

Information-theoretic analysis of electronic and printed document authentication

S. Voloshynovskiy, O. Koval, R. Villan, E. Topak,
J. Vila, F. Deguillaume, Y. Rytsar and T. Pun

Stochastic Image Processing (SIP) Group
University of Geneva

- 1. Introduction**
- 2. Document authentication**
- 3. Main practical scenarios and channel models**
- 4. Security analysis of document authentication**
- 5. System implementation and concept validation**
- 6. Conclusions**

1. Introduction



Main goal: study of information-theoretic limits of text document authentication stored in both electronic and printed forms.

Main requirements:

- § **preservation of document layout (no bar codes);**
- § **document storage and circulation in electronic (.doc, .tex, .pdf, .ps) and printed forms;**
- § **on-line authentication (no access to document database);**
- § **b&w document reproduction using cheap ink jet/laser printers;**
- § **document acquisition: scanner, cameras of PDAs and mobile phones;**
- § **robustness to legitimate distortions with:**
 - § **resistance against copy attack;**
 - § **detection of document duplication.**

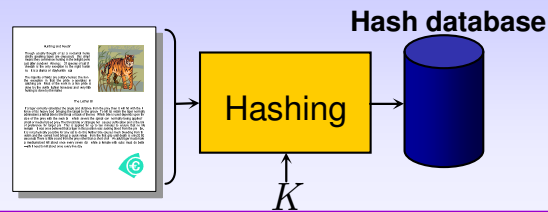
Common solution: generation of document hash computed using a secret key K .

1. Introduction



Three main protocols of hash storage and system implementation:

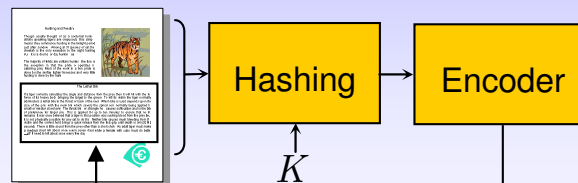
§ Hash storage in database;



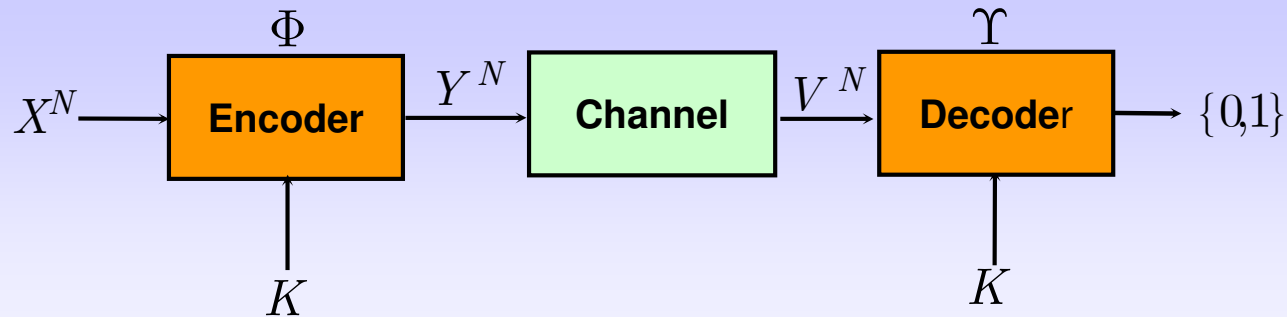
§ Direct storage of hash onto the documents (dense and sparse bar codes);



§ Hash storage in the document itself (self-embedding using data-hiding).



