Tight Lower Bound on Linear Authenticated Encryption

Charanjit S. Jutla
IBM T. J. Watson Research Center,
Yorktown Heights, NY 10598

Abstract

We show that any scheme to encrypt m blocks of size n bits while assuring message integrity, that apart from using $m + k$ invocations of random functions (from n bits to n bits) and vn bits of randomness, is linear in $(GF2)^n$, must have $k + v$ at least $\Omega(\log m)$. This lower bound is proved in a very general model which rules out many promising linear modes of operations for encryption with message integrity. This lower bound is tight as [8] shows a linear scheme to encrypt m blocks while assuring message integrity by using only $m + 2 + \log m$ invocations of random permutations.

1. Introduction

Recently new modes of operation for block ciphers (IAPM, IACBC) were described in [8], which in addition to assuring confidentiality of the plaintext, also assure message integrity. A related new mode of operation (XCBC-XOR) which accomplished simultaneous encryption and authentication was independently described in [5]. Prior to this, two separate passes were required; first to compute a cryptographic MAC (e.g. CBC-MAC [2]) and then to encrypt the plaintext with the MAC appended to the plaintext (e.g. using CBC [15]). Following up on works of [8], and [5], another authenticated encryption mode (OCB) was described in [17].

Before the modes in [8] and [5] many unsuccessful attempts were made to do authenticated encryption in one pass (e.g. [4]). Most of these attempts try to use a simple checksum (or not so non-linear code) instead of a cryptographic MAC as the tag appended to the plaintext before encryption. Other attempts try to do additional chaining, on top of the cipher block chaining in CBC (see figure 2 for one such mode called PCBC - plaintext ciphertext block chaining). In essence, all these proposed modes try to do authenticated encryption by using only exclusive-or operations (i.e. operations linear in $(GF2)^n$, where n is the block cipher size), or operations which can be approximated by such operations with reasonably high probability, and without generating any extra randomness using the block cipher or some pseudo-random function. A successful mode for authenticated encryption was described in [9], however it increased the length of the ciphertext by a constant factor.

The mode in [8] is proven to be secure for both encryption and authentication even though it only uses operations linear in $(GF2)^n$ (apart from block cipher invocations), but it actually generates $\log m$ extra blocks of randomness (where m is the number of blocks to be encrypted) by $\log m$ extra block cipher invocations.

In this paper we show a matching lower bound to the construction in [8] (see Figure 1). In other words, we show that the $\log m$ additional cryptographic operations in IAPM/IACBC scheme are essentially the least one has to do to assure message integrity along with message secrecy in any scheme linear in $(GF2)^n$.

We prove our lower bound in a very general model. We assume that the block cipher is modeled as a length preserving random function on n bits. Any invocation of such a random function constitutes one application of a cryptographic function. The only other operations allowed are linear operations over $(GF2)^n$ (i.e. n-bit exclusive-or), or testing an n bit quantity for zero. There is no other restriction on the scheme, apart from it being one to one (i.e. no two plaintexts generate the same ciphertext). There is no assumption about whether the scheme is actually invertible (which is the surprising part). The scheme is also allowed to be probabilistic with v blocks of randomness.

As our main result, we prove that any such linear scheme which encrypts m blocks of plaintext while assuring message integrity, using v blocks of randomness, and only $m + k$ cryptographic operations, must have $k + v$ at least $\Omega(\log m)$.

We use well known theorems from linear algebra to prove our lower bound. Specifically we analyze the ranks of matrices and solution spaces of linear system of equations to prove the lower bound. Such a linear algebra technique like analysis of rank of matrices has been used previously by [11] to show attacks on a whole class of schemes (double block length hash functions), although the matrices involved in [11] were of constant ranks (three or four).

We again emphasize that our lower bound is a very general result, as it rules out many potential schemes for authenticated encryption by just an inspection of the number of cryptographic operations, and the mixing operations used (i.e. regardless of the structure of the scheme). Figures 3 and 4 (in addition to Figure 2) describe some other modes which by this lower bound turn out to be insecure for authenticated encryption. The mode in Figure 3 tries to use the structure of both the counter mode[14], and the CBC mode. All mixing operations in Figure 1

