Related Key Attacks on Reduced Round KASUMI

Mark Blunden[1] and Adrian Escott[2*]

[1] Kryptosec Ltd, UK.
mark@kryptosec.co.uk
[2] Hutchison 3G, UK.
Adrian.Escott@Hutchison3G.com

Abstract. This paper describes related key attacks on five and six round KASUMI. The five round attack requires the encryption of approximately 2^{19} chosen plaintext pairs X and X^* under keys K and K^* respectively where K and K^* differ in only one bit, and requires a maximum of a little over 2^{33} trials to recover the entire key. The six round attack requires a smaller number of chosen plaintext encryptions than the five round attack, and recovers the entire key in a maximum of 2^{112} trials.

1 Introduction

KASUMI [4] is an eight round, 64-bit block cipher with a 128-bit key. It is based upon MISTY1 [7], and was designed to form the basis of the 3GPP (3rd Generation Partnership Project) confidentiality and integrity algorithms. 3GPP is the body standardizing the next generation of mobile telephony.

This paper describes differential related key attacks on five and six rounds of KASUMI. Differential attacks were introduced by Biham and Shamir in [3]. Related key attacks were first introduced by Biham [2], although a similar idea can be found in Knudsen [6]. Differential related key attacks are discussed in Kelsey *et al* [5].

The attacks require the encryption of plaintext pairs X and X^* under keys K and K^* respectively, where K and K^* differ in only one bit. The attack on five round KASUMI requires on average 2^{19} chosen plaintext pairs, and a maximum of a little over 2^{33} trials to find the entire key. The six round attack requires the encryption of on average $3 \cdot 2^{17}$ chosen plaintext pairs, and a maximum of 2^{112} trials to find the entire key.

Note that the attacks described in this paper present no security threat to KASUMI as used in 3GPP as firstly the attacks do not cover all rounds of KASUMI, and secondly it should not be possible to manipulate the keys in the required way within the 3GPP environment. The motivation for investigating such attacks comes from the fact that one significant difference between KASUMI and MISTY1 is the design of the key schedule, with the key schedule of KASUMI being much simpler than that of MISTY1.

2 KASUMI

KASUMI shares with MISTY1 the design goals of having a numerical basis for its security and of being sufficiently fast when implemented in hardware. More details of the design rationale of KASUMI can be found in [1]. To meet the first goal, the design of KASUMI is based on the theory of provable security of block ciphers against linear and differential cryptanalysis introduced by Nyberg and Knudsen [8][9]. To meet the second design goal, both the S-boxes and the key schedule were carefully chosen to optimize the hardware performance.

KASUMI is an eight round Fiestel type cipher. Each round is made up of an FL function and an FO function. In odd numbered rounds the FL function precedes the FO function, whereas in even numbered rounds the FO function precedes the FL function. The FL function is a simple function that is linear for a fixed key. The FO function contains the non-linearity in each round. The FO function is made up from three FI functions, which in turn are made up from two applications of two different S-boxes and a key addition. The S-boxes are chosen to have optimal resistance to linear and differential cryptanalysis. The overall structure is shown in Figure 1.

KASUMI has a simple, linear key schedule in order to make the hardware significantly smaller and to reduce key set-up time. Every bit of the key is used once in every round. For the purposes of this paper it is sufficient to consider every ith round key as being made up of three parts KL_i, KO_i and KI_i. These in turn are divided into a total of eight 16-bit parts such that $KL_i = KL_{i,1} \| KL_{i,2}$, $KO_i = KO_{i,1} \| KO_{i,2} \| KO_{i,3}$ and $KI_i = KI_{i,1} \| KI_{i,2} \| KI_{i,3}$, where $\|$ denotes concatenation. The $KI_{i,j}$ are further divided into two parts (first one of seven bits, second of nine) with $KI_{i,j} = KI_{i,j,1} \| KI_{i,j,2}$. The round keys are derived by splitting the key K into eight 16-bits parts $K_1 \| \cdots \| K_8$. Each part of the key is used to derive exactly one round key part in each round. The key schedule has the property that changing one bit of K changes exactly one key bit of each round key.

