GCM-AES-VR : A Scheme for Cloud Data Confidentiality and Authenticity

Rajani S. Sajjan1*, Vijay R. Ghorpade2

1Department of Computer Science & Engineering, VVPIET, Solapur, MH-India
2Principal, Bharathi Vidyapeeth’s College of Engineering, Kolhapur, MH-India

*Corresponding Author: raajanisajjan78@gmail.com, Tel.: 08308602157

Available online at: www.ijcseonline.org

Abstract— Cloud data security is recognized as making the data confidential along with proper authentication. The Galois/Counter Mode (GCM) is used to provide data confidentiality with associated data as authentication. It aims to provide birthday bound security i.e. it is secure up to $2^{n/2}$ adversarial queries where $n$ is a block size. But in some cases this much security is not sufficient. In this paper, we have proposed a new approach to authenticated encryption with associated data (AEAD), an improved AEAD scheme which can be secure up to approximately $2^n / p$ adversarial queries where, $p = (n/m)$, where $n$ is a block size and $m$ is a bit variance. This bit variance is introduced in the encryption process. In the proposed nonce-respecting AEAD scheme a new pseudorandom function is defined and used for implementation. To generate authentication tag universal hash function is used. In this paper security proofs of proposed scheme are given by presenting its construction and its security model.

Keywords—Authenticated encryption with associated data, beyond birthday bound security, cloud data confidentiality, data authentication

I. BACKGROUND

Cloud data security is recognized as making the data confidential along with proper authentication. This leads to many solutions. Traditionally data was encrypted and message authentication code glued with it [1]. Even though this solution seems feasible it has drawbacks. It contains two parts for the same solution. This wasn’t sufficient to make it applicable as a perfect solution for providing confidentiality and authenticity to data. This brought recognition to authenticated encryption with associated data (AEAD) schemes. Jutla [2], V. Gligor and P. Donescu [3] and P. Rogaway, M. Bellare, J. Black, and T. Krovetz [4] have come up with more standard solution for data privacy and authenticity as an integrated solution named as AEAD schemes.

In this paper we have proposed an AEAD scheme, Galois/Counter Mode (GCM) using variant of advanced encryption standard (AES-VR), named GCM-AES-VR. This scheme aims to achieve beyond birthday bound security. The paper is organized in different sections, the first section gives brief introduction of AEAD, GCM and discusses about the background. The second section defines the problem statement followed by the preliminaries used throughout the paper. The construction of proposed scheme is given in the fourth section where the key derivation process and the working of proposed scheme are discussed in detail. Algorithms are given with block diagram of proposed scheme. Further security analysis is done for data privacy and data authenticity. The proposed scheme is compared with existing systems and it is described in detail in the last section of the paper followed by conclusion and future work.

A. Authenticated Encryption with Associated Data (AEAD): AEAD scheme provides data confidentiality, integrity and authenticity under one umbrella. It applies encryption process and hash function to achieve this. It takes input as plaintext and provides a packet containing ciphertext and hash tag. There are various authenticated encryption modes developed. All of them work with symmetric block ciphers. In general AEAD scheme is secure if encryption scheme used is semantically secure under chosen plaintext attack and MAC function used is unforgeable under chosen message attack [5]. We can categorize various AEAD schemes into three classes depending on their working paradigm.

a. Encrypt then MAC
b. Encrypt and MAC
c. MAC then encrypt

Table 1 provides brief information about these schemes. Out of these three, encrypt then MAC is more secure AEAD approach, if chosen MAC scheme is unforgeable.
CAESAR competition is started in 2012 which promoted many researchers to develop AE schemes. These schemes have been applied in various fields. Depending on application requirements AE schemes can be classified as
a. AE with associated data
b. Parallelizable AE
c. Online AE
d. tweakable AE
e. Deterministic AE
f. wide block AE
g. XOR based AE
h. Dedicated AE algorithms

<table>
<thead>
<tr>
<th>Table 1: Different approaches to AEAD scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AEAD Approach</strong></td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Encrypt then MAC</td>
</tr>
<tr>
<td>Encrypt and MAC</td>
</tr>
<tr>
<td>MAC then encrypt</td>
</tr>
</tbody>
</table>

Depending on the design approach used AE schemes can be classified as
a. Generic composed AE
b. Block cipher based AE
c. Stream cipher based AE
d. Permutation based AE
e. Keyed function based AE
f. tweakable block cipher based AE
g. Hybrid AE


- **Birthday bound security:** The given encryption scheme is said to be birthday bound secure if it is secure up to $2^{n/2}$ adversarial queries, where $n$ is the block size.
- **Beyond birthday bound (BBB) security:** The given encryption scheme is said to be BBB secure if it is provably secure up to approximately $2^{r(n/(r+1))}$ adversarial queries, where $r >= 2$ is an integer.
- **Optimal security:** The given encryption scheme is said to be optimal secure if it is provably secure up to $2^n$ adversarial queries.