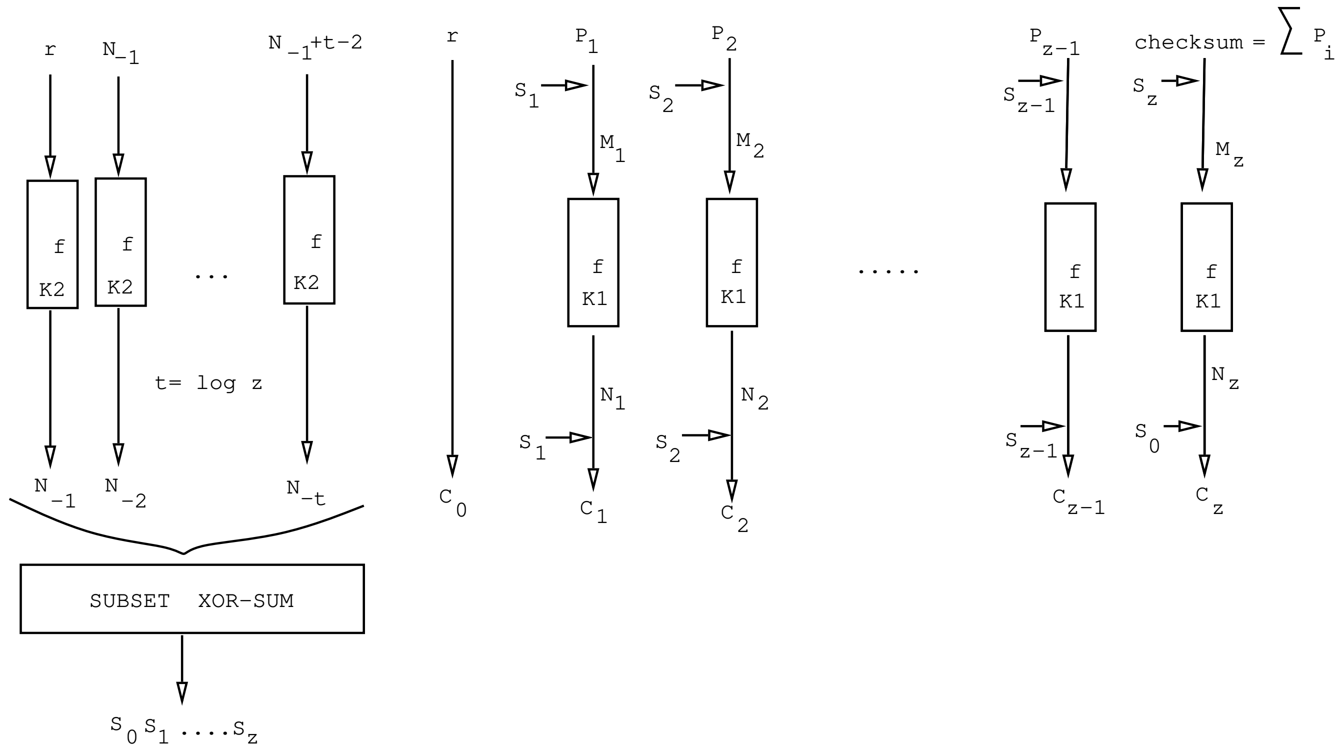

Figure 1: Authenticated Encryption Mode IAPM

to 4 are n-bit exclusive or operations.

Note that there are versions of IAPM/IACBC in [8], and modes for authenticated encryption in [5],[17] which are proven secure while using only one or two extra cryptographic operations. However, these schemes are not linear in $(GF2)^n$. In fact, the main theorem in [8] shows that authenticated encryption can be achieved by generating and using (linearly) a sequence of random numbers which are only pairwise -differentially uniform (or XOR-universal, a property slightly weaker than pairwise independence). Such a sequence can be generated by only one additional cryptographic operation if additional operations in GFp or $GF(2^n)$ are allowed, but such a sequence does indeed require $\log m$ extra cryptographic operations, if only linear in $(GF2)^n$ operations are allowed.

The rest of the paper is organized as follows. In section 2 we state some lemma from linear algebra which are used in proving the main theorem. In section 3 we describe the model of authenticated encryption. In section 4 we prove the main lower bound theorem.

