Attacking the Knudsen-Preneel Compression Functions

Onur Özen[1] and Thomas Shrimpton[2] and Martijn Stam[1]

[1]Ecole Polytechnique Fédérale de Lausanne

[2]Portland State University

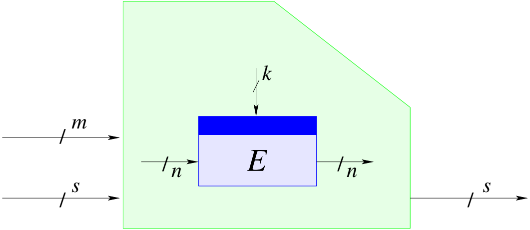
Fast Software Encryption, 2010

Outline

The Compression Function

Most well known constructions use (single call) blockcipher based compression functions

E.g. SHA-1, MD5, Whirlpool, Tiger ...

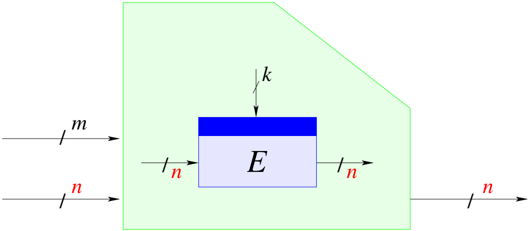


$$E : \{0,1\}^n \times \{0,1\}^k \longrightarrow \{0,1\}^n$$

The Compression Function

Most well known constructions use (single call) blockcipher based compression functions

E.g. PGV Compression Functions, SHA-1, MD5, Tiger ...

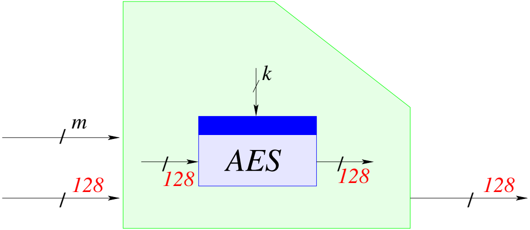


$$E : \{0,1\}^n \times \{0,1\}^k \longrightarrow \{0,1\}^n$$

The Compression Function

Most well known constructions use (single call) blockcipher based compression functions.

When the blockcipher is instantiated by AES

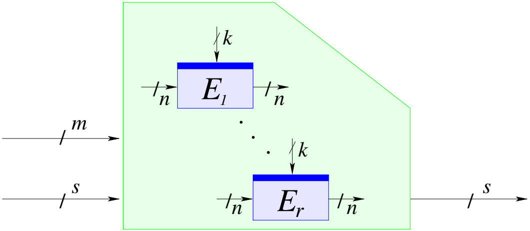


It takes 2^{64} operations to find a collision (due to birthday attack).

Considered to be insufficient!

Compression Functions

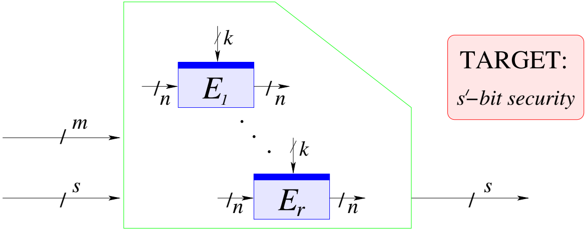
Multi-length blockcipher based compression functions:
Based on small blockciphers running (in general) in parallel, outputs more than n bits ($s > n$).



E.g. MDC-2, MDC-4, Abreast-DM, Tandem-DM, KP ...

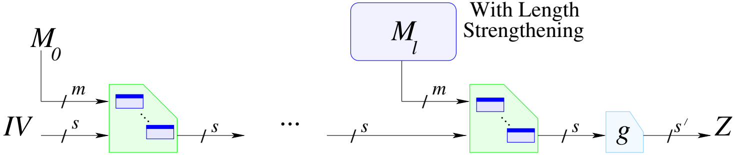
The Approach of Knudsen-Preneel

1. Let the output size and the number of blockcipher calls vary in order to guarantee a particular security target (say $s' \leq s$ bits).



