Wire-Tap Channel System and Dedicated Coding

Miodrag Mihaljevic

RCIS-AIST, Tokyo

Spring Working Meeting of

Japan-India Project

Indian Statistical Institute, Kolkata, April 14, 2009

Roadmap

- Introduction – Certain References
- Motivations and the Goals
- Wire-Tap Channel Coding
- Coding for Binary Erasures Wire-Tap Channel
- Applications of Wire-Tap Channel Coding for Design of Stream Ciphers

I. Introduction

Certain References on Wire-Tap Channel

A.D. Wyner, “The wire-tap channel”, *Bell Systems Technical Journal*, vol. 54, pp. 1355-1387, 1975.

- A different approach for achieving secrecy of communication based on the noise has been reported by Wyner in 1975 assuming that the channel between the legitimate parties is with a lower noise in comparison with the channel via which a wire-tapper has access to the ciphertext.
- **The proposed method does not require any secret**. It is based on a specific coding scheme which provides a reliably communications within the legitimate parties and prevents, at the same time, the wire-tapper from learning the communication's contents.

Some Recent References

- M. Mihaljevic, “**Generic framework for secure Yuen 2000 quantum-encryption employing the wire-tap channel approach**”, *Physical Review A*, vol. 75, no. 5, pp. 052334-1-5, May 2007.
- A. Thangaraj, S. Dihidar, A.R. Calderbank, S.W. McLaughlin, and J.-M. Merolla, “**Applications of LDPC Codes to the Wiretap Channel”**, IEEE Trans. Information Theory, vol. 53, no. 8, pp. 2933-2945, August 2007 .
- M. Bloch, J. Barros, M.R.D. Rodrigues, S.W. McLaughlin, **“Wireless Information-Theoretic Security”,** IEEE Trans. Information Theory, vol. 54, no. 6, pp. 2515-2534, June 2008.

II. Motivations and Goals for Involvement of Wire-Tap Channel Coding into Certain Crypto Techniques

Trade-Off between Security and Communications & Implementations Overheads

Main Goals

- A framework for design of stream ciphers which provides opportunity for design **the security as high as possible based on the employed secret key**, i.e. ***complexity of recovering the key as close as possible to*** $O(2^K)$
- A ***trade-off between the security and the communications rate***: Increase the security up to the upper limit at the expense of a moderate decrease of the communications rate.

The Main Underlying Ideas

- Employ **physical noise** which an attacker must face, in order to strengthen the stream cipher.
- **Strengthen the stream cipher employing a dedicated encoding** following the homophonic or wire-tap channel encoding approaches.

Notes (1): Novel Paradigm

- Traditional stream ciphers do not include any randomness: Basically, they are based on the deterministic operations which expand a short secret seed into a long pseudorandom sequence.
- **This talk proposes an alternative approach yielding a novel paradigm for design of stream ciphers.**
- The proposed framework employs a **dedicated coding and a deliberate noise** which, assuming the appropriate code and noise level, at the attacker's side provides **increased confusion up to the limit determined by the secret key length**.
- **Decoding complexities with and without the secret key are extremely different**

Notes (2): Security-Overhead Trade-Off

In order to achieve the main security goal, **the proposed stream ciphering approach includes the following two encoding schemes with impacts on the communications overhead**:

