Journal of Engineering and Applied Sciences 12 (Special Issue 3): 6576-6580, 2017
ISSN: 1816-949X

Contributor-Content Verified for Establishing Trust and Privacy in Content-Centric Environment

Norliza Katuk, Hatim Mohamad Tahir, Mohd. Hasbullah Omar and Shahrudin Awang Nor
School of Computing, Universiti Utara Malaysia (UUM), 06010 Sintok, Kedah, Malaysia

Abstract: The emerging of Web 2.0 technology (e.g., Wiki, social networking sites, video-sharing sites and blogs) has allowed users to share their knowledge, experience and expertise with other unknown users in the cyberspace. Users share them by creating web content that is discoverable online. The landscape of computing and communication that rich with User-Generated Content (UGC) is known as the content-centric environment. The UGC is increasing exponentially which leads to trust and privacy issues. Content contributors are usually anonymous to keep their privacy. Unfortunately, it makes the reliability of information obtained from UGC and the credibility of the content's contributor always unknown. To solve this issue, we propose a Contributor-Content Verification (CCV) approach to achieve trust and privacy within the content-centric environment.

Key words: Discoverable content, digital content, digital identity, online content, tnternet of content, content-centric

INTRODUCTION

The emerging of Web 2.0 technologies has remodeled online information sharing where User-Generated Content (UGC) are more efficiently created and distributed compared to the web initial version (Daugherty *et al*., 2008). It includes content that appears on blogs, personal websites, wikis, social network sites, video-sharing sites, and others. The technology allows users to create contents that are discoverable by other users through a simple searching over the internet. This computing and networking landscape is also a synonym to content-centric environment (Passarella, 2012) information-centric environment (Bari *et al*., 2012) and Internet of Content (IoC) (Ziccardi, 2012). UGC is increasing exponentially as any users can contribute and access content through the web. Users share their profile, experience, knowledge and expertise with other unknown users. Therefore, this communication model is exposed to privacy and trust issues. Users who contribute contents will require revealing their personal information to establish trust on the content they created. Personal information disclosure could lead to losing the contributor's privacy.

Privacy and trust are always a tradeoff (Seigneur and Jensen, 2004). In other words to achieve trustworthiness of UGC, user's must give up their privacy and vice versa. Many UGC available on the Internet is anonymously authored because contributors are not willing to disclose their identity due to privacy issue. The main question arises from this scenario is "how reliable is a content and its contributor in this communication environment?" It leads to digital identity management perspective; particularly the needs for a mechanism or technique for establishing trust and at the same time protect contributor's privacy which has been the main objective of the research.

This study explains a conceptual representation of Contributor-Content Verified (CCV) approach that intends to establish trust on the content and contributor of discoverable UGC on the Internet and at the same time maintain the privacy of the content contributor. The next section explains in brief about privacy and trust in the content-centric environment. Then, it is followed by the conceptual design of CCV and its suggested implementation.

MATERIALS AND METHODS

Privacy and trust on the content-centric environment: The prominent role that the internet plays in global communication infrastructure has proven that it was a great success of a research experiment started more than 40 years back (Rexford and Dovrolis, 2010). However, the rapid growth of the Internet and the emergence of other internet-related technology have increased the number of challenges in security, mobility, reliability, performance and content-distribution (Pan *et al*., 2011). These issues

Corresponding Author: Norliza Katuk, School of Computing, Universiti Utara Malaysia (UUM), 06010 Sintok, Kedah, Malaysia

are vital to be addressed appropriately to ensure the sustainability of the communication infrastructure in the future. As a result a growing number of studies on the future internet have evolved recently including the security and privacy aspects of UGC within the content-centric environment.

Content-centric is a promising model for future internet that changes the communication approach from the host-based (using IP address) to data-centric paradigm (Abidi *et al*., 2012). As the variety of communication devices are now can be attached to a network, the current Internet infrastructure might not be able to accommodate this increase in the near future. Hence, identifying host for the purpose of information exchange might not be feasible. This communication model requires some changes to support the fast-growing information exchange demand. An extensible and scalable approach can be achieved through the content-centric model. It should also be able to support the fast growing and increasing volume of UGC.

