Information-theoretic results for hypothesis testing across networks

Michèle Wigger
Telecom ParisTech
Joint work with R. Timo, S. Salehkalaibar, and L. Wang

Telecom ParisTech Research Day
October 10, 2017
Example 1: Earthquake Alert Systems

- Distributed sensors measure ground vibrations

- **Seismic activities:** correlated vibrations at distributed sensors

- **Local disturbances (cars, metro):** independent vibrations at various sensors

**Task of Distributed Alert System**

Decide on joint distribution underlying the sensors’ observations
Example 2: Distributed Control-System for Smart Cars

- Smart cars measuring speed, distance, road conditions
- Fixed road-side sensors measuring same parameters

- Intact car system: measurements highly correlated
- Erroneous car system: measurements independent

Task of Distributed Control-System

Decide on joint distribution underlying the observations
“Historical” Distributed Hypothesis Testing

\[ X^n \quad \text{Sensor} \quad nR \text{ bits} \quad Y^n \quad \text{Decision Center} \quad \text{Alarm/No alarm} \]

- Normal situation: \((X^n, Y^n) \sim \text{i.i.d. } P_{XY}\)
- Hazardous event (to be detected): \((X^n, Y^n) \sim \text{i.i.d. } Q_{XY}\)
- False alarm probability should be small
- Missed detection probability should be \textit{exponentially small}: \(2^{-n\theta}\)

Largest possible error exponent \(\theta^*(R)\) for given rate \(R > 0\)
Local Hypothesis Testing

- Rate $R$ is so large that sensor can send all $X^n$ to decision center.
- Decision center applies Neyman-Pearson test to both $(X^n, Y^n)$.

$$\theta^*(R = \infty) = D(P_{XY} \parallel Q_{XY})$$

- Alternative: Decision center raises alarm if statistics of $(X^n, Y^n)$ does not correspond to $P_{XY}$. 
Back to Distributed Hypothesis Testing (Han’87)

- Sensor quantizes $X^n$ to $S^n(j)$

- Center raises alarm unless $(X^n, S^n(j))$ and $(S^n(j), Y^n)$ have correct statistics for $P_{S|X} P_{XY}$

Achievable Rate-Exponent Tradeoff

$$\theta^*(R) \geq \max_{P_{S|X} : R \geq I(S;X)} \text{min}_{\tilde{P}_{SXY} : \tilde{P}_{SX} = P_{SX}, \tilde{P}_{SY} = P_{SY}} D(\tilde{P}_{SXY} \| P_{S|X} Q_{XY})$$
• Communication channel now is noisy and memoryless (can model fast fading, additive noise, etc.)
• Use the hypothesis testing scheme from before

• Channel code: *unequal error protection* code that specially protects the “alarm”-message (1111111)

• New error events related to erroneous decoding (cases where noise-free comm. leads to correct decision)
Result on Testing over Noisy Channels

Achievable Exponent (Salehkalaibar&W’2017)

\[ \theta^* \geq \max_{P_{S|X}, P_{QW}} \min \left\{ \theta_{\text{standard}}, \theta_{\text{wrong-dec.}}, \theta_{\text{missed-(111111)}} \right\}, \]

\[ I(S;X|Y) \leq I(W;V|Q) \]

where

\[ \theta_{\text{standard}} = \min_{\tilde{P}_{SXY}} D(\tilde{P}_{SXY} \parallel Q_{XY} P_{S|X}), \]

\[ \tilde{P}_{SX}=P_{SX}, \quad \tilde{P}_{SY}=P_{SY} \]

\[ \theta_{\text{wrong-dec.}} = \min_{\tilde{P}_{SXY}} D(\tilde{P}_{SXY} \parallel P_{S|X} Q_{XY}) + I(V;W|Q) - I(S;X|Y), \]

\[ \tilde{P}_{SX}=P_{SX}, \quad \tilde{P}_{SY}=P_{SY}, \quad H(S|Y)\leq H_P(S|Y) \]

\[ \theta_{\text{missed-(111111)}} = D(P_{Y} \parallel Q_{Y}) + I(V;W|Q) - I(S;X|Y) + E_Q[D(P_{V|Q} \parallel P_{V|W=Q})] \]

