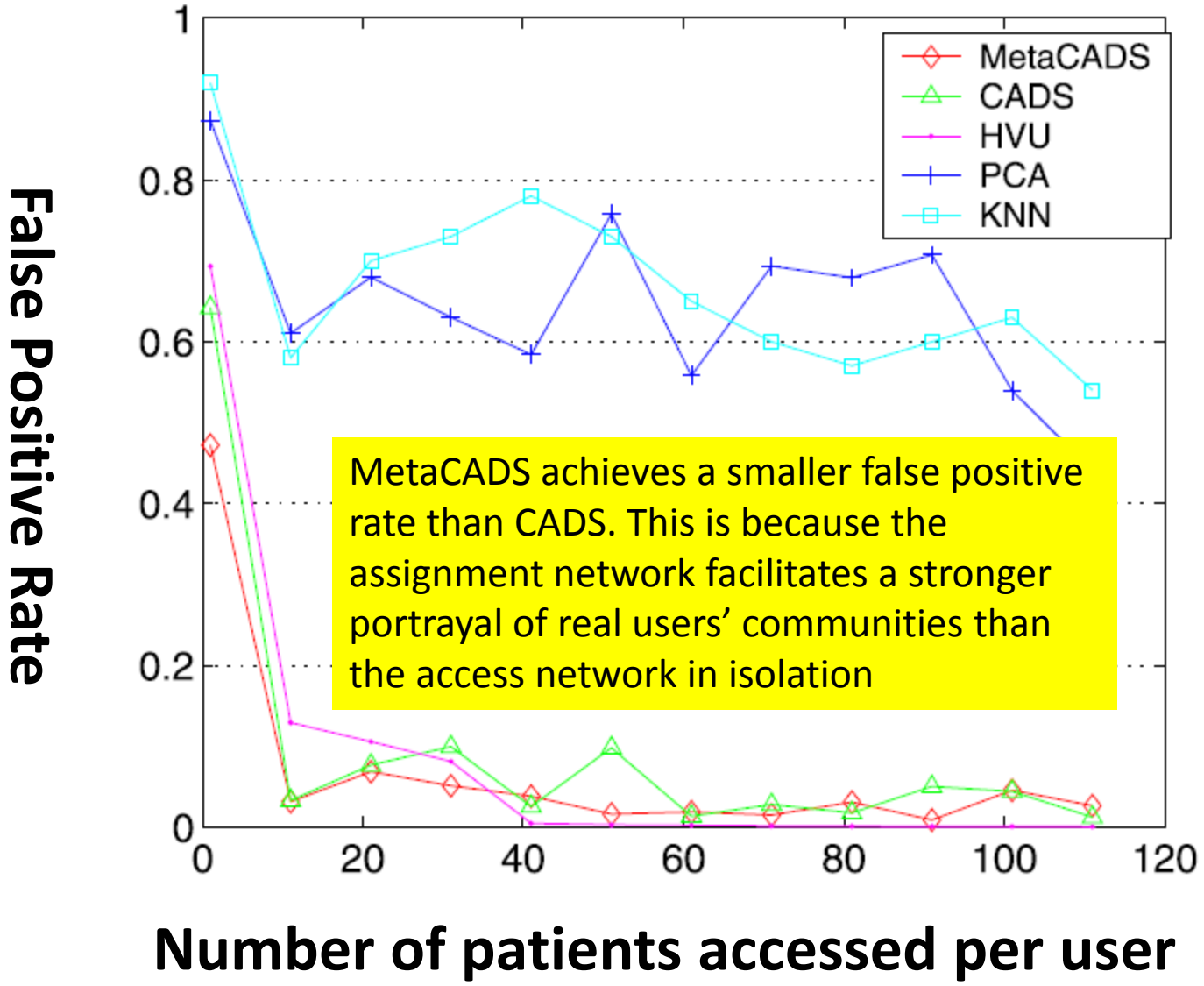
- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)

Experimental Design

- Datasets are not annotated for illicit behavior
- We simulated users in several settings to test:
  - Sensitivity to number of records accessed of a specific users
    - Range from 1 to 120
  - Sensitivity to number of anomalous users
    - simulated users correspond to 0.5% to 5% of total users
    - Number of records accessed fixed to 5
  - Sensitivity to diversity
    - Random number of users(0.5%~5%) and records accessed (1~150)

Exp1: False Positive Rate Decreases, when the Number of Subjects Accessed Increases



Exp2: Detection Rate With Various Mix Rates of Real and Simulated Users

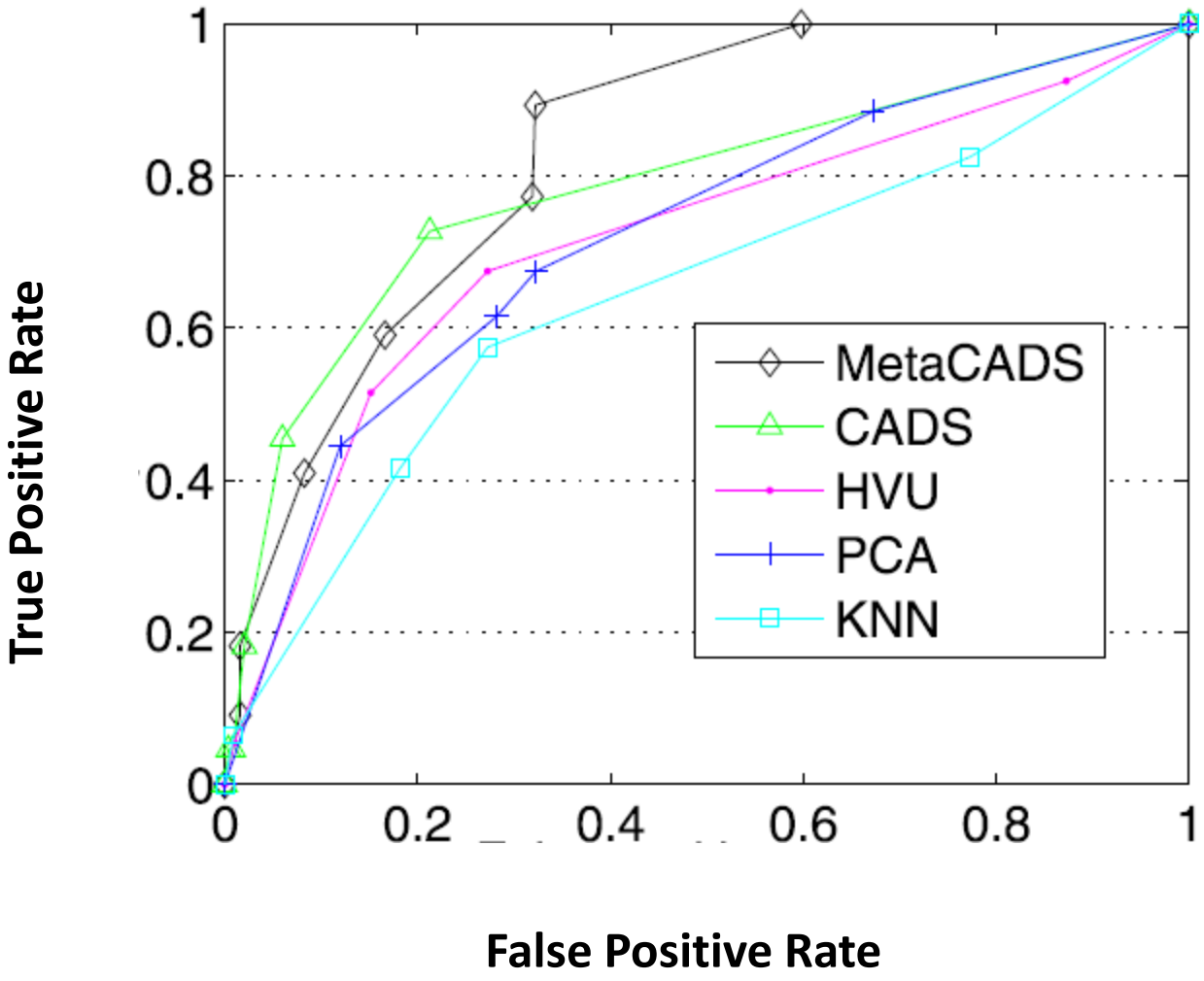
	MIX RATE		
MODEL	0.5%	2%	5%
MetaCADS	**0.92±0.02**	0.90±0.01	0.87±0.03
CADS	0.91±0.01	**0.94±0.02**	**0.94±0.01**
KNN	0.75±0.02	0.73±0.03	0.72±0.04
PCA	0.72±0.03	0.74±0.02	0.75±0.03
HVU	0.68±0.03	0.68±0.03	0.68±0.03

when the number of simulated users is low (i.e., 0.5 percent), MetaCADS yields a slightly higher AUC than CADS (0.92 versus 0.91)

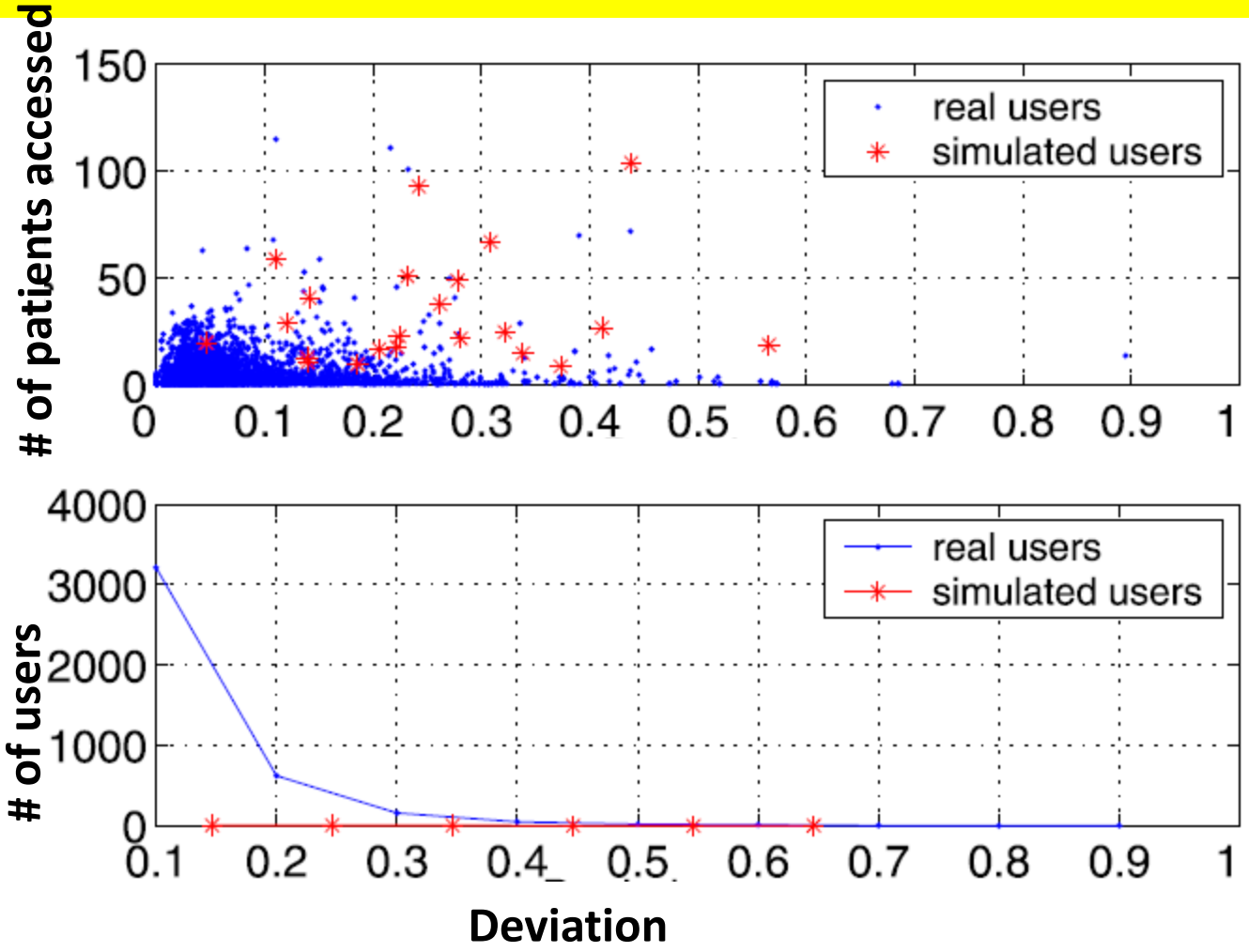
As the number of simulated users increases, CADS clearly dominates MetaCADS. The performance rate of CADS increases from 0.91 to 0.94, while MetaCADS decreases from 0.92 to 0.87.