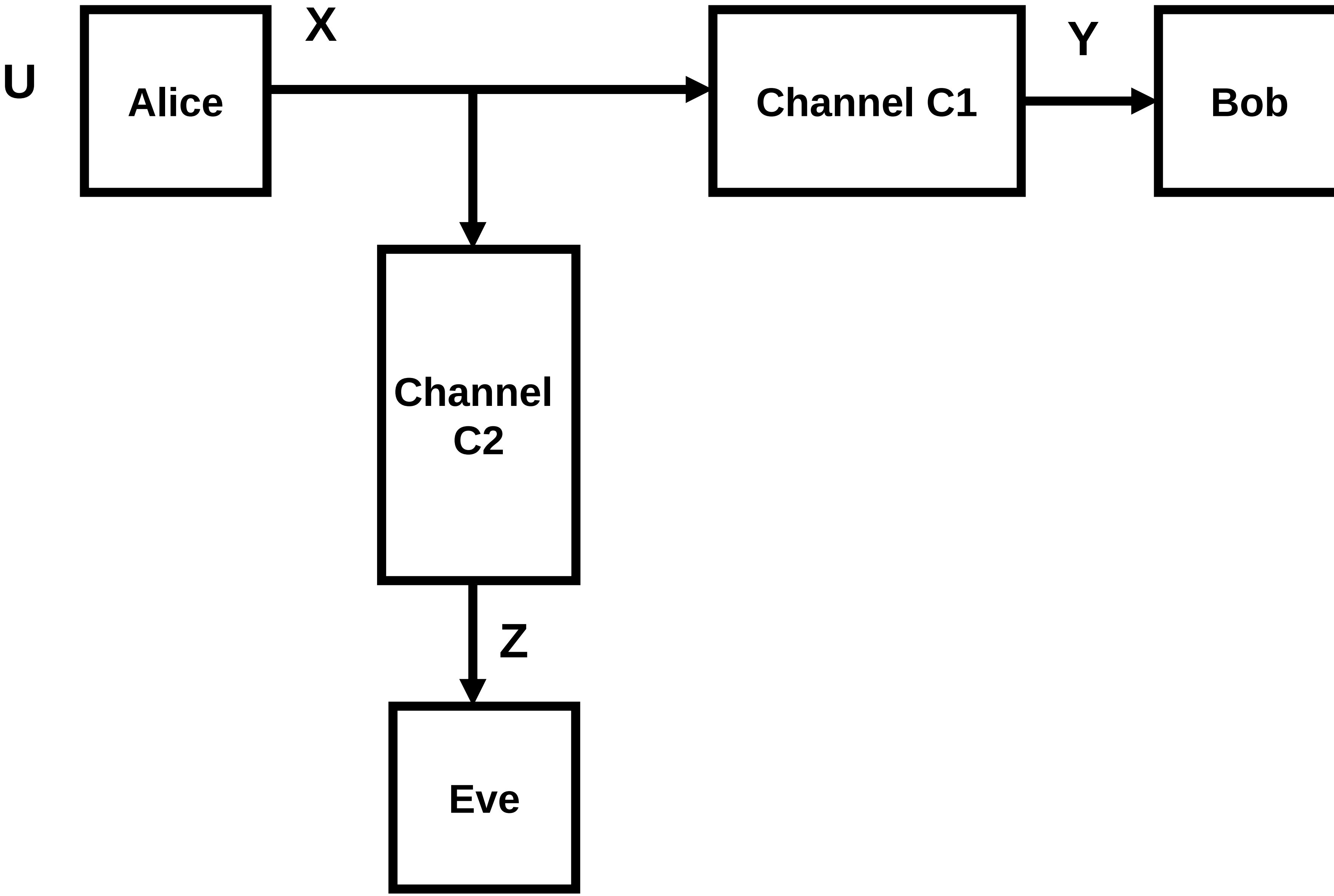
- error-correction encoding of the messages;
- dedicated homophonic/wire-tap channel coding which performs expansion of the initial ciphertext.

- Both of these issues imply the communications overhead: Accordingly, the proposed stream ciphers framework includes certain **trade-off between the security and the communications overhead** which in a number of scenarios can be considered as very appropriate.

III. Wire-Tap Channel

A. D. Wyner, “The wire-tap channel”,

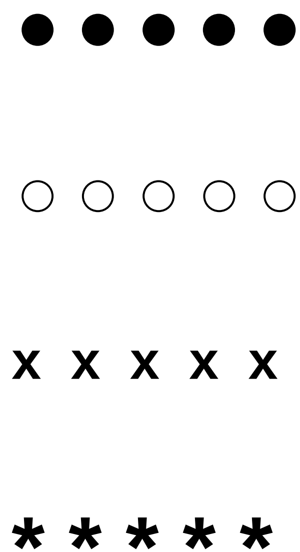
Bell Systems Technical Journal, vol. 54, pp. 1355-1387, 1975.


U
Alice
X
Channel C1
Y
Bob
Channel C2
Z
Eve

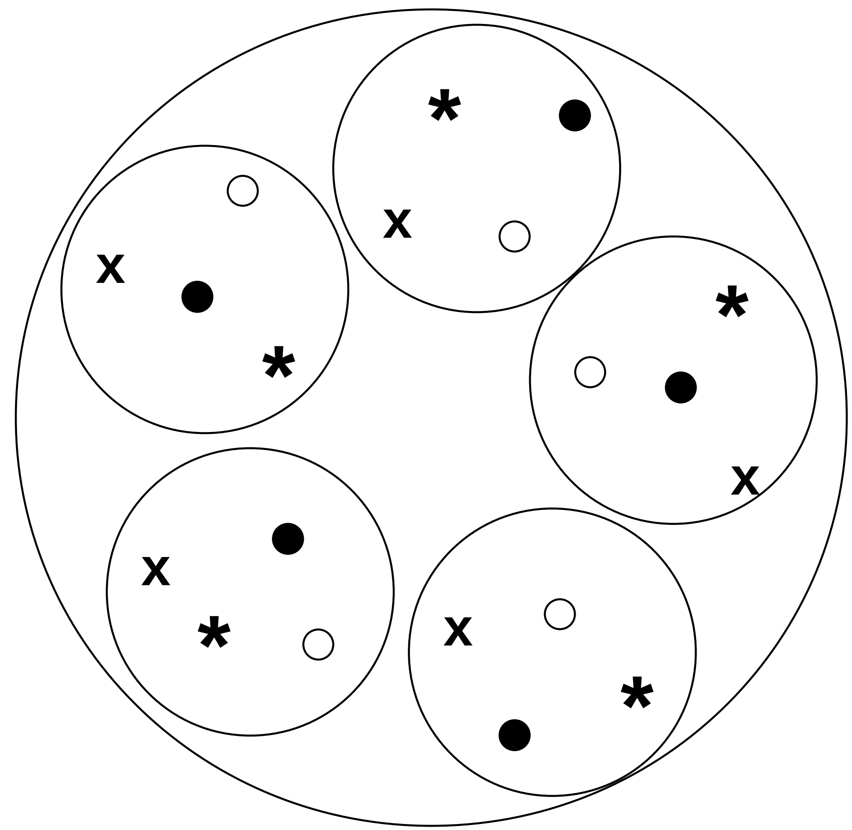
Coding Strategy for the Wire-Tap Channel

- Goal of encoding paradigm for the wire-tap channel is to **make the noisy data available to Eve (across the wire tap channel) useless** and achieving this goal is based on adding the randomness in encoding algorithm.

