# Evaluation of Effects on Retrieval Performance for an Adaptive User Model

Hien Nguyen, Eugene Santos Jr., Qunhua Zhao and Chester Lee

University of Connecticut

Department of Computer Science and Engineering

# Outline

- Motivation
- Background of our adaptive user model
- Evaluation
  - Standard procedure
  - New procedure
- Conclusion and Lessons learned

# Motivation

- Problems with evaluation the effectiveness of user model for information retrieval
  - Information Retrieval (IR): standard evaluation procedures and collections available, but user models constructed using IR techniques are short-lived.
  - User Modeling (UM): evaluating long-term effects of user model but each technique uses its own collections, users, procedures and thus makes it difficult to compare.

- Main idea of our approach:
  - Using procedures, collections and metrics of IR to evaluate short-term and long-term effects of user model by
    - Simulate standard procedure from IR
    - Create a new procedure for evaluating special features of user model

# Novelty of our evaluation approach

- Objectivity
- Inexpensiveness
- Comparability

# Our user modeling approach

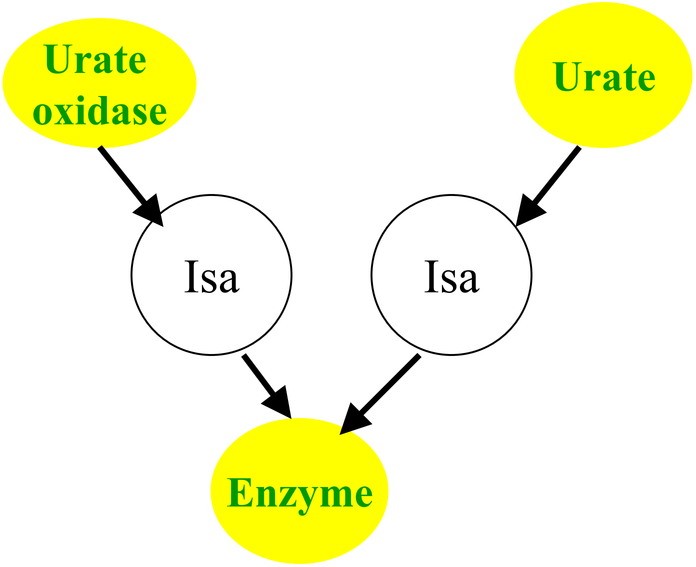
We try to improve retrieval performance and user satisfaction by developing a cognitive user model to capture user intent dynamically by analyzing behavioral information of retrieved relevant documents *(IAT01, UM03, HFES03, HFES04, AH04)*

# IPC User Model
## (*ICAI00, IAT01*)

- Captures ***user intent***.
- Consists of 3 components:
  - **User interests (I)**: *"**What** needs to be done or accomplished?"*
  - **User preferences (P)**: *"**How** is something done or accomplished?"*
  - **User context (C)**: *"**Why** is the user trying to accomplish something?"*

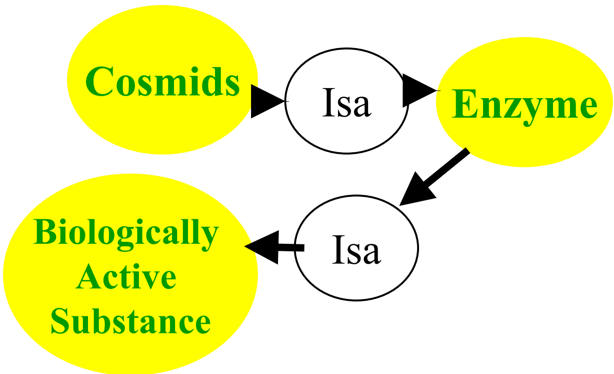
# Context Network (C)

- Captures user knowledge. It contains *concept nodes* and *relation nodes*.

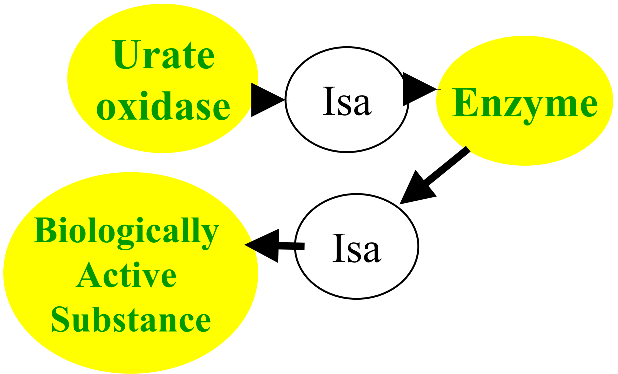



- Is constructed "*on-the-fly*" by finding intersections of all retrieved relevant document graphs.

