The Collision Security of Tandem-DM in the Ideal Cipher Model

Jooyoung Lee[1] Martijn Stam[2] John Steinberger[3]

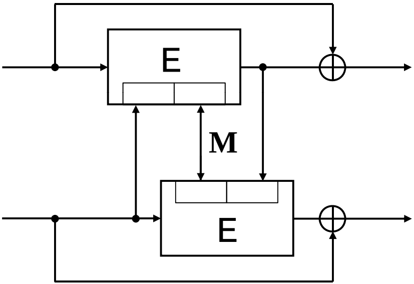
[1]Faculty of Mathematics and Statistics, Sejong University, Seoul, Korea

[2]Department of Computer Science, University of Bristol, Bristol, United Kingdom

[3]Institute of Theoretical Computer Science, Tsinghua University, Beijing, China

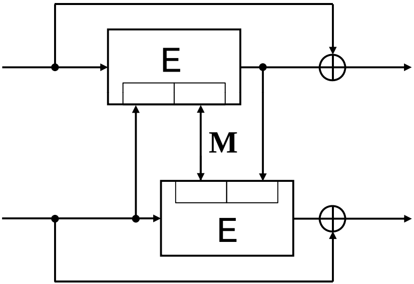
August 18, 2011

Tandem-DM



- A 3*n*-bit to 2*n*-bit compression function making two calls to a blockcipher using 2*n*-bit keys
- Proposed by Lai and Massey in Eurocrypt 1992
- The first security proof given in FSE 2009, its extension given in ProvSec 2010

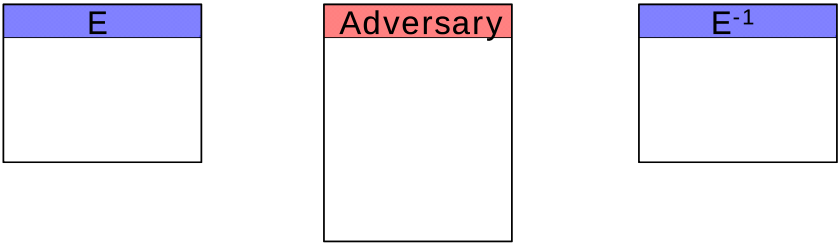
Tandem-DM



Contribution

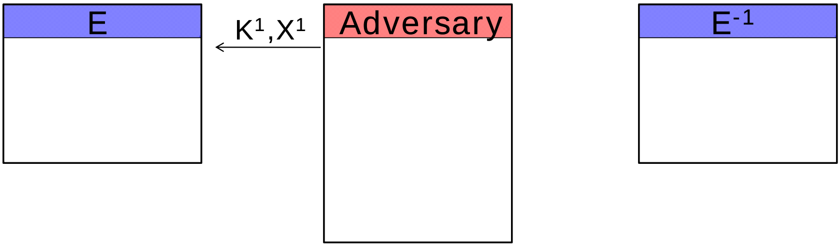
- Shows the prior proofs are flawed
- Presents a novel proof for the collision resistance of Tandem-DM in the ideal cipher model
- Mostly historical interest, rather than practical interest

Ideal Cipher Model & Query History



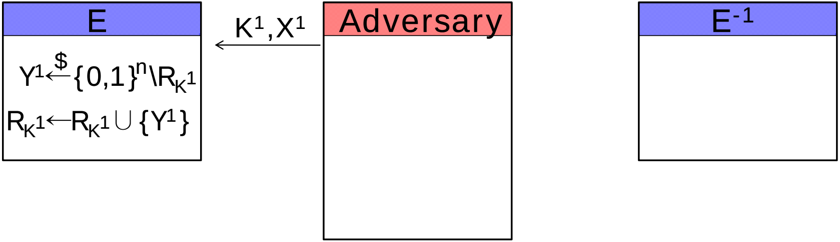
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



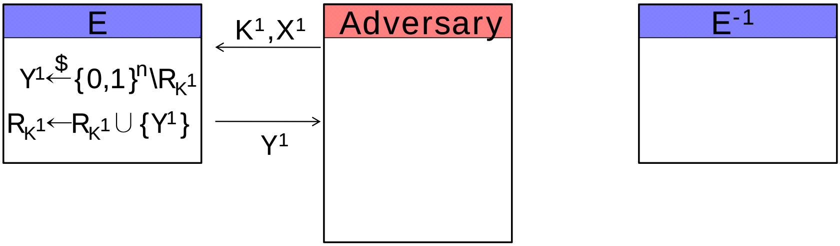
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



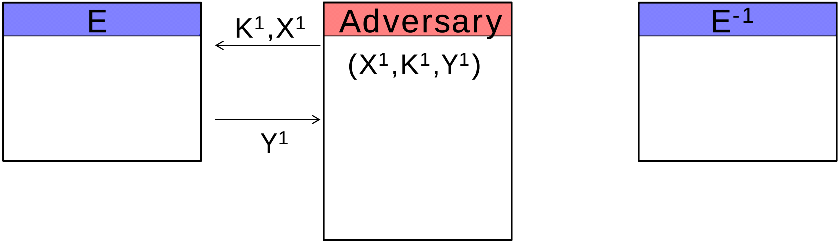
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



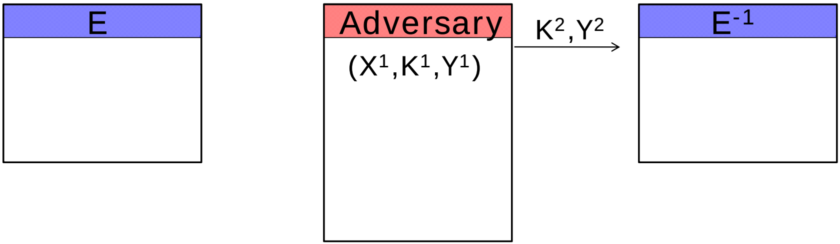
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