In general, Advanced Encryption Standard (AES) is used as the block cipher encryption algorithm which has block size of $n = 128$ bits, and when it is used with AE schemes its security degrades to $n/2$ i.e. 64 bits. AE modes involve multiple block invoking operation during execution; this causes extra overhead on existing hardware and software. This overhead increases with increase in block invoking. In case of BBB secure AE schemes number of block invoking may vary and accordingly overhead also varies.

**B. Galois/Counter Mode (GCM):**
GCM is a block cipher mode of operation designed by McGrew and Viega [11]. It is nonce based AEAD scheme. GCM uses universal hash function for authentication part and counter mode in encryption part. It performs two operations, authenticated encryption and authenticated decryption. The authenticated encryption operation takes four inputs viz., secret key K, initialization vector IV whose length can vary from 1 to 264, 96 bit IV values are recommended in critical cases due to its efficient processing. The authenticated encryption process produces ciphertext C of length P and authentication tag T of any length between 0 to 128. The authenticated decryption operation takes five inputs: K, IV, C, A, T. It produces only one output i.e. plaintext P or FAIL otherwise. FAIL output indicates that the operation is not authentic means whatever inputs were provided to authenticated decryption operation were not generated by authenticated encryption operation. The A, additional authenticated data (AAD) provides protection to the data which needs to be authenticated.

GCM is proved secure in concrete security models against chosen plaintext attack. Security of GCM relies on the fact that underlying block cipher is a random permutation. GCM uses universal hash function for strong authentication, which is a strong element used in many cryptosystems. Additionally GCM uses counter mode which is proven as strongly secure.

**II. PROBLEM STATEMENT**

GCM provides security up to $2^{n/2}$ adversarial queries. When used with AES its security drops down to 64 bits which is not acceptable in some critical operations. This security obtained using GCM is based on the assumption that underlying block cipher is a secure pseudorandom permutation. In this chapter we are proposing a GCM AEAD scheme which addresses improvement in the security bits as compared to GCM-AES AEAD scheme.

**III. PRELIMINARIES**

In this section we are going to discuss the terms and notations used throughout the paper.
A. Notations:
Assume \{0,1\} be the set containing strings including 0, empty string. For any string y, |y| indicates length of string y, such that y \in \{0,1\}^n. \|y\| indicates number of blocks of size n bits each, of input string y, where block size n >= 1. The n bit zero string is denoted as 0^n \in \{0,1\}^n. For any two strings y and z, concatenation is denoted by y \oplus z i.e. yz. To denote XOR operation of y and z strings y \oplus z notation is used.

For finite set Y, y is drawn from Y and it is denoted as y \leftarrow Y, here y is randomly drawn from Y. The number of elements in Y is denoted as |Y|. Adversary A produces output as 1 after oracle interaction, where oracle is indicated by O.

Let K_e be the key space, such that Ke \in \{0,1\}^n. Assume K is a key used in authenticated encryption process, where K \in K_e. Plaintext M is given for processing where M \in \{0,1\}^n. The authenticated encryption operation can be written as: Ke \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n. For any given K \in K_e, a fixed value, the encryption operation E_K : \{0,1\}^m \rightarrow \{0,1\}^n is n-bit permutation function. Its inverse i.e. D_K is written as D_K = E_K^{-1}. Assume that Permutation(n) is the set of all n-bit permutations. Adversary A has access to encryption oracle.

Let K \leftarrow K_e and \rho \leftarrow Permutation(n) then pseudorandom permutation (PRP) advantage of adversary A against E is defined as

\[Adv_{PRP}^E (A) = \left| \Pr[K \leftarrow K_e : E_K \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n] - \Pr[K \leftarrow K_e, \rho \leftarrow Permutation(n) : E_K \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n] \right|\]

Here probabilities are taken randomly over K and \rho and adversary A. If \[Adv_{PRP}^E (A)\] is negligible then the underlying block cipher E_K is considered as a secure pseudorandom permutation.

A function F_u is a keyed function, such that F_u : Kfu \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n. Where it takes input as a key K and plaintext M and returns ciphertext C; K \in Kfu, M \in \{0,1\}^m and C \in \{0,1\}^n.