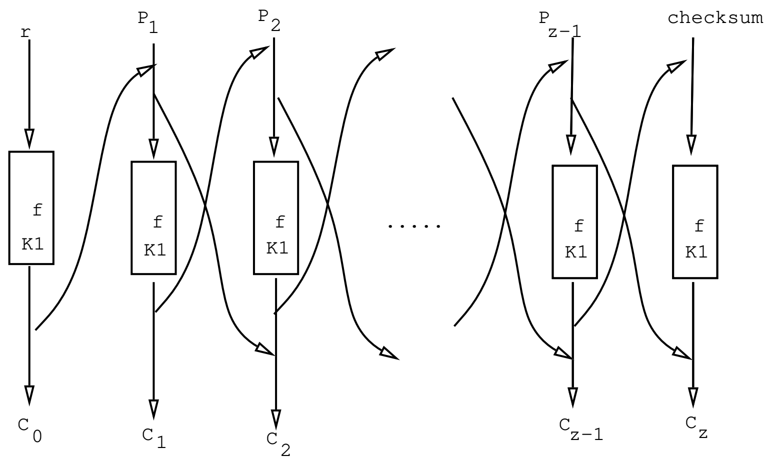

Figure 2: Encryption Mode PCBC

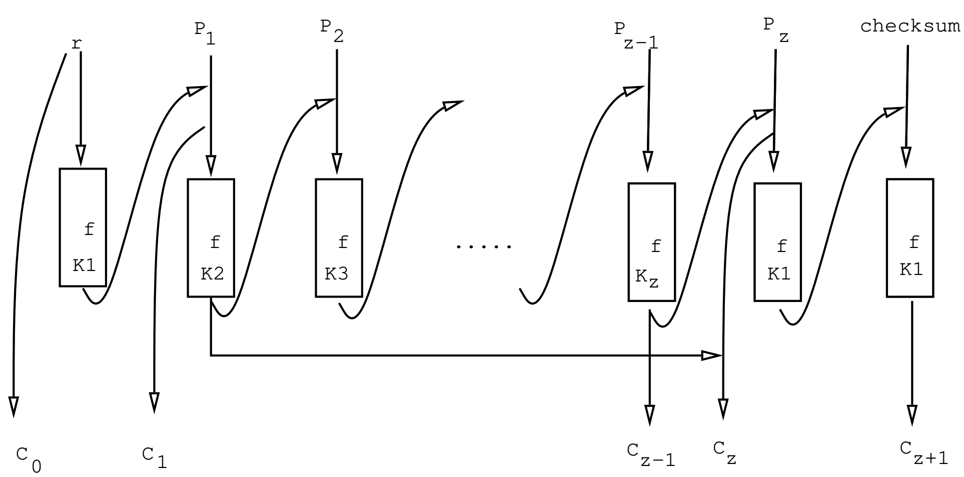

Figure 3: Erroneous Mode 2

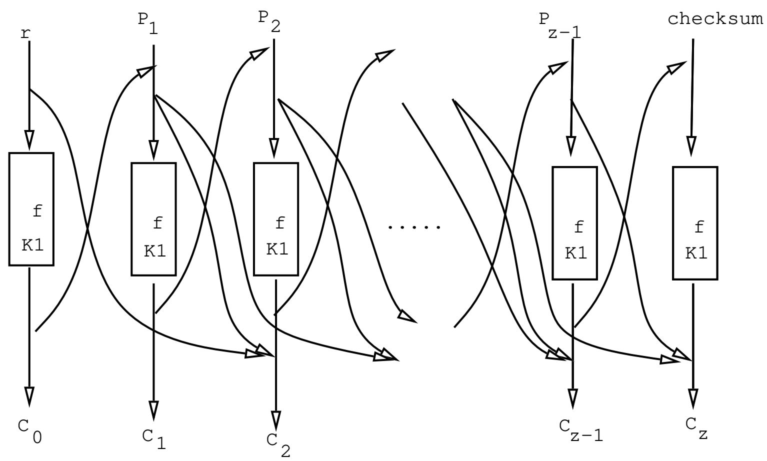

Figure 4: Erroneous Mode 3

2. Linear Algebra Basics

Proofs of the following lemmas of linear algebra can be found in any basic linear algebra book (e.g. [10]).

Lemma 2.1: Let

$$[X_1 ... X_q] \cdot \mathbf{A} \; = [Y_1 ... Y_m]$$

where $\mathbf{A}$ is a $q \times m$ binary matrix of rank m, and all the variables represent elements of $(GF2)^n$. If for some subset $\mathbf{B}$ of rows of $\mathbf{A}$, rank($\mathbf{B}$) ¡ m, then there is a non-trivial linear (over GF2) relation between the variables $Y_1 ... Y_m$, and variables $\{X_i | i \in [1..q]$, and i not index of some row in $\mathbf{B}\}$.