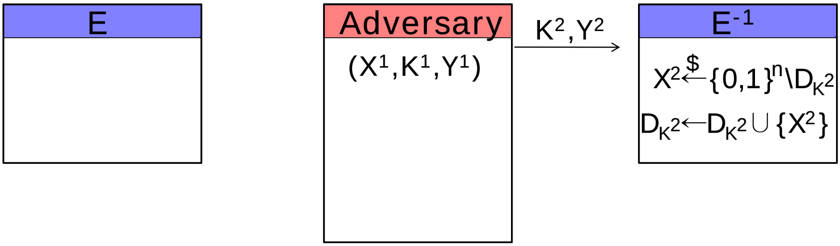
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



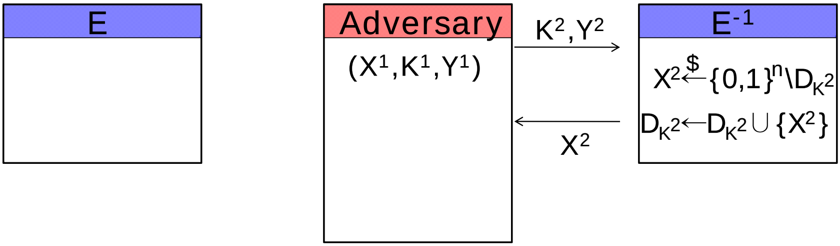
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



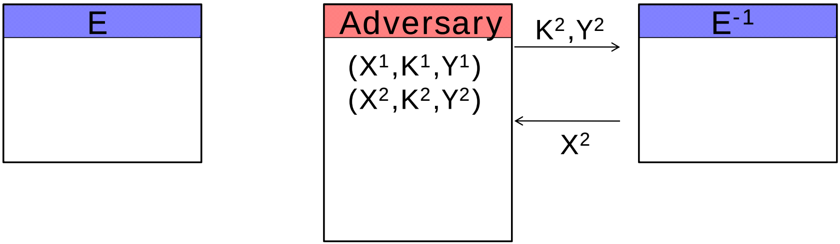
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



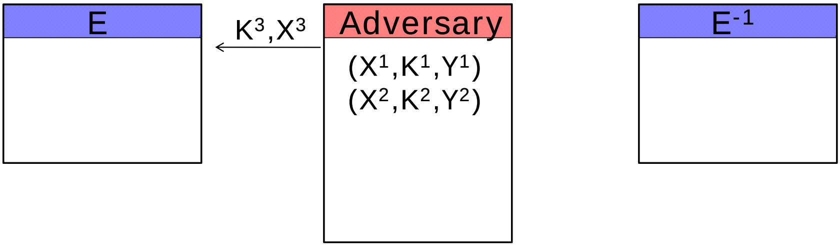
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



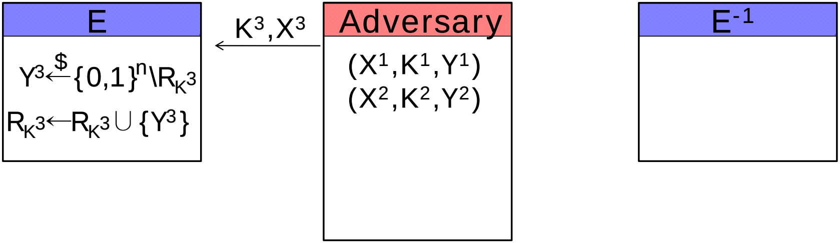
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



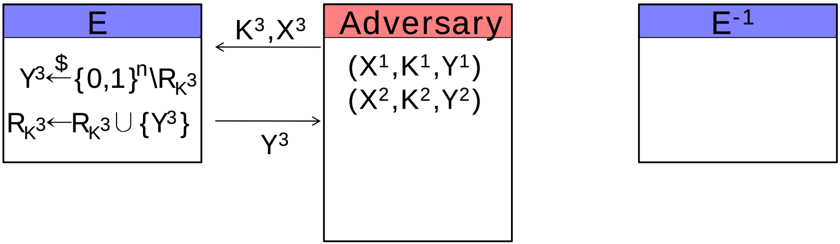
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



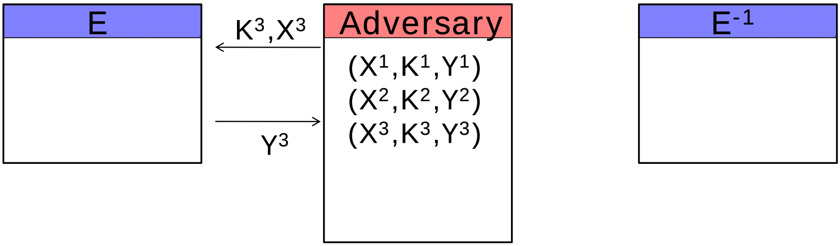
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



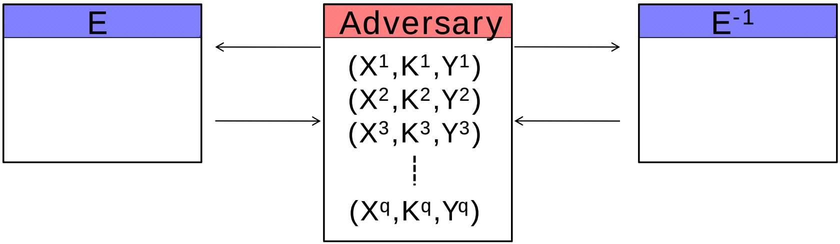
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



