KFC - The Krazy Feistel Cipher

Thomas Baignères   Matthieu Finiasz

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

ASIACRYPT 2006

Short State of the Art in Block Ciphers

Block Ciphers' specialists are very good at designing extreme constructions

On the one hand: Feistel scheme with 3 perfectly random functions.

- Provably secure in the Luby-Rackoff model (computationally unbounded adversary with limited queries)
- Unpractical $\approx 2^{70}$ random bits are necessary to instantiate a 128-bit block scheme.

On the other hand: AES and friends.

- Incredibly fast
- Only practically secure: none of the smart cryptanalysts who attacked them was able to break them (yet).
- $\rightsquigarrow$ don't miss today's new cryptanalytic results on IDEA!

Short State of the Art in Block Ciphers

Block Ciphers' specialists are very good at designing extreme constructions

On the one hand: Feistel scheme with 3 perfectly random functions.

- Provably secure in the Luby-Rackoff model (computationally unbounded adversary with limited queries)
- Unpractical $\approx 2^{70}$ random bits are necessary to instantiate a 128-bit block scheme.

On the other hand: AES and friends.

- Incredibly fast
- Only practically secure: none of the smart cryptanalysts who attacked them was able to break them (yet).
- $\rightsquigarrow$ don't miss today's new cryptanalytic results on IDEA!

Yet another Block Cipher?

KFC lies in-between both extremes:

- It comes with security proofs in the Luby-Rackoff model,
- and is practical (we mean, it can be implemented in practice).

More precisely, depending on the parameters choice:

- KFC is provably secure against d-limited adversaries for values of d ranging from 2 up to 70.
- This is enough to resist several statistical attacks.
- This includes Linear and Differential Cryptanalysis (taking hull/differentials effects in consideration), higher order differential cryptanalysis, etc.
- KFC's speed ranges from "not-very-fast" to "outrageously-slow".

Yet another Block Cipher?

KFC lies in-between both extremes:

- It comes with security proofs in the Luby-Rackoff model,
- and is practical (we mean, it can be implemented in practice).

More precisely, depending on the parameters choice:

- KFC is provably secure against d-limited adversaries for values of d ranging from 2 up to 70.
- This is enough to resist several statistical attacks.
- This includes Linear and Differential Cryptanalysis (taking hull/differentials effects in consideration), higher order differential cryptanalysis, etc.
- KFC's speed ranges from "not-very-fast" to "outrageously-slow".

Yet another Block Cipher?

KFC lies in-between both extremes:

- It comes with security proofs in the Luby-Rackoff model,
- and is practical (we mean, it can be implemented in practice).

More precisely, depending on the parameters choice:

- KFC is provably secure against d-limited adversaries for values of d ranging from 2 up to 70.
- This is enough to resist several statistical attacks.
- This includes Linear and Differential Cryptanalysis (taking hull/differentials effects in consideration), higher order differential cryptanalysis, etc.
- KFC's speed ranges from "not-very-fast" to "outrageously-slow".

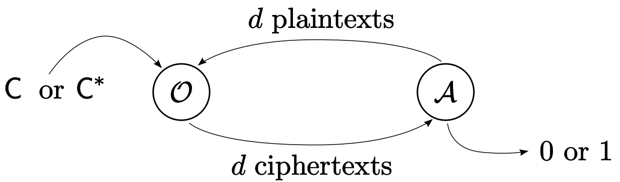
Outline

Outline

The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:

- computationally unbounded
- limited to d queries to an oracle $\mathcal{O}$ implementing either
  - a random instance C of the block cipher
  - or a random instance C^* of the perfect cipher
- the objective of $\mathcal{A}$ being to guess which is the case.



Advantage of $\mathcal{A}$

$$\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = |\Pr[\mathcal{A}(\mathsf{C}) = 0] - \Pr[\mathcal{A}(\mathsf{C}^*) = 0]|.$$

Computing $\mathrm{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

A block cipher C is secure if $\mathrm{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ is negligible for all $\mathcal{A}$'s.
Problem: computing this advantage is not a trivial task in general.

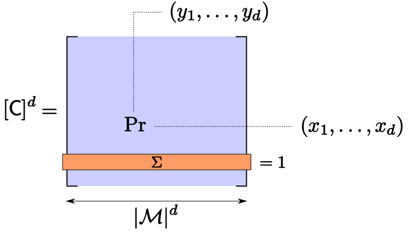
Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

A block cipher C is secure if $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ is negligible for all $\mathcal{A}$'s.