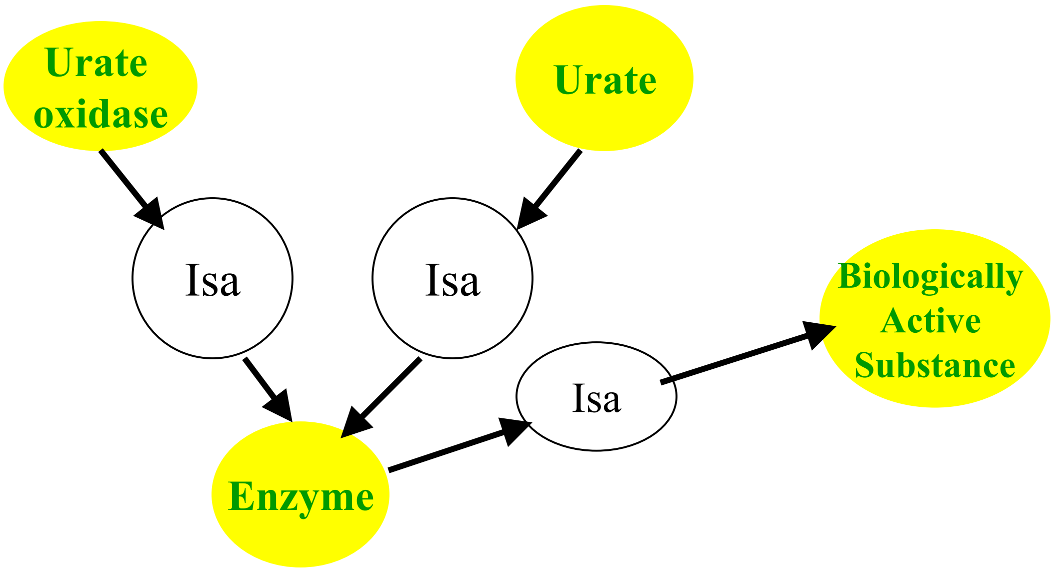
# Context network (C)(cont.)




(a)




(b)



# Interest Set (I)

- Determines what is currently relevant to a user.

- Constructs from user query, set of subgraphs in the intersections of retrieved relevant documents.

- Each element of interest set consists of ***interest concept*** *(a)* and ***interest level*** *L(a)*.

- Fading mechanism:

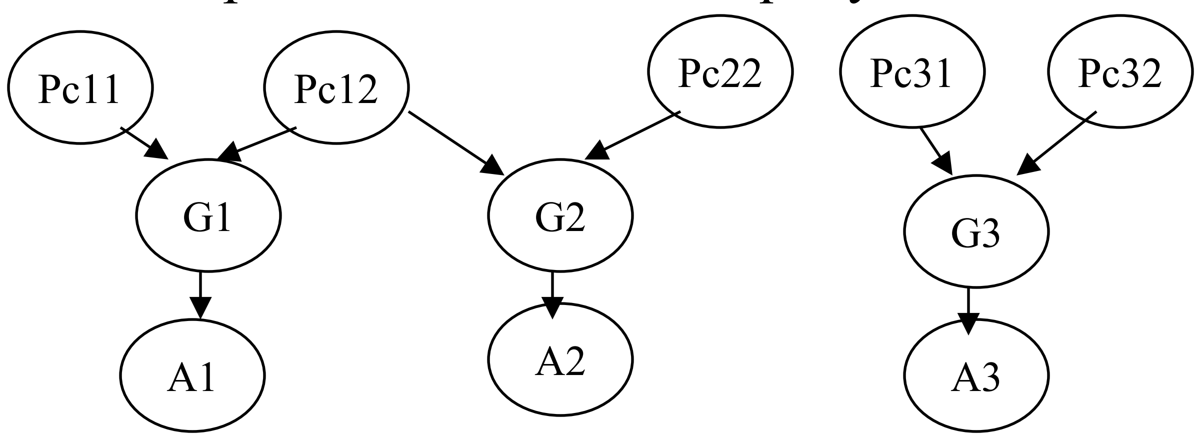
  $L(a) = 0.5*(L(a) + n/m)$

  n: number of retrieved relevant document with *a*

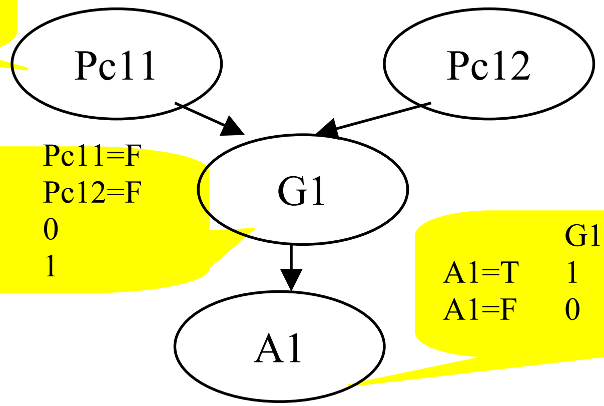
  m: number of retrieved documents

# Preference Network (P)

- Represents how a user wants to form a query
- Is represented using Bayesian networks.
- Consists of pre-condition, goal and action nodes
  - Pre-condition: represents the requirement of a tool used to form a query
  - Goal: represents a tool to form a query (filter/expander)
  - Action: represents the modified query