The Approach of Knudsen-Preneel

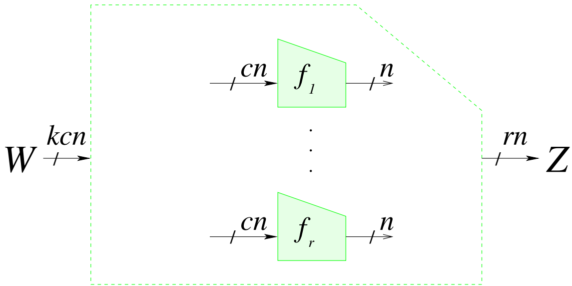
1. Let the output size and the number of blockcipher calls vary in order to guarantee a particular security target (say $s' \leq s$ bits).
2. When iterated, one could compress the final state to a desired length for the security target.



Knudsen-Preneel Compression Functions

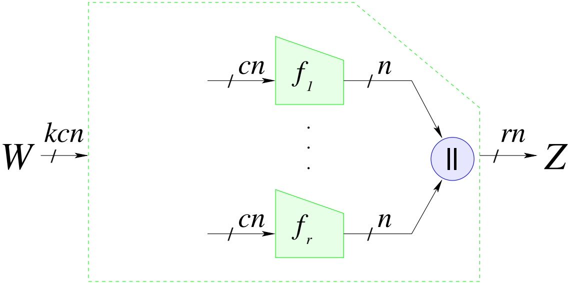
$$H : \{0,1\}^{kcn} \rightarrow \{0,1\}^{rn}$$

Knudsen-Preneel Compression Functions



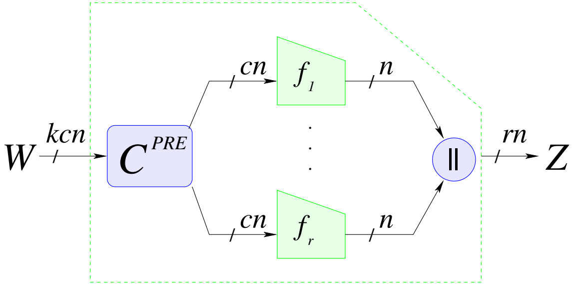
$f_1, \ldots, f_r : \{0,1\}^{cn} \to \{0,1\}^{n}$ running in parallel,

Knudsen-Preneel Compression Functions



$$Z = y_1||...||y_r$$

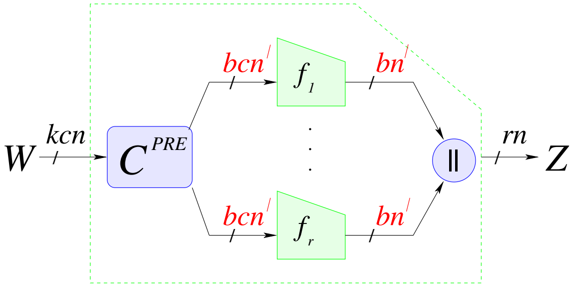
Knudsen-Preneel Compression Functions



$$(x_1, ..., x_r) = C^{PRE}(W)$$

C^{PRE} is based on a generator matrix of an $[r, k, d]$ error-correcting code over $\mathbb{F}_{2^c}$.

Knudsen-Preneel Compression Functions



$$(x_1, ..., x_r) = C^{PRE}(W)$$

C^{PRE} is based on a generator matrix of an $[r, k, d]$ error-correcting code over $\mathbb{F}_{2^{bc}}$ (where $bn' = n$).

An Example: $KP[5,3,3]_4$

Given $W = (W_1||W_2||W_3||W_4||W_5||W_6)$, $W_i \in \{0,1\}^n$

$$(W_1 \,||\, W_2) = x_1 \xrightarrow{2n} f_1 \xrightarrow{n} y_1$$

$$(W_3 \,||\, W_4) = x_2 \xrightarrow{2n} f_2 \xrightarrow{n} y_2$$

$$(W_5 \,||\, W_6) = x_3 \xrightarrow{2n} f_3 \xrightarrow{n} y_3$$

$$(W_1 \oplus W_3 \oplus W_5 \,||\, W_2 \oplus W_4 \oplus W_6) = x_4 \xrightarrow{2n} f_4 \xrightarrow{n} y_4$$

$$(W_1 \oplus W_3 \oplus W_5 \oplus W_6 \,||\, W_2 \oplus W_3 \oplus W_4 \oplus W_6) = x_5 \xrightarrow{2n} f_5 \xrightarrow{n} y_5$$

Knudsen-Preneel Compression Functions

Security Claims:

Collision Resistance

Any collision attack needs at least $2^{(d-1)n/2}$ time.
Intuition : The minimum number of small compression functions for which the simultaneous collisions need to be found.
Update by Watanabe : An attack of time complexity 2^n.

