

Lets suppose that a statement s belongs both to $MSG(t)$ and $MSG(u)$, where t and u are distinct statements. Then $MSG(s) = MSG(t)$ and $MSG(s) = MSG(u)$, so $MSG(t) = MSG(u)$, i.e. they are the actually the same MSG.

Corollary 1. An RDF model has an unique decomposition in MSGs.

This is a consequence of the fact that all the black nodes, in the MSG definition, are “properly surrounded” by actual URIs (or literals). As a consequence, a graph can be properly reconstructed between 2 peers by transferring and merging one or more MSG at a time.

Definition 4. The RDF Neighborhood (RDFN) of a resource is the graph composed by all the MSGs involving the resource itself.

It is straightforward to see that a graph can be transferred by moving the RDFN of all the involved URIs. Example MSGs and RDFN involving a resource are illustrated in image 1.

3. Signing MSGs

The MSG definition and properties highlighted in the previous sections say that it is possible to sign a MSG attaching the signature information to a single, arbitrary triple composing it. Along with the signature, an indication of the public key to use for verification might be provided. This indication is itself covered by the signing procedure. By “attach” we mean by using a reification procedure. Using the same procedure more signatures can be attached to the same MSG either independently or “layered” thus providing a mechanism for countersigning. The following example shows how a ground MSG (the triple <http://dbin.org/Home/Panaoli> dbin:student “Panaoli Fabio”) is signed by our implementation:

```
<rdf:RDF
  xmlns:dbin="http://dbin.org#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <rdf:Description rdf:about="http://dbin.org/Home/Panaoli">
    <dbin:student>Panaoli Fabio</dbin:student>
  </rdf:Description>
  <rdf:Description rdf:nodeID="A0">
    <rdf:predicate rdf:resource="http://dbin.org#student"/>
    <dbin:PGPCertificate>http://public.dbin.org/cont/238785872.asc</dbin:PGPCertificate>
    <dbin:Base64SigValue>MCwCFOPX....
A7xlaUgBzhkjcB5w==</dbin:Base64SigValue>
    <rdf:subject rdf:resource="http://dbin.org/Home/Panaoli"/>
    <rdf:object>Panaoli Fabio</rdf:object>
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
  </rdf:Description>
</rdf:RDF>
```

Given the MSG properties, this “information patch” can be merged into any existing model and the signature properties will be retained, checking the signature on any statement can be performed computing the MSG it belongs to (as this will not overlap with any triple previously or subsequently inserted in the model) and to check if any of the statements carry a MSG signature on it.

4. Supporting information revision in highly replicated P2P environments

In our RDFGrowth P2P semantic web model [4], peers synchronize the RDFN about URIs they're interested in with those coming from other peer in a fully monotonic (Ever growing

knowledge) model. This way of operating, which required the development of this theory, can be defined “resource centric” since a Peer is only interested about information strictly related to the requested URI, usually the collection of the MSGs involving it. Information so obtained is then passed and replicated by others who are specifically interested in the same URI or in any other involved by the MSG. The only connection between those who produce and consume information in this model is therefore the digital signature transferred along the MSGs.

Other than authenticating provenience, this methodology has been successfully used for providing capabilities of “information revision” in RDFGrowth. This not only supports trust at the client level by individually filtering MSGs from untrusted sources, but also allows information revision.

In short, once a MSG has been signed, the hash can be used as a IFP, that is, as a unique way to name to the MSG itself. This in turn can be used in a subsequent MSG to indicate the one that it substitutes. Given that the paternity of this subsequent MSG can be verified to be identical, the client can safely perform the information update, no matter where it received the update patch from.

5. Notes and conclusions

The RDFN definition is similar to the Concise Bound Description (CBD) as used in the URIQUA semantic web agent model [5], albeit more extended than the one that was available at the time when MSGs were first introduced in the RDFGrowth P2P algorithm. Recent modifications of the CBD have also addressed the case where IFP are used on bnodes and include reifications. The methodology presented here can be extended to encompass all this cases, although details cannot be included here.

Since this methodology uses reifications as a way to attach the signature to the MSGs, it is subject to all the shortcomings of this standard RDF construct. In particular, care should be used when using this proposed method in OWL FULL reasoners as the owl:sameAs property might cause substitutions inside MSGs. Given the digital signatures however, this change would immediately be detected and proper measures could be taken. Reification has also been often accused of being inefficient, that is, of causing “Triple bloat”. While this method does in fact see a consistent increase of triples when applied to very small MSGs (as in the previous example), this side effect becomes negligible as the MSG size grows, as only one statement needs reification.

This methodology has been implemented and is available as OS Java library. This library is also deployed in the SW P2P application DBin (www.dbin.org) where it provides the foundations mentioned above.

6. Acknowledgments

Our gratitude goes to Mauro Mazzieri for the theorem formalization, to Fabio Panaoli for the implementation and to Johan Johansson for the general support.

7. References

- [1] J.Carroll, "Signing RDF graphs", HP technical report 2003
- [2] J.Carroll, C. Bizer, P. Hayes, P. Stickler, "Named Graphs, Provenance and Trust", HP technical report 2004
- [3] E. Dumbill, "Signign FOAF files" [http:// usefulinc.com/foaf/signingFoafFiles](http://usefulinc.com/foaf/signingFoafFiles) personal communication
- [4] RDF Semantics, W3C Recommendation , 2004
- [5] G. Tummarello, C. Morbidoni, J. Petersson, P. Puliti, F. Piazza, "RDFGrowth, a P2P annotation exchange algorithm for scalable Semantic Web applications", P2PKM , Boston 2004
- [6] P.Stickler URIQUA The URI Query Agent Model, NOKIA 2003