Because when the number of simulated users increases, they have more frequent categories in common. In turn, these categories enable simulated users to form more communities than those based on patients alone, thus lowering their deviation scores.

Exp3: MetaCADS dominates when the mix rate is low (mix rate = 0.5%)



MetaCADS deviation scores of real and simulated users as a function of the number of subjects accessed. This system was generated with a mix rate of 0.5 percent and a random number of subjects accessed per simulated user



Where are We Going?

- User Level: CADS and MetaCADS
  - Framework of CADS and MetaCADS
  - An Example of CADS
  - Experimental Evaluation
  - **Limitation**

- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)

Some Limitations

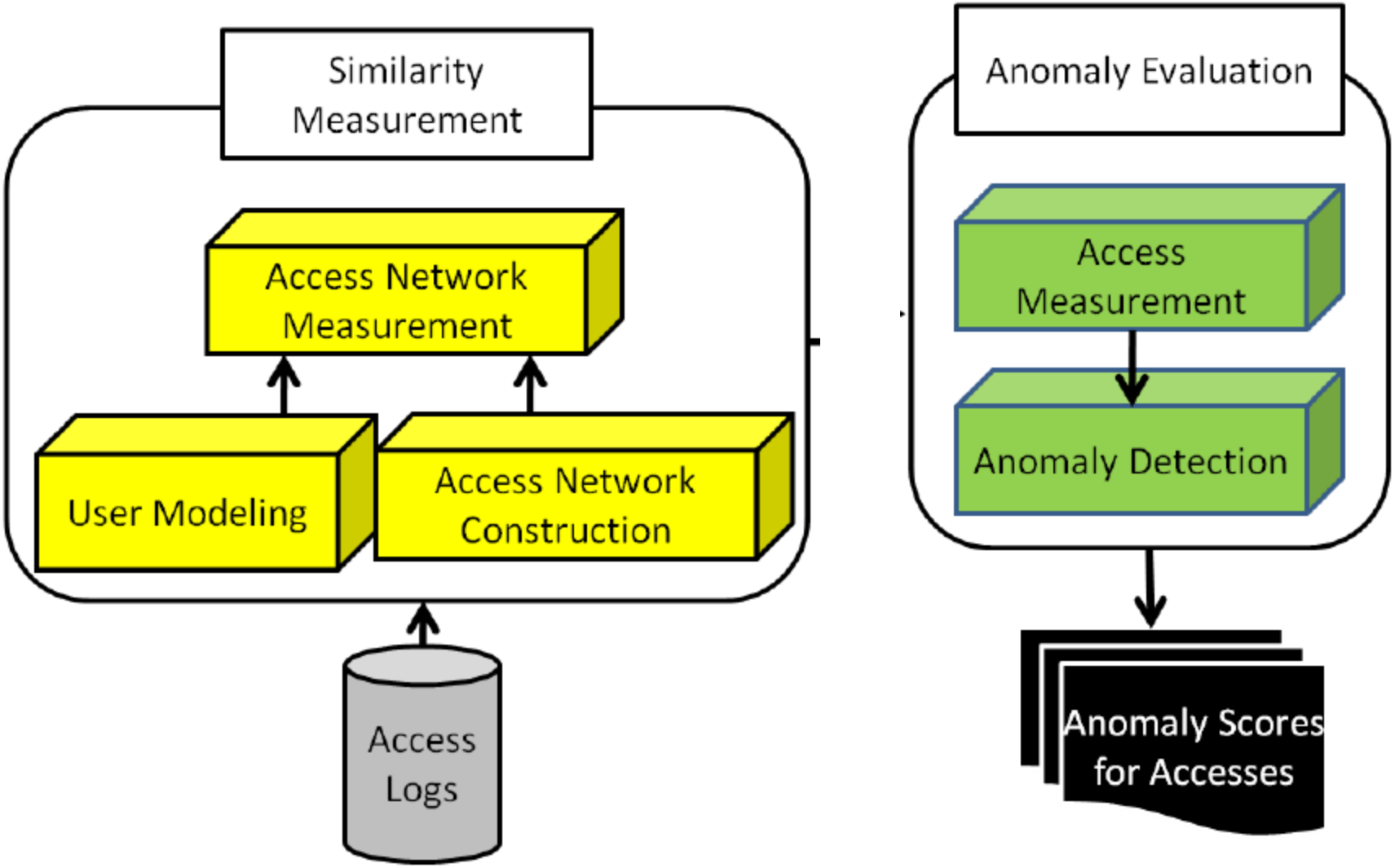
- Simulated users are indicative of misuse of the system…
  …but actual illicit behavior may be more directed.
- "False positives" are not necessarily false!
  (Adjudication by EHR privacy experts under way)
- Need to specialize tool to account for semantics of users and subjects
  - User: {Role, Department, Residence}
  - Patient: {Diagnosis, Procedure, Demographics, Residence}
- Anomalous users… not anomalous accesses
  - Need to account for insiders that deviate by only a couple of actions

Where are We Going?

- User Level: CADS and MetaCADS
- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)
  - **Framework of SNAD**
  - An Example of CADS
  - Experimental Evaluation
  - Limitation

SNAD Framework



Where are We Going?

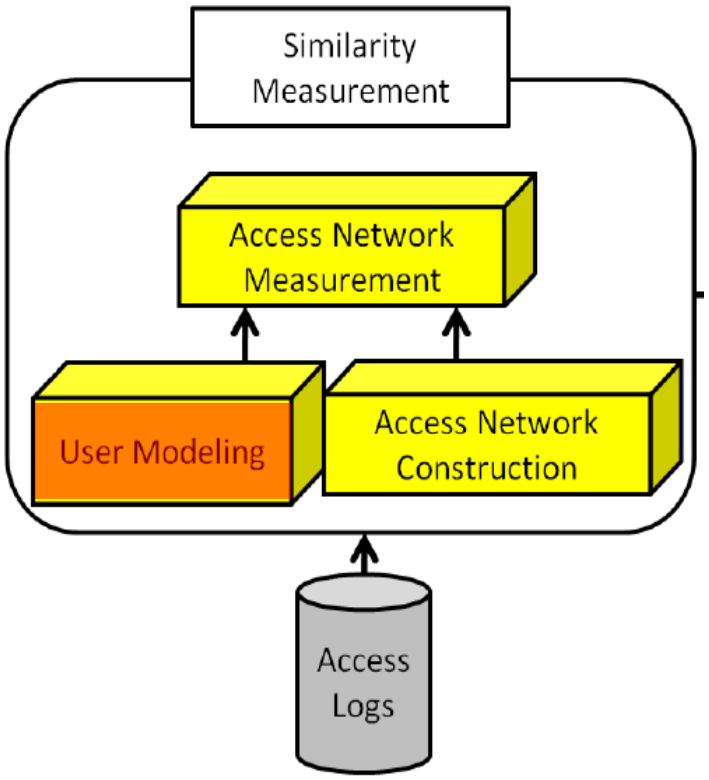
- User Level: CADS and MetaCADS

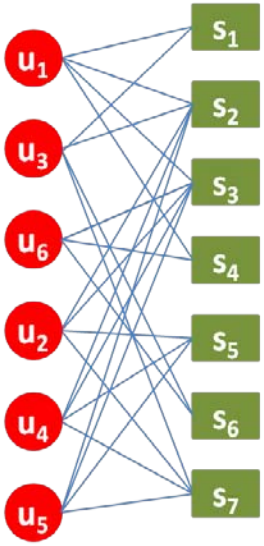
- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)

  - Framework of SNAD
  - **An Example of SNAD**
  - Experimental Evaluation
  - Limitation

User Modeling



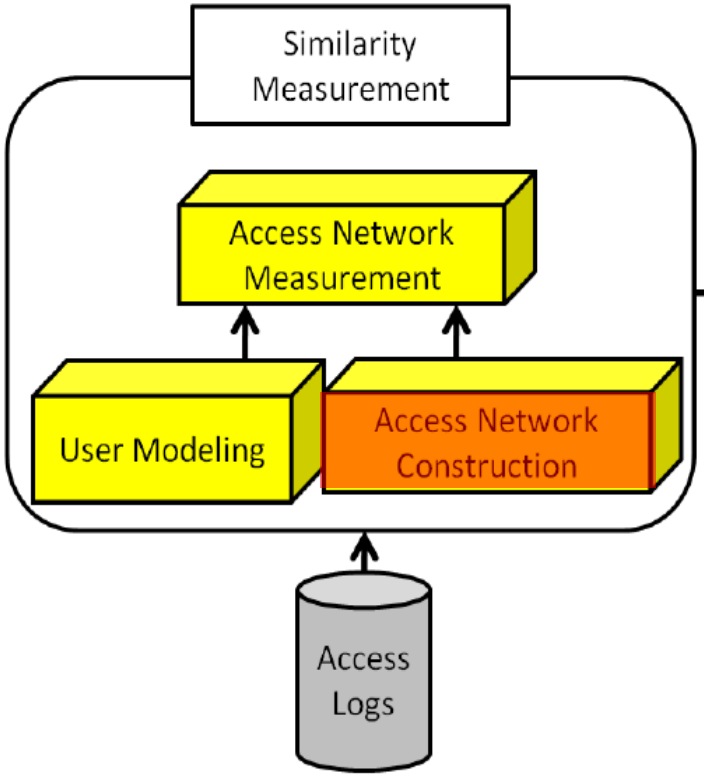


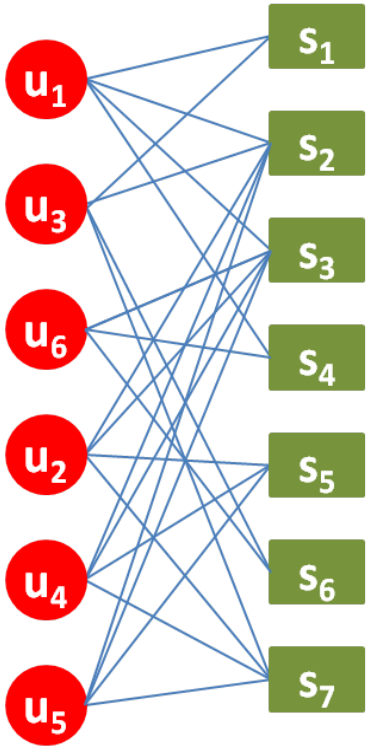
	u_1	u_2	u_3	u_4	u_5	u_6
s_1	1	0	1	0	0	0
s_2	1	1	1	1	1	0
s_3	1	1	0	1	1	1
s_4	1	0	0	0	0	1
s_5	0	1	0	1	1	0
s_6	0	0	1	0	0	1
s_7	0	1	1	1	1	0