Knudsen-Preneel Compression Functions

Security Claims:

Collision Resistance

Any collision attack needs at least $2^{(d-1)n/2}$ time.
Intuition : The minimum number of small compression functions for which the simultaneous collisions need to be found.
Update by Watanabe : An attack of time complexity 2^n.

Preimage Resistance

Conjecture: Any preimage attack requires at least $2^{(d-1)n}$ time.
Update: Today's talk!

Our Contribution

New Security Analysis of KP Constructions

1. A precise formalization of the KP transform and, more generally, blockwise-linear schemes.

Our Contribution

New Security Analysis of KP Constructions

1. A precise formalization of the KP transform and, more generally, blockwise-linear schemes.
2. A security proof for preimage resistance of the KP compression functions in the information-theoretic model.

Our Contribution

New Security Analysis of KP Constructions

1. A precise formalization of the KP transform and, more generally, blockwise-linear schemes.
2. A security proof for preimage resistance of the KP compression functions in the information-theoretic model.
3. New preimage attacks going well *below* the conjectured lower bound by Knudsen and Preneel!
   - With minimum number of queries.
   - Optimal time complexity for 9 out of 16 schemes.
   - Better time complexity than the one given by KP in every case but two where we get the same complexity.

Outline

1. Introduction
2. Information-Theoretic Considerations
3. Our Preimage Attacks on KP-Constructions
4. Conclusion

Security notion

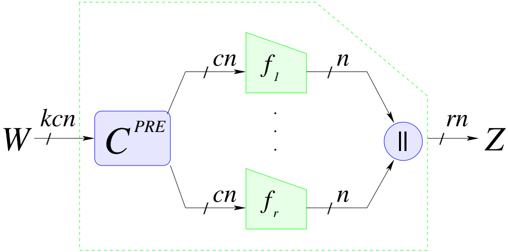
Definition (Everywhere preimage resistance)

Let $c, r, s, t > 0$ be integer parameters, and fix a blocksize $n > 0$. Let $H : \{0,1\}^{tn} \to \{0,1\}^{sn}$ be a PuRF-based compression function taking r oracles $f_1, \ldots, f_r \in \mathcal{F}(cn, n)$. The everywhere preimage-finding advantage of adversary $\mathcal{A}$ is defined to be

$$\mathrm{Adv}_H^{\mathrm{epre}}(\mathcal{A}) = \max_{Z \in \{0,1\}^{sn}} \left\{ \Pr \left[f_1 \ldots f_r \xleftarrow{\$} \mathcal{F}(cn, n), (Z') \leftarrow \mathcal{A}^{f_1 \ldots f_r}(Z) : Z = H^{f_1 \ldots f_r}(Z') \right] \right\}$$

Define $\mathrm{Adv}_H^{\mathrm{epre}}(q)$ and $\mathrm{Adv}_H^{\mathrm{epre}}(t)$ as the maximum advantage over all adversaries making at most q queries to each of their oracles respectively running in time at most t.

Information Theoretic Security Proof

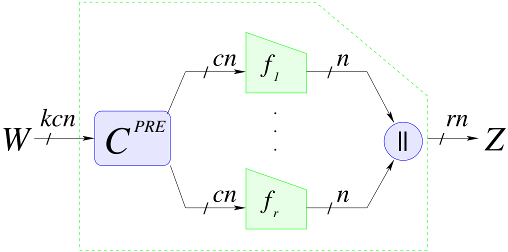


Corollary

Let $H = KP^b[r, k, d]_e$. Then asymptoticallly for n (with $b|n$) and $q \leq g(n) \left(\frac{2^n}{e}\right)^{r/k}$ with $g(n) = o(1)$, $\mathrm{Adv}_H^{epre}(q) = o(1)$.

So, $\Omega(2^{rn/k})$ queries are necessary to win the epre experiment.

