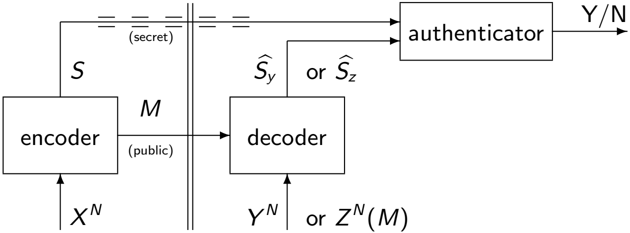
Authentication Based on Secret-Key Generation

Frans M.J. Willems
(joint work w. Tanya Ignatenko)

Eindhoven University of Technology

IEEE EURASIP Spain Seminar on Signal Processing, Communication and Information Theory,
Universidad Carlos III de Madrid,
December 11, 2014

INTRODUCTION: Scenario Secret-Based Authentication



- **ENROLLMENT:** An individual presents his biometric sequence X^N to an encoder. From this enrollment sequence X^N a secret S is generated. Also a public helper message M is produced.
- **LEGITIMATE PERSON:** The person presents a legitimate observation sequence Y^N to a decoder. The decoder produces an estimated secret $\widehat{S_y}$ using helper message M.
- **IMPOSTOR:** An impostor who has access to the helper message M presents an impostor sequence $Z^N(M)$ to the decoder that now forms estimated secret $\widehat{S_z}$ using M.
- **AUTHENTICATOR:** Checks whether the estimated secret $\widehat{S_y}$ or $\widehat{S_z}$ equals the enrolled secret S, and outputs yes or no.

The symbols of the enrollment and legitimate observation sequences assume values in the finite alphabets $\mathcal{X}$ and $\mathcal{Y}$ respectively.
The joint probability

$$\Pr\{X^N = x^N, Y^N = y^N\} = \Pi_{n=1}^{N} Q(x_n, y_n), \text{ for all } x^N \in \mathcal{X}^N, x^N \in \mathcal{Y}^N. \quad (1)$$

where $Q(x, y)$ for $x \in \mathcal{X}, y \in \mathcal{Y}$ is a probability distribution, hence the pairs (X_n, Y_n) for $n = 1, 2, \cdots, N$ are independent and identically distributed (i.i.d.).

Also the symbols of the impostor sequences assume values in the alphabet $\mathcal{Y}$.

INTRODUCTION: Encoder, Decoder, and Authenticator

Encoding function:

$$(S, M) = e(X^N), \tag{2}$$

where $S \in \{\phi_e, 1, 2, \cdots, |\mathcal{S}|\}$ is the generated secret and $M \in \{1, 2, \cdots, |\mathcal{M}|\}$ the public helper message. Here ϕ_e is the secret-value if the encoder could not assign a secret.

Decoding function:

$$\widehat{S_y} = d(M, Y^N), \tag{3}$$

where $\widehat{S_y} \in \{\phi_d, 1, 2, \cdots, |\mathcal{S}|\}$ is the estimated secret. Again ϕ_d is the estimated secret-value if the decoder could not find an estimated secret.

Note that an **impostor** can choose

$$Z^N = i(M), \tag{4}$$

depending on the helper data M. This impostor sequence $z^N \in \mathcal{Z}^N$ is then presented to the decoder that forms

$$\widehat{S_z} = d(M, Z^N) = d(M, i(M)). \tag{5}$$

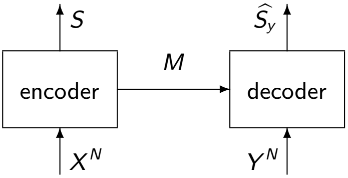
The **authenticator** checks whether the output of the encoder, i.e. the secret S, and the output of the decoder, i.e. the estimated secret $\widehat{S_y}$ or $\widehat{S_z}$, are equal.

INTRODUCTION: False-Reject and False-Accept Rates

The False Reject Rate (FRR) and False Accept Rate (FAR) are typical performance measures for authentication systems. They are defined as follows:

$$\begin{aligned} \text{FRR} &\triangleq \Pr\{\widehat{S_y} \neq S\}, \text{ and} \\ \text{FAR} &\triangleq \Pr\{\widehat{S_z} = S\}. \end{aligned} \tag{6}$$

NOTE that, given the probability distribution $Q(\cdot,\cdot)$, the FRR depends only on the encoder and decoder functions $e(\cdot)$ and $d(\cdot,\cdot)$. The FAR moreover depends on the impostor strategy $i(\cdot)$.



Both the enrolled and estimated secret assume values in $\{1, 2, \cdots, |\mathcal{S}|\}$.
A: The secret must be recoverable by the decoder. **B:** It should be large and uniform. **C:** The helper message should be uninformative about the secret.

