Characterizing Collision and Second-Preimage Resistance in Linicrypt

Ian McQuoid Trevor Swope Mike Rosulek



An Introduction

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1} + v_{2}$
 $return(v_{8}, v_{5})$
 $\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$

Which program is collision resistant?

Linicrypt programs are a class of	of
algorithms	

$$v_4 := H(v_1)$$
$$v_5 := H(v_3)$$

 $\mathcal{P}^{H}(v_{1},v_{2},v_{3}):$

Introduced by Carmer and Rosulek,

$$v_6 := v_4 + v_5 + v_2$$

 $v_7 := H(v_6)$

What can they do?

$$v_7 := H(v_6)$$
 $v_6 := v_7 + v_4$

Crypto 2016

 $v_8 := v_7 + v_1$ $return(v_8, v_5)$

Take field elements as input	$\mathcal{P}^H(v_1,v_2,v_3):$
	$v_4 := H(v_1)$
	$v_5 := H(v_3)$
	$v_6 := v_4 + v_5 + v_2$
	$v_7 := H(v_6)$
	$v_8 := v_7 + v_1$
	$return(v_8, v_5)$

Take field elements as input

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$

Query the random oracle $v_5 := H(v_3)$

$$v_6 := v_4 + v_5 + v_2$$

 $v_7 := H(v_6)$

 $v_8 := v_7 + v_1$

Often we write \mathcal{P}^H to be explicit

 $return(v_8, v_5)$

 $\mathcal{P}^{H}(v_{1},v_{2},v_{3}):$ Take field elements as input $v_4 := H(v_1)$ $v_5 := H(v_3)$ Query the random oracle $v_6 := v_4 + v_5 + v_2$ Use a fixed linear combination $v_7 := H(v_6)$ $v_8 := v_7 + v_1$ $return(v_8, v_5)$

 $\mathcal{P}^{H}(v_{1},v_{2},v_{3}):$ Take field elements as input $v_4 := H(v_1)$ $v_5 := H(v_3)$ Query the random oracle $v_6 := v_4 + v_5 + v_2$ Use a fixed linear combination $v_7 := H(v_6)$ $v_8 := v_7 + v_1$ $return(v_8, v_5)$ Return any number of elements

Modeling Linicrypt Programs

Algorithmically

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$

Modeling Linicrypt Programs

Algorithmically

$$\mathcal{P}^H(v_1,v_2,v_3)$$
:

$$v_4 := H(v_1)$$

$$v_5 := H(v_3)$$

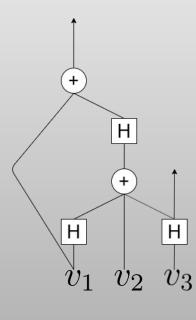
$$v_6 := v_4 + v_5 + v_2$$

$$v_7 := H(v_6)$$

$$v_8 := v_7 + v_1$$

$$return(v_8, v_5)$$

Graphically



Modeling Linicrypt Programs

Algorithmically

$$\mathcal{P}^H(v_1,v_2,v_3)$$
:

$$v_4 := H(v_1)$$

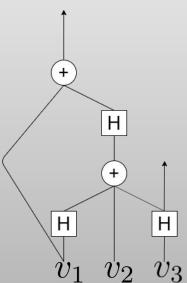
$$v_5 := H(v_3)$$

$$v_6 := v_4 + v_5 + v_2$$

$$v_7 := H(v_6)$$

$$a_1 := a_1 + a_2$$

$$v_8 := v_7 + v_1$$
$$return(v_8, v_5)$$



Algebraically

Previous Work

When are two randomized Linicrypt programs indistinguishable?

Carmer and Rosulek 2016

But what about collision resistance?

Informal Main Theorem

When is this class of linicrypt programs not resistant to collisions/second preimages?

Characterizable by algebraic properties!

Informal Main Theorem

When is this class of linicrypt programs not resistant to collisions/second preimages?

Characterizable by algebraic properties!