The purpose of content-centric networking is to switch the role of machines (hosts) to content in the current communication model (Perino and Varvello, 2011). Rather than identifying a machine to get access to information, it recognizes and identifies the content itself as an entity in the network. To allow effective information exchange a single content must have a digital identity that describes basic information of it including unique identifier, contribut's credentials and the content ownership.

Recent studies have proposed some content-centric approaches such as DONA (Koponen *et al*., 2007) DHT-based solutions (Xing and Wang, 2010) and PSIRP (Lagutin *et al*., 2010). These approaches demonstrate practical implementation for storing and locating content from the technical view of communication infrastructure. However, it is important to note that anyone who has access to the Internet is free to contribute and access content. Hence, this has raised two important questions related to security and privacy issues; how trustworthy is a particular content and its contributor? and how to ensure that the privacy of content contributor is protected?

In the future internet model, trust and privacy are two important aspects that will appear in any usage scenario (Rooy and Bus, 2010). Any communication protocol designed using this approach should include privacy and trust model in the first place. Exploratory research on privacy and trust in the content-centric environment has not gained much attention among researchers. Limited studies were found in this area. A study by Weber *et al*. (2010) proposed an Identity and Access Management (IAM) Model for creating a trustworthy content-centric environment. This model utilizes encryption techniques for protecting the digital identity and creates a secure communication channel for information exchange. A study by Abidi *et al*. (2012) proposed an approach named privacy-aware content-centric networking using federated digital identity and identity contract. These two studies are an early attempt that provides a good insight towards encouraging other researchers to explore on a similar topic of this area.

Table 1: The entities of CCV

Entities	Definition
Contributor (c)	A user that creates content
User (u)	A user that consumes content
Digital Identity authority (IdP)	A server that holds contributor's identity
Content (co)	Information that is created by a contributor, consumed by users and stored by a service provider
Service provider (sp)	A server that hosts the content
Digital identity (d)	The profile information of the content's contributor

This has inspired us to explore other potential approaches for privacy protection of digital identity in the content-centric environment. The main theme of this research is "how to create a content-centric environment with an acceptable level of privacy and trustworthy". Specifically, this research is intended to study how to establish trust on UGC and how to protect content's contributor at the same time.

Contributor-content verified; the proposed approach: Privacy and trust are two vital components of secure and reliable communication. In the content-centric environment where anyone can be a content contributor, establishing trust is a great challenge. The reliability of the source and the content is difficult to be identified; hence requiring an identifier. Although, content may have an identifier it does not necessarily giving true information, unless the contributor is known and trusted. In order to be a trusted contributor, one might need to reveal some personal identity together with the content where a violation of privacy occurred. In this model of communication, we can see that we might need to give up privacy to obtain a trustworthy communication model. On the other hand, trustworthy is jeopardized when the model chooses to have contributor privacy protection.

With a view to achieve both privacy and trust, this research uses a simple approach towards establishing both components in content-centric communication model. We initially named this approach as Contributor-Content Verified (CCV). The main entities of this approach are the contributor (c), the user (u), the digital Identity authority (IdP) the content (co), the service provider (sp), the digital identity (d) and the peer-group (optional). Table 1 provides the definition of each entity.

