DRAGOON: An Information Accountability System for High-Performance Databases

Kyriacos E Pavlou and Richard T Snodgrass
The University of Arizona Department of Computer Science

Motivation

Motivation

Corporate abuses by Enron and WorldCom have given rise to recent regulations which require many corporations to ensure trustworthy long-term retention of their routine business documents.

- Health Insurance Portability and Accountability Act: HIPAA (1996)
- Sarbanes-Oxley Act (2002)
- U.S. Food and Drug Administration regulation "21 CFR Part 11" (2003)

Due to widespread news coverage of collusion between auditors and the companies they audit, and a lack of tools to address such corruption, there has been interest within the file systems and database communities in built-in mechanisms to detect or even prevent tampering.

Compliant records are those required by law to follow certain "processes by which they are created, stored, accessed, maintained, and retained." It is common to use Write-Once-Read-Many (WORM) storage devices to preserve such records.

Information Accountability vs Restriction

Information restriction entails rendering retained records immutable and controlling access to them. This approach appears to be the prevailing viewpoint for achieving privacy and security.

Information accountability assumes that information should be transparent so as to easily determine whether a particular use is appropriate under a given set of rules.

Information accountability has been tried and tested successfully since ancient times.

Fig. 1. Modern Tamper-Indicating Seals *(left)*. *Bulla*, 14th c. Byzantium *(top)*. *American Scientist*, 94(6):515–524, Nov–Dec 2006

Fair Credit Reporting Act

Objectives

DRAGOON: Database foRensic Analysis safeGuard Of arizONa

DRAGOON is a prototype *continuous assurance* auditing system that is highly customizable in terms of offering a tunable trade-off between level of security and forensic cost. A beta version of DRAGOON is available at:
`http://www.cs.arizona.edu/projects/tau/dragoon/`
It is lightweight, scalable, and adequately addresses aspects of information accountability.

DRAGOON can effectively realize *appropriate use* (i.e., guarantee no unauthorized modifications—insertions, deletions, updates) in high-performance databases. It protects against a variety of threats (including *insider threats*) via *tamper detection* and *forensic analysis algorithms*. DRAGOON can also successfully deal with the aftermath of information restriction failure thereby rendering complex security problems tractable.

DRAGOON is a valuable information accountability solution in the particular area of correct storage, use, and maintenance of relational databases.

Funded by NSF grants IIS-0415101 and IIS-0803229 and a grant from Surety, LLC.

Reference Architecture

The Total Chain Computation Phase

Figure 2 shows the reference architecture of DRAGOON with the colored arrows showing the flow of information during the *Total Chain Computation* phase. All records of the monitored database are hashed and the resulting crypto-strong hash value is periodically notarized. The hash value and the returned unique notary ID are stored in a secure database called *DragoonDB*.

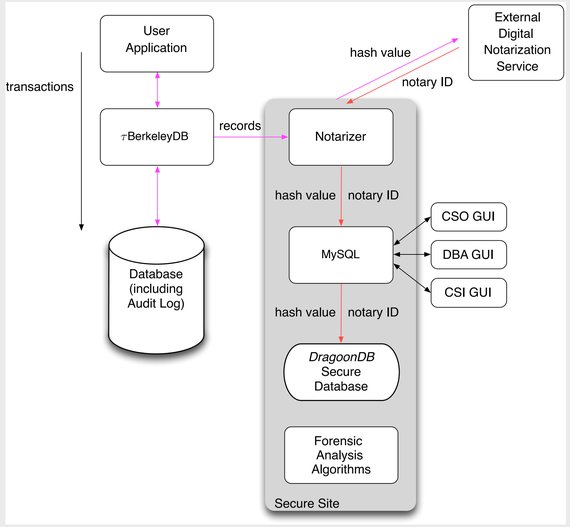


Fig. 2. The Total Chain Computation Phase

The Tamper Detection and Forensic Analysis Phases

During the *Tamper Detection* phase the contents of the monitored database are rehashed and the new hash value is compared against the old one (retrieved using the stored notary ID) by the Notarization Service. A hash value mismatch denotes data corruption. If tampering is detected then forensic analysis algorithms are used to compute spatial and temporal bounds for the corruption.