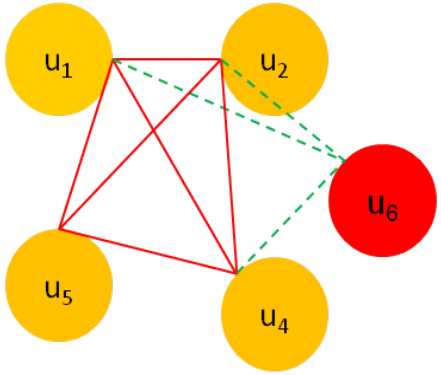
	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

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

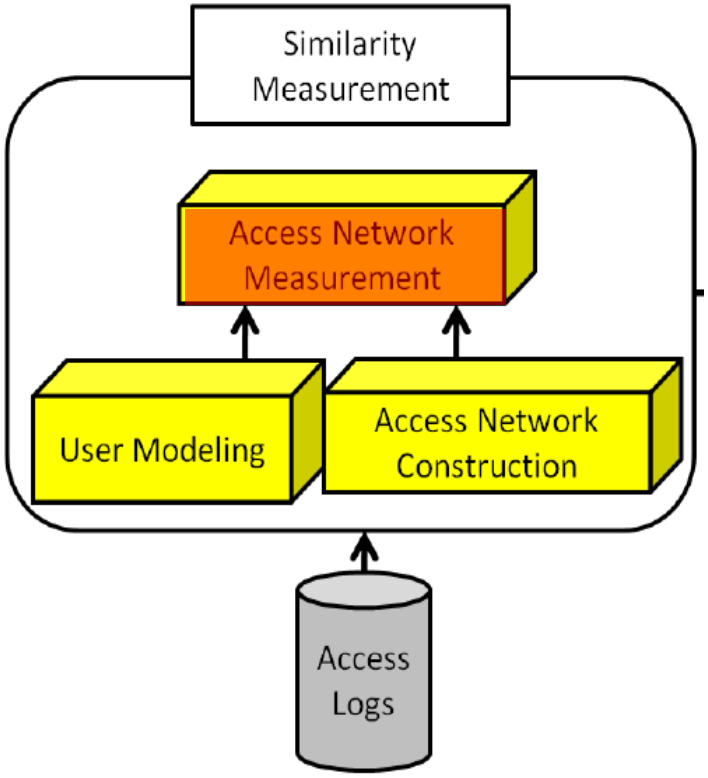
Access Network Construction







Access Network Measurement

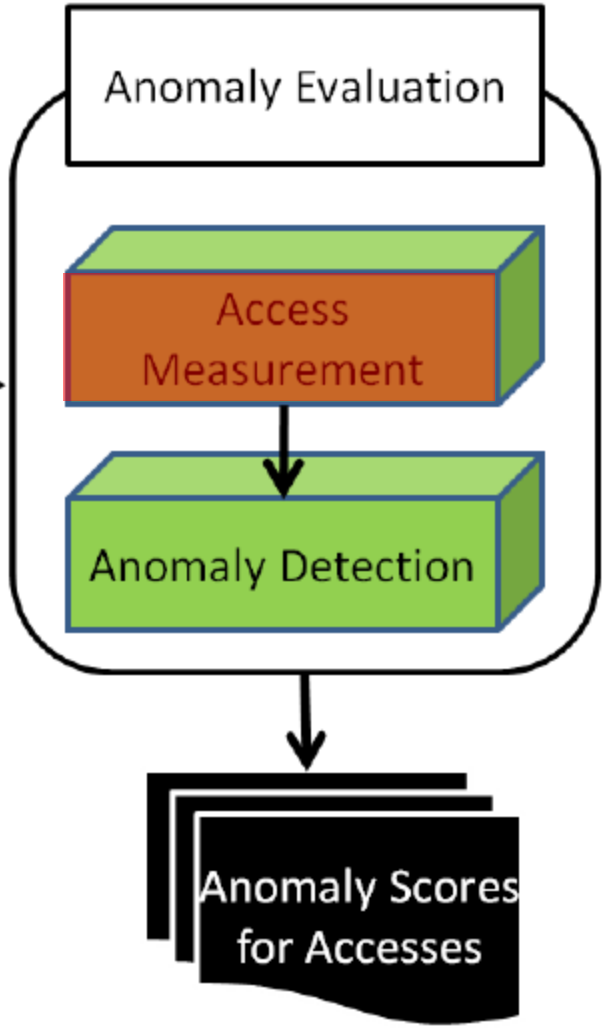


	u_1	u_2	u_3	u_4	u_5	u_6
s_1	0.15	0	0.15	0	0	0
s_2	0.15	0.15	0.15	0.15	0.15	0
s_3	0.15	0.15	0.00	0.15	0.15	0.24
s_4	0.15	0	0	0	0	0.24
s_5	0	0.15	0	0.15	0.15	0
s_6	0	0	0.15	0	0	0.24
s_7	0	0.15	0.15	0.15	0.15	0

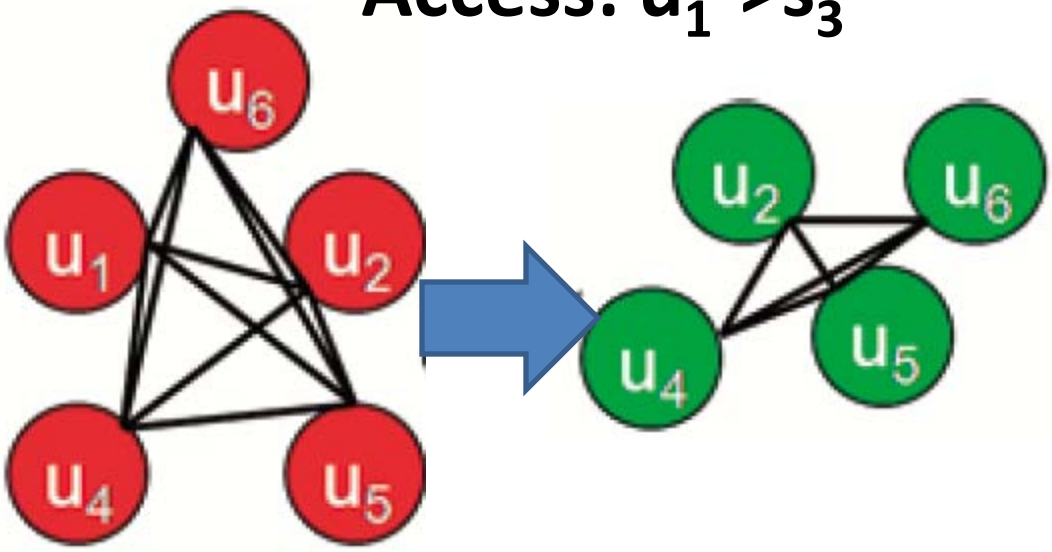
	u_1	u_2	u_4	u_5	u_6
u_1	1.00				
u_2	0.50	1.00			
u_4	0.50	1.00	1.00		
u_5	0.50	1.00	1.00	1.00	
u_6	0.58	0.29	0.29	0.29	1.00

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

Measuring Accesses for Changes in Network Similarity



Access: u_1->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

Access	Score	Size
u1-s3	0.05	4

Where are We Going?

- CADS and MetaCADS

- Access Level: Specialized Network Anomaly Detection (SNAD)

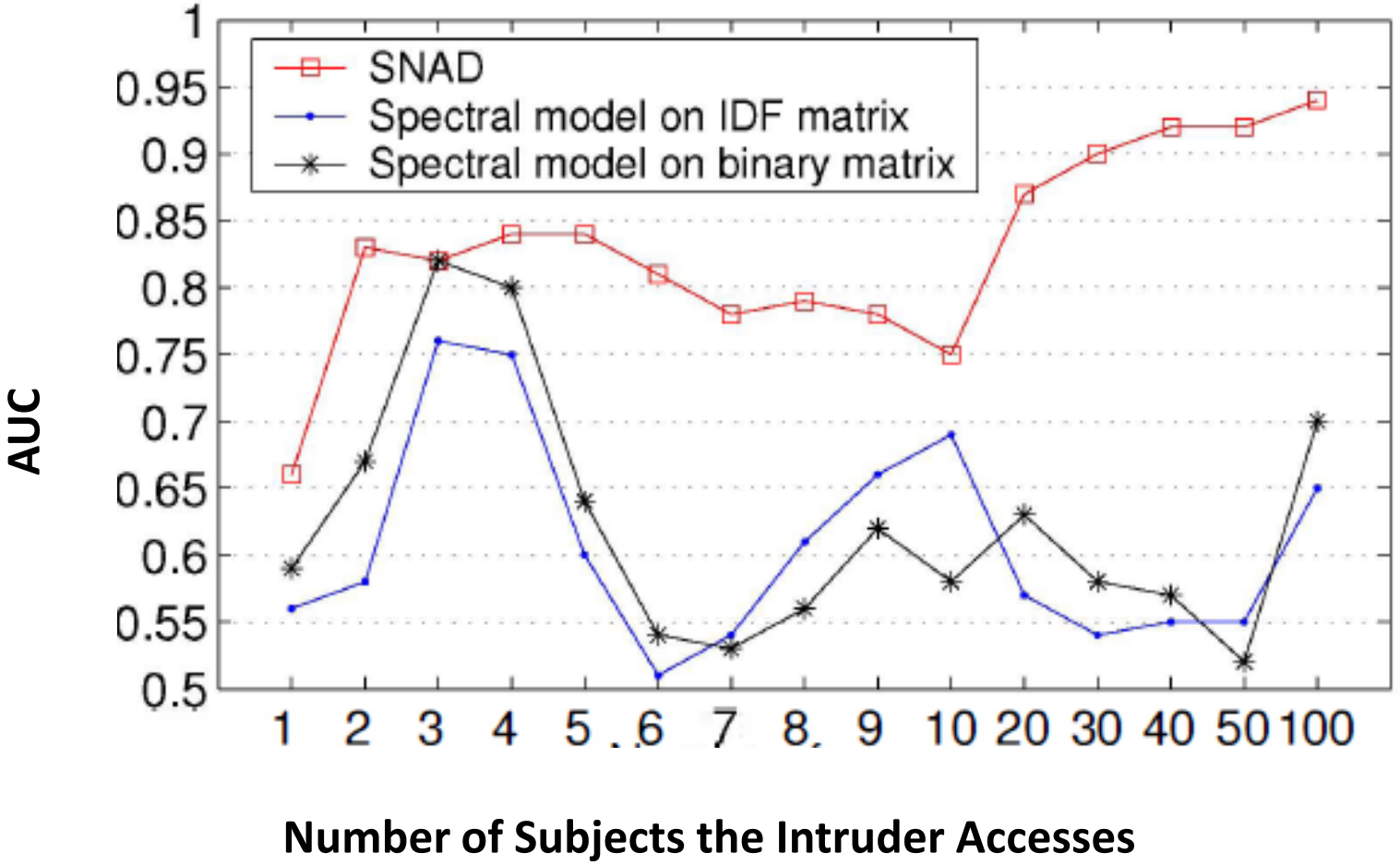
  (SI)