Groups of the codewords: *Same symbol denote different codewords belonging to the same group*



Codewords and N-dim Sphere



Dedicated Wire-Tap Channel Coding

Coding Method and

Selection of the Code

Wire-Tap Chanel Coding Preliminaries: Standard Array and Cosets

$M=2^k$

0	$\boldsymbol{c_1}$ … $\boldsymbol{c_i}$… $\boldsymbol{c_M}$
$\mathbf{e_1}$	
.	← ***coset e+C*** →
.	
$\mathbf{e_j}$	$\boldsymbol{c_i+e_j}$
.	
.	
.	
$\mathbf{e_N}$	

$N=2^{n-k}-1$

Coding Method (1)

We consider a generic approach for wire-tap channel coding as follows.

- To transmit m-bit message we first select a (n, k) code C such that $m \leq n - k$.

- Out of the 2^{n-k} cosets of C, we choose 2^m cosets and let each message correspond to a chosen coset.

- The selection of the cosets is done in a linear fashion as follows:

(a) Suppose $\mathbf{G}$ is a generator matrix for C with rows $\mathbf{g}_1$, $\mathbf{g}_2$, ..., and $\mathbf{g}_k$.

(b) We select m linearly independent vectors $\mathbf{h}_1$, $\mathbf{h}_2$, ..., $\mathbf{h}_m$, from $\{0, 1\}^n \backslash C$.

(c) The coset corresponding to a m-bit message $\mathbf{s} = [s_1, s_2, ..., s_m]$ is determined as follows:

$$\mathbf{s} \rightarrow s_1\mathbf{h}_1 \oplus s_2\mathbf{h}_2 \oplus ... \oplus s_m\mathbf{h}_m \oplus C. \qquad (1)$$

Coding Method (2)

The above correspondence is deterministic, but the encoding has a random component in the selection of the employed codeword. The transmitted word $\mathbf{c}$ is specified as follows:

$$\mathbf{c} = s_1\mathbf{h}_1 \oplus s_2\mathbf{h}_2 \oplus ... \oplus s_m\mathbf{h}_m \oplus u_1\mathbf{g}_1 \oplus u_2\mathbf{g}_2 \oplus ... \oplus u_k\mathbf{g}_k \tag{1}$$

where $\mathbf{u} = [u_1, u_2, ..., u_k]$ is an uniformly random k-bit vector and in a particular case $k = n - m$.

The overall encoding operation can be described as the following. Let $\mathbf{G}^*$ be the $m \times n$ matrix with rows $\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_m$. Then

$$\mathbf{c} = [\mathbf{su}] \begin{bmatrix} \mathbf{G}^* \\ \mathbf{G} \end{bmatrix} \tag{2}$$

Selection of the Code

For an arbitrary m-bit message $\mathbf{S} = \mathbf{s}$, the transmitted word belongs to $\mathbf{sG}^* \oplus C$. Since the cosets of C cover the entire space $\{0,1\}^n$, the attacker receives vector $\mathbf{Z}$ which belongs to some coset of C for example $\mathbf{rG}^* \oplus C$. If $\mathbf{e}$ denotes the error vector introduced by the wire-tapper's BSC(p), we have for $1 \leq i \leq 2^k$:

$$\text{Prob}\{Z \in \mathbf{rG}^* \oplus C\} = \text{Prob}\{\mathbf{e} \in (\mathbf{u} \oplus \mathbf{s})\mathbf{G}^* \oplus C\} = \text{Prob}\{\mathbf{e} \in \mathbf{w} \oplus C\} \ , \tag{1}$$