2. Document authentication: problem formulation



$$\Phi^N: \mathcal{X}^N \times \mathcal{K} \rightarrow \mathcal{Y}^N$$

$$\Upsilon^N: \mathcal{V}^N \times \mathcal{K} \rightarrow \{0,1\}$$

$$\begin{cases} H_0: V^N \sim p_{V_0^N|K}(v^N | k) - \text{authentic,} \\ H_1: V^N \sim p_{V_1^N|K}(v^N | k) - \text{modified.} \end{cases}$$

$$\ell(v^N | k) \triangleq \log_2 \frac{p_{V_1^N|K}(v^N | k)}{p_{V_0^N|K}(v^N | k)} \geq T$$

$$D(\hat{p}_{V^N} \| p_{V_1^N|K}) - D(\hat{p}_{V^N} \| p_{V_0^N|K}) \geq T$$

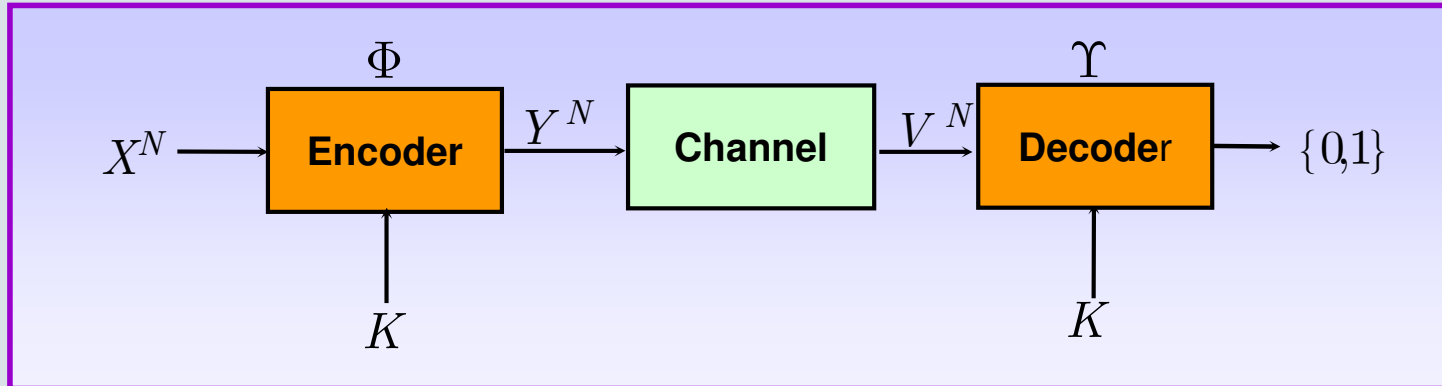
$$P_M \triangleq E_{p_K} [\Pr[\ell(v^N | k) < T | H_1]],$$

$$P_F \triangleq E_{p_K} [\Pr[\ell(v^N | k) > T | H_0]]$$

$$\text{Game: } \min_{\Phi, \Upsilon} \max_{p_{V^N|Y^N}(\cdot, \cdot)} P_M(\Phi, \Upsilon, p_{V^N|Y^N}(\cdot, \cdot))$$

for fixed P_F

2. Document authentication: problem formulation



Stein's lemma suggests the exponential decay of probabilities of errors:

$$P_F \sim 2^{-E_{pK} \left[D \left(p_{V_1^N|K} \parallel p_{V_0^N|K} \right) \right]}, \text{ for a fixed } P_M$$