# Preference network (P) (cont.)




| Pc11=T | Pc11=F |
|---|---|
| 0.9 | 0.1 |

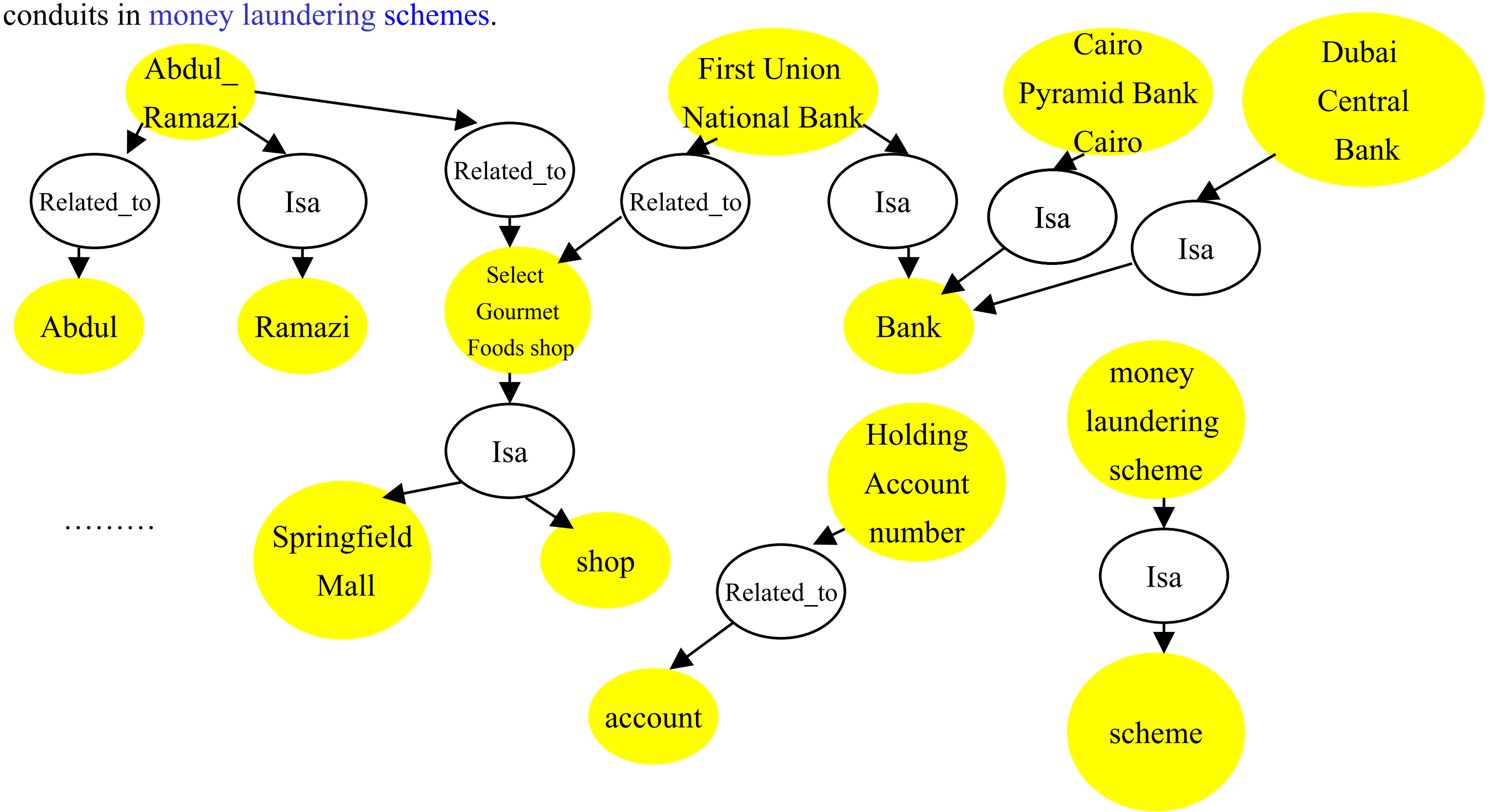
| | Pc11=T<br>Pc12=T | Pc11=T<br>Pc12=F | Pc11=F<br>Pc12=T | Pc11=F<br>Pc12=F |
|---|---|---|---|---|
| G1=T | 1 | 0 | 0 | 0 |
| G1=F | 0 | 1 | 1 | 1 |

