IMPROVING WIRELESS SECURITY THROUGH NETWORK DIVERSITY

Darryl Veitch[1]
&
Tao Ye[2] Jean Bolot[2]

[1]CUBIN, Department of Electrical & Electronic Engineering, University of Melbourne, Australia

[2]Sprint Advanced Technology Laboratories, Burlingame CA, USA

IPAM Program on Internet Multi-Resolution Analysis:
Foundations, Applications and Practice

Workshop II: Applications of Internet MRA to Cyber-Security

The Idea

Backdrop

- Mobile devices are ubiquitous
- Mobile devices have become a critical service
- Availability and security will increase in importance
- Data confidentiality a key component, especially for wireless

Aim

To exploit multiple channels to increase confidentiality

The Idea

Backdrop

- Mobile devices are ubiquitous
- Mobile devices have become a critical service
- Availability and security will increase in importance
- Data confidentiality a key component, especially for wireless

Opportunity: Modern wireless devices have multiple interfaces

Aim

To exploit multiple channels to increase confidentiality

The Idea

Backdrop

- Mobile devices are ubiquitous
- Mobile devices have become a critical service
- Availability and security will increase in importance
- Data confidentiality a key component, especially for wireless

Opportunity: Modern wireless devices have multiple interfaces

Aim

To exploit multiple channels to increase confidentiality

Benefits of Multiple Channels

- 'Physical Difficulty'
  - Cost (eg. CDMA)
  - Technical (perfect sniffing difficult)
- Leverage link heterogenity
  - Small, 'key' portion over expensive channel
  - Combine redundancy with higher security
- Bandwidth/Latency benefits
- Multichannel encryption

But We Already Have Strong Encryption?

Four Possible Objections

- Can't have too much security - may already be compromised
- Vulnerabilities may be known but still there - WEP still common
- Strong encryption not always available - Webmail
- Strong encryption can be too expensive - battery limitations

If we can add an extra layer of protection at low cost, we should

But We Already Have Strong Encryption?

Four Possible Objections

- Can't have too much security - may already be compromised
- Vulnerabilities may be known but still there - WEP still common
- Strong encryption not always available - Webmail
- Strong encryption can be too expensive - battery limitations

If we can add an extra layer of protection at low cost, we should

Literature

Using Multiple Links

- Multiple links suggested for bandwidth
- We focus on confidentiality
- Multipath routing used in MANET based on physical difficulty
  – redundancy focus → bandwidth overhead
  – expensive encoding/decoding
- We consider small number of reliable paths
  – look for low cost methods of data splitting
  – add security on top of physical difficulty (unlike Secret Sharing)