Information Theoretic Security Proof



Corollary

Let $H = KP^b[r,k,d]_e$. Then asymptoticallly for n (with $b|n$) and $q \leq g(n)\left(\frac{2^n}{e}\right)^{r/k}$ with $g(n) = o(1)$, $\mathrm{Adv}_H^{epre}(q) = o(1)$.

So, $\Omega(2^{rn/k})$ queries are necessary to win the epre experiment.

It also serves as the best case time complexity!

The Picture so far

Code	Query Low. Bound	KP-Conjec. Low. Bound	KP-Attack Time
$[r, k, d]_{2^e}$	$2^{rn/k}$	$2^{(d-1)n}$	
$[5, 3, 3]_4$	$2^{5n/3}$	2^{2n}	2^{2n}
$[8, 5, 3]_4$	$2^{8n/5}$	2^{2n}	2^{3n}
$[12, 9, 3]_4$	$2^{4n/3}$	2^{2n}	2^{3n}
$[9, 5, 4]_4$	$2^{9n/5}$	2^{3n}	2^{4n}
$[16, 12, 4]_4$	$2^{4n/3}$	2^{3n}	2^{4n}
$[6, 4, 3]_{16}$	$2^{3n/2}$	2^{2n}	2^{2n}
$[8, 6, 3]_{16}$	$2^{4n/3}$	2^{2n}	2^{2n}
$[12, 10, 3]_{16}$	$2^{6n/5}$	2^{2n}	2^{2n}
$[9, 6, 4]_{16}$	$2^{3n/2}$	2^{3n}	2^{3n}
$[16, 13, 4]_{16}$	$2^{16n/13}$	2^{3n}	2^{3n}

$$(f_i : \{0,1\}^{2n} \rightarrow \{0,1\}^n)$$

The Picture so far

Code	Query Low. Bound	KP-Conjec. Low. Bound	KP-Attack Time
$[r,k,d]_{2^e}$	$2^{rn/k}$	$2^{(d-1)n}$	
$[4,2,3]_8$	2^{2n}	2^{2n}	2^{2n}
$[6,4,3]_8$	$2^{3n/2}$	2^{2n}	2^{2n}
$[9,7,3]_8$	$2^{9n/7}$	2^{2n}	2^{2n}
$[5,2,4]_8$	$2^{5n/2}$	2^{3n}	2^{3n}
$[7,4,4]_8$	$2^{7n/4}$	2^{3n}	2^{3n}
$[10,7,4]_8$	$2^{10n/7}$	2^{3n}	2^{3n}

$(f_i : \{0,1\}^{3n} \rightarrow \{0,1\}^n)$

Outline

A Warm-up Example of our Attack on $KP[5,3,3]_4$

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r,k,d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[5,3,3]_4$	$2^{5n/3}$		2^{2n}	2^{2n}

$$(W_1 \,\|\, W_2) = x_1 \xrightarrow{2n} f_1 \xrightarrow{n} y_1$$

$$(W_3 \,\|\, W_4) = x_2 \xrightarrow{2n} f_2 \xrightarrow{n} y_2$$

$$(W_5 \,\|\, W_6) = x_3 \xrightarrow{2n} f_3 \xrightarrow{n} y_3$$

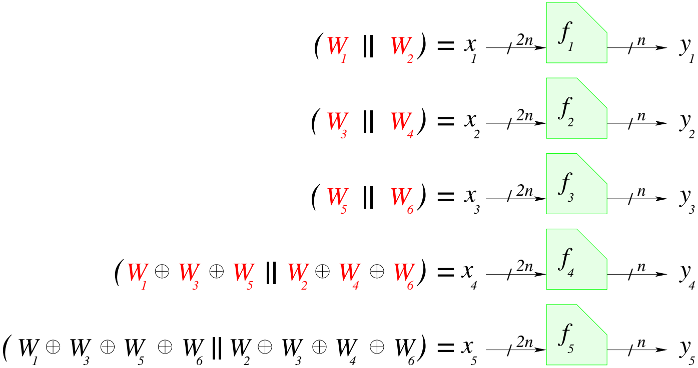
$$(W_1 \oplus W_3 \oplus W_5 \,\|\, W_2 \oplus W_4 \oplus W_6) = x_4 \xrightarrow{2n} f_4 \xrightarrow{n} y_4$$

$$(W_1 \oplus W_3 \oplus W_5 \oplus W_6 \,\|\, W_2 \oplus W_3 \oplus W_4 \oplus W_6) = x_5 \xrightarrow{2n} f_5 \xrightarrow{n} y_5$$