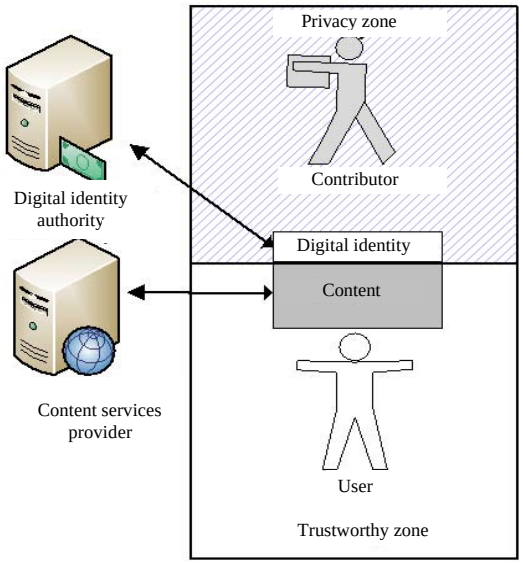

Fig. 1: The communication model of CCV approach

In Fig. 1, we illustrate the communication model for CCV that contains the entities and communication protocol of this approach. To create a trustworthy environment, content must have a verified digital identity attached to it. The digital identity is managed by a digital identity authority that provides valid and authenticated lists of content contributors. In terms of protecting the privacy of the contributors, only necessary information is provided in the digital identity as approved by the contributors. As long as the digital identity is verified by the authority, the content is trusted. IdP also performs peer verification for validating contributor's background if UGC involves professional content. Through this approach, both privacy and trust could be established.

In Fig. 1, a trustworthy and privacy zone is created for the user who accesses the discoverable content and the contributor of the content, respectively. By incorporating a verified digital identity into the content, sufficient support is given to prove that the contributor of the content is reliable and credible. Further, the digital identity will not disclose much personal information about the contributor, hence, protecting his/her privacy. Hence, both privacy and trust can be achieved at the same time for UGC.

The CCV communication model comprises of three main processes; enrollment, service request and digital identity verification. Sequence diagrams in Fig. 2-4 present the process respectively. In Fig. 2, the enrollment process starts when a contributor sends a request to IdP to enroll (to create a digital identity) for the service.

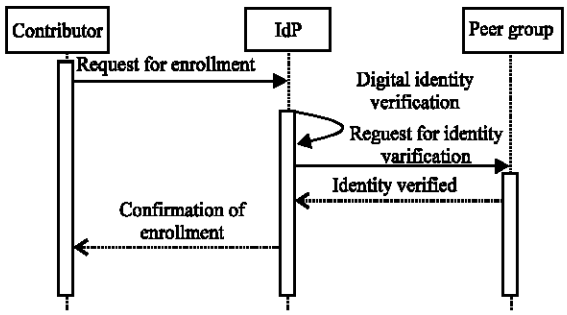

Fig. 2: A sequence diagram for enrollment

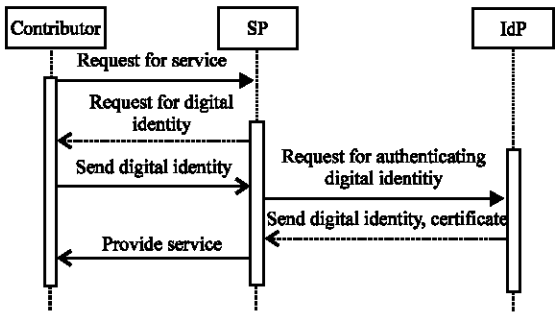

Fig. 3: A sequence diagram for service request

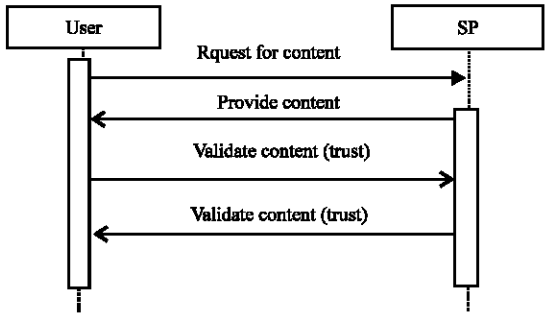

Fig. 4: A sequence diagram for digital identity verification

Verification can be done either by the IdP itself or through peer group. Verification of the contributor's profile through peer group is an optional process if the professional profile is required.

In Fig. 3, a contributor will request service from SP. In this context, SP could be any service provider that provides hosting for UGC. SP will request the contributor's digital identity. Then the SP will communicate with the IdP to authenticate the digital identity. Once authenticated, the IdP will send the contributor's digital identity and certificate to SP. The user communicates with SP to request for content as shown in Fig. 4. The user can validate the trustworthiness

of the content by communicating with the SP. As the contributor's digital identity is attached to the content, SP can verify the identity of the contributors when requested.