MEO: Multichannel Encryption Overlay

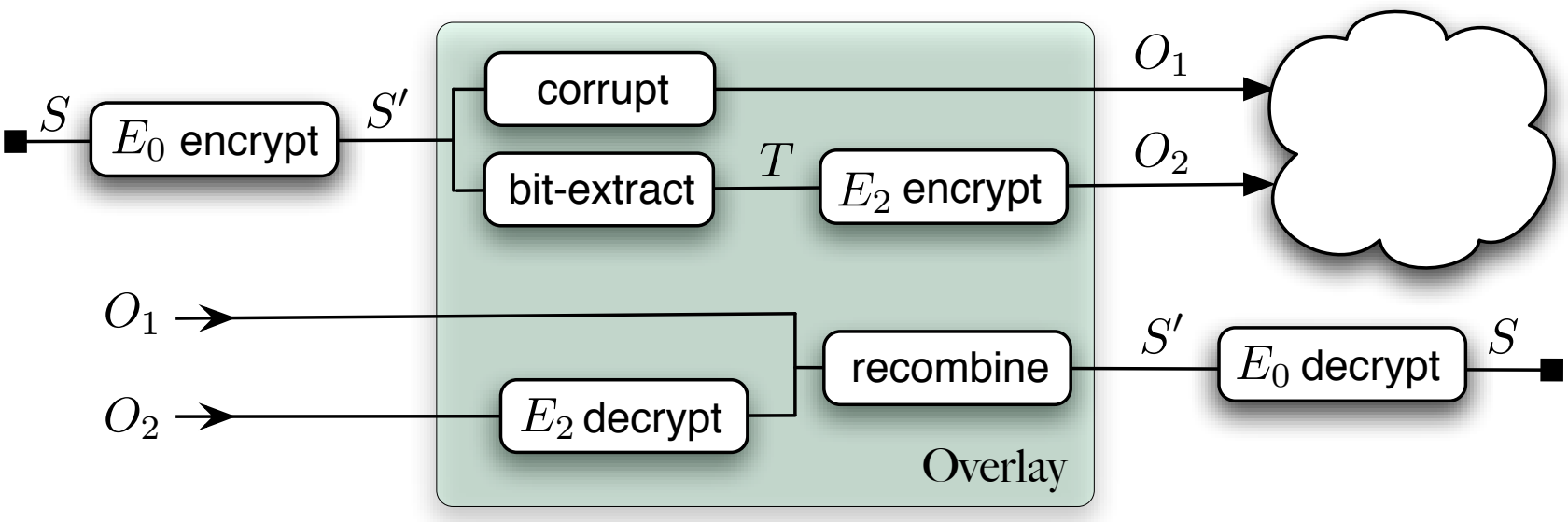
Based on Two Principles

- Data Corruption to thwart cracking
- Information Reduction to slow cracking

MEO: Multichannel Encryption Overlay

Based on Two Principles

- Data Corruption to thwart cracking
- Information Reduction to slow cracking



MEO: Definition

Input to Overlay

- Packet stream S' of packet rate λ_0, size p_0 bytes
- Data rate $r_0 = 8p_0\lambda_0$

Channel 1

- Each packet from S' corrupted by randomly removing b bits.
- Output stream O_1 of packet rate $\lambda_1 = \lambda_0$, size $p_1 = p_0 - b/8$
- Data rate $r_1 = r_0 - b\lambda_0$

Channel 2

- Extracted bits assembled into packets of size p_2, then encrypted
- Output stream O_2 of packet rate $\lambda_2 = b\lambda_0/8p_2$
- Data rate $r_2 = br_0/8p_0$

Note: MEO is bandwidth preserving: $r_0 = r_1 + r_2$

MEO: Definition

Input to Overlay

- Packet stream S' of packet rate λ_0, size p_0 bytes
- Data rate $r_0 = 8p_0\lambda_0$

Channel 1

- Each packet from S' corrupted by randomly removing b bits.
- Output stream O_1 of packet rate $\lambda_1 = \lambda_0$, size $p_1 = p_0 - b/8$
- Data rate $r_1 = r_0 - b\lambda_0$

Channel 2

- Extracted bits assembled into packets of size p_2, then encrypted
- Output stream O_2 of packet rate $\lambda_2 = b\lambda_0/8p_2$
- Data rate $r_2 = br_0/8p_0$

Note: MEO is bandwidth preserving: $r_0 = r_1 + r_2$

MEO: Definition

Input to Overlay

- Packet stream S' of packet rate λ_0, size p_0 bytes
- Data rate $r_0 = 8p_0\lambda_0$

Channel 1

- Each packet from S' corrupted by randomly removing b bits.
- Output stream O_1 of packet rate $\lambda_1 = \lambda_0$, size $p_1 = p_0 - b/8$
- Data rate $r_1 = r_0 - b\lambda_0$

Channel 2

- Extracted bits assembled into packets of size p_2, then encrypted
- Output stream O_2 of packet rate $\lambda_2 = b\lambda_0/8p_2$
- Data rate $r_2 = br_0/8p_0$

Note: MEO is bandwidth preserving: $r_0 = r_1 + r_2$

Cracking Model

Requirements

- Meaningful even for 'plug-in' ciphers E_0 and E_2
- Tractable

Consider class vulnerable to sniffed cipher-text

Definition

Let the r.v. $N \geq 0$ give the number of packets needed to recover the message. Define 'cracking time' by $T = N/\lambda$.

(Consistent with Shannon's *equivocation measure*.)

Cracking of Augmented System

(I) Cracking O_1 on channel 1 to S (channel 2 not needed)
(II) Cracking O_2 on channel 2 to S (channel 1 not needed)
(III) Cracking the overlay, then cracking E_0 to S.

Analysis ignores physical difficulty!

Cracking Model

Requirements

- Meaningful even for 'plug-in' ciphers E_0 and E_2
- Tractable

Consider class vulnerable to sniffed cipher-text

Definition

Let the r.v. $N \geq 0$ give the number of packets needed to recover the message. Define 'cracking time' by $T = N/\lambda$.

(Consistent with Shannon's *equivocation measure*.)

Cracking of Augmented System

(I) Cracking O_1 on channel 1 to S (channel 2 not needed)
(II) Cracking O_2 on channel 2 to S (channel 1 not needed)
(III) Cracking the overlay, then cracking E_0 to S.

Analysis ignores physical difficulty!

Cracking Model

Requirements

- Meaningful even for 'plug-in' ciphers E_0 and E_2
- Tractable

Consider class vulnerable to sniffed cipher-text

Definition

Let the r.v. $N \geq 0$ give the number of packets needed to recover the message. Define 'cracking time' by $T = N/\lambda$.