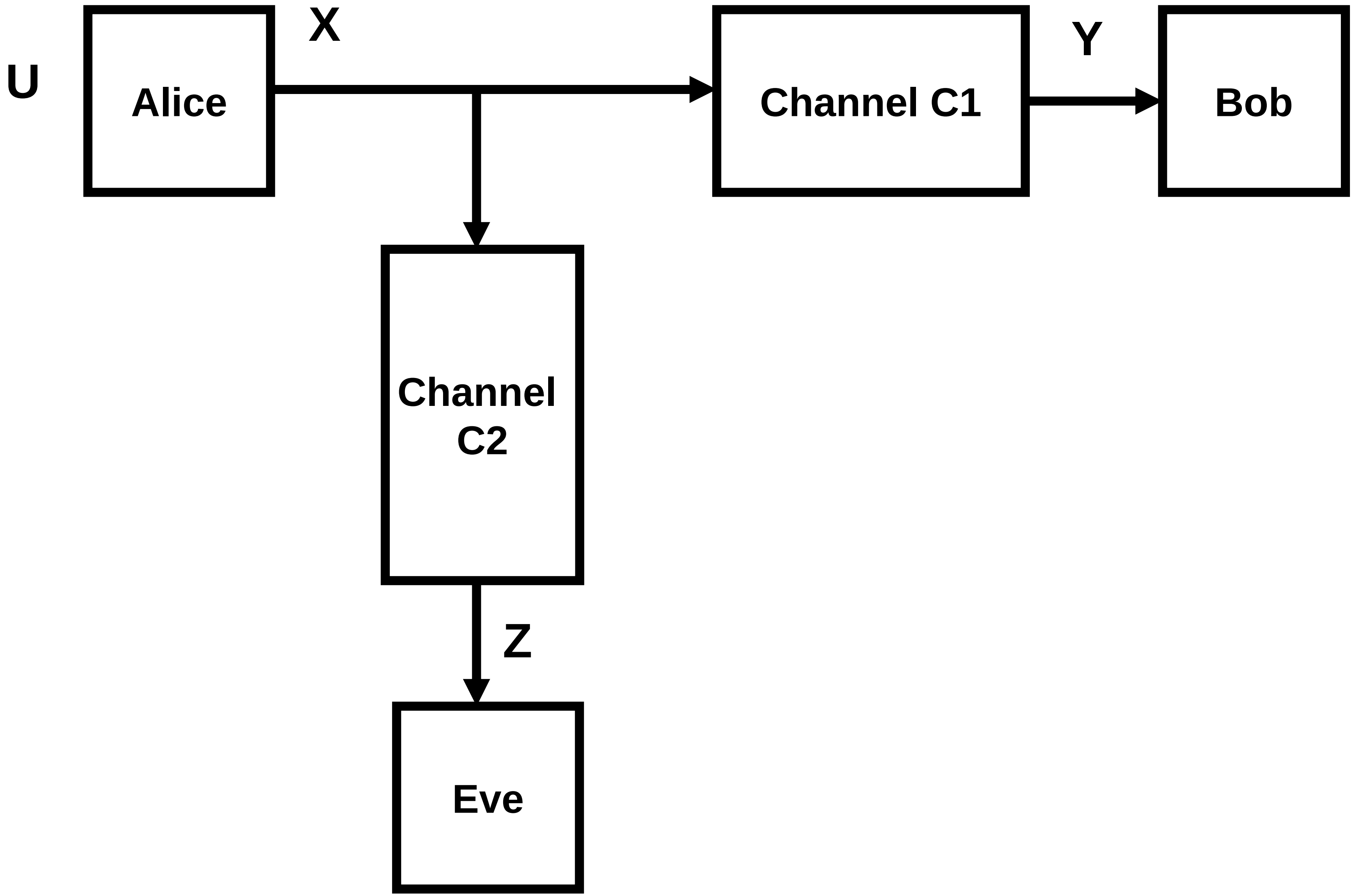
for some n-tuple $\mathbf{w}$. Accordingly, the following criterion for selecting the code C provides security of the message: Select C such that for any n-tuple $\mathbf{w}$, the following is valid:

$$\text{Prob}\{\mathbf{e} \in \mathbf{w} \oplus C\} \to 2^{-k}, \ \ as \ \ n \to \infty. \tag{2}$$

The above condition in conjunction with (10) implies that for an attacker it is equally likely to find $\mathbf{Z}$ in any coset of C given any message $\mathbf{S}$. Note that, assuming all $\mathbf{S} = \mathbf{s}$ are equally likely *a priori*, $\text{Prob}\{\mathbf{Z} \in \mathbf{rG}^* \oplus C\}$ is independent of $\mathbf{r}$: Hence,

$$\text{Prob}\{\mathbf{S} = \mathbf{s} | \mathbf{Z} \in \mathbf{rG}^* \oplus C\} \to 2^{-k} \ , \tag{3}$$

implies the security.


U
Alice
X
Channel C1
Y
Bob
Channel C2
Z
Eve

Implications of Capacity Approaching Codes over the Wire-Tap Channel

- $I(Z; UX) = I(U; Z) + I(X; Z|U)$
- $\quad\quad\quad\quad = I(X; Z) + I(U; Z|X)$
- Since $U \to X \to Z$ is a Markov chain,

$I(U; Z|X)=0$

Therefore

$I(U;Z)/n = I(X;Z)/n - I(X; Z|U)/n < C - (C-e) = e$

IV. Encoding for BEC Wiretapper's Channel and Noiseless Main Channel

We consider a scenario where a wire-tapper can observe the ciphertext via a binary erasure channel (BEC) where a bit erasure appears with the probability $1 - \epsilon$.

The transmitted n-tuple is denoted by the random variable $\mathbf{X} = [X_1, X_2, ..., X_n]$. Note that the message S can be seen as a syndrome of C with respect to a carefully constructed $k \times n$ parity-check matrix $\mathbf{H}$.

