On Certain Problems of Cryptology and Computational Complexity of Processing over Uncertain Data

Miodrag J. Mihaljević, Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia and Chuo University, Tokyo, Japan
Hideki Imai, Chuo University, Tokyo, Japan

Abstract. A link between certain problems of cryptology and mathematical logic is pointed out. Computational complexity of the Learning Parity in Noise (LPN) problem is discussed as an example of the reasoning over uncertain data.

Introduction

This paper links a real-life problem of growing importance and a topic of mathematical logic. We consider certain issues of security within cyber-space and a topic of mathematical logic dedicated to reasoning over uncertain data and corresponding computational complexity. Accordingly, an application of the topics of mathematical logic to cryptology is pointed out. Section 2 summarizes some emergency issues of cyber-security which imply request for employment of low-complexity cryptographic techniques. A mathematical problem called Learning Parity in Noise (LPN) relevant for design of low-complexity and highly secure cryptographic primitives is summarized in Section 3. Finally, section 4 addresses some issues of the LPN problem complexity which are also challenges regarding reasoning over uncertain data.

Preliminaries

Overheads Implied by Cryptographic Techniques. Our society strongly depends on information-communications technologies (ICT) and the security of ICT has been recognized as one of the top priorities in order to minimize impacts of potential attempts regarding misuse of ICT with disastrous consequences. Accordingly, we face an extensive employment of the security mechanisms and as a consequence we face significant overheads to the main functionality of the systems implied by the employed security mechanisms. Reduction of these security related overheads is of a top interest because cumulative effect of all these overheads has a (very) significant cost. A part of these overheads corresponds to the cryptographic techniques employed in the security mechanisms. Accordingly, reduction of the security overheads implied by cryptographic algorithms appear as an issue of very high importance. On the other hand, minimization of the "cryptographic overheads" should not jeopardize the cryptographic security, and design of highly secure cryptographic algorithms and protocols which minimize the overheads is still a challenge and an emergency issue.

Lightweight and Provably Secure Cryptographic Primitives. The main overheads implied by cryptographic techniques correspond to: (i) implementation overheads (required additional software/hardware); (ii) computational overheads for performing cryptographic operations; (iii) power-consumption overheads regarding cryptographic processing. Cryptographic techniques which provide minimization of the overheads are called light-weight cryptographic techniques. On the other hand side, a claim that a cryptographic primitive is provably secure means that assumption of its insecurity implies that certain hard mathematical problem can be solved (employing certain algorithm for cryptanalysis) implying a contradiction and a justification of the security. Note that When instead of a provably secure construction a heuristically secure approach is employed we could face iterative improvements of exploring the vulnerabilities with

serious security consequences (as an illustration see [7]-[9])

Learning Parity in Noise (LPN) Problem

LPN problem has been recognized as an underlying approach for constructions of light-weight and provably secure cryptographic primitives (see [10], for example). Informally, the LPN problem a problem of solving a probabilistic overdefined consistent system of linear equations over GF(2) where the right side of each equation is true with the known probability $p > 1/2$ (typically $p < 0.25$). One of its incarnations is the problem of decoding of a random linear binary block code.

Definition: LPN Search Problem. Let s be a random binary string of length l. We consider the Bernoulli distribution $\mathcal{B}_\theta$ with parameter $\theta \in (0, 1/2)$. Let $\mathcal{Q}_{s,\theta}$ be the following distribution:

$$\{(a, < s, a > \oplus e)| \leftarrow \{0,1\}^l, e \leftarrow \mathcal{B}_\theta\} .$$

For an adversary $\mathcal{A}$ trying to discover the random string s, we define its advantage as

$$\mathrm{Adv}_{\mathrm{LPN}_{\theta,\mathcal{A}}}(l) = \Pr[\mathcal{A}^{\mathcal{Q}_{s,\theta}} = s | s \leftarrow \{0,1\}^l] .$$

The LPN$_\theta$ problem with parameter θ is hard if the advantage of adversaries $\mathcal{A}$ that make a polynomial number of oracle queries is negligible.

In [5] a distinguishing variant of the problem has been introduced, which is more useful in the context of encryption schemes. Roughly speaking, the decisional LPN problem asks to distinguish a number of noisy samples of a linear function (specified by a secret vector x) from uniform random. The problem is, given $\mathbf{A}$ and $\mathbf{y}$, to decide whether $\mathbf{y}$ is distributed according to $\mathbf{A} \cdot \mathbf{x} \oplus \mathbf{e}$ or chosen uniformly at random.