(Consistent with Shannon's *equivocation measure*.)

Cracking of Augmented System

(I) Cracking O_1 on channel 1 to S (channel 2 not needed)
(II) Cracking O_2 on channel 2 to S (channel 1 not needed)
(III) Cracking the overlay, then cracking E_0 to S.

Analysis ignores physical difficulty!

Impact of Corruption

Since packet is cipher text, don't know when guess is good $\Longrightarrow$ brute force

Assume failed with $n = 1$ packets, trying n:

Positions known:

$$(2^b)^n = 2^{bn}$$

Positions unknown:

$$\left(\binom{8p_0}{b} 2^b\right)^n$$

Example: $n = 1, b = 3$ bit, processing time increased by 9422443520 ...

Conclude that cracking via (i) not feasible.

Impact of Information Reduction

Recall that packet rate λ_2 is reduced, hence

Average cracking time for E_2:

$$\mu_{T_2} = \frac{\mu_{N_2}}{\lambda_2} = \frac{\mu_{N_2}}{\lambda_0} \cdot \frac{8p_2}{b}$$

A similar slow down in cracking holds for quantiles of cracking time r.v. T_2

Example: $p_2 = 32$ byptes, $b = 1$ bit, increases cracking time by $257\times$.

Conclude that cracking via (iii) too slow to be practical.

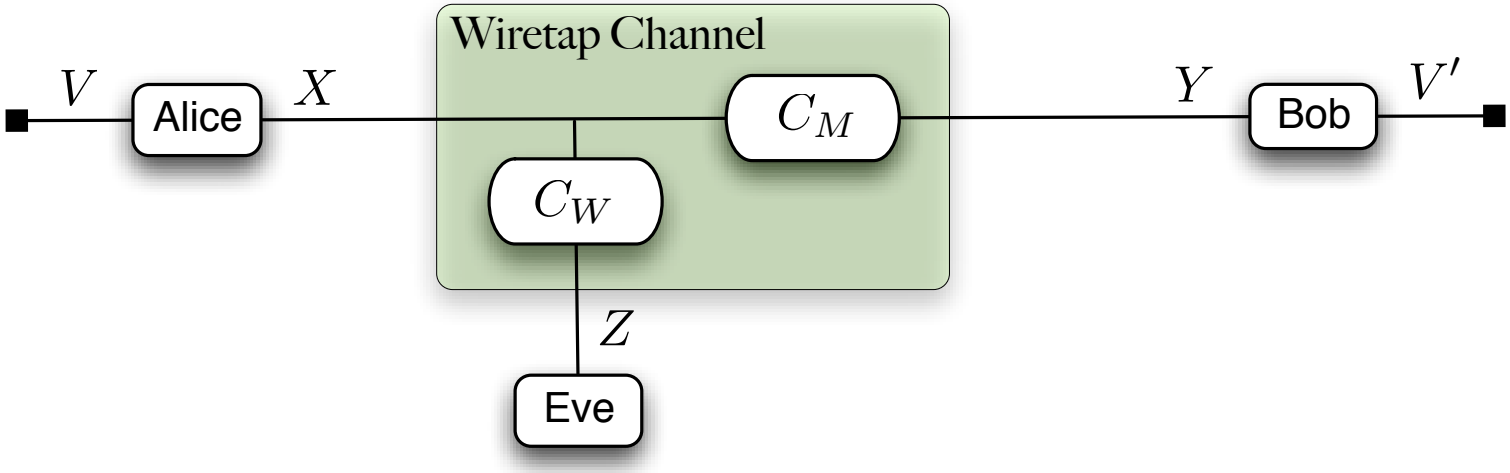
WEP Illustration

Testing the Corruption Principle

- Use Bittau's simulator: generates encrypted packets and feeds them to Aircrack software
- Measure distribution of N before: 75% cracked
- Corrupt by shifting payload $b = 2$ bits to left, zero padding
- Measure after: 0% cracked

Information Theoretical Security

Based on the Wiretap Channel (Wyner 1975)