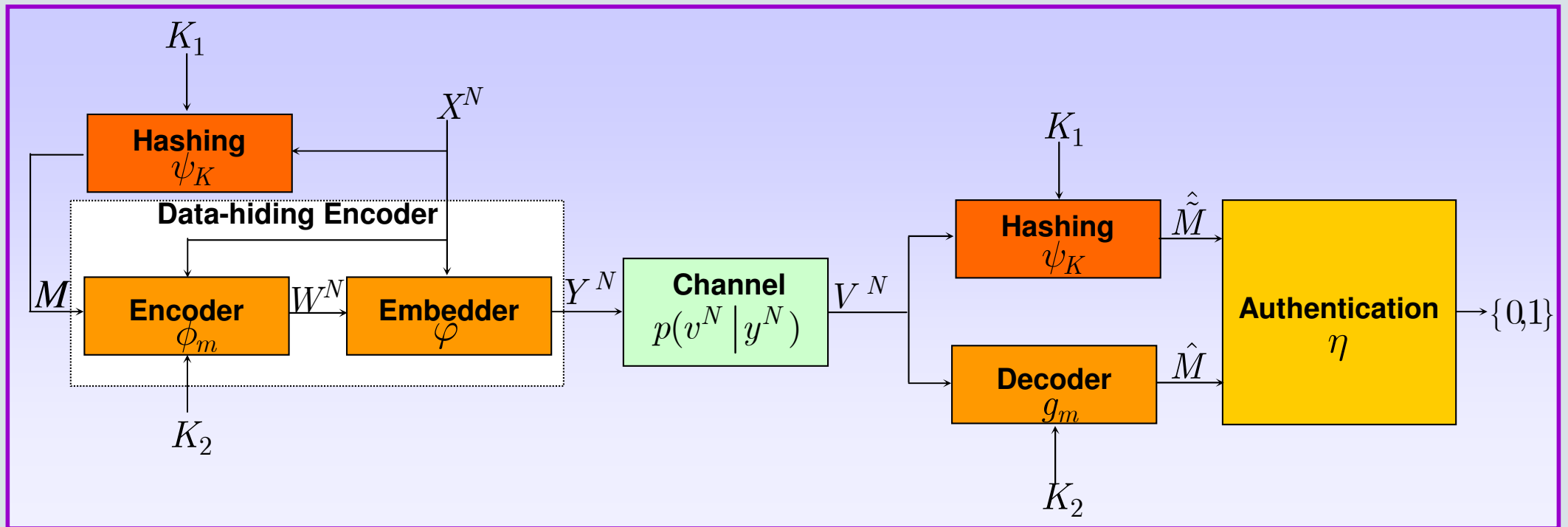
$$P_M \sim 2^{-E_{pK} \left[D \left(p_{V_0^N|K} \parallel p_{V_1^N|K} \right) \right]}, \text{ for a fixed } P_F$$

2. Document authentication: problem formulation



Hashing-data-hiding problem:

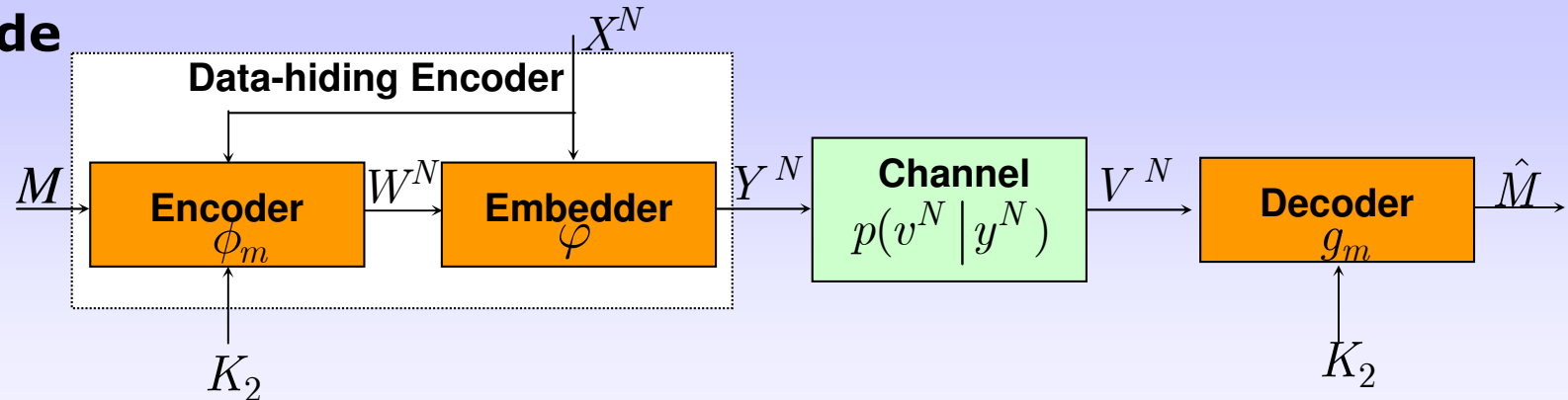
- § Separation approach (separate data-hiding and hashing: rate matching);
- § Joint approach (extreme case – uncoded transmission).