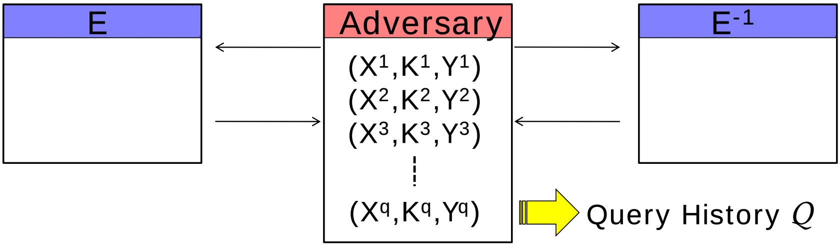
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



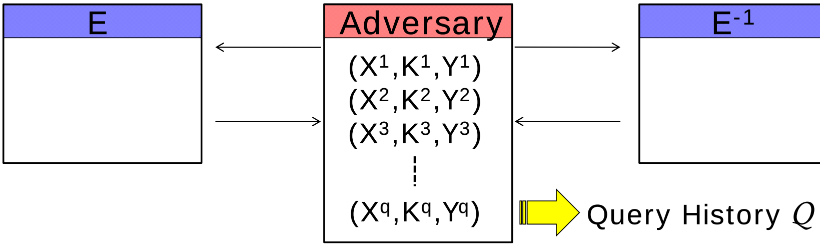
- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History



- An ideal cipher is simulated by lazy sampling
- The query history Q determines every evaluation of a blockcipher-based compression function

Ideal Cipher Model & Query History

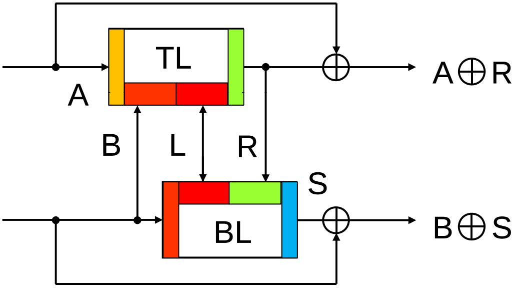


- An ideal cipher is simulated by lazy sampling
- The query history $\mathcal{Q}$ determines every evaluation of a blockcipher-based compression function

Evaluation of Tandem-DM

$(A, B||L, R), (B, L||R, S) \in \mathcal{Q}$ determine

$$\begin{array}{rcl} TDM^E : \{0,1\}^{3n} & \longrightarrow & \{0,1\}^{2n} \\ A||B||L & \longmapsto & A \oplus R||B \oplus S \end{array}$$

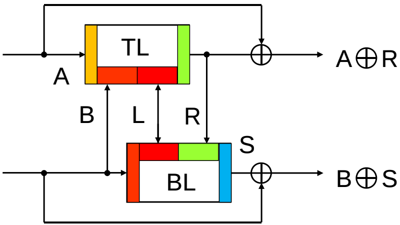


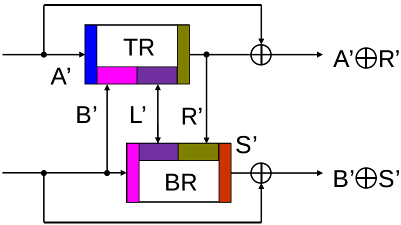
Collisions in Tandem-DM

The goal of a collision-finding adversary $\mathcal{A}$

To find $(A, B||L, R)$, $(B, L||R, S)$, $(A', B'||L', R')$, $(B', L'||R', S')$ such that $A||B||L \neq A'||B'||L'$, $A \oplus R = A' \oplus R'$, $B \oplus S = B' \oplus S'$

Predicate Coll($\mathcal{Q}$) is true if and only if such queries exist in $\mathcal{Q}$



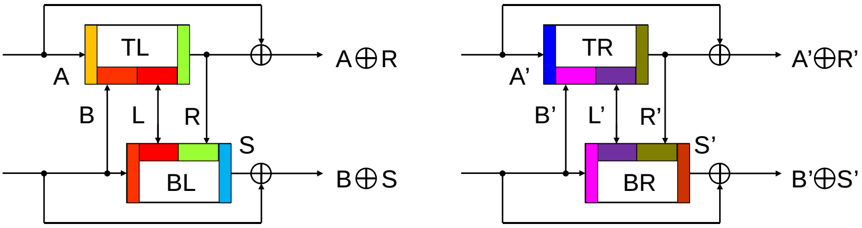


Collisions in Tandem-DM

The goal of a collision-finding adversary $\mathcal{A}$

To find $(A, B||L, R)$, $(B, L||R, S)$, $(A', B'||L', R')$, $(B', L'||R', S')$ such that $A||B||L \neq A'||B'||L'$, $A \oplus R = A' \oplus R'$, $B \oplus S = B' \oplus S'$

We want to upper bound $\mathbf{Pr}[\text{Coll}(\mathcal{Q})] = \mathbf{Adv}^{\text{Coll}}_{TDM^E}(\mathcal{A})$

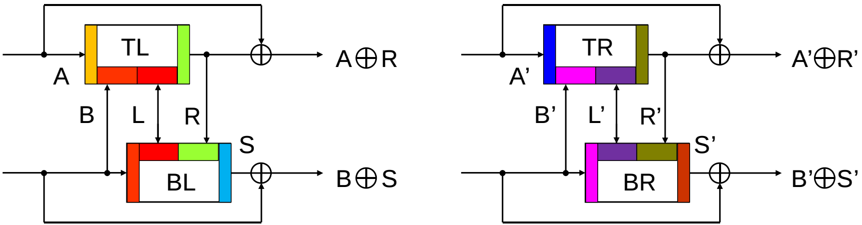


Collisions in Tandem-DM

The goal of a collision-finding adversary $\mathcal{A}$

To find $(A, B||L, R)$, $(B, L||R, S)$, $(A', B'||L', R')$, $(B', L'||R', S')$ such that $A||B||L \neq A'||B'||L'$, $A \oplus R = A' \oplus R'$, $B \oplus S = B' \oplus S'$