Very Different From Cryptographic Model

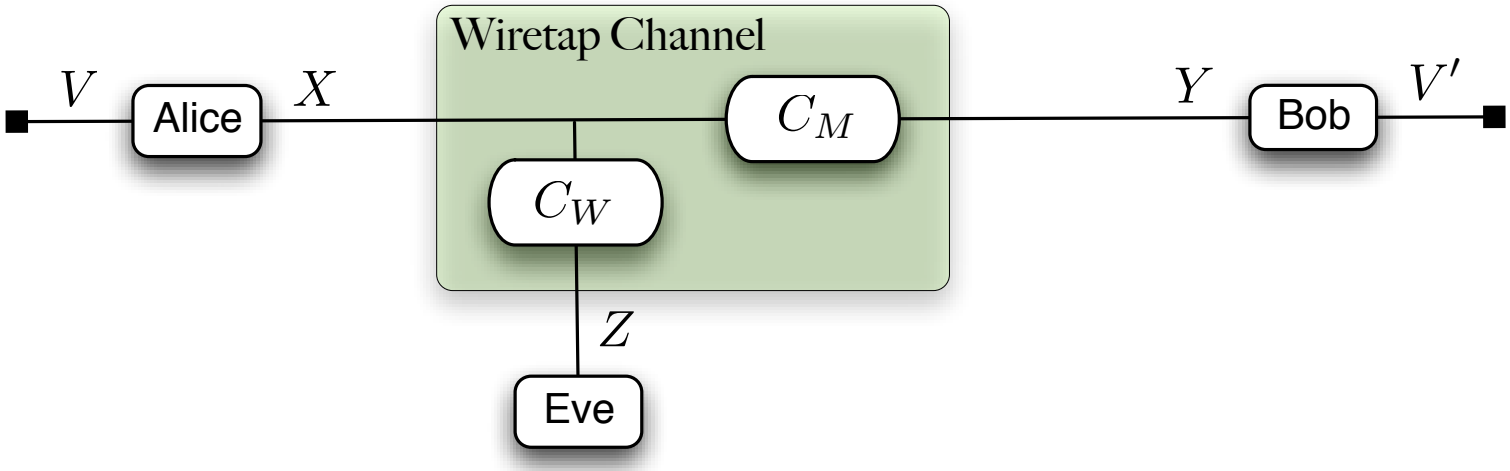
- Not based on keys, but security capacity: $C_s = \max_{p(x)} [I(X;Y) - I(X;Z)]$
- Goals are: Reliability: $P\{V \neq V'\} \to 0$
  Security: $I(V;Z)/n \to 0$, as $n \to \infty$

When the main channel is less noisy than the wiretapper's channel, and $I(X;Y)$, $I(X;Z)$ maximized by the same $p(x)$, then (van Dijk 1997):

$$C_s = \text{Capacity}\,(C_M) - \text{Capacity}\,(C_W)$$

Information Theoretical Security

Based on the Wiretap Channel (Wyner 1975)



Very Different From Cryptographic Model

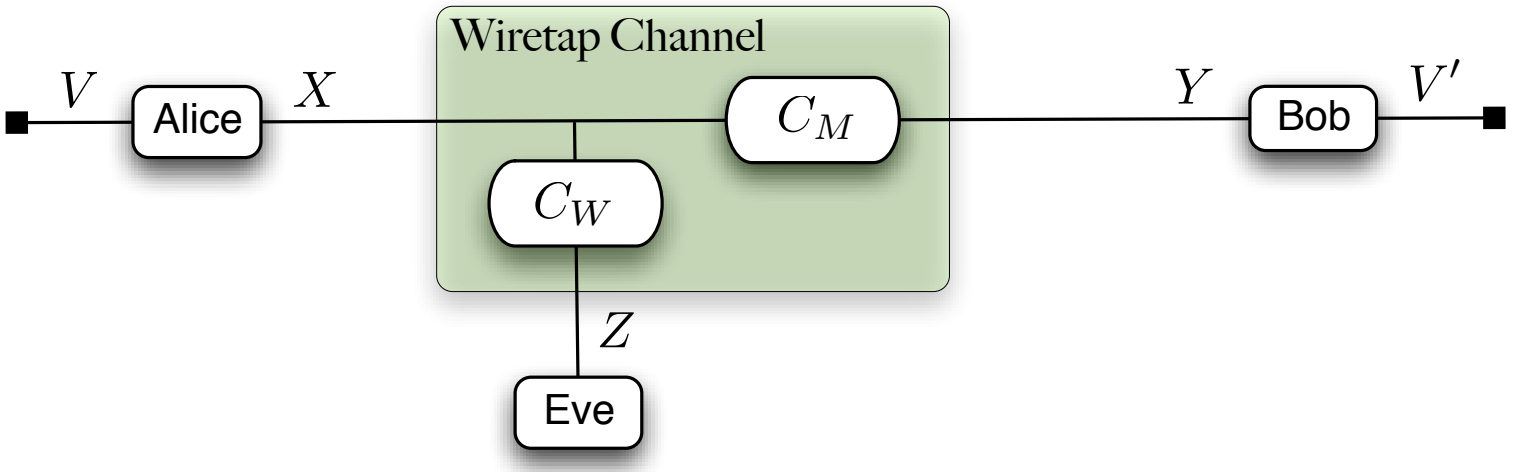
- Not based on keys, but security capacity: $C_s = \max_{p(x)} [I(X;Y) - I(X;Z)]$
- Goals are: Reliability: $P\{V \neq V'\} \to 0$
  Security: $I(V;Z)/n \to 0$, as $n \to \infty$

