

# ESF 算法的不可能差分密码分析

陈玉磊 卫宏儒

(北京科技大学数理学院 北京 100083)

**摘要** 分析研究了分组密码算法 ESF 抵抗不可能差分的能力,使用 8 轮不可能差分路径,给出了相关攻击结果。基于一条 8 轮的不可能差分路径,根据轮密钥之间的关系,通过改变原有轮数扩展和密钥猜测的顺序,攻击了 11 轮的 ESF,改善了关于 11 轮的 ESF 的不可能差分攻击的结果。计算结果表明:攻击 11 轮的 ESF 所需要的数据复杂度为  $O(2^{53})$ ,时间复杂度为  $O(2^{32})$ ,同时也说明了 11 轮的 ESF 对不可能差分是不免疫的。

**关键词** 分组密码,不可能差分,ESF,轮密钥

中图分类号 TN918.1 文献标识码 A DOI 10.11896/j.issn.1002-137X.2016.8.018

## Impossible Differential Cryptanalysis of ESF

CHEN Yu-lei WEI Hong-ru

(School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract** This paper studied and analyzed the ability of the block cipher algorithm ESF resisting the impossible difference, a 8-round impossible differential route was used and the related results were given. On the basis of the 8-round impossible differential route, according to the relationship of the round keys, by changing the original order of round number extension and key guessing, the paper attacked 11-round ESF, improving the result of the 11-round ESF impossible differential. Computing result shows that the attack of 11-round ESF needs  $O(2^{53})$  chosen plaintext operations and  $O(2^{32})$  encrypting computations. At the same time, it also shows that the 11-round ESF is not immune to the impossible difference.

**Keywords** Block cipher, Impossible differential, ESF, Round key

## 1 引言

近年来,随着电子科技的迅猛发展,无线传感网络、射频识别技术(RFID)等微型计算设备的应用日益广泛,轻量级分组密码算法逐渐成为人们关注的热点。在特定的情况下,受微型计算设备资源和计算能力的限制,传统的分组密码算法如 AES、DES 等虽然执行效率得到大大改进,但是没有足够的安全保证。相对而言,轻量级分组密码对资源的要求不高,占用的存储空间较少,处理的数据规模不大,对安全性要求也比较低。为了满足实际需求,大量轻量级分组密码算法逐渐被人们研制出来,如 MIBS, LBlock, CLEFIA, SNAKE, TWIS 等。

刘宣等人在研究了吴文玲学者设计的算法 LBlock<sup>[1]</sup>之后,发现要想使置换层在较少的轮数中扩散得更快,可以将其按位进行置换,这样密码抵抗攻击的安全性就会得到提升。同时, P 层置换的形式借鉴了算法 PRESNT<sup>[2]</sup>, 根据数学公式对 P 层进行优化,他们又提出了“八阵图算法”ESF, 对 P 层进行了改善,这些操作使 P 层达到了相当好的预期效果。ESF 算法具有典型的 Feistel 结构,轮函数  $F$  基于 SPN 代换置换

网络结构,它的迭代次数是 32 轮。ESF 的密钥编排及数据处理都有严格的准则,而且加解密速度快,结构完全一致,可实现性非常强,不同的软硬件平台都可以应用。

不可能差分密码分析是利用概率为 0 的差分(不可能差分)排除那些导致概率为 0 的差分出现的候选密钥,这个概念是由 Biham 和 Knudsen 两人提出的<sup>[3,4]</sup>,它是差分密码分析的一个变种。Knudsen 研究发现,在分析 DEAL 算法的安全性时,如果具有 Feistel 结构的密码的轮函数是双射,则算法存在“天然 5 轮不可能差分”,这样就能够严重威胁到 6 轮密码的安全性。不可能差分的概念<sup>[4,5]</sup>是由 Biham 等人在研究 Skipjack 算法的安全性时提出来的<sup>[6,7]</sup>。后来, Kim 等<sup>[8]</sup>对不可能差分的结果进行了总结,并且提出了新的方法,他们指出这些一定存在的不可能差分仅仅与算法结构有关,而与具体的非线性变换无关,这就是不可能截断差分<sup>[6]</sup>。