| | G1=T | G1=F |
|---|---|---|
| A1=T | 1 | 0 |
| A1=F | 0 | 1 |

# Example of Document Graph

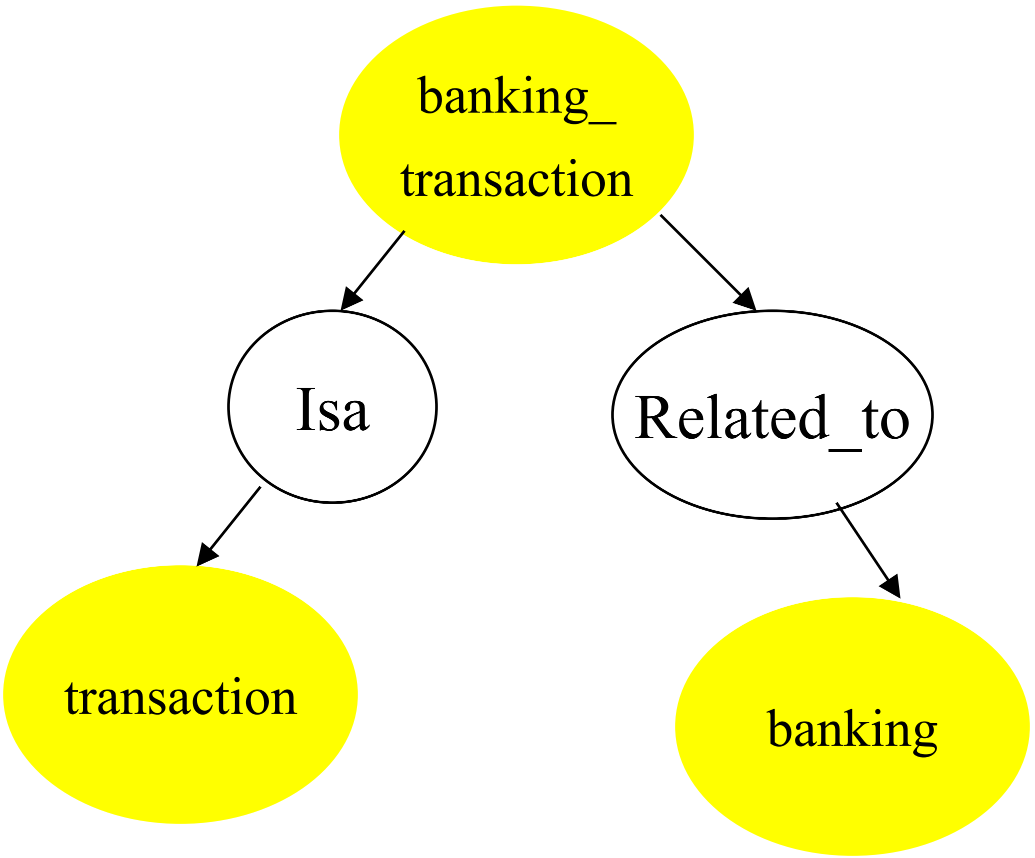
FBI 1) Report Date: 1 April, 2003.

FBI: Abdul Ramazi is the owner of the Select Gourmet Foods shop in Springfield Mall. First Union National Bank lists Select Gourmet Foods as holding account number. Six checks totaling $35.000 have been deposited in this account in the past four months and are recorded as having been drawn on accounts at the Pyramid Bank of Cairo, Egypt and the Central Bank of Dubai, United Arab Emirates. Both of these banks have just been listed as possible conduits in money laundering schemes.