Since the channel between Alice and Bob is error-free, Bob finds the message as follows: $\mathbf{S} = \mathbf{H}\mathbf{X}^T$.

Wire-Tap Chanel Coding Preliminaries: Standard Array and Cosets

$M=2^k$

0	C_1	…	C_i…	C_M
$\mathbf{e_1}$				
.	←	***coset e+C***	→	
.				
$\mathbf{e_j}$			C_i+e_j	
.				
.				
.				
$\mathbf{e_N}$				

$N=2^{n-k}-1$

The eavesdropper learns X_i with probability ϵ. That is, the random variable $\mathbf{Z} = [Z_1, Z_2, ..., Z_n]$ is such that $Z_i = X_i$ with probability ϵ, and $Z_i =?$ (unknown or erasure) with probability $1 - \epsilon$.

If a coset of C contains at least one vector that agrees with $\mathbf{z} \in \{0, 1, ?\}^n$ in the unerased positions, the **coset is consistent** with $\mathbf{z}$.

Each consistent coset corresponds to a possible message for the eavesdropper.

Let $\mathbf{v}$ be a vector consistent with $\mathbf{z}$ in the coset $\mathbf{v} + C$.

Let S be the set of all vectors in $\mathbf{v} + C$ consistent with $\mathbf{z}$.

Then, $\mathbf{v} + S$ is the set of all vectors in C with zeros in the positions revealed in $\mathbf{z}$. That is,

$$\mathbf{v} + S = \{u \in C : u_i = 0 \text{ whenever } z_i \neq ?\} .$$

Since $|S| = |\mathbf{v} + S|$, the number of vectors consistent with $\mathbf{z}$ in each consistent coset is a constant.

Let $N(C, \mathbf{z})$ denote the total number of cosets of C consistent with $\mathbf{z}$.

Since each message is equally likely a priori, we get

$$H(S|\mathbf{Z} = \mathbf{z}) = \log_2 N(C, \mathbf{z}) \ .$$

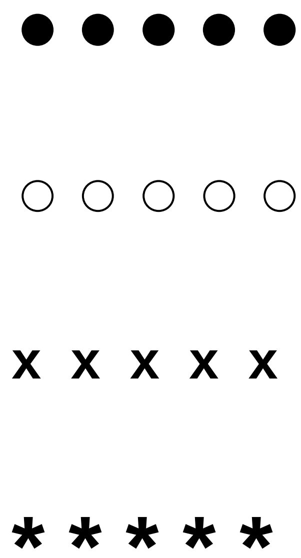
For an $(n, n-k)$ code C, the maximum possible value for $N(C, \mathbf{z})$ is the total number of cosets which is equal to 2^k.

If $N(C, \mathbf{z}) = 2^k$, we say that **z** is secured by C since the eavesdroppers

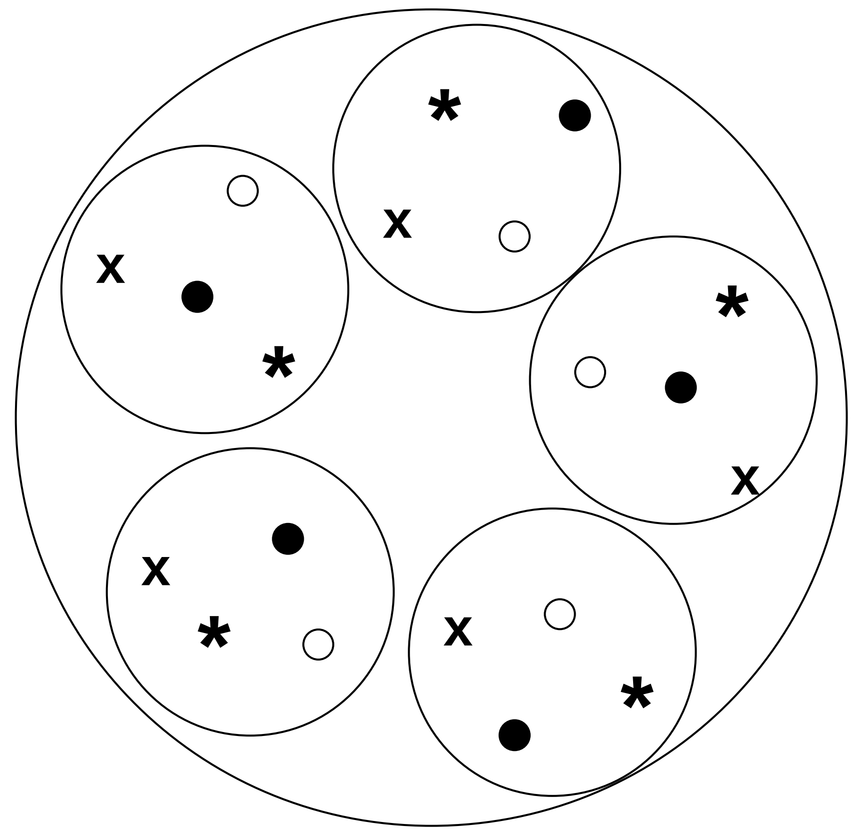
$$\text{Prob}\{\mathbf{S} = \mathbf{s} | \mathbf{Z} = \mathbf{z}\} = 1/2^k$$

for every possible message $\mathbf{s}$.