  - Framework of SNAD
  - An Example of SNAD
  - **Experimental Evaluation**
  - Limitation

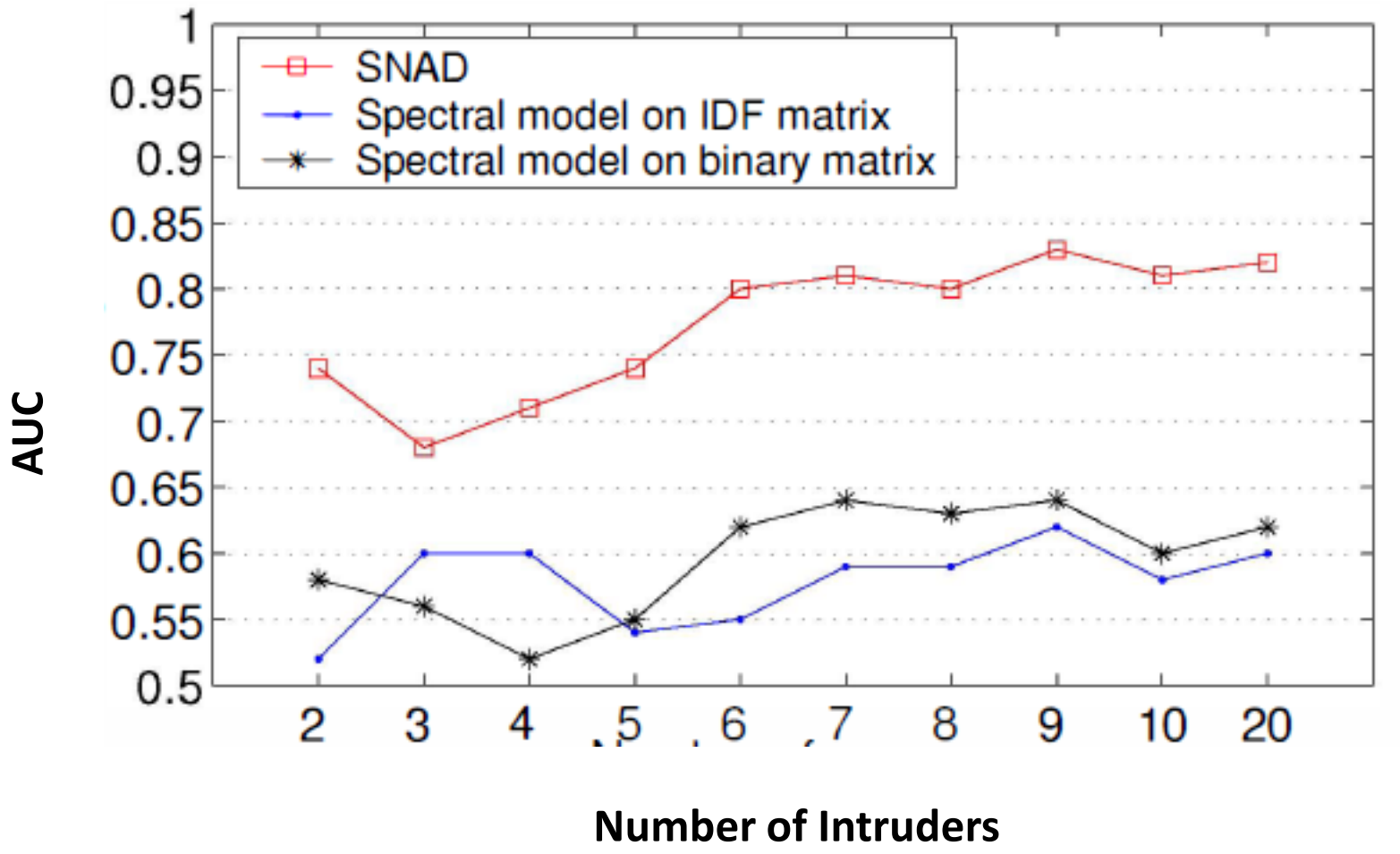
Experimental Design

- Datasets are not annotated for illicit behavior
- We simulated users in several settings to test:
  - Sensitivity to number of subjects accessed
    - Range from 1 to 1,00
  - Sensitivity to number of anomalous users
    - Range from 2 to 20
    - Number of subjects accessed fixed to 5
  - Sensitivity to diversity
    - Random number of users and subjects accessed

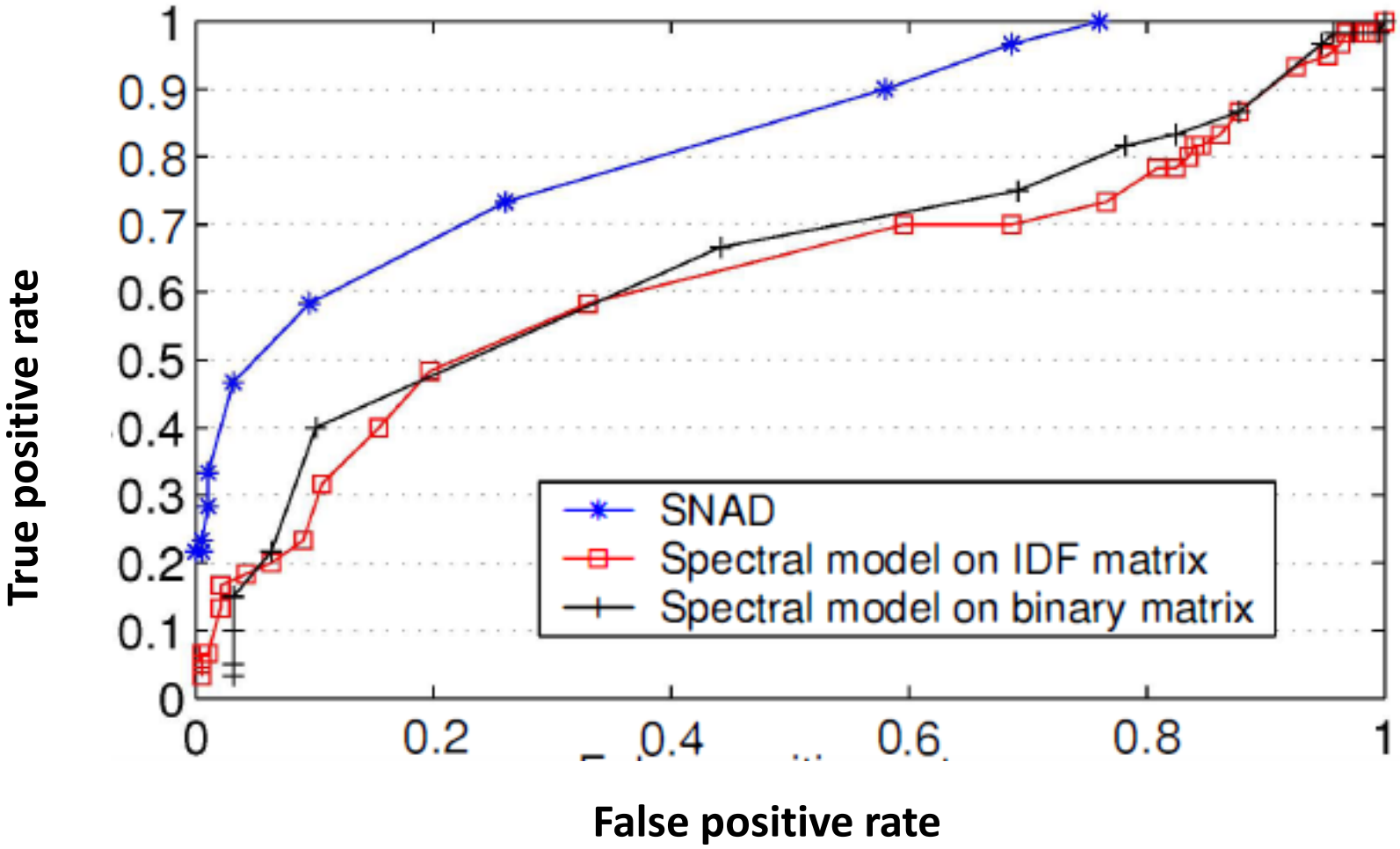
SNAD: Detection Rate Increase with Number of Subjects Accessed



SNAD: Detection Rate Increases with Number of Intruders



SNAD Outperforms Competitors When the Number of Intruders & Accessed Subjects is Random



Where are We Going?

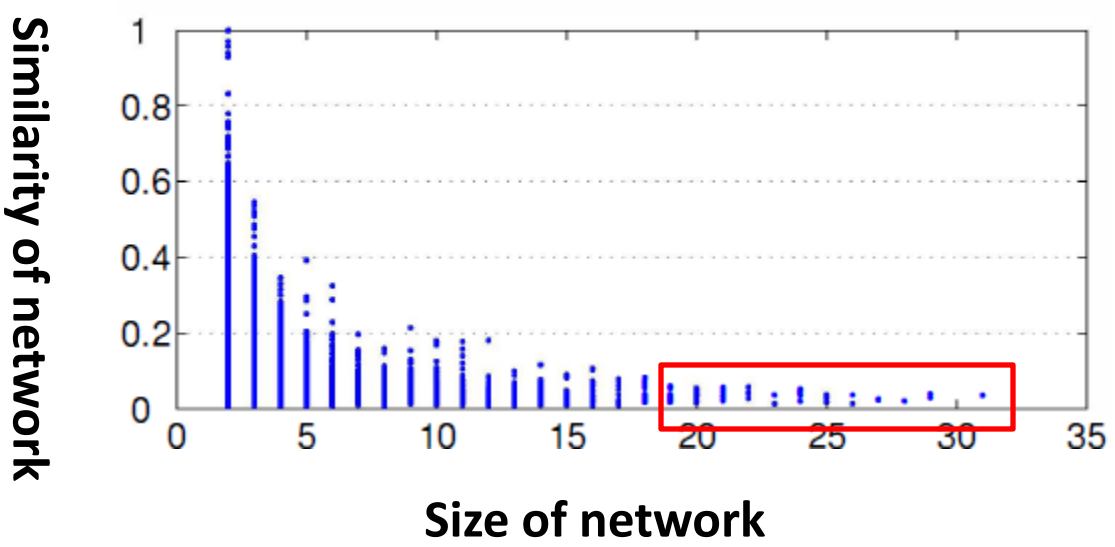
- CADS and MetaCADS
- Access Level: Specialized Network Anomaly Detection (SNAD)

  (SI)
  - Framework of SNAD
  - An Example of SNAD
  - Experimental Evaluation
  - **Limitation**

Limitations

- SNAD has high performance in Vanderbilt's EHR system because
  - organization is collaborative
  - access networks have high network similarity

- SNAD may not be appropriate for large access network with low network similarity
  - Absence of a user has little influence on the similarity.



