Protection of Database Security via Collaborative Inference Detection[1]

Yu Chen and Wesley W. Chu

Computer Science Department
University of California, Los Angeles, CA 90095
Email: {chenyu,wwc}@cs.ucla.edu

Abstract

Malicious users can exploit the correlation among data to infer sensitive information from a series of seemingly innocuous data accesses. Thus, we develop an inference violation detection system to protect sensitive data content. Based on data dependency, database schema and semantic knowledge, we constructed a semantic inference model (SIM) that represents the possible inference channels from any attribute to the pre-assigned sensitive attributes. The SIM is then instantiated to a semantic inference graph (SIG) for query-time inference violation detection. For a single user case, when a user poses a query, the detection system will examine his/her past query log and calculate the probability of inferring sensitive information. The query request will be denied if the inference probability exceeds the pre-specified threshold. For multi-user cases, the users may share their query answers to increase the inference probability. Therefore, we develop a model to evaluate collaborative inference based on the query sequences of collaborators and their task-sensitive collaboration levels. Experimental studies reveal that information authoritativeness, communication fidelity and honesty in collaboration are three key factors that affect the level of achievable collaboration. An example is given to illustrate the use of the proposed technique to prevent multiple collaborative users from deriving sensitive information via inference.

1 This research is supported by NSF grant number IIS-03113283

1. Introduction

Access control mechanisms are commonly used to protect users from the divulgence of sensitive information in data sources. However, such techniques are insufficient because malicious users may access a series of innocuous information and then employ inference techniques to derive sensitive data using that information.

To address this inference problem, we develop an inference detection system that resides at the central directory site. Because inference channels can be used to provide a scalable and systematic sound inference, we need to construct a semantic inference model (SIM) that represents all the possible inference channels from any attribute in the system to the set of pre-assigned sensitive attributes. The SIM can be constructed by linking all the related attributes which can be derived via attribute dependency from data dependency, database schema and semantic related knowledge. Based on the semantic inference model, the violation detection system keeps track of a user's query history. When a new query is posed, all the channels where sensitive information can be inferred will be identified. If the probability to infer sensitive information exceeds a prespecified threshold, the current query request will then be denied. Therefore, our system can prevent malicious users from obtaining sensitive information.

This inference detection approach is based on the assumption that users are isolated and do not share information with one another. This assumption, however, may not be the case in a real-life situation. Most users usually work as a team, and each member can access the information independently. Afterwards, the members may merge their knowledge together and jointly infer the sensitive information. Generalizing from a single-user to a multi-user collaborative system greatly increases the complexity of the inference detection system.

For example, one of the sensitive attributes in the system can be inferred from four different

inference channels. There are two collaborators and each poses queries on two separate channels. Based on individual inference detection, neither of the users violates the inference threshold from their query answers. However, if the two users share information, then the aggregated knowledge from the four inference channels can cause an inference violation (see Section 7.2).

This motivates us to extend our research from a single user to the multiple user case, where users may collaborate with each other to jointly infer sensitive data. We have conducted a set of experiments, using our inference violation detector as a test bed to understand the characteristics in collaboration as well as the effect on collaborative inference. From the experiments, we learn that for a given specific task, the amount of information that flows from one user to another depends on the closeness of their relationships and the knowledge related to the task. Thus, collaborative inference for a specific task can be derived by tracking the query history of all the users together with their collaboration levels.

This paper is organized as follows. Section 2 presents related work. Section 3 introduces a general framework for the inference detection system, which includes the knowledge acquisition module, semantic inference model and violation detection module. Section 4 discusses how to acquire and represent knowledge that could generate inference channels. Section 5 integrates all possible inference channels into a Semantic Inference Model which can be instantiated and then mapped into a Bayesian network to reduce the computation complexity for data inference. As shown in Section 6, we are able to detect inference violation at query time for both individual user and multiple collaborative users. Section 7 presents an example to illustrate the use of the proposed technique for collaboration inference detection. Section 8 presents collaboration level experiments and their estimations. Section 9 discusses the robustness of inference detection and threshold determination via sensitivity analysis. Section 10 presents the conclusion.

2. Related Work

Database inferences have been extensively studied. Many approaches to address the inference problem were presented in [FJ02]. Particularly, Delugach and Hinke used database schema and human-supplied domain information to detect inference problems during database design time [DH96, HD92, HDW96]. Garvey *et al.* developed a tool for database designers to detect and remove specific types of inference in a multilevel database system [GLQ92]. Both approaches use schema-level knowledge and do not infer knowledge at the data level. These techniques are also used during database design time and not at run time. However, Yip *et al*. pointed out the inadequacy of schema-level inference detection, and he identifies six types of inference rules from the data level that serve as deterministic inference channels [YL98]. In order to provide a multilevel secure database management system, an inference controller prototype was developed to handle inferences during query processing. Rule-based inference strategies were applied in this prototype to protect the security [TFC93]. Further, since data update can affect data inference, [FTE01] proposed a mechanism that propagates update to the user history files to ensure no query is rejected based on the outdated information. To reduce the time in examining the entire history log in computing inference, [TFE05] proposed to use a prior knowledge of data dependency to reduce the search space of a relation and thus reduce the processing time for inference. Open inference techniques were proposed to derive approximate query answering when network partitions occurred in distributed databases. Feasible open inference channels can be derived based on query and database schema [CCH94].

The previous work on data inference mainly focused on deterministic inference channels such as functional dependencies. The knowledge is represented as rules and the rule body exactly determines the rule head. Although such rules are able to derive sound and complete inference,

much valuable non-deterministic correlation in data is ignored. For example, salary ranges may not deterministically depend on the ranks. Further, many semantic relationships, as well as data mining rules, can not be specified deterministically. To remedy this shortcoming, we propose a probabilistic inference approach to treat the query-time inference detection problem. The contribution of the paper consists of: 1) Derive probabilistic data dependency, relational database schema and domain-specific semantic knowledge and represent them as probabilistic inference channels in a Semantic Inference Model. 2) Map the instantiated Semantic Inference Model into a Bayesian network for efficient and scalable inference computation. 3) Propose an inference detection framework for multiple collaborative users.

3. The Inference Framework

The proposed inference detection system consists of three modules, as shown in Figure 1: knowledge acquisition, semantic inference model (SIM), and security violation detection including user social relation analysis.

The *Knowledge Acquisition* module extracts data dependency knowledge, data schema knowledge and domain semantic knowledge. Based on the database schema and data sources, we can extract data dependency between attributes within the same entity and among entities. Domain semantic knowledge can be derived by semantic links with specific constraints and rules. A semantic inference model can be constructed based on the acquired knowledge.

The *Semantic Inference Model (SIM)* is a data model that combines data schema, dependency and semantic knowledge. The model links related attributes and entities as well as semantic knowledge needed for data inference. Therefore SIM represents all the possible relationships among the attributes of the data sources. A *Semantic Inference Graph* (*SIG*) can be constructed

by instantiating the entities and attributes in the SIM. For a given query, the SIG provides inference channels for inferring sensitive information.

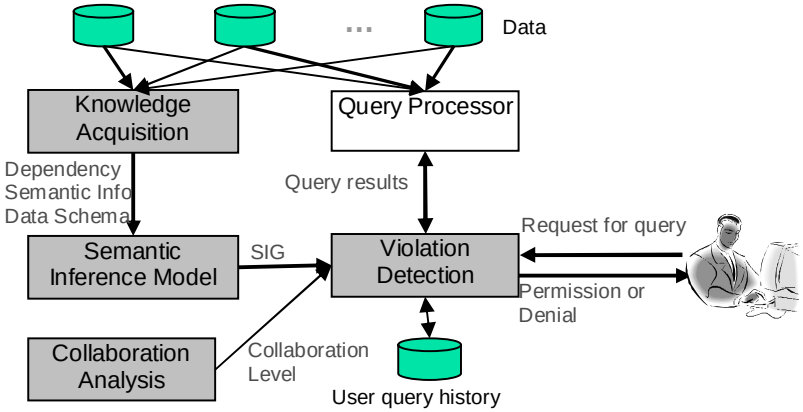


Fig. 1. The framework for an Inference Detection System

Based on the inference channels derived from the SIG, violation detection combines the new query request with the request log, and it checks to see if the current request exceeds the pre-specified threshold of information leakage. If there is collaboration according to collaboration analysis, the *Violation Detection* module will decide whether to answer a current query based on the acquired knowledge among the malicious group members and their collaboration level to the current user.

4. Knowledge Acquisition for Data Inference

Since users may pose queries and acquire knowledge from different sources, we need to construct a semantic inference model for the detection system to track user inference intention. The semantic inference model requires the system to acquire knowledge from data dependency, database schema and domain-specific semantic knowledge. This section will discuss how to acquire that knowledge.

4.1 Data Dependency

Data dependency represents causal relationships and non-deterministic correlations between at-

tribute values. Because of the non-deterministic nature, the dependency between two attributes A and B is represented by conditional probabilities $p_{i|j}=Pr(B=b_i|A=a_j)$. Thus, the non-deterministic data dependency is a more general representation than the relational functional dependency or other types of deterministic relationships. There are two types of non-deterministic data dependencies as defined in the Probabilistic Relational Model [FGK99, GFK01]: *dependency-within-entity* and *dependency-between-related-entities*, as defined in the following.

Dependency-within-entity: Let A and B be two attributes in an entity E; if B depends on A, then for each instance of E, its value of attribute B depends on its value of attribute A with a probability value. To learn the parameter of dependency-within-entities from relational data, from a relational table that stores entity E, we can derive the conditional probabilities $p_{i|j}=Pr(B=b_i|A=a_j)$ via a sequential scan of the table with a counting of the occurrences of A, B, and co-occurrences of A and B.

Dependency-between-related-entities: Let A be an attribute in entity E_1 and C be an attribute in E_2, and E_1 and E_2 are related by R, which is a relation that can be derived from database schema. If C depends on A, then only for related instances of E_1 and E_2, the value of attribute C in E_2 instances depends on the value of attribute A in related instances of E_1. Such dependency-between-related-entities only exists for related instances of entities E_1 and E_2. The parameters of dependency-between-related-entities can be derived by first joining the two entity tables based on the relation R and then scanning and counting the frequency of occurrences of the attribute pair in the joined table. If two entities have an m-to-n relationship, then the associative entity table can be used to join the related entity tables to derive dependency-between-related-entities [Dat95].

4.2 Database Schema

In relational databases, database designers use data definition language to define data schema.

The owners of the entities specify the primary key and foreign key pairs. Such pairing represents a relationship between two entities. If entity E_1 has primary key *pk*, entity E_2 has foreign key *fk*, and $e_1.pk=e_2.fk$, then dependency-between-related-entities from attribute A (in e_1) to attribute C (in e_2) can be derived.

