1. REPORT DATE  
18 NOV 2009

2. REPORT TYPE

3. DATES COVERED
00-00-2009 to 00-00-2009

4. TITLE AND SUBTITLE
Quantitative Trust Management: QuanTM, Reputation, and PreSTA

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania, School of Engineering and Applied Science, 220 South 33rd Street, Philadelphia, PA, 19104-6391

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
PRECISE Research Group Presentation, Nov. 2009

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:
   a. REPORT unclassified
   b. ABSTRACT unclassified
   c. THIS PAGE unclassified

17. LIMITATION OF ABSTRACT
   Same as Report (SAR)

18. NUMBER OF PAGES
   48

19a. NAME OF RESPONSIBLE PERSON

Standard Form 298 (Rev. 8-98)  
Proscribed by ANSI Std Z39-18
• Introducing QTM
• QuanTM [1] Model
  – TDG: Glue between security & reputation
• Fundamentals of Reputation Management
• PreSTA [3] Reputation Model
  – Partial QTM use-case
  – Applicable for fighting spam, Wikipedia...
• Conclusions
'Trust Management' (TM) aspect

- **STATIC** delegation of access rights between principals using policy/credentials/conditions
- Implemented by a **Policy-based TM (PTM)** language (*i.e.*, KeyNote) and evaluator ('compliance checker')

'Quantitative' (Q) aspect

- **DYNAMIC** weighting of above delegations, based on reputations of those involved
- Implemented by a **Reputation Management (RTM)** algorithm (PreSTA [3], TNA-SL [4], EigenTrust [5])
**QTM DEFINED**

**Policy-Based Trust Mgmt. (PTM)**
- Effective for delegated credentials and access enforcement
- Can’t handle uncertainty and partial information
- Foundation: Cryptography

**Rep-Based Trust Mgmt. (RTM)**
- Quantifies trust relationships
- No delegation (non-transferable)
- No enforcement
- Foundation: Aggregation of past behavior via feedback.

**QUANTITATIVE TRUST MANAGEMENT (QTM)**
- Combine PTM and RTM
- Dynamic interpretation of authorization policies for access control decisions based on evolving reputations of the entities involved, and environmental context at evaluation-time [6].
QTM for CPS

• MAIN GOAL
  – Integrating cyber and physical trusts

• ISSUES FORESEEN
  – Authentication/provenance of physical stimuli
  – Environmental uncertainty

• POTENTIAL USE-CASES
  – Voting machines
  – Emergency management
QuanTM Model

Combining TM and RM [1]
BUILDING A TDG

Authorizer: Alice
Licensees: (Bob && Charles)
Conditions:
  operation ==
    “read” -> ALLOW
    “execute” -> MAYBE
    “write” -> DENY
Signature: “rsa-sig:3850...”

Trust Dependency Graph (TDG): Data structure gluing Policy and Reputation based TM.

Above: An example KeyNote credential
Authorizer: Alice
Licensees: (Bob && Charles)
Conditions:
    operation ==
    "read" -> ALLOW
    "execute" -> MAYBE
    "write" -> DENY
Signature: “rsa-sig:3850…”

Authorizer: The person who is “saying” a particular delegation
Authorizer: Alice
Licensees: (Bob && Charles)
Conditions:
  operation ==
    “read” -> ALLOW
    “execute” -> MAYBE
    “write” -> DENY
Signature: “rsa-sig:3850...”

Binary Operator: Nature of the delegation. Here, “AND” implies both parties must be present. KeyNote also supports “OR”
Authorizer: Alice
Licensees: (Bob && Charles)
Conditions:
  operation ==
    “read” -> ALLOW
    “execute” -> MAYBE
    “write” -> DENY
Signature: “rsa-sig:3850...”

Licensees: Those parties the 'Authorizer' is delegating trust to, as constrained by the binary operator
Authorizer: Alice
Licensees: (Bob && Charles)
Conditions:
  operation ==
      “read” -> ALLOW
      “execute” -> MAYBE
      “write” -> DENY
Signature: “rsa-sig:3850...”

Compliance values: Output of the evaluator. Varies based on evaluation of conditions. Could be a binary YES/NO.
CREDENTIAL GROUPS:

We divide portions of the graph based on the credentials from which they were derived.
NULL NODES:

(1) Used to make graph explicitly binary

(2) Overwrite principals mentioned in credentials, but not 'present' in a particular request
- TDG: Excellent representation of trust dependencies in a KEYNOTE request
  - Other TM languages?
- We would like to have a TDG structure which can encapsulate the features of all/general trust management langs.
BIG IDEA:

Each graph arc can be weighted with a value speaking to the reputation of connecting parties.