When any fixed value is taken as a K \in Kfu; Fu is a function from \{0,1\}^m to \{0,1\}^n which can be represented as Fu_{uk} : \{0,1\}^m \rightarrow \{0,1\}^n. Assume that Fu(m,n) is a set of all functions from \{0,1\}^m to \{0,1\}^n. We write Fu(m) if m = n occurs. Adversary A has access to encryption oracle.

Let K \leftarrow K_e and \eta \leftarrow Fu(m,n) then pseudorandom function (PRF) advantage of adversary A against F_u is defined as

\[Adv_{PRF}^u (A) = \left| \Pr[K \leftarrow K_e : F_{uK} \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n] - \Pr[K \leftarrow K_e, \eta \leftarrow Fu(m,n) : F_{uK} \cdot X \cdot \{0,1\}^m \rightarrow \{0,1\}^n] \right|\]

Here probabilities are taken randomly over K and \eta and adversary A. If \[Adv_{PRF}^u (A)\] is negligible then the underlying block cipher F_{uK} is considered as a secure pseudorandom function.

B. Hash function:
The hash function used here is universal hash function. It is denoted by H. This keyed hash function takes two inputs: one is key and another is message. It returns one output. Input key K_h \in K_h and input message y \in \{0,1\}^m is given and output z \in \{0,1\}^n is given by H. The H is said to be almost XOR universal hash function if for y \in \{0,1\}^m and z \in \{0,1\}^n,

\[\Pr[K_h \leftarrow K_h : H_{kh} (y) = z] \leq \alpha\]

For y \in \{0,1\}^m and y' \in \{0,1\}^m where y and y' are two distinct values and z \in \{0,1\}^n

\[\Pr[K_h \leftarrow K_h : H_{kh} (y) \oplus H_{kh} (y') = z] \leq \beta\]

Where H(\alpha,\beta) is almost XOR universal hash function.

C. Galois field (GF):
The finite field GF(2^n) can be the set from \{0,1\}^n. Assume there is a n bit string b = b_n b_{n-1} \ldots \ldots b_0 \in \{0,1\}^n string b can be defined with the help of polynomial b(x) \in \mathbb{R}[x] by b(x) = b_n x^{n-1} + \ldots + b_1 x + b_0 where b_i \in \{0,1\} for any i \in \{0,n-1\}. All integers between 0 and 2^n-1 can be viewed as polynomial of binary coefficients with maximum degree of n-1. As example, 1 is represented as x, 3 is represented as x+1, 6 is represented as x^2 + x.

In GF addition in the field GF(2^n) is the polynomial addition operation over GF(2^n). This operation is denoted by \oplus. For example addition of p and q is denoted as p \oplus q where p,q \in GF(2^n).

In GF multiplication over GF(2^n) is quite complicated as compared with addition operation. Here we have to use irreducible polynomial f(y) over the GF(2^n) of degree n. For n = 128, f(y) = y^{128} + y^7 + y^2 + y + 1. The multiplication of B and D: B, D \in GF(2^n); is multiplication over GF(2^n) reduced modulo f(y), i.e. B(y)D(y) mod f(y).

D. Authenticated encryption and authenticated decryption operation:
An authenticated encryption with associated data scheme consists of an authenticated encryption and decryption operation.

\[E : K \cdot XN \cdot XM \cdot XH \rightarrow C \cdot XT\]
\[D : K \cdot XN \cdot XH \cdot XC \cdot XT \rightarrow M \cdot U \{ \square \}\]

i.e.
\[E(K, N, A, M) = E_u(N, A, M) \rightarrow (C, T)\]
\[D(K, N, A, C, T) = D_u(N, A, C, T) \rightarrow M \square\]

Where,
Key K \in K,
Nonce N \in N,
Associated data A \in H, where H \in \{0,1\}^*
Plaintext M \in M, where M \in \{0,1\}^*
Tag T \in T, where T \subseteq \{0,1\}^*
Error symbol \square, indicates failure of decryption operation.
\[E_u(N, A, M) \rightarrow (C, T) \text{ iff } D_u(N, A, C, T) \rightarrow M.\]
When \( D_k \) returns \( \square \), it indicates that input parameters given to decryption operation are not generated by encryption operation for the same key. This sort of output proves that AEAD scheme is secure.