4.3 Domain-Specific Semantic Knowledge

Other than data dependencies inside relational data sources, outside information such as domain knowledge can also be used for inferences. Specifically, domain-specific semantic relationships among attributes and/or entities can supplement the knowledge of malicious users and help their inference. For example, the semantic knowledge "can land" between Runway and Aircraft implies that the length of Runway should be greater than the minimum Aircraft landing distance, and the width of Runway should be greater than the minimum width required by Aircraft. If we know the runway requirement of aircraft *C-5*, and *C-5* "can land" in the instance of runway *r*, then the values of attributes *length* and *width* of *r* can be inferred from the semantic knowledge. Therefore, we want to capture the domain-specific semantic knowledge as extra inference channels in the Semantic Inference Model.

Semantic knowledge among attributes is not defined in the database and may vary with context. However, from a large set of semantic queries posed by the users, we can extract the semantic constraints [ZC99]. For example, in the WHERE clause of the following query, clauses #3 and #4 are the semantic conditions that specify the semantic relation "can land" between entity Runways and entity Aircrafts. Based on this query, we can extract semantic knowledge "can land" and integrate it into the Semantic Inference Model shown in Figure 3.[2]

2 Clearly, the set of the semantic queries may be incomplete, which can result in the semantic knowledge being incomplete as well. However, additional semantic knowledge can be appended to the Semantic Inference Model as the system gains more semantic queries. The system can then reset to include the new knowledge. Otherwise, this will result in inference with knowledge update and is beyond the scope of this paper.

- **Query: Find airports that *can land* a C-5 cargo plane.**

```
SELECT AP.APORT_NM
FROM AIRCRAFTS AC, AIRPORTS AP, RUNWAYS R
WHERE AC.AC_TYPE_NM = 'C-5' and                                  #1
AP.APORT_NM = R.APORT_NM and                                     #2
AC.WT_MIN_AVG_LAND_DIST_FT <= R.RUNWAY_LENGHT_FT and             #3
AC.WT_MIN_RUNWAY_WIDTH_FT <= R.RUNWAY_WIDTH_FT;                  #4
```

5. Semantic Inference Model

The Semantic Inference Model (SIM) represents dependent and semantic relationships among attributes of all the entities in the information system. As shown in Figure 2, the related attributes (nodes) are connected by three types of relation links: dependency link, schema link and semantic link.

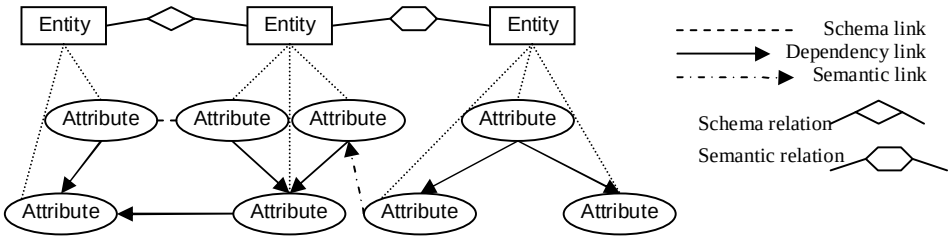


Fig. 2. A Semantic Inference Model. Entities are interconnected by schema relations (diamond) and semantic relations (hexagon). The related attributes (nodes) are connected by their data dependency, schema and semantic links.

Dependency link connects dependent attributes within the same entity or related entities. Consider two dependent attributes A and B. Let A be the parent node and B be the child node. The degree of dependency from B to A can be represented by the conditional probabilities $p_{i|j} = Pr(B=b_i|A=a_j)$. The conditional probabilities of the child node given all of its parents are summarized into a conditional probability table (CPT) that is attached to the child node. For instance, Figure 3b shows the CPT of the node "TAKEOFF_LANDING_CAPACITY" of the SIM in Figure 3a. The conditional probabilities in the CPT can be derived from the database content [FGK99, GFK01]. For example, the conditional probability $Pr(B=b_i|A=a_j)$ can be derived by counting the co-occurrence frequency of the event $B=b_i$ and $A=a_j$ and dividing it by the occur-

rence frequency of the event $A=a_j$.

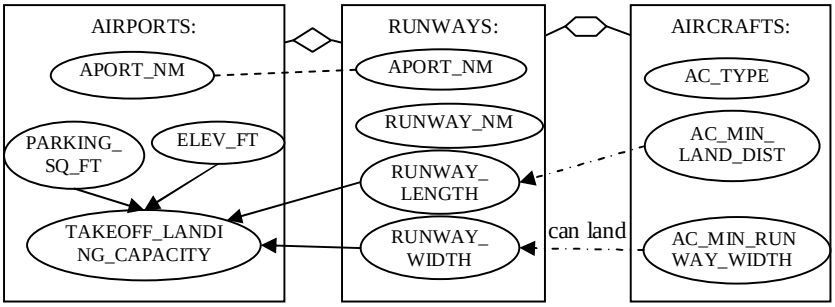


Fig. 3a. A Semantic Inference Model example for Airports, Runways and Aircraft

	Conditional Probability of TAKEOFF_LANDING_CAPACITY																							
Cond	parking_sq_ft	small												large										
	elev_ft	low						high						low						high				
	runway_length	short		medium		long		short		medium		long		short		medium		long		short		medium		long	
	runway_width	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide	narrd	wide
Takeoff_Landing_Cap	small	0.9	0.8	0.8	0.7	0.7	0.6	0.95	0.85	0.85	0.75	0.75	0.65	0.85	0.75	0.75	0.65	0.4	0.3	0.8	0.7	0.55	0.5	0.4	0.25
	large	0.1	0.2	0.2	0.3	0.3	0.4	0.05	0.15	0.15	0.25	0.25	0.35	0.15	0.25	0.25	0.35	0.6	0.7	0.2	0.3	0.45	0.5	0.6	0.75

Fig. 3b. Conditional probability table (CPT) for the attribute "TAKEOFF_ LANDING_CAPACITY" summarizes its dependency on the four parent nodes. For example, *Pr(Takeoff_landing_capacity=small | Parking_sq_ft=small, Elev_ft =low, Runway_length=short, Runway_width=narrow)=0.9*. The conditional probabilities in the CPT can be derived from the database content.

Schema link connects an attribute of the primary key to the corresponding attribute of the foreign key in the related entities. For example, in Figure 3a, APORT_NM is the primary key in AIRPORTS and foreign key of RUNWAYS. Therefore, we connect these two attributes via schema link.

Semantic link connects attributes with a specific semantic relation. To evaluate the inference introduced by semantic links, we need to compute the CPT for nodes connected by semantic links. Let T be the target node of the semantic link, P_S be the source node, and P_1, …, P_n be the other parents of T, as shown in Figure 4a. The semantic inference from a source node to a target node can be evaluated as follows.

If the semantic relation between the source and the target node is unknown or if the value of the source node is unknown, then the source and target node are independent. Thus, the semantic link between them does not help inference. To represent the case of the unknown semantic relationship, we need to introduce the attribute value "unknown" to the source node and set the value

of the source node to "unknown." In this case, the source and target node are independent, i.e., $Pr(T=t_i|P_1=v_1, \ldots P_n=v_n, P_S=unknown) = Pr(T=t_i|P_1=v_1, \ldots P_n=v_n)$. When the semantic relationship is known, the conditional probability of the target node is updated according to the semantic relationship and the value of the source node. If the value of the source node and the semantic relation are known, then $Pr(T=t_i| P_1= v_1, \ldots P_n= v_n, P_S=s_j)$ can be derived from the specific semantic relationship. For example, in Figure 4b, the semantic relationship determines that $Pr(T=t_1| P_1, \ldots P_n, P_S=s_1)=0.6$ and $Pr(T=t_1| P_1, \ldots P_n, P_S=s_2)=0.8$.

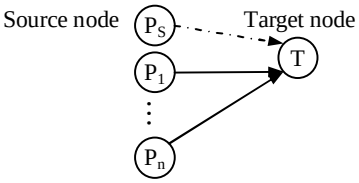


Fig. 4a. Target node T with semantic link from source node P_S and dependency links from parents P_1, …, P_n.

		Conditional Probability of T											
Cond	Ps	unknown				s1				s2			
	P1	v11		v12		v11		v12		v11		v12	
	Pn	vn1	vn2	vn1	vn2	vn1	vn2	vn1	vn2	vn1	vn2	vn1	vn2
T	t1	0.5	0.3	0.4	0.2	0.6	0.6	0.6	0.6	0.8	0.8	0.8	0.8
	t2	0.5	0.7	0.6	0.8	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0.2

Fig. 4b. The CPT of target node T summarizes the conditional probabilities of T given values of P_S and P_1, …, P_n. For example, $Pr(T=t_1 | P_S=unknown, P_1=v_{11}, P_n= v_{n1})=0.5$.

For example, the semantic relation "can land" between Runway and Aircraft (Figure 5a) implies that the length of Runway is greater than the minimum required Aircraft landing distance. So the source node is aircraft_min_land_dist, and the target node is runway_length. Both attributes can take three values: "short," "medium" and "long." First, we add value "unknown" to source node aircraft_min_land_dist and set it as a default value. Then we update the conditional probabilities of the target node to reflect the semantic relationship. Here, we assume that runway_length has an equal probability of being short, medium or long. When the source node is set to "unknown," the runway_length is independent of aircraft_min_land_dist; when the source node has a known value, the semantic relation "can land" requires runway_length is greater than or equal to aircraft_min_land_dist. Thus, the corresponding CPT for the node runway_length is shown in Figure 5b.

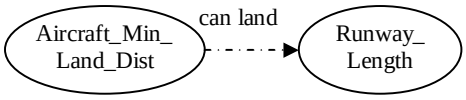


Fig. 5a. The semantic link "can land" between "Aircraft_Min_Land_Dist" and "Runway_Length"

	Conditional Probability of Runway_length				
Cond	aircraft_min_	unknown	short	medium	long
Runway_ Length	short	0.33	0.33	0	0
	medium	0.33	0.33	0.5	0
	long	0.33	0.33	0.5	1

Fig. 5b. Conditional Probability Table for Runway_length

5.1 Computation Complexity of Constructing Semantic Inference Model

A SIM consists of linking related attributes (structure) and their corresponding conditional probabilities (parameters). Given a relational database, the learning of a SIM can be decomposed into two tasks: parameter learning and structure learning. In the first task, we assume that the structure of the SIM is known, i.e., the links between attributes are fixed, and our goal is to derive the conditional probability tables for each attribute. Since the parameters of semantic link are determined by its semantic constraint, let us now consider the computation complexity on learning parameters of data dependencies. Consider that given structure S has m attributes, each attribute A_i in table T_j has a set of parents $P(A_i)$. If all parents of Ai are in the same table with Ai, then the CPT of Ai can be derived by a single scan of T_j. If attribute A_i has a parent from related entity table T_k, then scanning on the joined table of T_j and T_k is needed to derive the CPT of A_i. In the worst case, the parameters can be learned in $O(m\prod_i n_i)$ time, where m is the total number of attributes in the model and n_i is the size of the i^{th} table. When the number of dependency-between related-entities is limited, the parameter learning can be reduced to approximately $O(\sum_i m_i n_i)$ where m_i (< m) is the number of attributes in the i^{th} table. If the structure of the SIM is not given by domain experts, we can generate a set of candidate structures with their corresponding parameters, and select the one that best matches the data sources. Algorithms for searching good dependency structures can be found in [FGK99,GTK01].