Problem: computing this advantage is not a trivial task in general.

Possible Solution: Use Vaudenay's Decorrelation Theory as a toolbox.

For a given cipher $\mathsf{C} : \mathcal{M} \to \mathcal{M}$, the distribution matrix at order d is:

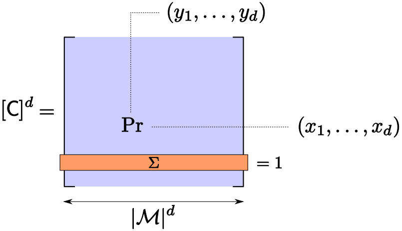


$$\Pr = \Pr_{\mathsf{C}}[\mathsf{C}(x_1) = y_1, \ldots, \mathsf{C}(x_d) = y_d]$$

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

A block cipher C is secure if $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ is negligible for all $\mathcal{A}$'s.
Problem: computing this advantage is not a trivial task in general.
Possible Solution: Use Vaudenay's Decorrelation Theory as a toolbox.

For a given cipher $\mathsf{C} : \mathcal{M} \to \mathcal{M}$, the distribution matrix at order d is:



$$\mathrm{Pr} = \Pr_{\mathsf{C}}[\mathsf{C}(x_1) = y_1, \ldots, \mathsf{C}(x_d) = y_d]$$

Link between $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ and $[\mathsf{C}]^d$

$$\max_{\mathcal{A}} \mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = \frac{1}{2} \| [\mathsf{C}]^d - [\mathsf{C}^*]^d \|.$$

⚠ $|\mathcal{M}|^d \approx 2^{128d}$ for a 128-bit block cipher! ⚠

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

There are at least two ways to deal with distribution matrix size:

- Use decorrelation modules as building blocks (drawback: may lead to "algebraic" constructions)
- Exploit the symmetries of the cipher (as done in [Baignères, Finiasz SAC06] and here)

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

There are at least two ways to deal with distribution matrix size:

- Use decorrelation modules as building blocks (drawback: may lead to "algebraic" constructions)
- Exploit the symmetries of the cipher (as done in [Baignères, Finiasz SAC06] and here)

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$ using the Decorrelation Theory

There are at least two ways to deal with distribution matrix size:

- Use decorrelation modules as building blocks (drawback: may lead to "algebraic" constructions)
- Exploit the symmetries of the cipher (as done in [Baignères, Finiasz SAC06] and here)

Link with "Practical" Attacks

The security proofs we provide not only induce resistance against LC and DC but against a wider class of attacks.

Most of statistical attacks (LC, DC, Higher order differentials, etc.) belong to the family of iterated attacks of order d.

For example:

- LC is an iterated attack of order 1, and
- DC is an iterated attack of order 2.

Provable security against d-limited adversaries $\Rightarrow$ Provable security against iterated attacks of order $\frac{d}{2}$ [Vaudenay JOC03].

$\rightsquigarrow$ one security proof leads to provable security against several attacks.

Link with “Practical” Attacks

The security proofs we provide not only induce resistance against LC and DC but against a wider class of attacks.
Most of statistical attacks (LC, DC, Higher order differentials, etc.) belong to the family of iterated attacks of order d.

For example:

- LC is an iterated attack of order 1, and
- DC is an iterated attack of order 2.

Provable security against d-limited adversaries $\Rightarrow$ Provable security against iterated attacks of order $\frac{d}{2}$ [Vaudenay JOC03].

$\rightsquigarrow$ one security proof leads to provable security against several attacks.

Link with "Practical" Attacks

The security proofs we provide not only induce resistance against LC and DC but against a wider class of attacks.
Most of statistical attacks (LC, DC, Higher order differentials, etc.) belong to the family of iterated attacks of order d.

For example:

- LC is an iterated attack of order 1, and
- DC is an iterated attack of order 2.

Provable security against d-limited adversaries $\Rightarrow$ Provable security against iterated attacks of order $\frac{d}{2}$ [Vaudenay JOC03].

$\rightsquigarrow$ one security proof leads to provable security against several attacks.

Link with “Practical” Attacks

The security proofs we provide not only induce resistance against LC and DC but against a wider class of attacks.
Most of statistical attacks (LC, DC, Higher order differentials, etc.) belong to the family of iterated attacks of order d.