Lemma 2.2: Let

$$[X_1 ... X_q] \cdot \mathbf{A} \; = [Y_1 ... Y_m]$$

where $\mathbf{A}$ is a $q \times m$ binary matrix of rank $m' \leq m$, and $m \leq q$, and all the variables represent elements from $(GF2)^n$. Then for a fixed $Y = [Y_1 ... Y_m]$ the solution space of $[X_1 ... X_q]$ is a $q - m'$ dimensional affine space, namely

$$[X_1 ... X_q] = [< f(Y) >] + \alpha_1 \cdot V_1 + ... + \alpha_{q-m'} \cdot V_{q-m'}$$

where $< f(Y) >$ is a row of q linear functions determined by $\mathbf{A}$, each of $\alpha_1 ... \alpha_{q-m'}$ is a scalar ranging over all elements in $(GF2)^n$, and $V_1 ... V_{q-m'}$ are $q - m'$ linearly independent binary row vectors determined by $\mathbf{A}$.

3. Authenticated Encryption Model

We consider the following model. We assume a fixed block size n for a block cipher (or random permutations or length preserving random functions). Any application of one of these will constitute one application of a cryptographic operation. From now on we will assume that the block cipher is modeled as a length preserving n bit random function, although a scheme may use many different random functions. The only other operations allowed are linear operations over $(\mathrm{GF2})^n$, i.e. bit-wise exclusive-or. Of course, operations of testing whether an n bit quantity is zero is also allowed. Since, the scheme could be probabilistic, as IACBC/IAPM [8] is, we also allow v blocks of randomness, $r_1, ..., r_v$.

Let the message P to be encrypted be of size m blocks, i.e. mn bits. Call the input blocks $P_1, ..., P_m$. Let there be $m+k$ invocations of random functions, and let the inputs to these functions be $M_1, M_2, ..., M_{m+k}$. Similarly, let the outputs of these random functions be $N_1, N_2, ..., N_{m+k}$. Let, $C = C_1, C_2, ...C_{m+t}$ be linear functions of P's, r's, and N's. Here $t \geq 0$.

Our aim is to show that either the scheme is not secrecy secure, or it is not message integrity secure, or it is not one to one (not just not invertible), or $k+v = \Omega(\log m)$. More formally, the scheme is not *secrecy secure* if an adversary can correctly predict a non-trivial linear combination of the plaintext blocks, given the corresponding ciphertext, with probability more than $1 - O(2^{-n})$, in time polynomial in m and n. Note that we do not need an epsilon-delta definition of security, as we will always be able to demonstrate attacks which work with high probability. Also, the adversary need not see many ciphertexts before predicting the plaintext. Thus, we have a weak adversarial model.

For *message integrity*, let there be $u > 0$ MDC functions $\mu_1, ..., \mu_u$. Without loss of generality assume that they are linearly independent. During encryption of plaintext P, using randomness r, each μ_i is computed as a linear combination of P's, M's, N's, and r's. During decryption of the corresponding ciphertext C, another set of functions $\mu'_1, ..., \mu'_u$ is computed as a function of C's, M's, and N's. The decryption process passes the message integrity test if for all i, $\mu_i = \mu'_i$. For example in IAPM (fig 1), $\mu_1 = \Sigma P$, and $\mu'_1 = M_z \oplus S_z$, where S_z is some linear combination of $N_{-1}...N_{-t}$. Thus, define $D_i = \mu_i \oplus \mu'_i$, a linear function of P,M,N,r, and C. Since C can be written as a linear combination of P, N, and r, each D_i is a linear function of P,M,N, and r. On a valid decryption all the D_i should evaluate to zero.

A scheme is not message integrity secure, if given P and corresponding C, an adversary can produce a $C' \neq C$ in time polynomial in m and n, such that on inversion, all the functions D_i evaluate to zero. Once again, we do not require many plaintext, ciphertext combinations before a forged ciphertext is demonstrated.

A scheme is *one-to-one* if it is not the case that there are two plaintext messages P^1 and P^2, and two random strings r^1, and r^2, such that ¡r^1, P^1¿ generates ciphertext C, and ¡r^2, P^2¿ generates the same ciphertext C.

Since, the ciphertext is a linear combination of P, r and N, and similarly, each M_i is a linear combination of P, r and N, let

$$[P_1 ... P_m r_1 ... r_v N_1 ... N_{m+k}\ 1] \cdot \mathbf{B} \ = \ [C_1 ... C_{m+t}]$$

$$[P_1 ... P_m r_1 ... r_v N_1 ... N_{m+k}\ 1] \cdot \mathbf{E} \ = \ [M_1 ... M_{m+k}]$$

where each of **B** and **E** is a binary matrix except for the last row which can have arbitrary elements in $(GF2)^n$.