5.2 Semantic Inference Graph

To perform inference at the instance level, we instantiate the SIM with specific entity instances

and generate a semantic inference graph (SIG), as shown in Figure 6. Each node in the SIG represents an attribute for a specific instance. To highlight the attributes of an entity instance, we group all the attributes of the instance into a rectangular box. Related attributes are then connected via instance-level dependency links, instance-level schema links and instance-level semantic links. The attribute nodes in SIG have the same CPT as in SIM because they are just instantiated versions of the attributes in entities. As a result, the SIG represents all the instance-level inference channels in the SIM.

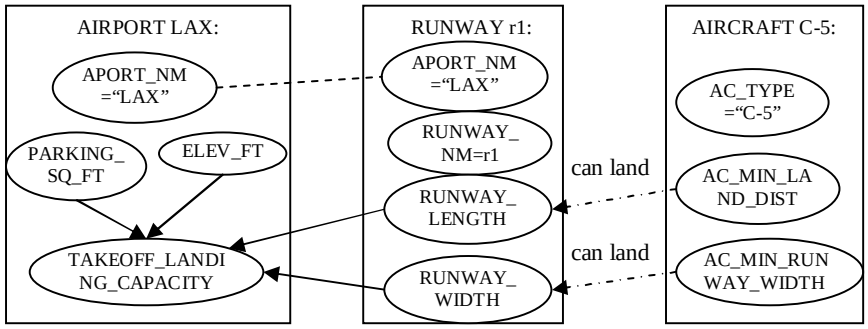


Fig. 6. The Semantic Inference Graph for airport instance (LAX), with runway r1 and aircraft C-5.

Instance-level dependency link: When a SIM is instantiated, the dependency-within-entity is transformed into dependency-within-instance in the SIG. Similarly, the dependency-between-related-entities in the SIM is transformed into a dependency between two attributes in the related instances. This type of dependency is preserved only if two instances are related by the instantiated schema link. That is, if attribute B in instance e_2 depends on attribute A in instance e_1, and instances e_1 and e_2 are related by R denoted as R(e_1, e_2), then there is a dependency-between-related-instances from B to A.

Instance-level schema link: The schema links between entities in the SIM represent "key, foreign-key" pairs. At instance level, if the value of the primary key of an instance e_1 is equal to the value of the corresponding foreign key in the other instance e_2 which can be represented as R(e_1, e_2), then connecting these two attributes will represent the schema link at the instance level.

Otherwise, these two attributes are not connected.

Instance-level semantic link: At the instance level, assigning the value of the source node to "unknown" disconnects the semantic link between the attributes of two instances. On the other hand, if two instances have a specific semantic relation, then the inference probability of the target node will be computed based on its CPT and the value of the source node.

5.3 Evaluating Inference in Semantic Inference Graph (SIG)

For a given SIG, there are attribute dependencies within an entity, between related entities, and semantic relationships among attributes. As a result, there are many feasible inference channels that can be formed via linking the set of dependent attributes. Therefore, we propose to map the SIG to a Bayesian network to reduce the computational complexity in evaluating user inference probability for the sensitive attributes.

For any given node in a Bayesian network, if the value of its parent node(s) is known, then the node is independent of all its non-descending nodes in the network [HMW95, Hec96, Jen96, Pea88, Pea01]. This independence condition greatly reduces the complexity in computing the joint probability of nodes in the network. More specifically, let x_i be the value of the node X_i, pa_i be the values of the parent nodes of X_i, then $P(x_i|pa_i)$ denotes the conditional probability of x_i given pa_i where i=1,2,…,n. Thus, the joint probability of the variables x_i is reduced to the product of P ($x_i|pa_i$):

$$P(x_1, \ldots, x_n) = \prod_i P(x_i \mid pa_i) \qquad (1)$$

The probability for users to infer the sensitive node $S=s$ given evidences $D_i=d_i$, $i=1, 2,\ldots, n$ is:

$$P(s \mid d_1, d_2 \ldots, d_n) = \frac{P(s, d_1, d_2 \ldots, d_n)}{P(d_1, d_2 \ldots, d_n)} \qquad (2)$$

which can be further computed using Equation (1). Thus, the probability of inferring a sensitive node can be computed from the conditional probabilities in the network. Many algorithms have been developed to efficiently perform such calculations [Dec96, JLO90, LS88, ZP96, ZP94].

The Probabilistic Relational Model (PRM) is an extension of the Bayesian network that integrates schema knowledge from relational data sources [FGK99, GTK01, GFK01]. Specifically, PRM utilizes a relational structure to develop *dependency-between-related-entities*. Therefore, in PRM an attribute can have two distinct types of parent-child dependencies: *dependency-within-entity* and *dependency-between-related-entities*, which match the two types of dependency links in the SIM. Since the semantic links in the SIM are similar to dependency links, we can convert each SIM to a PRM-based model. The corresponding Bayesian network can be generated after instantiating the model to instance level. Thus, for a given network, the probability of inferring a specific sensitive attribute can be evaluated via efficient Bayesian inference algorithms. In our test bed, we use SamIam [Sam], a comprehensive Bayesian network tool developed by the Automated Reasoning Group at UCLA, to compute the inference. The computation complexity for exact inference is mostly $O(n \cdot \exp(w))$, where n is number of nodes and w is the tree-width of the network [CD05, CAD05, Dar01, Dec99, JLO90, ZP94] and is scalable.

We have measured the elapse time for inference computation from our test bed. Since all the attribute nodes in the Bayesian network need to be re-evaluated, after posing each query, the time required for inference evaluation is almost constant[3]. For a sample Bayesian network with 40 nodes and 28 edges, the elapse time for inference evaluation after a single user poses a random query is around 16ms on a Dell desktop running Windows XP with 3.20GHz CPU and 2GB of RAM.

[3] For a sequence of queries, only a selected subset of all attribute nodes in the Bayesian network needs to be re-evaluated after each query, thus the computation time may be optimized for a large-sized network. The current implementation of the SamIam does not provide this optimization feature.

6. Inference Violation Detection for Individual User

Semantic inference graphs provide an integrated view of the relationships among data attributes, which can be used to detect inference violation for sensitive nodes. In such a graph, the values of the attributes are set according to the answers of the previous posted queries. Based on the list of queries and the user who posted those queries, the value of the inference will be modified accordingly. If the current query answer can infer the sensitive information greater than the prespecified threshold, then the request for accessing the query answer will be denied [CC06].

Consider the example in Figure 3. Let the TAKEOFF_LANDING_ CAPACITY of any airport be the sensitive attribute, and it should not be inferred with probability greater than 70%. If the user has known that: 1) Aircraft C-5 can land in airport LAX runway r1; 2) C-5 has aircraft_min_land_dist = long and aircraft_min_runway_width = wide. Then this user is able to infer the sensitive attribute "LAX's TAKEOFF_LANDING_ CAPACITY = large" via Equation (2) and (1) with probability 58.30%, as shown in Figure 7a.

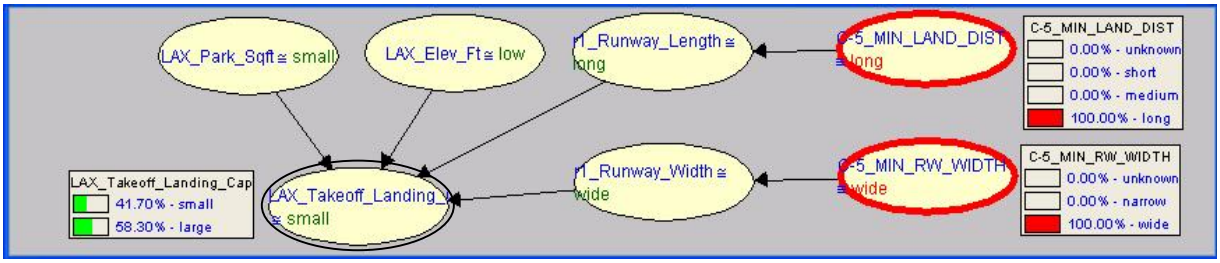


Fig. 7a. Example of inference violation detection for single user. This is a portion of the Bayesian network for the transportation mission planning. The probability distribution of each node is shown in a rectangular box. The values of the bold nodes are given by previous query answers; the probability values of sensitive nodes are inferred.

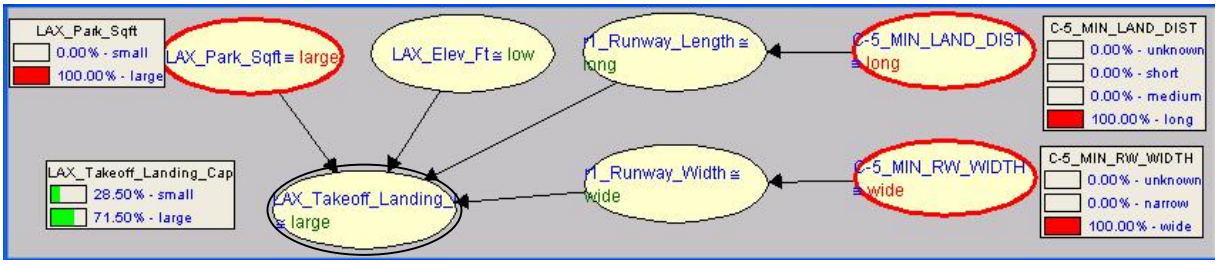


Fig. 7b. Given the additional knowledge "LAX_Parking_Sq_Ft=large", the probability for inferring the sensitive information "LAX_TAKEOFF_LANDING_ CAPACITY =large" is increased to 71.50%.

Now if the same user poses another query about the “Parking_sq_ft of LAX” and if this query is answered (as shown in Figure 7b, LAX_Parking_Sq_Ft=large), then the probability of inferring LAX_TAKEOFF_LANDING_CAPACITY = large by this user will increase to 71.50%, which is higher than the pre-specified threshold. Thus, this query request should be denied.

7. Inference Violation Detection for Collaborative Users

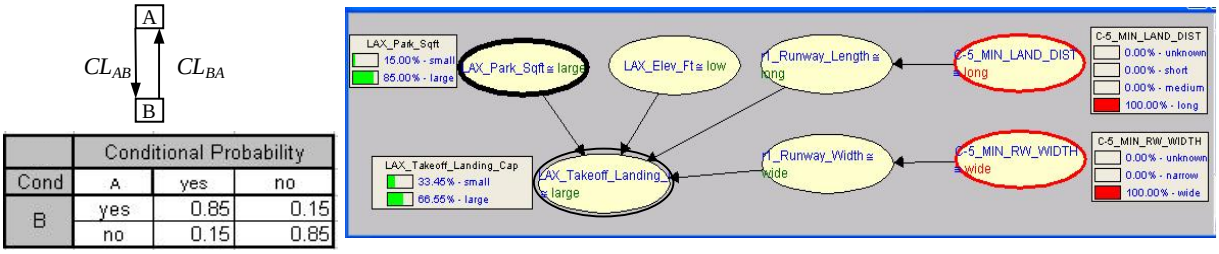
Generalizing from single-user to multi-user collaborative system greatly increases the complexity and present two challenges for building the inference detection system. First, we need to estimate the effectiveness of collaboration among users, which involves such factors as the authoritativeness of the collaborators, the communication mode among collaborators and the honesty of the collaboration. In addition, we need to properly integrate the knowledge from collaborators on the inference channels for the inference probability computation. In this section, we will address these two challenges.