Conclusions

- It is an effective way by using social network analysis to detect anomalous usages of electronic health records, such as CADS and SNAD
- Adding semantic information of users and subjects will make social network analysis be more understandable

Protecting Patients through Dynamic Network Analysis of Hospital Department Relationships

You Chen, Steve Nyemba and Bradley Malin. Auditing Medical Records Accesses via Healthcare Interaction Networks. AMIA 2012 Annual Symposium. p93-102

Patient information needs to be protected from insiders

- Traditional security practices (e.g., role-based access control) are insufficient to ensure EMR security
  - Common for >100 employees to access a patient's medical record during their visit
  - Often difficult to determine who the members of a care team are and who will need access to what information at which time

EHRs have adopted collaborative capabilities to facilitate interaction between teammates and coordinate care

- We hypothesize that HCO departments will exhibit predictable interaction behavior
- Our goals:
  1. Investigate if such behavior exists
  2. If so, determine if it is stable
     - If stable interactions become unstable → associated patients will be anomalous

Our goal is to retrieve the dependent relations of departments and determine whether the dependencies among departments touching that patient are expected?

Expected behavior:

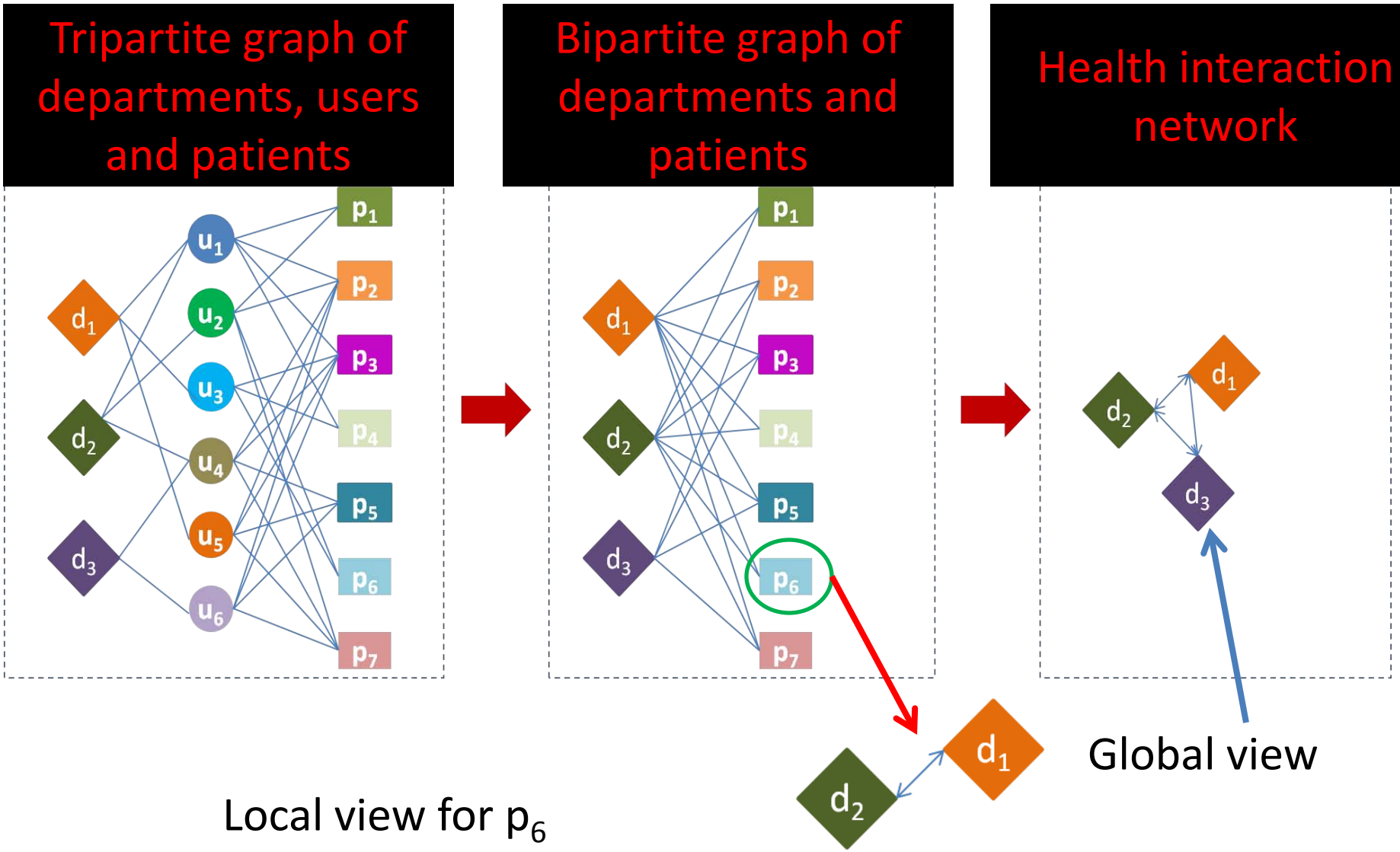
Pediatric housestaff; pediatric cardiology; Vanderbilt Children's hospital

Utilized behavior:

Mental health center; burn center; breast cancer center; pediatric housestaff; pediatric cardiology; Vanderbilt Children's hospital

The dependent relations between green departments and red departments are very low

Healthcare Interaction Networks



Where are We Going?

A Global Network of Departments

Two metrics: certainty and reciprocity

Stable status in terms of the two metrics

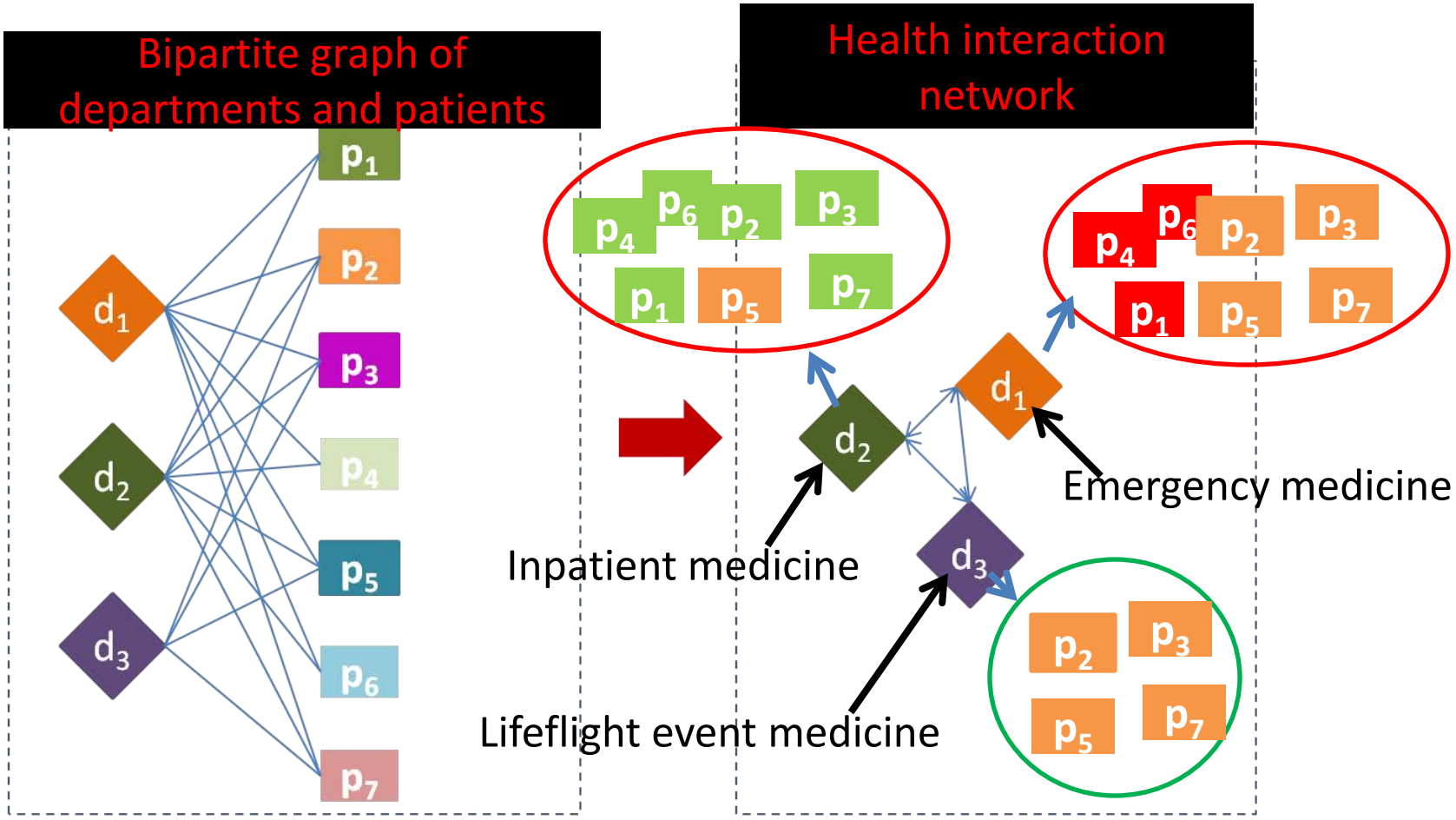
Local Network-for a specific patient

Two metrics: local network score and reciprocity

Application of the Networks

Detecting patients with anomalous medical records accesses

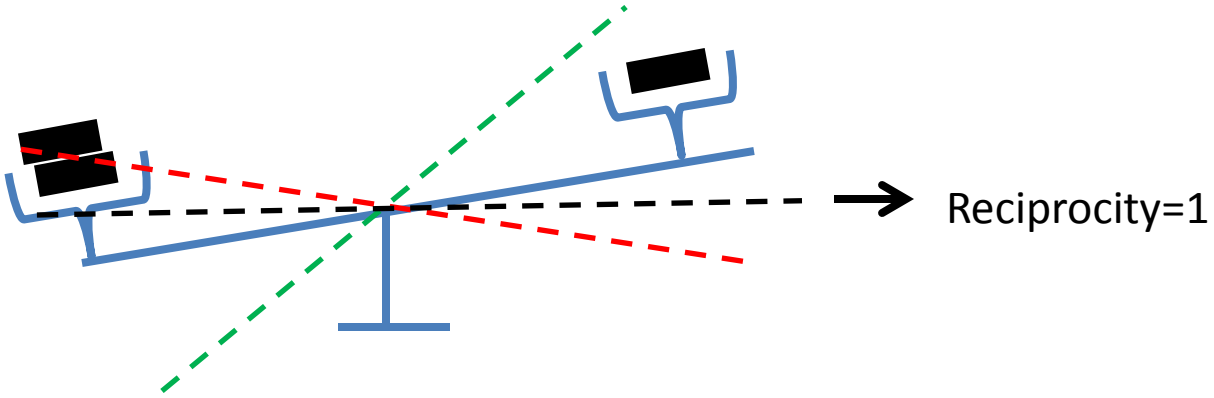
Certainty to Model Relationship of Global Network



Cert(Lifeflight event medicine (d_3)->Emergency medicine (d_1)) = 4/4

Cert(Inpatient medicine(d_2)->Inpatient medicine(d_2)) = 6/7

Using reciprocity to characterize the mutual interaction between all pairs of departments in the global network



Pediatric Emergency Dept -> Peds Respiratory Care =0.57

Peds Respiratory Care -> Pediatric Emergency Dept = 0.037

Where are We Going?