3 Four Round Differentials of KASUMI

This section describes a family of four round differentials that can be used to attack up to six rounds of KASUMI. The core idea behind this family of differentials is to choose small key differences and control the effect of these differences as they progress through the cipher using either the properties of the FL function or appropriately chosen plaintext differences.

The differential is formed by encrypting plaintexts X and X^* under keys K and K^* respectively, where $K_j = K_j^*$ for $j = 1, 2, 4, \ldots, 8$, $K_3 = K_3^* \oplus k$ for any 16-bit string k, $X = 0^{16}1^{16}a$ for any 32-bit string a, $X^* = 0^{16}1^{16}(a \oplus (k <<< 5)0^{16})$ and $<<< n$ denotes a left circular rotation by n bits. With probability of $2^{-wt(k)}$, where $wt(k)$ is the number of non-zero bits of k, the output difference at the end of the fourth round is $c(k <<< 5)0^{16}$, where c is any 32 bit string. In practice, k is chosen such that $wt(k) = 1$ in order to maximize the probability

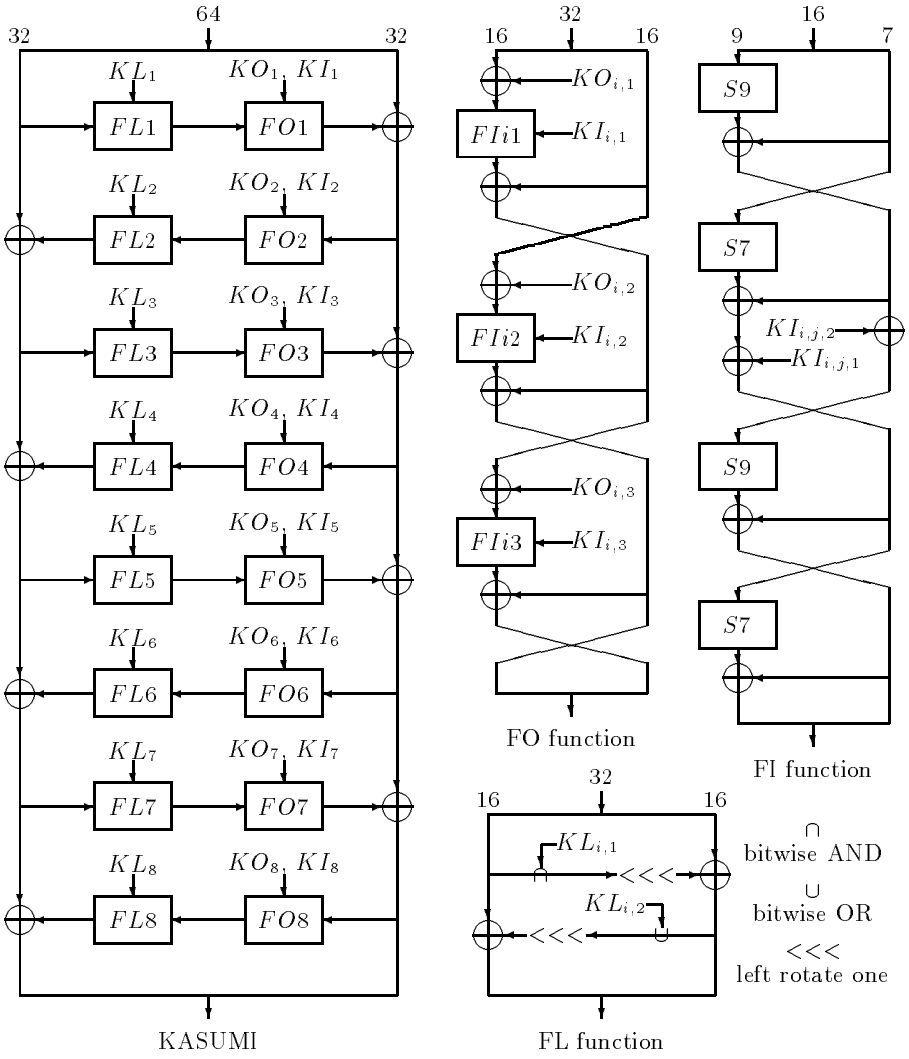
64
32
32
KL_1
KO_1, KI_1
$FL1$
$FO1$
KL_2
KO_2, KI_2
$FL2$
$FO2$
KL_3
KO_3, KI_3
$FL3$
$FO3$
KL_4
KO_4, KI_4
$FL4$
$FO4$
KL_5
KO_5, KI_5
$FL5$
$FO5$
KL_6
KO_6, KI_6
$FL6$
$FO6$
KL_7
KO_7, KI_7
$FL7$
$FO7$
KL_8
KO_8, KI_8
$FL8$
$FO8$
KASUMI
32
16
16
$KO_{i,1}$
$FIi1$
$KI_{i,1}$
$KO_{i,2}$
$FIi2$
$KI_{i,2}$
$KO_{i,3}$
$FIi3$
$KI_{i,3}$
FO function
16
9
7
$S9$
$S7$
$KI_{i,j,2}$
$KI_{i,j,1}$
$S9$
$S7$
FI function
32
16
16
$KL_{i,1}$
$KL_{i,2}$
<<<
<<<
∩
bitwise AND
∪
bitwise OR
<<<
left rotate one
FL function

