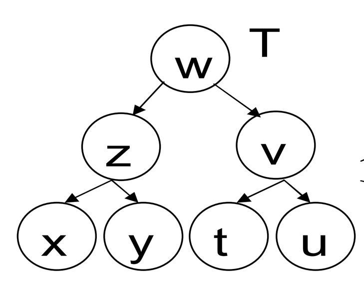
CERIAS

the center for education and research in information assurance and security

Completely-Secure Sharing of Trees and Hierarchical Content

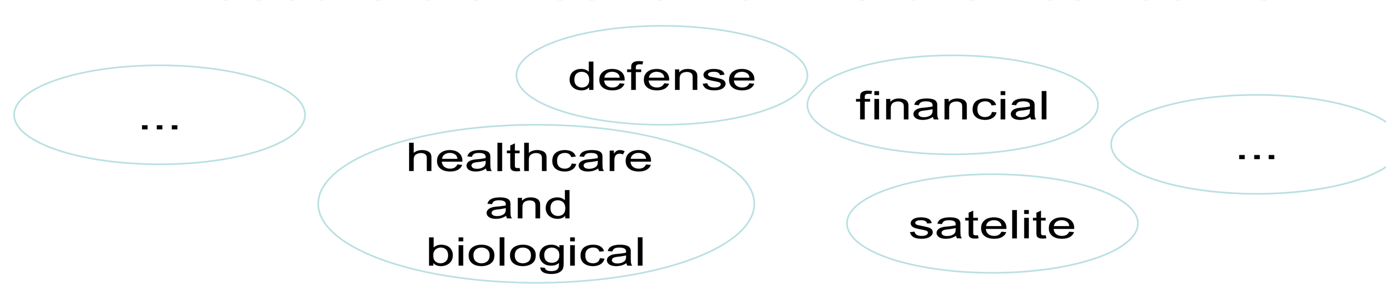
Ashish Kundu, Elisa Bertino CERIAS, Purdue University