Corollary:
Second preimage resistance and collision resistance are the same (asymptotically)

Second Preimages in Linicrypt

$$(\boldsymbol{x} \neq \boldsymbol{x}') \wedge (\mathcal{P}^H(\boldsymbol{x}) = \mathcal{P}^H(\boldsymbol{x}'))$$

Second Preimages in Linicrypt

$$(\boldsymbol{x} \neq \boldsymbol{x}') \land (\mathcal{P}^H(\boldsymbol{x}) = \mathcal{P}^H(\boldsymbol{x}'))$$

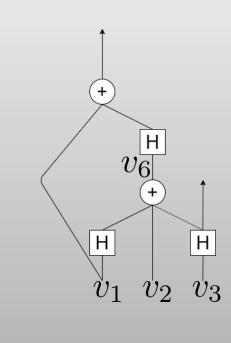
1. The set of input variables are different

Second Preimages in Linicrypt

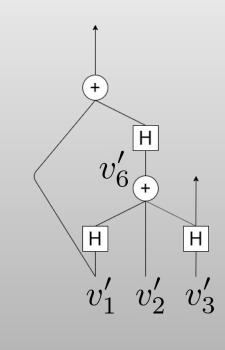
$$(\boldsymbol{x} \neq \boldsymbol{x}') \wedge (\mathcal{P}^H(\boldsymbol{x}) = \mathcal{P}^H(\boldsymbol{x}'))$$

- 1. The set of input variable are different
- 2. The outputs are the same

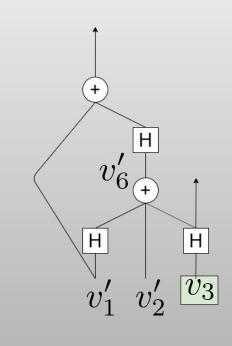
$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$



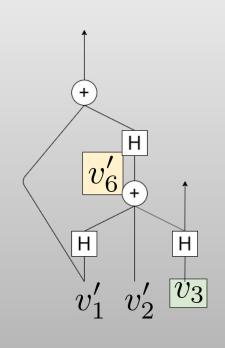
$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$



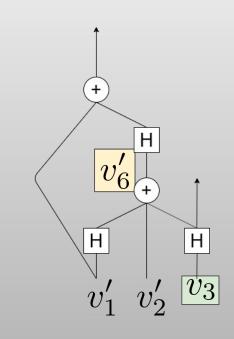
$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$



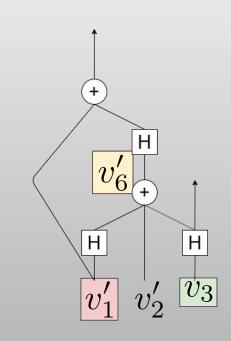
$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$



$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$



$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $v_{1} \neq v_{1}$
 $v_{1} \neq v_{1}$



$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $v_{1} \neq v_{1}$
 $v_{2} = v_{3}$
 $v_{1} \neq v_{2}$

 $(\mathcal{P}^H(oldsymbol{x}) = \mathcal{P}^H(oldsymbol{x}'))$

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}):$$
 $v_{4} := H(v_{1})$
 $v_{5} := H(v_{3})$
 $v_{6} := v_{4} + v_{5} + v_{2}$
 $v_{7} := H(v_{6})$
 $v_{8} := v_{7} + v_{1}$
 $return(v_{8}, v_{5})$
 $v_{1} \neq v_{1}$
 $v_{2} \quad v_{3}$

Finding Collisions

1. Identify oracle queries that are the same between runs

$$v_3$$

2. Identify an oracle query that is different

$$v_6'$$

3. Solve backwards using linear algebra until all queries are defined

$$v_1' v_2'$$

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}): \ v_{4} := H(v_{1}) \ v_{5} := H(v_{3}) \ v_{6} := v_{4} + v_{5} + v_{2} \ v_{7} := H(v_{6}) \ v_{8} := v_{7} + v_{1} \ return(v_{8}, v_{5}) \ M = egin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 0 & 0 \ 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 0 & 0 & 1 & 0 & 0 \ 0 & 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ \end{bmatrix}; \ M = egin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 0 & 1 \ 1 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 1 & 0 \ \end{bmatrix};$$

$$\mathcal{P}^{H}(v_{1}, v_{2}, v_{3}): \ v_{4} := H(v_{1}) \ v_{5} := H(v_{3}) \ v_{6} := v_{4} + v_{5} + v_{2} \ v_{7} := H(v_{6}) \ v_{8} := v_{7} + v_{1} \ return(v_{8}, v_{5}) \ M = egin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \ 0 & 1 & 0 & 0 & 0 & 0 \ 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 0 & 0 & 1 & 0 & 0 \ 0 & 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 1 & 0 \ \end{bmatrix};$$

Every intermediate value is a linear combo of base

vars

The outputs of a linicrypt program are held in $oldsymbol{M}$

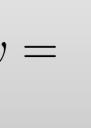
But how do we represent queries to the oracle?