Fig. 1. KASUMI

of the differential holding. The following paragraph provides more detail as to why the differential works.

If K and K^* are as defined above, then the only non zero round key part differences in the first three rounds are $\Delta KL_{1,2} = k$, $\Delta KO_{2,1} = k <<< 5$ and $\Delta KL_{3,1} = k <<< 1$. Now, the FL function has the property that for any key KL, an input of $0^{16}1^{16}$ always gives an output of 1^{32}. Hence when encrypting plaintexts X and X^* as defined above under keys K and K^* respectively, the input and output differences of $FL1$ are both 0^{32}. The round key difference in the second round combines with the difference in the right hand side of the plaintexts such that the output difference of $FO2$ is always 0^{32}. Since the input difference to $FL3$ is 0^{32}, and the AND of a message bit and key bit is always zero when the message bit is zero, the difference in the third round key has no effect on the output of $FL3$ with probability $2^{-wt(k)}$. Therefore, with probability $2^{-wt(k)}$ the output difference at the end of the fourth round is $c(k <<< 5)0^{16}$.

Figure 2 illustrates this family of four round differentials over six round KASUMI.

Note that the restriction on the inputs to $FL1$ to always be of the form $0^{16}1^{16}$ is stronger than necessary when $wt(k) < 16$, but is defined this way for simplicity.

4 An Attack on Five Round KASUMI

This section describes a related key attack on five round KASUMI. The attack requires an average of 2^{15} chosen plaintext encryptions under the original key and 2^{19} chosen plaintext encryptions under a chosen related key. The attack is expected to recover the entire key in a maximum of a little over 2^{33} trials, where each trial is equivalent to approximately 1·6 five round KASUMI encryptions.

When encrypting a plaintext pair X and X^* under keys K and K^* respectively, where X, X^*, K and K^* are as described in Section 3, with probability 2^{-16} the ciphertext difference is $Y \oplus Y^* = cge$, where c is any 32 bit string, e is any 16 bit string and $g = e \oplus (k <<< 5)$. When $wt(k) = 1$ the attack requires an average of eight ciphertext pairs of this form, and six ciphertext pairs whose inputs are of this form but whose output difference can be of any form. It is expected that these ciphertext pairs can be found from a set of 2^{19} ciphertext pairs generated from inputs of the appropriate form. One method of generating this set is to obtain the encryptions of 2^{15} chosen plaintexts under the key K and for each such encryption obtain the sixteen related encryptions that result from using the different values of k such that $wt(k) = 1$. Note that with K and K^* as defined above, the only non zero round key part difference in round five is $\Delta KO_{5,3} = k <<< 13$. Note also that the rightmost 32 bits of the plaintexts do not need to be known.