## 2 ESF 算法

### 2.1 符号及术语说明

$M$ : 64bits 明文;

$C$ : 64bits 密文;

到稿日期:2015-07-05 返修日期:2015-09-10 本文受 2013 年国家自然科学基金(61272476),内蒙古自治区科技创新引导奖励资金项目(2012)资助。

陈玉磊(1989—),女,硕士生,主要研究方向为密码学;卫宏儒(1963—),男,副教授,主要研究方向为数学、信息安全与密码学、物联网关键技术, E-mail: weihr@ustb.edu.cn(通信作者)。

$K$ : 80bits 主密钥;  
 $K_r$ : 32bits 轮密钥;  
 $F$ : 轮函数;  
 $S$ :  $4 \times 4$  S盒;  
 $\oplus$ : 按位异或;  
 $P$ : 扩散层;  
 $\lll 7$ : 循环左移 7 位;  
 $\parallel$ : 二进制字符联接;  
 $[i]_2$ : 轮常数  $i$  的二进制表示。

## 2.2 ESF 算法简介<sup>[9]</sup>

轻量级分组密码 ESF 包括加密和解密两部分, 它的数据处理流程是基于 SPN 轮函数<sup>[10]</sup> 和 2 分支的 Feistel 变种结构<sup>[11]</sup>, 它的密钥长度是 80bits, 分组长度是 64bits。ESF 的轮函数包括轮密钥加、线性置换层和非线性混淆层。

设  $M, C \in \{0, 1\}^{64}$  是相应的明密文。32 轮的 ESF 每一轮都含有异或操作、置换和非扩散。扩散层由 8 个 S 盒构成, 并且是 4 进 4 出的。图 1 示出 ESF 算法的加密流程。

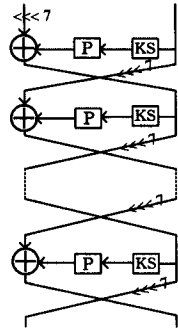


图 1 ESF 算法的加密流程

设  $L_i, R_i$  是第  $i$  轮的输入, 它们分别为 32bits, 则第  $i$  轮的迭代公式为:  $L_{i+1}=R_i, R_{i+1}=(L_i \lll 7) \oplus F(R_i, K_i)$ , 其中  $K_i$  是轮密钥。在经过 31 轮变换之后, 输出结果为  $C=L_{32} \parallel R_{32}$ 。

密钥加: 先把每轮输入的右 32 bits 循环左移 7 位, 然后与轮函数异或。

非线性混淆层: 算法 ESF 的轮函数  $F$  具有典型的 SPN 结构, 它的非线性混淆层是非线性变换, 非线性变换  $S$  函数由 8 个  $4 \times 4$  并行的 S 盒构成:

$$S=(S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8)$$

$P$  置换层: 该层为线性变换, 其构成如下:

$$(1) P: \{0, 1\}^{32} \rightarrow \{0, 1\}^{32}$$

$$b=b_7 \parallel b_6 \parallel b_5 \parallel b_4 \parallel b_3 \parallel b_2 \parallel b_1 \parallel b_0 \rightarrow c=c_7 \parallel c_6 \parallel c_5 \parallel c_4 \parallel$$

$$c_3 \parallel c_2 \parallel c_1 \parallel c_0$$

$$(2) \text{当 } 0 \leq i < 8 \text{ 时}$$

$$b_{4i} \parallel b_{4i+1} \parallel b_{4i+2} \parallel b_{4i+3} \rightarrow c_i \parallel c_{i+8} \parallel c_{i+16} \parallel c_{i+24}$$