We want $\mathbf{Pr}[\mathrm{Coll}(\mathcal{Q})]$ to be small

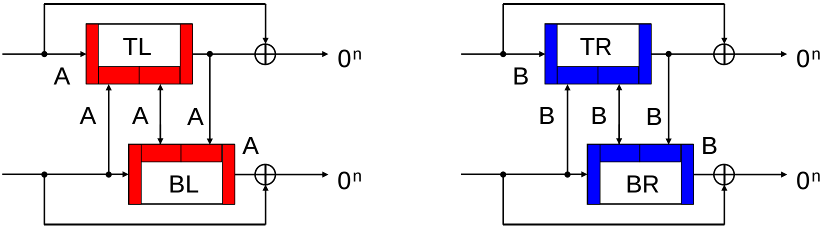


Case Analysis

$\mathrm{Coll}(\mathcal{Q}) \Rightarrow \mathrm{Coll}_1(\mathcal{Q}) \vee \mathrm{Coll}_2(\mathcal{Q}) \vee \mathrm{Coll}_3(\mathcal{Q})$, where

- $\mathrm{Coll}_1(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL, BL, TR, BR distinct
- $\mathrm{Coll}_2(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL = BL or TR = BR
- $\mathrm{Coll}_3(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL = BR or BL = TR

Ex) $\mathrm{Coll}_2(\mathcal{Q})$ occurs if $(A, A||A, A)$, $(B, B||B, B)$ s.t. $A \neq B$ exist



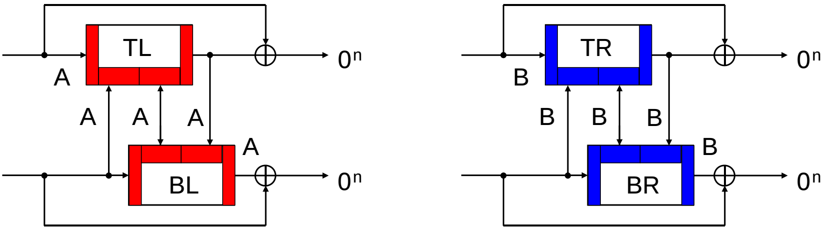
Case Analysis

$\mathsf{Coll}(\mathcal{Q}) \Rightarrow \mathsf{Coll}_1(\mathcal{Q}) \vee \mathsf{Coll}_2(\mathcal{Q}) \vee \mathsf{Coll}_3(\mathcal{Q})$, where

- $\mathsf{Coll}_1(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL, BL, TR, BR distinct
- $\mathsf{Coll}_2(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL = BL or TR = BR
- $\mathsf{Coll}_3(\mathcal{Q}) \Leftrightarrow \mathcal{Q}$ has a collision with TL = BR or BL = TR

We are going to focus on upper bounding $\mathbf{Pr}[\mathsf{Coll}_1(\mathcal{Q})]$

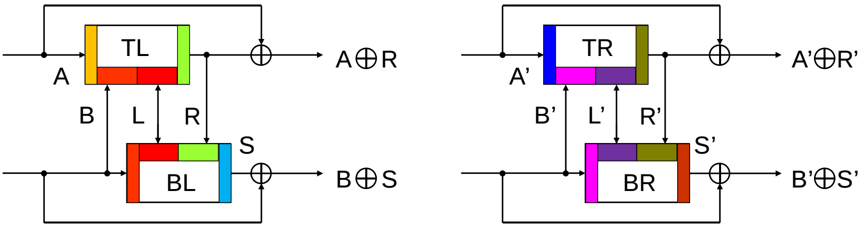
Ex) $\mathsf{Coll}_2(\mathcal{Q})$ occurs if $(A, A||A, A)$, $(B, B||B, B)$ s.t. $A \neq B$ exist



Upper bounding $\mathbf{Pr}[\mathrm{Coll}_1(\mathcal{Q})]$

General Framework

1. Upper bound the probability of $\mathrm{Coll}_1^i(\mathcal{Q})$ that the i-th query completes a collision
2. Union bound by summing the upper bounds over all possible queries $i = 1, \ldots, q$ (If the upper bounds are independent of each query, then we can just multiply q)

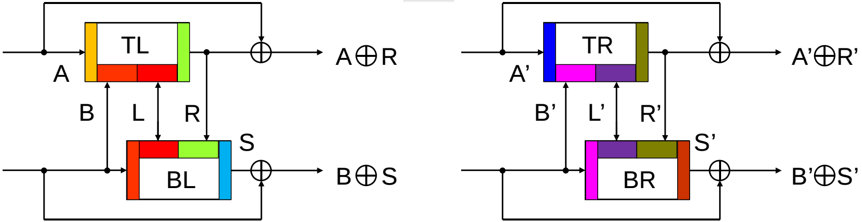


Upper bounding $\mathbf{Pr}[\mathsf{Coll}_1(\mathcal{Q})]$

General Framework

1. Upper bound the probability of $\mathsf{Coll}_1^i(\mathcal{Q})$ that the i-th query completes a collision
2. Union bound by summing the upper bounds over all possible queries $i = 1, \ldots, q$ (If the upper bounds are independent of each query, then we can just multiply q)

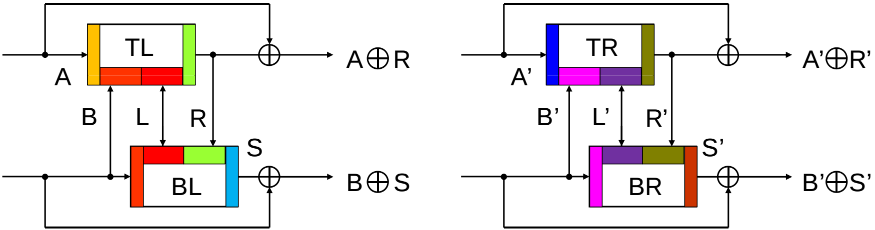
How can we upper bound $\mathbf{Pr}[\mathsf{Coll}_1^i(\mathcal{Q})]$?