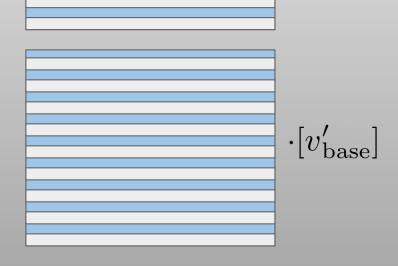
$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_5 \\ v_7 \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_5 \\ v_7 \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_7 \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_7 \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0$$

 $\mathcal C$ holds the oracle queries a program makes

This query corresponds to $|v_4| := |H(v_1)|$

Here are the internals of two runs of \mathcal{P}^H



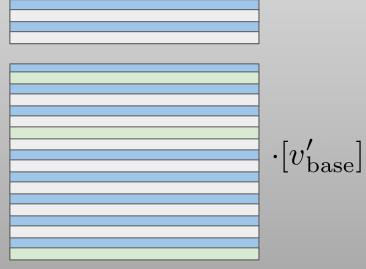


 $\cdot [v_{\mathrm{base}}]$

Here are the internals of two runs of \mathcal{P}^H

Identify base variables shared

 $v = \int_{0}^{\infty}$

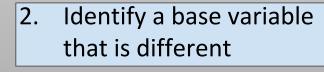


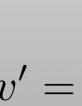
 $\cdot [v_{\mathrm{base}}]$

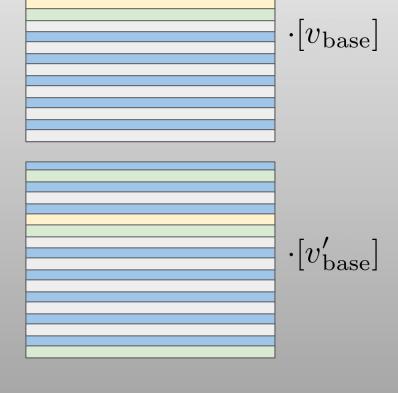
Here are the internals of two runs of \mathcal{P}^H



 Identify base variables shared



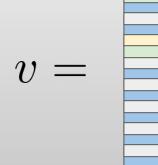


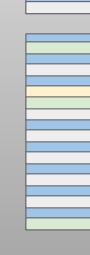


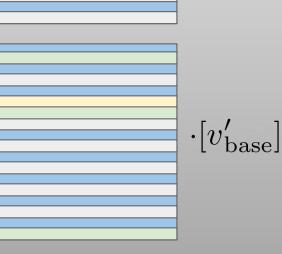
Here are the internals of two runs of \mathcal{P}^H

- Identify base variables shared
- 2. Identify a base variable that is different

Corresponding query must be lin. indep!







 $\cdot [v_{\mathrm{base}}]$

Here are the internals of two runs of \mathcal{P}^H

$$v =$$



3. Solve backwards using linear algebra until all queries are defined





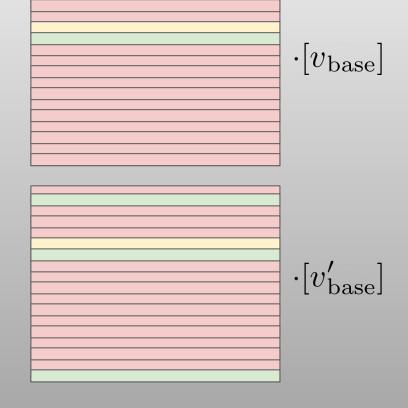
Here are the internals of two runs of \mathcal{P}^H

3. Solve backwards using linear algebra until all queries are defined

$$v_7' = H(v_6')$$

What if either is fixed?





Let $\mathcal{P} = (\mathbf{M}, \mathcal{C})$ be a Linicrypt program.

A collision structure for \mathcal{P} is a tuple $(i^*; c_1, \ldots, c_n)$, where:

Let $\mathcal{P} = (\mathbf{M}, \mathcal{C})$ be a Linicrypt program.

