Redactable Signatures for Verification and Minimal Disclosure in Health Information Exchange

Doug Blough, Georgia Tech

Trust and Control in Health Information Exchange

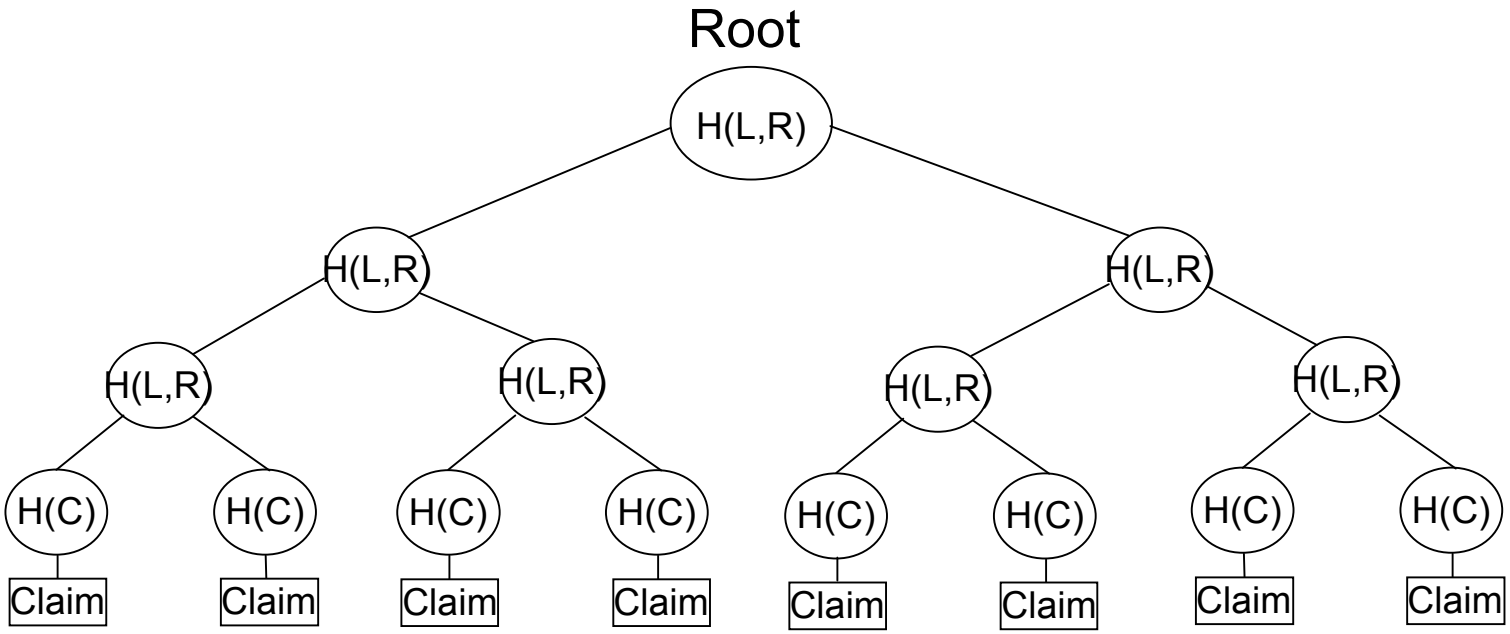
- HIE participants exchange patient info using HL7 standard Continuity of Care Documents (CCDs)
- CCDs are XML docs with standard format including sections for purpose, payers, advance directives, problems, family history, social history, alerts, medications, medical equipment, immunizations, vital signs, test results, procedures, etc.
- Example CCD from John Halamka available on his blog: http://geekdoctor.blogspot.com/
- CCDs are passed between providers with info filled in by different parties along the way

Trust and Control in HIE, continued

- ***Problems***
  1. How to know the source of particular info in the CCD and verify that it has not been modified by any other party?
  2. How to allow selective disclosure of CCD info to respect patient's privacy wishes?

- ***Motivating example***
  1. Parents send their daughter to religious summer camp, where they need to show proof of immunizations
  2. Parents get child's CCD including immunization record from their health care provider
  3. Parents forward CCD to camp but do not want to reveal any record of the child's HPV vaccination
  4. Camp receives partial CCD and needs to verify that it came from a recognized health care provider and was not modified

Redactable Signatures for Selective Disclosure



- sign root of hash tree
- release arbitrary subset of claims plus hash values necessary to reconstruct root hash
- Johnson, et. al., "Homomorphic Signature Schemes“, 2002.