2. Document authentication: Data-hiding code



Data-hiding code



Assumptions: $M \in \{1, 2, \dots, |\mathcal{M}|\}, |\mathcal{M}| = 2^{NR}$ $K_2 \in \{1, 2, \dots, |\mathcal{K}_2|\}$

$$w^N \in \mathcal{W}^N, x^N \in \mathcal{X}^N, y^N \in \mathcal{Y}^N, v^N \in \mathcal{V}^N \quad p_{V^N|Y^N}(v^N | y^N) = \prod_{i=1}^N p_{V|Y}(v_i | y_i)$$

Encoding: $\phi_m^N : \mathcal{M} \times \mathcal{X}^N \times \mathcal{K}_2 \rightarrow \mathcal{W}^N$

Embedding: $\varphi^N : \mathcal{W}^N \times \mathcal{X}^N \rightarrow \mathcal{Y}^N$

Decoding: $g_m^N : \mathcal{V}^N \times \mathcal{K}_2 \rightarrow \mathcal{M}$

Distortion criteria:

$$d^N(x^N, y^N) = \frac{1}{N} \sum_{i=1}^N d(x_i, y_i)$$

2. Document authentication: Data-hiding code



Constraints:

$$\frac{1}{|\mathcal{K}_2| |\mathcal{M}|} \sum_{k_2 \in \mathcal{K}_2} \sum_{m \in \mathcal{M}} \sum_{x^N \in \mathcal{X}^N} d^N(x^N, \varphi^N(\phi_m^N(m, x^N, k), x^N)) p_{X^N}(x^N) \leq D^E$$

$$\sum_{y^N \in \mathcal{Y}^N} \sum_{v^N \in \mathcal{V}^N} d^N(y^N, v^N) p_{V^N|Y^N}(v^N | y^N) p_{Y^N}(y^N) \leq D^A$$

Probability of error:
$$P_e^{(N)} = \frac{1}{|\mathcal{K}_2| |\mathcal{M}|} \sum_{k_2 \in \mathcal{K}_2} \sum_{m \in \mathcal{M}} \Pr[g^N(V^N, K_2) \neq m | M = m]$$

Data-hiding capacity for a fixed channel:

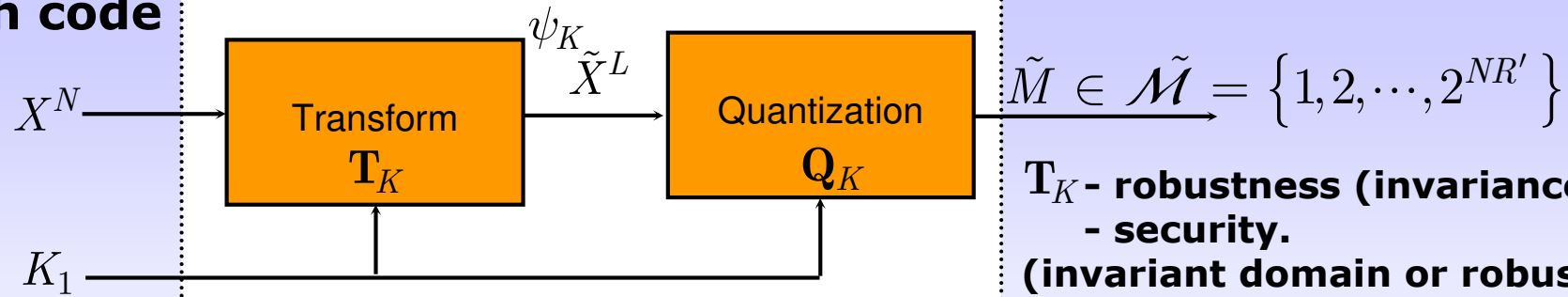
$$C = \frac{1}{N} \max_{p(u^N, w^N | x^N)} [I(U^N; V^N) - I(U^N; X^N)] \quad u^N \in \mathcal{U}^N (K_2 = k_2)$$

$$p(k_2, x^N, u^N, w^N, y^N, v^N) = p(k_2) p(x^N) p(u^N | x^N) \mathbf{1}\{u^N \in \mathcal{U}(K_2 = k_2)\} p(w^N | u^N, x^N) p(y^N | x^N, w^N) p(v^N | y^N)$$