解密: 与加密具有相同的结构, 可参考上文, 不再叙述。

## 2.3 ESF 的密钥编排

密钥编排算法是把 80 bits 的主密钥经过一系列的运算后产生的 32bits 的轮密钥。将主密钥取为  $K=k_{79}k_{78} \dots k_0$ , 并存入密钥寄存器, 最左端存入  $k_{79}$ , 最右端存入  $k_0$ 。最左端的 32 bits 密钥取作轮密钥  $k_1$ , 然后更新每轮密钥。对  $i=1, 2, 3, \dots, 31$ , 如下更新密钥寄存器  $K$ :

$$(1) K \lll 13;$$

$$(2) \text{密钥寄存器 } [k_{79} k_{78} k_{77} k_{76}] = S_0 [k_{79} k_{78} k_{77} k_{76}], [k_{75} k_{74} k_{73} k_{72}] = S_0 [k_{75} k_{74} k_{73} k_{72}];$$

$$(3) [k_{47} k_{46} k_{45} k_{44} k_{43}] = [k_{47} k_{46} k_{45} k_{44} k_{43}] \oplus \text{轮常数 } i;$$

(4) 输出轮密钥  $K_{i+1}$  是当前寄存器  $K$  最左端的 32bits。

## 3 ESF 算法的不可差分分析

### 3.1 算法 ESF 的 8 轮不可能差分路径

ESF 的 8 轮不可能差分路径已经在文献<sup>[9]</sup> 中由刘宜等人给出, 本文利用已有的 8 轮不可能差分区分器对轻量级分组密码算法 ESF 进行 11 轮不可能差分攻击。8 轮不可能差分路径为:  $(000a \mid bcd0 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000) \rightarrow (0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000 \mid 0000)$ , 其中  $a, b, c, d, e, f, g$  是非零值, 且  $h=1$ 。图 2 示出 ESF 的 8 轮不可能差分路径, 它的推导过程可以参考文献<sup>[9]</sup>。

