Revisiting Ontology Based Access Control: The Case for Ontology Based Data Access

Ozgu Can[a] and Murat Osman Unalir[b]

Department of Computer Engineering, Ege University, 35100, Bornova-Izmir, Turkey

Keywords: Access Control, Data Access, Privacy, Ontology, Semantic Web, Knowledge Engineering.

Abstract: Ontology Based Data Access (OBDA) is a semantic paradigm to perform a mapping between an ontology and a data source for querying heterogeneous data sources. The result of this mapping ensures data access and data integration. Therefore, OBDA allows to query various datasets and provides data virtualization by integrating multiple and varied data sources. Ontology Based Access Control (OBAC) enables the realization of an access control mechanism by using Semantic Web technologies. OBAC allows to model the access control knowledge and uses domain knowledge to create policy ontologies. This paper revisits the OBAC approach by considering the OBDA to query legacy data that are stored in different types of data sources. For this purpose, OBAC is examined within the scope of OBDA and a conceptual model is proposed to extend OBAC with OBDA to provide data virtualization and to consolidate users' access privileges. Thus, security and management of complex information systems could be carried out by using Semantic Web technologies.

1 INTRODUCTION

In recent years, continuous improvements in information technology and the interconnectedness of systems led to the rapid growth in data volume. Moreover, this huge amount of data are continuously growing as data continues to come from many sources like social media, the web, sensors, devices and etc. In the information technology, bringing information together, extracting information, and using this extracted information to make strategic decisions are major challenges. In this context, one of the main problem is the lack of semantics in data representation.

In today's information systems, most of the data and information are stored in relational databases. The relational model is the de facto database system for data accumulation and data processing applications. In a relational database, data items and their relationships are organized as a set of tables with columns and rows. However, relational databases do not provide the semantic meaning of concepts. For this purpose, information should be represented with a conceptualized model and the relationship between concepts should be specified accurately (Haw et al., 2017). In Semantic Web, information is semantically annotated and the conceptualization of information is provided (Martinez-Cruz et al., 2012). The core of the Semantic Web is ontology. An ontology which is defined as "*a formal, explicit specification of a shared conceptualization* (Gruber, 1995)" is a semantic data schema paradigm and provides the conceptualization. An ontology defines concepts and relationships between these concepts that are used to describe a domain. Therefore, information is represented in a machine-readable form and it can be shared, reused, distributed, and used to make deductions.

The mapping between a database and an ontology is considered as a case of data integration (Haw et al., 2017). In the mapping process, concepts of the ontology are linked with the relevant entities in the database. As most of the data and information are stored and represented in databases, the mapping of legacy relational data to ontology concepts enables to retrieve more enriched query results and provide effective data analytics. The mapping between a relational database and Semantic Web is as follows: a record is an RDF node, the field (column) name is RDF property type, and the record field (table cell) is a value (Berners-Lee, 2009). Ontology-Based Data Access (OBDA) concerns answering queries over the target ontology by using a source relational database,

[a] https://orcid.org/0000-0002-8064-2905
[b] https://orcid.org/0000-0003-4531-0566

Can, O. and Unalir, M.
Revisiting Ontology Based Access Control: The Case for Ontology Based Data Access.
DOI: 10.5220/0010898100003120
In *Proceedings of the 8th International Conference on Information Systems Security and Privacy (ICISSP 2022)*, pages 515-518
ISBN: 978-989-758-553-1; ISSN: 2184-4356
Copyright © 2022 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

a target OWL ontology and a mapping from the relational database to the ontology (Sequeda, 2017). Thus, OBDA presents a conceptual representation of a domain and achieves data virtualization which integrates data without moving and transforming them (Xiao et al., 2019). As data virtualization enables a centralized point of access, it brings security requirements such as privacy, access control, authentication, authorization, data integrity, and effective mechanisms to ensure the security and the privacy of the data. Ontology Based Access Control (OBAC) aims to provide an access control mechanism for the security requirements in Semantic Web. OBAC allows to create, modify and query semantically-rich policies (Can et al., 2010; Can, 2009; Can and Unalir, 2010). Therefore, ontology based policies are specified over domain knowledge and access to information is achieved by authorized entities.