Definition

Secrecy rate R is achievable if, for all $\delta > 0$ and all N large enough, there exist encoders and decoders such that

$$\begin{aligned} \Pr\{\widehat{S_y} \neq S\} &\leq \delta, \\ \frac{1}{N}H(S) + \delta \geq \frac{1}{N}\log_2 |\mathcal{S}| &\geq R - \delta, \\ \frac{1}{N}I(S;M) &\leq \delta. \end{aligned} \tag{7}$$

Theorem (Ahlswede-Csiszar, 1993)

For a secret-generation system the maximum achievable secrecy rate is equal to $I(X;Y)$. We call this largest rate the **secrecy capacity** C_s.

QUESTION and REMARK:

- Only statement about FRR. What is the consequence of this theorem in terms of FAR?
- Note that an impostor has access to the helper data M.

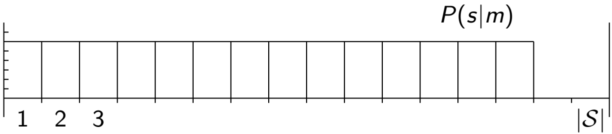
Next we will consider two distributions $P(m,s)$ realized by an encoder. The distributions satisfy the achievability constraints, hence

$$\begin{aligned} \frac{1}{N} I(S;M) &\leq \delta, \\ \frac{1}{N}H(S)+\delta \geq \frac{1}{N}\log_2|\mathcal{S}| &\geq R-\delta. \end{aligned} \tag{8}$$

INTRODUCTION: A distribution $P(s,m)$ with a small FAR

Authentication Based on Secret-Key Generation

Frans M.J. Willems (joint work w. Tanya Ignatenko)



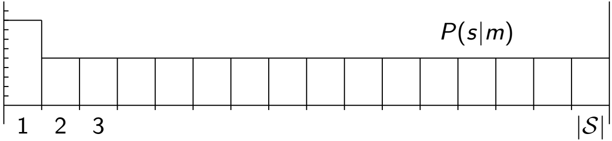
For each m let $P(s|m) = 1/(\alpha|\mathcal{S}|)$ or 0. Then an impostor can achieve

$$\begin{aligned} \log_2 \frac{1}{\text{FAR}} &= \log_2(\alpha|\mathcal{S}|) \\ &= H(S|M) \\ &= H(S) - I(S;M) \\ &\geq N(R-2\delta) - N\delta \\ &= N(R-3\delta). \end{aligned} \tag{9}$$

Therefore

$$\frac{1}{N}\log_2 \frac{1}{\text{FAR}} \geq R - 3\delta. \tag{10}$$

INTRODUCTION: A distribution $P(s, m)$ with a large FAR



For each m let $P(s|m) = 1 - \beta$ for a single s, and $\beta/(|\mathcal{S}| - 1)$ for all the others. Then

$$\begin{aligned} H(S|M) &= H(S) - I(S;M) \geq H(S) - N\delta = (H(S) + N\delta) - 2N\delta \\ H(S|M) &= h(\beta) + \beta \log_2(|\mathcal{S}| - 1) \\ &\leq 1 + \beta \log_2 |\mathcal{S}| \leq 1 + \beta(H(S) + N\delta). \end{aligned} \quad (11)$$

Hence

$$(1-\beta)(H(S) + N\delta) \leq 1 + 2N\delta, \quad (12)$$

$$\text{FAR} = (1-\beta) \leq \frac{1 + 2N\delta}{H(S) + N\delta} \leq \frac{3\delta}{R - \delta}, \quad (13)$$

for large enough N, and for a MAP-impostor

$$\frac{1}{N} \log_2 \frac{1}{\text{FAR}} \geq \frac{1}{N} \log_2 \frac{R - \delta}{3\delta}. \quad (14)$$

INTRODUCTION: Questions

- **CONCLUSION** is that, in the Ahlswede-Csiszar setting, a small $I(S;M)$ does not guarantee an exponentially small FAR.
- **QUESTION** is whether FAR $\approx 2^{-NC_s} = 2^{-NI(X;Y)}$ can be guaranteed in an authentication system based on secret-generation for all impostors.
- **QUESTION** is whether $I(X;Y)$ is a fundamental limit for the false-accept exponent, just as is it the fundamental limit for secret-key rate.
- **QUESTION** is (a) how to define achievability, (b) how to construct an achievability proof and a (c) converse that support the statement that $I(X;Y)$ is maximal false-accept exponent.

RESULT: Achievability and Result

Definition

False-accept exponent E is achievable if for all $\delta > 0$ and all N large enough there exists an encoder and a decoder such that

$$\text{FRR} \leq \delta, \tag{15}$$

while all impostor strategies will result in

$$\frac{1}{N} \log_2 \frac{1}{\text{FAR}} \geq E - \delta. \tag{16}$$

We will prove here the following result:

Theorem

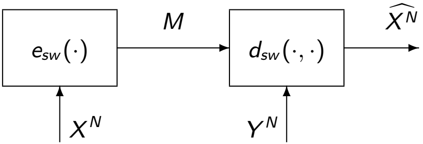
For a biometric authentication model based on secret-generation the maximum achievable false-accept exponent E is equal to $I(X;Y)$.