The attack assumes that the eight ciphertext pairs whose output differences are of the correct form all satisfy the differential described in Section 3. Consequently, for each of these eight ciphertext pairs it follows that the output difference of $FO5$ is $ge \oplus (k <<< 5)0^{16} = ee$, so that the output difference

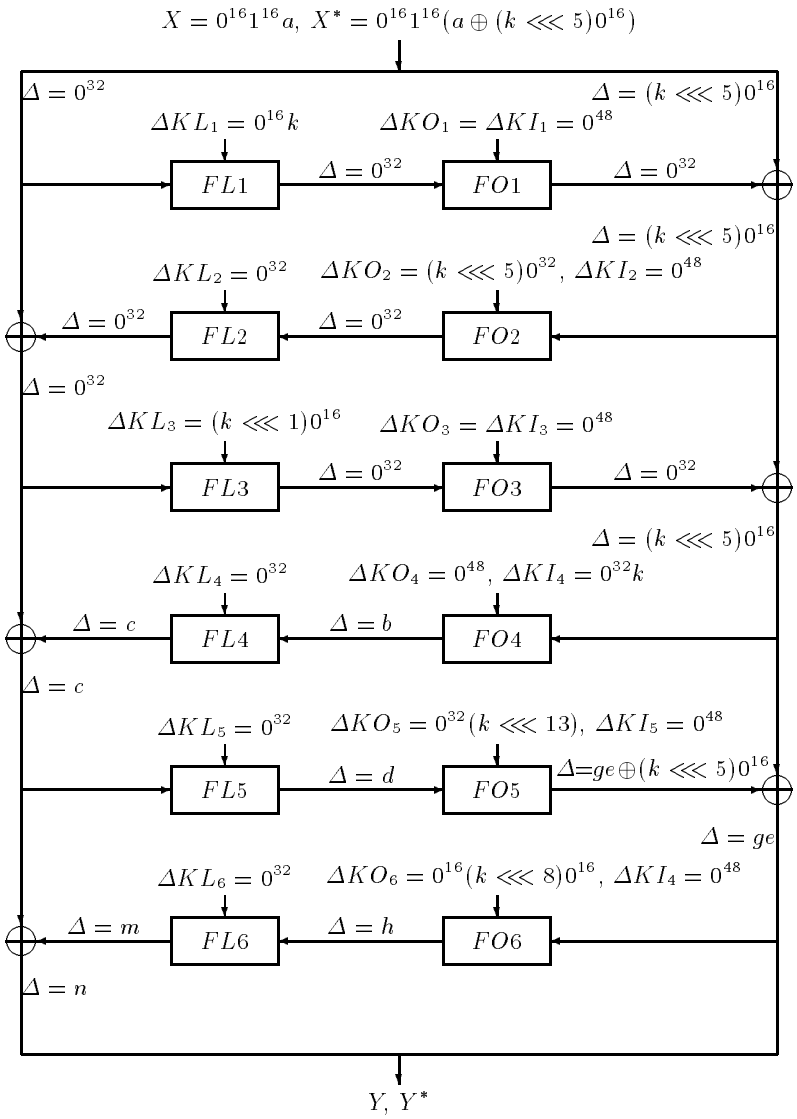

Fig. 2. A family of four round differentials over six-round KASUMI which hold with probability one half, where k, e and g are any 16-bit strings and a, b, c, d, h and m are any 32 bits strings.

of $FI53$ is $e \oplus e = 0^{16}$. Since $\Delta KI_{5,3} = 0^{16}$, the input difference to $FI53$ is also zero. Therefore, as $\Delta KO_{5,3} = k <<< 13$, the output difference of $FI52$ is $e \oplus (k <<< 13)$. Furthermore, for each guess of $KL_{5,1}$ and $KO_{5,2}$, the inputs to $FI52$ are known. Hence, for each ciphertext pair, the difference tables of $S7$ and $S9$ can be used to suggest values of $KI_{5,2}$.