A collision structure for \mathcal{P} is a tuple $(i^*; c_1, \ldots, c_n)$, where:

 $|1, c_1, \ldots, c_n|$ is an ordering of C

Let $\mathcal{P} = (\mathbf{M}, \mathcal{C})$ be a Linicrypt program.

A collision structure for \mathcal{P} is a tuple $(i^*; c_1, \ldots, c_n)$, where:

- 1. c_1, \ldots, c_n is an ordering of \mathcal{C}
- 2. The inputs to c_{i^*} are not in the span of the queries

before it union the total output.

Let $\mathcal{P} = (\mathbf{M}, \mathcal{C})$ be a Linicrypt program.

A collision structure for \mathcal{P} is a tuple $(i^*; c_1, \ldots, c_n)$, where:

- 1. c_1, \ldots, c_n is an ordering of \mathcal{C}
- 2. The inputs to c_{i^*} are not in the span of the queries before it union the total output.
- 3. All following queries $(c_j \mid j \geq i^*)$ outputs are not in

the span of the queries before them union the total output.

1. Identify oracle queries that are the same between runs

$$v_3$$

2. Identify an oracle query that is different

$$v_6'$$

3. Solve backwards using linear algebra until all queries are defined

$$v_1' v_2'$$

1. Identify oracle queries that are the same between runs

$$c_1,\ldots,c_{i^*-1}$$

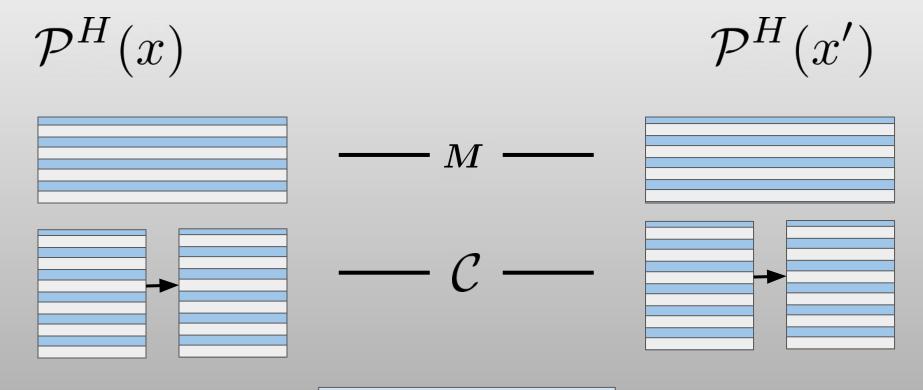
2. Identify an oracle query that is different