ACHIEVABILITY: Objective

Note that in our achievability proof we must demonstrate

- that there exist encoders and decoders that achieve the FRR constraint (15),
- and that guarantee, for all impostor strategies, that the FAR constraint (16) is met for $E = I(X;Y)$.

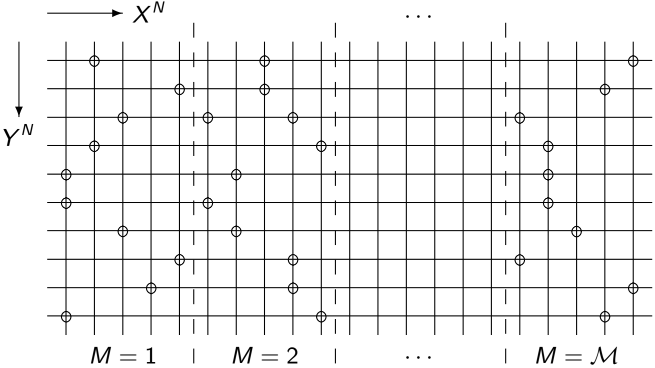
First we show that there exist a (Slepian-Wolf) code for reconstruction of $\widehat{X}^N$ by the decoder, see figure below.



This code defines the M-labeling.

It guarantees that $\Pr\{\widehat{X}^N \neq X^N\} \leq \delta$ for $|\mathcal{M}| = 2^{N(H(X|Y)+3\epsilon)}$ and N large enough.

PROOF:



Fix $\varepsilon > 0$ and an N. Consider the typical set $\mathcal{A}_{\varepsilon}^{N}(XY)$.
To each $x^N \in \mathcal{X}^N$ a label m that is **uniformly chosen** from $\{1, 2, \cdots, \mathcal{M}\}$ is assigned. Denote this label by $m(x^N)$. See figure above.

ENCODING: Upon observing x^N the encoder sends $m(x^N)$ to the decoder.
DECODING: The decoder chooses the unique $\widehat{x^N}$ such that $m(\widehat{x^N}) = m(x^N)$ and $(\widehat{x^N}, y^N) \in \mathcal{A}_\varepsilon^N(XY)$. If such an $\widehat{x^N}$ cannot be found, the decoder declares an error[1].
ERROR PROBABILITY: Averaged over the ensemble of labelings

$$\begin{aligned} &\Pr\{\widehat{X^N} \neq X^N\} \\ &= \Pr\left\{(X^N, Y^N) \notin \mathcal{A}_\varepsilon^N \cup \bigcup_{x^N \neq X^N, (x^N, Y^N) \in \mathcal{A}_\varepsilon^N} M(x^N) = M(X^N)\right\} \\ &\leq \Pr\{(X^N, Y^N) \notin \mathcal{A}_\varepsilon^N\} + \Pr\left\{\bigcup_{x^N \neq X^N, (x^N, Y^N) \in \mathcal{A}_\varepsilon^N} M(x^N) = M(X^N)\right\} \end{aligned} \quad (17)$$

First term, for N large enough, is

$$\Pr\{(X^N, Y^N) \notin \mathcal{A}_\varepsilon^N\} \leq \varepsilon. \quad (18)$$

[1] It is not important what value $\widehat{x^N}$ gets in that case.

Second term, again for N large enough, is

$$
\begin{aligned}
\Pr\left\{ \bigcup_{x^N \neq X^N, (x^N, Y^N) \in \mathcal{A}_\varepsilon^N} M(x^N) = M(X^N) \right\} \\
&\leq \sum_{x^N, y^N} P(x^N, y^N) \sum_{\tilde{x}^N \neq x^N, (\tilde{x}^N, y^N) \in \mathcal{A}_\varepsilon^N} \Pr\{M(\tilde{x}^N) = M(x^N)\} \\
&\leq \sum_{x^N, y^N} P(x^N, y^N) |\mathcal{A}_\varepsilon^N(X|y^N)| \frac{1}{|\mathcal{M}|} \\
&\leq 2^{N(H(X|Y)+2\varepsilon)} \frac{1}{2^{N(H(X|Y)+3\varepsilon)}} \\
&= 2^{-N\varepsilon} \\
&\leq \varepsilon, \qquad (19)
\end{aligned}
$$

when we take

$$|\mathcal{M}| = 2^{N(H(X|Y)+3\varepsilon)}. \qquad (20)$$

Averaged over the ensemble of M-labelings, the error probability is smaller than or equal to 2ε, for N large enough, hence **there exists** an M-labeling with

$$\Pr\{\widehat{X^N} \neq X^N\} \leq 2\varepsilon. \tag{21}$$

S-Labeling used by the encoder during enrollment:
ANY labeling $s(x^N) : \mathcal{X}^N \rightarrow \{1, 2, \cdots, |\mathcal{S}|\}$ for $x^N \in \mathcal{A}_\varepsilon^N(X)$, and $s(x^N) = \phi_e$ for $x^N \notin \mathcal{A}_\varepsilon^N(X)$.