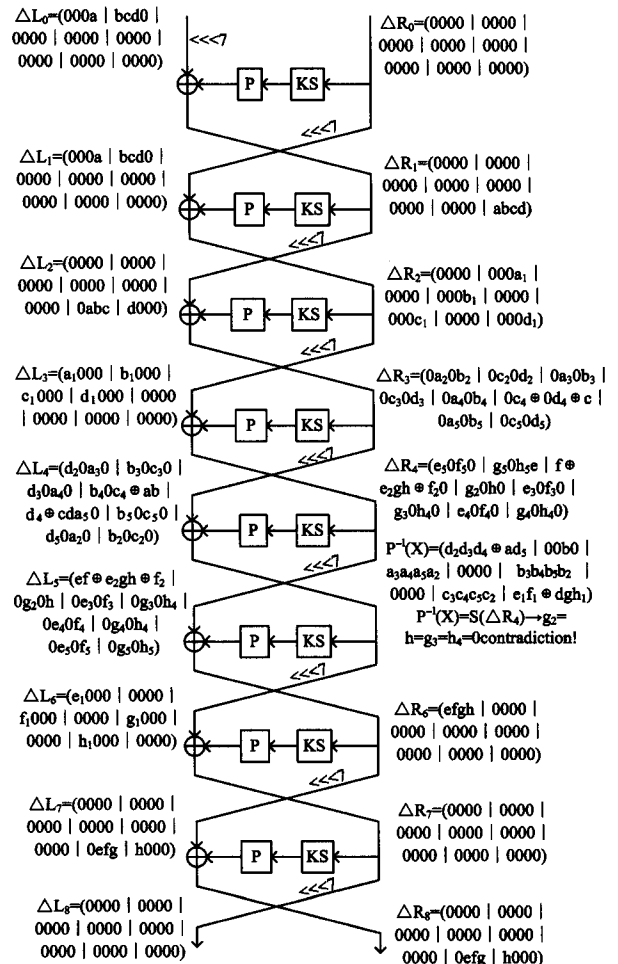


图 2 ESF 算法的 8 轮不可能差分路径

### 3.2 11 轮 ESF 算法的不可差分分析

利用已有的 8 轮不可能差分路径, 对 11 轮的 ESF 进行不可能差分攻击。其基本思想是: 在 8 轮不可能差分路径后面加 2 轮, 前面加 1 轮, 对 11 轮的 ESF 算法进行不可能差分<sup>[10]</sup> 分析, 通过不可能差分把错误的密钥淘汰。11 轮不可能差分攻击的过程如图 3 所示。

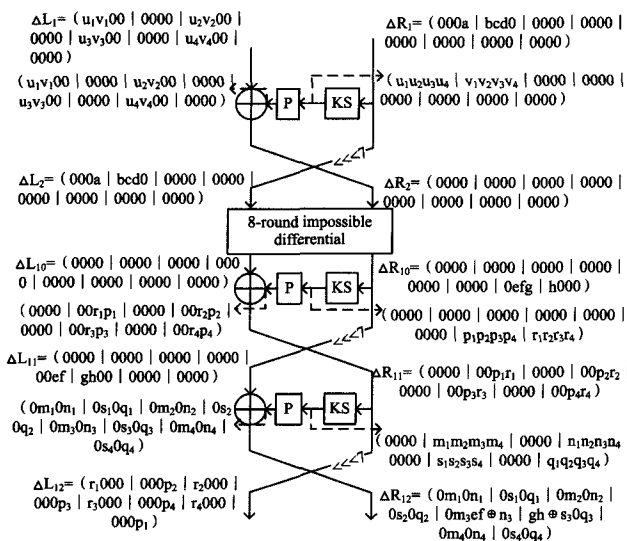


图3 11轮ESF算法的不可能差分路径

(1)选择如下明文组:  $L_0 = (x_1, x_2, a_3, a_4 | a_5, a_6, a_7, a_8 | x_9, x_{10}, a_{11}, a_{12} | a_{13}, a_{14}, a_{15}, a_{16} | x_{17}, x_{18}, a_{19}, a_{20} | a_{21}, a_{22}, a_{23}, a_{24} | x_{25}, x_{26}, a_{27}, a_{28} | a_{29}, a_{30}, a_{31}, a_{32})$ ,  $R_0 = (b_1, b_2, b_3, y_4 | y_5, y_6, y_7, b_8 | b_9, b_{10}, b_{11}, b_{12} | b_{13}, b_{14}, b_{15}, b_{16} | b_{17}, b_{18}, b_{19}, b_{20} | b_{21}, b_{22}, b_{23}, b_{24} | b_{25}, b_{26}, b_{27}, b_{28} | b_{29}, b_{30}, b_{31}, b_{32})$ 。其中  $a_i$  为任意常数,  $x_i (i=1, 2, 9, 10, 17, 18, 25, 26)$  和  $y_i (i=4, 5, 6, 7)$  取所有值。符合上述这些条件的明文一共有  $2^{12}$  个, 那么这些明文可以组成的明文对一共有  $2^{12} \times 2^{12} \times 2^{-1} = 2^{23}$  对。在这些明文对中选择  $2^N$  组进行 11 轮加密, 于是可以得到  $2^{N+23}$  个明文对。

(2)选择满足如下差分的密文对:  $\Delta L_{11} = (r_1, 000 | 000, p_2 | r_2, 000 | 000, p_3 | r_3, 000 | 000, p_4 | r_4, 000 | 000, p_1)$ ,  $\Delta R_{11} = (0, m_1, 0, n_1 | 0, s_1, 0, q_1 | 0, m_2, 0, n_2 | 0, s_2, 0, q_2 | 0, m_3, 0, n_3 | 0, s_3, 0, q_3 | 0, m_4, 0, n_4 | 0, s_4, 0, q_4)$ 。其中的  $e, f, g, h$  和  $r_i, p_i, m_i, n_i, s_i, q_i (i=1, 2, 3, 4)$  为非零值, 共有  $2^{8+18} = 2^{26}$  个这样的密文对, 满足上述条件的概率是  $2^{26} \times 2^{-64} = 2^{-38}$ 。所有这样的密文对共有大约  $2^{N+23} \times 2^{-38} = 2^{N-15}$  对。