Groups of the codewords: *Same symbol denote different codewords belonging to the same group*



Codewords and N-dim Sphere



Theorem, [1]. Let an $(n, n-k)$ code C have a generator matrix $\mathbf{G} = [a_1, ..., a_n]$, where a_i is the i-th column of $\mathbf{G}$. Consider an instance of the eavesdroppers observation $\mathbf{z} \in \{0, 1, ?\}^n$ with μ unerased positions given by $\{i : z_i \neq ?\} = \{i_1, i_2, ..., i_\mu\}$. $\mathbf{z}$ is secured by C if the matrix $\mathbf{G}_\mu = [a_{i_1}\ a_{i_2}\ ...\ a_{i_\mu}]$ has rank μ.

[1] L.H. Ozarow and A.D. Wyner, "Wire-tap channel II", *AT&T Bell Systems Technical Journal*, vol. 63, pp. 2135-2157, Oct. 1984.

Dedicated LDPC for Wire-Tap Channel Coding

- Just to mention -

V. Applications of Wire-Tap Channel Coding for Design of Cryptographic Primitives

Very Recent References (1)

- [1] M. Mihaljevic and H. Imai, **“An Approach for Stream Ciphers Design Based on Joint Computing over Random and Secret Data”**, *COMPUTING*, accepted for publication, 2009. (Impact Factor: 0.949)
- [2] M. Mihaljevic, **“A Framework for Stream Ciphers Based on Pseudorandomness, Randomness and Error-Correcting Coding**”, in *Enhancing Crypto-Primitives with Techniques from Coding Theory*, Editors B. Preneel and S. Dodunekov, Vol. in the Series *Information and Communication Security*, IOS Press, Amsterdam, 23 pages, to appear 2009.
- [3] M. Mihaljevic and H. Imai, **“A Stream Cipher Design Based on Embedding of Random Bits”**, *IEEE 2008 Int. Symp. on Inform. Theory and its Appl. - ISITA2008*, Auckland, New Zealand, Dec. 7-10, 2008, Proceedings, pp. 1497-1502. (ISBN: 978-1-4244-2069-8; copyright2008 IEEE)

