IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 4, Issue 1, Feb - March, 2016
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

Trustworthy URI Enhancing The Reliability Of Data On The Web Adopting Base 64 Encryption

A.Vignesh, V.B.Sathya, M.Anand

Computer Science & Engineering, GKM College of Engineering & Technology,
Chennai, Tamil Nadu-600 063, India

Abstract

To make digital artifacts such as datasets, code, texts, and images verifiable and permanent. Furthermore there is no commonly accepted method to enforce this immutability of digital artifacts. To solve this problem, we propose trusty URIs containing base 64 encryption values. Base 64 encoding can be used when identifying a content which is used in HTTP background.

Keywords: Digital Artifacts, Base64, Encoding, Immutable, HTTP

1. Introduction

Data mining is the extraction of information from large databases which is al new technology which aids companies focus on the most distinct information from their data warehouses. Data mining derives its name from the correlation between searching business information from a large database — for example, finding linked data in gigabytes.

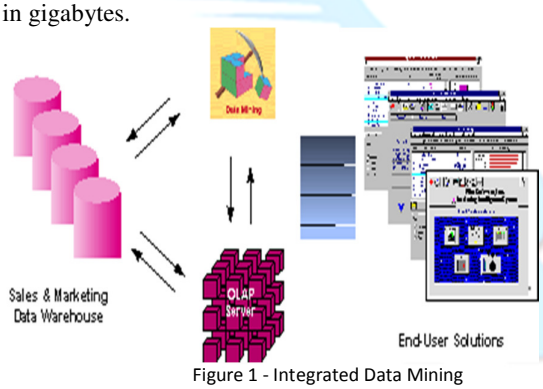

Figure 1 - Integrated Data Mining Architecture

To apply the advanced techniques, data must be fully integrated with a data warehouse as well as the business analysis tools.Some data mining tools which operate outside the warehouse require extra steps to extract and import. It requires operational implementation, Integration with the warehouse that corrects the application of results from data mining. The resulting analytic data warehouse can be applied in organization, in areas such as promotional campaign management, fraud detection, for advanced analysis in a huge data warehouse. This warehouse can be implemented in a variety of relational database systems should be optimized for flexible and fast data access. Growth of warehouse with new decisions and results, the organization can continuously implement the best practices and apply them to decisions.

It is used for maintaining the database with large set of collections which can be retrieved easily and enhances the productivity of information and is used to ensure that the user publishes the data through URI which is reliable or not.It is used to improve the certainty of data which is effective when reusing the data.

2. EXISTING SYSTEM

In many areas and in particular in science, reproducibility is important.

Verifiable, immutable, and permanent digital artifacts are the important elements for making the results of automated processes reproducible, but the current Web offers no commonly accepted methods to ensure these properties.

Endeavours such as the Semantic Web to publish complex knowledge in a machine-interpretable manner aggravates this problem, as automated algorithms operating on enormous amounts of data can be expected to be even more vulnerable than humans to manipulate or corrupted content.

Without appropriate countermeasures, malicious actors can destruct or trick such algorithms by adding just carefully manipulated items to large sets of input data

2.1PROBLEMS IN EXISTING SYSTEM

1. Data content corrupted by human beings.
2. in existing, no methods to make data content immutable.

3. PROPOSED

This approach includes cryptographic hash values in the Web URI's, especially acceptance and

decentralized architecture. Our proposed approach boils down to the idea that references can be verified if it contains a hash value of the referenced Digital artifact.
This method does not apply for every URIs, of course, but only to those which is to show a specific and immutable digital artifact.
We also propose trusty URI's for the web artifacts to be reliable and more secure.

3.1ALGORITHM

Base64 encoding is used to identify the information in an HTTP environment. For instance, a database persistence framework might use Base64 encoding to encode a relatively lengthy unique id into a string for use as an HTTP parameter in HTTP forms or HTTP GET URLs. Also, many applications need to encode binary data in a way that is convenient to be included in URLs or hidden web form fields, and Base64 is a convenient encoding to render them in a compact way.The algorithm which is used in this module is converts the ASCII value to base64 String which gives security to the data that is to be sent as a reliable data. This prevents from unauthorized decoding of data.

4. MODULE DISCRIPTION

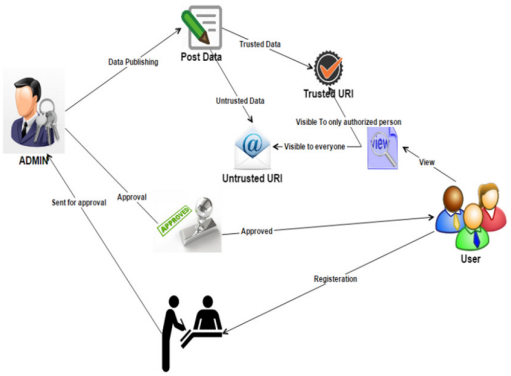

Figure 2 - Architecture

4.1Authentication & Authorization:

Authentication is a process in which the credentials provided are verified to those on file in a database of authorized users' information in a local database or within an authentication server. Authorization is the function of specifying access to resources related to information security and computer security in general and to access control.
In this project authentication is done to provide more security for the users to have their own credentials to log in. The admin approves the users who are registered and provide rights to login to the process.

4.2Cache of the data:

Cache which is widely used and very stable, but has not changed in years and is no longer actively developed.
The Cache is designed to assist a developer in persisting data for a specified period of time.
In this project, it is used as the collection of data to store which is used for various processing.

4.3Secured Distribution (Encoding & Decoding):

Encoding is the process of making a sequence of characters such as letters, numbers, punctuation, and certain symbols etc. into a specialized format for efficient transmission or storage. Decoding is the inverse process -- the conversion of an encoded format back into the original sequence of characters.