(3)先猜测轮密钥  $k_{11}$  的 8 比特值  $k_{11,1}, k_{11,8}, k_{11,9}, k_{11,16}, k_{11,17}, k_{11,24}, k_{11,25}, k_{11,32}$ 。对上一步留下的任意一对密文, 通过计算选择满足  $s(L_{11,i} \oplus K_{11,i}) \oplus s(L'_{11,i} \oplus K_{11,i}) = R_{11,i} \oplus R'_{11,i}$  的对, 其中  $i=1, 8, 9, 16, 17, 24, 25, 32$ 。满足上述条件的概率是  $2^{-8}$ 。故共有约  $2^{N-15} \times 2^{-8} = 2^{N-23}$  对。

(4)猜测轮密钥  $k_{10}$  的第 26, 27, 28, 29 这 4 比特值  $k_{10,26}, k_{10,27}, k_{10,28}, k_{10,29}$ , 对第(3)步留下来的任意一对密文 ( $L_{10}, R_{10}$ ) 与 ( $L'_{10}, R'_{10}$ ), 计算  $s_7(L_{10,i} \oplus k_{10,i}) \oplus s_7(L'_{10,i} \oplus k_{10,i}), i=26, 27, 28, 29$ , 选择差分满足如下条件的数据对:  $s_7(L_{10,26} \oplus k_{10,26}) \oplus s_7(L'_{10,26} \oplus k_{10,26}) = R_{10,26} \oplus R'_{10,26}, s_7(L_{10,27} \oplus k_{10,27}) \oplus s_7(L'_{10,27} \oplus k_{10,27}) = R_{10,27} \oplus R'_{10,27}, s_7(L_{10,28} \oplus k_{10,28}) \oplus s_7(L'_{10,28} \oplus k_{10,28}) = R_{10,28} \oplus R'_{10,28}, s_8(L_{10,29} \oplus k_{10,29}) \oplus s_8(L'_{10,29} \oplus k_{10,29}) = R_{10,29} \oplus R'_{10,29}$ 。此步的概率为  $p=2^{-4}$ 。因此剩余对数共有  $2^{N-23} \times 2^{-4} = 2^{N-27}$  对。

(5)最后来猜测  $k_1$  的第 4, 5, 6, 7 这 4 比特值, 对第(4)步留下来的每一对 ( $L_1, R_1$ ) 与 ( $L'_1, R'_1$ ), 其中  $L_1 = (y_1, y_2, a_3, a_4 | a_5, a_6, a_7, a_8 | y_9, y_{10}, a_{11}, a_{12} | a_{13}, a_{14}, a_{15}, a_{16} | y_{17}, y_{18}, a_{19}, a_{20} | a_{21}, a_{22}, a_{23}, a_{24} | y_{25}, y_{26}, a_{27}, a_{28} | a_{29}, a_{30}, a_{31}, a_{32})$ ,  $R_1 = (r_1, r_2, r_3, z_4 | z_5, z_6, z_7, r_8 | r_9, r_{10}, r_{11}, r_{12} | r_{13}, r_{14}, r_{15}, r_{16} | r_{17}, r_{18}, r_{19}, r_{20} | r_{21}, r_{22}, r_{23}, r_{24} | r_{25}, r_{26}, r_{27}, r_{28} | r_{29}, r_{30}, r_{31}, r_{32})$ ,