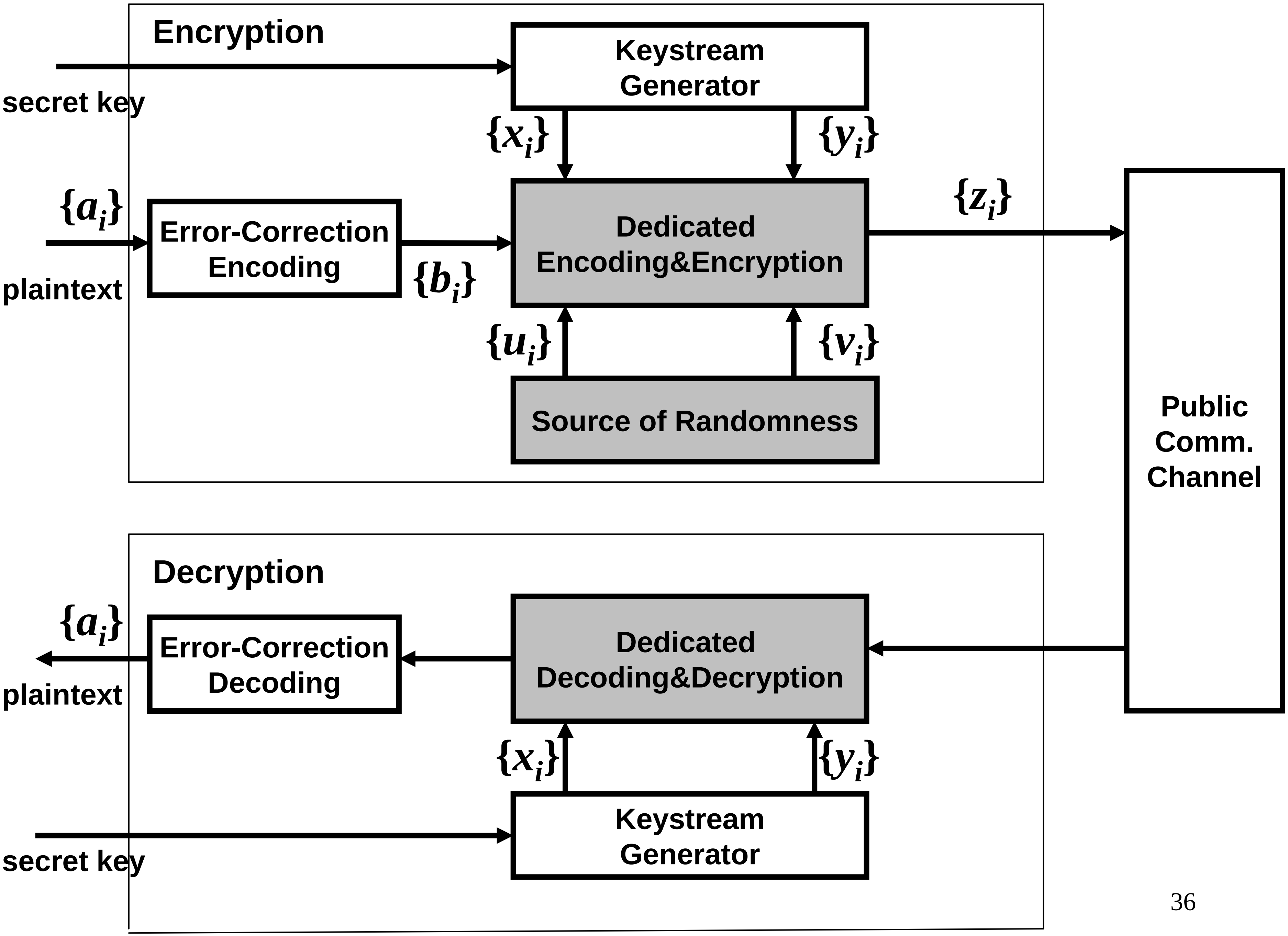
Some Earlier Results on Crypto&Coding

(there is a number of other results achieved in the period 2005-2008)

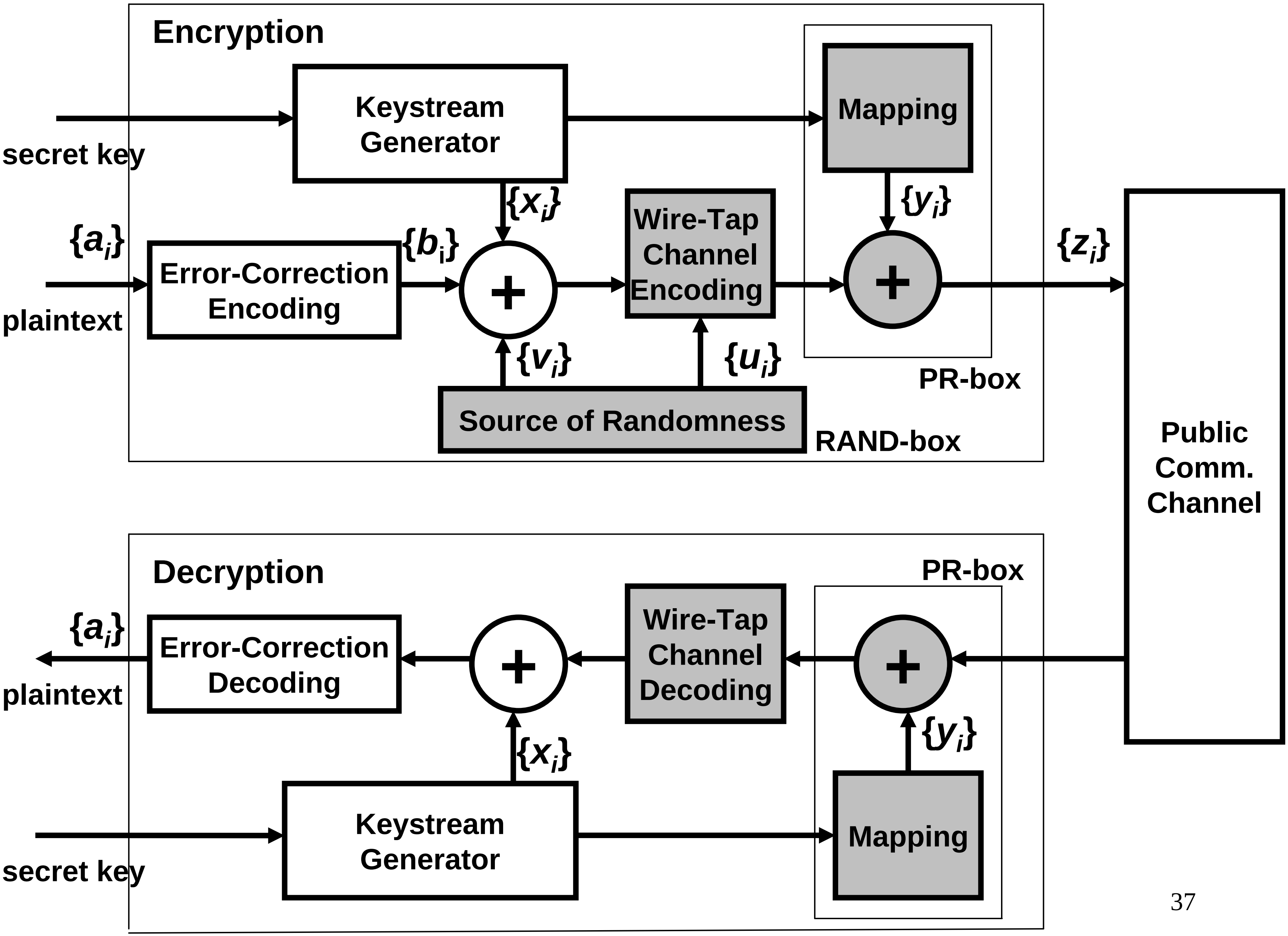
- [7] M. Mihaljevic, M. Fossorier and H. Imai, **“Key Management with Minimized Secret Storage Employing an Erasure Channel Approach”**, *IEEE Communications Letters*, vol. 9, pp. 741-743, Aug. 2005. (Impact Factor: 0.922)
- [8] M. Fossorier, M. Mihaljevic, H. Imai, Y. Cui and K. Matsuura, **“An Algorithm for Solving the LPN Problem and its Application to Security Evaluation of the HB Protocols for RFID Authentication”,** *Lecture Notes in Computer Science*, vol. 4329, pp. 48-62, Dec. 2006. (Impact Factor: ~ 0.5)
- [9] M. Mihaljevic, **“Generic framework for secure Yuen 2000 quantum-encryption employing the wire-tap channel approach”,** *Physical Review A*, vol. 75, no. 5, pp. 052334-1-5, May 2007. (Impact Factor: ~ 3.0)
- [10] M. Fossorier, M. Mihaljevic 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. (Impact Factor: 2.183)