When the main channel is less noisy than the wiretapper's channel, and $I(X;Y)$, $I(X;Z)$ maximized by the same $p(x)$, then (van Dijk 1997):

$$C_s = \text{Capacity}\,(C_M) - \text{Capacity}\,(C_W)$$

Information Theoretical Security

Based on the Wiretap Channel (Wyner 1975)



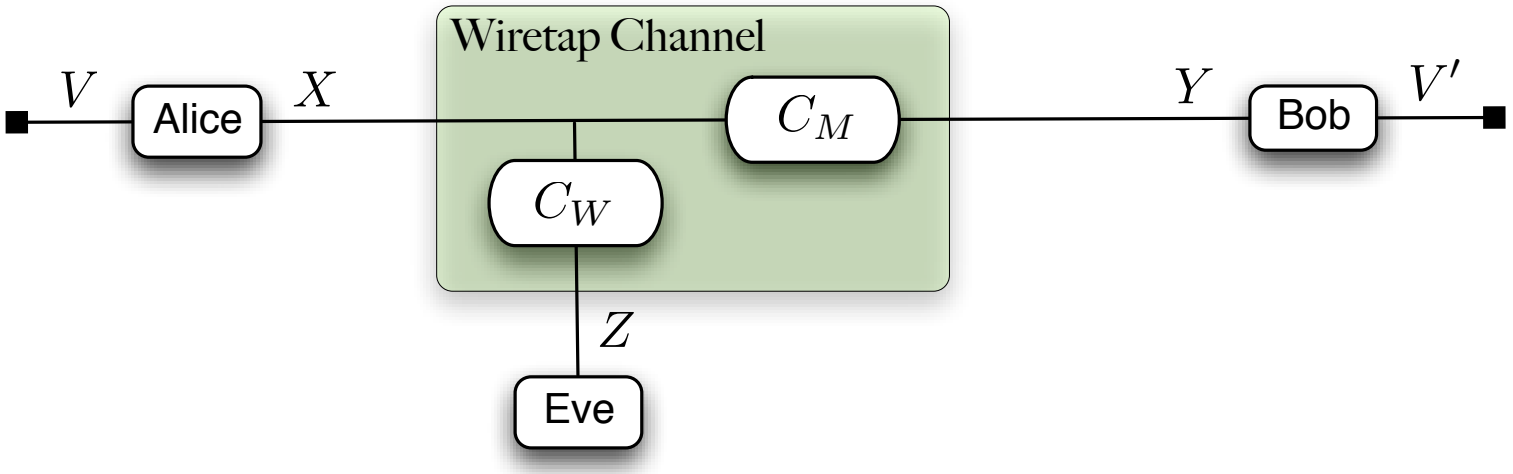
Very Different From Cryptographic Model

- Not based on keys, but security capacity: $C_s = \max_{p(x)} [I(X;Y) - I(X;Z)]$
- Goals are: Reliability: $P\{V \neq V'\} \to 0$
  Security: $I(V;Z)/n \to 0$, as $n \to \infty$

