

What are the optimum quasi-identifiers to re-identify medical records?

Yong Ju LEE*, Kyung Ho LEE*

**School of Information Security, Korea University, Korea*

sky4uni@korea.ac.kr, kevinlee@korea.ac.kr

Abstract—Recently, medical records are shared to online for a purpose of medical research and expert opinion. There is a problem with sharing the medical records. If someone knows the subject of the record by using various methods, it can result in an invasion of the patient's privacy. To solve the problem, it is important to carefully address the tradeoff between data sharing and privacy. For this reason, de-identification techniques are applicable to address the problem. However, de-identified data has a risk of re-identification. There are two problems with using de-identification techniques. First, de-identification techniques may damage data utility although it may decrease a risk of re-identification. Second, de-identified data can be re-identified from inference using background knowledge. The objective of this paper is to analyze the probability of re-identification according to inferable quasi-identifiers. We analyzed factors, inferable quasi-identifiers, which can be inferred from background knowledge. Then, we estimated the probability of re-identification from taking advantage of the factors. As a result, we determined the effect of the re-identification according to the type and the range of inferable quasi-identifiers. This paper contributes to a decision on de-identification target and level for protecting patient's privacy through a comparative analysis of the probability of re-identification according to the type and the range of inference.

Keyword—Privacy, Re-identification, De-identification, Medical records



Yong Ju LEE, was born in Republic of Korea, October 7, 1989. Yong Ju Lee earned Master's degree from School of Information Security at Korea University. His main research interests include risk management, privacy policy, de-identification and re-identification of personal information.



Kyung Ho LEE, was born in Republic of Korea, September 9, 1967. Kyung Ho Lee earned his Ph.D. degree from Korea University. He is now a Professor in School of Information Security at Korea University, and leading the Risk management Laboratory in Korea University since 2011. He was the former CISO in Naver corporation and CEO of Secubase corporation. His main research interests include information security management system(ISMS), risk management, information security consulting, privacy policy, and privacy impact assessment(PIA).