7.1 Collaboration Effectiveness

We shall define *Collaboration Level* (*CL*) as a metric for measuring the percentage of useful information flow from the information source to the recipient. The range of Collaboration Level is from 0 to 1, CL=0 and CL=1 represent none and all of the information is received by the recipient.

Consider users A and B in Figure 8. User B has a collaborative level of 85% for the information from A. Let Q_A and Q_B be the query answer set of user A and user B. User B can combine Q_A with his own previous query answer set Q_B and yield a higher inference probability for the sensitive node. For the example in Figure 7a, user B has past query answers Q_B = {C5_min_land_dist = long, C-5_min_rw_width = wide} and then combines this with his acquired knowledge from user A: Q_A = {LAX_Park_Sqft = large}. Such collaboration increases the

inference probability for the sensitive node from 58.30% to 66.55%, as shown in Fig. 8. Note that because the collaborative level of B for information from A is 85%, it yields a lower inference probability than the case where user B queries directly about LAX_Parking_Sq_Ft, as in Fig. 7b.



	Conditional Probability		
Cond	A	yes	no
B	yes	0.85	0.15
	no	0.15	0.85

Fig. 8. Example of inference violation detection for multiple users. User B knows "C5_min_ land_dist=long" and "C5_min_rw_width=wide" from his past query answers. User B also has the knowledge from A "LAX_Park_Sqft =large" with collaborative level 85%. Thus, the probability for user B to infer the sensitive information (shown in double ellipses) "LAX_Takeoff_Landing_Capacity=large" increases to 66.55%.

By a series of experimental studies, we find that the collaboration level depends on three components: *Authoritativeness* of the information provider, A; *Honesty* of the collaboration, H; and *Fidelity* of the communication channel between the provider and recipient, F.

Authoritativeness of the information provider represents how accurate is the information. If a provider is knowledgeable and has high reputation in the field related with the task, then he/she can provide more accurate information.

Honesty represents the honesty level of the provider and the willingness of releasing his/her knowledge to the recipient. For example, if user A is very knowledgeable, in addition, A and B have a good communication channel; then both authoritativeness and fidelity of user A are high. However, A is not willing to release his full knowledge to B, as a result, the useful information cannot reach B for inference. Further, A can deceive B with false information. Thus, we shall use honesty measure as an indication of the honesty in collaboration.

Fidelity measures the effectiveness of the communication between the provider and recipient. Poor mode of communication can cause information loss during the transmission, which reduces

the effectiveness of the collaboration.

Authoritativeness measures how accurate the provider can supply information; honesty describes the willingness of the provider to release the accurate information; and fidelity measures the percentage of information transferred to the recipient due to the limitation of the communication mode. Once we estimate these three components for a set of users on a specific task, we can derive the collaboration level, as will be described in Section 8.3.

7.2 Combining Knowledge from Collaborators

In this section, we study the combination of knowledge from collaborators on different types of inference channels. Based on the users' query history, there are two different types of collaborative user pairs: *Collaboration with non-overlap inference channels* and *Collaboration with overlap inference channels*, as shown in Figure 9.

Collaboration with non-overlap inference channels: In this case, the two users pose queries on different non-overlap inference channels. The inference probability will be computed based on their combined knowledge discounted by their collaborative level.

For example, two collaborators A and B ask queries on non-overlap inference channels. In addition, the collaboration level from user A to user B is given by CL_{AB}, and the collaboration level from B to A is CL_{BA}. Therefore, to compute the inference probability to security attribute of user A, the query answers acquired by B (Q_B) can be combined with his/her own query answers (Q_A), but discounted by the collaborative level CL_{BA}. On the other hand, to derive the inference probability of user B, A's query answers (Q_A) are discounted by collaboration level CL_{AB} and then combined with Q_B. Because Q_A and Q_B are from independent non-overlap inference channels, their inferences to sensitive node S are independent and can be directly combined. Thus the inference probability for the sensitive node can be computed based on the user's knowledge from his past queries combined with his collaborator's query answers discounted by their respective

collaborative level.

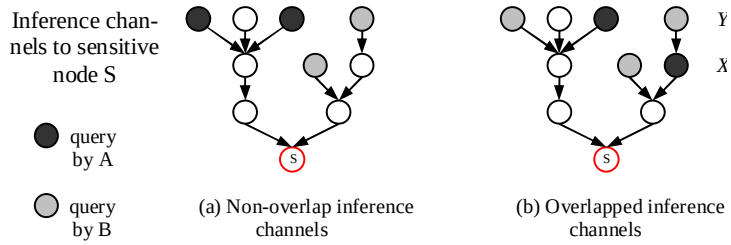


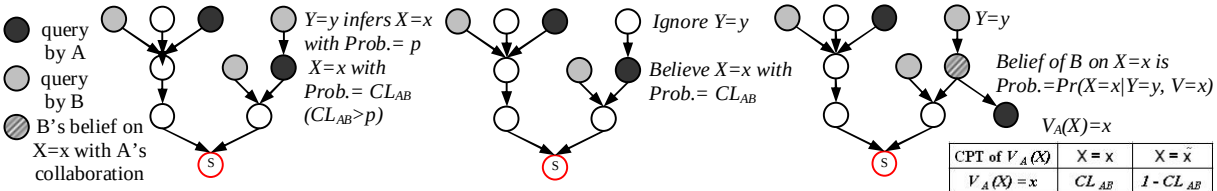
Fig. 9. Types of collaborative user pairs in the social network posing query sequence on the inference channels.

For example, two collaborators A and B ask queries on non-overlap inference channels. In addition, the collaboration level from user A to user B is given by CL_{AB}, and the collaboration level from B to A is CL_{BA}. Therefore, to compute the inference probability to security attribute of user A, the query answers acquired by B (Q_B) can be combined with his/her own query answers (Q_A), but discounted by the collaborative level CL_{BA}. On the other hand, to derive the inference probability of user B, A's query answers (Q_A) are discounted by collaboration level CL_{AB} and then combined with Q_B. Because Q_A and Q_B are from independent non-overlap inference channels, their inferences to sensitive node S are independent and can be directly combined. Thus the inference probability for the sensitive node can be computed based on the user's knowledge from his past queries combined with his collaborator's query answers discounted by their respective collaborative level.

Collaboration with overlap inference channels: In this case, the query sets posed by the two users overlap on inference channels. Such overlap may cause the users to have inconsistent belief in the same attribute on the inference channel. Thus, we need to integrate the overlapping knowledge according to the collaborative level to compute the appropriate inference probability.

For collaboration with overlap inference channels, the queries posed by user A and user B overlap on their inference channels. Since Q_A and Q_B may cause inconsistent belief on some attribute nodes, these two query answer sets cannot be simply combined. For example, in Figure

10(a), for attribute node X, Q_A indicates A has known $X=x$ and B can believe it with collaboration level CL_{AB} ($CL_{AB} \leq 1$). On the other hand, Q_B includes $Y=y$ which can infer $X=x$ with probability p. If $p \neq CL_{AB}$, then Q_A and Q_B can cause B to have inconsistent belief on attribute X. Without loss of generality, we assume $p < CL_{AB}$ for this example.



(a) Q_A and Q_B cause inconsistent belief on X. From $Q_B(Y=y)$, B infers $X=x$ with Prob.= p. From $Q_A(X=x)$, B infers $X=x$ with Prob.= CL_{AB}.

(b) Believe $X=x$ with Prob.= CL_{AB} yields inaccurate inference probability to S

(c) Introduce soft evidence for $Q_A(X=x)$ by adding a virtual node $V_A(X)$ as the child of X. User B now believes $X=x$ with $Prob.=Pr(X=x|Y=y, V_A(X)=x)$

Fig. 10. A virtual node can be used in user B's inference network to resolve inconsistent belief when user B and A overlap on their inference channels.

One approach to reconcile such inconsistent belief is to assume B will always choose to maximize his inference probability. Therefore, as shown in Figure 10(b), B only follows A's advice ($X=x$ with $prob.=CL_{AB}$) and ignore his own acquired knowledge ($Y=y$ infers $X=x$ with $prob.=p$). However, such a "max-inference" approach is not always correct, since people's belief is often strengthened by the confirmation and reduced by the conflicting knowledge.

To represent the integration of inconclusive belief, we introduce the concept of *soft evidence* in probability calculus [Dar03]. Soft evidence is inconclusive or unreliable information, as in the previous example, A tells B that $X=x$ and B only believes it with CL_{AB} ($CL_{AB} < 1$). For user B, $X=x$ is inconclusive knowledge, and therefore it needs to be set as soft evidence. To specify the soft evidence, we use the *Virtual Evidence* method developed in [Dar03]. As shown in the Figure 10(c), this method first adds a virtual attribute node $V_A(X)$ to be the child of the attribute node X to represent the event of receiving the soft evidence of X, that is, A tells B about $X=x$. Then the conditional probability of the virtual node is determined by the reliability of the soft evidence. In

our example, both $\Pr(V_A(X) = x \mid X = x)$ and $\Pr(V_A(X) = \bar{x} \mid X = \bar{x})$ are determined by user B's *collaboration level* of information from A: CL_{AB}. Thus, the soft evidence can be integrated into the user's own knowledge. In the example, if originally B is ignorant about X, once A tells B about $X=x$, B will believe $X=x$ with probability CL_{AB}. If originally B can infer X with knowledge $Y=y$, then his current belief in $X=x$ can be computed as $\Pr(X = x \mid Y = y, V_A(X) = x)$. Thus, we are able to integrate query answers on overlapped inference channels from multiple collaborators based on their corresponding collaboration levels.

7.3 An Example of Inference Violation Detection for Collaborative Users

A set of data sources for transportation and related facilities is available for mission planning. Due to the classified nature of the data sources, users can only access limited amounts of information. Malicious users want to identify whether a specific facility is capable of launching certain classified missions. However, the users are unable to access all the information that is required to derive the conclusion. Therefore, they apply inference techniques to infer the otherwise inaccessible information. In the following example, we shall demonstrate how our detection system prevents these users from accessing the relevant information.

As shown in Figure 11, the transportation and facility data sources consist of four types of information: 1) takeoff and landing activities and capacity of the airport, such as parking_area, runway_length, runway_width, aircraft landing requirements etc.; 2) equipment handling capacity, such as weapons, human experts, loading facility; 3) airport cargo and warehouse capacity and activities, such as daily cargo handling capacity, warehouse space; and 4) fueling storage and consumption. Based on these entities and attributes, we can derive the dependency links between attributes, the schema links that join different aspects of information together for each airport. Furthermore, based on the following set of semantic queries:

- Query1: which airports *can land* a C-5 cargo plane?
- Query2: which airports have the loading facility that *can load* weapon type HC-1?
- Query3: which aircraft *can carry* weapon type HC-1?