For each guess of $KL_{5,1}$ and $KO_{5,2}$, a count is made of the number of times each possible value of $KI_{5,2}$ is suggested by a ciphertext pair. The number of times a key $KI_{5,2}$ is suggested for a particular choice of $KL_{5,1}$ and $KO_{5,2}$ is taken as the number of times the key triple $KL_{5,1}$, $KO_{5,2}$, $KI_{5,2}$ is suggested. Any key triple suggested as many as four times is kept.

The correct value of the triple $KL_{5,1}$, $KO_{5,2}$ and $KI_{5,2}$ will be included in the values suggested by each of the ciphertext pairs that do genuinely follow the differential of Section 3. Such pairs are referred to as right pairs. Since it is expected that four of the eight pairs satisfy the differential, the correct value of the triple $KL_{5,1}$, $KO_{5,2}$ and $KI_{5,2}$ is expected to be suggested at least four times. The probability of an incorrect triple being suggested by one ciphertext pair is 2^{-16}, and the probability of an incorrect triple being suggested as many as four times is less than 2^{-57}. Therefore, the probability that at least one incorrect triple is suggested as many as four times is less than 2^{-9}. Thus the correct triple $KL_{5,1}$, $KO_{5,2}$ and $KI_{5,2}$ is expected to be uniquely identified. This requires one round of KASUMI to be unwound for each choice of values for $KI_{5,1}$ and $KO_{5,2}$ and each of the eight differential pairs. Hence the work factor is equivalent to at most 1.6×2^{32} five round KASUMI encryptions.

Once the subkey $KL_{5,1}$ has been recovered, the sixteen rightmost bits h of the input difference to $FO5$ are known for each ciphertext pair. Therefore for each right pair the output difference $(k <<< 13) \oplus h$ of $FI51$ is known. Guessing the subkeys $KL_{5,2}$ and $KO_{5,1}$ allows the input values to $FI51$ to be computed. Hence, using only those ciphertext pairs that suggested the triple $KL_{5,1}$, $KO_{5,2}$ and $KI_{5,2}$, the triple $KL_{5,2}$, $KO_{5,1}$ and $KI_{5,1}$ can be recovered using a similar technique to that used to recover $KL_{5,1}$, $KO_{5,2}$ and $KI_{5,2}$.

The subkeys $KO_{5,3}$ and $KI_{5,3}$ are then easily recovered using the already recovered round key parts and the six ciphertext pairs whose output differences are not of the form $Y \oplus Y^* = cge$, where $g = e \oplus (k <<< 5)$. For each guess of $KO_{5,3}$, the input values to $FI53$ can be computed for each ciphertext pair. Since the output difference of $FI_{5,3}$ is also known for each right pair, a similar method to that described above can be used to recover the actual values of $KO_{5,3}$ and $KI_{5,3}$. The correct values of $KO_{5,3}$ and $KI_{5,3}$ are expected to be uniquely identified in at most 2^{16} trials. Note that the subkeys $KO_{5,3}$ and $KI_{5,3}$ can also be recovered by a brute force search in at most 2^{32} five round KASUMI encryptions.

Note that if all four right pairs have identical bit values in one or more common locations within the leftmost sixteen bits of the ciphertexts then there are at least two key pairs $KL_{5,1}$ and $KO_{5,2}$ giving the correct input values to $FI52$ for all four right pairs. A similar property applies to $KL_{5,2}$ and $KO_{5,1}$. Hence with probability ≈ 0.22 the correct key will not be uniquely identified.

However, the attack is easily modified to cope with such an event, and the expected impact on running time is minimal.

This attack was implemented on a 400 Mhz Alpha and took approximately five hours to complete.