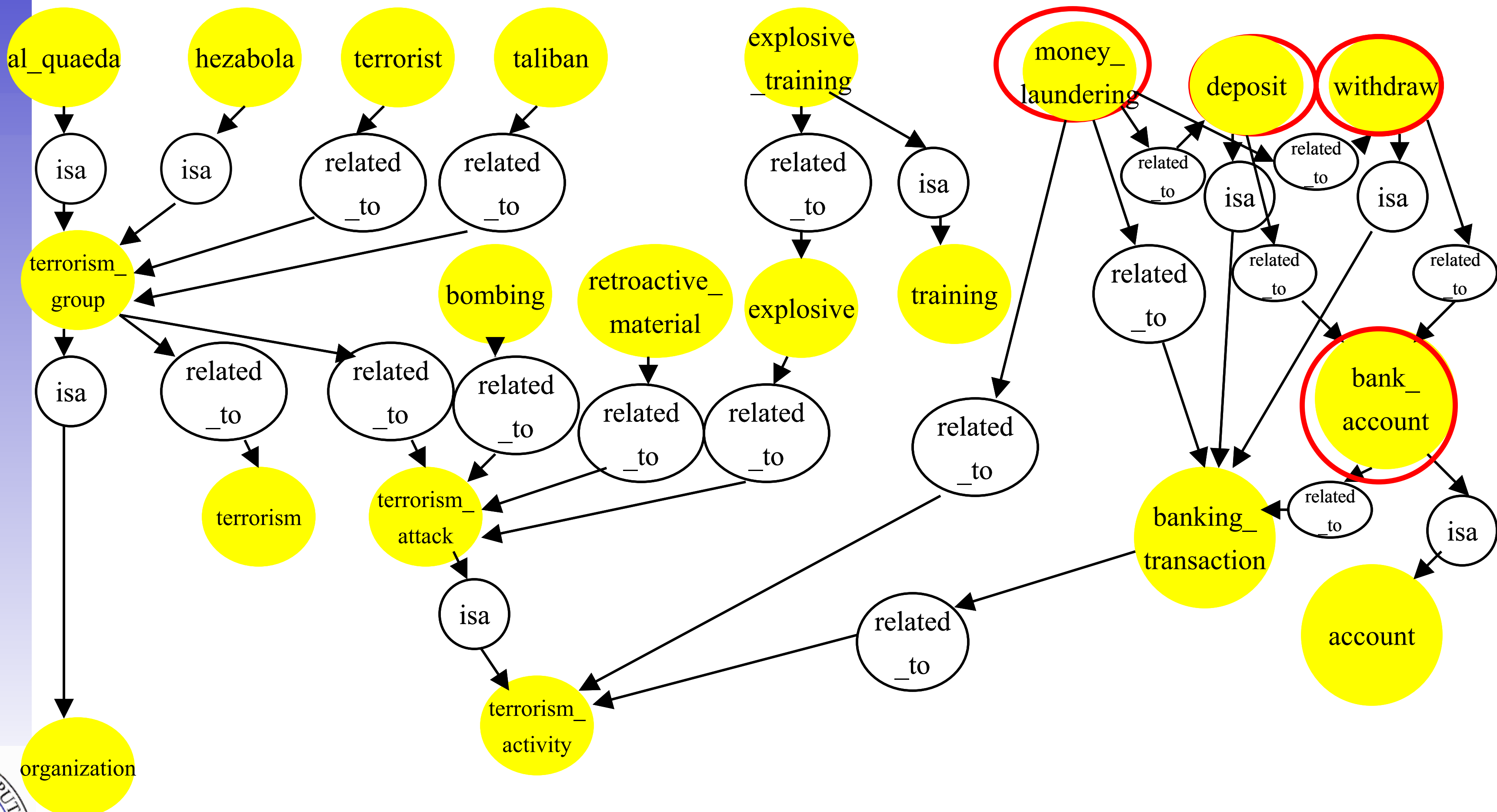


# Example of Query Graph

Query:1. banking transaction:



# Context Network



# Interest Set

Interest set:

terrorism, level=0.96

terrorist, level=0.9

money_laundering, level=0.93

al_quaeda, level=0.85

deposit, level=0.82

withdraw, level=0.81

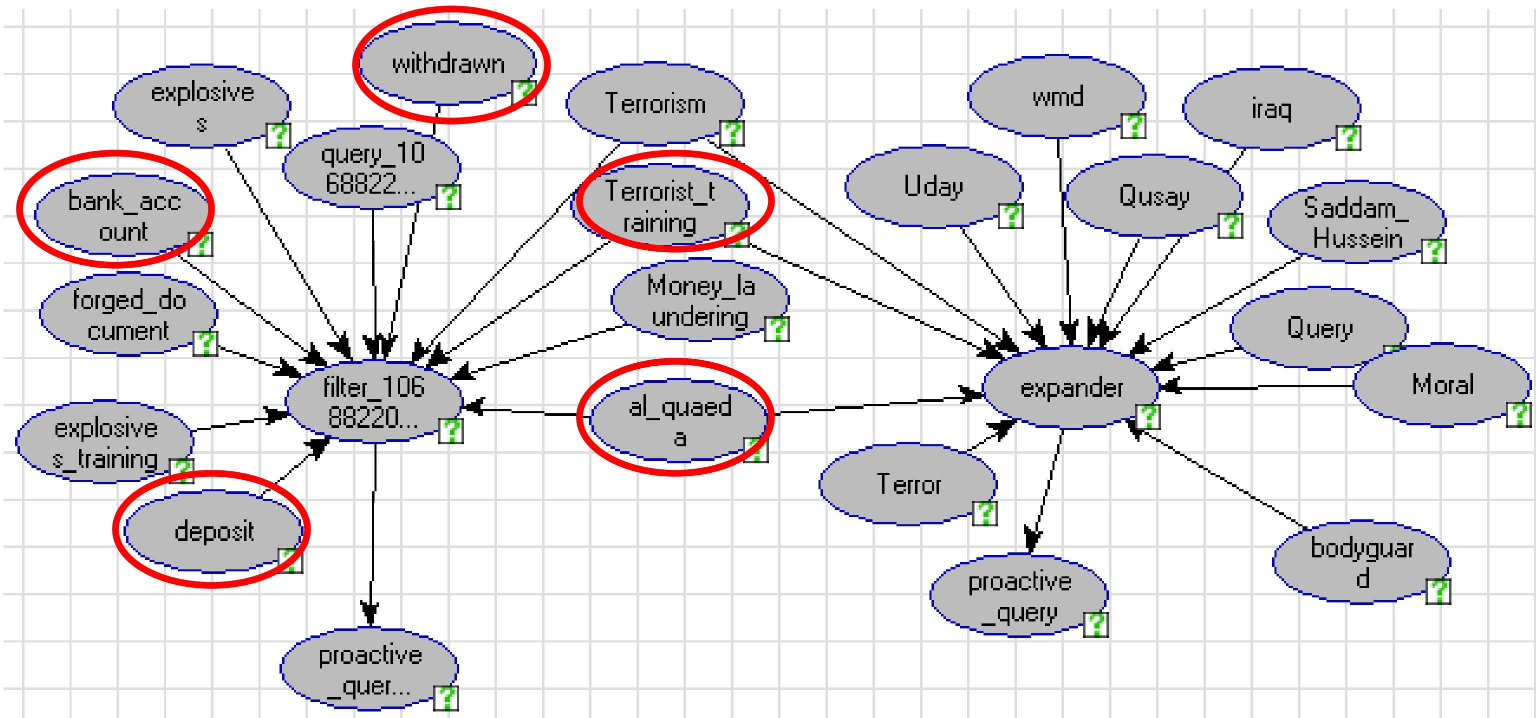
explosion, level=0.8

terrorist_training, level=0.79

forged_document, level=0.77

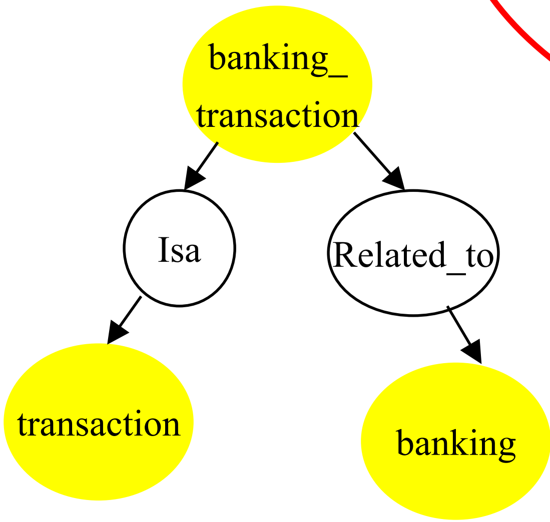
bank_account, level =0.76

# Preference Network

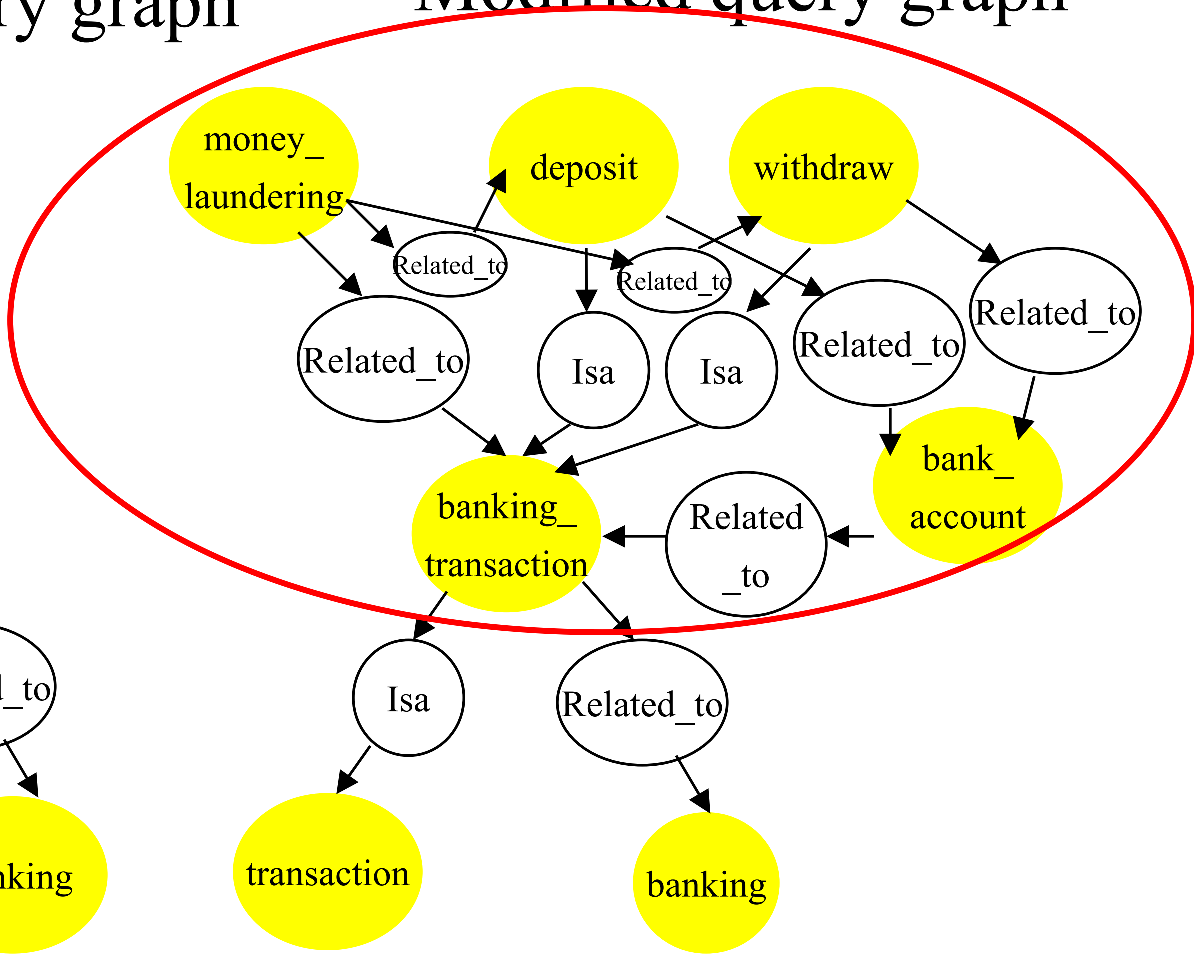


# Modified Query Graph

Original query graph




Modified query graph



# Evaluation

- Objectives:
  - Comparability
  - Coverage
  - New procedure to evaluate special features offered by our long-lived user model.
    - Use of prior knowledge
    - Effects of relevance feedback long-term and short-term.