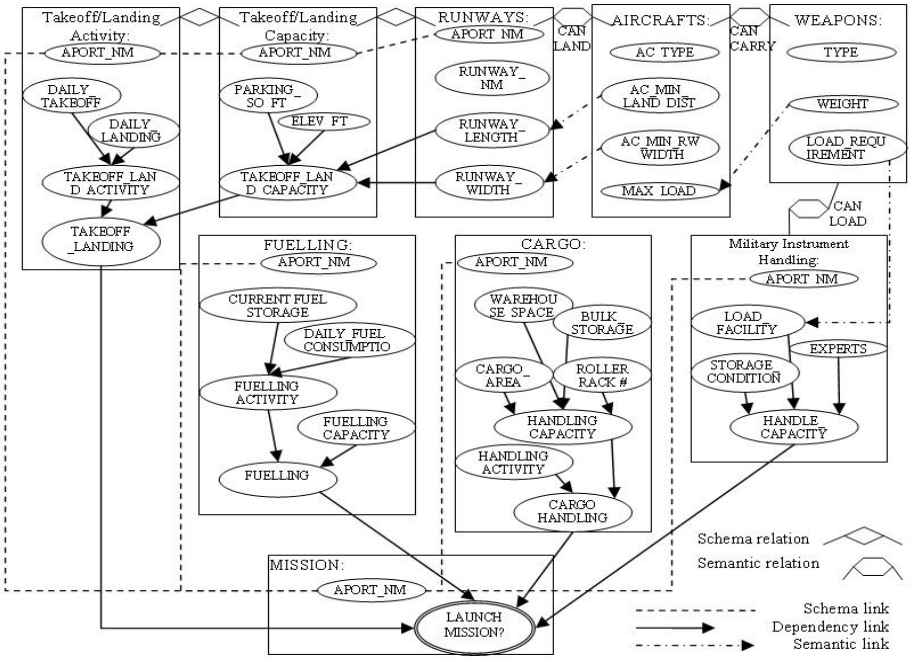


Fig. 11. The SIM for a transportation mission planning example

We can extract the semantic knowledge for "can land," "can load" and "can carry" for semantically linking the related attributes, as shown in Figure 11.

Based on these dependency links, schema links and semantic links, a reduced semantic inference model was constructed (Figure 11) to represent all the possible inference channels between data attributes for all the entities. There are four data sources which yield four main inference channels to the mission entity: takeoff_landing to launch_mission; fueling to launch_mission; cargo_handling to launch_mission and handle_capacity to launch_mission. Each of the main inference channels consists of many local inference channels. To carry out the inference computation, we need to generate a semantic inference graph (SIG) by substituting the specific instance to the semantic inference model. The corresponding Bayesian network representation mapped

from the SIG for airport “LAX” is shown in Figure 12.

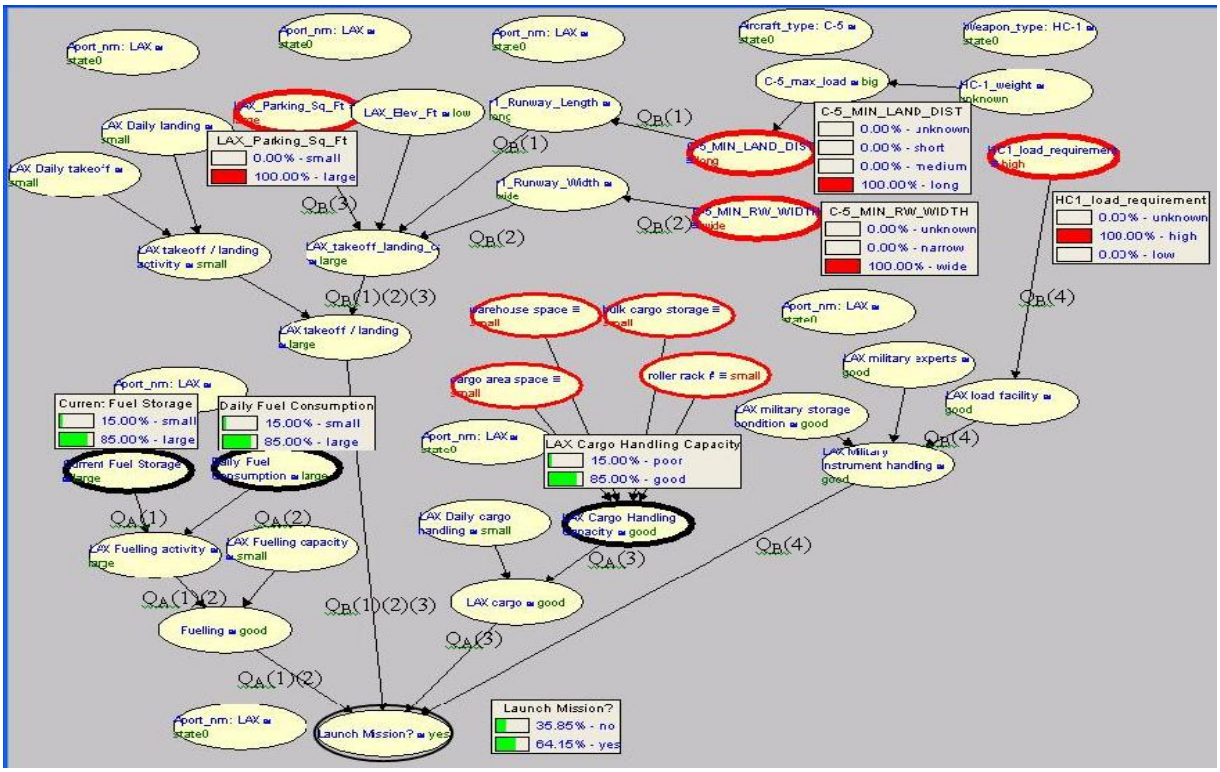


Fig. 12. The Bayesian network for the mission planning example. The bold nodes represent user queried attributes. Knowledge from the query answers can be accumulated along the inference channels towards the sensitive attribute. The inference channels used by each query are labeled by its query identifier. The collaborative level from user A of 85% is shown in the probability distribution boxes of $Q_A(1)$, $Q_A(2)$ and $Q_A(3)$. When all the seven queries are answered, user B can infer the sensitive attribute (shown in double ellipses) with a certainty of 64.15%.

Let “Launch Mission?” be the sensitive attribute. The violation detection module examines each user’s past query log, as well as the current query request. The probability to infer “Launch Mission?” in the Bayesian network will be evaluated before answering each query. If answering the current query increases the certainty of inferring the sensitive attribute above the pre-specified threshold, then the query will be denied. Let the pre-specified threshold for launch mission be 60%, and the users have prior knowledge of: 1) Aircraft C-5 can land in airport LAX; 2)Airport LAX can load weapon HC-1. When user A poses the sequence of queries shown in Table1, each query answer will update his certainty of inferring the “Launch Mission? = yes” (as shown in the table). The same is true for user B when he poses the queries in Table2.

Table 1. The inference probability of "Launch Mission? = yes" after answering user A's queries. The probabilities are computed from the Bayesian network in Figure 12.

Query Set of A $Q_A(i)$	$Answer_i$	Pr(Launch_mission? = yes | $answer_1$,…, $answer_i$)
(1) What is current_fuel_storage of airport LAX?	large	52.01%
(2) What is current_fuel_consumption of LAX?	large	56.50%
(3) What is cargo_handling_capacity of LAX?	good	59.80%

Table 2. The inference probability of "Launch Mission? = yes" after answering user B's queries. The probabilities are computed from the Bayesian network in Figure 12.

Query Set of B $Q_B(i)$	$Answer_i$	Pr(Launch_mission?=yes | $answer_1$,…, $answer_i$)
(1) What is the min_land_dist of aircraft C-5?	long	50.31%
(2) What is the min_rw_width of aircraft C-5?	wide	50.85%
(3) What is the parking_area_sq_ft of airport LAX?	large	52.15%
(4) What is the load_requirement of weapon type HC-1?	high	57.15%

Tables 1 and 2 are assuming that user A and user B do not collaborate. Neither A or B are getting enough information to infer the sensitive attribute above the threshold, thus all the queries are answered. However, based on the questionnaires collected from these two users, we notice that they are collaborators with an 85% collaborative level from B to A for this specific "airport mission" task. Therefore, the knowledge from their query answers can be combined for collaborative inference. If we examine their query set Q_A and Q_B on the SIM, we notice that they do not have overlapping inference channels. This is because Q_A focused on the fueling and cargo storage of the airport while Q_B focused on the takeoff and landing activities and military instrument handling. Thus, users A and B belong to the "non-overlap inference channels" case as shown in Figure 9. We can directly integrate their knowledge from query set answers based on their social relation. Thus user B can integrate Q_A into Q_B and adjust the inference probability using their respective collaborative level, as shown in Table3.

From Table3, we note that the last query posed by user B will infer sensitive information with probability higher than the pre-specified threshold of 60%. Therefore, $Q_B(4)$ should be denied by the violation detection module. In contrast, in the non-collaborative case as shown in Table1 and Table2, all the above queries can be answered.

Table 3. User B integrates user A's query set Q_A into his own query set Q_B. The Bayesian network is used to compute the inference probability in accordance with the posed query sequence and adjusted by the collaborative levels of the corresponding answers.