In this work, the OBAC model is enhanced with OBDA. The aim is ensuring security and preserving privacy while providing an efficient processing of data that exists in different heterogeneous sources. For this purpose, the OBAC model is revisited with the OBDA approach. In the proposed revised model: (i) OBDA provides the abstraction of how data sources are maintained in the data layer of the system itself (Poggi and et al., 2008), and (ii) OBAC provides the security and privacy of data by preventing unauthorized access requests. Therefore, data virtualization will be provided by achieving access control.

The remainder of this paper is organized as follows: in Section 2, the recent studies in the field of ontology-based data access and ontology-based access control are presented, the proposed conceptual model is detailed in Section 3. Finally, Section 4 concludes the paper and summarizes the future directions of the presented study.

2 RELATED WORK

The relation between database and Semantic Web is a frequently studied impressive topic in the literature. In order to access the existing data sources flexibly and efficiently, databases are mapped to ontology representations. In (Haw et al., 2017), steps to transform relational databases to ontology representation are outlined and a review of some of the mapping tools is presented by highlighting their requirements. A method is proposed in (Dadjoo and Kheirkhah, 2015) for automatic ontology construction based on a relational database. The presented method generates an ontology data model from the relational database schema. The relationship between relational databases and the Semantic Web is investigated in (Sequeda, 2017). In this study, the specific research question that is tried to be answered is "*How and to what extent can Relational Databases be integrated with the Semantic Web?*". A survey is presented in (Spanos et al., 2012) to review methods and tools that bring relational databases into Semantic Web. Moreover, the survey study also explores the future perspectives of the field. Ontology Based Data Access (OBDA) is a prominent approach to establish a mapping between a database and an ontology. Thus, it simplifies the process of data access and enhances the quality of query results. In (Kharlamov et al., 2017), data access challenges in the petroleum company Statoil are presented, and an OBDA based solution is developed. Similar to this study, OBDA is applied to the energy technology database within the technology forecasting information system in (Mikheev, 2018). In (Hoehndorf et al., 2015), a framework named Aber-OWL is developed to provide reasoning services for bio-ontologies by enabling ontology-based semantic access to biological data. The developed reasoning infrastructure uses OBDA to access information. In an EU FP7-funded project named Optique (Kharlamov et al., 2013; Giese et al., 2013), an end-to-end OBDA system is developed to provide scalable end-user access to industrial Big Data stores. The project focuses on two use cases: the first use case is provided by Siemens and the second use case is provided by Statoil. In the Semantic Web, access control is a challenging problem and access to resources should be controlled to secure the Semantic Web. An access control mechanism allows to define, manage and enforce access conditions for resources. In (He et al., 2010), the Role-Based Access Control (RBAC) model is extended to implement an access control mechanism for Semantic Web services. A Semantic Based Access Control model (SBAC) is presented in (Javanmardi et al., 2006) to authenticate users based on their credentials when requesting an access right. In (Kagal et al., 2003), a policy language and a security framework based on this language to address security issues in Semantic Web are presented. An Ontology Based Access Control (OBAC) model is proposed in (Can et al., 2010; Can, 2009; Can and Unalir, 2010) to define and enforce semantically rich access control policies. The OBAC models both the requestor and the requested by using the Rei policy language (Kagal et al., 2003).

In this work, the goal is to revisit the OBAC model with the concepts of OBDA to improve security and to preserve privacy while providing data virtualization. To the best of our knowledge, this paper is the first one that proposes a privacy framework for OBDA that is based on a fully semantic ontology based access control model. The proposed model addresses the issues arising from the security and privacy needs of the OBDA approach.