Sometimes optimal. Sometimes noisy channel causes no penalty.
Result on Testing over Noisy Channels

Achievable Exponent (Salehkalaibar&W’2017)

$$\theta^* \geq \max_{P_{S|X}, P_{QW}} \min \left\{ \theta_{\text{standard}}, \theta_{\text{wrong-dec.}}, \theta_{\text{missed-(111111)}} \right\},$$

where

$$\theta_{\text{standard}} = \min_{\tilde{P}_{SXY}} D(\tilde{P}_{SXY} || Q_{XY} P_{S|X}),$$

$$\tilde{P}_{SX} = P_{SX}, \quad \tilde{P}_{SY} = P_{SY}$$

$$\theta_{\text{wrong-dec.}} = \min_{\tilde{P}_{SXY}} D(\tilde{P}_{SXY} || P_{S|X} Q_{XY}) + I(V; W|Q) - I(S; X|Y),$$

$$\tilde{P}_{SX} = P_{SX}, \quad \tilde{P}_{SY} = P_{SY}, \quad H(S|Y) \leq H_P(S|Y)$$

$$\theta_{\text{missed-(111111)}} = D(P_Y || Q_Y) + I(V; W|Q) - I(S; X|Y) + E_Q[D(P_{V|Q} || P_{V|W=Q})]$$

Sometimes optimal. Sometimes noisy channel causes no penalty.
Two Simultaneous Hypothesis Tests

Setup can model:

- Two different decision centers
- Single decision center with uncertain $P_{XY}$
- Single decision center raising two types of alarms

Tension: Communication needs to serve both decisions!
Tradeoff Between Exponents at Two Centers

- Optimal exponents region for a Gaussian example based on adapted “quantization” method (Salehkalaibar/W’/Timo’2017)
Multi-Hop Channel

- Tension on first communication link: communication serves relay and decision center
- Intermediate processing/decisions at relay
Markov chain $X^n \rightarrow Y^n \rightarrow Z^n$

Accumulation of Exponents (Salehkalaibar/W’/Wang 2017)

$$\theta(R, T) = \theta_{\text{Sensor} \rightarrow \text{Relay}}(R) + \tilde{\theta}_{\text{Relay} \rightarrow \text{Decision}}(T)$$

$\tilde{\theta}_{\text{Relay} \rightarrow \text{Decision}}(T)$ modified exponent where $Y^n$ always $\sim P_Y$

- Achieved by: “raise-alarm” as soon as someone says “alarm”
Communication Constraint Requires Relay Processing

\[ X^n, Y^n, \text{ i.i.d. } \mathcal{B}(1/2) \]

and

\[ Z_t = \begin{cases} 
(X_t, Y_t) & \text{if } Y_t = 1 \text{ and under normal situation} \\
(X'_t, Y_t) & \text{otherwise} 
\end{cases} \]

- If \( T \geq R \): \( Y^n \) “useless” and relay simply forwards message
Communication Constraint Requires Relay Processing

$X^n, Y^n, \quad \text{i.i.d. } \mathcal{B}(1/2)$

and

$$Z_t = \begin{cases} (X_t, Y_t) & \text{if } Y_t = 1 \text{ and under normal situation} \\ (X'_t, Y_t) & \text{otherwise} \end{cases}$$

- If $T \geq R$: $Y^n$ “useless” and relay simply forwards message
- If $T < R$: relay extracts relevant info, i.e., positions where $Y_t = 0$. 
Summary

- Distributed hypothesis testing for multi-hop and one-to-many networks and for noisy channels

- Schemes based on improved quantization methods, unanimous-decision forwarding, and unequal error protection
  - Accumulation of error exponents
  - Competition for network resources → tradeoff in exponents
  - Intermediate processing required for optimal communication

- Derived error exponents achieve fundamental limits in some cases