Redactable Signatures for Health Info

- Multi-authority redactable signatures (*Bauer, Blough, and Cash, ACM CCS Workshop on Digital Identity Management, 2008*) allow a CCD to be signed by all providers that contribute to it

- Redactable signatures with disclosure dependencies (*Bauer, Blough, and Mohan, ACM CCS Workshop on Privacy in the Electronic Society, 2009*) allow providers to put limited constraints on how information can be hidden after they release a CCD

Redactable Signatures with Disclosure Dependencies

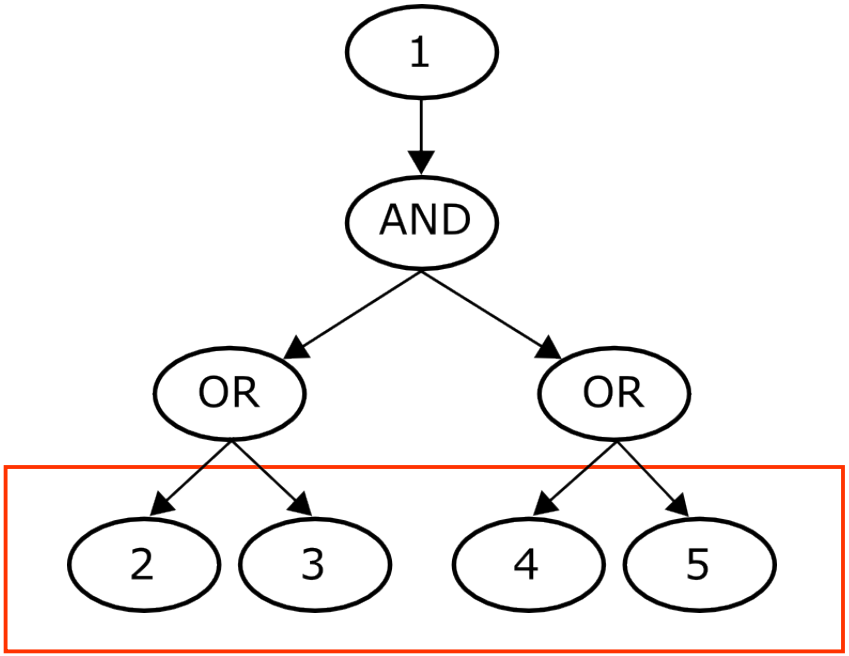
- Motivation: provider of signed record, e.g. health care provider, wants to enforce some release constraints
- Example: do not release a diagnosis without also releasing the test result on which it was based
- Example: do not release medical image without its metadata (however, might want to release metadata without image for searching/browsing)

Allowable Subsets

- Release policies may allow many options
  - "Release A only if also releasing B or C" gives 4 options: "AB", "AC", "B", and "C"
  - Adding the rule "Release B or C only if releasing D or E" gives 10 options

- Number of allowable subsets can be exponential (enumerating possibilities in hash tree is extremely inefficient)

Dependencies as Logical Expressions

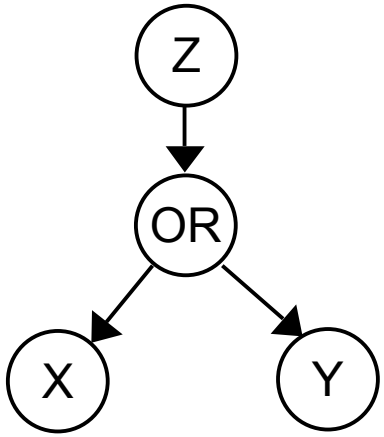


"1" cannot be released without also releasing either "2" or "3" along with either "4" or "5"

- Release policy is a **graph**
  - Each claim is a **node**
  - Each AND/OR is a node
  - No limit on fan-out or fan-in
- May have many top-level and bottom-level nodes
- Bottom (leaf) nodes are signed directly by a redactable signature
  - Other nodes are not

