Integrated Provenance Data for Access Control in Group-centric Collaboration

Dang Nguyen, Jaehong Park and Ravi Sandhu
Institute for Cyber Security
University of Texas at San Antonio

August 09, 2012
International Conference on Information Reuse and Integration
Provenance of a digital data object is defined as the documentation of its origin and all the processes that influence and lead to its current state.

In a provenance-aware system, related provenance information of system transactions/events are captured, stored, and maintained.

Provenance potentially provides many enhanced benefits: usage tracking, workflow control, versioning, trustworthiness, repeatability, access control, etc.

Provenance information may be more sensitive than the underlying data.
Is necessary: Integrity, Confidentiality, Availability, Privacy

Our focus: Access Control.

Two aspects: Provenance-based Access Control and Provenance Access Control.

Provenance data naturally forms a Directed-Acyclic Graph (DAG), aligned with information flow and causality.

World-Leading Research with Real-World Impact!
Difference in the incorporation of base model for PBAC in uni vs. multi-provenance systems.

Group-centric collaboration provides secure information sharing.

Support **administrative** and **usage** operations.

Focus on usage operations such as: Add, Merge, Update, Create.
PBAC MODEL

World-Leading Research with Real-World Impact!
OPEN PROVENANCE MODEL (OPM)

World-Leading Research with Real-World Impact!
A SIMPLIFIED SCENARIO

An object is created in org1 and modified locally into versions in accordance to the versioning system.

At some point in time, a version of this object is added to a collaboration group cg1 so users from a different organization can participate in updating the object content (now represented as a different object with its own versions).

Meanwhile, users in the org1 also perform updates on local versions.

At some point, a version of the object from cg1 is merged back to the version tree of the original object in org1.
A SIMPLIFIED OPM SCENARIO

World-Leading Research with Real-World Impact!
Req (ad1, merge, CG1.o2v3, Org1.o1v2) ?

\[
allow(au, merge, o_{from}, o_{to}) \Rightarrow o_{to} \in (o_{from}, wasDerivedVersionOfCopyOf)
\]

(\text{wasDerivedVersionOfCopyOf} :: [g(Update).u]^{*}.g(Add).u)

Req(au2.1, update, CG1.o2v3) ?

\[
allow(au, update, o) \Rightarrow au \notin (o, creatorOfOriginalVersionOf)
\]

(\text{creatorOfOriginalVersionOf} :: [g(Update).u]^{*}.g(Add).u.[g(Update).u]^{*}.g(Create).c)

World-Leading Research with Real-World Impact!
TAXONOMY OF APPROACHES

Integrated Use of Provenance Data (PD)

Integration Approaches:
- Using Cascading subquery
- Using Sticky Provenance Data (SPD)

Response to cascading query:
- Y/N
- Resulting Nodes
- Sticky PD

World-Leading Research with Real-World Impact!
A reconstructed query from the local query.

Can be transmitted and evaluated in remote system.

Three types of response, each require different additional information:
1. Y or N: (startingNode_new,dPath_new,rule_new) must be transmitted.
2. Resulting Nodes: (startingNode_new,dPath_new) must be transmitted.
3. Provenance Data Set: (startingNode_new) must be transmitted.
CASCADING QUERY RESPONSE

ORG1

(\text{startingNode\_new, dPath\_new, rule\_new}) \quad \rightarrow \quad \text{Y or N}

\leftarrow \quad \text{(startingNode\_new, dPath\_new)}

\rightarrow \quad \text{Resulting Nodes}

\leftarrow \quad \text{(startingNode\_new)}

\rightarrow \quad \text{Provenance Data Set}

CG1

World-Leading Research with Real-World Impact!
The sticky provenance data of an object/version contains all the provenance information of that object/version up to the point in time when the information flow takes place.

Allows a locally generated query to be fully evaluated for decision making.

Demonstrate with a modified scenario next.
A “STICKY” SCENARIO

World-Leading Research with Real-World Impact!
A “STICKY” SCENARIO

CG1 Policy: No Update actions can be performed on a group version if the original organization object version had been updated.

World-Leading Research with Real-World Impact!
A “STICKIER” SCENARIO

Should SPD(o3v1) contain:
SPD(o2v1,o2v2) ?
SPD(o2v1,o2v2) + SPD(o2v1) ?

World-Leading Research with Real-World Impact!
CONCLUSION

Demonstrated the incorporation of PBAC in a Group-centric collaboration environment.

Identified the issue in a multi-provenance systems setting.

Proposed two approaches to address such issue.