Clearly, given P, r and N, the resulting C using the first set of equations is a valid ciphertext. However, given C, is every solution of the first set of equations a *consistent decryption*? This is not to be confused with a *valid decryption*, which is a consistent decryption with all D_i evaluating to zero.

It is not possible in general to determine if a solution in terms of P, r and N (given C) of the first set of equations is consistent, *unless* we know what each of N's decrypt to, in which case we can use the second set of equations to check for consistency.

Formally, let f_i be a random function modeling the block cipher or a Pseudorandom function which is used to compute N_i from M_i. Then, given C, a solution of the two set of equations (**B** and **E**) is a **consistent decryption** if $f_i(M_i) = N_i$ (for $i \in [1..m+k]$).

Let

$$[M_1 ... M_{m+k} P_1 ... P_m r_1 ... r_v N_1 ... N_{m+k}\ 1] \cdot \mathbf{F} \ = \ [D_1 ... D_u]$$

We combine these three systems of equations to write a big system as follows:

$$[M_1 ... M_{m+k} P_1 ... P_m r_1 ... r_v N_1 ... N_{m+k}\ 1] \cdot \mathbf{G} \ = \ [C_1 ... C_{m+t} D_1 ... D_u\ 0...0]$$

where there are $m+k$ 0's in the R.H.S vector corresponding to the matrix **E** (i.e. second system of equations). More precisely,

$$\mathbf{G} \ = \ \begin{bmatrix} \mathbf{0} & \mathbf{F} & \mathbf{I} \\ \mathbf{B} & & \mathbf{E} \end{bmatrix}$$

We will refer to a given *authenticated encryption scheme* by the matrix **G**.

4. Lower Bound

Theorem 1: If the scheme $\mathbf{G}$ is secrecy secure, message integrity secure, and one-to-one, then $k+v$ is at least $\Omega(\log m)$.

We begin by proving some lemmas.

Lemma 1: Either the scheme $\mathbf{G}$ is not secrecy secure or $k+v \geq t+u-1$

Proof: Use the identity matrix in the top right corner of $\mathbf{G}$ to zero out the first $m+k$ rows of the the first $m+t+u$ columns. We call this new matrix $\mathbf{G}'$. We now write the columns corresponding to D as $[\mathbf{0}\ \mathbf{F'}^{\top}]^{\top}$. Thus,

$$\mathbf{G}' = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{B} & \mathbf{F}' & \mathbf{E} \end{bmatrix}$$

We first show that w.l.o.g. the rank of the matrix $\mathbf{G}'$ is $m+t+u$. Suppose the rank of the matrix $\mathbf{G}'$ is $m' < m+t+u$. Clearly the columns corresponding to D are linearly independent, as we assumed earlier. Thus, there are $m+t+u-m'$ columns corresponding to C which are linear combinations of the other columns corresponding to C, and the columns corresponding to D. However, on a valid encryption all the D_i are zero. This means, that these $m+t+u-m'$ C_i's corresponding to the columns mentioned can be computed by a linear combination of the other $m+t-(m+t+u-m') = m'-u$ C_i's. Thus, there need only be $m'-u$ C_is, in the big equation above. Thus, we can assume w.l.o.g that the rank of the matrix $\mathbf{G}'$ is indeed $m+t+u$.

In fact, the rank of the sub-matrix of $\mathbf{G}'$ consisting of the first $m+t+u$ columns is $m+t+u$.

Now, let's focus on the matrix $\mathbf{G}''$ comprising of the rows of $\mathbf{G}'$ excluding the rows corresponding to P, and only the first $m+t+u$ columns. If the rank of the sub-matrix $\mathbf{G}''$ is less than $m+t+u$, then there is a non-trivial linear relationship between C's, D's, and P. Once again, since on a valid encryption D's are zero, we would get a non-trivial linear relation between P's and C's, contradicting that the scheme is secrecy-secure. Since the $m+k$ rows of the first $m+t+u$ columns of $\mathbf{G}''$ are zero, we have that $(v+(m+k)+1) \geq m+t+u$, or $k+v \geq t+u-1$. $\square$

Going back to $\mathbf{G}$, it is useful to reduce the rows corresponding to P in $\mathbf{B}$ and $\mathbf{F}$ to zero, if possible. In other words, we would like to express P in terms of M, N, and r, if possible. So, by doing column operations, let the the rows in $[\mathbf{I}\mathbf{E}^{\top}]^{\top}$ corresponding to P be reduced to

$$\begin{pmatrix} 0 & X \\ 0 & \mathbf{I} \end{pmatrix}$$

where the identity matrix is of size w, $0 \leq w \leq m$, resulting in the new equation

$$[M_1...M_{m+k}P_1...P_m r_1...r_v N_1...N_{m+k}\ 1] \cdot \begin{bmatrix} \mathbf{0} & \mathbf{F} & \mathbf{E}' \\ \mathbf{B} & & \end{bmatrix} = [C_1...C_{m+t}D_1...D_u\ 0...0]$$

Consequently we can assume, w.l.o.g., that the bottom w rows corresponding to P in $\mathbf{F}'$ are zero.