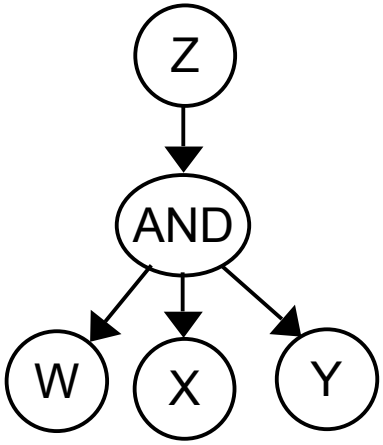
OR Dependencies

- Consider z → x OR y
- S(x) is called the string for node x
- S(x) = H(S(z) + x)
  - H is a hash function
  - "+" is concatenation
  - x is the actual data
  - S(z) is the string for node z
- S(y) = H(S(z) + y)
- S(z) = z

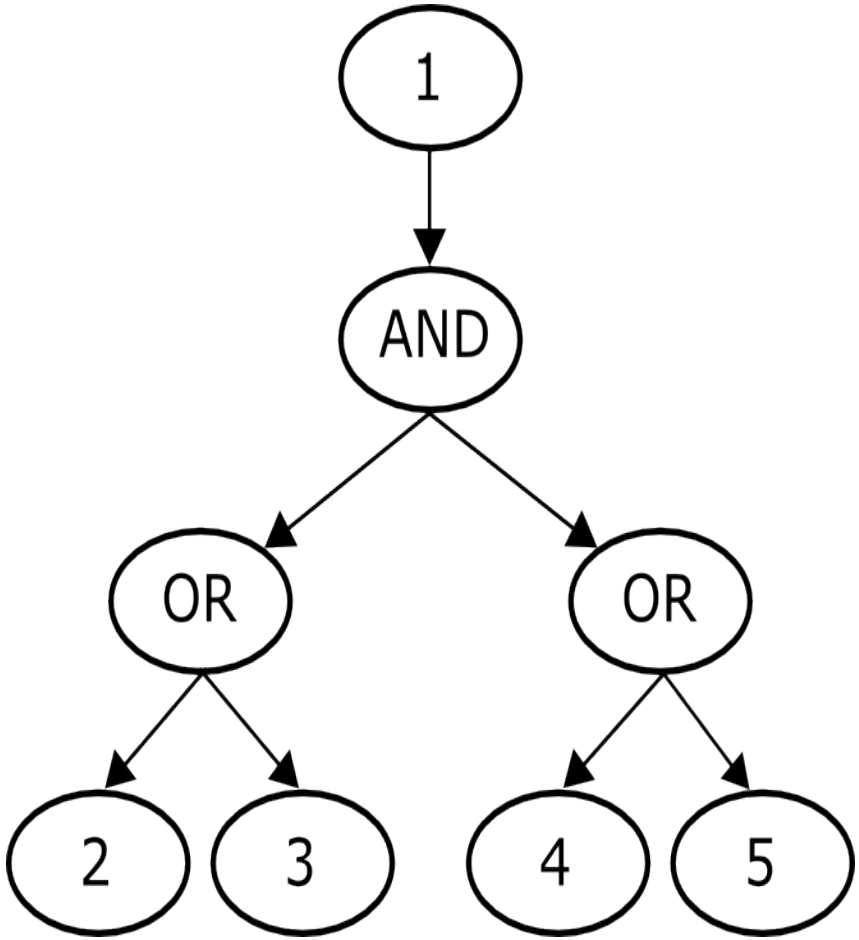


AND Dependencies

- OR Nodes disappear; AND nodes don't
- AND nodes handled with ***secret sharing***
- Consider z → w AND x AND y
- Generate random strings A1, A2
- S(AND) = H(S(z) + A1 + A2)
- A3 = S(AND) XOR A1 XOR A2
- S(w) = H(A1 + w)
- S(x) = H(A2 + x)
- S(y) = H(A3 + y)



Prior Example



- S(1) = 1
- A1 = random string
- S(AND) = H(S(1) + A1)
- A2 = S(AND) XOR A1
- S(2) = H(A1 + 2)
- S(3) = H(A1 + 3)
- S(4) = H(A2 + 4)
- S(5) = H(A2 + 5)

