Enhancing Information Leakage in Multi-Cloud Storage Services

Sanket Aher[1], Rohit Raut[2], Akash Tidke[3], Shubham Vanarase[4], Prof. H. R. Agashe[5]

Students, Department of Information Technology[1,2,3,4]

Professor, Department of Information Technology[5]

Matoshri College of Engineering and Research Center, Nashik, Maharashtra, India

Abstract: *Data leakage is a growing insider threat in information security among organizations and individuals. A series of methods have been developed to address the problem of data leakage prevention (DLP). However, large amounts of unstructured data need to be tested in the Big Data era. As the volume of data grows dramatically and the forms of data become much complicated, it is a new challenge for DLP to deal with large amounts of transformed data. We propose an Adaptive Weighted Graph Walk model (AGW) to solve this problem by mapping it to the dimension of weighted graphs. Our approach solves this problem in three steps. First, the adaptive weighted graphs are built to quantify the sensitivity of tested data based on its context. Then, the improved label propagation is used to enhance the scalability for fresh data. Finally, a low-complexity score walk algorithm is proposed to determine the ultimate sensitivity. Experimental results show that the proposed method can detect leaks of transformed or fresh data fast and efficiently.*

Keywords: Data Leakage, Big data, Adaptive Weighed Graph, Low-complexity, etc..

REFERENCES

[1]. R. Agrawal and J. Kiernan, "Watermarking Relational Databases,"Proc. 28th Int'l Conf. Very Large Data Bases (VLDB '02), VLDBEndowment, pp. 155-166, 2002. IEEE Transaction and knowledge and data engineering, Vol.23, No.1, January 2011.

[2]. P. Bonatti, S.D.C. di Vimercati, and P. Samarati, "An Algebra for Composing Access Control Policies," ACM Trans. Information and System Security, vol. 5, no. 1, pp. 1-35, 2002.

[3]. P. Buneman, S. Khanna, and W.C. Tan, "Why and Where: Characterization of Data Provenance," Proc. Eighth Int'l Conf. Database Theory (ICDT '01), J.V. den Bussche and V. Vianu, eds., pp. 316-330, Jan. 2001.

[4]. P. Buneman and W.-C.Tan, "Provenance in Databases," Proc.ACM SIGMOD, pp. 1171-1173, 2007.

[5]. Ms. Aishwarya Potdar1, Ms. Rutuja Phalke2, Ms. Monica Adsul3, Ms.Prachi Gholap4 B.E, Department of Computer Engineering, KJCOEMR, Pune University, Pune, India, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 4, April 2013.

[6]. R. Agrawal and J. Kiernan. Watermarking relational databases. In VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases, pages 155–166. VLDB Endowment, 2002.

[7]. P. Bonatti, S. D. C. di Vimercati, and P. Samarati. An algebra for composing access control policies. ACM Trans. Inf. Syst. Secur., 5(1):1–35, 2002.

[8]. P. Buneman, S. Khanna, and W. C. Tan. Why and where: A characterization of data provenance. In J. V. den Bussche and V. Vianu, editors, Database Theory - ICDT 2001, 8th International Conference, London, UK, January 4-6, 2001, Proceedings, volume 1973 of Lecture Notes in Computer Science, pages 316–330. Springer, 2001.

[9]. P. Buneman and W.-C. Tan. Provenance in databases. In SIGMOD '07: Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pages 1171–1173, New York, NY, USA, 2007. ACM.

[10]. Y. Cui and J. Widom. Lineage tracing for general data warehouse transformations. In The VLDB Journal, pages 471–480, 2001.

224