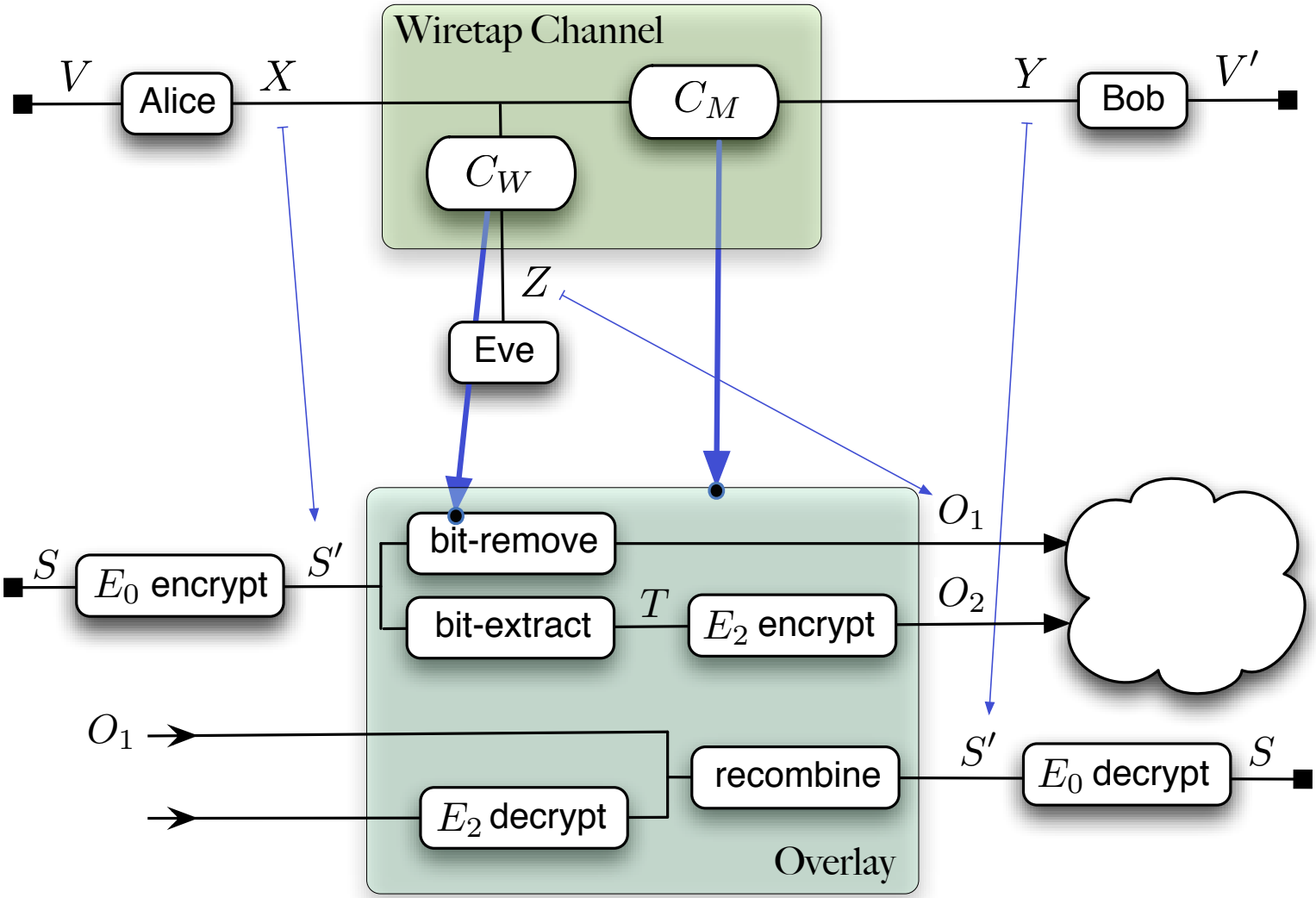
When the main channel is less noisy than the wiretapper's channel, and $I(X;Y)$, $I(X;Z)$ maximized by the same $p(x)$, then (van Dijk 1997):

$$C_s = \text{Capacity}\ (C_M) - \text{Capacity}\ (C_W)$$

USING THE WIRETAP CHANNEL IN THE MEO



USING THE WIRETAP CHANNEL IN THE MEO



$$C_s = \text{Capacity}\,(C_M) - \text{Capacity}\,(C_W) = 1 - C_{\text{BRC}}$$

Capacity of the Bit Removal Channel (BRC)

A deletion channel is one where symbols are deleted with probability d.
Example: 1100100 $\rightarrow$ 11010, note erasure gives 11?01?0 !!

Simpler 'blocked' version (Diggavi et al. 2006) is deletion-like within a block width n, memoryless across blocks. Proportion θ of bits removed is random.

Upper Capacity Bound for BRC

- Diggavi proof can be adapted with $\theta = b/8p_0$ constant
- Yields: $C_{\text{BRC}} \geq 1 - H(\theta),\ \theta \geq 0.5$
- Use looser general deletion channel result for $\theta < 0.5$, obtain ($A = 0.1185$)

$$C_{\text{BRC}} \geq \begin{cases} 1 - H(\theta), & \theta \geq 0.5 \\ A\theta, & \theta < 0.5. \end{cases}$$

Lower Bound for BRC

- Use argument based on side information
- Yields: $C_{\text{BRC}} \leq \theta$

Capacity of the Bit Removal Channel (BRC)

A deletion channel is one where symbols are deleted with probability d.
Example: 1100100 $\rightarrow$ 11010, note erasure gives 11?01?0 !!

Simpler 'blocked' version (Diggavi et al. 2006) is deletion-like within a block width n, memoryless across blocks. Proportion θ of bits removed is random.