Behavior of decoder:
The decoder outputs as estimated secret $s(\widehat{x^N})$, where $\widehat{x^N}$ is the output of the SW-decoder, if this decoder didn't declare an error, and ϕ_d if an error was declared by the SW-decoder.

Note that if no error occurred the SW-encoder input x^N and equal SW-decoder output $\widehat{x^N} \in \mathcal{A}_\varepsilon^N(X)$. This implies, that for an authorized individual, our encoder and decoder guarantee that

$$\text{FRR} = \Pr\{\widehat{S_y} \neq S\} \quad \leq \quad \Pr\{\widehat{X^N} \neq X^N\} \leq 2\varepsilon. \tag{22}$$

ACHIEVABILIY: FAR, S-Labeling

Fix a Slepian-Wolf code constructed before, and define for all $m \in \mathcal{M}$ the sets of typical sequences

$$\mathcal{A}(m) \triangleq \{x^N \in \mathcal{A}_\varepsilon^N(X) \text{ for which } m(x^N) = m\}. \tag{23}$$

Now consider an $m \in \mathcal{M}$.

- An impostor, knowing the helper message m, tries to pick a sequence z^N such that the resulting estimated secret $\widehat{S_z}$ is equal to the secret key S of the individual he claims to be.
- The impostor, knowing m, can decide for the most promising secret-key $\widehat{S_z}$ and then choose a z^N that results, together with m, in this most promising key.
- The impostor, knowing m, need only consider secrets $\widehat{S_z}$ that result from typical sequences, i.e. from $x^N \in \mathcal{A}(m)$. Other such sequences can not be output of the SW-decoder.

ACHIEVABILITY: Uniform S-labeling

For each m, we distribute all the sequences $x^N \in \mathcal{A}(m)$ roughly uniform over the s labels. All non-typical sequences get label ϕ_e.

- The number of typical sequences with label m is upper bounded by

$$\Pr\{M = m\}/2^{-N(H(X)+\varepsilon)}.$$

- Distributing these sequences over all s-labels uniformly leads to at most

$$\lceil \Pr\{M = m\}/(2^{-N(H(X)+\varepsilon)}|\mathcal{S}|)\rceil$$

  typical sequences having a certain secret label.
- The joint probability that m occurs and an impostor, knowing m, chooses the correct secret, is therefore upper-bounded by

$$\left\lceil \frac{\Pr\{M = m\}}{2^{-N(H(X)+\varepsilon)}|\mathcal{S}|} \right\rceil \cdot 2^{-N(H(X)-\varepsilon)}.$$

An upper bound for the FAR follows if we carry out the summation over all m. This results in

$$\begin{aligned}
\text{FAR} &\leq \sum_{m=1,|\mathcal{M}|} \left\lceil \frac{\Pr\{M=m\}}{2^{-N(H(X)+\varepsilon)}|\mathcal{S}|} \right\rceil \cdot 2^{-N(H(X)-\varepsilon)} \\
&\leq \sum_{m=1,|\mathcal{M}|} \left(\frac{\Pr\{M=m\}}{2^{-N(H(X)+\varepsilon)}|\mathcal{S}|} + 1 \right) 2^{-N(H(X)-\varepsilon)} \\
&= \sum_{m=1,|\mathcal{M}|} \frac{\Pr\{M=m\}}{2^{-N(H(X)+\varepsilon)}|\mathcal{S}|} 2^{-N(H(X)-\varepsilon)} + \sum_{m=1,|\mathcal{M}|} 2^{-N(H(X)-\varepsilon)} \\
&= 2^{-N(I(X;Y)-4\varepsilon)} + 2^{-N(I(X;Y)-4\varepsilon)} \\
&\leq 2^{-N(I(X;Y)-5\varepsilon)}, \qquad (24)
\end{aligned}$$

for large enough N, for all impostors, if we take the number of s-labels

$$|\mathcal{S}| = 2^{N(I(X;Y)-2\varepsilon)}. \qquad (25)$$

The upper bound (22) on the FRR and the upper bound (24) on the FAR, results in the **achievability of false-accept exponent $E = I(X;Y)$.**

CONVERSE: Definition set $\mathcal{B}(m)$ and B-function

We will show that for all encoders and decoders that achieve the FRR constraint (15), there is at least one impostor that does NOT satisfy the FAR constraint (16) for $E > I(X;Y)$.