Hierarchical data forms: trees

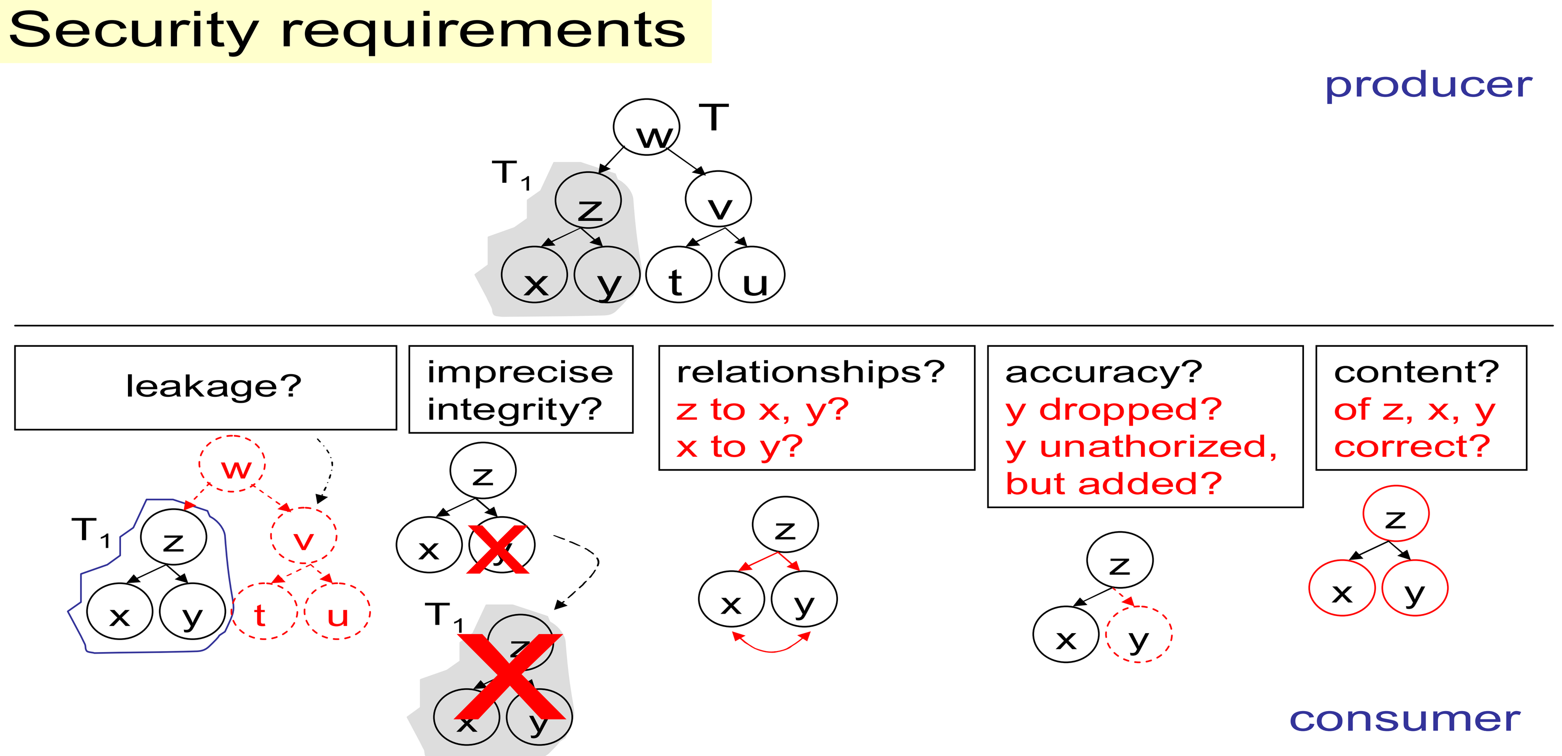


- XML
- VoiceXML
- Composite data objects
- Serialized objects
- Views of mobile applications

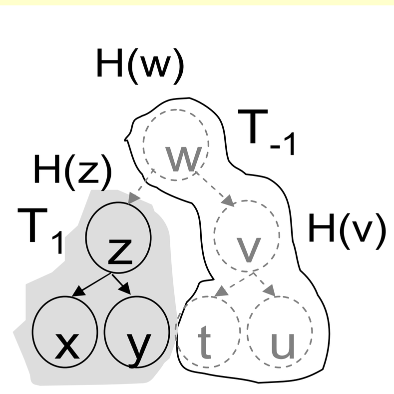
Problem: secure distribution of hierarchical content.



Security requirements



Merkle Hash: Leaks information



- H(z) = G(H(x) | H(y))
- H(w) = G(H(z) | H(v))
- H, G: hash functions for leaves, non-leaves
- Uses cascaded hashing: non-associative

Leakage

- Structural ordering between z, v and w.
- H(V), H(w)

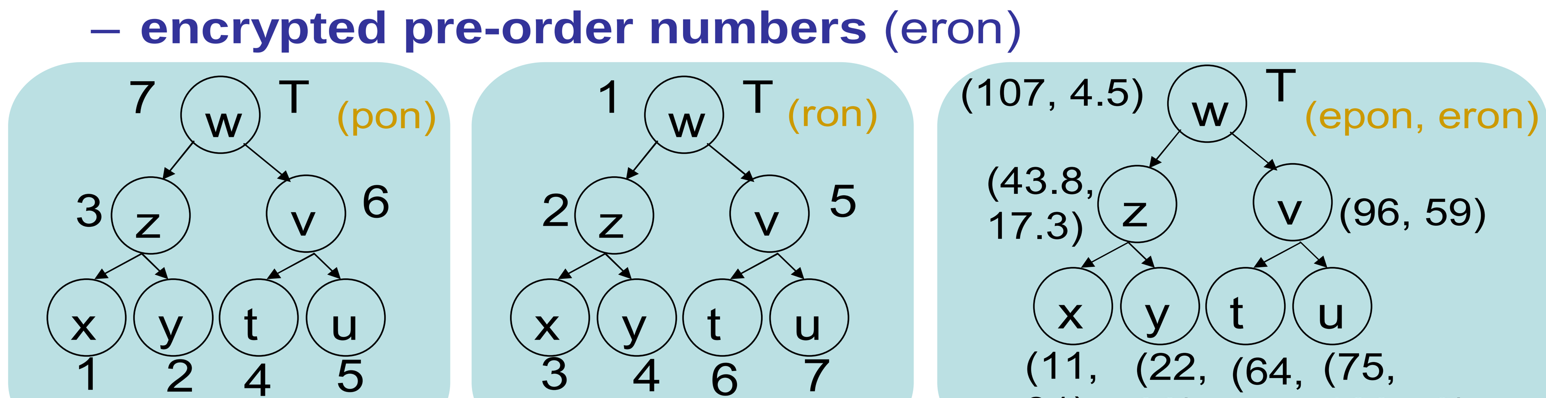
Inference attacks

✓ Node v is on the right of x, y, z: *semantic relationship between v and x, y, z*

✓ Knowledge of H(v) and H(z): *tree is larger, semantic information about source data*

Simple tree traversals

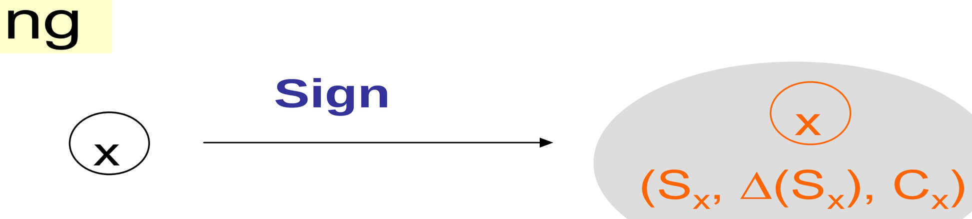
- **Unique re-construction** of a tree: from its post-order and pre-order traversals
- Post-order numbers (pon) and pre-order numbers (ron)
- Anonymize structural information
  - **encrypted post-order numbers** (epon)
  - **encrypted pre-order numbers** (eron)