Upper Capacity Bound for BRC

- Diggavi proof can be adapted with $\theta = b/8p_0$ constant
- Yields: $C_{\text{BRC}} \geq 1 - H(\theta),\ \theta \geq 0.5$
- Use looser general deletion channel result for $\theta < 0.5$, obtain ($A = 0.1185$)

$$C_{\text{BRC}} \geq \begin{cases} 1 - H(\theta), & \theta \geq 0.5 \\ A\theta, & \theta < 0.5. \end{cases}$$

Lower Bound for BRC

- Use argument based on side information
- Yields: $C_{\text{BRC}} \leq \theta$

Capacity of the Bit Removal Channel (BRC)

A deletion channel is one where symbols are deleted with probability d.
Example: 1100100 $\rightarrow$ 11010, note erasure gives 11?01?0 !!

Simpler 'blocked' version (Diggavi et al. 2006) is deletion-like within a block width n, memoryless across blocks. Proportion θ of bits removed is random.

Upper Capacity Bound for BRC

- Diggavi proof can be adapted with $\theta = b/8p_0$ constant
- Yields: $C_{\text{BRC}} \geq 1 - H(\theta),\ \theta \geq 0.5$
- Use looser general deletion channel result for $\theta < 0.5$, obtain ($A = 0.1185$)

$$C_{\text{BRC}} \geq \begin{cases} 1 - H(\theta), & \theta \geq 0.5 \\ A\theta, & \theta < 0.5. \end{cases}$$

Lower Bound for BRC

- Use argument based on side information
- Yields: $C_{\text{BRC}} \leq \theta$

Capacity of the Bit Removal Channel (BRC)

A deletion channel is one where symbols are deleted with probability d.
Example: 1100100 $\rightarrow$ 11010, note erasure gives 11?01?0 !!

Simpler 'blocked' version (Diggavi et al. 2006) is deletion-like within a block width n, memoryless across blocks. Proportion θ of bits removed is random.

Upper Capacity Bound for BRC

- Diggavi proof can be adapted with $\theta = b/8p_0$ constant
- Yields: $C_{\text{BRC}} \geq 1 - H(\theta),\ \theta \geq 0.5$
- Use looser general deletion channel result for $\theta < 0.5$, obtain ($A = 0.1185$)

$$C_{\text{BRC}} \geq \begin{cases} 1 - H(\theta), & \theta \geq 0.5 \\ A\theta, & \theta < 0.5. \end{cases}$$

Lower Bound for BRC

- Use argument based on side information
- Yields: $C_{\text{BRC}} \leq \theta$

Locating the Missing Pieces

$$\begin{aligned} \Pr(\text{channel 1 is sniffed}) &= q_1 \\ \Pr(\text{channel 2 is sniffed}) &= q_2 \\ \Pr(E_2 \text{ cracked} \mid \text{channel 2 sniffed}) &= q_{E_2} \end{aligned}$$

Cracking scenarios:

(1) Channel sniffed? 1-YES, 2-NO.
Probability $p_1 = q_1(1-q_2)$

(2) Channel sniffed? 1-YES, 2-YES, E_2 cracked.
Probability $p_2 = q_1 q_2 q_{E_2}$

(3) Channel sniffed? 1-YES, 2-YES, E_2 not cracked.
Probability $p_3 = q_1 q_2 (1-q_{E_2})$

(4) Channel sniffed? 1-NO, 2-YES, E_2 cracked.
Probability $p_4 = (1-q_1)(q_2)q_{E_2}$

(5) Channel sniffed? 1-NO, 2-YES, E_2 not cracked.
Probability $p_5 = (1-q_1)q_2(1-q_{E_2})$

(6) Channel sniffed? 1-NO, 2-NO
Probability $p_6 = (1-q_1)(1-q_2)$.

PUTTING IT TOGETHER

Expected Secrecy Capacity is

$$C_s = q_1(1 - q_2 q_{E_2})C_{s_1} + q_2 q_{E_2}(1 - q_1)C_{s_2} + (1 - q_1)(1 + q_2 q_{E_2})$$