Upper bounding **Pr**[$\mathrm{Coll}_1^i(\mathcal{Q})$]

By symmetry, we can assume the last query is either TL or BL.

The last query:	TL	BL
Backward	Case 1	Case 3
Forward	Case 2	Case 4



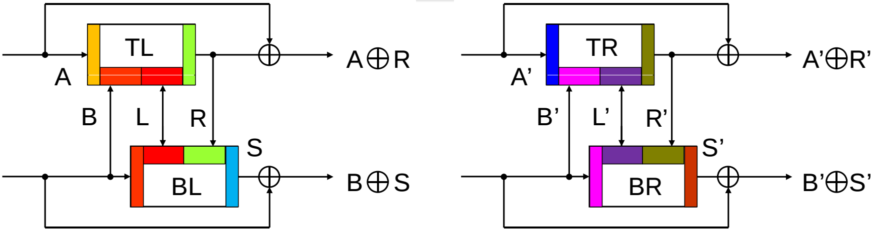
Upper bounding $\mathbf{Pr}[\mathrm{Coll}_1^i(\mathcal{Q})]$

By symmetry, we can assume the last query is either TL or BL.

The last query:	TL	BL
Backward	Case 1	Case 3
Forward	Case 2	Case 4

Union bound

$$\mathbf{Pr}[\mathrm{Coll}_1^i(\mathcal{Q})] \leq \mathbf{Pr}[\mathrm{Case1}] + \mathbf{Pr}[\mathrm{Case2}] + \mathbf{Pr}[\mathrm{Case3}] + \mathbf{Pr}[\mathrm{Case4}]$$



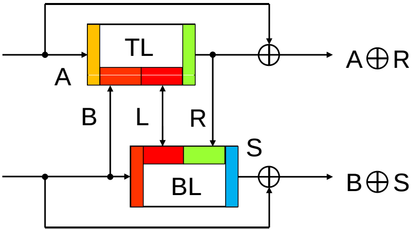
Upper bounding $\mathbf{Pr}[\mathrm{Coll}_1^i(\mathcal{Q})]$

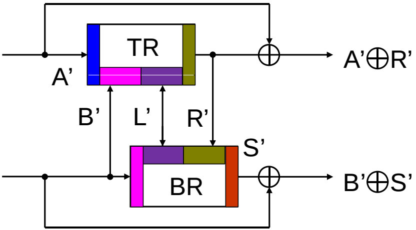
By symmetry, we can assume the last query is either TL or BL.

The last query:	TL	BL
Backward	Case 1	Case 3
Forward	Case 2	Case 4

Union bound

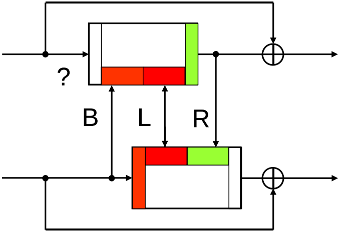
$$\mathbf{Pr}[\mathrm{Coll}_1^i(\mathcal{Q})] \leq \mathbf{Pr}[\mathrm{Case1}] + \mathbf{Pr}[\mathrm{Case2}] + \mathbf{Pr}[\mathrm{Case3}] + \mathbf{Pr}[\mathrm{Case4}]$$

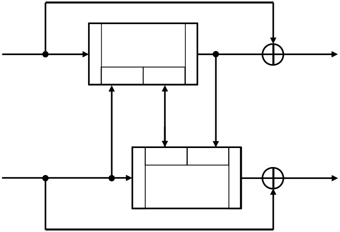




Case 1: The Last Query is TL and Backward

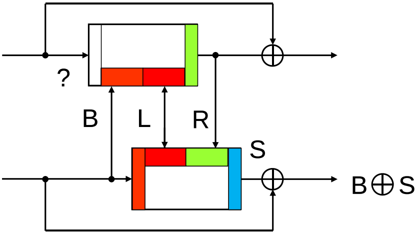
1. At the point when TL is queried, B, L, R are fixed
2. B, L, R uniquely determine BL, and $B \oplus S$
3. The number of BR-queries $(B', L' || R', S')$ such that $B' \oplus S' = B \oplus S$ is at most α except with small probability
4. Each of BR-queries uniquely determines TR, and $A' \oplus R'$
5. The response should be $A' \oplus R' \oplus R$, so **Pr**[Case1] $\leq \frac{\alpha}{2^n - q}$ (except with the "bad event")

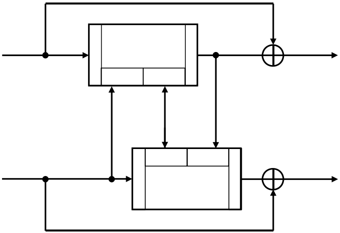




Case 1: The Last Query is TL and Backward

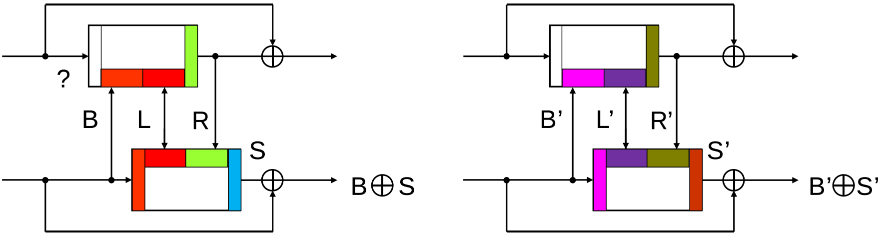
1. At the point when TL is queried, B, L, R are fixed
2. B, L, R uniquely determine BL, and $B \oplus S$
3. The number of BR-queries $(B', L'||R', S')$ such that $B' \oplus S' = B \oplus S$ is at most α except with small probability
4. Each of BR-queries uniquely determines TR, and $A' \oplus R'$
5. The response should be $A' \oplus R' \oplus R$, so **Pr**[Case1] $\leq \frac{\alpha}{2^n - q}$ (except with the “bad event")