A Warm-up Example of our Attack on $KP[5, 3, 3]_4$

Observation

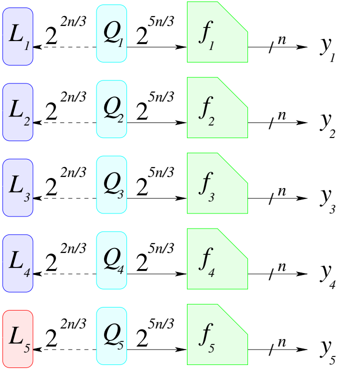
1. $(0^a||x) \oplus (0^a||y) = (0^a||x \oplus y)$
2. $x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0$



A Warm-up Example of our Attack on $KP[5,3,3]_4$

Query Phase: Takes $\mathcal{O}(2^{5n/3})$ time!

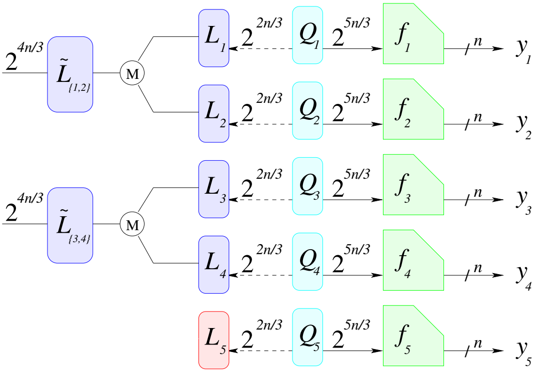
- Let $x_i = (x_i^1 || x_i^2)$. Ask $x_i^1, x_i^2 \in 0^{n/6} \times \{0,1\}^{5n/6}$ to each f_i.
- Keep the lists L_i containing partial preimages.



A Warm-up Example of our Attack on $KP[5,3,3]_4$

Merge Phase: (Takes $\mathcal{O}(n2^{4n/3})$ time!) Construct

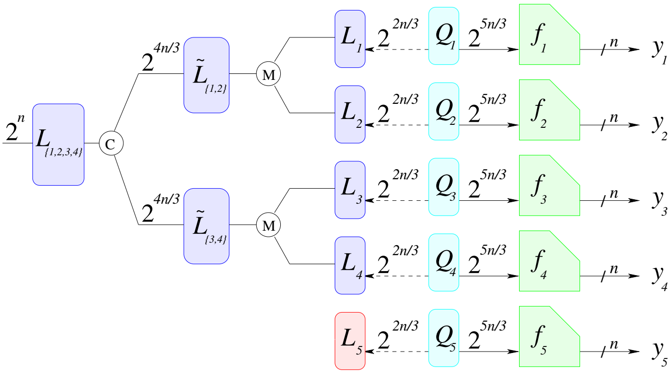
$$\tilde{L}_{\{1,2\}} = \{((x_1,x_2), x_1 \oplus x_2) | (x_1,x_2) \in L_1 \times L_2\},$$
$$\tilde{L}_{\{3,4\}} = \{((x_3,x_4), x_3 \oplus x_4) | (x_3,x_4) \in L_3 \times L_4\}$$



A Warm-up Example of our Attack on $KP[5,3,3]_4$

Join Phase: (Takes $\mathcal{O}(n2^{4n/3})$ time!)

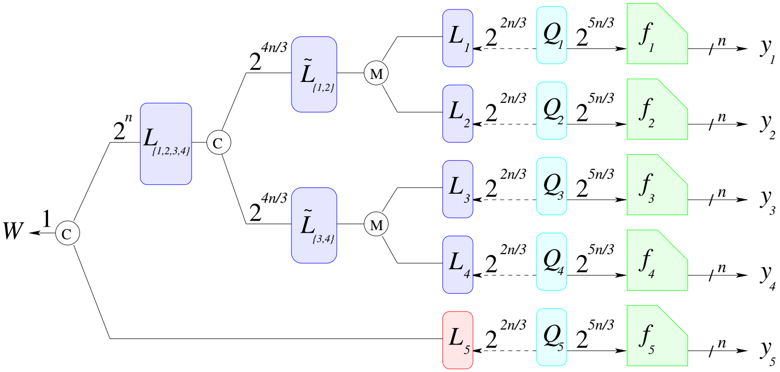
Keep all solutions of $x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0$ in
$L_{\{1,2,3,4\}} = \{(x_1, x_2, x_3, x_4) \in L_1 \times \ldots \times L_4 | x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0\}$