where the C_{s_i} are given by

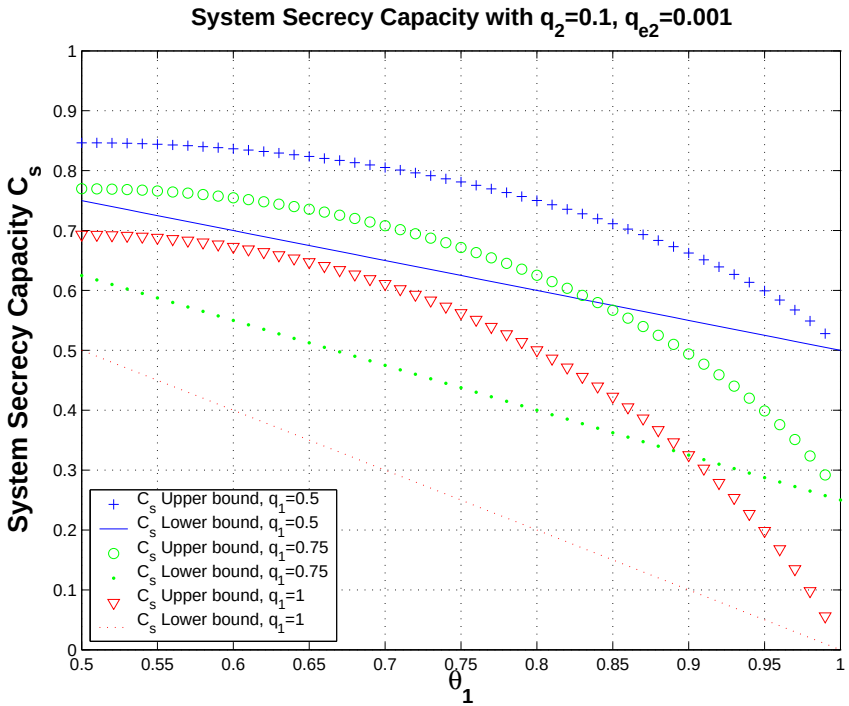
$$C_{s_i} = 1 - C_{\mathrm{BRC}}(\theta_i).$$

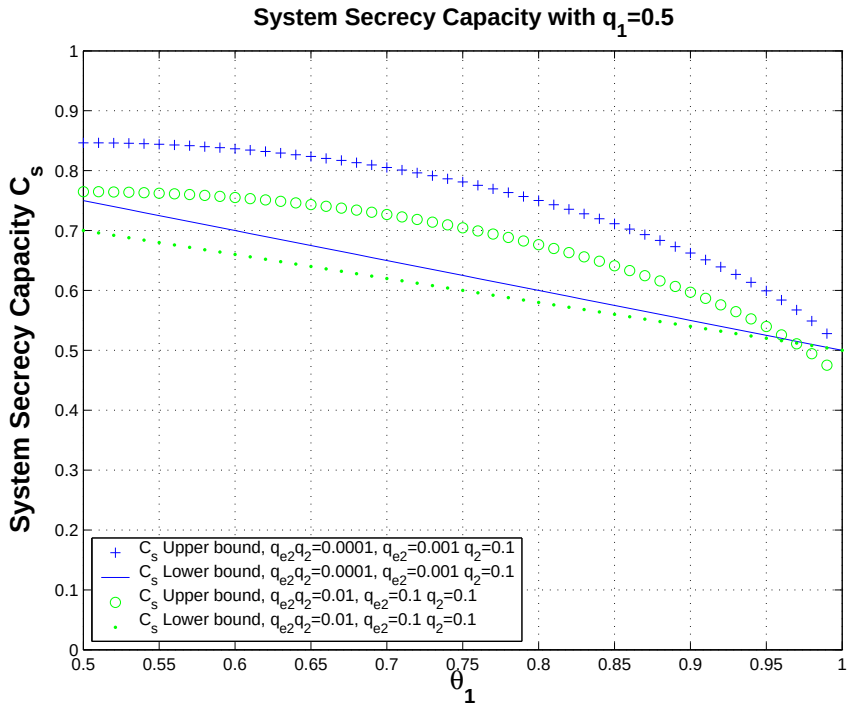
Assume that $\theta_1 > 0.5$, then:

$$\begin{aligned} C_s &\geq q_1(1 - q_2')(1 - \theta_1) + q_2'(1 - q_1)(1 - \theta_2) + (1 - q_1)(1 + q_2') \\ C_s &\leq q_1(1 - q_2')H_0(\theta_1) + q_2'(1 - q_1)A\theta_2 + (1 - q_1)(1 + q_2'), \end{aligned}$$

where $q_2' = q_2 q_{E_2}$ is Pr(adversary has access to BRC on channel 2).

How Much Secrecy?





- Even WiFi hotspots (think $q_1 = 1, \theta_1 = 0.9$) have usable secrecy
- Secrecy capacity greatest when $\theta_1 = 0.5$
- Capacity insensitive to q_{E_2}

CONCLUSIONS

- MEO is a novel scheme to split traffic over multiple channels whilst increasing data confidentiality over and above physical difficulty.
- The MEO can be applied as a modular overlay to existing schemes, and is lightweight (in principle).
- It enormously increases cracking times (under fairly strong assumptions).
- Through the Wiretap Channel, we can drop those assumptions and show a non-zero (expected) security capacity.