FLAIM: A New Tool for Log Anonymization

A. Slagell, K. Lakkara, J. Luo

Vision:
A Privacy-preserving Tool for all Data Sources

The need for Anonymization
1. Meet Privacy and Security Needs of Data Owner: Logs, databases, and audit trails all contain sensitive information whose protection is often mandated by legislation.
2. Meet the Utility Needs of the Data Analyzer: Those doing analysis often need some part of the data to do the task, but not everything.
3. Anonymization: Finds the Middle Ground: The goal is to find the appropriate trade-off between information loss and privacy that meets the needs of both parties; so sharing can happen.
4. Example Scenarios: Share data for research; share data for educators, creating useful projects, Share data in real time for an investigation.

FLAIM: A new kind of Anonymizer

The state of the art anonymization tools lack several features. They usually can only anonymize a single type of log and they do not offer many anonymization algorithms.

We have developed FLAIM with the purpose of allowing users to easily allow new logs to be anonymized, and to provide users of anonymization primitive to use.

FLAIM has been developed with these four objectives in mind:

1. Broad Set of Anonymization Algorithms
2. Multi-Level Anonymization
3. Multi-Log Support
4. Modular and Extensible

FLAIM uses anonymization algorithms that have been individually implemented for over a dozen common data types in network logs.

To use FLAIM the user must specify an anonymization policy, an XML document that indicates which anonymization algorithms are applied to which field. The anonymization policy is parsed by the policy manager at run time. This allows users to specify which fields to anonymize and in what order. The policy manager is written in C without having to recompile the entire tool. In addition, users can switch between different policies easily, depending upon the circumstances. This allows users to easily change the level of anonymization for a log.

Below is a diagram illustrating the interaction between the various components of FLAIM:

FLAIM Anonymization Algorithms

Black Marker - replaces all values of a field with a single canonical value for the field
P... records by the value, but then remove the specific value and instead assign an arbitrary, but increasing, number.

Netfilter Example

We have developed an I/O module for the Netfilter log format. We have also described several anonymization algorithms that can be used to compact Netfilter logs. Using FLAIM, one can anonymize Netfilter logs in minutes. The possible anonymization algorithms that can be applied to Netfilter logs are listed below. In addition we have shown a couple of records from a Netfilter log in its unformatted and anonymized form. One possible anonymization policy that could have been used to anonymize these records is also shown.

http://flaim.ncsa.uiuc.edu

www.iti.uiuc.edu

University of Illinois at Urbana-Champaign