Other Forms of Constraints

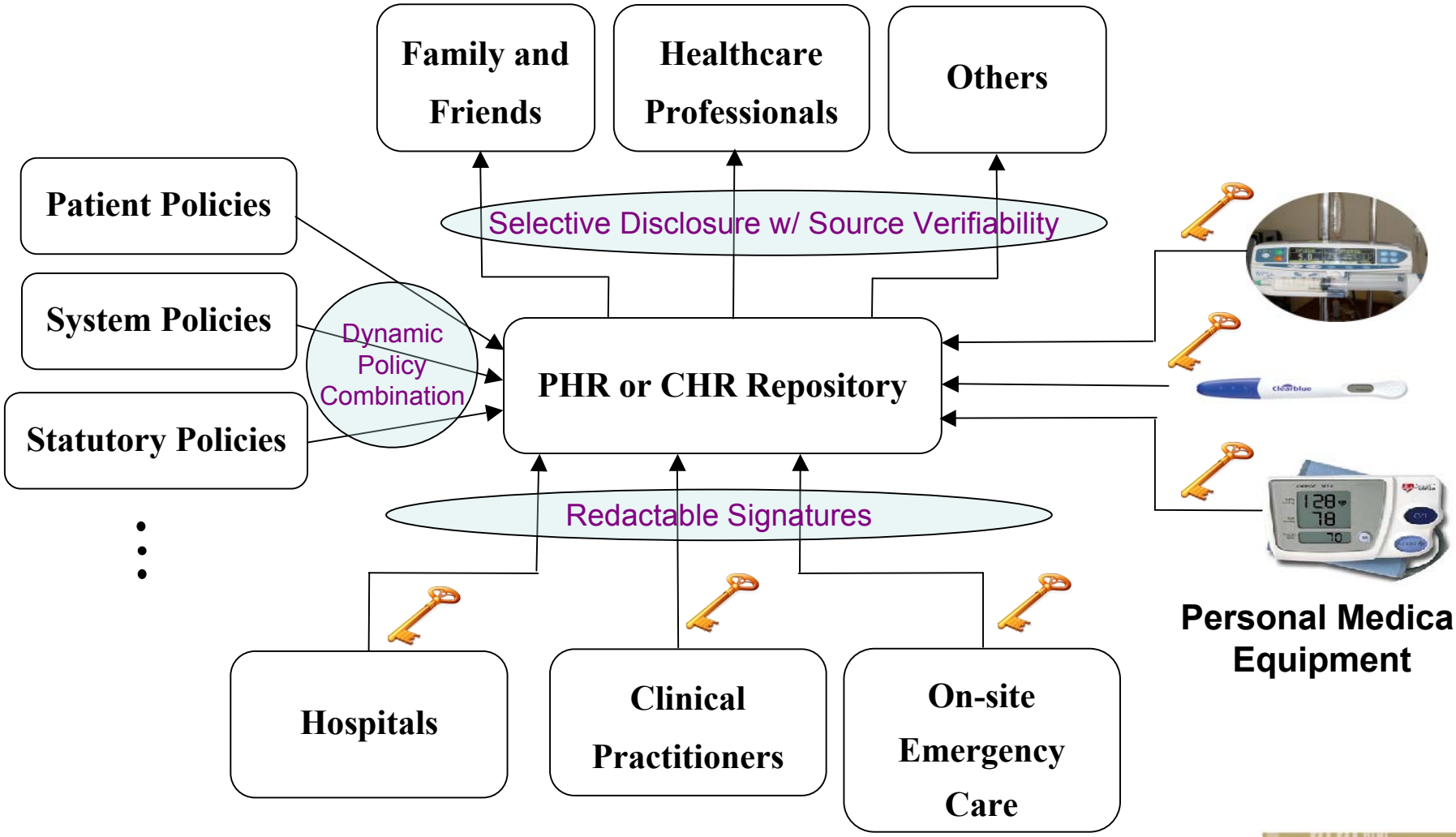
- Threshold functions, e.g. release A only if at least k out of n other claims are also released: can be handled by replacing xor secret sharing with Shamir's threshold secret sharing scheme

- Other functions - open problem (of course, can generate all allowable subsets and treat as a logical sum of products)

Other Application Areas

- Allow quoting from documents while preventing out of context quotes
- Verifying source tree signatures while ensuring that software module dependencies are honored (have done extensive testing on Ubuntu source trees)
- Any area where source signs data and wants to permit limited redaction once data leaves their sphere of control

A Patient-centric, Source-verifiable Health Records Repository



Questions??