For example:

- LC is an iterated attack of order 1, and
- DC is an iterated attack of order 2.

Provable security against d-limited adversaries $\Rightarrow$ Provable security against iterated attacks of order $\frac{d}{2}$ [Vaudenay JOC03].

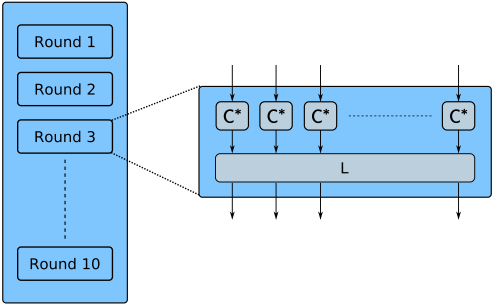
$\rightsquigarrow$ one security proof leads to provable security against several attacks.

Outline

1. Security Model
2. From the SPN of C to the Feistel scheme of KFC
3. Overview of Security Proofs on KFC

The block cipher C

C is a block cipher based on a Substitution-Permutation Network (SPN) [Baignères, Finiasz SAC06].



Each round is made of:

- A layer of substitution boxes ⇝ confusion
- A linear layer ⇝ diffusion

- The C*'s are mutually independent and perfectly random permutations on $\{0,1\}^8$
- The linear layer L is exactly the one used in AES

Security Results on the block cipher C

We showed that C is provably secure against 2-limited adversaries:

- Instead of directly computing the $2^{256} \times 2^{256}$ distribution matrix $[\mathrm{C}]^2$...
- we took advantage of the fact that symmetries of the cipher induce symmetries in the distribution matrix $[\mathrm{C}]^2$.
- $\rightsquigarrow$ computation on 625×625 matrices:

$$\max_{\mathcal{A}} \mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = 2^{-185.5}$$

Problem: we could not exhibit similar symmetries in $[\mathrm{C}]^d$ for $d > 2$.

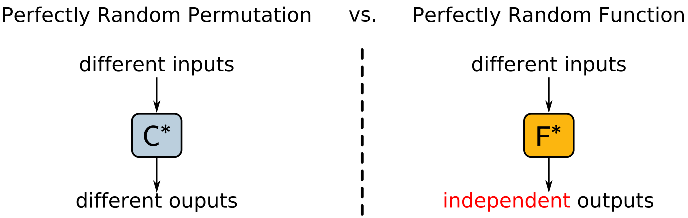
The Main Idea that lead us to the KFC Construction

Instead of computing the advantage of the best d-limited adversary, we will bound it by a function of the advantage of the best $(d-1)$-limited adversary.

The Main Idea that lead us to the KFC Construction

Instead of computing the advantage of the best d-limited adversary, we will bound it by a function of the advantage of the best $(d-1)$-limited adversary.

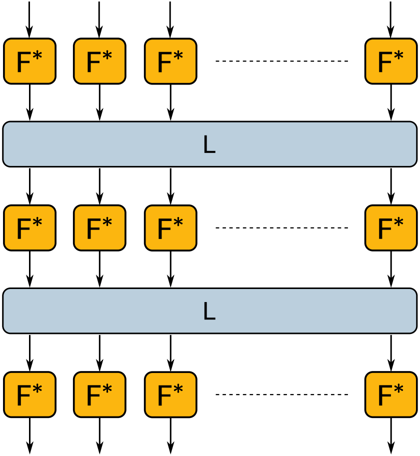
This approach is problematic with layers of random permutations:



- two correlated inputs of a random permutation always lead to two correlated outputs,
- two different inputs of a random function lead to two independent outputs.

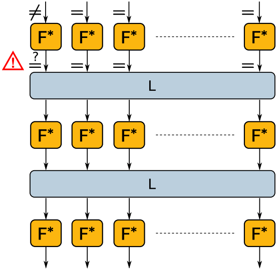
The Main Idea that lead us to the KFC Construction

Idea: Replace the layers of mutually independent and perfectly random permutations by layers of mutually independent and perfectly random functions.