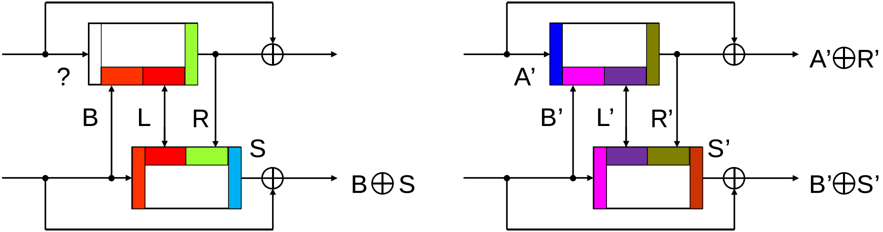
Case 1: The Last Query is TL and Backward

1. At the point when TL is queried, B, L, R are fixed
2. B, L, R uniquely determine BL, and $B \oplus S$
3. The number of BR-queries $(B', L' || R', S')$ such that $B' \oplus S' = B \oplus S$ is at most α except with small probability
4. Each of BR-queries uniquely determines TR, and $A' \oplus R'$
5. The response should be $A' \oplus R' \oplus R$, so $\mathbf{Pr}[\text{Case1}] \leq \frac{\alpha}{2^n - q}$ (except with the "bad event")



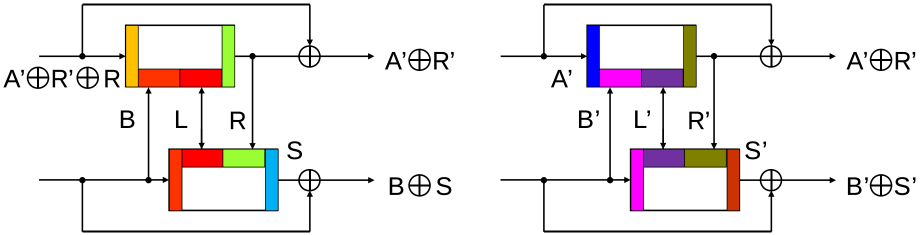
Case 1: The Last Query is TL and Backward

1. At the point when TL is queried, B, L, R are fixed
2. B, L, R uniquely determine BL, and $B \oplus S$
3. The number of BR-queries $(B', L' || R', S')$ such that $B' \oplus S' = B \oplus S$ is at most α except with small probability
4. Each of BR-queries uniquely determines TR, and $A' \oplus R'$
5. The response should be $A' \oplus R' \oplus R$, so **Pr**[Case1] $\leq \frac{\alpha}{2^n - q}$ (except with the "bad event")



Case 1: The Last Query is TL and Backward

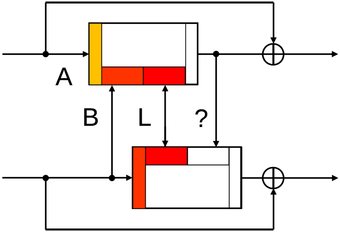
1. At the point when TL is queried, B, L, R are fixed
2. B, L, R uniquely determine BL, and $B \oplus S$
3. The number of BR-queries $(B', L' || R', S')$ such that $B' \oplus S' = B \oplus S$ is at most α except with small probability
4. Each of BR-queries uniquely determines TR, and $A' \oplus R'$
5. The response should be $A' \oplus R' \oplus R$, so **Pr**[Case1] $\leq \frac{\alpha}{2^n - q}$ (except with the "bad event")

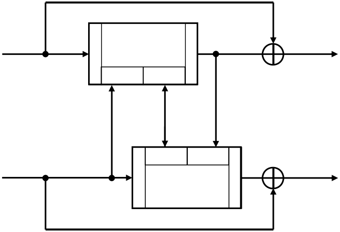


Case 2: The Last Query is TL and Forward

Subcase 2a: BL-query is Backward

1. At the point when TL is queried, A, B, L are fixed
2. The number of backward queries whose answer is B is at most α except with small probability
3. Since each of such backward queries uniquely determines R, $\mathbf{Pr}[\text{Subcase2}a] \leq \frac{\alpha}{2^n - q}$ (except with the "bad event")

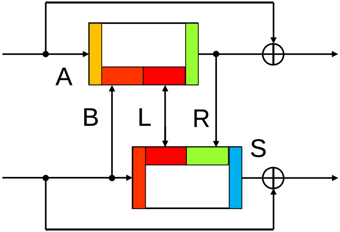


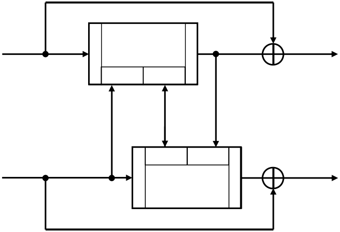


Case 2: The Last Query is TL and Forward

Subcase 2a: BL-query is Backward

1. At the point when TL is queried, A, B, L are fixed
2. The number of backward queries whose answer is B is at most α except with small probability
3. Since each of such backward queries uniquely determines R, $\mathbf{Pr}[\text{Subcase2}a] \leq \frac{\alpha}{2^n - q}$ (except with the “bad event")