A Global Network of Departments

Two metrics: certainty and reciprocity

Stable status in terms of the two metrics

Local Network-for a specific patient

Two metrics: local network score and reciprocity

Application of the Networks

Detecting Patients with Anomalous Medical Records Accesses

Dataset used for this study

- Vanderbilt University Medical Center “StarPanel”
- 3 months in 2010
- Arbitrary Week
  - ≈ 9,200 users
  - ≈ 99,000 patient records
  - ≈ 400,000 accesses
  - ≈ 450 departments

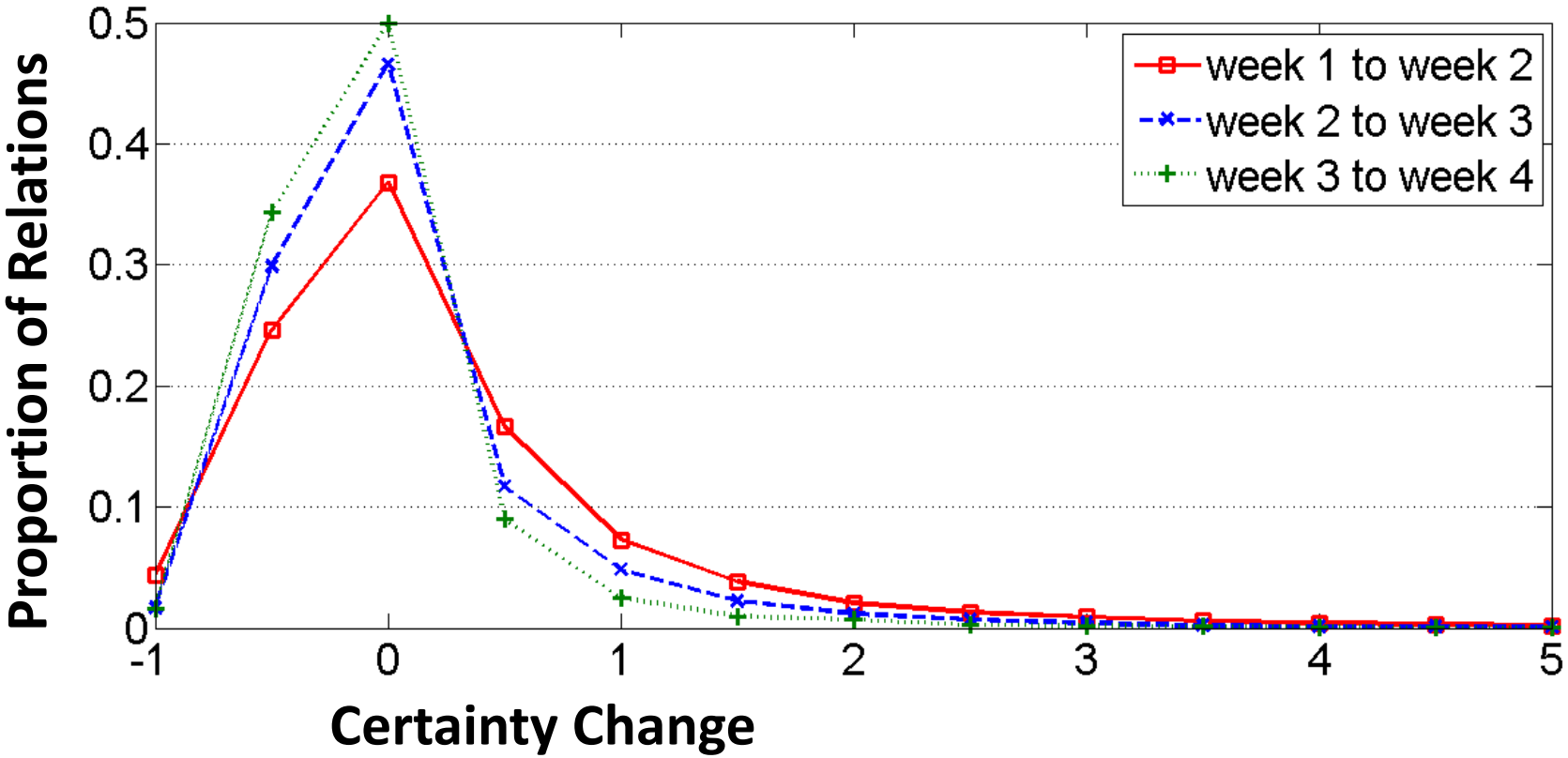
Although the relations of the network are very unbalanced, the unbalance is stable over time

Time	Week 1	Week 2	Week 3	Week 4
Reciprocity	0.267	0.2814	0.2858	0.2871

$$(0.2814 - 0.267) / 0.267 = 0.05$$

Week 1 to week 2

The changes become smaller over time (centralization: green > blue > red)



Degree of relations between departments changes little over time

>82.5% of the change resides in [-0.25, 0.25]

Strong relations between VUMC departments over a four week period

Department (d_i)	Department (d_j)	Min Certainty	Max Certainty
Intradepartmental Relations			
4East OB/GYN	4East OB/GYN	0.74319	0.7669
Adult Emergency Medicine	Adult Emergency Medicine	0.74024	0.78453
Cancer Infusion Center	Cancer Infusion Center	0.73171	0.844
8N Inpatient Medicine	8N Inpatient Medicine	0.7197	0.80909
Newborn Nursery	Newborn Nursery	0.70406	0.72727
Interdepartmental Relations			
DOT Radiology	Orthopaedics	0.99621	1
Nursing Education and Development	Medical Information Services	0.95833	1
Main OR - Trauma/Renal	Medical Information Services	0.94444	1
Life Flight Event Medicine	Emergency Medicine	0.90805	1
Emergency Medicine Admin	Adult Emergency Medicine	0.91489	0.94186

Where are We Going?

A Global Network of Departments

Two metrics: certainty and reciprocity

Stable status in terms of the two metrics

Local Network-for a specific patient

Two metrics: local network score and reciprocity

Application of the Networks

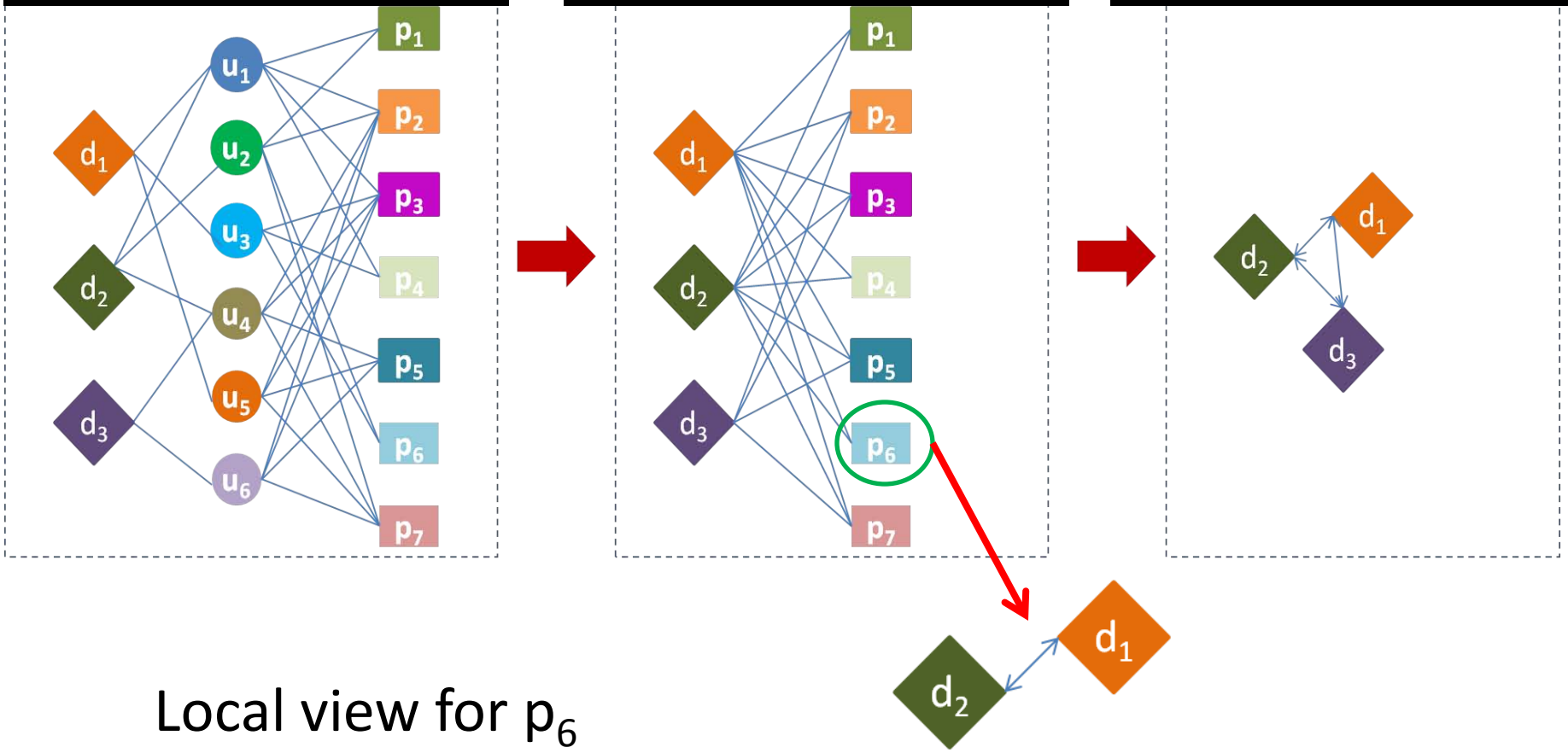
Detecting patients with anomalous medical records accesses