$L'_1 = (y'_1, y'_2, a_3, a_4 | a_5, a_6, a_7, a_8 | y'_9, y'_{10}, a_{11}, a_{12} | a_{13}, a_{14}, a_{15}, a_{16} | y'_{17}, y'_{18}, a_{19}, a_{20} | a_{21}, a_{22}, a_{23}, a_{24} | y'_{25}, y'_{26}, a_{27}, a_{28} | a_{29}, a_{30}, a_{31}, a_{32})$ ,  $R'_1 = (r'_1, r'_2, r'_3, z'_4 | z'_5, z'_6, z'_7, r'_8 | r'_9, r'_{10}, r'_{11}, r'_{12} | r'_{13}, r'_{14}, r'_{15}, r'_{16} | r'_{17}, r'_{18}, r'_{19}, r'_{20} | r'_{21}, r'_{22}, r'_{23}, r'_{24} | r'_{25}, r'_{26}, r'_{27}, r'_{28} | r'_{29}, r'_{30}, r'_{31}, r'_{32})$ 。计算  $s_1(z_4 \oplus k_{1,4}) \oplus s_1(z'_4 \oplus k_{1,4}) = R_{1,4} \oplus R'_{1,4}, s_2(z_5 \oplus k_{1,5}) \oplus s_2(z'_5 \oplus k_{1,5}) = R_{1,5} \oplus R'_{1,5}, s_2(z_6 \oplus k_{1,6}) \oplus s_2(z'_6 \oplus k_{1,6}) = R_{1,6} \oplus R'_{1,6}, s_2(z_7 \oplus k_{1,7}) \oplus s_2(z'_7 \oplus k_{1,7}) = R_{1,7} \oplus R'_{1,7}$ 。选择差分形式满足上述条件的所有对, 此步概率为  $(2^{-1})^4 \times (2^4 - 1) / 2^4 = 2^{-4}$ 。所以, 满足上述条件的数据对一共有  $2^{N-27} \times 2^{-4} = 2^{N-31}$  对。

(6)对上一步所剩的每一对, 判断  $\Delta R_2$  是否等于  $(0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000)$ 。

如果这一密钥猜测满足不可能差分路线, 就会使第(6)步成立, 这就表明所猜测的密钥肯定是错误密钥, 将其舍去。

在分析了  $2^{N-31}$  对明文对后, 还剩余  $2^{64} \times (1 - 2^{-4})^{N-31}$  个错误密钥, 如果将错误密钥全部淘汰, 有  $2^{64} \times (1 - 2^{-4})^{2^{N-31}} < 1, N \approx 41$ 。唯一正确密钥将会输出。

### 3.3 复杂度分析

对 11 轮 ESF 算法攻击的时间和数据复杂度的计算如下:

第(3)步需要  $2 \times 2^8 \times 2^{N-15} = 2^{N-6}$  次一轮加密; 第(4)步需要  $2 \times 2^8 \times 2^4 \times 2^{N-23} \times 2/8 = 2^{N-12}$  次一轮加密; 第(5)步需要  $2 \times 2^8 \times 2^4 \times 2^8 \times 2^{N-27} \times 4/8 = 2^{N-7}$  次一轮加密。所以该攻击总的时间复杂度为  $(2^{N-16} + 2^{N-12} + 2^{N-7})/11 \approx 2^{32}$ 。

若按上述分析, 总的时间复杂度为  $O(2^{32})$ , 明文量为  $2^{12} \times 2^N = 2^{53}$ 。

**结束语** 本文对 ESF 算法进行了不可能差分分析。在已有 8 轮不可能差分路径的基础上, 向后增加 2 轮, 向前增加 1 轮, 由此构成了 11 轮不可能差分攻击, 获得了较好的结果。其与已有分析方法相比较的结果如表 1 所列。

表1 ESF算法分析结果总结

Attack Type	Round	Data	Time	Source
Impossible DC	11	$2^{64}$	$2^{75.5}$	Ref. [9]
Impossible DC	11	$2^{53}$	$2^{32}$	This paper