2. Document authentication: Hash code



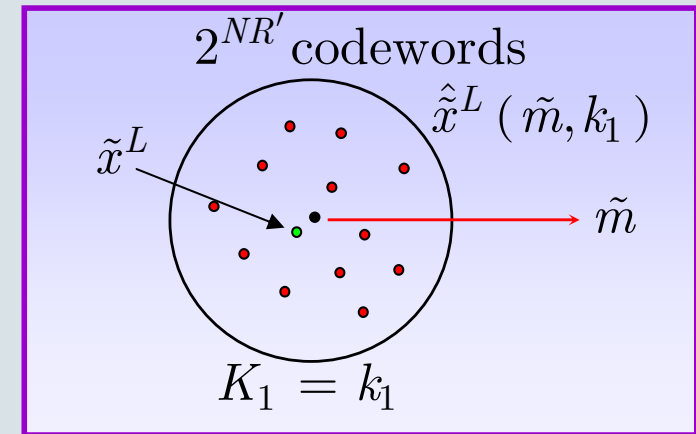
Hash code



\mathbf{T}_K - **robustness (invariance);**
 - **security.**
 (invariant domain or robust features)

\mathbf{Q}_K - **quantization with "robust"**
binary labeling.

Hashing: $\psi_K^N : \mathcal{X}^N \times \mathcal{K}_1 \rightarrow \tilde{\mathcal{M}}$

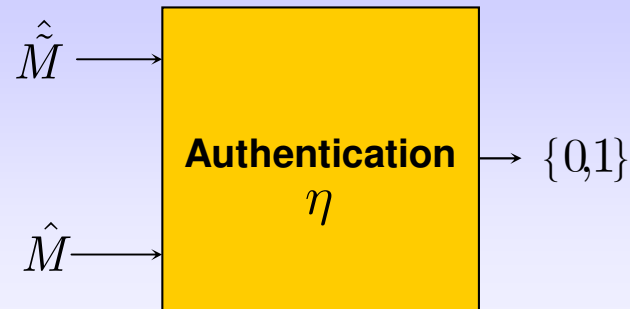


$$R'(D^A) = \min_{p_{\hat{X}|\tilde{X}} \text{ (i.l.)}: E[d(\hat{X}, \tilde{X})] \leq D^A} I(\hat{X}; \tilde{X})$$

2. Document authentication: Authentication decision



Authentication



Binary representation of computed hash:

$$\hat{m} \triangleq \hat{\mathbf{b}}$$

decoded message:

$$\hat{m} \triangleq \hat{\mathbf{b}}$$

Decision: $\eta : \mathcal{M} \times \tilde{\mathcal{M}} \rightarrow \{0,1\}$ Decision is taken wrt threshold T defined by P_F .

Conjecture (Authentication based on hashing-data-hiding principle): if X^N is a finite alphabet stochastic process that satisfies the asymptotic equipartition property (AEP) then there is a hashing-data-hiding code with specified P_F and $P_M \rightarrow 0$, if the rate of the hash code R' satisfies $R' < C$. Conversely, for any stationary process, if $R' > C$, the P_M is bounded away from zero, and it is not possible to authenticate X^N with arbitrarily low probability of error. **⇒ Rate matching!**

3. Main practical scenarios and channel models



Electronic document authentication

Channel model: intensity conversions among electronic formats are deterministic and well-defined that can be easily taken into account at the encoder.

Hybrid electronic-analog document authentication

Channel model: additionally halftoning process (mapping of modulated signals from intensity space to halftone space for printing and back for decoding).

Two models:

- § **Parallel binary symmetric channel;**
- § **Non-stationary Generalized Gaussian channel approximation.**

Analog document authentication

Channel model: direct encoding using halftone modulation.