We now have the system of equations

$$[M_1...M_{m+k}P_1...P_m r_1...r_v N_1...N_{m+k}\ 1] \cdot \mathbf{H} = [C_1...C_{m+t}D_1...D_u\ 0...0]$$

where the bottom w rows corresponding to P are zero in the columns corresponding to D.

Lemma 2: Either the scheme $\mathbf{G}$ is not message integrity secure or $u \geq (m - w)$

Proof: Let c be a ciphertext which is computed based on a given p and r. Consider the sub-matrix of $\mathbf{H}$ which consists of the first $m - w$ rows corresponding to the P's and the u columns corresponding to D. If this sub-matrix has rank less than $m - w$, then there is a $p' \neq p$ (with p' different from p only in the first $m - w$ indices (blocks)), such that D's remain same, i.e. zero. Because, of the identity matrix in $\mathbf{E}'$ we can arrive at a p'', which is identical to p' in the first $m - w$ blocks, but possibly different in the remaining w blocks, so that none of the M's and N's are affected (i.e. p'' is consistent with same Ms and Ns as computed from p). The new p'' still keeps all the Ds zero (as the bottom w rows corresponding to P were zeroed out in $\mathbf{F}$). This new p'' results in a new c'' which is different from c (as the scheme is 1-1). Thus, we have a different c''. Note that, $p'' \oplus p$, does not depend on p; and similarly, $c'' \oplus c$ does not depend on p (and not even c). Thus, an adversary with access to a valid c, can come up with a c'' which on decryption leads to all the D's being zero. Thus, $u \geq m - w$. □

We will need yet another combination of equations to prove the next lemma. This time, using the identity matrix in $\mathbf{E}'$ corresponding to the P rows, we now also zero out the corresponding entries in $\mathbf{B}$ and let the new matrix be $\mathbf{H}'$. Thus, $\mathbf{H}'$ is different from $\mathbf{H}$ in only the columns corresponding to C.

Using $\mathbf{H}'$ let's rewrite the equations for D more conveniently:

For $i = 1..u$, let

$$D_i = \sum_{j=1}^{m+k} (a_j^i \cdot M_j) \oplus \sum_{j=1}^{m+k} (b_j^i \cdot N_j) \oplus \sum_{j=1}^{v} (c_j^i \cdot r_j) \oplus \sum_{j=1}^{m-w} (d_j^i \cdot P_j)$$

In the matrix $\mathbf{E}'$, the first $(m-w)$ columns have the rows corresponding to P equal to zero. In a way these columns also work as hidden integrity checks, though not always. So, for $i = u+1..u+m-w$ define D'_i similar to above, using the $(m-w)$ columns of $\mathbf{E}'$ or $\mathbf{H}$.

$$D'_i = \sum_{j=1}^{m+k}(a^i_j \cdot M_j) \oplus \sum_{j=1}^{m+k}(b^i_j \cdot N_j) \oplus \sum_{j=1}^{v}(c^i_j \cdot r_j)$$

We say that N_i and N_j *resolve* if $N_i \oplus N_j$ can be written as a linear combination of only the C's and P's. Similarly, we say that M_i and M_j resolve if $M_i \oplus M_j$ can be written as a linear combination of only the C's and P's.

We will later show that there exists a pair $i, j, i \neq j, i, j \in [1..m+k]$ such that

1. N_i and N_j resolve
2. M_i and M_j resolve
3. For all $x \in [1..u+m-w]$, $a^x_i \oplus a^x_j = 0$, and $b^x_i \oplus b^x_j = 0$
4. exists $y \in [1..m+t]$, $\mathbf{H}'_{2m+k+v+i,y} \oplus \mathbf{H}'_{2m+k+v+j,y} = 1$

In item (4), $\mathbf{H}'_{2m+k+v+i,y}$ is the entry in $\mathbf{H}'$ in row corresponding to N_i and in column corresponding to C_y. Essentially, it says that the rows corresponding to N_i and N_j are not identical (for the first $m+t$ columns).

