Privacy-Preserving Accountable Accuracy Management Systems (PAAMS)

Roshan K. Thomas[1], Ravi Sandhu[2], Elisa Bertino[3], Budak Arpinar[4],
and Shouhuai Xu[5]

[1] SPARTA, Inc., 5875 Trinity Parkway, Suite 300, Centreville, VA 20120
[2] Institute for Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249
[3] Computer Science Department and CERIAS, Purdue University, West Lafayette IN 47907
[4] Department of Computer Science, 415 GSRC, University of Georgia Athens, GA 30602
[5] Department of Computer Science, University of Texas at San Antonio, TX 78249

Abstract. We argue for the design of "Privacy-preserving Accountable Accuracy Management Systems (PAAMS)". The designs of such systems recognize from the onset that accuracy, accountability, and privacy management are intertwined. As such, these systems have to dynamically manage the tradeoffs between these (often conflicting) objectives. For example, accuracy in such systems can be improved by providing better accountability links between structured and unstructured information. Further, accuracy may be enhanced if access to private information is allowed in controllable and accountable ways. Our proposed approach involves three key elements. First, a model to link unstructured information such as that found in email, image and document repositories with structured information such as that in traditional databases. Second, a model for accuracy management and entity disambiguation by proactively preventing, detecting and tracing errors in information bases. Third, a model to provide privacy-governed operation as accountability and accuracy are managed.

Keywords: privacy, accuracy, accountability, entity disambiguation.

1 Introduction

We present our initiative on "Privacy-preserving Accountable Accuracy Management Systems (PAAMS)" that aims to advance the accuracy management of intelligence information collected, produced and disseminated by government as well as commercial systems. Our approach is to develop unified proactive and reactive, both backward (trace to the source) and forward (trace to derivates) accuracy management techniques. Collectively, these will enable timely detection, identification and correction of errors in source information and finished intelligence while ensuring accountability and privacy preservation. Such information may span multiple repositories in varying formats including documents, email, and databases across multiple administrative domains. Our effort has direct and immediate application to well-known difficult problems such as entity disambiguation, as manifested in the compilation, merging and correction of terror watch lists.

C.S. Gal, P.B. Kantor, and M.E. Lesk (Eds.): ISIPS 2008, LNCS 5661, pp. 115–121, 2009.

Our approach is based on the recognition that accountability is a prerequisite for better accuracy management and further, that trading off privacy in controlled and accountable ways may yield significant increases in accuracy. Hence, the approach involves three key elements:

- A model to link unstructured information such as that found in email, image and document repositories with structured information such as that in traditional databases. Recognizing the need for this linkage, it forms the foundation for our development of a unified accountability and audit model that tracks the provenance, state changes and information flows across information bases.
- With the accountability model as an enabler, we propose to build a model for accuracy management and entity disambiguation to (1) proactively prevent errors and reactively detect errors in information bases and (2) to trace and correct the impact of such errors in source and target (derivate) information bases.
- Finally, both the accountability and accuracy models are governed by a privacy model that provides the principles and associated mechanisms for preserving privacy as accountability and accuracy are managed.

We are developing an architectural framework to realize these models and demonstrate their utility and viability.

2 Technical Approach

We believe that significant improvements in the accuracy of intelligence community (IC) information bases can be achieved *only* with a fundamental realignment of the information processing architecture – one that recognizes from the outset that accuracy, accountability and privacy are intertwined. Any comprehensive solution must manage the dependencies and tradeoffs between these elements. Our key assertion is that these tradeoffs cannot be hardwired but instead must be dynamically managed based on ongoing threats and acceptable individual tolerances for these elements as driven by application needs.

Current systems do not recognize this dynamic interrelationship and most existing research efforts have treated these elements in isolation. To illustrate the above, consider the Terrorist Identities Datamart Environment (TIDE) and associated compilation, dissemination and use of watch lists by intelligence and law enforcement agencies as reported in the popular press [1]. The high number of misidentifications and occurrences of repeated misidentification of the same individual indicate inadequate accuracy management. Public perception is that the system has poor accountability as it cannot quickly trace the source of errors with consequent slow and cumbersome redress procedures. Our claim is that accuracy can only be improved in such systems by addressing and improving the accountability and privacy management dynamics as they interrelate with accuracy. To elaborate, better provenance tracking and auditing of actions on content can help in the rapid location of sources of errors. Once errors are located, the error correction process can be significantly improved if better entity disambiguation is provided. However, this may call for access to more discriminating and potentially privacy-sensitive information about source intelligence as well as individuals.