These can be collapsed to produce a single TRUST VALUE for an entire request.
BIG IDEA:

Each graph arc can be weighted with a value speaking to the reputation of connecting parties.

These can be collapsed to produce a single TRUST VALUE for an entire request.
Reputation $R_1$:

Arcs from operators to principals

Weight with service providers (BANK) reputation valuation of sink principal
Reputation $R_1$:

Arcs from operators to principals

Weight with service providers (BANK) reputation valuation of sink principal

* Magic numbers
Reputation $R_2$:

Arcs from principals to operators

Weight with service provider's (BANK) trust in 'the ability of the source principal to delegate'
Reputation $R_2$:

Arcs from principals to operators

Weight with service provider's (BANK) trust in 'the ability of the source principal to delegate'

* Mention $R_3$
USING THE TDG

Graph Collapse:
Graph Collapse:

* Swap out binary operators for numeric binary functions

* Start at TDG-bottom, perform functions, pass resulting values up the graph. Transitivity handled by multiply.
Graph Collapse:

* Swap out binary operators for numeric binary functions
* Start at TDG-bottom, perform functions, pass resulting values up the graph. Transitivity handled by multiply.
Graph Collapse:

* Swap out binary operators for numeric binary functions

* Start at TDG-bottom, perform functions, pass resulting values up the graph. Transitivity handled by multiply.
Graph Collapse:

* Swap out binary operators for numeric binary functions

* Start at TDG-bottom, perform functions, pass resulting values up the graph. Transitivity handled by multiply.

\[
0.62 = 0.65 \times 0.95
\]

\[
1.0 = 0.65
\]
Graph Collapse:

* Swap out binary operators for numeric binary functions

* Start at TDG-bottom, perform functions, pass resulting values up the graph. Transitivity handled by multiply.
There was an action request made...

- The TM language evaluator outputs some compliance value, *e.g.*, “MAYBE”
- We generated a TDG, and collapsed it using magic numbers, *e.g.*, “0.62”

... Combining these two things, and sufficient hand-waving -> binary access decision

- Cost-benefit analyses
• TM: Revocation difficult - One shouldn't delegate unless they completely trust.
  - QTM: Dynamic revocation using reputation
  - QTM: Safe to delegate in partial trust situations
• TM: Rights can be delegated to principals that service provider knows nothing about
  - QTM: Can check these new principals at the reputation stage
• RM: Lacks enforcement/delegation
Rep. Management

Aggregating Behavioral Feedback

(and testing these strategies [2])
• **DYNAMIC** valuation using (in)direct interaction history between parties
  - Loose interpretation: probability that A trusts B
  - Informal; produces values in [0,1]
  - Many different logics/systems to aggregate feedback
    - EigenTrust (Garcia-molina) and Subjective-Logic (Jøsang)
EIGENTRUST [5]

- Normalized vector-matrix multiply aggregation towards globally convergent view.
  - Feedbacks viewed in matrix, normalized

\[ A = \begin{bmatrix}
  \text{pos : 0} & 0 & \text{pos : 3} & 2 & \text{pos : 3} & 1 \\
  \text{neg : 0} & \text{neg : 1} & \text{neg : 2} & 7 \\
  \text{pos : 9} & \text{pos : 0} & \text{pos : 8} & 0 \\
  \text{neg : 3} & \text{neg : 0} & \text{neg : 1} & 0 \\
  \text{pos : 2} & \text{pos : 5} & \text{pos : 0} & 1 \\
  \text{neg : 4} & \text{neg : 4} & \text{neg : 0} & 0
\end{bmatrix} \]

\[ A' = \begin{bmatrix}
  0/6 & 2/3 & 1/8 \\
  6/6 & 0/3 & 7/8 \\
  0/6 & 1/3 & 0/8
\end{bmatrix} \quad p = \begin{bmatrix}
  1/3 \\
  1/3 \\
  1/3
\end{bmatrix} \quad t_\infty = \begin{bmatrix}
  0.35 \\
  0.49 \\
  0.16
\end{bmatrix} \]

\[ t_{k+1} = (0.5 \times A'^T \times t_k) + 0.5 \times p \]
• Trust 4-tuples (belief, disbelief, uncertainty, ...)
• User-centric trust-graph decomposition
• Advantages: Absolute interpretation (beta-PDF), user-centric views, negative trust
• Disadvantages: Scalability, sparse scenarios

Opinion: \((b, d, u, a)\)

