Characterizing Collision and Second-Preimage Resistance in Linicrypt

Ian McQuoid    Trevor Swope    Mike Rosulek

Oregon State University
An Introduction

Which program is collision resistant?

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

\[
\begin{align*}
v_4 & := H(v_1) \\
v_5 & := H(v_3) \\
v_6 & := v_4 + v_5 + v_2 \\
v_7 & := H(v_6) \\
v_8 & := v_7 + v_1 + v_2
\end{align*}
\]

\text{return}(v_8, v_5)

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

\[
\begin{align*}
v_4 & := H(v_1) \\
v_5 & := H(v_3) \\
v_6 & := v_4 + v_5 + v_2 \\
v_7 & := H(v_6) \\
v_8 & := v_7 + v_1
\end{align*}
\]

\text{return}(v_8, v_5)
Linicrypt programs are a class of algorithms

Introduced by Carmer and Rosulek, Crypto 2016

What can they do?

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]
What can we do?

<table>
<thead>
<tr>
<th>Take field elements as input</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{P}^H(v_1, v_2, v_3)$ :</td>
</tr>
<tr>
<td>$v_4 := H(v_1)$</td>
</tr>
<tr>
<td>$v_5 := H(v_3)$</td>
</tr>
<tr>
<td>$v_6 := v_4 + v_5 + v_2$</td>
</tr>
<tr>
<td>$v_7 := H(v_6)$</td>
</tr>
<tr>
<td>$v_8 := v_7 + v_1$</td>
</tr>
<tr>
<td>$return(v_8, v_5)$</td>
</tr>
</tbody>
</table>
What can we do?

Take field elements as input

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

\[ v_4 := H(v_1) \]

\[ v_5 := H(v_3) \]

\[ v_6 := v_4 + v_5 + v_2 \]

\[ v_7 := H(v_6) \]

\[ v_8 := v_7 + v_1 \]

\[ \text{return } (v_8, v_5) \]

Query the random oracle

Often we write \( \mathcal{P}^H \) to be explicit
What can we do?

Take field elements as input

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

\[ v_4 := H(v_1) \]

Query the random oracle

\[ v_5 := H(v_3) \]

Use a fixed linear combination

\[ v_6 := v_4 + v_5 + v_2 \]

\[ v_7 := H(v_6) \]

\[ v_8 := v_7 + v_1 \]

\[ \text{return}(v_8, v_5) \]
What can we do?

Take field elements as input

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]

Query the random oracle

\[ v_5 := H(v_3) \]

Use a fixed linear combination

\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

Return any number of elements

return\((v_8, v_5)\)
Algorithmically

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]
Modeling Linicrypt Programs

Algorithmically

\[ P^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]

Graphically

```
+ -> H
+ |-> H
|   |-> H
v_1 | v_2 | v_3
```
Modeling Linicrypt Programs

Algorithmically

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]

Graphically

Algebraically

\[ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} ; \]
\[ \begin{cases} [1 0 0 0 0 0], [0 0 0 1 0 0], \\ [0 0 1 0 0 0], [0 0 0 0 1 0], \\ [0 1 0 1 1 0], [0 0 0 0 0 1] \end{cases} \]
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

\[(x \neq x') \land (P^H(x) = P^H(x'))\]
Second Preimages in Linicrypt

\[(x \neq x') \land (\mathcal{P}^H(x) = \mathcal{P}^H(x'))\]

1. The set of input variables are different
Second Preimages in Linicrypt

\[(x \neq x') \land (\mathcal{P}^H(x) = \mathcal{P}^H(x'))\]

1. The set of input variable are different

2. The outputs are the same
Collisions in Linicrypt

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

return \((v_8, v_5)\)
\[ \mathcal{P}^H(v_1, v_2, v_3) : \\
v_4 := H(v_1) \\
v_5 := H(v_3) \\
v_6 := v_4 + v_5 + v_2 \\
v_7 := H(v_6) \\
v_8 := v_7 + v_1 \\
\text{return}(v_8, v_5)\]
Collisions in Linicrypt

\[ P^H(v_1, v_2, v_3) : \]

\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

\[ \text{return}(v_8, v_5) \]
\[ P^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]
\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]
\[ \mathcal{P}^H(v_1, v_2, v_3) : \]

\[
\begin{align*}
v_4 & := H(v_1) \\
v_5 & := H(v_3) \\
v_6 & := v_4 + v_5 + v_2 \\
v_7 & := H(v_6) \\
v_8 & := v_7 + v_1 
\end{align*}
\]

\[ \text{return} (v_8, v_5) \]

\[ v'_1 \neq v_1 \]
Collisions in Linicrypt

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

\[ v_1' \neq v_1 \]

\[ \text{return}\((v_8, v_5)\) \]

\[ (\mathcal{P}^H(x) = \mathcal{P}^H(x')) \]
Collisions in Linicrypt

\[ P^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

\[ v_1' \neq v_1 \]

\[ (P^H(x) = P^H(x')) \]
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

   \[ v'_1 \ v'_2 \]
\[ P^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return } (v_8, v_5) \]