Problem #1

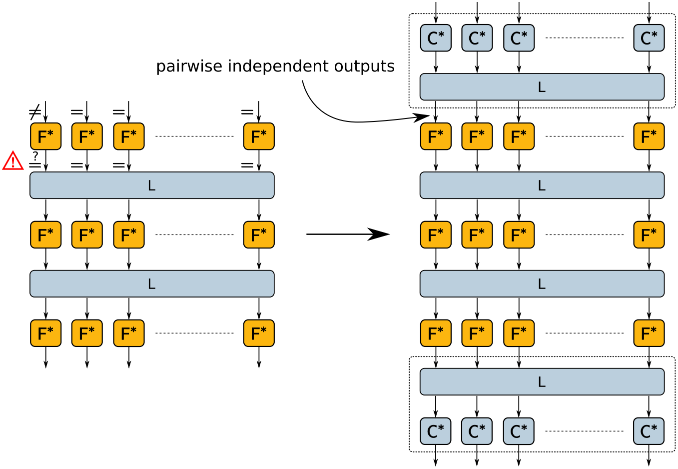
Problem: If two inputs are equal on all F* inputs but one $\rightsquigarrow$ non-negligible probability to obtain a full collision.



Problem #1

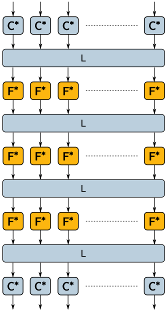
Problem: If two inputs are equal on all F* inputs but one $\rightsquigarrow$ non-negligible probability to obtain a full collision.

Solution: The Sandwich Technique



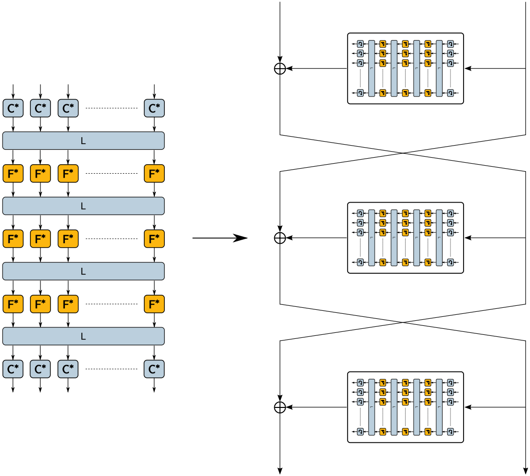
Problem #2

Problem: Our construction is not invertible.

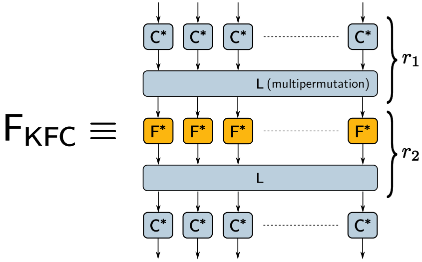


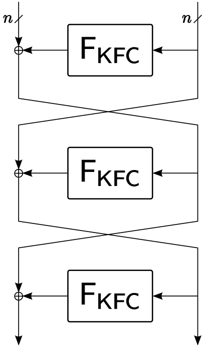
Problem #2

Problem: Our construction is not invertible.
Trivial Solution: Plug it in a Feistel Scheme



KFC: The Big Picture





$F_{KFC} : \{0,1\}^n \to \{0,1\}^n$ such that $\mathrm{Adv}_{\mathcal{A}}(F_{KFC}, F^*) \leq \epsilon$

$$\mathrm{Adv}_{\mathcal{A}}(\mathrm{KFC}, C^*) \leq 2\epsilon + \frac{d^2}{2^n}$$

Objective: Prove that ϵ is negligible.

Outline

Computing the Advantage of the Best 2-limited Adversary

We denote $\mathsf{F}_{\mathsf{KFC}} = \mathsf{S} \circ (\mathsf{L} \circ \mathsf{F})^{r_2} \circ (\mathsf{L} \circ \mathsf{S})^{r_1}$ so that

$[\mathsf{F}_{\mathsf{KFC}}]^2 = [\mathsf{S} \circ (\mathsf{L} \circ \mathsf{F})^{r_2} \circ (\mathsf{L} \circ \mathsf{S})^{r_1}]^2 = ([\mathsf{S}]^2 \times [\mathsf{L}]^2)^{r_1} \times ([\mathsf{F}]^2 \times [\mathsf{L}]^2)^{r_2} \times [\mathsf{S}]^2.$