The framework for our novel approach to dynamically navigating the accuracy-accountability-privacy tradeoff triad is to formulate this as variants of an optimization problem. For example, one variant would be to maximize accuracy subject to pre-specified limits on accountability overhead and privacy intrusion. In a different scenario an 80% accuracy with 70% statistical confidence may be acceptable as long as privacy exposure is kept below a specified threshold.

2.1 Better Accountability by Linking Unstructured and Structured Data

To improve accountability, it is necessary to link structured information such as watch lists stored in modern relational database systems with unstructured information that can provide a corpus of supporting documentation and evidence. The latter can consist of text files, images and web pages etc. A database management system tracks the audit history of a record or a field, but is mostly unaware of the accuracy of the stored data, the data sources that produced the data and the means to correct the accuracy of such data when an error is discovered. The modern approach to linking structured and unstructured data is based on the following three layers: (i) a set of information sources that need to be integrated in a common view, (ii) a semantic layer providing common interfaces and ontology for the information sources as well as a high-level language/notation through which the user can express queries on the semantic model; and (iii) a knowledge/information access layer which translates queries on the semantic model into the languages supported by the information sources.

We are developing a semantic model and specialized ontology centered around two key concepts that are at the core of realizing the vision of PAAMS - an *accountable knowledge unit* and an *accountable knowledge activity model.*

The Accountable Knowledge Unit (AKU)

From a semantic modeling standpoint, an AKU represents an identifiable and attributable piece of intelligence information, such as “Joe Smith should be on the suspect watch list.” The AKU goes beyond traditional database transactions and system level notions for structuring activities and updates to records and instead provides the basis to reason about the collection, production, and dissemination of intelligence “knowledge” pieces at a higher level. The AKU can thus link unstructured information with structured database records. This linking can happen in many combinations depending on how specific unstructured and structured information bases serve as sources, intermediary storage and sinks (derivates).

The Accountable Knowledge Activity Model (AKAM)

The heart of our novel accountability approach is the AKAM. This is an activity model that considers how information-related activities for knowledge units need to be organized and tracked across information bases. In particular, several AKUs may be dependent on each other. Thus the AKU which says “Joe Smith should be on the suspect watch-list” may be dependent on facts from another AKU which says: “Joe Smith was arrested in July 2005 with weapons possession at Atlanta airport.” These dependencies are essentially information flow, accountability and integrity dependencies across the information bases touched by AKUs. The AKAM keeps track of these dependencies. Thus if a source document or email is now considered to be suspect or false, the relevant and encompassing AKU and integrity dependency will point us to the dependent database records that need to be corrected or retracted. In this case, the DBMS has a backward-looking source-to-sink information flow dependency in that the DBMS records were derived from the source documents. The dependencies also exist in the forward direction when a DBMS becomes the source for subsequent documents or emails.

2.2 Improved Accuracy Management through Better Entity Disambiguation

Given the accountability substrate based on AKUs and the AKAM, we pursue better accuracy management through improved entity disambiguation techniques. Accuracy management in the context of correcting data errors has been proposed in the context of data quality [2]. Such an approach, known as record matching, consists of correcting data by comparison with other sources, presumably of better quality. Record matching has also been used for the purpose of integrating data. However, our goal in PAAMS for entity disambiguation is to go beyond record matching and combine this with ontology-based approaches.

Figure 1 provides an overview of the approach. Disambiguation utilizes background knowledge stored in the form of one or more populated ontologies. It does *not* rely solely on the existence of data items that can provide strong and what appears to be obvious evidence such as email addresses or affiliations. In addition, it also uses any available relationships that may be provided as evidence, as well as those from the ontology to provide clues in determining the correct entity.

Similar to what we have done in [3], the task of disambiguating entities can utilize the types of relationships that connect similar entities to determine if they are the same entity or different. For example, in the domain of computer science researchers, the affiliation of researcher is commonly used to indicate that the computer scientist “John Miller” we are talking about is the one at the University of Georgia, rather than the one at the University of Montana, with a large confidence level. In the same domain, other types of relationships can also be used such as publications, or research interests to disambiguate two entities. Of course the confidence level of the disambiguation process will greatly depend on the types of relationships used. For example, the “affiliation” might be a more accurate indication rather than “research interests”.

