On Privacy and Anonymity in Knowledge Externalization

Yuen-Yan Chan and Chi-Hong Leung
The Chinese University of Hong Kong
rosannachan@cuhk.edu.hk ,
leung_chi_hong@yahoo.com.hk

Secure Knowledge Management (SKM 2004)
24 September 2004
Buffalo USA

Outline

- Knowledge Externalization
  - Approaches for Knowledge Externalization
  - Privacy Issues in Knowledge Externalization
- Solution to Anonymity
  - Network-Layer Solutions
  - Use of Alias
  - Zero-knowledge Proofs
- Our Protocol
- Security Analysis
- Efficiency

Knowledge Externalization

- Knowledge management includes the capability to collect, archive, manage, evaluate, and distribute knowledge across an organization
- Tacit knowledge – "internal knowledge", personal beliefs, perspective and value
- Explicit knowledge – exists in form of readable documents, records
- Knowledge Externalization: the transformation of tacit knowledge to explicit knowledge

Approaches for Knowledge Externalization

- Knowledge externalization involves the harnessing of tacit knowledge to explicit knowledge.
- Typical ways of achieving it includes:
  - postings on an electronic discussion board
  - the uploading of one's own records
  - generating reports on one's experience
  - recording of meetings, interviews and phone calls
- usually involve the sharing and exposing of one's internal knowing

Privacy Issues in Knowledge Externalization

- Technically, achieving knowledge externalization requires logging-on to a system with authorized identity
- However, to log-on to knowledge management systems using the true identity is often subjected to the following threats:
  - erroneous and obtrusive requests for information
  - botheration by potential information seekers
  - traceability of the information providers (for example, when an employee is speaking out against the management)
- Therefore, it is desirable that the source of information is *hidden* but *authorized* at the same time
- Solution: Anonymous Authentication

Solutions to Anonymity

- Unconditional anonymity, or complete anonymity (in which no authentication is required and the users are freely login to a system), hinders knowledge exchange as the source of information cannot be traced
- Also, such anonymity may be abused as one is not responsible for his or her actions
- Therefore, what we are interested is a more challenging solution: *conditional anonymity* (hereinafter refer to as “anonymity”).
- Three main approaches to provide anonymity: network-layer approach, use of alias, and zero-knowledge proofs approach

Solution to Anonymity – Network Layer Approach

- Achieve anonymity in the network layer
- A user's action is hidden within the actions of many others
- Examples:
  - Crowds (Reiter and Rubin, 1997)
  - MIX (Chaum, 1981)
  - Onion Routing (Syverson, Goldschlag, and Reed, 1997)

Solution to Anonymity - Alias

- Use alias, or a pseudonym, to substitute for the true identity of an entity
- Usually, a pre-authentication phrase prior to the authentication is required
  - in pre-authentication phrase, the entity proves its identity to a server using conventional authentication methods
  - afterward, the server randomly generates a number which is uncorrelated to the true identity of the entity and digitally signs on it
- Examples:
  - Pseudonym Systems proposed by Lysyanskaya et al. (Lysyanskaya, Rivest, Sahai, and Wolf, 1999)
  - Temporary identity (TID) to achieve anonymity in wireless communication systems (Go and Kim, 2001)

Solution to Anonymity – Zero Knowledge Proofs

- Is an interactive proof with a prover P and a verifier V
- P convinces V of the knowledge of a secret, without revealing any information about the secret or how to go about proving this secret
- Examples:
  - Electronic Cash proposes by David Chaum (Chaum, Fiat, and Naor, 1988)
  - Non-transferable electronic voting pass proposed by Chan et al. (Chan, Wong and Chan, 2000)
  - Our Protocol

Our Protocol - Preliminary

- A – User
- B – Knowledge management server
- C - Central repository

- (d, e) – be the private-public key pair of B
- n = p•q where p and q are two large prime numbers. d•e = 1 mod (p-1)(q-1)
- u – the identity of A
- a, b, c and r – some random integers
- k – security parameter
- f() and g() – two-argument one-way collision free hash functions
- R – a set, and R' be mutually exclusive to R'.