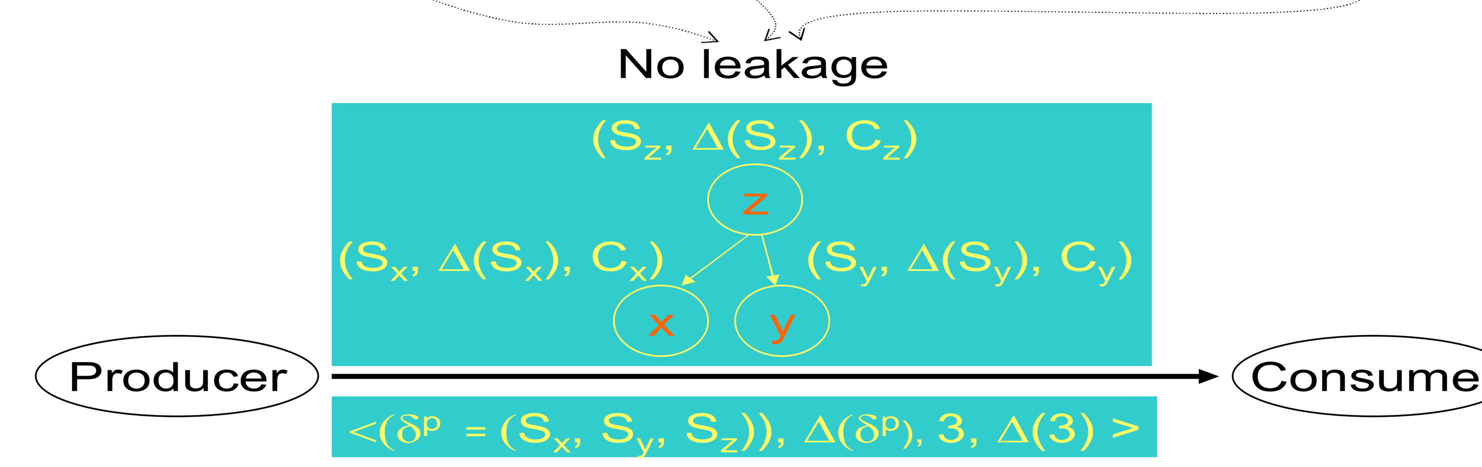
Structural Signature

S_x: Structural signature of node x	(p^e_x, r^e_x) – p^e_x , r^e_x : EPON, ERON of x
C_x: Content signature of node x	$\Delta(S_x, H(g_x))$ – g_x: content of node x – H: one-way collision-resistant hash function – Δ: MAC operator
$\delta^h(V_i)$: Signature of a sub-set of nodes V_i in tree T	$\rho^h(S_x$, x in $V_i)$ – ρ^h : *h*-order sequence of nodes in V_i - *h*: p: post-order, r: pre-order

Signing & Sharing



Share: < signed nodes, signature of the set of nodes, number >



Consumer-side Integrity Verification

Signature integrity	Δ(received S_x) ≠ received $\Delta(S_x)$: S_x in-authentic Δ(received δ^h) ≠ received $\Delta(\delta^h)$: δ^h in-authentic Δ(received number) ≠ received Δ(number): number of nodes in-authentic
Relationships	$(p^e_x \ge p^e_z)$ OR $(r^e_x \ge r^e_z)$: (z, x) order incorrect $(p^e_x \ge p^e_y)$ OR $(r^e_x \ge r^e_y)$: (x, y) order incorrect
Content Integrity	$\Delta(S_x, H(g_x)) \neq C_x$: content of x, g_x in-authentic
Accuracy	number of received nodes (≠ | ≥ | ≤) received number: received (not exact, more | less) nodes
Authenticity	S_x not in δ^h : x in-authentic δ^h(received nodes) ≠ received δ^h: data in-authentic

Conclusions

- We showed that Merkle hash technique leaks information and does **not** support complete confidentiality

Our approach: first such technique for complete security of trees

✓ **No leakage of information:** complete confidentiality with encrypted transmission

✓ **Precise verification of integrity:** efficient data-recovery and failure-oblivious computing

✓ **Worst case - O(n):** n is the number of nodes

✓ **Efficient:** no cascaded hashing

✓ **Easy to implement:** post-order, pre-order and in-order traversals are simple to understand and implement

Reference
Secure Dissemination of XML Content Using Structure-based Routing, *Ashish Kundu & Elisa Bertino, in the Proceedings of IEEE EDOC 2006.*

PURDUE UNIVERSITY

CERIAS

Discovery Park e-Enterprise Center