CONTRÔLE D'ACCÈS BASÉ SUR LA PROVENANCE

GESTION DES MÉTADONNÉES RELATIONNELLES DE SÉCURITÉ

François LESUEUR, Romuald THION

francois.lesueur@liris.cnrs.fr
romuald.thion@liris.cnrs.fr

Plan

Politiques d'autorisation

Problème

Permettre aux sources de données de contrôler la diffusion et les traitements effectués sur leurs données.

Proposition

- On considère une annotation (métadonnée) de sécurité associée à chaque tuple (*sticky policy*) ;
- Chaque source définit des groupes de sécurité (discrétionnaire) ;
- Faire en sorte que la politique de la source soit respectée :
  - Filtrage lors de l'évaluation des requêtes ;
  - Matériel cryptographique que seul le propriétaire peut délivrer ;
- Définir le comportement des groupes lors des requêtes.

Gestion des politiques

Gestion des annotations relationnelles

- Travaux sur la *provenance* (Green, Tannen)
- Modèle des K-relations (tuples annotées avec des K) ;
- K doit être équipé d'une certaine structure ;
- un *semi-anneau commutatif* $(K, \oplus, \otimes, 0, 1)$

Intuition : les propriétés algébriques de SPJRU se reflètent dans les annotations et leur donnent une structure algébrique

Utilisation pour la sécurité (état de l'art)

- $\mathbb{L}$ = labels totalement ordonnés de confidentialité
- $\mathbb{L} = P < C < S < TS < O$
- $(\mathbb{L}, min, max, O, P)$

Gestion des politiques

Modélisation des autorisations

- Modèles mandataires
- Modèles discrétionnaires
- Modèles à groupes, à rôles, ...

Cette contribution

- Montrer que l'on peut avoir une gestion des droits plus souple que les labels de confidentialité dans le cadre *provenance* ;
- Applications de la cryptographie pour réduire les hypothèses de confiance sur le code.

Combinaisons de politiques de sécurité

R	A	B	
r_0	a	1	les ABs ou les Cs
r_1	a	2	les Bs ou les Cs

S	A	B	
s_0	a	2	les ABs
s_1	b	1	les As ou les Bs

$R \cup S$	A	B		
t_0	a	1	les ABs ou les Cs	(r_0)
t_1	a	2	les ABs, les Bs ou les Cs	$(r_1 \oplus s_0)$
t_2	b	1	les As ou les Bs	(s_1)

$R \bowtie_A S$	A	B	B'		
u_0	a	1	2	les ABs ou les ABCs	$(r_0 \otimes s_0)$
u_1	a	2	2	les ABs ou les ABCs	$(r_1 \otimes s_0)$

Structure des groupes de sécurité

Structure $\langle \mathcal{P}(\mathcal{P}(G)), \cup, \Cup, \emptyset, \{\emptyset\} \rangle$

- G l'ensemble des groupes de sécurité
- $\mathcal{P}(\mathcal{P}(G))$: ensembles d'ensembles d'annotations
  par exemple $\{\{g_0, g_1\}, \{g_2\}, \{g_0, g_3\}\}$
- $X \Cup Y = \{x \cup y \mid x \in X \wedge y \in Y\}$

Évaluation

- consommateur exhibe un ensemble d'accréditations $Cr \subseteq G$
  e.g., des certificats signés par les sources
- pour chaque tuple du résultat, on compare Cr à son annotation :

$$eval : \mathcal{P}(G) \times \mathcal{P}(\mathcal{P}(G)) \to \mathbb{B}$$

$$eval(Cr, X) = \exists C \in X . C \subseteq Cr$$

Combinaisons de politiques de sécurité

R	A	B	
r_0	a	1	$\{\{g_0, g_1\}, \{g_2\}\}$
r_1	a	2	$\{\{g_2\}, \{g_1\}\}$

S	A	B	
s_0	a	2	$\{\{g_0, g_1\}\}$
s_1	b	1	$\{\{g_0\}, \{g_1\}\}$

$R \cup S$	A	B		
t_0	a	1	$\{\{g_0, g_1\}, \{g_2\}\}$	(r_0)
t_1	a	2	$\{\{g_2\}, \{g_1\}, \{g_0, g_1\}\}$	$(r_1 \oplus s_0)$
t_2	b	1	$\{\{g_0\}, \{g_1\}\}$	(s_1)