5 An Attack on Six Round KASUMI

This section describes a related key attack on six round KASUMI. The idea behind the attack is to obtain ciphertext pairs with a property such that it is possible to work back through round six without having to guess all of the round six round key bits, then use round five and the differential of Section 3 to suggest values for the missing key bits, and finally to use round four to help eliminate incorrect values for the key. The attack requires the encryption of on average 3×2^{13} chosen plaintexts under the original key and 3×2^{17} chosen plaintexts under a chosen related key, and is expected to recover the entire key in a maximum of 2^{112} trials, equivalent to 15×2^{113} six round KASUMI encryptions.

When encrypting plaintexts X and X^* under keys K and K^* respectively, where X, X^*, K and K^* are as described in Section 3, with probability 2^{-16} the ciphertext difference is $Y \oplus Y^* = nge$, where n is any 32-bit string, g is any 16-bit string and $e = k <<< 8$. When $wt(k) = 1$ the attack requires an average of six ciphertext pairs of this form, and it is expected that these ciphertext pairs can be found from a set of 3×2^{17} ciphertext pairs generated from inputs of the appropriate form. One method of generating this set is to obtain 2^{15} chosen plaintext encryptions under the key K and for each such encryption obtain the sixteen related encryptions that result from using the different values of k such that $wt(k) = 1$. Note that with K and K^* as defined above, the only non zero round key part differences in rounds five and six are $\Delta KO_{5,3} = k <<< 13$ and $\Delta KO_{6,2} = k <<< 8$. Note also that the rightmost 32 bits of the plaintexts do not need to be known.

Note that for each of the six ciphertext pairs, the input difference to $FI62$ is 0^{16}, since $\Delta KO_{6,2} = k <<< 8 = e$. As $\Delta KI_{6,2} = 0^{16}$, the output difference of $FI62$ is also 0^{16}, so that there are only 2^{16} possible outputs of $FI62$ for each pair. Therefore, if the 16-bit output t of $FI62$ is known for any of these ciphertext pairs, it is not necessary to know the round keys $KO_{6,2}$ and $KI_{6,2}$ in order to compute the output of round six.

The attack guesses the ninety six key bits KL_6, $KO_{6,1}$, $KO_{6,3}$, $KI_{6,1}$ and $KI_{6,3}$ (so that the only unknown parts of the fifth round key are $KI_{5,1}$ and $KO_{5,3}$), and then takes each possible value of the string t in turn and assumes it to be the output of $FI62$ for all six ciphertext pairs. These values are then used along with the ciphertext values to compute the input values of $FI51$ and the output values of $FI52$ for all ciphertext pairs. The attack then assumes that all ciphertext pairs follow the differential of Section 3. This enables the output difference of $FI51$ to be computed for each ciphertext pair. The difference tables of $S7$ and $S9$ can then be used to suggest values for $KI_{5,1}$. Note that key bits

must be suggested by both $S7$ and $S9$ in order for a value to be suggested for $KI_{5,1}$. On average one value of $KI_{5,1}$ will be suggested by each ciphertext pair.

Each value of $KI_{5,1}$ suggested by a ciphertext pair is then used to compute the two outputs of $FI51$ for that pair. These outputs are then XORed with the sixteen rightmost input bits to $FO5$ and the resulting values used along with the output difference $(k <<< 5) \oplus (k <<< 8) \oplus g$ of $FI53$ to suggest values for $KO_{5,3}$ by considering $FI53$ as a key dependent S-box, since the key $KI_{5,3}$ is equivalent to the already guessed key $KO_{6,3}$. On average one value of $KO_{5,3}$ will be suggested by each ciphertext pair.

If values are suggested for both $KI_{5,1}$ and $KO_{5,3}$ then each suggested pair $KI_{5,1}$, $KO_{5,3}$ is used along with the ninety six guessed key bits to compute from the two ciphertext values the output difference of $FL4$. If this output difference is equal to the input difference to $FL5$ then the pair $KI_{5,1}$, $KO_{5,3}$ is checked against a list of possible values for the key pair suggested already by that ciphertext pair for different values of t but the same value of the ninety six key bits KL_6, $KO_{6,1}$, $KO_{6,3}$, $KI_{6,1}$ and $KI_{6,3}$ and, if not already there, added to the list. If the suggested key pair already appears in the list, then it should not be added again. This is to prevent the same key pair being suggested more than once by a given ciphertext pair and value of KL_6, $KO_{6,1}$, $KO_{6,3}$, $KI_{6,1}$ and $KI_{6,3}$. The probability of wrong keys generating the correct output difference of $FL4$ is 2^{-32}.