RESULTS AND DISCUSSION

Implementation: The conceptual representation of CCV explained in the previous chapter could have different ways of implementation. In this study, we discuss our implementation of CCV approach at user interface layer of UGC technology. We give priority to user interface design because it is the most important component in the communication model between a user and discoverable UGC. The physical representation of the user interface is the point where users justify the reliability of a particular content. We limit our discussion here to the implementation of CCV on the design of the user interface only.

In this example of implementation, consider a contributor who contributes text content and publishes the content on a blog. When a user searches content through a search engine using the appropriate keyword, the search engine crawls for the content and show the result to the user. As the user selects the result, the browser renders the content in which is hosted by the SP (i.e., blog hosting provider). In the design of the user interface, we use a simple html button that sends http request to IdP for verifying the identity of the contributor. Figure 5 shows an example of the user interface.

Behind this interface, it contains an event handler associated with the button that will call the digital identity of the contributor. Once the client (user's browser) receives instruction through a button click it will send the http request (with the digital identity as the attribute) to the SP to verify the identity of the contributor. The digital identity is encrypted using asymmetric key encryption. Once the identity is verified, the confirmation message is rendered to the user. The confirmation message contains simple notification window with a minimum amount of the contributor's personal information disclosed. Figure 6 shows the example of the confirmation message.

Prior to the process, the contributor must create a digital identity with IdP. IdP in this instance could be any providers including social network sites that provide social login facility and single sign-on services such as Facebook, Google and Yahoo!. The past studies investigated the use of social network credentials as a single sign-on mechanism from user's behavioral perspectives (Katuk *et al*., 2015ab). The social network credentials are seen providing a good potential for the contributor's digital identity. We will discuss this implementation in other studys.

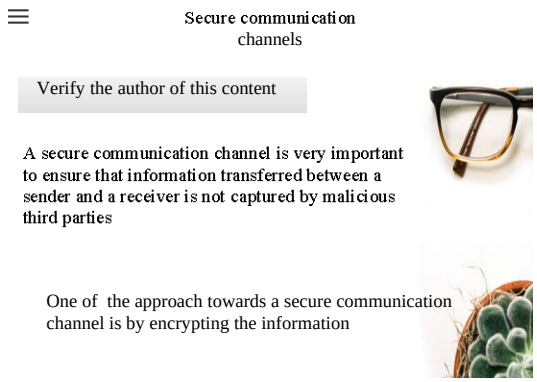

Fig. 5: An example of a user interface for CCV implementation

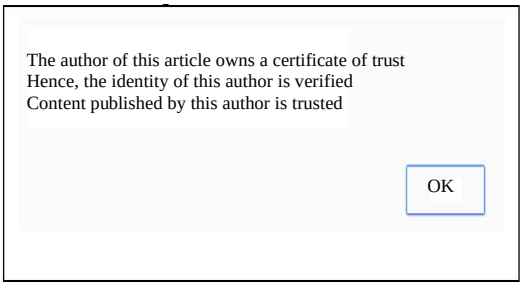

Fig. 6: An example of a verification message of Fig. 5 and 6. Contributor's digital identity

CONCLUSION

Privacy and trust are two important aspects for users in contributing and consuming content from the Internet. However, they are always a tradeoff of each other. In other words, users must reveal their personal information to create a trustworthy content and the other way round. In this study, we discussed an approach known as Contributor-Content Verified (CCV) approach to improve trust and privacy issues in the content-centric environment. In this approach, we impose a verified digital identity to protects contributor's privacy and embeds it in the content. The verified digital identity is an entity that confirms contributo's identity and reliability.

Currently, we study the implementation of CCV approach and measure the applicability of the approach in various instances of UGC. We also intend to study how users perceive trustworthiness of UGC when CCV is embedded in different user interface designs.