### 参考文献

- [1] Wu W, Zhang L. LBlock: a lightweight block cipher[C]// Applied Cryptography and Network Security. Springer Berlin Heidelberg, 2011:327-344
- [2] Bogdanov A, Knudsen L R, Leander G, et al. PRESENT: An ultra-lightweight block cipher[M]. Springer Berlin Heidelberg, 2007
- [3] Dunkelman O. Techniques for cryptanalysis of block ciphers [D]. Technion-Israel Institute of Technology, Faculty of Computer Science, 2006
- [4] Biryukov A, Wagner D. Slide attacks[M]// Fast Software Encryption. Springer Berlin Heidelberg, 1999:245-259
- [5] Biham E, Dunkelman O, Keller N. Improved slide attacks [M]// Fast Software Encryption. Springer Berlin Heidelberg, 2007: 153-166
- [6] 吴文玲, 冯登国, 张文涛. 分组密码的设计与分析[M]. 北京: 清华大学出版社, 2009:68-72

(下转第 99 页)

表7 SMCTBRK 文本相似性度量结果

算法	实际相似的 文本数量	实际不相似的 文本数量
检测结果为相似的文本数量	25	0
检测结果为不相似的文本数量	5	30

#### 4.4.2 实验结果及分析

根据 4.3 节的相关评价标准,分别计算 3 种算法的 3 个主要性能指标,结果如表 8 所列。

表8 3种算法的性能比较

算法	准确率(%)	召回率(%)	平衡系数
Shingling	75.68	93.33	0.84
Simhash	86.96	66.67	0.75
SMCTBRK	100	83.33	0.91

表 8 比较了 3 种文本相似性度量方法的准确率、召回率以及平衡系数,可以看出,SMCTBRK 在召回率方面略低于 Shingling 算法,但在准确率和平衡系数方面高于 Shingling 算法和 Simhash 算法;而且从图 3—图 5 可以看出,SMCTBRK 方法的相关文本之相似系数与无关文本的相似系数之可区分性高于其它两种参与对比的算法。另外,虽然 3 种算法的时间复杂度相同,均为  $O(n^2)$ ,但是由于 SMCTBRK 方法在计算时所依赖的关键词数量比 Shingling 算法、Simhash 算法要少,因此在应用时,SMCTBRK 方法实际所消耗的时间要比另外两种算法少。所以,从这个角度而言,SMCTBRK 方法的效率要高于其它两种算法的。

**结束语** 本文提出了 SMCTBRK 方法,主要工作来自 3 个方面:1)优化了原有的关键词权重度量方法,在权重度量过程中考虑了关键词的范围属性,使关键词的度量在篇幅和文章结构差异性较大的文本中更加公平,并能显著减少在文本度量过程中所需要依赖的关键词数量,同时不影响最后的度量效果;2)通过更合理地计算关键词的权重,通过抽取更能代表文本主要内容的关键词,并根据它们构建文档的密文索引,有效地防止了数据信息泄漏;3)通过计算文档之密文索引向量之间的余弦夹角而不是具体的关键词来度量文档之间的相似性,起到了保护信息隐私的作用。在真实文档上进行了实验,结果表明 SMCTBRK 方法在准确率和平衡系数方面高于 Shingling 算法和 Simhash 算法,但在召回率方面略低于 Shingling 算法,需要进一步改进。

### 参考文献

[1] Wang C, Cao N, Li J, et al. Secure ranked keyword search over

(上接第 91 页)

[7] Biham E, Biryukov A, Shamir A. Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials [M]// Advances in Cryptology—Eurocrypt'99. Springer Berlin Heidelberg, 1999:12-23

[8] Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems[C]//Proc of CRYPTO'90. 1991:2-21