A Warm-up Example of our Attack on $KP[5,3,3]_4$

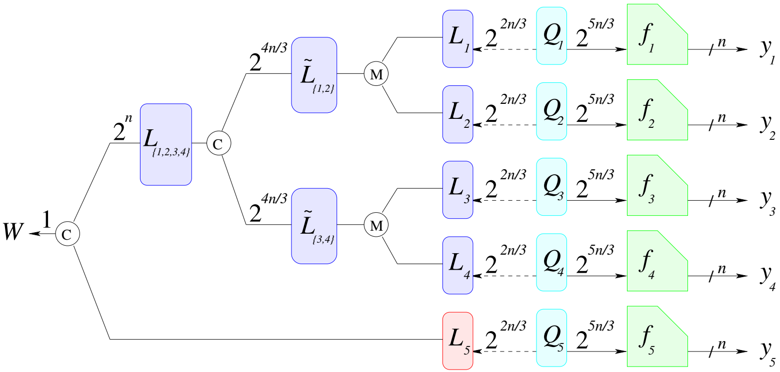
Finalization: Takes $\mathcal{O}(n2^n)$ time!

- Check L_5 membership!



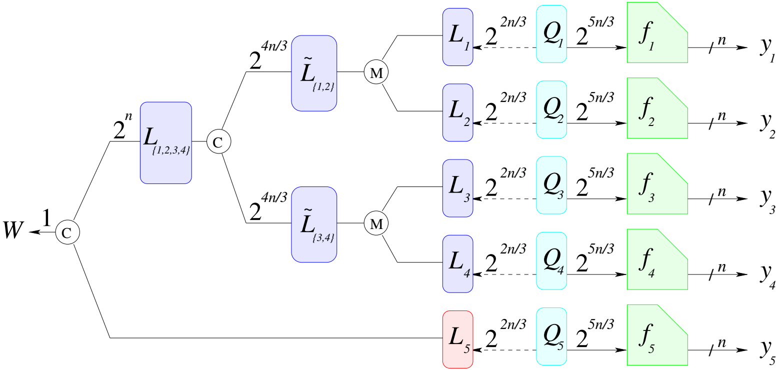
Overall Comparison

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r, k, d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[5, 3, 3]_4$	$2^{5n/3}$	$2^{5n/3}$	2^{2n}	2^{2n}



The Core Observations

1. The Relation $x_1 \oplus x_2 \oplus x_3 \oplus x_4 = 0$ is defined by a dual codeword: $h = (11110)_4$.
2. The complexity of Merge and Join Phases are directly related with the Hamming weight of h.



Extending Our Attack to all MDS-Schemes

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r, k, d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[5, 3, 3]_4$	$2^{5n/3}$	$2^{5n/3}$	2^{2n}	2^{2n}
$[8, 5, 3]_4$	$2^{8n/5}$	$2^{8n/5}$	2^{2n}	2^{3n}
$[12, 9, 3]_4$	$2^{4n/3}$	$2^{4n/3}$	2^{2n}	2^{3n}
$[9, 5, 4]_4$	$2^{9n/5}$	$2^{11n/5}$	2^{3n}	2^{4n}
$[16, 12, 4]_4$	$2^{4n/3}$	$2^{7n/3}$	2^{3n}	2^{4n}
$[6, 4, 3]_{16}$	$2^{3n/2}$	$2^{3n/2}$	2^{2n}	2^{2n}
$[8, 6, 3]_{16}$	$2^{4n/3}$	$2^{4n/3}$	2^{2n}	2^{2n}
$[12, 10, 3]_{16}$	$2^{6n/5}$	$2^{6n/5}$	2^{2n}	2^{2n}
$[9, 6, 4]_{16}$	$2^{3n/2}$	2^{2n}	2^{3n}	2^{3n}
$[16, 13, 4]_{16}$	$2^{16n/13}$	2^{2n}	2^{3n}	2^{3n}

$$(f_i : \{0,1\}^{2n} \rightarrow \{0,1\}^n)$$

Extending Our Attack to all MDS-Schemes (Cont.)

