Querying Different Configurations Simultaneously to Conduct Database Forensic Examinations

Martin S Olivier, PhD, CCFP, Department of Computer Science, University of Pretoria, Pretoria, South Africa

This paper explores the possibilities of simultaneously querying multiple configurations of a database for forensic purposes. A configuration here is defined as a particular combination of selected ANSI SPARC intension-extension layers. The presentation will demonstrate the power of such queries and recommend implementation strategies.

This presentation impacts digital forensics by extending database forensics - an area where very little research exists.

It is common practice to, where possible, examine digital evidence in a 'clean' environment. This is typically achieved by imaging the media and examining the image in a laboratory or by booting the untrusted device with a known (clean) operating system – often from a read-only medium. In previous work it was demonstrated that the notion of a clean environment in the database forensics context is ambiguous given that the data in the database is interpreted through layers of metadata from which the data derives meaning.[1] In previous work it was also demonstrated that different configurations of clean and found layers may be assembled to be able to best answer the forensic questions of interest.[2] However it is obvious that there will not always be a clear 'best' configuration. In fact, often hypotheses will be best tested by asking questions about differences in query results provided by different configurations. The thesis in earlier work was that normal query languages (such as SQL) provide a powerful mechanism to obtain evidence (or leads) from a database. The current work claims that such query languages are even more useful in the case where different configurations are examined simultaneously and proposes that this technique should be explored in detail.

In order to formulate queries across configurations such configurations need to be amalgamated. Two primary options exist: Either the various configurations need to be integrated (in a way similar to normal schema integration) or the configurations have to be configured as a distributed database. In the latter case a number of alternatives exist (which may be systematically explored using Özsu and Valduriez's taxonomy of such databases).[3]

It is demonstrated that the ideal approach for such comparisons depends on the configurations that are to be processed. Where the data models or data dictionaries of the configurations differ, a distributed architecture is indicated. It will be shown that a federated architecture is the best general solution, based on the fact that nodes in a federated database retain autonomy and user permissions will form an inherent part of many database forensic examinations. However, other distributed database architectures may be better suited for specific categories of analyses.

In contrast, where the data model and data dictionary are consistent amongst the configurations to be examined, schema integration is indicated. (In this case configurations will differ on the schema and/or data layers - possibly in a temporal dimension.) Schema integration is facilitated by the fact that the configuration schemas will be identical or very similar in these cases, and a simple tagging approach may be used to distinguish artefacts from the different configurations. Such integration allows a more natural use of the query language (that will facilitate communication of results to non-technical parties in a case). In contrast, the distributed database approaches indicated in the other cases call for a more complex query language. In the case of SQL, in particular, extensions to the language may be required to adequately express forensic queries where configurations are combined as a distributed database.

---

[1] MS Olivier MS. On Metadata Context in Database Forensics. Digital Investigation, 5, 3-4, 115-123, 2009.

[2] Beyers H, Olivier MS and Hancke GP. An Approach to Examine the Metadata and Data of a Database Management System by making use of a Forensic Comparison Tool, in Venter HS, Coetzee M and Loock M (eds.). Proceedings of the 2011 Information Security for South Africa (ISSA 2011) Conference, August 2011, Johannesburg, South Africa.

[3] Özsu MT and Valduriez P. Principles of Distributed Database Systems, 3rd ed., Springer, 2011.

Database Forensics, Forensic Queries, Forensic Configurations