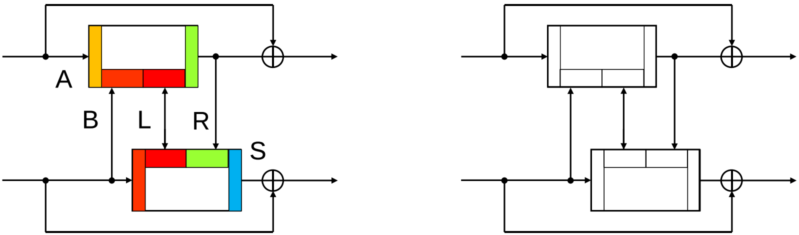




Case 2: The Last Query is TL and Forward

Subcase 2a: BL-query is Backward

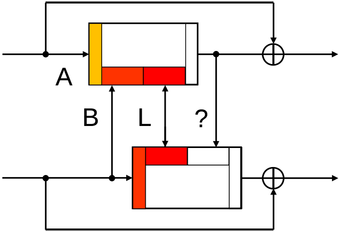
1. At the point when TL is queried, A, B, L are fixed
2. The number of backward queries whose answer is B is at most α except with small probability
3. Since each of such backward queries uniquely determines R, **Pr**[Subcase2a] $\leq \frac{\alpha}{2^n - q}$ (except with the "bad event")

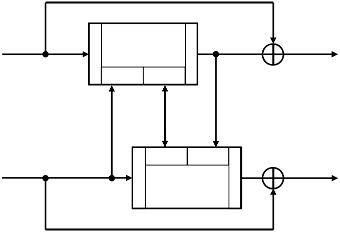


Case 2: The Last Query is TL and Forward

Subcase 2b: BL-query is Forward

1. At the point when TL is queried, A, B, L are fixed
2. The number of forward queries whose input block is B?

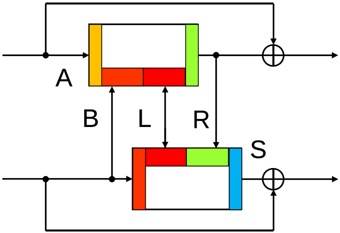


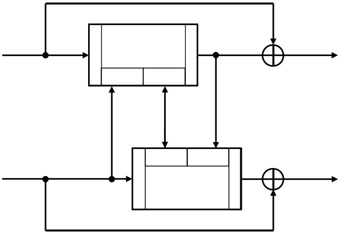


Case 2: The Last Query is TL and Forward

Subcase 2b: BL-query is Forward

1. At the point when TL is queried, A, B, L are fixed
2. The number of forward queries whose input block is B?



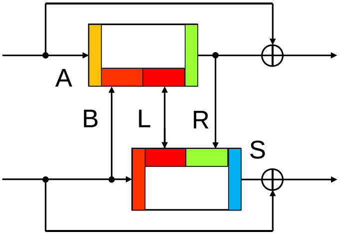


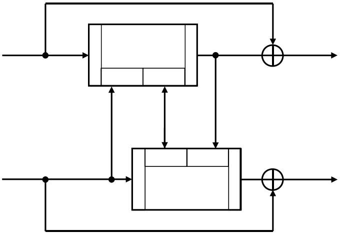
Case 2: The Last Query is TL and Forward

Subcase 2b: BL-query is Forward

1. At the point when TL is queried, A, B, L are fixed
2. The number of forward queries whose input block is B?

It is hard to probabilistically restrict this number!



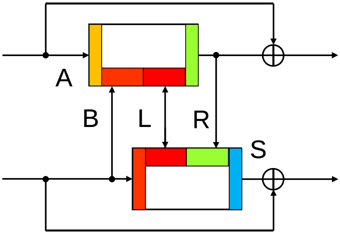


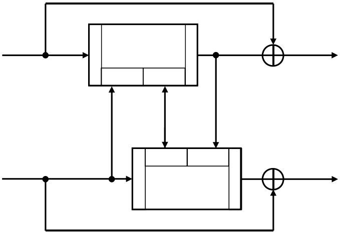
Case 2: The Last Query is TL and Forward

Subcase 2b: BL-query is Forward

1. At the point when TL is queried, *A*, *B*, *L* are fixed
2. The number of forward queries whose input block is *B*?

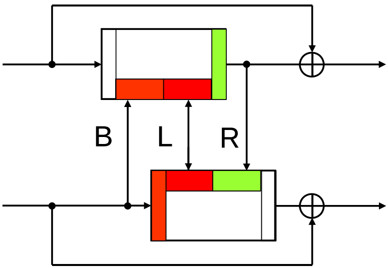
We want to eliminate this case





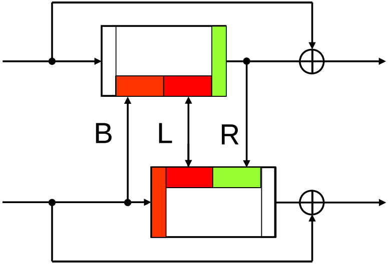
Main Idea: Modified Adversary $\mathcal{A}'$

- $\mathcal{A}'$ runs $\mathcal{A}$ as a subroutine and records its query history $\mathcal{Q}'$
- If $\mathcal{A}$ makes a forward query $E_{L||R}(B)$, then $\mathcal{A}'$ makes a query $E_{L||R}(B)$, and an additional query $E^{-1}_{B||L}(R)$
- If $\mathcal{A}$ makes a backward query $E^{-1}_{B||L}(R)$, then $\mathcal{A}'$ makes a query $E^{-1}_{B||L}(R)$, and an additional query $E_{L||R}(B)$



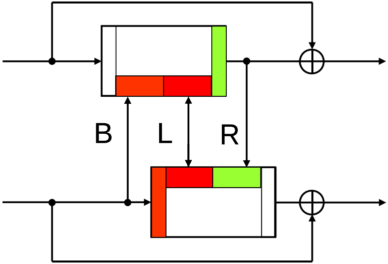
Main Idea: Modified Adversary $\mathcal{A}'$

- $\mathcal{A}'$ runs $\mathcal{A}$ as a subroutine and records its query history $\mathcal{Q}'$
- If $\mathcal{A}$ makes a forward query $E_{L||R}(B)$, then $\mathcal{A}'$ makes a query $E_{L||R}(B)$, and an additional query $E^{-1}_{B||L}(R)$
- If $\mathcal{A}$ makes a backward query $E^{-1}_{B||L}(R)$, then $\mathcal{A}'$ makes a query $E^{-1}_{B||L}(R)$, and an additional query $E_{L||R}(B)$