In Encoding, the data which are to be published is being encoded and it is being transformed into encoded values and stored it in the database. In Decoding, the converted data is being decoded back only if the valid user enters otherwise, it shows that you cannot access the file.

4.4Publishing the data:

Data publishing is the process of making the data available on the Internet, so that they can be accessed, analysed and reused by anyone for research or other purposes.
The data are being published where the appropriate level has the permission to access the file which is determined by the admin.

5. Data Flow

LEVEL1:

Login:

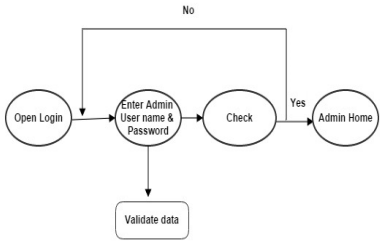

Figure 3-Login

LEVEL 2:

Admin:

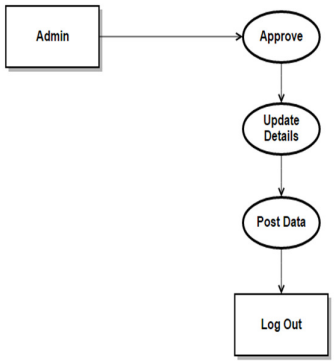

Figure 4-Admin

LEVEL 3:

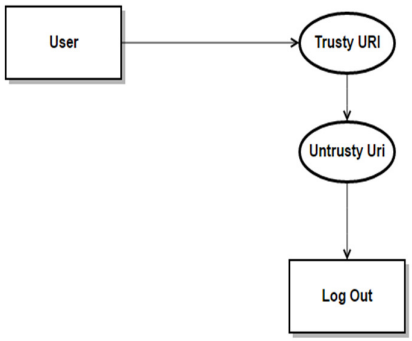

Figure 5-User

5. SCOPE OF THE PROJECT:

The project aims in providing the content of the data verifiable immutable and permanent so that it can be used in many areas particularly in science because reproducibility is important for making the results of automated process. It makes the contents of the data trustworthy which is sent as a URI to the user and it make sure whether it is trusted or not.It can be used mainly in research centres where scientists publish research papers the papers are not to be changed but to be reproduced and to verify that it is trusted data posted by the authorized person.

6. FUTURE WORK

To develop a decentralized Nano publications network.

In addition to that the theory of Nano publication indexes that allow for the definition as well as identification of small or large sets of Nano publications can be implemented.

7. CONCLUSION

We have given a proposal for explicit URI references to make digital artifacts on the (Semantic) Web to be verifiable, immutable as well as permanent. If adopted, it could have a considerable impact on the structure and functioning of the Web, could enhance the efficiency and accuracy of tools using Web resources, which becomes an important technical pillar for the Semantic Web, especially for scientific data, where provenance and verifiability are important.Further; we have started to develop a decentralized nanopublication server network. Nanopublications are distributed and replicated among such servers and identified by trusty URIs, thereby assuring that these artifacts remains accesable even if individual servers are terminated. In addition, we are working on the concept of nanopublication indexes that allow for the definition and identification of small or large sets of nanopublications.

8. ACKNOWLEDGEMENT

We would like to express our gratitude and greatest appreciation towards Assistant Professor M.Anand for giving us an opportunity to work under his guidance. We would also like to express our sincere thanks to IJREAT who provide us a chance for publishing this paper.

REFERENCE

[1] T. Kuhn and M. Dumontier, “Trusty URIs: Verifiable, immutable, and permanent digital artifacts for linked data,” in Proc. 11th Extended Semantic Web Conf., 2014, pp. 395–410.

[2] P. Growth, A. Gibson, and J. Velterop, “The anatomy of a nanopublication,” Inf. Serv. Use, vol. 30, no. 1, pp. 51–56, 2010.

[3] Farrell, Kutscher, C.Dannewitz, B.Ohlman, A.Keranen, and. Hallam-Baker, “Naming things with hashes,” Internet Engineering Taskforce(IETF),StandardsTrackRFC6920,Apr.2013.

[4] R. Hoekstra, "The MetaLex document server," in Proc. 10th Int. Conf. The Semantic Web, 2011, pp. 128–143.

[5] M. Altman and G. King, "A proposed standard for the scholarly citation of quantitative data," D-lib Mag., vol. 13, no. 3, p. 5, 2007
.

[6] M. Bartel, J. Boyer, B. Fox, B. LaMacchia, and E. Simon, (2008, Jun.). XML signature syntax and processing. W3C, Recommendation. [Online]. Available: http://www.w3.org/TR/xmldsig-core/

[7] J. Carroll, "Signing RDF graphs," in Proc. 2nd Int. Semantic Web Conf., The Semantic Web, 2003, pp. 369–384.

[8] E. H€ofig and I. Schieferdecker, "Hashing of RDF graphs and a solution to the blank node problem," in Proc. 10th Int. Workshop Uncertainty Reasoning Semantic Web, 2014, pp. 55.

[9] M. Bellare, O. Goldreich, and S. Goldwasser, "Incremental cryptography: The case of hashing and signing," in Proc. 14th Annu. Int. Cryptol. Conf., Adv. Cryptol., 1994, pp. 216–233.

[10] C. Sayers and A. Karp, "Computing the digest of an RDF graph," Mobile and Media Systems Laboratory, HP Laboratories, Palo Alto, USA, Tech. Rep. HPL-2003-235(R.1), 2004.

[11] R. Phan and D. Wagner, "Security considerations for incremental hash functions based on pair block chaining," Compute. Security, vol. 25, no. 2, pp. 131–136, 2006.