In the next lemma we show that if such a pair exists with the above four conditions holding then the scheme $\mathbf{G}$ is not message integrity secure.

Lemma 3: If there exists a pair $i, j, i \neq j, i, j \in [1..m+k]$ such that the above four conditions hold than the scheme $\mathbf{G}$ is not message integrity secure.

Proof: We will show that with probability greater then $1 - O(2^{-n})$ there exists a c' (different from a given c) which can easily be computed (given c and the corresponding p) such that

- $N'_i = N_j$
- $N'_j = N_i$
- for z different from i, j, $N'_z = N_z$
- the first $m-w$ blocks of P remain same

We have a similar set of relations for M, and hence given (3), all the D functions would evaluate to zero, leading to $\mathbf{G}$ being insecure for message integrity.

To demonstrate such a c', using $\mathbf{H}'$, we evaluate Δc, for ΔN and ΔM, where

- $\Delta N_j = \Delta N_i = N_i \oplus N_j$
- $\Delta M_j = \Delta M_i = M_i \oplus M_j$

Because of (3) all the D' remain zero, which means there is no change in any other N or M. Moreover the changes above in M and N do not cause any change in the first $m - w$ plaintext blocks (all the changes can be incorporated in the lower w blocks because of the identity matrix in $\mathbf{E}'$). Since the rows corresponding to the bottom w rows of P in $\mathbf{B}$ were zeroed out, these changes in the plaintext do not affect Δc.

Now, ΔN_j is non-zero with probability $1 - 2^{-n}$ (at least). Since M_i is related to N_i by a random function, the probability that ΔM cancels out ΔN in computing Δc is at most 2^{-n}. This leads to a non-zero Δc because of (4) above with probability at least $1 - O(2^{-n})$.

Since conditions (1) and (2) hold as well, an adversary can compute such a c' from c and p. $\Box$

Lemma 4: Either $k + v + u + m - w$ is $\Omega(\log m)$, or the scheme $\mathbf{G}$ is not secrecy secure, or there exists a pair i, j satisfying (1),(2), (3) and (4)

Proof: Recall that,

$$[P_1 ... P_m r_1 ... r_v N_1 ... N_{m+k}\ 1] \cdot \mathbf{B} \ = \ [C_1 ... C_{m+t}]$$

The rank of the matrix $\mathbf{B}$ is at least m, say m'. Now, we call a pair of rows from the rows corresponding to N in $\mathbf{B}$ dependent if one row can be expressed linearly in terms of other using the bottom w rows corresponding to P. This is clearly an equivalence relation. From each such pairwise dependent set (including sets with only one row), pick only one row, and push the remaining rows to the bottom. Let q be the number of rows so picked. The rank of the top $m + v + q$ rows is still at least m'.

Now if we also ignore the top m rows (corresponding to P), the rank of the remaining $v + q$ rows is still m', for otherwise we have a non-trivial linear relationship between C and P, and hence the scheme is not secrecy secure.

This implies (by lemma 2.2) that

$$[r_1 ... r_v N_1 ... N_q] \ = \ [< f(C, P) >] + (\mathrm{GF2})^n \cdot V_1 + ... + (\mathrm{GF2})^n \cdot V_{q+v-m'}$$

where $< f(C,P) >$ is a set of linear functions of C and P, and V_i are linearly-independent binary row-vectors. For a subset of N's with indices a set $J \subseteq [1..q]$ to be pair-wise "non-resolving" thus requires $q+v-m' \geq \log|J|$. In other words, there exists $i,j \in J, i \neq j$, N_i and N_j resolve if $q+v-m' < \log|J|$. Stated differently, there is a set $J1$ of size $|J1| = (q)/2^{q+v-m'}$ in which all pairs of N's resolve with each other.

Now each M_i can be written as a linear combination of r, N and P (using matrix $\mathbf{E}$). Once again (using lemma 2.2) we have

$$[r_1...r_v N_1...N_{m+k}] = [< f'(C,P) >] + (\mathrm{GF2})^n \cdot V_1' + ... + (\mathrm{GF2})^n \cdot V_{m+k+v-m'}'$$

where $V_1'...V_{m+k+v-m'}'$ are linearly-independent binary row vectors. Thus, for any set of indices $J' \subseteq [1..m+k]$, there is a a set $J'' \subseteq J'$ of size $|J''|$ at least $|J'|/2^{m+k+v-m'}$, such that all pairs of Ms in this set J'' resolve with each other.

