Risk Assessment Based Access Control with Text and Behavior Analysis for Document Management

Zhuo Lu, University of South Florida
Yalin Sagduyu, Intelligent Automation Inc.
Outline

• Motivation
  – Risk assessment based access control

• Models and Methods
  – Two risk assessment modules

• Evaluation Results

• Conclusion
Document Management

• Documents with sensitive information for business, government, and military operations must be classified and accessible only to appropriate personnel

• Example:
  – Unclassified (U), Confidential (C)
  – Secret (S) and Top Secret (TS)
Traditional Access

• Binary decision:
  – Alice has a ‘Secret’ level.
    • Able to access ‘Secret’ or lower level documents.
  – Potential security issues:
    • Account hacking
    • Insider threat
    • Classification error (due to human or machine errors)
    • ...

• Potential security issues:
Motivation

• We are motivated to design a new access control mechanism to protect sensitive information from unintentional or malicious access and disclosure
  – Traditional:
    • granting document access when a user has such access
  – Our method:
    • assess the risk of document disclosure to such a user
      – Scanning textual content and analyzing behavior
Research Question

• Research Question:
  – Given all $N$ documents $\{D_k\}$, $k \in [1,N]$ stored in the system and a user’s security clearance level $C_u$,
  – Determine whether to grant the user’s current request to access documents $\{P_k\}$, $k \in [1,L]$, where $L$ is the number of currently requested documents.
Models

• Each document can be classified into one level in an ordered set of classification levels.
  – E.g., \{U, C, S, TS\} with U < C < S < TS

• Classification
  – is a function, mapping from a document to a security classification level
    • either by human being or a classification algorithm.
Risk Assessment based Access Control

- Two key components:
  - Analysis of Erroneous Disclosure (AED)
  - Analysis of Abnormal Behavior (AAB)
Analysis of Erroneous Disclosure (AED)

• The role of the AED module is to determine the risk of these documents being erroneously classified.

• Such a risk contains two major factors:
  – Risk due to classification errors/mismatch, called Type I risk.
  – Risk due to information similarity, called Type II risk.
Analysis of Erroneous Disclosure (AED)

- **Type I risk**: Risk due to classification errors:
  - the risk of information leakage
    - A high level document is classified as a low level
  - the risk of information blocking
    - A low level document is classified as a high level
Bayes Classifier

\[ D: \quad f \quad \rightarrow \quad \text{mission} \quad \text{locations} \quad \text{time} \quad \text{specification} \]

Chose ‘S’ because the probability is maximum.

Information blocking

Information leakage

Document Storage/Training
Two Sub-Risks under Bayes Rule

• The risk of information leakage

\[ R_d(P_{in}) = \sum_{s > c} \pi(s) = \sum_{s > c} \mathbb{P}(f|s), \]

• The risk of information blocking

\[ R_b(P_{in}) = \sum_{s < c} \pi(s) = \sum_{s < c} \mathbb{P}(f|s). \]

• Type I risk:

\[ R_I(P_{in}) = R_d + R_b \]
Type II Risk

- **Type II risk**: Risk due to information similarity:
  - similar content between documents with different classification levels
Type II Risk

• Model the similar textual content?
  – Using a topic modeling method
  – Latent Dirichlet Allocation (LDA): gives the probability distribution that a document is associated with a certain topic.
Modeling of Type II Risk

- Cosine similarity of two vectors of topic distribution:

\[ C(P_1, P_2) = \frac{\sum_{i=1}^{T} |X_1(i)X_2(i)|}{\sqrt{\sum_{i=1}^{T} X_1(i)^2} \sqrt{\sum_{i=1}^{T} X_2(i)^2}} \]
Overall Risk Assessment

- Combing Type I and Type II:

\[ R_o(P_{in}) = w_1 R_I(P_{in}) + w_2 R_{II}(P_{in}), \]

weights
Risk Assessment based Access Control

• Two key components:
  – Analysis of Erroneous Disclosure (AED)
  – Analysis of Abnormal Behavior (AAB)
Analysis of Abnormal Behavior (AAB)

- Intuition: Due to the nature of a user’s job, the documents that the user requests should have similar, if not the same, topics in routine operations.
- Hacking: the hacker uses a user’s account to download all possible documents that the account has the access to, which should exhibit quite distinct topic
Evaluation

• 100 documents from the Internet
  – Treated as notional ‘classified’ documents.
  – Classified into levels: A(lowest), B, C, D (highest).
  – Intentionally misclassified some documents
Analysis of Erroneous Disclosure (AED) Evaluation

Overall Risk

\[ R_0 \]

Intentionally misclassified

0.5

threshold

index of document

1 2 3 4 5 6 7
Analysis of Abnormal Behavior (AAB) Evaluation

risk of abnormal behavior $R_a$

account hacked, maliciously.document access, AAB detection

0.15

threshold

index of user request

1 2 3 4 5 6
Conclusions

• We proposed an access control mechanism:
  – Two relatively independent AED and AAB modules
  – Text analysis and behavior analysis to quantify the risk of access to certain documents
  – grant the user access only if the risk is assessed low with respect to the user’s credentials.

• Rudimentary evaluation results
  – More comprehensive tests
  – better testing using public declassified documents.
Thank you!

Q/A?