Integrated Query Set of B (i)	$Answer_i$	Collaborative Level t_i (%)	Pr(Launch_mission? = yes $\|t_1*answer_1,\ldots, t_i*answer_i$)
Q_B(1)What is min_land_dist of aircraft C-5?	long	100%	50.31%
Q_B(2)What is min_rw_width of aircraft C-5?	wide	100%	50.85%
Q_A(1)What is current_fuel_storage of LAX?	large	85%	52.39%
Q_A(2)What is current_fuel_consumption of LAX?	large	85%	55.54%
Q_B(3)What is parking_area_sq_ft of LAX?	large	100%	56.84%
Q_A(3)What is cargo_handling_capacity of LAX?	good	85%	59.15%
Q_B(4)What is load_requirement of weapon HC-1?	high	100%	64.15%

7.4 N-Collaborators

Therefore, for any two collaborative users, we can integrate one's knowledge to the other and detect their inference towards sensitive data. When any user poses a query, the system not only checks to see if the query requester can infer sensitive data above the threshold with a query answer, it also checks the other team members to guarantee that the query answer will not indirectly let them infer the sensitive attribute. We can iteratively generalize the above approach to an N-collaborator case. In general, when there are N collaborative users in the team, the violation detection system tracks the query posed by every team member. A query should be denied if the query answer will increase the certainty of any team member to infer the sensitive data above the pre-specified threshold.

```
1.  Assume: current query request Q, malicious team M, sensitive data S, threshold of S is T;
2.  List(M) = sort team members M in descending order of inference probability to S;
3.  While(List(M) is not empty) {
4.     m = first member in List(M) with highest inference probability;
5.     max_col = the maximum collaborative level from any member in List(M) to the query requester;
6.     real_col = m's collaborative level to query requester;
7.        If (m integrate answer to Q with max_col can get inference probability < T)
8.        Then {answer query Q; goto end;}
9.        Else
10.          If (m integrate answer to Q with real_col can get inference probability >= T)
11.          Then {deny query Q; goto end;}
12.          Else { List(M) = List(M) – m;}
13. }
```

An inference violation detection algorithm for N collaborative users

We can use the above greedy algorithm to efficiently decide to either answer or deny a query request from any team member. We first sort all N members by their inference probability to the

sensitive attribute and start with the member having the highest inference probability. We also compute every member's collaborative level to the query requester and determine the max collaborative level. Suppose that the member with the highest inference probability integrates the current query answer adjusted by the maximum collaborative level and still cannot infer sensitive data above the threshold. Then we can stop checking the rest of the team members and answer the query. This is because no other member in the team will be able to make a higher inference. If the member with the highest inference probability integrates the query answer adjusted by his collaborative level to the requester and can infer the sensitive data above or equal to the threshold, then we can stop checking and deny this query. Otherwise, we continue on to another member with the next highest inference probability until a decision can be made.

Since the system needs to evaluate the inference probability for every collaborator, the time required for inference evaluation increases as the number of collaborator increases. In our test bed on a sample Bayesian network with 40 nodes, after any user in a group of collaborators poses a random query, the time for inference evaluation ranges from 15ms for a single user to 281ms for five collaborators when their CL=1. The inference evaluation time almost doubles when the CL is less than one because the system requires extra computation to insert virtual nodes.

8. Collaboration Level

As stated in Section 7, information authoritativeness, honesty and communication channel fidelity are three components of the collaboration level metrics. In this section we shall first conduct a set of experiments to validate the premise of the proposed metrics and then propose to integrate these three components to estimate the collaboration level.

8.1 Experimental Study of Collaboration Level

Since authoritativeness, honesty and fidelity are user-sensitive, we used the students in one of the authors' classes as test subjects. The experiment was used as homework for the class. A web interface was developed for our inference test bed so that students could pose queries directly to the test bed and receive the answers.

Before posing queries for inference, each student needed to register in the system and fill in the necessary background information, including their age, gender, major, year in school, courses taken, GPA, skills, interests, teamwork ability, social activities, friends in the class, etc. The information gave us clues about the information authoritativeness and certain aspect of the fidelity of the test subjects.

8.1.1 Study the Effect of Authoritativeness and Fidelity

In the first experiment, our goal is to see how authoritativeness and fidelity affect the collaboration of multiple users. Based on the collected background information, we divided the class into five teams of four students to perform collaborative inference. The first team consisted of Ph.D. students with good knowledge in the database area, which should have good authoritativeness. The second team members were good friends, which provide good communication fidelity. The other three teams are randomly formed.

In the test 1a, the teams were given the SIG structure based on the database, the security attribute, but not the SIG parameters (CPTs) nor the threshold (75%) of the security attribute. Then we allowed each team to pose a fixed number of queries to infer the security attribute. The test bed computed their inference probability after each member posed the query. The system denied the query request if the posed queries exceeded the threshold. The four members in the team could collaborate in the best way possible to increase their inference probability of the security attribute. As shown in Figure 13 (1a), we observed that team2 reached the highest inference

probability. This is because they held meetings to discuss strategies of posing queries and voted if there was disagreement; therefore, their queries leveraged on each other to get better inference. This result reveals that communication fidelity plays a positive role in determining collaboration effectiveness.

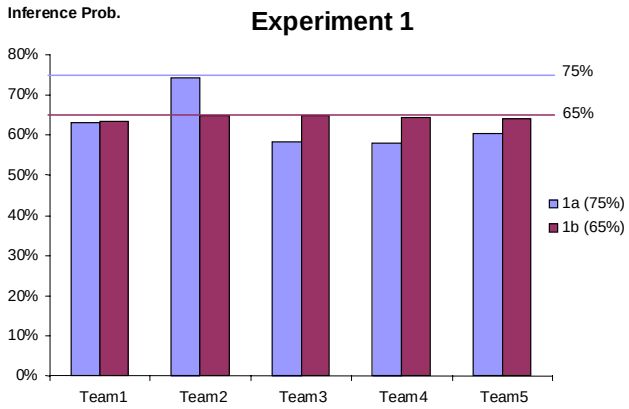


Fig. 13. The inference results for Experiment 1. In experiment 1a, the teams were given the SIG structure but without the parameters (CPTs) and the threshold (75%) of the security node. In experiment 1b, the teams were given both the SIG structure and the CPTs and the inference threshold (65%) of the sensitive node.

In the test 1b, we repeated the same experiment over these five teams. However, we let all the teams know the SIG structure, CPTs and the threshold value (65%) of the security attribute. As a result, this increased the authoritativeness of each team. With the same fixed number of queries, we noticed that with the additional knowledge of the CPTs and threshold of the security attribute, all the teams were able to ask better queries to improve their inference probability as close to the threshold as possible. Six queries were denied for four of the five teams due to the excess of the threshold during the process of this experiment. This experimental result reveals that the CL increases with authoritativeness.

In the second experiment, we investigated the collaboration effectiveness under controlled authoritativeness and communication fidelity. This experiment was carried out similar to experiment 1, except it was conducted in another graduate class in the following quarter. Because of the small class size, we divided the students into three teams, each having three members.

To control the authoritativeness of the players, we gave each team different SIG. The first team was given the full SIG with four inference channels leading to the security attribute; the second team received half of the SIG with two channels which can maximally infer the security attribute with probability 75%; and the third team only had knowledge of one inference channel in the SIG that can infer the security attribute up to 60% inference probability. Therefore, for this inference task, the first team had authoritativeness value 1; the second team had authoritativeness 0.75; and the authoritativeness of the third team was 0.6.

To study the effect of fidelity, we controlled the mode of communication in each team. The first team was allowed to have "full fidelity." Members were required to meet and discuss query strategies and exchange their query answers in making their selection of queries. The second team was allowed "limited fidelity" in which only member1 can tell member2 and member3 about his/her query answers, but member2 and member3 could not communicate with each other and with member1. The third team had "restricted fidelity" because only member1 is allowed to tell member2 the query answers, but member2 and member 3 are forbidden to talk to each other and to member1.

As shown in Figure 14, the inference result of team1 is higher than team2 and team2 is higher than team3. This set of experimental results confirm our premise that information authoritativeness and communication fidelity can both positively affect the collaboration performance, and therefore are two key components of the collaboration level.

Among the six queries by each team, team1 had two queries denied; team2 and team3 had one query denied. These denied queries would have been answered if the collaboration within each team is unknown and the team members were treated as separate individuals.

8.1.2 Study the Effect of Honesty in Collaboration

In the past two experiments, players in the same team are willing to release true knowledge to their teammates. The "honesty" for every information provider is one, i.e. totally willing to release true information. In the third experiment, we want to introduce a less honest scenario to test the effect of honesty to collaboration level. The third experiment was conducted based on the same three teams as in the second experiment.

A proxy is introduced to study the "honesty" in the experiment. Users pose queries and receive answers from the proxy. As a result, the player cannot directly exchange the query answers with his/her team members. Thus, the proxy becomes the communication conduit among team members and can alter the query answers to control the "honesty" in the experiment.

For example, in the three teams of this experiment, the proxy used three different level of honesty to transfer the query answers. For the first team, the proxy does not send the exact answer of the original query, but send the answer of the parent node of the original query in the inference channel. Thus it simulates the scenario of "unwilling collaboration" since the collaborators will have less knowledge for inference than the query requester himself. The loss of information can be measured as the inference probability from the "answered node" to the "original query node". For the first group, based on the tested Bayesian network, the average loss of information by answering "parent" node is 0.7418. For the third group, the proxy answers "grandparent" node of the original query, and the loss of information due to the less honesty can be measured as 0.6107. For the second group, the proxy gives the "opposite" answer of the original query node to the collaborators. For example, if the answer to the original query is "Fueling=good", then the proxy sends the opposite answer "Fueling=poor" to the collaborators. This simulates the scenario that the information source is "deceptive" and misleads the collaborators to believe in the opposite.

Since we use the same teams as in experiment 2, they have the same authoritativeness. Further, the proxy sends query answers directly to members in each team, the fidelity values of the three teams are equal to one.

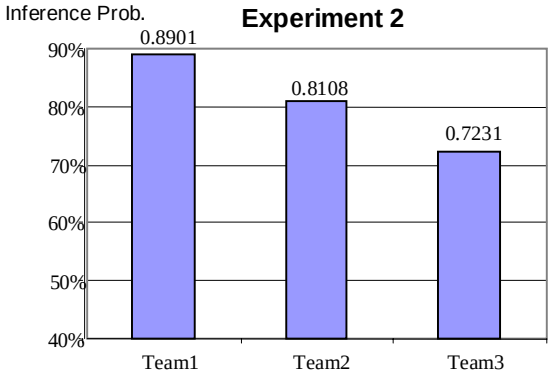


Fig. 14. Inference result for Experiment 2. The inference threshold of the security node was set at 90%.

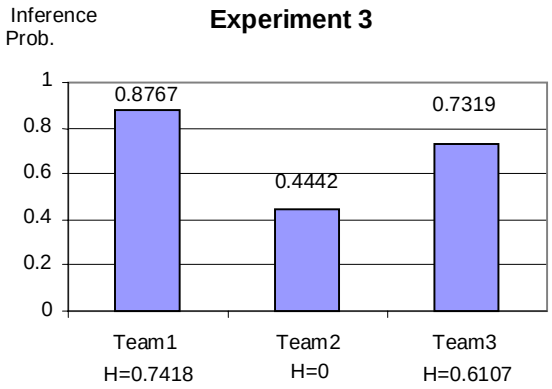


Fig.15. Inference result for Experiment 3. The inference threshold of the security node was set at 90%.

As shown in Figure 15, the first team has the highest "honesty in collaboration" among the three teams and yields the best inference result. The collaboration of the third team is less honest than team1; therefore their collaboration yields lower inference than team one. The members of the second team deceive each other, thus yields the lowest inference result. This experimental result reveals that the collaboration honesty does affect Collaboration Level. Less willingness to release information between collaborators reduces the collaboration effectiveness; and deception between collaborators causes negative effect in inference. Note H=1 represents honest collaboration which maximizes the collaboration effectiveness and can be used for conservative collaborative inference detection. Such over-protection may introduce false denial of legitimate queries.