The facts which can be used to disambiguate entities can have varying “sensitivity levels” to limit their access and avoid privacy exposure. By “sensitivity level” we mean certain information can be classified as public knowledge, like address and telephone number, while others such as credit-rating may be considered more “private” or sensitive. Yet other information, such as medical history may be considered “very sensitive”. Attempts to regulate access to these types of documents have been undertaken. For example, [4] describes the research and prototyping of a system that takes an ontological approach and is primarily targeted for use by the intelligence community. The approach utilizes the notion of semantic associations and their discovery among a collection of heterogeneous documents.

The basic input to the entity disambiguation module (DM) in PAAMS is a request with input parameters that specify the desired accuracy, level of confidence, and the tolerable limits on accountability overhead and privacy-intrusion. The DM then attempts to find an answer to a feasibility or optimization problem. It may come back and give an answer for accuracy with some confidence but indicate that further accuracy improvements with higher confidence intervals are possible if, say, the privacy constraint can be relaxed. Thus the analyst or TSA agent (as an example) can have a series of interactions with PAAMS to refine the desired result, provided the needed level of access can be justified.

2.3 Privacy-Governed Operation

The third element in our technical approach provides for a privacy-governed operation so as to meaningfully tradeoff privacy with accuracy and accountability. Popular techniques to ensure accuracy and improve data quality require access to several data sources often containing personally identifying information and hold the potential for privacy breaches. To address such issues, privacy-preserving data matching techniques in the database context have been proposed [5, 6]; such techniques however have some major drawbacks. They use protocols based on secure set intersection [8] and their costs are prohibitive. They only perform exact matching [7, 8]. This is a major drawback when data across different sources have heterogeneous quality, as an exact match may not be very successful and thus will likely result in very few matches. In such cases, approximate matching techniques are the only viable approach.

To address such issues, we plan to explore a technique recently proposed in [7]. This technique is based on embedding the records to be matched in an Euclidean space, that is, a vector space having the Euclidean distance as norm and to perform the comparison in such a space. To ensure privacy, the SparseMap embedding method [8] is used. Preliminary experiments have shown that the approach is also very efficient for very large data sets. We will investigate a variety of issues such as how to deal with the problem of heterogeneous data schema to determine how the approach can be used for disambiguating entities, as well as how to perform record matching between structured data (like relational DB data) and unstructured data.

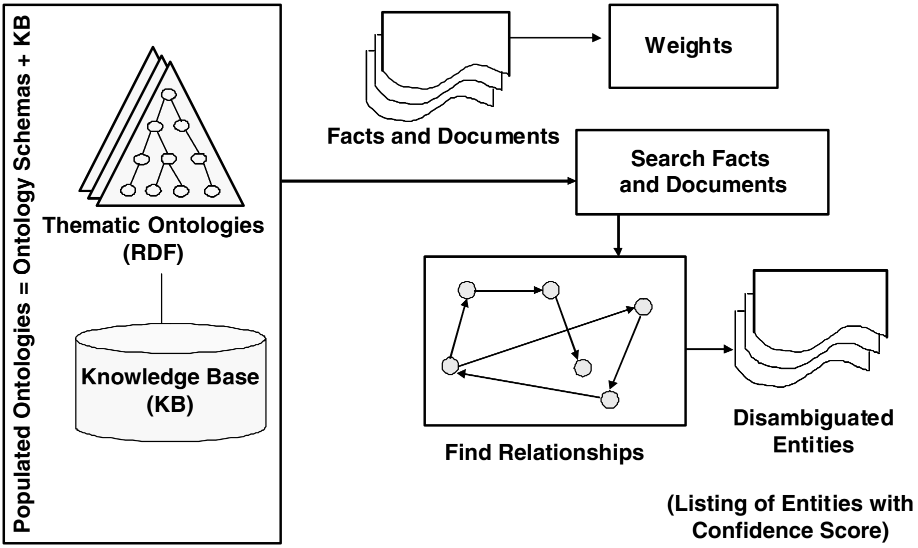
Populated Ontologies = Ontology Schemas + KB
Thematic Ontologies (RDF)
Knowledge Base (KB)
Facts and Documents
Weights
Search Facts and Documents
Find Relationships
Disambiguated Entities
(Listing of Entities with Confidence Score)

Fig. 1. Main parts of the entity disambiguation module

In essence, our goal is to explore how entity disambiguation techniques for both structured and unstructured information can be tuned to navigate privacy and accountability tradeoffs while seeking greater accuracy in information bases.