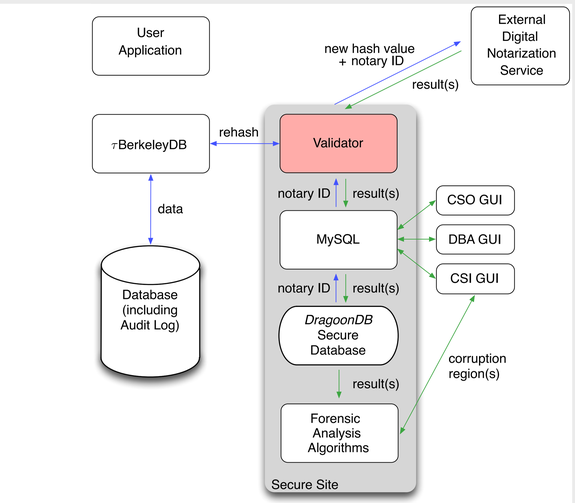


Fig. 3. The Tamper Detection and Forensic Analysis Phases

Forensic Analysis Algorithms

Corruption Diagrams

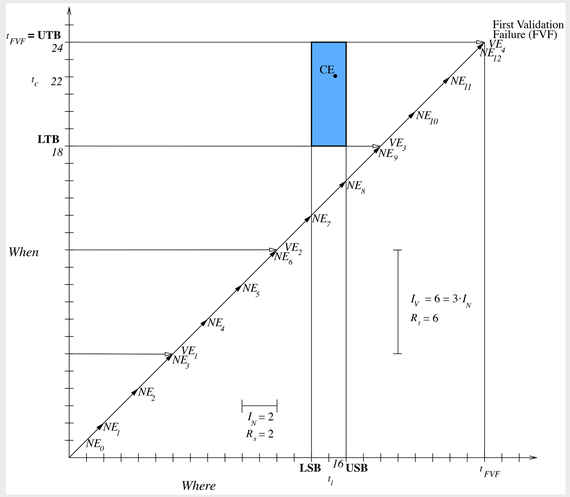


Fig. 4. The Monochromatic Forensic Analysis Algorithm

Symbol	Name	Definition
CE	Corruption event	An event that compromises the database
VE	Validation event	The validation of the audit log by the notarization service
NE	Notarization event	The notarization of a document (hash value) by the notarization service
I_V	Validation interval	The time between two successive *VE*s
I_N	Notarization interval	The time between two successive *NE*s
R_t	Temporal detection resolution	Finest granularity chosen to express temporal bounds uncertainty of a CE
R_s	Spatial detection resolution	Finest granularity chosen to express spatial bounds uncertainty of a CE
t_{FVF}	Time of first validation failure	Time instant at which the CE is first detected
USB	Upper spatial bound	Upper bound of the spatial uncertainty of the corruption region
LSB	Lower spatial bound	Lower bound of the spatial uncertainty of the corruption region
UTB	Upper temporal bound	Upper bound of the temporal uncertainty of the corruption region
LTB	Lower temporal bound	Lower bound of the temporal uncertainty of the corruption region

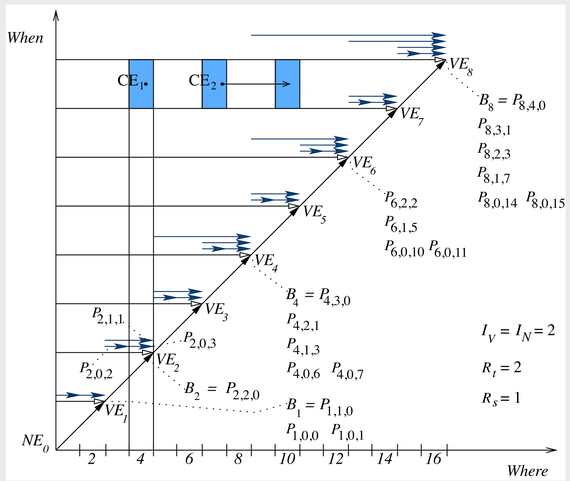
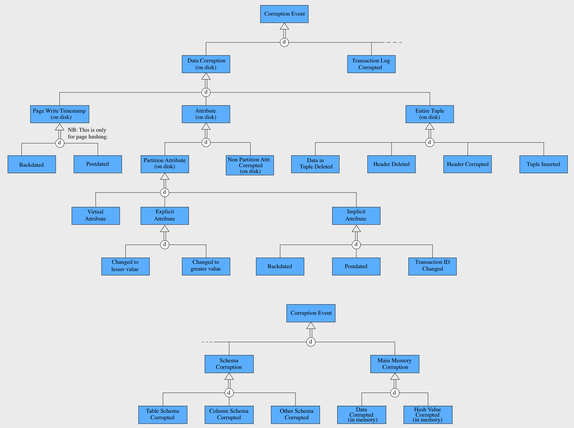


Fig. 5. The a3D Forensic Analysis Algorithm

Taxonomy of Corruption Types



Contributions

The DRAGOON prototype system with advanced tamper detection capabilities and forensic analysis tools is useful in a variety of sectors.
DRAGOON can:

- ensure record compliance for financial and medical institutions,
- serve as an unbiased witness to databases storing sensitive information, e.g., court-submitted data from police databases,
- ensure non-deviation from standard operating procedures in biosciences labs (provenance of results),
- detect bugs silently corrupting databases,
- can be deployed in the cloud (vid. DMC'12)
- automate some of the forensic work required in the aftermath of a database corruption saving both time and money,
- provide advantages over information restriction approaches which rely on special hardware (prohibitive costs for small institutions, limited shelf-life, relatively complex), and
- mirror the relationship between the law and human behavior more closely.

References

K. E. Pavlou and R. T. Snodgrass. Forensic Analysis of Database Tampering. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 109–120, June 2006.

K. E. Pavlou and R. T. Snodgrass. Forensic Analysis of Database Tampering. *ACM Transactions on Database Systems*, 33(4):1–47, November 2008.

K. E. Pavlou and R. T. Snodgrass. The Tiled Bitmap Forensic Analysis Algorithm. *IEEE Transactions on Knowledge and Data Engineering*, 22(4):590–601, April 2010.

R. T. Snodgrass, S. S. Yao, and C. Collberg. Tamper Detection in Audit Logs. In *Proceedings of the International Conference on Very Large Databases*, pages 504–515, September 2004.