8.2 Estimating Authoritativeness, Fidelity and Honesty

Estimation of Authoritativeness: Information authoritativeness represents the capability of the provider to supply accurate task-specific knowledge. This measurement can be derived based on the provider's profile, such as education level, profession, experience with the specific area of

the task, and reputation etc. For example, the answers to a carefully designed questionnaire at registration can be used as input for the estimation. We can take a weighted sum of the answers to related questions to estimate the authoritativeness of a user. In addition, the authoritativeness can also be enhanced by information derived from the relationship between users. If many individuals (especially highly authoritative ones) indicate user u_i as their reference in this field, then u_i has a significant impact on others and therefore has a higher authority. The link-based analysis (such as page rank) can be used to derive the reputation of people [PB98, KSG03]. In general, authoritativeness of the information provider can be derived from user profiles and/or questionnaire answers; when such background information is not available or incomplete, we can complement the authoritativeness from the provider's reputation among users.

Estimation of Fidelity: The fidelity measures the percentage of information sent by the provider that reaches the recipient side. Thus, fidelity depends on the quality of the communication channel and depends on the communication mode (e.g., face-to-face meeting, email or through a third party, etc.)

Estimation of Honesty: The honesty represents the willingness and truthfulness of the information release from the provider to the recipient. This is related with evaluation of trust in P2P networks [AD01, CDV02, KSG03, YW03a, YW03b, WL04, SBG05, XL04, DSC05 and MG06], which can be categorized by reputation-based or evidence-based approaches [LS07]. One approach is to use the reputation-based method as proposed in [YS02, YS03]. Honesty level is recipient-dependent. Therefore, for a given task, the honesty between two collaborators should also be estimated based on their closeness or friendship for a specific task. Therefore the specific honesty of a provider to a specific recipient needs to be adjusted by the closeness between the collaborators for a given task. Further research in this area is needed.

8.3 Estimating Collaboration Level from a Training Set

Since the collaboration level is user and task sensitive, we propose to use regression method to study the task and user specific relation between the Collaboration Level and its parameters. Specifically, for a group of users and a specific task, we can treat A, H, and F as the predictors and the Collaboration Level (CL) as the response variable. We can then learn the coefficients of these variables from the regressional model via the set of training data.

As an example, let the results of collaborative inference from experiment2 and experiment3 under controlled environment with selected A, H and F values (Table 4) be a training set. Since the inference result obtained by a team reflects the collaboration effectiveness under the corresponding controlled environment, we can normalize the inference result (i.e. inference result of the security attribute divided by the threshold) as the estimate of CL. Using the six entries as the input for regression analysis, the CL can be fit by multiple regression method with residual sum of squares 8.124×10^{-3} as shown in Table 4. Thus, we can estimate future collaboration level by substituting the parameters A, H and F into the regression model for similar users and tasks.

Table 4. Based on the six training data points, regression analysis can integrate the three components into Collaboration Level for future prediction.

	A	H	F	CL
From Experiment2 Team1:	1	1	1	0.989
From Experiment2 Team2:	0.75	1	0.33	0.901
From Experiment2 Team3:	0.6	1	0.17	0.8034
From Experiment3 Team1:	1	0.7418	1	0.9741
From Experiment3 Team2:	0.75	0	1	0.4936
From Experiment3 Team3:	0.6	0.6107	1	0.8132
Regression model: $CL = 0.1449 \cdot A + 0.4948 \cdot H + 0.1988 \cdot F + 0.2075$ Residual Sum of Squares: $rss = 8.124 \cdot 10^{-3}$				

9. Robustness in Inference Detection

Usually security experts or database administrators have some idea of the required level of protection for each security attribute, but they can hardly give a quantitative measurement to describe such protection requirements. Further, in a large database system, the dependency relationship between the security attribute and other attributes is complicated. The inference towards security attribute computed from a Bayesian network depends on both the network topology (qualitative attribute dependencies) and the parameter of the network (conditional probabilities). If a small variation of a given parameter can trigger the inference probability to exceed the threshold, then the inference detection may not satisfy the robustness requirements. This motivates us to find a methodology to systematically quantify the robustness of the threshold.

Sensitivity measures the impact of small changes in a network parameter on a target probability value or distribution [Las95]. In other words, a small change in the more sensitive attribute will cause a large impact on the inference probability. Therefore, the sensitivity values of attributes in the network provide an insight to the robustness of inference with respect to the changes in attribute parameter value. In this section, we propose to use the sensitivity analysis results to study the interrelationship between sensitive nodes and the security threshold.

9.1 Sensitivity Definition

"Sensitivity values are partial derivatives of output probabilities with respect to parameters being varied in the sensitivity analysis. They measure the impact of small changes in a network parameter on a target probability value or distribution" [Las95]. More formally, for a function *f(x)*, the quantity: $\lim_{(x-x_0)\to 0} \frac{(f(x)-f(x_0))/f(x_0)}{(x-x_0)/x_0}$ is typically known as the *sensitivity* of *f* to *x* at x_0, which is the ratio of relative change in output probability over the relative change in the parameter, where x_0 is the initial value of *X*. If we consider the function to be the probability of security

node *Y* given the change of attribute node *X*, then the sensitivity for attribute *X* at probability x_0 in a given network *N* with the initial probability of the security node y_{init} can be represented as:

$$\underset{x_0,N,y_{init}}{Sen}(X,Y) = \lim_{(x-x_0)\to 0}\left|\frac{(y-y_0)/y_0}{(x-x_0)/x_0}\right| = \lim_{\Delta x\to 0}\left|\frac{\Delta y/y_0}{\Delta x/x_0}\right| \qquad (3)$$

The initial probability of the security node is the probability of Y at the state when the set of evidence was given in the network. y_{init} represents the initial probability of *Y*, which is different from y_0 that represents the probability of *Y* when $X=x_0$. According to this definition, in a Bayesian network, if minor changes to an attribute node's probability can result in a significant change in the output probability of the security node, then this attribute node is considered highly sensitive.

9.2 Adjust Security Threshold by Attribute Sensitivity Analysis

To compute the sensitivity of attributes in an inference network, we first identify all inference channels toward each security node so that the sensitivity values for the attributes along the inference channels can be computed. The inference channels include channels coming into the security node and those going out of the security node. For those out-going inference channels, we can treat them as if the channels are coming into the security node by reversing the edges along such channels and revising the corresponding conditional probabilities. This is because, in terms of inference, the security node can be thought of as the "sink" of all information. Regardless of whether the attribute is the ancestor or descendent of the security node, the inference is always from the attribute towards the security node. Thus, we can compute the attribute sensitivities on both in-coming and out-going inference channels.

In a large-scale network, because of the large number of attributes, it is time-consuming to compute the sensitivity value for each attribute on the inference channels. However, for two attribute nodes on the same inference channel, the node that is closer to the security node is more sensitive than the node that is farther from the security node at the same probability value. This

difference of sensitivity value between closer and farther nodes is intuitive as closer nodes generally contain more sensitive information and are more influential on the security node than that of farther nodes. The farther node influences the security node through the inference channel which includes the closer node. Therefore, the amount of change at the farther node has the equivalent effect of inferring the security node as a smaller (or equal) amount of change at the closer node. For example, in the inference channel in Figure 15(a), the closest attribute to security node "*LaunchMission?*" is "*Fueling*." The sensitivity of "*Fueling*" is greater than sensitivity of its parents "*LAX_Fueling_Activity*" for all x_0, as shown in Figure 15(b). Similarly, the sensitivity of "*LAX_Fueling_Activity*" is greater than the sensitivity of "*Daily_Fuel_Consumption*."

Each value of the security node is protected by a threshold. For example, we need threshold for "Launch_Mission=Yes" and another threshold for "Launch_Mission=No" so that the malicious user cannot infer the exact value of this attribute above thresholds. When the data administrator proposes a threshold value based on the required protection level, he/she can check the sensitivity values of closest attributes on inference channels. If one of these inference channels is too sensitive which means a small change in the attribute value can resulted in exceeding the threshold, then the threshold needs to be tightened to make it less sensitive. In the case that the threshold cannot be further lowered to satisfy the sensitivity constraints, we can block the access to the closest attribute to the security node on the most sensitive inference channel, so that the accessible nodes on that inference channel are less sensitive to the threshold of the security node.

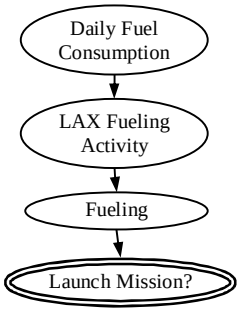


Fig. 15(a). A portion of an inference channel in the Bayesian network from the example.

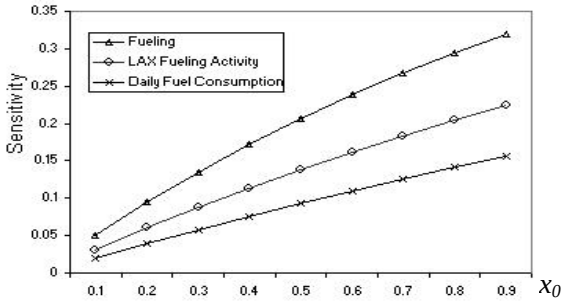


Fig. 15(b). The sensitivity of corresponding attribute nodes in 15(a) to the security node at selected initial values x_0.

10. Conclusion

In this paper we present a technique that prevents users from inferring sensitive information from a series of seemingly innocuous queries. Compared to the deterministic inference approach in previous works, we include non-deterministic relations into inference channels for query-time inference detection. Specifically, we extract possible inference channels from probabilistic data dependency, the database schema and the semantic knowledge and construct a semantic inference model (SIM). The SIM links represent the possible inference channels from any attribute to the set of pre-assigned sensitive attributes. The parameters of attributes in SIM can be computed in polynomial time in terms of the rows and columns of the relational table. The SIM is then instantiated by specific instances and reduced to a semantic inference graph (SIG) for inference violation detection at query time. To reduce computation complexity for inference, the SIG can be mapped into a Bayesian network so that available Bayesian network tools can be used for evaluating the inference probability along the inference channels. Therefore, our proposed approach can be scalable to large systems.

When a user poses a query, the detection system will examine his/her past query log and calculate the probability of inferring sensitive information from answering the posed query. The query request will be denied if it can infer sensitive information with probability exceeding the

pre-specified threshold. We find Bayesian network is able to preserve the structure of inference channels, which is very useful in providing accurate and scalable inference violation detection.

In the multiple-user inference environment, the users can share their query answers to collaboratively infer sensitive information. Collaborative inference is related to the collaboration level as well as the inference channels of the user-posed queries. For inference violation detection, we developed a collaborative inference model that combines the collaborators' query log sequences into inference channels to derive the collaborative inference of sensitive information.

Sensitivity analysis of attributes in the Bayesian network can be used to study the sensitivity of the inference channels. Our study reveals that the nodes closer to the security node have stronger inference effect on the security node. Thus sensitivity analysis of these close nodes can assist domain experts to specify the threshold of the security node to ensure its robustness.