4. Security analysis of document authentication



Lemma 1 (Equivocation for the data-hiding code)

$$Q_A^{DH} = 2^{h(U^N|Y^N)} = 2^{[h(U^N) - I(U^N; Y^N)]}$$

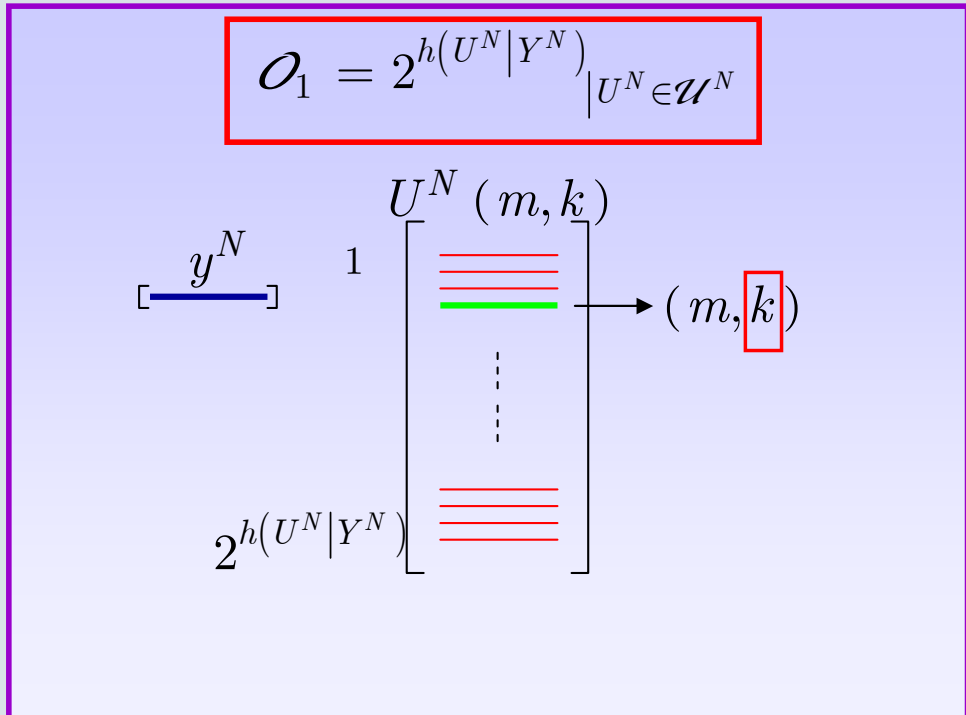
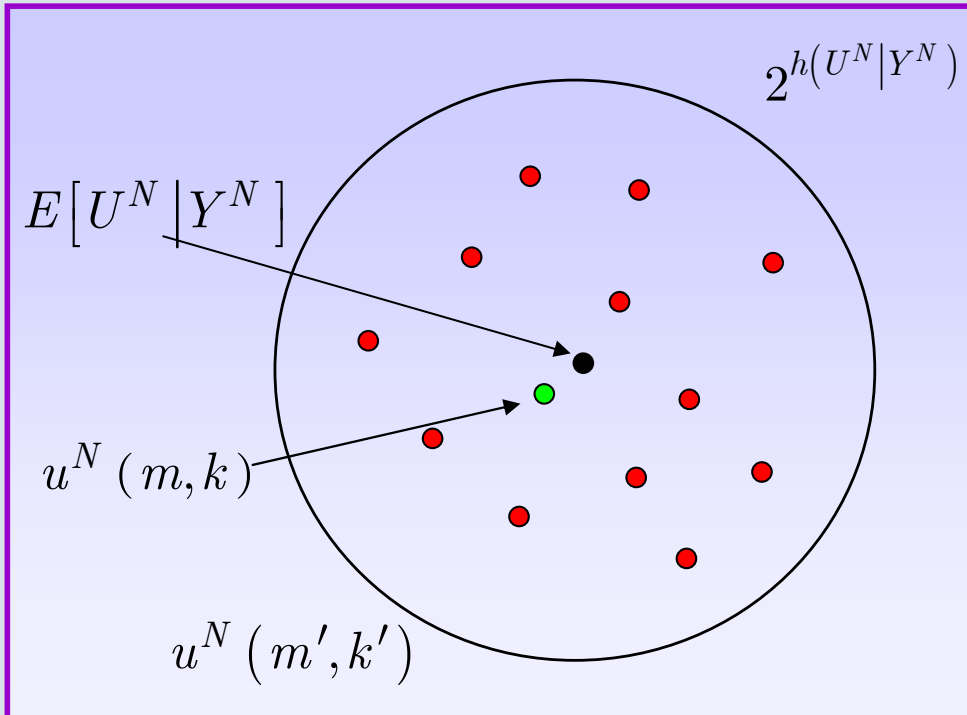
Lemma 2 (Equivocation for the hash code)