- Testbeds
  - CACM (computer science and engineering): 3024 documents, 64 queries.
  - Medline (Medical): 1033 documents, 30 queries.

- Comparing technique
  - Ide dec-hi using term frequency inverted document frequency (TFIDF) (Salton & Buckley 90) (Lopez-Pujalte et al 03)

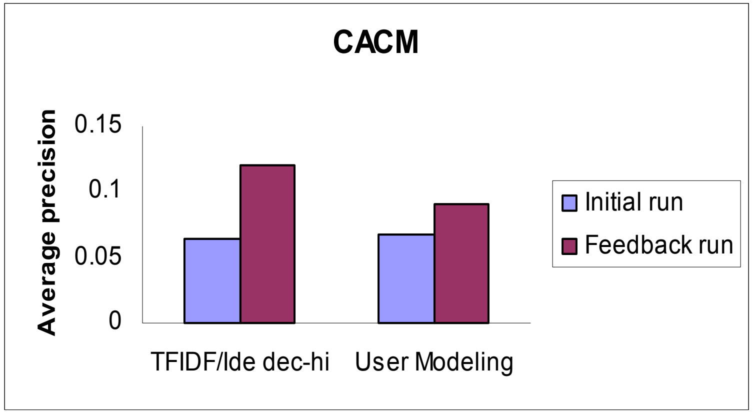
# Evaluation procedures

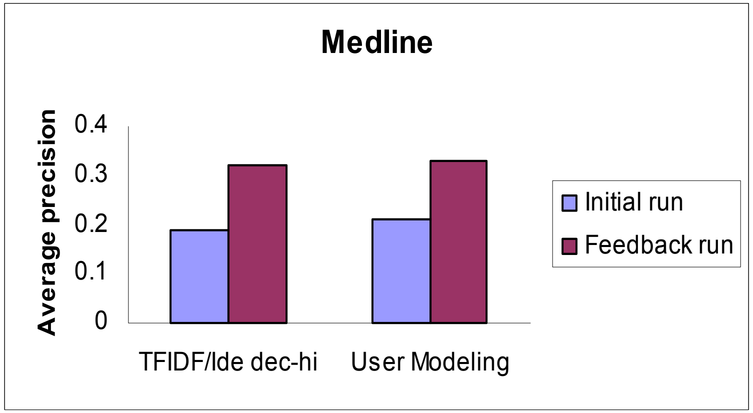
- Standard procedure
  - Ide dec-hi with TFIDF:
    - Convert documents, query to vectors.
    - Compute weight for each term (Salton & Buckley 90) and compute cosine similarity between a document and a query
    - Run each query ***(initial run)*** . Take top 15 for relevancy assessment and feedback. Run query again (***feedback run***).
  - Our user modeling approach
    - Start with empty user model
    - Each query is converted to a query graph (QG)
    - Run each query twice as the same as Ide dec-hi/TFIDF
    - Reset the user model to empty

# Evaluation procedures

- New procedure:
  - Assess the long-term effects of our user model
  - Assess the effects of prior knowledge
  - Assess the  effects of combination between prior knowledge and short-term/long-term effects
- Experiments:
  - Experiment 1:
    - Start with empty user model. Run each query twice but don't reset user model. Saved as the *seed* user model.
  - Experiment 2:
    - Start with seed user model. Don't update user model.
  - Experiment 3:
    - Start with seed user model. Run each query twice. Reset user model to seed user model.
  - Experiment 4:
    - Start with seed user model. Run each query twice. Don't reset user model