First consider an encoder and decoder achieving (15). Now

$$\mathcal{B}(m) \stackrel{\Delta}{=} \{s : \text{there exists an } y^N \text{ such that } d(m, y^N) = s\}, \tag{26}$$

hence $\mathcal{B}(m)$ is the set of secrets that can be reconstructed from m.

Moreover let $B(\cdot,\cdot)$ be a function of s and m, such that $B(s,m) = 1$ for $s \in \mathcal{B}(m)$ and $B(s,m) = 0$ otherwise.
Next note that

$$\begin{aligned} \delta \geq \Pr\{\widehat{S_y} \neq S\} &\geq \sum_m \Pr\{M = m, S \notin \mathcal{B}(m)\} \\ &= P(B = 0), \end{aligned} \tag{27}$$

since $S \notin \mathcal{B}(M)$ will always lead to an error.

CONVERSE: A Conditional-MAP Impostor Strategy

An impostor chooses, knowing m, a target secret $\widehat{s_z} \in \mathcal{B}(m)$ with maximum conditional probability, i.e.,

$$\widehat{s_z}(m) = \arg \max_{s \in \mathcal{B}(m)} P(s|m). \tag{28}$$

Since the target secret can be realized, this impostor achieves

$$\begin{aligned} \text{FAR} &= \sum_m P(m) \max_{s \in \mathcal{B}(m)} P(s|m) \\ &= \sum_m P(m)P(B=1|m) \max_{s \in \mathcal{B}(m)} \frac{P(s|m)}{P(B=1|m)} \\ &= \sum_m P(m)P(B=1|m) \max_{s \in \mathcal{B}(m)} \frac{P(s,B=1|m)}{P(B=1|m)} \\ &= \sum_m P(m,B=1) \max_s P(s|m,B=1). \end{aligned} \tag{29}$$

CONVERSE: Conditional Entropy and FAR

Next we consider a relation between conditional entropy and FAR.

$$
\begin{aligned}
& H(S|M, B=1) \\
&= \sum_m P(m|B=1) \sum_s P(s|m, B=1) \log_2 \frac{1}{P(s|m, B=1)} \\
&\geq \sum_m P(m|B=1) \sum_s P(s|m, B=1) \log_2 \frac{1}{\max_s P(s|m, B=1)} \\
&= \sum_m P(m|B=1) \log_2 \frac{1}{\max_s P(s|m, B=1)} \\
&\geq \log_2 \frac{1}{\sum_m P(m|B=1) \max_s P(s|m, B=1)} \\
&= \log_2 \frac{P(B=1)}{\text{FAR}}. \qquad (30)
\end{aligned}
$$

See Feder and Merhav [1994], Ho and Verdu [2009].

CONVERSE: Conditional Entropy and Mutual Information

Now can combine

$$
\begin{aligned}
P(B=1)H(S|M,B=1) &\leq H(S|M,B) \\
&\leq H(S|M) \\
&\leq I(S;Y^N|M)+F \\
&\leq H(Y^N)-H(Y^N|M,S,X^N)+F \\
&= H(Y^N)-H(Y^N|X^N)+F \\
&= NI(X;Y)+F, \qquad (31)
\end{aligned}
$$

where $F = 1 + \Pr\{\widehat{S_y} \neq S\} \log_2 |\mathcal{X}|^N$, is the Fano-term.

CONVERSE: Wrap Up

Combining (30) and (31) we get

$$\begin{aligned} P(B=1)\log_2 \frac{P(B=1)}{\text{FAR}} &\leq P(B=1)H(S|M,B=1) \\ &\leq NI(X;Y)+1+\Pr\{\widehat{S_y}\neq S\}\log_2|\mathcal{X}|^N. \end{aligned} \quad (32)$$

Consider an achievable exponent E. Then

$$\begin{aligned} &P(B=1)N(E-\delta)+P(B=1)\log_2 P(B=1) \\ &\leq \quad NI(X;Y)+1+\Pr\{\widehat{S_y}\neq S\}\log_2|\mathcal{X}|^N. \end{aligned} \quad (33)$$

If we now let $\delta\downarrow 0$ and $N\to\infty$ then since $\Pr\{\widehat{S_y}\neq S\}\leq\delta$, and $P(B=1)\geq 1-\Pr\{\widehat{S_y}\neq S\}\geq 1-\delta$, we get that

$$E\leq I(X;Y). \quad (34)$$

PRIVACY LEAKAGE: Introduction

Consider the mutual information $I(X^N; M)$ of the biometric sequence X^N and the helper data M. This mutual information is what we call the privacy-leakage. We can write for our code that demonstrates the achievability of $E = I(X;Y)$ that

$$\begin{aligned} I(X^N; M) &\leq H(M) \\ &\leq \log_2 |\mathcal{M}| \\ &= N(H(X|Y) + 3\epsilon) \end{aligned} \tag{35}$$

QUESTION: What is the trade-off between false-accept exponent and privacy-leakage rate?