It is expected that three of the six ciphertext pairs follow the differential of Section 3. Therefore, taking the number of times a pair $KI_{5,1}$, $KO_{5,3}$ (equivalently $KO_{6,2}$, $KI_{6,2}$) is suggested for a particular choice of KL_6, $KO_{6,1}$, $KO_{6,3}$, $KI_{6,1}$ and $KI_{6,3}$ as the number of times the key KL_6, KO_6 and KI_6 is suggested, the correct value of the key KL_6, KO_6 and KI_6 is expected to be suggested at least three times. An incorrect key has probability 2^{-64} of being suggested by any one particular ciphertext pair and value of t, so that the probability of an incorrect key being suggested by any one particular pair is less than 2^{-48}. Therefore, the probability that an incorrect key is suggested by at least three out of the six pairs is less than 2^{-139}, and the probability that at least one incorrect key is suggested as many as three times is less than 2^{-11}. It is therefore expected that the correct key is the only key to be suggested as many as three times.

6 Conclusion

This paper describes related key attacks on both five and six round KASUMI. Under the assumption that it is possible to obtain the encryptions of approximately 2^{19} chosen plaintext pairs X and X^* under keys K and K^* respectively, where K and K^* differ in only one bit, the five round attack is practical in terms of computing power. The six round attack, while requiring fewer chosen plaintext encryptions than the five round attack, is of theoretical interest only due to the computing power required to implement the attack.

The attacks described above both rely on the same differential that predicts differences at the end of the third round with probability one half. This differen-

tial arises from the use of a very simple key schedule and the locations of where round key parts dependent upon K_3 are fed into the first three rounds. Changing the locations of where these round key parts are fed into the rounds may destroy this particular differential, but doesn't necessarily eliminate the possibility of constructing similar differentials.

Note that the attacks described in this paper present no real security threat to KASUMI as used in 3GPP as firstly they are only valid against reduced round variants of KASUMI, and secondly they require plaintexts to be encrypted under related keys in a way that should not be possible within the 3GPP environment.

References

1. S. Babbage, "Design of Security Algorithms for Third Generation Mobile Telephony." In *Information Security Technical Report (Elsevier)*, (5), 2000.
2. E. Biham, "New Types of Cryptanalytic Attacks Using Related Keys." In *Advances in Cryptology - EUROCRYPT '93*, Lecture Notes in Computer Science (LNCS 765), Springer-Verlag, 1994.
3. E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems." In *Journal of Cryptology*, (4), 1991.
4. ETSI, see http://www.etsi.org/dvbandca/3GPP/3gppspecs.htm
5. J. Kelsey, B. Schneier and D. Wagner, "Key-Schedule Cryptanalysis of IDEA, G-DES, GOST, SAFER and Triple-DES." In *Advances in Cryptology - CRYPTO '96*, Lecture Notes in Computer Science (LNCS 1109), Springer-Verlag, 1996.
6. L. Knudsen, "Cryptanalysis of LOKI91." In *Advances in Cryptology - AUSCRYPT '92*, Lecture Notes in Computer Science (LNCS 718), Springer-Verlag, 1994.
7. M. Matsui, "New Block Encryption Algorithm MISTY." In *Fast Software Encryption: 4th International Workshop*, Lecture Notes in Computer Science (LNCS 1267), Springer-Verlag, 1997.
8. K. Nyberg, "Linear Approximation of Block Ciphers." In *Advances in Cryptology - EUROCRYPT '94*, Lecture Notes in Computer Science (LNCS 950), Springer-Verlag, 1995.
9. K. Nyberg and L. Knudsen, "Provable Security Against a Differential Attack." In *Journal of Cryptology*, (8), 1995.