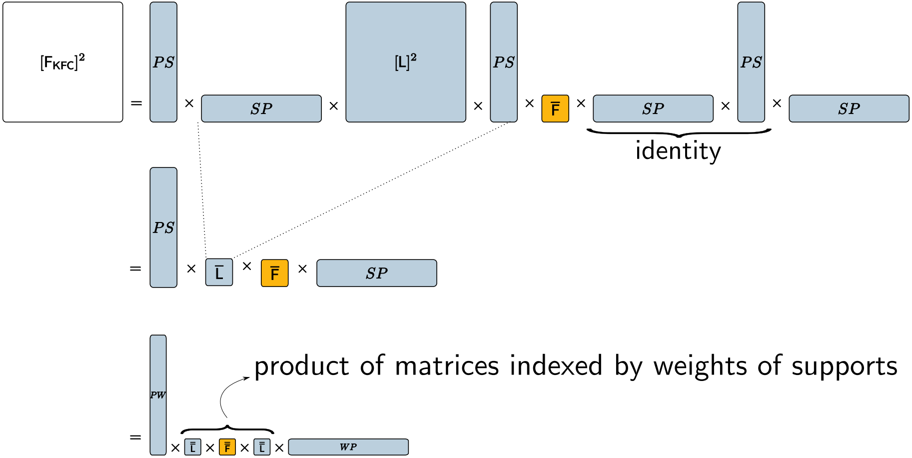
⚠These are $2^{2n} \times 2^{2n}$ matrices... The shape of the confusion layers allows to write

$$\underbrace{[\mathsf{F}]^2}_{\#\text{ pair of texts}} = PS \times \underbrace{\bar{\mathsf{F}}}_{\#\text{ support of pair}} \times SP$$

$$[\mathsf{S}]^2 = PS \times SP$$

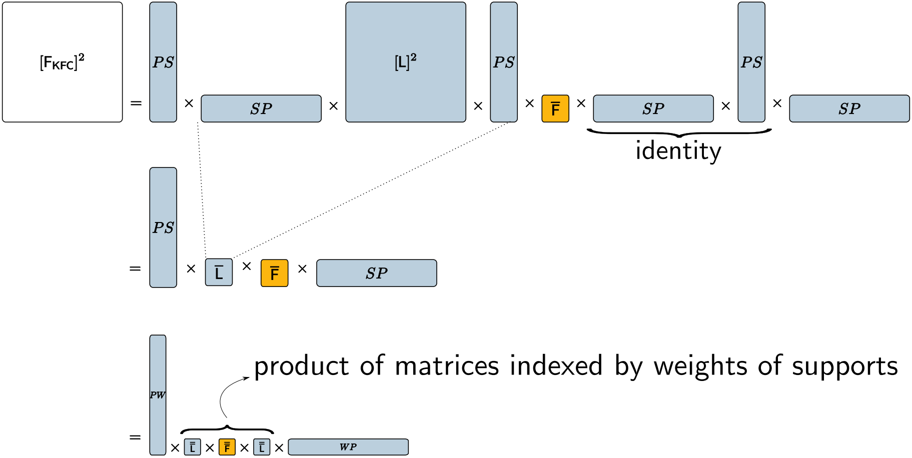
Computing the Advantage of the Best 2-limited Adversary

Putting things together, we have (with $r_1 = r_2 = 1$):



Computing the Advantage of the Best 2-limited Adversary

Putting things together, we have (with $r_1 = r_2 = 1$):



$$[\mathsf{F}_{\mathsf{KFC}}]^2 = PW \times (\overline{\overline{\mathsf{L}}})^{r_1} \times (\overline{\overline{\mathsf{F}}} \times \overline{\overline{\mathsf{L}}})^{r_2} \times WP$$

Computing the Advantage of the Best 2-limited Adversary (at last)

In the end...

$$\| \underbrace{[\mathsf{F}_{\mathsf{KFC}}]^2 - [\mathsf{F}^*]^2}_{2^{256} \times 2^{256} \text{ matrices}} \| = \| \underbrace{(\overline{\overline{\mathsf{L}}})^{r_1} \times (\overline{\overline{\mathsf{F}}} \times \overline{\overline{\mathsf{L}}})^{r_2} - U}_{9 \times 9 \text{ matrices}} \|$$

so that one can easily compute

$$\boxed{\mathsf{Adv}_{\mathcal{A}}(\mathsf{F}_{\mathsf{KFC}}) = \frac{1}{2} \|(\overline{\overline{\mathsf{L}}})^{r_1} \times (\overline{\overline{\mathsf{F}}} \times \overline{\overline{\mathsf{L}}})^{r_2} - U\|}$$

Computing the Advantage of the Best 2-limited Adversary (at last)

In the end...