User profiles and questionnaire data provide a good starting point for learning collaboration levels among collaborative users. However, gathering such information is complicated by the fact that the information may be incomplete and incorrect. In addition, the accuracy of such information is task-specific and user-community sensitive. We have constructed a testbed on the inference violation detection system to study the collaboration level for collaborative users. Our preliminary study reveals that information provider authoritativeness, communication fidelity and honesty in collaboration play key roles in the collaboration level. Further research and experiments in generating training set to estimate and validate collaboration level are needed.

Acknowledgements

The authors wish to thank the anonymous reviewers for their helpful comments and suggestions. These have led to valuable improvements to this paper.

References

[AD01] K. Aberer, Z. Despotovic, "Managing Trust in a Peer-2-Peer Information System", *Proceedings of the tenth international conference on Information and knowledge management*, October 05-10, 2001, Atlanta, Georgia, USA.

[CAD05] M. Chavira, D. Allen, and A. Darwiche. "Exploiting Evidence in Probabilistic Inference." In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 112-119, 2005.

[CD02a] H. Chan and A. Darwiche. "A Distance Measure for Bounding Probabilistic Belief Change." In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI)*, pages 539-545, Menlo Park, California, 2002. AAAI Press.

[CD02b] H. Chan and A. Darwiche. "When Do Numbers Really Matter?" *Journal of Artificial Intelligence Research*, 17:265-287, 2002.

[CD03] H. Chan and A. Darwiche. "Reasoning about bayesian network classifiers." In *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, pages 107-115, 2003.

[CD04] H. Chan and A. Darwiche. "Sensitivity analysis in Bayesian networks: From single to multiple parameters." In *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 67-75, Arlington, Virginia, 2004. AUAI Press.

[CDV02] F. Cornelli , E. Damiani , S. De Capitani di Vimercati , S. Paraboschi , P. Samarati, "Choosing reputable servents in a P2P network", *Proceedings of the 11th international conference on World Wide Web*, May 07-11, 2002, Honolulu, Hawaii, USA

[CD05] Mark Chavira and Adnan Darwiche. "Compiling bayesian networks with local structure." In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1306-1312, 2005.

[CC06] Yu Chen and Wesley W. Chu. "Database Security Protection via Inference Detection." *IEEE International Conference on Intelligence and Security Informatics*, 2006.

[CCH94] Wesley W. Chu, Qiming Chen and Andy Y. Hwang. "Query Answering via Cooperative Data Inference." *Journal of Intelligent Information Systems* (JIIS), Volume 3(1): 57-87, 1994.

[CYC96] Wesley W. Chu, Hua Yang, Kuorong Chiang, Michael Minock, Gladys Chow, and Chris Larson. "CoBase: A Scalable and Extensible Cooperative Information System." Journal of Intelligence Information Systems (JIIS). Vol 6, 1996, Kluwer Academic Publishers, Boston, Mass.

[Dat95] C. J. Date: An Introduction to Database Systems, 6th Edition. Addison-Wesley 1995.

[Dar01] Adnan Darwiche. "Recursive conditioning". *Artificial Intelligence*, 126(1-2):5-41, 2001.

[Dar03] Adnan Darwiche. Class notes for CS262A: Reasoning with Partial Beliefs, UCLA, 2003.

[DSC05] Duma, C., Shahmehri, N., Caronni, G., "Dynamic trust metrics for peer-to-peer systems", *Proceedings. Sixteenth International Workshop on Database and Expert Systems Applications, 776-781*, 2005.

[Dec96] Rina Dechter. "Bucket elimination: A unifying framework for probabilistic inference." In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 211-219, 1996.

[Dec99] R. Dechter. "Bucket elimination: A unifying framework for reasoning". *Artificial Intelligence*, 113:41-85, 1999.

[DH96] Harry S. Delugach and Thomas H. Hinke. "Wizard: A Database Inference Analysis and Detection System." In *IEEE Trans. Knowledge and Data Engeneering*, vol. 8, no. 1, 1996. pp. 56-66.

[FGK99] N. Friedman, L. Getoor, D. Koller and A. Pfeffer. "Learning Probabilistic Relational Models." *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI)*, Stockholm, Sweden, August 1999, pages 1300--1307.

[FJ02] C. Farkas and S. Jajodia "The Inference Problem: A Survey," *SIGKDD Explorations*, 4(2): 6-11, 2002.

[FTE01] C. Farkas, T. Toland, C. Eastman, "The Inference Problem and Updates in Relational Databases," *Proc. 15th IFIP WG11.3 Working Conference on Database and Application Security*, 181-194, 2001.

[GLQ92] T.D. Garvey, T.F. Lunt, X. Quain, and M. Stickel, "Toward a Tool to Detect and Eliminate Inference Problems in the Design of Multilevel Databases." *6th Annual IFIP WG 11.3 Working Conference on Data and Applications Security*, 1992.

[GTK01] L. Getoor, B. Taskar, and D. Koller. "Selectivity Estimation using Probabilistic Relational Models." *Proceedings of the ACM SIGMOD (Special Interest Group on Management of Data) Conference*, 2001.

[GFK01] L. Getoor, N. Friedman, D. Koller, and A. Pfeffer. "Learning Probabilistic Relational Models." Invited contribution to the book *Relational Data Mining*, S. Dzeroski and N. Lavrac, Eds., Springer-Verlag, 2001.

[HCL06] J. He, W.W. Chu, and Z. Liu. "Inferring Privacy Information From Social Networks." *IEEE International Conference on Intelligence and Security Informatics*, 2006.

[HMW95]Guest Editors: David Heckerman, Abe Mamdani, and Michael P. Wellman. “Real-world applications of Bayesian networks.” *Communications of the ACM*, 38(3):24-68, March 1995.

[Hec96] David Heckerman. “A Tutorial on Learning with Bayesian Networks.” Techinical report, Microsoft Research, 1996.

[HD92] Thomas H. Hinke and Harry S. Delugach. “Aerie: An Inference Modeling and Detection Approach for Databases.” *6th Annual IFIP WG 11.3 Working Conference on Data and Applications Security*, 1992.

[HDW96] Thomas H. Hinke, Harry S. Delugach, and Randall Wolf. “Wolf: A Framework for Inference-Directed Data Mining.” *10th Annual IFIP WG 11.3 Working Conference on Data and Applications Security*, 1996.

[Jen96] Finn V. Jensen. “An Introduction to Bayesian Networks.” Springer, New York, 1996.

[JLO90] F. V. Jensen, S.L. Lauritzen, and K.G. Olesen. “Bayesian updating in recursive graphical models by local computation.” *Computational Statistics Quarterly*, 4:269-282, 1990.

[KSG03] Sepandar D. Kamvar, Mario T. Schlosser, Hector Garcia-Molina, “The Eigentrust algorithm for reputation management in P2P networks”, *Proceedings of the 12th international conference on World Wide Web*, May 20-24, 2003, Budapest, Hungary

[KSS97] H. Kautz, B. Selman, and M. Shah. “The Hidden Web.” In AI magazine, 1997.

[Las95] Laskey, K. B. “Sensitivity Analysis for Probability Assessments in Bayesian Networks.” *IEEE Transactions on Systems, Man and Cybernetics*, 25, 909-909, 1995.

[LS88] S.L. Lauritzen and D.J. Spiegelhalter. “Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems (with Discussion).” *Journal of the Royal Statistical Society*, Series B, 50(2): 157-224, 1988.

[LS07] Huaizhi Li and M. Singhal, “Trust Management in Distributed Systems”, IEEE Computer, Vol 40. No 2, February 2007, pp. 45-53.

[LSC01] Wenke Lee, Salvatore J. Stolfo, Philip K. Chan, Eleazar Eskin, Wei Fan, Matthew Miller, Shlomo Hershkop and Junxin Zhang. “Real Time Data Mining-based Intrusion Detection.” *Proceedings of DISCEX II*. June 2001.

[MG06] Sergio Marti, Hector Garcia-Molina: Taxonomy of trust: Categorizing P2P reputation systems. *Computer Networks* 50(4): 472-484, 2006.

[PB98] Lawrence Page and Sergey Brin. “The anatomy of a large-scale hypertextual web search engine.” In *Proceedings of the Seventh International World-Wide Web Conference*, Brisbane, Australia, April 1998.

[Pea88] Judea Pearl. “Probabilistic Reasoning in Intelligence Systems.” Morgan Kaufmann, San Mateo, CA, 1988.

[Pea01] Judea Pearl. “Bayesian Networks, Causal Inference and Knowledge Discovery.” UCLA Cognitive Systems Laboratory, Technical Report (R-281), March 2001. In *Second Moment*, March 1, 2001.

[Sam] SamIam by Automated Reasoning Group, UCLA. http://reasoning.cs.ucla.edu/samiam/

[SBG05] Basit Shafiq, Elisa Bertino, Arif Ghafoor, “Access control management in a distributed environment supporting dynamic collaboration”, In Workshop On Digital Identity Management *Proceedings of the 2005 workshop on Digital identity management*, 2005.

[TFC93] Bhavani M. Thuraisingham, William Ford, M. Collins, and J. O'Keeffe. “Design and Implementation of a Database Inference Controller.” *Data Knowl. Eng.* 11(3), page 271, 1993.

[TFE05] Toland, C. Farkas, and C. Eastman, “Dynamic Disclosure Monitor (D^2Mon): An Improved Query Processing Solution,” *the Secure Data Management Workshop*, 2005.

[WL04] Winsborough, W., Li, N. “Safety in automated trust negotiation”, In *Proceedings of the IEEE Symposium on Security and Privacy*, 2004, 147--160.

[XL04] Xiong, L., and Liu, L. “Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities”. *IEEE Transactions on Knowledge and Data Engineering 16*, 7 , 2004.

[YL98] Raymond W. Yip, and Karl N. Levitt. “Data Level Inference Detection in Database Systems.” PCSFW*: Proceedings of the 11th Computer Security Foundations Workshop*, 1998.

[YS02] Bin Yu and Munindar P. Singh. “An Evidential Model of Distributed Reputation Management.” *Proceedings of the 1st International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS).* July 2002, 294 -301.

[YS03] Bin Yu and Munindar P. Singh. “Detecting Deception in Reputation Management.” *Proceedings of the 2nd International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, Melbourne, ACM Press, July 2003.

[YW03a] Ting Yu, Marianne Winslett, “A Unified Scheme for Resource Protection in Automated Trust Negotiation”, *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, p.110, May 11-14, 2003

[YW03b] Ting Yu, Marianne Winslett, "Policy migration for sensitive credentials in trust negotiation", *Proceedings of the 2003 ACM workshop on Privacy in the electronic society*, October 30, 2003, Washington, DC

[ZC99] Guogen Zhang, Wesley W. Chu, Frank Meng, Gladys Kong. "Query Formulation from High-Level Concepts for Relational Databases." *User Interfaces to Data Intensive Systems (UIDIS)* 1999: 64-75.

[ZP96] N. Zhang and D. Poole. "Exploiting Causal Independence in Bayesian Network Inference." Journal of Artificial Intelligence Research 5: 301—328, 1996.

[ZP94] Nevin Lianwen Zhang and David Poole. "A simple approach to bayesian network computations." In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 171-178, 1994.