3 REVISITING OBAC FOR OBDA

In this work, the OBAC model is revised within the scope of the OBDA approach. OBAC is an access control mechanism that is used to secure Semantic Web based applications. OBDA allows querying a database that uses an ontology to expose data by abstracting away from the technical schema-level details of the underlying data (Kharlamov et al., 2017). In OBDA, domain knowledge is represented in the form of an ontology, and data virtualization is achieved through a mapping between the domain ontology and the data sources (Poggi and et al., 2008). When the desired mapping is established, the user can execute queries on the ontology and retrieve data from the mapped database (Spanos et al., 2012). Hence, OBDA presents a conceptual view of data ad the ontology acts as a mediator between the user and the data. The aim of revisiting the OBAC model within the scope of OBDA is to provide a privacy framework so that ontology-based data access can be performed in a privacy-aware manner. For this purpose, OBDA will preserve privacy based on the OBAC model. Thus, an access control mechanism will be enforced for the OBDA. Moreover, access control needs to be managed on different data models due to polyglot persistency. As the legacy data can be represented in an ontology, relational database, or non-relational database, queries should be translated into a form that will be understood by the legacy data. This process will be achieved by mapping. Also, RBAC mechanism will be integrated into the model. In the RBAC model, permissions are given directly to roles, not to the user. In the scope of this study, a query on a relational database should be transformed from RBAC to OBAC. Thus, each data source will be queried in its environment. Therefore, the RBAC model will be mapped to the OBAC model where data is represented semantically. The overall architecture of the proposed model is given in Fig. 1.

The proposed model is based on a materialization-based approach (forward chaining). In the materialization-base approach (Sequeda, 2017), the input is the database D, the target ontology is O and the mapping from D to O is M. The legacy data source is the ABox (A) and the ontology is the TBox (T). The SPARQL query Q is executed over the D, O, and M. The OBAC model is based on Rei policy ontologies (Kagal et al., 2003; Kagal, 2002). In the OBAC model, a Permission *Per* denotes what an entity can do, a Prohibition *Pro* states what an entity can't do, an Obligation *Obl* is what an entity should do and a Dispensation *Dis* indicates what an entity need no longer do. Along with the OBAC, access to the underlying data sources is abstracted independent of the mapping. The mapping between a database and an ontology will be achieved by Ontop (Ontop, 2009) which is an open-source OBDA framework. Also, Ontop is a query transformation module. Therefore, queries will be executed by using the Ontop framework. The proposed model will adhere to the following bottom-up architecture: (i) establishing an integrated framework model for access control and privacy, (ii) evaluation of personal and organizational identities within the scope of the multi-tenant model, (iii) conversion of RBAC entities to OBAC entities for relational data sources, (iv) realization of an access control independent of the data model, (v) comparative testing of query execution in both proactive and reactive structure, (vi) extending OBAC into a privacy-aware structure, (vii) implementing the proposed framework with Ontop, and (viii) evaluation of the proposed framework for a specific domain.

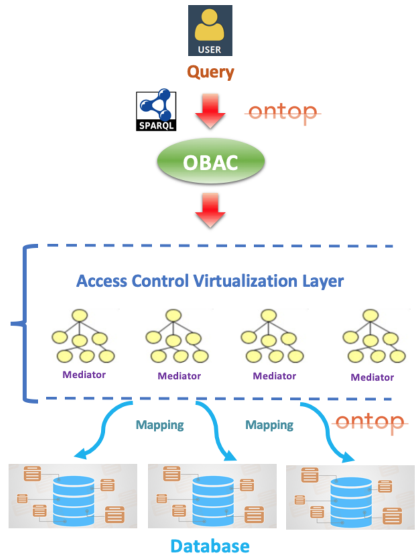

Figure 1: Architecture of the proposed model.

4 CONCLUSIONS

In today's conditions, most of the daily routines are dependent on information systems. Thus, large amounts of information are produced and stored by these systems. Most of these data are stored and represented in relational databases. In order to extract semantic information from a database, inference it and obtain valuable information, the database needs to be converted to the knowledge base (Dadjoo and Kheirkhah, 2015). Therefore, the mapping between databases and ontologies should be maintained to execute semantic queries and to discover new relationships by inference. Thus, the quality of data integration will be improved. On the other hand, the security and privacy of systems must be maintained. This work proposes a Semantic Web based model to improve the security and privacy of systems that may arise when applying the OBDA approach. In this work, a conceptual model for the proposed model is presented. As future work, the OBAC ontologies will be extended and mapped with the concepts of the RBAC model. As OBAC is an access control model, it will be also extended to preserve privacy. The desired mappings will be established and queries will be executed by using the Ontop framework.