Healthcare Interaction Networks

Tripartite graph of departments, users and patients

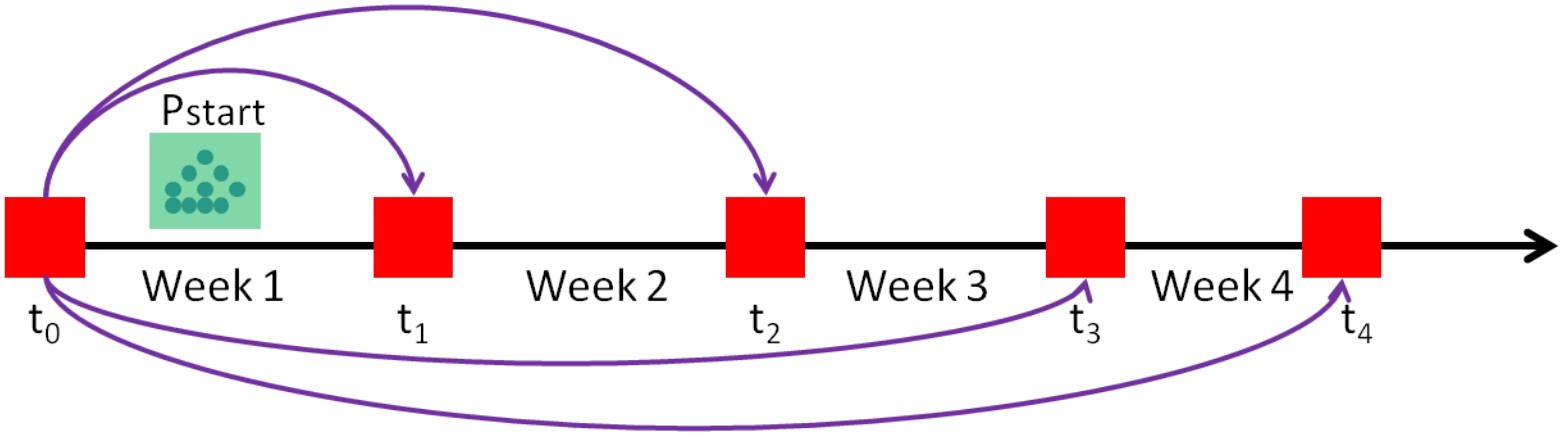
Bipartite graph of departments and patients

Health interaction network



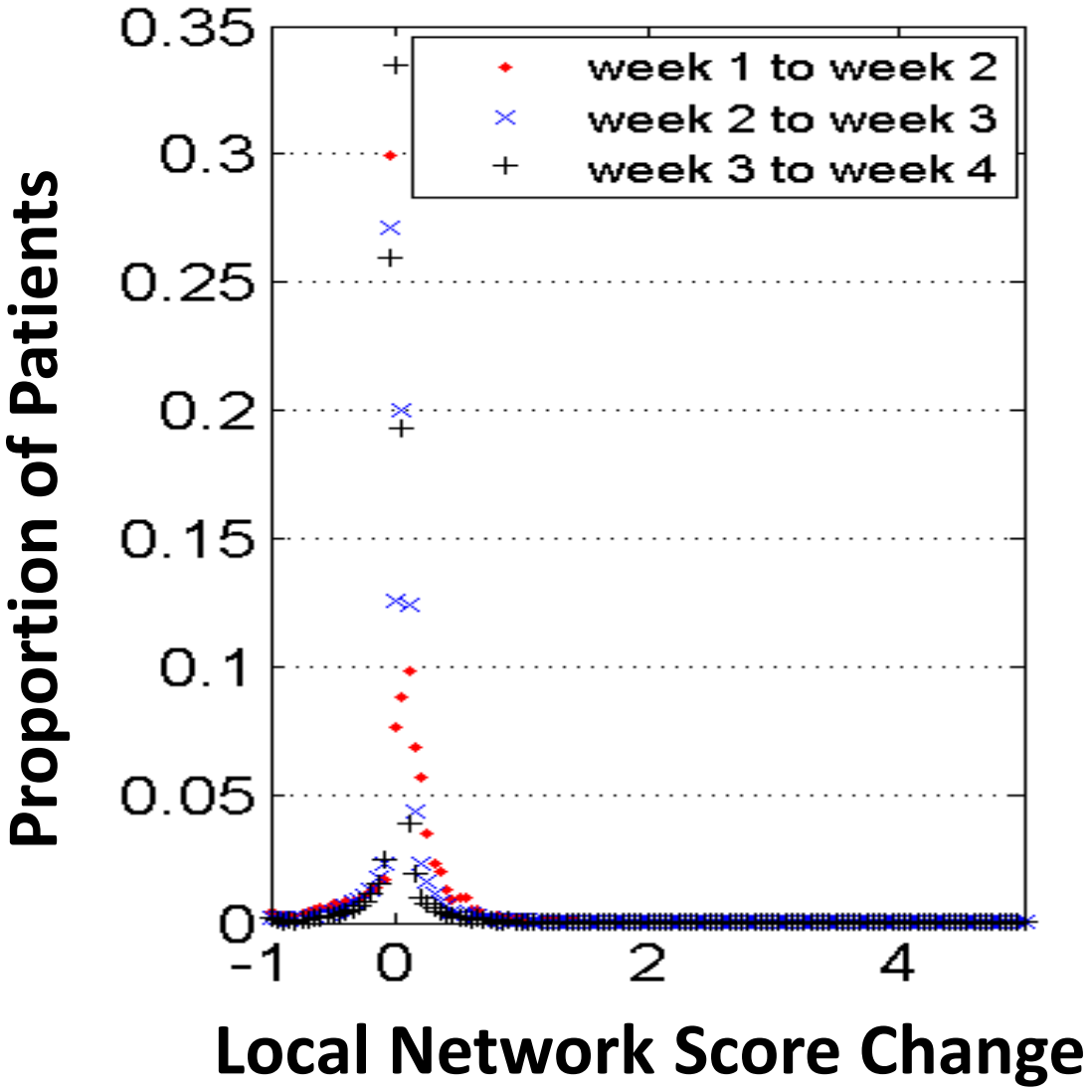
Local view for p_6

Evolution of Local Networks in Terms of Local Network Score and Local Network Reciprocity

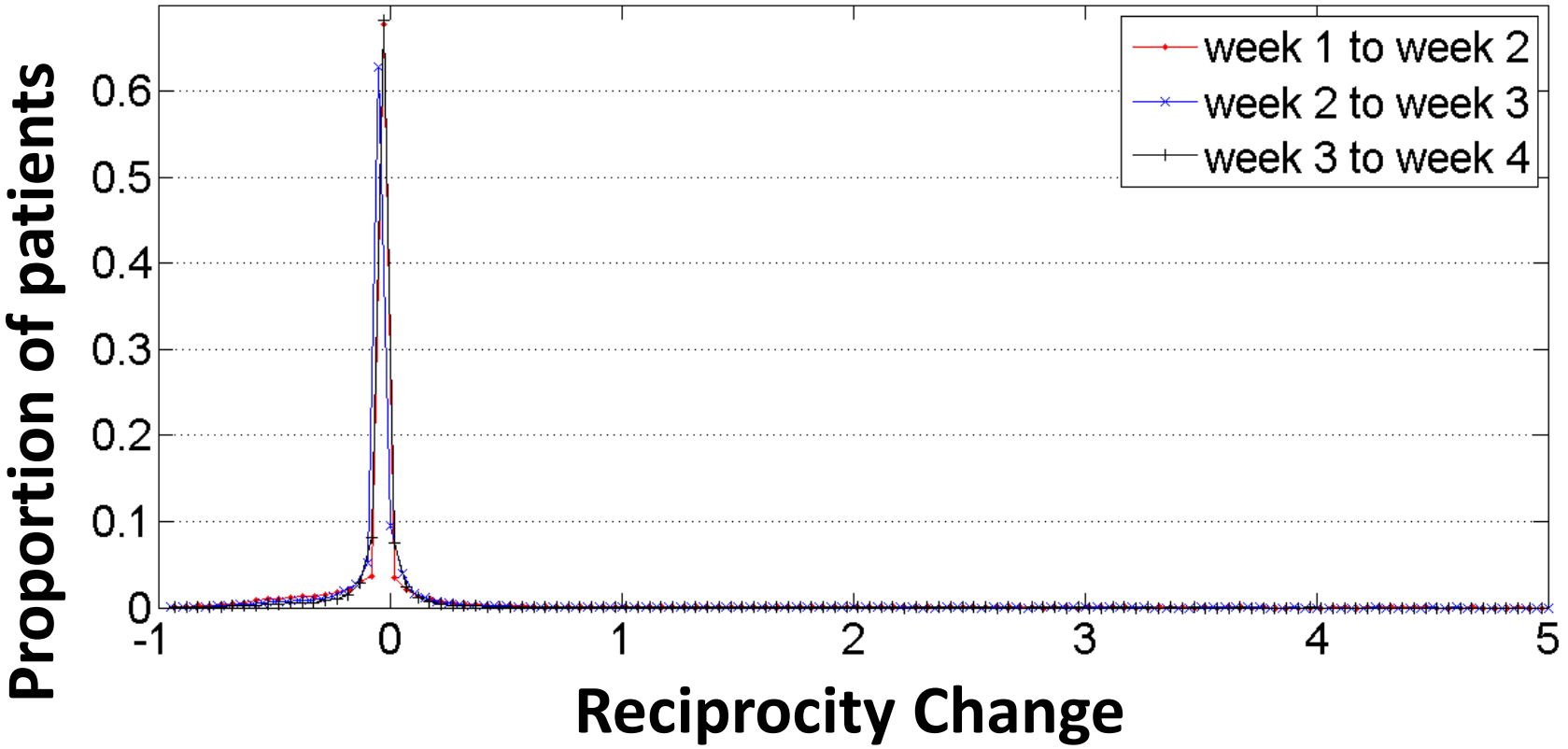


Each point in P_{start} corresponds to a local network

Over 98% of patients are normal because they exhibit a score change <0.05



Approximately 99% of patients are normal because they have a change of reciprocity <0.1



Where are We Going?