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r,k,d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[4,2,3]_8$	2^{2n}	2^{2n}	2^{2n}	2^{2n}
$[6,4,3]_8$	$2^{3n/2}$	$2^{3n/2}$	2^{2n}	2^{2n}
$[9,7,3]_8$	$2^{9n/7}$	$2^{9n/7}$	2^{2n}	2^{2n}
$[5,2,4]_8$	$2^{5n/2}$	2^{3n}	2^{3n}	2^{3n}
$[7,4,4]_8$	$2^{7n/4}$	$2^{9n/4}$	2^{3n}	2^{3n}
$[10,7,4]_8$	$2^{10n/7}$	2^{2n}	2^{3n}	2^{3n}

$(f_i : \{0,1\}^{3n} \to \{0,1\}^n)$

Extending Our Attack to Non-MDS-Schemes

- Since $d^{\perp} < k + 1$ for non-MDS codes, we can no longer reconstruct a unique W after the first Merge-Join phase.
- We require one more Merge and Join Phases using another dual codeword.

Choice of code

- Our attacks against the four non-MDS codes were based on the generator matrix given by Magma.
- Non-equivalent codes may perform differently under our attack (they might not have the same $d^{\perp}$)

Overall Results

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r,k,d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[5,3,3]_4$	$2^{5n/3}$	$2^{5n/3}$	2^{2n}	2^{2n}
$[8,5,3]_4$	$2^{8n/5}$	$2^{8n/5}$	2^{2n}	2^{3n}
$[12,9,3]_4$	$2^{4n/3}$	$2^{4n/3}$	2^{2n}	2^{3n}
$[9,5,4]_4$	$2^{9n/5}$	$2^{11n/5}$	2^{3n}	2^{4n}
$[16,12,4]_4$	$2^{4n/3}$	$2^{7n/3}$	2^{3n}	2^{4n}
$[6,4,3]_{16}$	$2^{3n/2}$	$2^{3n/2}$	2^{2n}	2^{2n}
$[8,6,3]_{16}$	$2^{4n/3}$	$2^{4n/3}$	2^{2n}	2^{2n}
$[12,10,3]_{16}$	$2^{6n/5}$	$2^{6n/5}$	2^{2n}	2^{2n}
$[9,6,4]_{16}$	$2^{3n/2}$	2^{2n}	2^{3n}	2^{3n}
$[16,13,4]_{16}$	$2^{16n/13}$	2^{2n}	2^{3n}	2^{3n}

$$(f_i : \{0,1\}^{2n} \rightarrow \{0,1\}^n)$$

Overall Results

Code	Query Low. Bound	Our Attack Time	KP-Conjec. Low. Bound	KP-Attack Time
$[r,k,d]_{2^e}$	$2^{rn/k}$		$2^{(d-1)n}$	
$[4,2,3]_8$	2^{2n}	2^{2n}	2^{2n}	2^{2n}
$[6,4,3]_8$	$2^{3n/2}$	$2^{3n/2}$	2^{2n}	2^{2n}
$[9,7,3]_8$	$2^{9n/7}$	$2^{9n/7}$	2^{2n}	2^{2n}
$[5,2,4]_8$	$2^{5n/2}$	2^{3n}	2^{3n}	2^{3n}
$[7,4,4]_8$	$2^{7n/4}$	$2^{9n/4}$	2^{3n}	2^{3n}
$[10,7,4]_8$	$2^{10n/7}$	2^{2n}	2^{3n}	2^{3n}

$$(f_i : \{0,1\}^{3n} \rightarrow \{0,1\}^n)$$

Outline

Conclusion

- We presented a new preimage attack whose time complexity is well below (nearly for all cases) the conjectured lower bound given by Knudsen and Preneel.
- We determined a lower bound on the query complexity to successfully find preimages.
- Based on our security proof, the query complexity of our new attack is essentially optimal (up to a small factor).
- For 9 out of the 16 schemes, our new preimage-finding attack is optimal.
- For the remaining seven schemes we leave a gap between the information-theoretic lower bound and the real-life upper bound.

Upcoming Work: Similar Analysis for the Collision Resistance!

감사합니다

(THANK YOU!)