TRADE-OFF False-Accept Exponent and Privacy-Leakage Rate

Consider again our authentication system based on secret generation.

Definition

False-accept exponent - privacy-leakage rate combination (E, L) is achievable if for all $\delta > 0$ and all N large enough there exist encoders and decoders such that

$$\begin{aligned} \text{FRR} &\leq \delta, \\ \frac{1}{N} I(X^N; M) &\leq L + \delta, \end{aligned} \tag{36}$$

while all impostor strategies will result in

$$\frac{1}{N} \log_2 \frac{1}{\text{FAR}} \geq E - \delta. \tag{37}$$

The region of achievable exponent-rate combinations is defined as $\mathcal{R}_{\text{EL}}$.

TRADE-OFF: Achievability and Result

We will prove here the following result:

Theorem

For a biometric authentication system based on secret-generation the region $\mathcal{R}_{EL}$ of achievable false-accept exponent - privacy-leakage combinations satifisfies

$$\mathcal{R}_{EL} = \{(E, L) \quad : \quad 0 \leq E \leq I(U;Y), \\ L \geq I(U;X) - I(U;Y), \\ \text{for } P(u,x,y) = Q(x,y)P(u|x)\} \tag{38}$$

(A) In the achievability part we will transform the biometric source (X, Y) into a source (Q, Y^N) with roughly $H(Y^N|Q) \leq NH(Y|U)$ and $Q \in \{\phi_q, 1, 2, \cdots, |\mathcal{Q}|\}$ with roughly $|\mathcal{Q}| = 2^{NI(U;X)}$. We can say that Q is a quantized version of X^N. For this new source we use the achievability part of the first theorem.
(B) The converse part is standard.

TRADE-OFF (Ach.): Modified Weakly Typical Sets

Fix a $P(u|x)$. Let $0 < \varepsilon < 1$. For the properties of $\mathcal{A}_{\varepsilon}^{N}$ we refer to Cover and Thomas [2006].

Definition

Assuming that transition probability matrix $P(u|x)$ determines the joint probability distribution $P(u,x,y) = Q(x,y)P(u|x)$ we define

$$\mathcal{B}_{\varepsilon}^{N}(UX) \triangleq \{(u^N, x^N) : \Pr\{Y^N \in \mathcal{A}_{\varepsilon}^{N}(Y|u^N, x^N)|(U^N, X^N) = (u^N, x^N)\} \geq 1 - \sqrt{\epsilon}\}, \quad (39)$$

where Y^N is the output sequence of a "channel" $Q(y|x) = Q(x,y)/\sum_x Q(x,y)$ when sequence x^N is input.

TRADE-OFF (Ach.): Two Properties

Property

If $(u^N, x^N) \in \mathcal{B}_\varepsilon^N(UX)$ *then also* $(u^N, x^N) \in \mathcal{A}_\varepsilon^N(UX)$.

Property

Let (U^N, X^N) *be i.i.d. according to* $P(u,x)$ *then*

$$\Pr\{(U^N, X^N) \in \mathcal{B}_\varepsilon^N(UX)\} \geq 1 - \sqrt{\epsilon} \tag{40}$$

for all large enough N.

TRADE-OFF (Ach.): Proofs of the Two Properties

1. Observe that $(u^N, x^N) \in \mathcal{B}_\varepsilon^N(UX)$ implies that at least one y^N exist such that $(u^N, x^N, y^N) \in \mathcal{A}_\varepsilon^N(UXY)$. Thus $(u^N, x^N) \in \mathcal{A}_\varepsilon^N(UX)$.
2. When (U^N, X^N, Y^N) is i.i.d. with respect to $P(u, x, y)$ then

$$\begin{aligned}
&\Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_\varepsilon^N(UXY)\} \\
&\le \sum_{(u^N,x^N)\in\mathcal{B}_\varepsilon^N(UX)} P(u^N, x^N) + \sum_{(u^N,x^N)\notin\mathcal{B}_\varepsilon^N(UX)} P(u^N, x^N)(1-\sqrt{\epsilon}) \\
&= 1-\sqrt{\epsilon}+\sqrt{\epsilon}\Pr\{(U^N, X^N) \in \mathcal{B}_\varepsilon^N(UX)\}, \qquad (41)
\end{aligned}$$

or

$$\begin{aligned}
&\Pr\{(U^N, X^N) \in \mathcal{B}_\varepsilon^N(UX)\} \\
&\ge 1 - \frac{1-\Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_\varepsilon^N(UXY)\}}{\sqrt{\epsilon}}. \qquad (42)
\end{aligned}$$

The weak law of large numbers implies that $\Pr\{(U^N, X^N, Y^N) \in \mathcal{A}_\varepsilon^N(UXY)\} \ge 1-\epsilon$ for large enough N. From (42) we now obtain the second property.