Some Previous Results on Randomized Encryption

- [11] R. Rivest and T. Sherman, **"Randomized Encryption Techniques",** *Advances in Cryptology: Proceedings of CRYPTO '82*, Plemum, New Yourk, pp. 145-163, 1983.
- [12] N.J.A. Sloane, **"Error-correcting codes and Cryptography",** *Cryptologia*, vol. 6, pp. 128-153, 1982.
- [13] O. Kara and I. Erguler, **"A New Approach to Keystream Based Cryptosystems",** *SASC 2008*, Workshop Record, pp. 205-221, Feb.



Encryption
secret key
Keystream Generator
$\{x_i\}$
$\{y_i\}$
$\{a_i\}$
plaintext
Error-Correction Encoding
$\{b_i\}$
Dedicated Encoding&Encryption
$\{z_i\}$
$\{u_i\}$
$\{v_i\}$
Source of Randomness
Public Comm. Channel
Decryption
$\{a_i\}$
plaintext
Error-Correction Decoding
Dedicated Decoding&Decryption
$\{x_i\}$
$\{y_i\}$
Keystream Generator
secret key


Encryption
secret key
Keystream Generator
Mapping
{x_i}
{y_i}
{a_i}
{b_i}
Wire-Tap Channel Encoding
{z_i}
plaintext
Error-Correction Encoding
+
+
{v_i}
{u_i}
PR-box
Source of Randomness
RAND-box
Public Comm. Channel
Decryption
PR-box
{a_i}
Error-Correction Decoding
+
Wire-Tap Channel Decoding
+
plaintext
{y_i}
{x_i}
Keystream Generator
Mapping
secret key

Algebraic Model of Stream Cipher II under Chosen Plaintext Attack

Corollary 1. Under the chosen plaintext attack which for each t implies $\mathbf{b}_t = \mathbf{0}$ (i.e. the all zeros vector), Proposition 1 implies:

$$\mathbf{z}_t = \mathbf{q}_t \mathbf{S} \oplus \vec{\nu}_t \ , \tag{1}$$

where

$$\mathbf{q}_t = (\bigoplus_{i=1}^{m} x_i \mathbf{h}_i \oplus \mathbf{y}_t)\mathbf{S}^{-1} \ , \ \ \vec{\nu}_t = (\bigoplus_{i=1}^{n-m} u_i \mathbf{g}_i) \oplus (\bigoplus_{i=1}^{m} v_i \mathbf{h}_i) \ , \tag{2}$$

and where $\mathbf{S}$ is an $n \times n$ binary matrix determined by the length k binary secret key $\mathbf{k}$, and $\mathbf{S}^{-1}$ is its inverse.

Assumption 1. For any $t = 1, 2, ...,$ $\vec{\nu}_t \leftarrow \text{Ber}_{n,\eta}$, where η is the parameter.

A Statement on Stream Cipher II Security

Theorem 1. Assume there is an adversary $\mathcal{A}$, running in time T , and attacking the Stream cipher II specified by Corollary 1 and Assumption 1 with parameters (ℓ, m, k, n, η), $k = n$, in the sense of IND with advantage δ by making at most q queries to the encryption oracle. Then there is an algorithm $\mathcal{M}$ making $O(q)$ oracle queries, running in time $O(T)$, and such that

$$\left|\Pr\left[\mathrm{s} \leftarrow \{0,1\}^k \ : \ \mathcal{M}^{\Pi_{s,\eta}}(1^k) = 1\right] - \Pr\left[\mathcal{M}^{\mathcal{U}_{k+1}}(1^k) = 1\right]\right| \geq \frac{\delta}{n} . \tag{1}$$

Thank You Very Much for the Attention,

and

QUESTIONS Please!