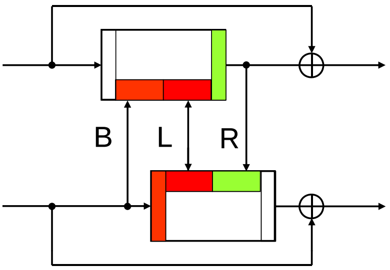
Main Idea: Modified Adversary $\mathcal{A}'$

- $\mathcal{A}'$ runs $\mathcal{A}$ as a subroutine and records its query history $\mathcal{Q}'$
- If $\mathcal{A}$ makes a forward query $E_{L||R}(B)$, then $\mathcal{A}'$ makes a query $E_{L||R}(B)$, and an additional query $E^{-1}_{B||L}(R)$
- If $\mathcal{A}$ makes a backward query $E^{-1}_{B||L}(R)$, then $\mathcal{A}'$ makes a query $E^{-1}_{B||L}(R)$, and an additional query $E_{L||R}(B)$



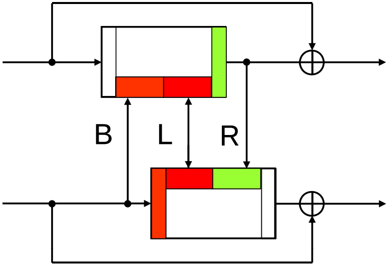
The Property of the Modified Adversary

- If $\mathcal{A}$ makes q queries, then $\mathcal{A}'$ makes at most $2q$ queries
- Since $\mathcal{Q} \subset \mathcal{Q}'$, $\mathbf{Adv}^{\mathrm{Coll}}_{TDM^E}(\mathcal{A}) \leq \mathbf{Adv}^{\mathrm{Coll}}_{TDM^E}(\mathcal{A}')$



The Property of the Modified Adversary

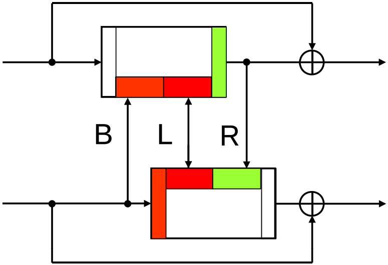
- If $\mathcal{A}$ makes q queries, then $\mathcal{A}'$ makes at most $2q$ queries
- Since $\mathcal{Q} \subset \mathcal{Q}'$, $\mathbf{Adv}^{\text{Coll}}_{TDM^E}(\mathcal{A}) \leq \mathbf{Adv}^{\text{Coll}}_{TDM^E}(\mathcal{A}')$



The Property of the Modified Adversary

- If $\mathcal{A}$ makes q queries, then $\mathcal{A}'$ makes at most $2q$ queries
- Since $\mathcal{Q} \subset \mathcal{Q}'$, $\mathbf{Adv}^{\mathsf{Coll}}_{TDM^E}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{Coll}}_{TDM^E}(\mathcal{A}')$

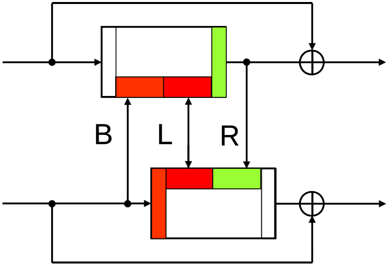
If $\mathcal{A}'$ obtains the BL position of a certain evaluation by a forward query, then $\mathcal{A}'$ will immediately make an additional backward query and place it at the TL position



The Property of the Modified Adversary

- If $\mathcal{A}$ makes q queries, then $\mathcal{A}'$ makes at most $2q$ queries
- Since $\mathcal{Q} \subset \mathcal{Q}'$, $\mathbf{Adv}^{\mathrm{Coll}}_{TDM^E}(\mathcal{A}) \leq \mathbf{Adv}^{\mathrm{Coll}}_{TDM^E}(\mathcal{A}')$

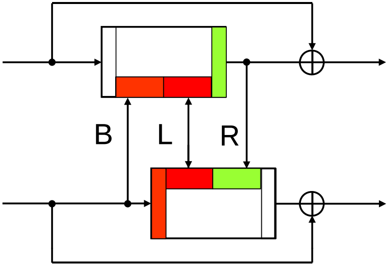
If the TL position of a certain evaluation is obtained by a forward query after the BL position is determined, then the BL query should have been obtained by a backward query



The Property of the Modified Adversary

- If $\mathcal{A}$ makes q queries, then $\mathcal{A}'$ makes at most $2q$ queries
- Since $\mathcal{Q} \subset \mathcal{Q}'$, $\mathbf{Adv}^{\mathsf{Coll}}_{TDM^E}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{Coll}}_{TDM^E}(\mathcal{A}')$

It means that $\mathcal{A}'$ does not create Subcase 2b



Main Result

Theorem

For $N = 2^n$, $q < N/2$ and $1 \leq \alpha \leq 2q$,

$$\mathbf{Adv}^{\mathrm{coll}}_{TDM}(q) \leq 2N \left(\frac{2eq}{\alpha(N-2q)} \right)^{\alpha} + \frac{4q\alpha}{N-2q} + \frac{4q}{N-2q}$$

Asymptotically, using $\alpha = n/\log n$

$$\lim_{n \to \infty} \mathbf{Adv}^{\mathrm{coll}}_{TDM}(N/n) = 0$$

Numerically, for $n = 128$, using $\alpha = 16$

$$\mathbf{Adv}^{\mathrm{coll}}_{TDM}(2^{120.87}) < \frac{1}{2}$$

Thank You