Using $J1$ for J', thus there is a set $J2$ of size $q/2^{q+v-m'+m+k+v-m'}$ such that for all $i,j \in J2$, M_i and M_j resolve, and so do N_i and N_j.

Similarly, there is a set $J3$ of size $|J3| = |J2|/2^{u+m-w}$ such that

$$\forall k \in [1..u+m-w],\ \forall i,j \in J3 : \ a_i^k \oplus a_j^k = 0$$

Thus, there exists a pair satisfying (1), (2) ,(3) and (4) if $2^{m+k+q+2v-2m'+u+m-w} < q$. Now, $q+v \geq m' \geq m$. Now, either v is $\Omega(\log m)$, or q is at least $\Omega(m)$. Thus, there exists a pair satisfying (1..4) if $m+k+q+2v-2m'+u+m-w < O(\log m)$. Since, $q-m < k$, the previous inequality is implied by $2(k+v)+u+m-w < O(\log m)+2(m'-m)$, which in turn is implied by $2(k+v)+2(u+m-w) < O(\log m)$. Thus, either there exists a pair with (1..4) holding or, $k+v+u+m-w$ is $\Omega(\log m)$. □

Finally, we are ready to prove the main theorem.

Proof (**Theorem 1**): By Lemma 3, since the scheme $\mathbf{G}$ is message integrity secure, there does not exist a pair with conditions (1..4) holding. Thus, by lemma 4, and the fact that $\mathbf{G}$ is secrecy secure, we have $k+v+u+m-w > \Omega(\log m)$. By lemma 1 and 2 it follows that $k+v$ is at least $\Omega(\log m)$. □

References

[1] ANSI X3.106, "American National Standard for Information Systems - Data Encryption Algorithm - Modes of Operation", American National Standards Institute, 1983.

[2] M. Bellare, J. Kilian, P. Rogaway, "The Security of Cipher Block Chaining", CRYPTO 94, LNCS 839, 1994

[3] M. Bellare, C. Namprempre, "Authenticated Encryption: Relations among notions and analysis of the generic composition paradigm", Proc. Asiacrypt 2000, T. Okamoto ed., Springer Verlag 2000

[4] V.D. Gligor, P.Donescu, "Integrity Aware PCBC Encryption Schemes", 7th Intl. Workshop on Security Protocols, Cambridge, LNCS, 1999

[5] V.D. Gligor, P. Donescu, "Fast Encryption Authentication: XCBC Encryption and XECB Authentication Modes", 2000, http://csrc.nist.gov/encryption/modes/workshop1

[6] ISO 8372, " Information processing - Modes of operation for a 64-bit block cipher algorithm", International Organization for Standardization, Geneva, Switzerland, 1987

[7] ISO/IEC 9797, "Data cryptographic techniques - Data integrity mechanism using a cryptographic check function employing a block cipher algorithm", 1989

[8] C. S. Jutla, "Encryption Modes with Almost Free Message Integrity", Eurocrypt 2001, LNCS 2045.

[9] J. Katz and M. Yung, "Unforgeable Encryption and Adaptively Secure Modes of Operation", Fast Software Encryption 2000.

[10] Saunders MacLane, "Algebra", New York : Macmillan (1967)

[11] L.R. Knudsen, X. Lai, B. Preneel, "Attacks on Fasst Double Block Length Hash Functions", Journal of Cryptology, Vol 11, No. 1, Winter 1998

[12] Hugo Krawczyk, "LFSR-based Hashing and Authentication", Proc. Crypto 94. LNCS 839, 1994

[13] C.H. Meyer, S. M. Matyas, "Cryptography: A New Dimension in Computer Data Security", John Wiley and Sons, New York, 1982

[14] National Institute of Standards and Technology, "Recommendation for Block Cipher Modes of Operation", SP 800-38A, http://csrc.nist.gov/publications/nistpubs/index.html

[15] National Bureau of Standards, NBS FIPS PUB 81, "DES modes of operation", U.S. Department of Commerce, 1980.

[16] RFC 1510,"The Kerberos network authentication service (V5)", J. Kohl and B.C. Neuman, Sept 1993

[17] Phillip Rogaway, Mihir Bellare, John Black, and Ted Krovetz, "OCB: A Block-Cipher Mode of Operation for Efficient Authenticated Encryption", Eighth ACM Conference on Computer and Communications Security (CCS-8), ACM Press, pp. 196-205, 2001.