A Global Network of Departments

Two metrics: certainty and reciprocity

Stable status in terms of the two metrics

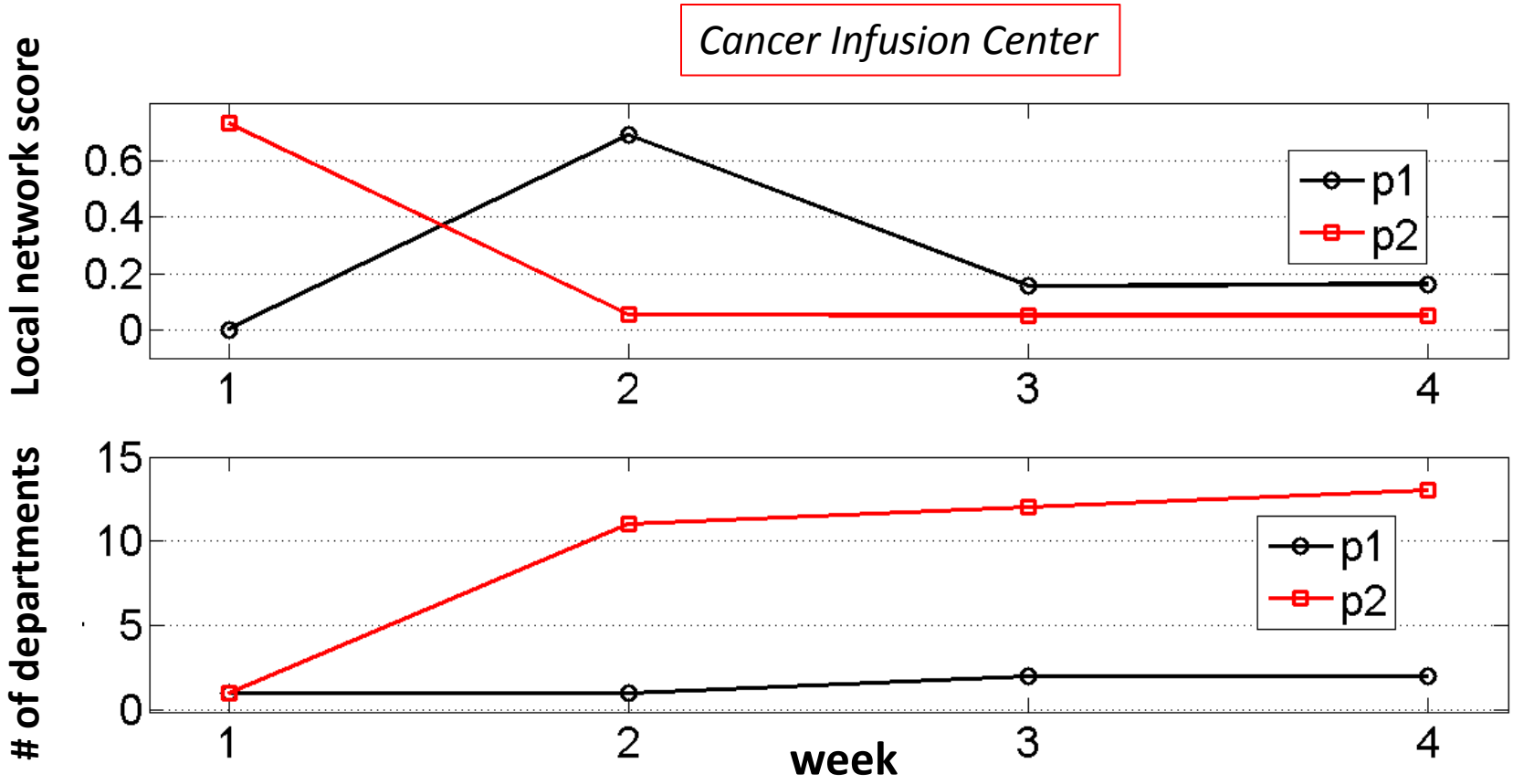
Local Network-for a specific patient

Two metrics: local network score and reciprocity

Application of the Networks

Detecting patients with anomalous medical records accesses

p2 has -0.93 change of local network score and -0.79 change of local reciprocity from the 1st to the 2nd week



Breast Center, [Anonymized Street Location], Care/Eskind Diab Acces, Disease Management Service, Eskind Diabetes - Adult, Free Stipends, Internal Medicine, VIM, VMG Physician Billing Services, Vanderbilt Home Care Primary

Conclusions

- We hypothesized an HCO would exhibit strong stability

  → **confirmed by our experiments**

- We can characterize how strange a patient's local network appears
  - Two groups of patients; those with **small changes** in local network score and reciprocity score and those with **significant changes**
  - The changes in the latter group do not justify the claim that the patient has been intruded upon, but may provide a reason for an investigation that incorporates more nuanced domain knowledge

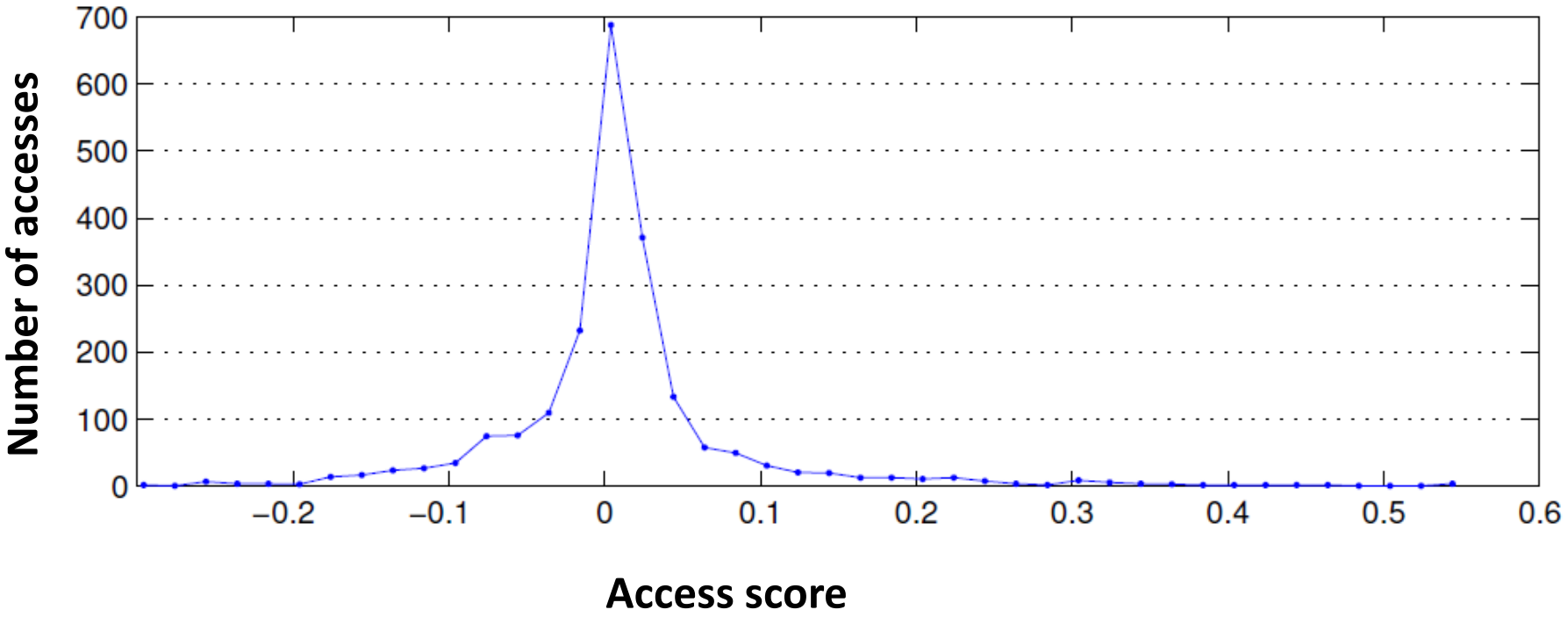
Some Limitations

- Global and local networks appear to represent the business processes of HCO departments
  - however, such claims must be confirmed with employees knowledge about the working of the medical center and its affiliated clinics

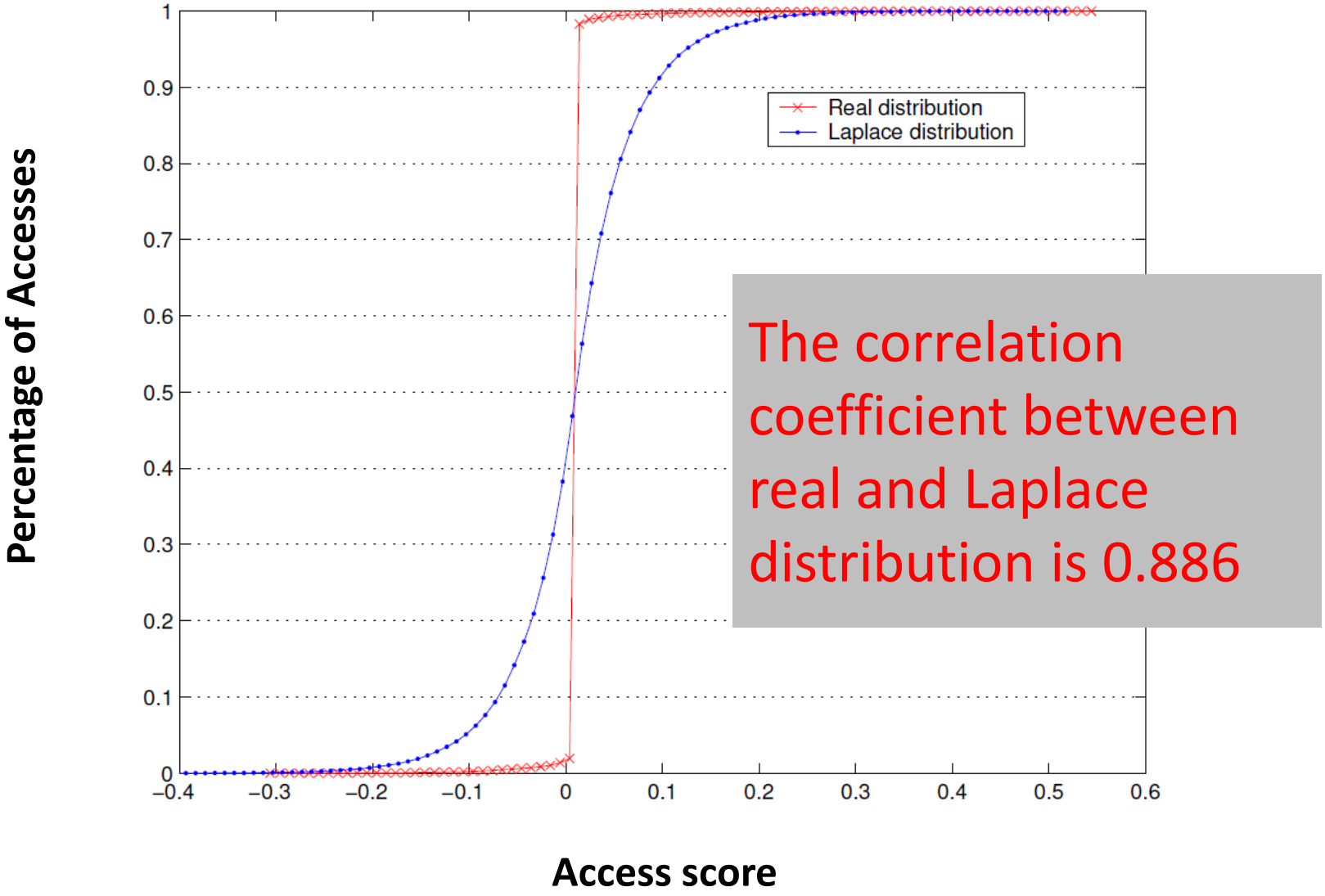
- Need to specialize tool to account for semantics of patients
  - Patient: {Diagnosis, Procedure, Demographics, Residence, physical location in a hospital}

  - Incorporating semantics about the patient, p_2 in the last figure may have no intrusion; rather it is likely a complex cancer patient, which could be confirmed by inspection of clinical documents in the medical record

Thanks!



SNAD assumes that access scores are approximately distributed around a well-centered mean.



The correlation coefficient between real and Laplace distribution is 0.886