TRADE-OFF (Ach.): A Quantizer of $\mathcal{X}^N$

- **Random coding:** For each index $q \in \{1, 2, \cdots, |\mathcal{Q}|\}$ generate an auxiliary sequence $u^N(q)$ at random according to $P(u) = \sum_{x,y} Q(x,y)P(u|x)$.
- **Quantizing:** When x^N occurs, let Q be the smallest value of q such that $(u^N(q), x^N) \in \mathcal{B}_\varepsilon^N(UX)$. If no such q is found set $Q = \phi_q$.
- **Events:** Let X^N and Y^N be the observed biometric source sequences, Q the index determined by the quantizer. Define, for $q = 1, 2, \cdots, |\mathcal{Q}|$, the events:

$$E_q \triangleq \{(u^N(q), X^N) \in \mathcal{B}_\varepsilon^N(UX)\}. \tag{43}$$

TRADE-OFF (Ach.): A Quantizer ...

As in Gallager [1968], p. 454, we write

$$
\begin{aligned}
\Pr\left\{\bigcap_q E_q^c\right\} &= \sum_{x^N\in\mathcal{X}^N} Q(x^N)\prod_q\Big(1-\sum_{u^N\in\mathcal{B}_\varepsilon^N(U|x^N)} P(u^N)\Big)\\
&\overset{(a)}{\leq} \sum_{x^N\in\mathcal{X}^N} Q(x^N)\Big(1-2^{-N(I(U;X)+3\varepsilon)}\cdot\sum_{u^N\in\mathcal{B}_\varepsilon^N(U|x^N)} P(u^N|x^N)\Big)^{|\mathcal{Q}|}\\
&\overset{(b)}{\leq} \sum_{x^N\in\mathcal{X}^N} Q(x^N)\Big(1-\sum_{u^N\in\mathcal{B}_\varepsilon^N(U|x^N)} P(u^N|x^N)\\
&\qquad\qquad + \exp(-|\mathcal{Q}|2^{-N(I(U;X)+3\varepsilon)})\Big)\\
&\leq \sum_{(u^N,x^N)\notin\mathcal{B}_\varepsilon^N(UX)} P(u^N,x^N)+\sum_{x^N\in\mathcal{X}^N} Q(x^N)\exp(-2^{N\varepsilon})\\
&\overset{(c)}{\leq} 2\sqrt{\varepsilon}, \qquad (44)
\end{aligned}
$$

for N large enough, if $|\mathcal{Q}| = 2^{N(I(U;X)+4\varepsilon)}$.

TRADE-OFF (Ach.): A Quantizer ...

Here (a) follows from the fact that for $(u^N, x^N) \in \mathcal{B}_\varepsilon^N(UX)$, using the first property, we get

$$\begin{aligned} P(u^N) &= P(u^N|x^N)\frac{Q(x^N)P(u^N)}{P(x^N,u^N)} \\ &\geq P(u^N|x^N)\frac{2^{-N(H(X)+\varepsilon)}2^{-N(H(U)+\varepsilon)}}{2^{-N(H(U,X)-\varepsilon)}} \\ &= P(u^N|x^N)2^{-N(I(U;X)+3\varepsilon)}, \end{aligned} \quad (45)$$

(b) from the inequality $(1-\alpha\beta)^K \leq 1-\alpha+\exp(-K\beta)$, which holds for $0 \leq \alpha, \beta \leq 1$ and $K > 0$, and (c) from the second property.

We have shown that, over the ensemble of auxiliary sequences, for N large enough, $\Pr\{Q = \phi_q\} \leq 2\sqrt{\varepsilon}$.
Therefore there exists a set of auxiliary sequences achieving

$$\Pr\{Q = \phi_q\} \leq 2\sqrt{\varepsilon}. \quad (46)$$

Consider such a set of auxiliary sequences (a quantizer).

TRADE-OFF (Ach.): A Quantizer ...

With probability $\geq 1 - 2\sqrt{\varepsilon}$ an x^N occurs for which there is a q such that $(u^N(q), x^N) \in \mathcal{B}_\varepsilon^N(UX)$.
Then, with probability $\geq 1 - \sqrt{\varepsilon}$ the observed y^N is in $\mathcal{A}_\varepsilon^N(Y|u^N(q), x^N)$ and consequently in $\mathcal{A}_\varepsilon^N(Y|u^N(q))$. Furthermore note that $|\mathcal{A}_\varepsilon^N(Y|u^N(q))| \leq 2^{N(H(Y|U)+2\varepsilon)}$.