ACKNOWLEDGEMENT

This research was funded by Research Acculturation Grant Scheme (RAGS), Ministry of Higher Education of Malaysia (UUM S/O Code: 12680).

REFERENCES

Abidi, A., G.B. Ayed and F. Kamoun, 2012. Towards constructing a trustworthy internet: Privacy-aware transfer of digital identity document in content centric internetworking. Proceedings of the International Conference on Security in Computer Networks and Distributed Systems, (SNDS'12), October 11-12, 2012, Springer, Trivandrum, India, pp: 85-96.

Bari, M.F., S.R. Chowdhury, R. Ahmed, R. Boutaba and B. Mathieu, 2012. A survey of naming and routing in information-centric networks. IEEE. Commun. Mag., 50: 44-53.

Daugherty, T., M.S. Eastin and L. Bright, 2008. Exploring consumer motivations for creating user-generated content. J. Interact. Advertising, 8: 16-25.

Katuk, N., H.M. Tahir, N.H. Zakaria and M.S. Halim, 2015a. Can Single Sign-on Improve Password Management? A Focus Group Study. In: Pattern Analysis, Intelligent Security and the Internet of Things, Abraham, A., A.K. Muda, Y.H. Choo (Eds.). Springer, Cham, Switzerland, ISBN:978-3-319-17397-9, pp: 85-93.

Katuk, N., K.L. Chun, N.H. Zakaria, H.M. Tahir and M.S. Halim, 2015b. Authenticate yourself once using OpenID. J. Inf. Assur. Secur., 10: 139-151.

Koponen, T., M. Chawla, B.G. Chun, A. Ermolinskiy and K.H. Kim *et al*., 2007. A data-oriented (and beyond) network architecture. ACM. SIGCOMM Comput. Commun. Rev., 37: 181-192.

Lagutin, D., K. Visala and S. Tarkoma, 2010. Publish-Subscribe for Internet: PSIRP Perspective. In: Towards the Future Internet, Tselentis, G. and A. Galis (Eds.). IOS Press, Amsterdam, Netherlands, pp: 75-84.

Pan, J., S. Paul and R. Jain, 2011. A survey of the research on future internet architectures. IEEE. Commun. Mag., 49: 26-36.

Passarella, A., 2012. A survey on content-centric technologies for the current Internet: CDN and P2P solutions. Comput. Commun., 35: 1-32.

Perino, D. and M. Varvello, 2011. A reality check for content centric networking. Proceedings of the 2011 ACM SIGCOMM Workshop on Information-Centric Networking, August 19, 2011, ACM, Toronto, Ontario, ISBN:978-1-4503-0801-4, pp: 44-49.

Rexford, J. and C. Dovrolis, 2010. Future internet architecture: Clean-slate versus evolutionary research. Commun. ACM., 53: 36-40.

Rooy, V.D. and J. Bus, 2010. Trust and privacy in the future internet: A research perspective. Identity Inf. Soc., 3: 397-404.

Seigneur, J.M. and C. Jensen, 2004. Trading privacy for trust. Proceedings of the 2nd International Conference on Trust Management (iTrust'04), March 29-April 1, 2004, Springer, Oxford, UK., pp: 93-107.

Weber, S.G., L.A. Martucci, S. Ries and M. Muhlhauser, 2010. Towards trustworthy identity and access management for the future internet. Proceedings of the 4th International Workshop on Trustworthy Internet of People, Things and Services (IoPTS'10) Vol. 29, November 29-December 1, 2010, Royal Park Hotel, Tokyo, Japan, pp: 1-8.

Xing, C. and H. Wang, 2010. DHT-Based namespace solution for grids. Proceedings of the 2010 IEEE 2nd International Conference on Information Engineering and Computer Science (ICIECS'10), December 25-26, 2010, IEEE, Wuhan, China, ISBN:978-1-4244-7939-9, pp: 1-3.

Ziccardi, G., 2012. Resistance, Liberation Technology and Human Rights in the Digital Age. Vol. 7, Springer, Berlin, Germany, ISBN:978-94-007-5275-7, Pages: 327.