Definition: LPNDP - LPN Decisional (Distingushing) Problem. Let s, a be binary strings of length l. Let further $\mathcal{Q}_{s,\theta}$ be as in Definition of the LPN search problem. Let $\mathcal{A}$ be a adversary. The distinguishing-advantage of $\mathcal{A}$ between $\mathcal{Q}_{s,\theta}$ and the uniform distribution $\mathcal{U}_{l+1}$ is defined as

$$\mathrm{Adv}_{\mathrm{LPNDP}_{\theta,\mathcal{A}}}(l) = \Pr[\mathcal{A}^{\mathcal{Q}_{s,\theta}} = s | s \leftarrow \{0,1\}^l] - \Pr[\mathcal{A}^{\mathcal{U}_{l+1}} = 1] .$$

The LPNDP$_\theta$ with parameter θ is hard if the advantage of adversaries $\mathcal{A}$ is negligible.

It has been shown in [5] that the distinguishing-problem is as hard as the search-problem with similar parameters.

Complexity of Reasoning over Uncertain Data

According to the definitions of the LPN search and distinguishing problems, they belong to a wider class of problems related to the reasoning over uncertain data.

It has been proved that in the worst-case, the problem of decoding a random liner binary block code is NP-complete [1] as well as the LPN problem in the worst case.

On the other hand side, it should be noted that the average case hardness of the LPN problems, cannot be reduced to the worst-case hardness of a NP-hard problem. The confidence on the hardness of solving LPN problems in average case appears from the lack of efficient solutions despite the efforts over the years. Currently, the best known algorithms for solving the LPN search problems are the one reported in [2] and its improvements/alternatives (see [6]. [3] and [4]). The BKW algorithm [2] has the complexity $2^{O(l/\log_2 l)}$ and its improvements/alternatives can provide further reduction of the exponent for a factor $\Delta(l, \theta)$ (see the Definitions of the LPN problems). The talk discuses complexity of the above mentioned algorithms and points out to the open challenges.

References

[1] E.R. Berlekamp, R.J. McEliece, and H.C.A. van Tilborg, "On the Inherent Intractability of Certain Coding Problems", *IEEE Trans. Info. Theory*, vol. 24, pp. 384-386, 1978.

[2] A. Blum, A. Kalai and H. Wasserman, "Noise-Tolerant Learning, the Parity Problem, and the Statistical Query Model", *Journal of the ACM*, vol. 50, no. 4, pp. 506-519, July 2003.

[3] M. Fossorier, M.J. Mihaljević, H. Imai, Y. Cui and K. Matsuura, "An Algorithm for Solving the LPN Problem and its Application to Security Evaluation of the HB Protocols for RFID Authentication", INDOCRYPT 2006, *Lecture Notes in Computer Science*, vol. 4329, pp. 48-62, Dec. 2006.

[4] M. Fossorier, M.J. Mihaljević and H. Imai, "Modeling Block Encoding Approaches for Fast Correlation Attack", *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4728-4737, Dec. 2007.

[5] J. Katz and J. Shin, "Parallel and Concurrent Security of the HB and HB+ Protocols" EUROCRYPT 2006, *Lecture Notes in Computer Science*, vol. 4004, pp. 73-87, 2006.

[6] E. Levieil and P.-A. Fouque, "An Improved LPN Algorithm", SCN 2006, *Lecture Notes in Computer Science*, vol. 4116, pp. 348-359, 2006.

[7] M.J. Mihaljević, S. Gangopadhyay, G. Paul and H. Imai, "State Recovery of Grain-v1 Employing Normality Order of the Filter Function", *IET Information Security*, vol. 6, no. 2, pp. 55-64, June 2012

[8] M.J. Mihaljević, S. Gangopadhyay, G. Paul and H. Imai, "Internal State Recovery of Keystream Generator LILI-128 Based on a Novel Weakness of the Employed Boolean Function", *Information Processing Letters*, vol. 112, no. 21, pp. 805-810, November 2012.

[9] M.J. Mihaljević, S. Gangopadhyay, G. Paul and H. Imai, "Generic Cryptographic Weakness of k-normal Boolean Functions in Certain Stream Ciphers and Cryptanalysis of Grain-128", *Periodica Mathematica Hungarica* (Selected Papers of 2011 Central European Conference on Cryptology), vol. 65, no. 2, pp. 205-227, Dec. 2012.

[10] K. Pietrzak, "Cryptography from Learning Parity with Noise", SOFSEM 2012, *Lecture Notes in Computer Science*, vol. 7147, pp. 99-114, 2012.