Our Protocol – Pre-authentication

- In this phrase, A obtains an anonymous pass from B and deposits the identity revocation value to C
- We employ the blind signature technique (Chaum, 1983) in the generation of the anonymous pass
- A pass signed by B is produced
- Please refer to paper section 4.2 for details

Our Protocol - Authentication

- When *A* logins to the knowledge management server *B*, *A* and *B* perform authentication
- To do this, *A* presents the pass that it obtained from the Pre-Authentication phase
- B undergo random challenges on the pass to check A's knowledge on the pass
- Please refer to paper section 4.3 for details

Our Protocol - Revocation

- In case *A* misbehaves, the revocation phase can be executed to revoke the identity of *A*.
- This phase is performed by *B* and *C*.

Security Analysis

- Authenticity
  - B authenticates A based on the pass verification
- Anonymity
  - A's identity is protected throughout the process
- Masquerade Prevention
  - During the pre-authentication phrase, A is not anonymous and B should make sure A's identity before signing on the pass
- Chance of Successful Cheating by A
  - $1/(2^{k/2})$, which decreases exponentially with the value of the security parameter k
- Stolen Pass
  - Suppose the pass is stolen by eavesdropper E during pre-authentication phrase, this will not bring any lost to A because E does not know the values of a_i, b_i, c_i,and a_i XOR u.

Efficiency

- Most operations involved in our protocol are hashes and random number generation
- Hashes are light in terms of computational power (O'Mahony, Peirce, and Tweari, 1997).

Conclusion

- We have studied the privacy issues involved in knowledge externalization
- In order to protect the privacy of the knowledge source, we have proposed a solution for conditional anonymous authentication
  - in which one can login to a knowledge management system anonymously, while the identity will be revoked conditionally
- We have also reviewed the methods for providing anonymity in general
- Security and efficiency analysis of our proposed protocol is also provided

Reference

- Y. Chan, J. C. Wong, and A. C. Chan (2000) Anonymous Electronic Voting System with Non-Transferable Voting Passes. In proceedings of SEC 2000, pp.321-330.
- D. Chaum. (1981) Untraceable electronic mail, return addresses, and digital pseudonyms, Communications of the ACM.
- D. Chaum. (1983) Blind Signature for Untraceable Payment. Advances in Cryptology – Proceedings of CRYPTO'82, pp.199-203.
- D. Chaum, A. Fiat and M. Naor. (1988) Untraceable Electronic Cash. Advances in Cryptology - Proceedings of Crypto '88, pp. 319-327.
- R. Hauser and G. Tsudik. (1996) On Shopping Incognito. Proceedings of the 2nd USENIX Workshop on Electronic Commerce, pp.251-7.
- J. Go and K. Kim. (2001) Wireless authentication protocol preserving user anonymity. In Proc. of the 2001 Symposium on Cryptography and Information Security (SCIS 2001).
- R. Molva, D. Samfat, and G. Tsudik. (1994) Authentication of Mobile Users. IEEE Network, Special Issue on Mobile Communications, March/April 1994, pp. 26-34.
- G. Ng-Kruelle, P. A. Swatman, D. S. Rebne and J. F. Hampe. (2002) The Price of Convenience: Privacy and Mobile Commerce. Proceedings of the 3rd World Congress on the Management of Electronic Commerce.
- A. Lysyanskaya, R. Rivest, A. Sahai, and S. Wolf (1999) Pseudonym Systems. Proceedings of Selected Areas in Cryptography 99, Lecture Notes in Computer Science, Springer-Verlag, vol. 1758, pp. 184-199.
- D. O'Mahony, M. Peirce, H. Tewari (1997) Electronic Payment Systems. Artech House, 1996.
- I. Nonaka and J. Reinmoeller. (1998) The ART of Knowledge: Systems to Capitalize on Market Knowledge, European Management Journal, vol. 16, no. 6, pp. 673-684.
- M. Polanyi. (1966) The Tacit Dimension. London: Routledge & Kegan Paul
- M. K. Reiter and A. D. Rubin (1997) Crowds: Anonymity for web transactions. Technical Report 97-15, DIMACS
- E. Sallis and G. Jones. (2002) Knowledge Management in Education: Enhancing Learning & Education, London: Kogan Page.
- A. L. Schirmer. (2003) Privacy and Knowledge Management: Challenges in the Design of the Lotus Discovery Server, IBM Systems Journal, vol 42, no 3, pp. 519-531.
- P. F. Syverson, D. M. Goldschlag, and M. G. Reed. (1997) Anonymous Connections and Onion Routing. Proceedings of the 18th Annual Symposium on Security and Privacy, pp. 44-54.