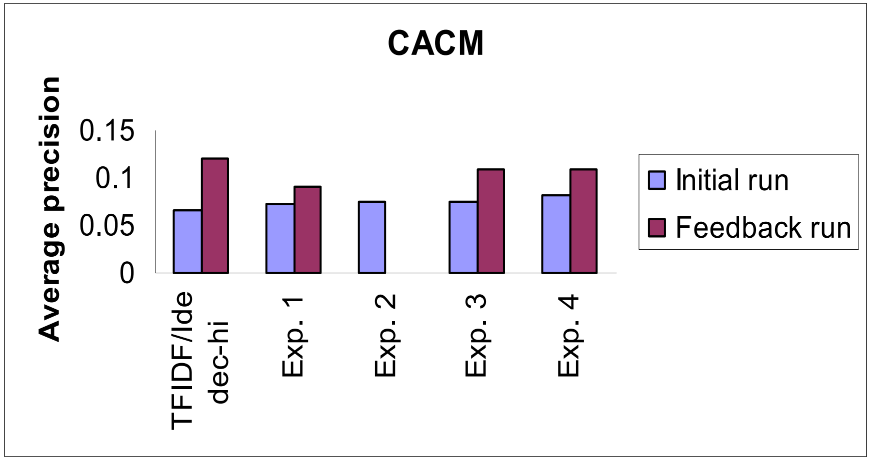
# Results for standard procedure

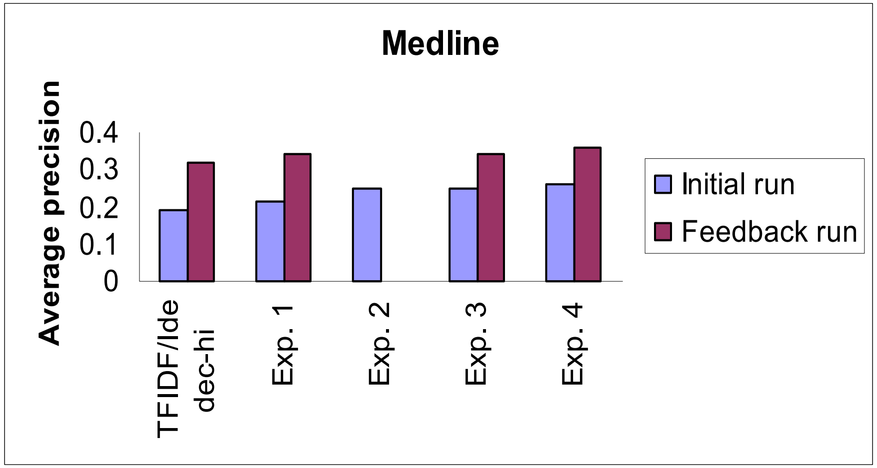



**Average precision at three point fixed recalls on residual collections**

# Results for special procedure




**Average precision at three point fixed recalls on residual collections**

# Analysis

- We achieve competitive results in both initial run and feedback run for standard procedures with CACM and Medline
- By using seed user model, the precision is increased in Experiment 2.
- By using seed user model and long-term user model, we achieve better performance in the initial runs than TFIDF and competitive performance in the feedback run compared to Ide dec-hi.

# Conclusion

- We can evaluate the retrieval effectiveness of a user model using procedures, collections and metrics of IR by:
  - Standard procedure: offering the comparability
  - New procedure: accessing long-term effects and effects of prior knowledge.

# Lessons learned from evaluation

- Re-ordering the questions with involvement from real users.
- Using different seed user model (manually constructed, constructed from training set)
- Combine the results of this phase with other two phases which assess how accurate user intent has been captured and how improved is the user performance.

# References
## (http://excalibur.engr.uconn.edu/)

### Published/to appear papers


H. Nguyen, E. Santos Jr., Z. Qunhua and C. Lee. Evaluation of effects on retrieval performance of an adaptive user model. *Workshops of Evaluation of adaptive user model AH 2004*. August 2004. Eindhoven, the Neithelands. To appear.

H. Nguyen, E. Santos Jr., Z. Qunhua and H. Wang. 2004. Capturing User Intent for Information Retrieval. To appear *in the Proceedings of the 48th Annual meeting for the Human Factors and Ergonomics Society (HFES-04)*. October 2004, New Orleans.

E. Santos Jr., H. Nguyen, Z. Qunhua, and H. Wang. 2003. User Modelling for Intent Prediction in Information Analysis. In *Proceedings of the 47th Annual Meeting for the Human Factors and Ergonomincs Society (HFES-03)*, Denver, CO. Pages 1034-1038.

E. Santos Jr., H. Nguyen, Z. Qunhua, and E. Pukinskis. 2003. Empirical Evaluation of Adaptive User Modeling in a Medical Information Retrieval Application. *Lecture Notes in Artificial Intelligence 2702: User Modeling 2003 (Eds. P. Brusilovsky, A. Corbett, and F. de Rosis), Springer.* Pages 292-296

E. Santos Jr., H. Nguyen and S.M. Brown. Kavanah: An Active User Interface Information Retrieval Application. *Proceedings of the 2nd Asia-Pacific Conference on Intelligent Agent Technology*. Maebashi, Japan, October 2001. Pages 412-423.

H. Nguyen, M. G. Saba, E. Santos, Jr. and S. M. Brown. 2000. Active user interface in a knowledge discovery and retrieval system. In *Proceedings of the International Conference on Artificial Intelligence (ICAI2000).* Pages 339-344.

E. Santos Jr., H. Nguyen. S. M. Brown. 2000. Medical Document Information Retrieval through Active User Interfaces. In *Proceedings of the International Conference on Artificial Intellgence (ICAI00).* Pages 323-329. (invited paper)

# Acknowledgement


- This research has been funded by AFRL Human Effectiveness Directorate Through Sytronics Inc. and Advanced Research and Development Activity (ARDA) – US Government.

# Questions?