Metrics-Oriented Privacy Management

We are developing methods for "metrics oriented privacy management" based on the following ideas. First, different information items (i.e., records, fields, documents) may be annotated with different privacy sensitivity levels, possibly real numbers between zero and one. Second, the decision as to whether a computation process (such as a query) is allowed to succeed would depend on how much privacy information may be disclosed by the answer to the query. Such a leakage may be quantified through the differential in the private information entropies before and after conducting the query. If such a leakage violates a policy, the query or computation process may be cancelled. However, in some extreme cases such as when a TSA supervisor needs to make a decision whether to allow a passenger to board a flight, perhaps with the consent of the passenger, the supervisor may be given access to some sensitive information about the user. Schemes such as private-governed computation using cryptography-based [9, 10, 11] and statistics-based approaches [12, 13] can be leveraged for this purpose.

3 Summary and Conclusions

We have discussed a vision and approach to better accuracy management in information systems. This approach recognizes from the outset that accuracy, accountability, and privacy management are intertwined. Many systems such as those maintaining terror watch lists have had a difficult time maintaining accuracy and consistency. Our thesis is that accuracy in such systems can be improved by providing better accountability links between structured and unstructured information. Further, accuracy may be enhanced if access to private information was allowed in controllable and accountable ways. Thus some access to private information can lead to better entity disambiguation. We thus argue for a metric-based approach to privacy management. We lay out a framework for providing better accountability by developing the notion of an Accountable Knowledge Activity Model (AKAM) that ties together Accountable Knowledge Unit (AKUs). AKUs span structured and unstructured information. Collectively, we lay out a vision and the related architectural concepts to build systems that can provide improved accuracy through better accountability across islands of information and through controlled management of privacy to reduce errors and ambiguity.

References

1. http://www.washingtonpost.com/wp-dyn/content/article/2007/03/24/AR2007032400944.html
2. Batini, C., Scannapieco, M.: Data Quality: Concepts, Methodologies and Techniques. Springer, Heidelberg (2006)

3. Hassell, J., Aleman-Meza, B., Arpinar, I.B.: Ontology-Driven Automatic Entity Disambiguation in Unstructured Text. In: 5th International Semantic Web Conference (ISWC 2006), Athens, GA, USA, November 5-9 (2006)
4. Aleman-Meza, B., Burns, P., Eavenson, M., Palaniswami, D., Sheth, A.P.: An Ontological Approach to the Document Access Problem of Insider Threat. In: Proceedings of the IEEE Intl. Conference on Intelligence and Security Informatics (ISI-2005), May 19-20 (2005)
5. Agrawal, R., Evfimievski, A., Srikant, R.: Information Sharing Across Private Databases. In: SIGMOD (2003)
6. Naor, M., Pinkas, B.: Oblivious Transfer and Polynomial Evaluation. STOC (1999)
7. Scannapirco, M., Figotin, I., Bertino, E., Elmagarmid, A.: Privacy Preserving Schema and Data Matching. In: SIGMOD (2007)
8. Hjaltason, G.R., Samet, H.: Properties of Embedding Methods for Similarity Searching in Metric Spaces. IEEE TPAMI 25(5) (2003)
9. Xu, S., Yung, M.: K-anonymous Multi-party Secret Handshakes. In: Financial Cryptography and Data Security, FC 2007 (2007)
10. Tsudik, G., Xu, S.: A Flexible Framework for Secret Handshakes. In: Danezis, G., Golle, P. (eds.) PET 2006. LNCS, vol. 4258, pp. 295–315. Springer, Heidelberg (2006)
11. Xu, S., Yung, M.: K-Anonymous Secret Handshakes with Reusable Credentials. In: The Proceedings of ACM Conference on Computer and Communications Security 2004 (ACM CCS 2004), pp. 158–167. ACM Press, New York (2004)
12. Sharkey, P., Tian, H., Zhang, W., Xu, S.: Privacy-Preserving Data Mining Through Knowledge Model Sharing. In: Proceedings of the First ACM SIGKDD International Workshop on Privacy, Security, and Trust in KDD, ACM PinKDD 2007 (2007)
13. Dowd, J., Xu, S., Zhang, W.: Privacy-Preserving Decision Tree Mining Based on Random Substitutions. In: Müller, G. (ed.) ETRICS 2006. LNCS, vol. 3995, pp. 145–159. Springer, Heidelberg (2006)