Now:

$$\begin{aligned} H(Y^N|Q) &\leq 2\sqrt{\varepsilon}\log_2|\mathcal{Y}|^N + (1-2\sqrt{\varepsilon}) + (1-2\sqrt{\varepsilon})\sqrt{\varepsilon}\log_2|\mathcal{Y}|^N \\ &\quad +(1-2\sqrt{\varepsilon})(1-\sqrt{\varepsilon})\log_2 2^{N(H(Y|U)+2\varepsilon)} \\ &\leq N(1-3\sqrt{\varepsilon}+2\varepsilon)H(Y|U) + N(3\sqrt{\varepsilon}-2\varepsilon)\log_2|\mathcal{Y}| \\ &\quad +(1-2\sqrt{\varepsilon}). \end{aligned} \tag{47}$$

By decreasing ε and increasing N, we can get $H(Y^N|Q)/N$ arbitrarily close to $H(Y|U)$, or

$$I(Q;Y^N)/N = H(Y) - H(Y^N|Q)/N \tag{48}$$

arbitrary close to $I(U;Y)$.

TRADE-OFF (Ach.): A Quantizer ...

Moreover in the same way we can get

$$\begin{aligned} H(Q|Y^N)/N &= H(Q)/N + H(Y^N|Q)/N - H(Y) \\ &\leq I(U;X) + 4\epsilon + H(Y^N|Q)/N - H(Y) \end{aligned} \tag{49}$$

arbitrary close to $I(U;X) - I(U;Y)$.

We apply the achievability proof for the basic theorem now. This leads to the achievability of the combination

$$(E, L) = (I(U;Y), I(U;X) - I(U;Y)). \tag{50}$$

TRADE-OFF: Converse

We only consider the basic steps. First we bound

$$\begin{aligned} H(S|M) &\leq I(S;Y^N|M) + H(S|Y^N,M) \\ &\leq I(S;Y^N|M) + H(S|\widehat{S_y}) \\ &\leq I(S,M;|Y^N) + F \\ &= \sum_{n=1,N} I(S,M;Y_n|Y^{n-1}) + F \\ &= \sum_{n=1,N} I(S,M,Y^{n-1};Y_n) + F \\ &\leq \sum_{n=1,N} I(S,M,Y_{n-1},X^{n-1};Y_n) + F \\ &= \sum_{n=1,N} I(S,M,X^{n-1};Y_n) + F, \end{aligned} \tag{51}$$

where $F \triangleq 1 + \delta \log |\mathcal{X}|^N$.
This is plugged into the FAR part of the basic converse.

Now we continue with the privacy leakage.

$$
\begin{aligned}
I(X^N;M) &= H(M) - H(M|X^N) \\
&\geq H(M|Y^N) - H(S,M|X^N) \\
&= H(S,M|Y^N) - H(S|M,Y^N,\widehat{S_y}) - H(S,M|X^N) \\
&\geq H(S,M|Y^N) - H(S|\widehat{S_y}) - H(S,M|X^N) \\
&\geq H(S,M|Y^N) - F - H(S,M|X^N) \\
&= I(S,M;X^N) - I(S,M;Y^N) - F \\
&= \sum_{n=1,N} I(S,M;X_n|X^{n-1}) - \sum_{n=1,N} I(S,M;Y_n|Y^{n-1}) - F \\
&= \sum_{n=1,N} I(S,M,X^{n-1};X_n) - \sum_{n=1}^{N} I(S,M,Y^{n-1};Y_n) - F \\
&\geq \sum_{n=1,N} I(S,M,X^{n-1};X_n) - \sum_{n=1,N} I(S,M,X^{n-1};Y_n) - F,
\end{aligned} \tag{52}
$$

where $(S,M,X^{n-1}) - X_n - Y_n$. Etc.

CONCLUSION

- **One-way function** can be used to convey the secret S safely to the authenticator.
- We extended the work of Ahlswede-Csiszar [1993] on the **secrecy capacity** to authentication with an impostor that has access to the helper-message. We found the fundamental limit on the false-accept exponent. As expected it is equal to the secrecy capacity.
- We also determined the fundamental **trade-off** between false-accept exponent and privacy-leakage rate. In this way we strengthened the results of Ignatenko-W [2008,2009] and Lai, Ho, and Poor [2008,2011] on the trade-off between secret-key rate and privacy-leakage rate. Again there is no difference in regions.
- Related to hypothesis testing literature (Ahlswede-Csiszar [1986]), ... , however impostor can use M.
- Extensions to **identification with helper data** and FAR with impostor?
- **Code constructions**. In the binary symmetric case **fuzzy commitment** (Juels and Wattenberg [1999]) could be fine.
- Comparison to **unprotected case** shows that same FAR-exponent can be achieved. Leakage case different.
- Authentication models **not based on secret generation**.