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

\[ M = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}; \]
Every intermediate value is a linear combo of base vars
Algebraic Representation

\[ \mathcal{P}^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]

\[ \text{return}(v_8, v_5) \]

\[ (v_1) \begin{pmatrix} v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_5 \\ v_7 \end{pmatrix} \]

\[ M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}; \]

The outputs of a linicrypt program are held in \( M \)
Algebraic Representation

\[ P^H(v_1, v_2, v_3) : \]
\[ v_4 := H(v_1) \]
\[ v_5 := H(v_3) \]
\[ v_6 := v_4 + v_5 + v_2 \]
\[ v_7 := H(v_6) \]
\[ v_8 := v_7 + v_1 \]
\[ \text{return}(v_8, v_5) \]

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

\[ M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}; \]

But how do we represent queries to the oracle?
Algebraic Representation

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

\[
M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{bmatrix};
\]

\[
C = \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
\end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
\end{pmatrix} \right\}
\]

\[C\] holds the oracle queries a program makes.
Algebraic Representation

\[
\begin{align*}
\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{pmatrix} \\
&= \begin{pmatrix} \color{blue}{1} & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} ; \\
C &= \left\{ \begin{pmatrix} [1 & 0 & 0 & 0 & 0] \\ [0 & 0 & 0 & 1 & 0 & 0] \end{pmatrix} , \begin{pmatrix} [0 & 0 & 1 & 0 & 0] \\ [0 & 0 & 0 & 1 & 0] \end{pmatrix} , \begin{pmatrix} [0 & 1 & 0 & 1 & 1] \\ [0 & 0 & 0 & 0 & 1] \end{pmatrix} \right\}
\end{align*}
\]

This query corresponds to \( v_4 := H(v_1) \)
What is a collision structure?

Here are the internals of two runs of $P^H$

\[ v = \cdot [v_{\text{base}}] \]

\[ v' = \cdot [v'_{\text{base}}] \]
What is a collision structure?

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

1. Identify base variables shared

\[ u = \cdot [u_{\text{base}}] \]

\[ u' = \cdot [u'_{\text{base}}] \]
What is a collision structure?

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

1. Identify base variables shared

2. Identify a base variable that is different

\[ v = \cdot [v_{base}] \]

\[ v' = \cdot [v'_{base}] \]
What is a collision structure?

Here are the internals of two runs of $P^H$

1. Identify base variables shared

2. Identify a base variable that is different

Corresponding query must be lin. indep!

$\mathbf{v} = \begin{bmatrix} v_{\text{base}} \end{bmatrix}$

$\mathbf{v}' = \begin{bmatrix} v'_{\text{base}} \end{bmatrix}$
What is a collision structure?

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

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

$$v = \cdot [v_{\text{base}}]$$

$$v' = \cdot [v'_{\text{base}}]$$
What is a collision structure?

Here are the internals of two runs of $P^H$

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

$$v'_7 = H(v'_6)$$

What if either is fixed?
What is a collision structure?

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:
What is a collision structure?

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 $\mathcal{C}$
What is a collision structure?

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.
What is a collision structure?

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 then union the total output.
What is a collision structure?

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

   \[ v'_1 \quad v'_2 \]
What is a collision structure?

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

\[ \mathcal{P}^H(x) \]

Which queries are green?
Finding Collision Structures From Second Preimages

\[ \mathcal{P}^H(x) \rightarrow M \rightarrow C \rightarrow \mathcal{P}^H(x') \]

What’s our special query?
Finding Collision Structures From Second Preimages

\[ \mathcal{P}^H(x) \]

\[ \mathcal{P}^H(x') \]

\[ M \]

\[ C \]
Finding Collision Structures From Second Preimages

$\mathcal{P}^H(x)$

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

Now we can create our Collision Structure
Theorem Statement

For a linicrypt program with distinct nonces:

Collision Structures in a Program $\iff$ No Collision Resistance*
$\iff$ 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
Distinct nonces:

\[
H(x, y) = H(2; H(1; x)) - H(3; y)
\]

\[
H(x, y) = H(H(x)) - H(y)
\]
Limitations and future work

Distinct nonces:

\[
H(x, y) = H(2; H(1; x)) - H(3; y)
\]

\[
H(x, y) = H(H(x)) - H(y)
\]

\[y := H(x)\]?

NP complete problem!
Limitations and future work

Distinct nonces:

\[ H(x, y) = H(2; H(1; x)) - H(3; y) \]

\[ H(x, y) = H(H(x)) - H(y) \]

\[ y := H(x)? \]

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!

\[ H(a + b) = c = H(d + (a + b - d)) \]

We have an entire space of collision inputs!