$R \bowtie_A S$	A	B	B'		
u_0	a	1	2	$\{\{g_0, g_1\}, \{g_0, g_1, g_2\}\}$	$(r_0 \otimes s_0)$
u_1	a	2	2	$\{\{g_0, g_1, g_2\}, \{g_0, g_1\}\}$	$(r_1 \otimes s_0)$

Remarque

D'après la définition de $eval_{Cr}$, on peut simplifier certaines expressions, e.g.

$$eval_{Cr}(\{\{g_1\},\{g_2\},\{g_0,g_1\}\}) = eval_{Cr}(\{\{g_1\},\{g_2\}\})$$

Le modèle se comporte « bien »

$eval_{Cr}$ est un homomorphisme
de $\langle \mathcal{P}(\mathcal{P}(G)), \cup, \Cup, \emptyset, \{\emptyset\}\rangle$ dans $\langle \mathbb{B}, \vee, \wedge, \bot, \top\rangle$:

$$eval_{Cr}(X \cup Y) = eval_{Cr}(X) \vee eval_{Cr}(Y) \qquad eval_{Cr}(\emptyset) = \bot$$
$$eval_{Cr}(X \Cup Y) = eval_{Cr}(X) \wedge eval_{Cr}(Y) \qquad eval_{Cr}(\{\emptyset\}) = \top$$

Mise en œuvre du modèle

Architecture #1 : confiance entre les acteurs du système

- Le SGBD collecte les crédentials Cr de l'utilisateur ;
- Le SGBD calcule le résultat de la requête et ses annotations ;
- Le SGBD filtre en calculant $eval_{Cr}$ pour chaque tuple.

Architecture #2 : minimum de confiance entre les acteurs du système

- À chaque annotation correspond un matériel cryptographique ;
- Le SGBD travaille sur des données chiffrées ;
- L'utilisateur collecte les clefs pour déchiffrer le résultat ;
- Une annotation $\cong$ un ensembles de « recettes » pour déchiffrer.

Conclusion

Modèle de sécurité

- Contrôle des flux d'informations par les sources ;
- Droits d'accès représentés par des annotations de provenance ;
- Conception en environnement de confiance.

Travaux en cours

- Chiffrement conforme aux annotations (commutatif, alternatif) ;
- Mise à contribution des sources (protocoles cryptographiques interactifs) ;
- Amélioration de l'expressivité des politiques de sécurité :
  - Annotations des attributs d'un tuple ;
  - Combinaison de politiques ;
- Autres modèles de données (arbres XML, graphes RDF, clefs/valeurs noSQL)

Semi-anneau$(K, \oplus, \otimes, 0, 1)$

- $(K, \oplus, 0)$ un monoïde commutatif ($\oplus$ associée à $\cup$ et π)
  - (associative) $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
  - (unitaire) $a \oplus 0 = a = 0 \oplus a$
  - (commutative) $a \oplus b = b \oplus a$
- $(K, \otimes, 1)$ un monoïde ($\otimes$ associée à $\bowtie$)
- (distribution) $a \otimes (b + c) = a \otimes b + a \otimes c$
- (distribution) $(b + c) \otimes a = b \otimes a + c \otimes a$
- (absorbant) $0 \otimes a = 0 = a \otimes 0$

Le seigneur des semi-anneaux

Le semi-anneau librement généré à partir d'un ensemble X
$\cong$ aux polynômes sur un ensemble de variables $\mathbb{N}[X]$

Intuition : cas général où chaque tuple a un id unique, et où les effets des requêtes sur les tags sont aussi différents que possible.

Modélisation avec les annotations

- Provenance (*where, how, why*)
  - d'où vient un tuple ?
  - comment a été obtenu ce tuple ?
- Uncertainty (quelle probabilité de présence, présence de nulls)
- Trust scores (quelle confiance dans le tuple)
- Multiplicity (bag semantics)
- Security (labels totalement ordonnés de confidentialité)

Sémantique formelle $(K, \oplus, \otimes, 0, 1)$

- Provenance (*where, how, why*)
- Uncertainty $(\mathcal{P}(\Omega), \cup, \cap, \emptyset, \Omega)$
- Trust scores $(\mathbb{R}_+^\infty, min, +, \infty, 0)$
- Multiplicity $(\mathbb{N}, +, \times, 0, 1)$
- Security $(\mathbb{L}, min, max, O, P)$ avec $\mathbb{L} = P < C < S < TS < O$