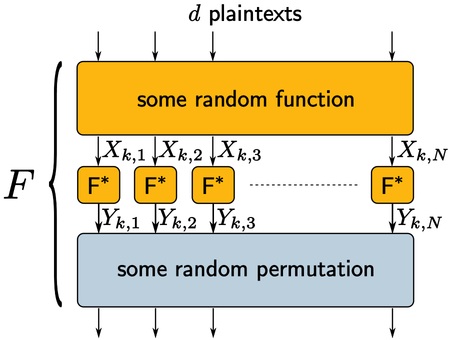
$$\| \underbrace{[\mathsf{F}_{\mathsf{KFC}}]^2 - [\mathsf{F}^*]^2}_{2^{256}\times 2^{256}\ \text{matrices}} \| = \| \underbrace{(\bar{\bar{\mathsf{L}}})^{r_1} \times (\bar{\bar{\mathsf{F}}} \times \bar{\bar{\mathsf{L}}})^{r_2} - U}_{9\times 9\ \text{matrices}} \|$$

so that one can easily compute

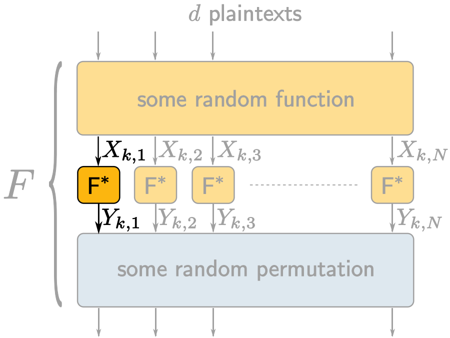
$$\boxed{\mathsf{Adv}_{\mathcal{A}}(\mathsf{F}_{\mathsf{KFC}}) = \frac{1}{2}\|(\bar{\bar{\mathsf{L}}})^{r_1} \times (\bar{\bar{\mathsf{F}}} \times \bar{\bar{\mathsf{L}}})^{r_2} - U\|}$$

	$N=8$ and $q=2^8$				$N=8$ and $q=2^{16}$				$N=16$ and $q=2^8$			
$r_1 \backslash r_2$	0	1	10	100	0	1	10	100	0	1	10	100
0	1	2^{-5}	2^{-8}	2^{-8}	1	2^{-13}	2^{-16}	2^{-16}	1	2^{-4}	2^{-8}	2^{-8}
1	2^{-5}	2^{-50}	2^{-52}	2^{-49}	2^{-13}	2^{-114}	2^{-116}	2^{-113}	2^{-4}	2^{-95}	2^{-104}	2^{-103}
2	2^{-46}	2^{-53}	2^{-52}	2^{-49}	2^{-110}	2^{-117}	2^{-116}	2^{-113}	2^{-87}	2^{-104}	2^{-104}	2^{-103}
3	2^{-62}	2^{-53}	2^{-52}	2^{-49}	2^{-128}	2^{-117}	2^{-116}	2^{-113}	2^{-120}	2^{-104}	2^{-104}	2^{-103}

Bounding the Advantage of the Best d-limited Adversary ($d > 2$)



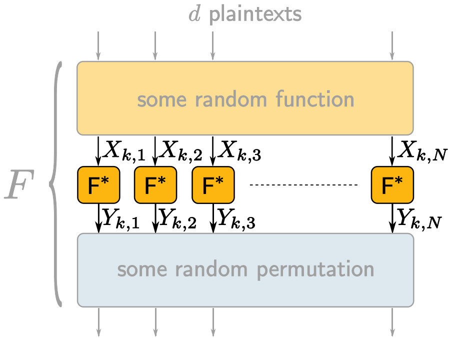
Bounding the Advantage of the Best d-limited Adversary ($d > 2$)