$$Q_A^{HS} = 2^{h(\hat{X}^L|Y^N)} = 2^{[h(\hat{X}^L) - I(\hat{X}^L; Y^N)]}$$

Two possible key management protocols:

- § the same keys $k_1 = k_2 = k$
- § different keys $k_1 \neq k_2$

4. Security analysis of document authentication



$$\hat{X}^N = E[X^N | Y^N, \hat{U}^N]$$

Produce a new hash M' from \hat{X}'^N and k

Produce a faked copy \hat{X}'^N based on \hat{X}^N

Perform embedding of M' into \hat{X}'^N using k

4. Security analysis of document authentication



Attacker complexity to reveal K based on data-hiding part of code:

$$\mathcal{O}_1 = 2^{h(U^N|Y^N)} \Big|_{U^N \in \mathcal{U}^N}$$

Attacker complexity to reveal K based on hash part of code:

$$\mathcal{O}_2 = 2^{h(\hat{X}^L|Y^N)} \Big|_{\hat{X}^L \in \tilde{\mathcal{X}}^L}$$

Attacker chooses strategy with lower complexity:

$$\mathcal{O} = \min \{ \mathcal{O}_1, \mathcal{O}_2 \}$$

Note: case of two different keys is presented in the paper.

5. System implementation and concept validation



Authentic physical document and correct content

Note: hash is computed from overlapping blocks on different hierarchical levels.

Geneva

Correct text

=> Document is validated

5. System implementation and concept validation



Authentic physical document but altered content

Paris V
Wrong text !

=> Modifications are localized

5. System implementation and concept validation



Document copied using copy machine

The screenshot shows a software interface with several windows. On the left is a document page with a photo of Frédéric Dage and contact information. In the center is a window titled 'Berkut v3.05' showing a scanned document 'scanned_document0.png' with the following text: 'Senior researcher at Stochastic Image Processing (SIP) group, Computer Vision and Multimedia Laboratory (CVML) at University of Geneva, Centre Universitaire d'Informatique. Teaching instructor at University of Geneva of the lecture "Multimedia Security". Development of data-hiding technologies and digital rights management (DRM) protocols for multimedia content.' Below this is a color calibration bar. On the right is an 'Authentication' dialog box with the message: 'This document was completely altered ! There are 8 modified text regions: 0 1 2 3 4 5 6 7 Document is not authentic !' with an 'OK' button. Below the dialog box is a window titled 'tampering_map - Aperçu des images et des télécopies Windows' showing the same text with red and purple diagonal lines overlaid on the characters, indicating tampering. A blue arrow points from the authentication dialog to the tampering map window.

Not original !

=> Document is fully rejected

6. Conclusions



- § **Information-theoretic analysis of authentication problem based on hashing-data-hiding is presented.**
- § **Security of such scheme is estimated based on analysis of information security leakages.**
- § **The first prototype is implemented to prove the considered concept.**

Extensions:

- § **Theoretic:**
 - § **error exponent analysis;**
 - § **“uncoded” authentication;**
 - § **security analysis extension to multiple documents.**
- § **Practical: experimental validation for various printing/scanning conditions.**