Private-Preserving Encoding and Decoding Using Variable-Length Coding Schemes

Yuki Seto Kunihiko Sadakane Kazunari Tozawa

Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo
sitoo@g.ecc.u-tokyo.ac.jp

Secure computation is a technology that provides calculations on data while keeping it encrypted, which reduces the risk of information leakage. In this paper, we propose efficient secure protocols for oblivious encoding and decoding using variable-length codes, i.e., encoding and decoding encrypted data without decryption.

Functionalities of Encoding and Decoding. For a variable-length code $C : X \to \{0,1\}^*$, we define the functionalities of oblivious encoding and decoding using C as follows:

Functionalities of Encoding and Decoding

- $(\llbracket \boldsymbol{w} \rrbracket, \llbracket l \rrbracket) \leftarrow \mathsf{Encode}(\llbracket \boldsymbol{x} \rrbracket, l_w)$: Given the ciphertext $\llbracket \boldsymbol{x} \rrbracket$ of the message $\boldsymbol{x} \in X^*$ and an upper bound of the length l_w of the encoded string as input, output the ciphertext $\llbracket \boldsymbol{w} \rrbracket$ of the encoded string $\boldsymbol{w} \in \{0,1\}^*$ of $\boldsymbol{x}$ and the ciphertext $\llbracket l \rrbracket$ of the value $l \coloneqq |\boldsymbol{w}|$.
- $\llbracket \boldsymbol{x} \rrbracket \leftarrow \mathsf{Decode}(\llbracket \boldsymbol{w} \rrbracket, l_x)$: Given the ciphertext $\llbracket \boldsymbol{w} \rrbracket$ of the encoded string $\boldsymbol{w}$ and the number of characters l_x of the original message as input, output the ciphertext $\llbracket \boldsymbol{x} \rrbracket$ of the original message $\boldsymbol{x}$.

Generic Methods for Encoding and Decoding. We first propose two protocols for encoding and decoding of variable-length prefix codes using lookup tables. The proposed protocols, denoted as Π_{Encode} and Π_{Decode}, respectively, realize the functionalities Encode and Decode. Both Π_{Encode} and Π_{Decode} are constructed under a model of secure computation schemes that contains several basic operations. The security of our protocols is derived from the Universally Composable security framework [1].

The protocol Π_{Encode} can also be directly applied to encoding using non-prefix codes. Furthermore, these methods can be adapted to make certain succinct data structures secure, such as Fully Indexable Dictionaries [3].

Decoding Protocols for Unary and Gamma Codes. In addition to generic protocols, we also propose efficient and specialized decoding protocols for Unary codes and Gamma codes [2]. These codes explicitly encode integers in the original message by the length of the codewords or their binary representations. Our specialized protocols use this feature to achieve decoding without using lookup tables, thereby making the protocols more efficient.

References

[1] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *42nd Annual Symposium on Foundations of Computer Science*. Las Vegas, NV, USA: IEEE Computer Society Press, Oct. 14–17, 2001, pp. 136–145.

[2] P. Elias, "Universal codeword sets and representations of the integers," *IEEE transactions on information theory*, vol. 21, no. 2, pp. 194–203, 1975.

[3] R. Raman, V. Raman, and S. R. Satti, "Succinct indexable dictionaries with applications to encoding k-ary trees, prefix sums and multisets," *ACM Transactions on Algorithms (TALG)*, vol. 3, no. 4, pp. 43–es, 2007.