Assume one of the d box inputs is different from all the others

Then one of the d box outputs is independent from all the others

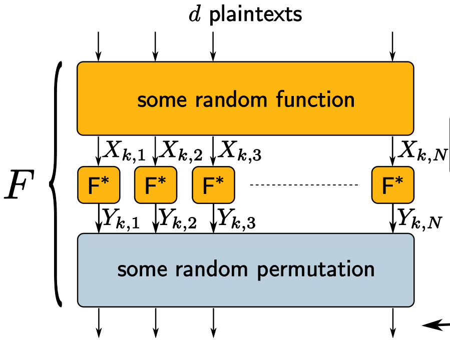
Bounding the Advantage of the Best d-limited Adversary ($d > 2$)



Assume one of the d ~~box~~ inputs is different from all the others on all boxes

Then one of the d ~~box~~ outputs is independent from all the others

Bounding the Advantage of the Best d-limited Adversary ($d > 2$)



event α

Assume one of the d ~~box~~ inputs is different from all the others on all boxes

Then one of the d ~~box~~ outputs is independent from all the others

$$\mathsf{Adv}_{\mathcal{A}_d}(F, \mathsf{F}^*) \leq \mathsf{Adv}_{\mathcal{A}_{d-1}}(F, \mathsf{F}^*) + \Pr[\overline{\alpha}]$$

Bounding the Advantage of the Best d-limited Adversary ($d > 2$)

Considering several α events on t successive rounds, one can bound the probability that none of them occurs:

$$\Pr[\overline{\alpha}_1, \ldots, \overline{\alpha}_t] \leq \left(1 - \left(1 - \frac{d-1}{q}\right)^N\right)^t$$

Bounding the Advantage of the Best d-limited Adversary ($d > 2$)

Considering several α events on t successive rounds, one can bound the probability that none of them occurs:

$$\Pr[\overline{\alpha}_1, \ldots, \overline{\alpha}_t] \leq \left(1 - \left(1 - \frac{d-1}{q}\right)^N\right)^t$$

Theorem

Assume $\mathsf{Adv}_{\mathcal{A}_2}(\mathsf{F}_{\mathsf{KFC}}, \mathsf{F}^*) \leq \epsilon$. For any d and set of integers $\{t_3, \ldots, t_d\}$ s.t. $\sum_{i=3}^{d} t_i \leq r_2$, we have

$$\mathsf{Adv}_{\mathcal{A}_d}(\mathsf{F}_{\mathsf{KFC}}, \mathsf{F}^*) \leq \epsilon + \sum_{i=3}^{d} \left(1 - \left(1 - \frac{i-1}{q}\right)^N\right)^{t_i}$$

KFC “in practice”

Regular KFC: $N = 8$, $q = 2^8$, $r_1 = 3$, $r_2 = 100$

- Provable security against 8-limited adaptive adversaries
- Thus against iterated attacks of order 4
- (Estimated) Speed of 15-25 Mbits/s

KFC “in practice”

Regular KFC: $N = 8$, $q = 2^8$, $r_1 = 3$, $r_2 = 100$

- Provable security against 8-limited adaptive adversaries
- Thus against iterated attacks of order 4
- (Estimated) Speed of 15-25 Mbits/s

Extra Crispy KFC: $N = 8$, $q = 2^{16}$, $r_1 = 3$, $r_2 = 1000$

- Provable security against 70-limited adaptive adversaries
- Thus against iterated attacks of order 35
- (Estimated) Speed $<$ ($\ll$?) 1 Mbit/s
- 4 GB of memory are required

Conclusion and Further Improvements

- KFC is the first “practical” block cipher with security proofs up to a large order.
- Bounds can be improved: the same security level can be achieved with fewer rounds (hint: improve the bound on α).
- It is possible to weaken the assumptions on the round functions of the Feistel scheme and obtain the same security level (see [Lucks FSE96] or [Maurer, Oswald, Pietrzak, Sjödin Eurocrypt06]).
- Use a faster diffusion layer (`ShiftRows+Mixcolumns`): increase r_1 but improve global speed.

Conclusion and Further Improvements

- KFC is the first "practical" block cipher with security proofs up to a large order.

- Bounds can be improved: the same security level can be achieved with fewer rounds (hint: improve the bound on α).
- It is possible to weaken the assumptions on the round functions of the Feistel scheme and obtain the same security level (see [Lucks FSE96] or [Maurer, Oswald, Pietrzak, Sjödin Eurocrypt06]).
- Use a faster diffusion layer (`ShiftRows+Mixcolumns`): increase r_1 but improve global speed.