$$C_{i^*}$$

3. Solve backwards using linear algebra until all queries are defined

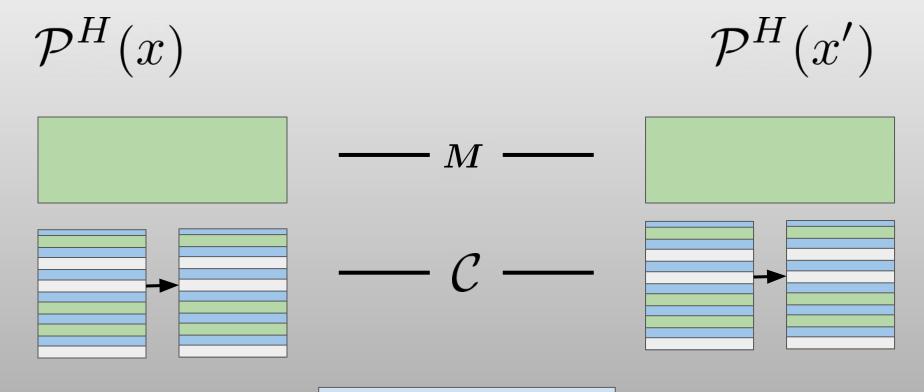
$$c_{i^*+1},\ldots,c_n$$

Finding Collision Structures From Second Preimages



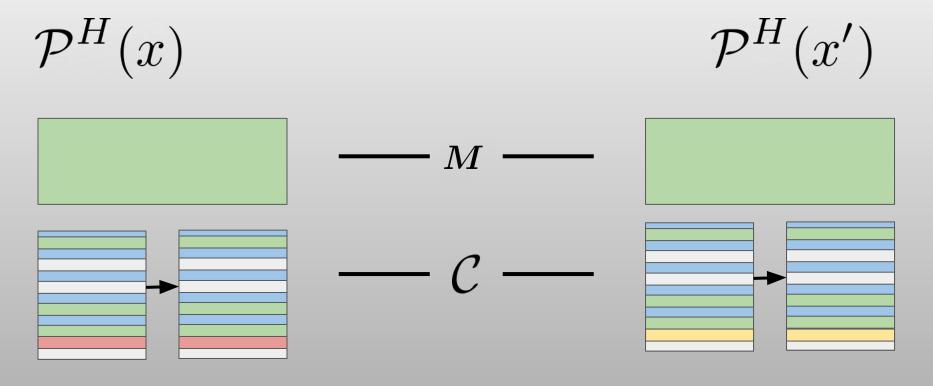
Which queries are green?

Finding Collision Structures From Second Preimages

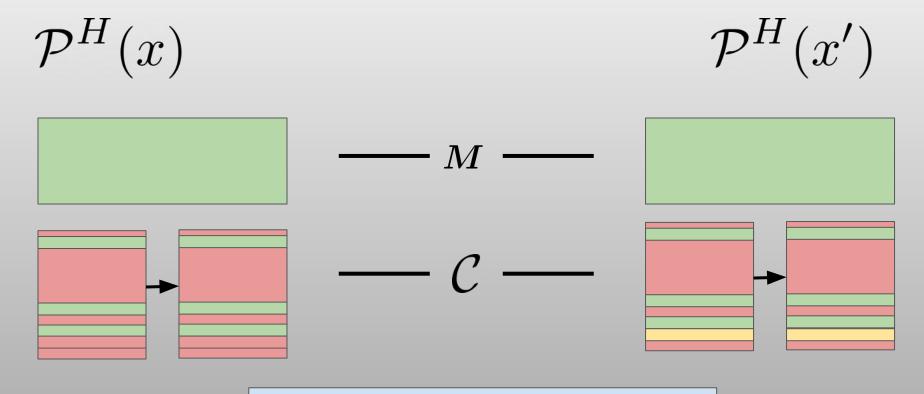


What's our special query?

Finding Collision Structures From Second Preimages



Finding Collision Structures From Second Preimages



Now we can create our Collision

Structure

Theorem Statement

For a linicrypt program with distinct nonces:

Collision Structures in a Program ← No Collision Resistance*

← No 2nd Preimage Resistance*

* modulo degeneracy

Wrap Up

1. Collisions and second preimages can be boiled down to algebra

2. Properties can be determined in poly time

3. We often only have to worry about second preimages

Limitations and future work

Distinct nonces:

$$H(\mathbf{x}, \mathbf{y}) = H(2; H(1; \mathbf{x})) - H(3; \mathbf{y})$$

$$H(\mathbf{x}, \mathbf{y}) = H(H(\mathbf{x})) - H(\mathbf{y})$$

Limitations and future work

Distinct nonces:

$$H(\mathbf{x}, \mathbf{y}) = H(2; H(1; \mathbf{x})) - H(3; \mathbf{y})$$
$$H(\mathbf{x}, \mathbf{y}) = H(H(\mathbf{x})) - H(\mathbf{y})$$

$$\mathbf{y} := H(\mathbf{x})$$
?

NP complete problem!

Limitations and future work

Distinct nonces:

$$H(\mathbf{x}, \mathbf{y}) = H(2; H(1; \mathbf{x})) - H(3; \mathbf{y})$$
$$H(\mathbf{x}, \mathbf{y}) = H(H(\mathbf{x})) - H(\mathbf{y})$$

$$\mathbf{y} := H(\mathbf{x})$$
?

NP complete problem!

Ideal cipher model?

Thank you

What is degeneracy?

Degeneracy? $\mathcal{P}^H(x,y)$: H(x+y)

All queries are identical, but inputs are different!

$$H(\mathbf{a} + \mathbf{b}) = \mathbf{c} = H(\mathbf{d} + (\mathbf{a} + \mathbf{b} - \mathbf{d}))$$

We have an entire space of collision inputs!