<table>
<thead>
<tr>
<th>Trust State</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>belief</td>
<td>((\text{pos}/(\text{pos} + \text{neg} + 2.0)))</td>
</tr>
<tr>
<td>disbelief</td>
<td>((\text{neg}/(\text{pos} + \text{neg} + 2.0)))</td>
</tr>
<tr>
<td>uncertainty</td>
<td>((2.0/(\text{pos} + \text{neg} + 2.0)))</td>
</tr>
<tr>
<td>base-rate</td>
<td>(\begin{cases} 1.0 &amp; \text{if user is pre-trusted} \ 0.5 &amp; \text{otherwise} \end{cases})</td>
</tr>
</tbody>
</table>

Transitivity:

\[
\omega^A_{C:B} = \omega^A_B \otimes \omega^B_C
\]

Average:

\[
\omega^A_{C:B} = \omega^A_C \oplus \omega^B_C
\]
• How to test effectiveness of RM systems?
• Simulator [2]: File exchange (i.e., P2P network)
  – Good files and corrupt files
  – Behaviors: Clean-up and honesty

\[ \text{Metric} = \frac{\# \text{ valid files received by ‘good’ users}}{\# \text{ transactions attempted by ‘good’ users}}. \]
Under naïve circumstances, all trust algorithms are very effective (a sanity check).

Under complex dishonesty and sparseness, PRE-TRUST becomes very important.
PreSTA Model

(Preventative Spatio-Temporal Aggregation)

Preventing Malicious Behavior (Spam) [3]
PreSTA: BIG IDEA

**PROBLEM**
- Traditional punishment mechanisms (*i.e.*, blacklists) are **reactive**
- PreSTA: Detect malicious users (*i.e.*, spammers) **before** harm is done

**SOLUTION**

**HYPO-THESIS:**
- Malicious users are **spatially** clustered (in any dimension)
- Malicious users are likely to repeat bad behaviors (**temporal**)

**GIVEN:**
- A historical record of those principals **known** to be bad, and
  the timestamp of this observation (**feedback**)

**PRODUCE:**
- An **extended** list of principals who are **thought** to be bad **now,**
  based on their past history, and history of those around them
• IP delegation hierarchy extremely similar to TDG

• Exploit this fact:
  – Calculate reputations at varying hierarchy levels
  – Feedback: IP blacklists
  – Combine granularities

• Can more malignants (spammers) be caught?
**TEMPORAL: Bad Guys Repeat Bad Behaviors**

- Maximize utilization: re-use
- Predictable blacklist duration
- 25% reappear within 10 days

**SPATIAL: Bad Guys Live Together**

- Corrupt ISPs: McColo, 3FN
- Geography -> IP space
- Intra-allocation spamming
TO CALCULATE REPUTATION FOR ENTITY $\alpha$:

$$\text{raw}_\text{rep}(\alpha) = \sum_{i=1}^{i \leq |\text{BL}(\alpha)|} \frac{\text{time}_\text{decay}(\text{BL}(\alpha)_i)}{\text{magnitude}(\alpha)}$$

$$\text{REP}(\alpha) = 1.0 - (\text{raw}_\text{rep}(\alpha) \times \phi^{-1})$$

- $\text{time}_\text{decay}(*):$ Returns on $[0,1]$, higher weight to more recent events
- $\text{magnitude}(\alpha):$ Number of IPs in grouping $\alpha$
- $\phi$: Normalization constant putting $\text{REP}()$ on $[0,1]$
We capture between 20-50% of spam that gets past current blacklists
  - By design our FP-rate is equivalent to BLs: ~0.4%

- Total blockage remains near constant: 90%
  - Blacklists are reactive, we are predictive. We can cover its slack
  - Cat and mouse. Graph should roll over time

Captures up to 50% of mail not caught by traditional blacklists with the same low false-positives
PreSTA Results

(LEFT) Temporal (single IP) example where our metric could mitigate spam

(RIGHT) Probable botnet attack which our metric could mitigate via both temporal/spatial means
**PreSTA: Wikipedia**

**PURPOSE:** Build a blacklist of user-names/IPS based on the probability they will vandalism

**TEMPORAL**
- Straightforward, vandals are probably repeat offenders
- Registered users have IDs indicating when they joined, are new users more likely to vandalize?

**SPATIAL**
- Geographical: Based on user location (i.e., Wash. D.C.)
- Topical: A user may vandalize one topic (Rush Limbaugh), while properly editing another (Barack Obama)
- Anonymous users: IP address properties

**FEEDBACK**
- Certain administrators have rollback (revert) privileges
- Comment: “Reverted edit by X to last edition by Y”
CONCLUDING (ALL)

• Quantitative Trust Management (QTM)
  – Combines Policy-based and Reputation-based TM

• QuanTM [1] framework
  – Theoretical underpinnings of combination
  – TDG as the shared data-structure
  – Partial applications:
    • Simulator [2] for reputation-component
    • PreSTA [3]: Reputation incorporating properties of a hierarchical delegation (as in PTM)
REFERENCES