[9] Liu Xuan, Liu Feng, Meng Shuai. Impossible differential cryptanalysis of lightweight block cipher ESF[J]. Computer Engineering & Science, 2013, 35(9): 89-95(in Chinese)

刘宣, 刘枫, 孟帅. 轻量级分组密码算法 ESF 的不可能差分分析

encrypted cloud data[C]//Proceedings of ICDCS. Genova, Italy, 2010:253-262

[2] Sebastiani F. Machine learning in automated text categorization, acmcs[J]. ACM Computing Surveys, 2002, 34(1): 1-47

[3] Hemalatha S, Raja K, Arasu T. Duplicate Detection of Query Results from Multiple Web Databases [J]. IJCA Special Issue on Computational Science—New Dimension & Perspectives, 2011(2):71-75

[4] Zhang Zu-ping, Xu Xin, Long Jun, et al. Parameters Correlation and optimization in Text Similarity Measurement[J]. Journal of Chinese Computer Systems, 2011, 32(5): 983-989(in Chinese)

张祖平, 徐昕, 龙军, 等. 文本相似性度量中参数相关性与优化配置研究[J]. 小型微型计算机系统, 2011, 32(5): 983-989

[5] Song Qin-bao, Yang Xiang-rong, Shen Jun-yi, et al. A Detection Algorithm for the Illegal Coping and Distributing of Digital Goods[J]. Chinese Journal of Computers, 2002, 25(11): 1207-1213(in Chinese)

宋擒豹, 杨向荣, 沈钧毅, 等. 数字商品非法复制的检测算法[J]. 计算机学报, 2002, 25(11): 1207-1213

[6] Li Ya-zhou. The research and improvement of an automatic construction system of text classification corpus[D]. Wuhan: Wuhan University of Technology, 2011(in Chinese)

李亚洲. 文本分类语料库自动构建系统的研究与改进[D]. 武汉: 武汉理工大学, 2011

[7] Ye Shao-zhi, Wen Ji-rong, Ma Wei-ying. A systematic study on parameter correlation in large scale duplicate document detection [J]. Knowledge and Information Systems, 2008, 14(2): 217-232

[8] Li Rui-lin, Sun Bing, Li Chao, et al. Differential Fault Analysis on SMS4 using a single fault[J]. Information Processing Letters, 2011, 111(4): 156-163

[9] Shi Kan-sheng, Liu Hai-tao, Song Wen-tao. A Text Clustering Method Based on Speech to Text and Improved Center Selection [J]. Pattern Recognition and Artificial Intelligence, 2012, 25(6): 996-1001(in Chinese)

施侃晟, 刘海涛, 宋文涛. 基于词性和中心点改进的文本聚类方法[J]. 模式识别与人工智能, 2012, 25(6): 996-1001

[10] Xu Ge, Wang Hou-feng. The Development of Topic Models in Natural Language Processing[J]. Chinese Journal of Computers, 2011, 34(8): 1423-1436(in Chinese)

徐戈, 王厚峰. 自然语言处理中主题模型的发展[J]. 计算机学报, 2011, 34(8): 1423-1436

[J]. 计算机工程与科学, 2013, 35(9): 89-95

[10] Chen Jie, Hu Yu-pu, Zhang Yue-yu. Impossible differential attack on the 17-round block cipher SMS4[J]. Journal of Xidian University, 2008, 35(3): 455-458(in Chinese)

陈杰, 胡于濮, 张跃宇. 用不可能差分法分析 17 轮 SMS4 算法[J]. 西安电子科技大学学报, 2008, 35(3): 455-458

[11] Liu Qing, Wei Hong-ru. New Related-key Rectangle Attack on Full ARIRANG Encryption Mode[J]. Computer Science, 2013, 40(8): 109-114(in Chinese)

刘青, 卫宏儒. 对完整轮数 ARIRANG 加密模式的新的相关密钥矩形攻击[J]. 计算机科学, 2013, 40(8): 109-114