ACKNOWLEDGEMENTS

This study is supported by Ege University Scientific Research Projects Committee under the grant number 18-MUH-036.

REFERENCES

Berners-Lee, T. (2009). Relational databases on the semantic web. https://www.w3.org/DesignIssues/RDB-RDF.html. Online; accessed 05 November 2021.

Can, O. (2009). *Personalizable Ontology Based Access Control for Semantic Web and Policy Management*. Phd thesis, Ege University, Department of Computer Engineering.

Can, O., Bursa, O., and Unalir, M.-O. (2010). Personalizable ontology based access control. *Gazi University Journal of Science*, 23(4):465–474.

Can, O. and Unalir, M.-O. (2010). Ontology based access control. *Pamukkale University Journal of Engineering Sciences*, 16(2):197–206.

Dadjoo, M. and Kheirkhah, E. (2015). An approach for transforming of relational databases to owl ontology. *Int. Journal of Web and Semantic Technology*, 6(1):19–28.

Giese, M. et al. (2013). *Big Data Computing*. Chapman and Hall/CRC, New York, 1st edition.

Gruber, T.-R. (1995). Towards principles for the design of ontologies used for knowledge sharing. *Int. Journal of Human-Computer Studies*, 43:907–928.

Haw, S.-C., May, J.-W., and Subramaniam, S. (2017). Mapping relational databases to ontology representation: A review. In *ICDTE'17*, pages 54–55.

He, Z. et al. (2010). Using semantic web techniques to implement access control for web service. In *ICICA 2010*, volume 105, pages 258–266. Springer.

Hoehndorf, R., Slater, L., Schofield, P.-N., et al. (2015). Aber-owl: A framework for ontology-based data access in biology. *BMC Bioinformatics*, 16(26).

Javanmardi, S., Amini, M., and Jalili, R. (2006). An access control model for protecting semantic web resources. In *SWPW'06*.

Kagal, L. (2002). Rei: A policy language for the me-centric project. Techreport.

Kagal, L., Finin, T., and Joshi, A. (2003). A policy based approach to security for the semantic web. In *ISWC 2003*, volume 2870, pages 402–418. Springer.

Kharlamov, E. et al. (2013). Optique: Towards obda systems for industry. In *The Semantic Web: ESWC 2013 Satellite Events*, volume 7955. Springer.

Kharlamov, E., Hovland, D., Jimenez-Rui, E., et al. (2017). Ontology based data access in statoil. *Journal of Web Semantics*, 44:3–36.

Martinez-Cruz, C., Blanco, I.-J., and Vila, M.-A. (2012). Ontologies versus relational databases: are they so different? a comparison. *Artificial Intelligence Review*, 38:271–290.

Mikheev, A.-V. (2018). Ontology-based data access for energy technology forecasting. In *IWCI 2018*, volume 158.

Ontop (2009). https://ontop-vkg.org. Online; accessed 05 November 2021.

Poggi, A. and et al. (2008). Linking data to ontologies. *Journal on Data Semantics X.*, 4900:133–173.

Sequeda, J.-F. (2017). Integrating relational databases with the semantic web: A reflection. In *Reasoning Web 2017*, volume 10370, pages 68–120. Springer.

Spanos, D.-E., Stavrou, P., and Mitrou, N. (2012). Bringing relational databases into the semantic web: A survey. *Semantic Web*, 3(2):169–209.

Xiao, G., Ding, L., Cogrel, B., and Calvanese, D. (2019). Virtual knowledge graphs: An overview of systems and use cases. *Data Intelligence*, 1(3):201–223.