E. Pseudorandom function:
In the proposed model the pseudorandom function (PRF), \( F_1 \), is used by using multi encrypted Davies Meyer model. Let \( Q_1 \) and \( Q_2 \) are the two pseudorandom permutations on \( n \) bits. These permutations work independently. The function \( F_1 \) is defined as

\[
F_1 : \{0,1\}^l \rightarrow \{0,1\}^{n\times S}, F_1(x) = (y_1, y_2, \ldots, y_s). y_i = Q_2(Q_1(inc^c(x \ || \ [0]_{n-1})) \oplus inc^c(x) \in \{0,1\}^n
\]

For \( i \in [1, s] \) and \( s \leq 2^{n/4} - 1, 1 \leq n \) and \( x \in \{0,1\}^l \)

The information theoretic security of function \( F_1 \) is given below.

Theorem \( F_1 \):
Let adversary has access to the function \( A \) is an adversary and \( F_1 \) is the function. Assuming that \( A \) makes \( q \leq 2^n / p \) where \( q \) represents oracle queries and generates \( \square = q \) blocks, then PRF-advantage of adversary, \( A \) against function, \( F_1 \), is upper bounded by,

\[
Adv_{F_1}^{PRF}(A) \leq \frac{n}{2^n}
\]

Above theorem \( F_1 \) proves function \( F_1 \) constructed using \( Q_1 \), \( Q_2 \) achieves beyond birthday bound security. The \( p = (n/m) \), where \( m \) can take values from the set \( \{8, 16, 32\} \) and \( 1 \leq p \leq 128 \). If \( m = 16 \) then PRF advantage of adversary against \( F_1 \) is a close to optimally secure PRF up to \( q \leq 2^n / p \) adversarial queries.

Proof of theorem \( F_1 \):
Assume \( S \leftarrow \text{fun} \ (l, ns) \) and \( s \leftarrow \text{fun} \ (l, n) \) are the functions. The PRF advantage of adversary \( A \) against function \( F_1 \) is given as,

\[
Adv_{F_1}^{PRF}(A) = \left| Pr[A^{F_1} = 1] - Pr[A^{S} = 1] \right|
\]

Consider EDM construction in the reduced form with \( n \cdot l \) fixed bits and it is denoted as \( f : \{0,1\}^l \rightarrow \{0,1\}^n \). Assume that there is another adversary \( G \) which has access to reduced EDM construction. Also \( G \) has access to \( s \), a random function, and it builds queries for \( f \).

According to EDM construction security, if \( q \leq 2^n / 67\sigma^2 \) and \( \sigma \geq 2 \) we have

\[
Adv_{f}^{PRF}(G) = \left| Pr[G^{f} = 1] - Pr[G^{s} = 1] \right|
\]

Then we construct a hybrid function \( H_j \), where \( i \) indicates \( f \) functions and \( j \) indicates \( s \).

\[
H_j = (f, \ldots, f, s, \ldots, s)
\]

The \( (f, \ldots, f, s, \ldots, s) \) is recognized as function \( f \), and \( (s, \ldots, s) \) is \( s \).

If \( i = 0 \) then, \( H_j = (s, \ldots, s) \) is nothing but \( s \).

If \( i = j \) then, \( H_j = (f, \ldots, f, s, \ldots, s) \) indicates \( F_1 \).

The PRF advantage of adversary \( A \) against function \( F_1 \) is upper bounded by,

\[
Adv_{F_1}^{PRF}(A) = \left| Pr[A^{F_1} = 1] - Pr[A^{S} = 1] \right|
\]

\[
= \left| Pr[A^{H_j = 1}] - Pr[A^{H_j = 1}] \right|
\]

\[
\leq \sum_i \left| Pr[G^{f} = 1] - Pr[G^{s} = 1] \right|
\]

\[
\leq \sum_i \frac{q}{2^n} + \frac{(\frac{q}{2^n})}{2^n}
\]

Where inequality is obtained by \( \sum_i q_i = \square \). Proof is over.

IV. CONSTRUCTION OF PROPOSED SCHEME

The proposed scheme has two major parts. The first part is implementation of variant of Advanced Encryption Standard (AES) i.e. AES-VR and second part is application of AES-VR with GCM i.e. GCM-AES-VR. The first part is already implemented, tested and its security is verified against the standard security measurement criteria; avalanche effect and strict avalanche criteria. This work is accepted for publication [12]. Here in this paper we are presenting the second part of using AES-VR with GCM to provide BBB security with authentication. In the following section AES-VR is discussed in short followed by construction of GCM-AES-VR.

A. AES-VR:
It is a variant of advanced encryption standard. This scheme is proposed and implemented to improve the data confidentiality in distributed environment. The security of AES is closely related with strength of key schedule process and S-box. Since AES S-box is tested thoroughly under various attacks of the form linear, differential and algebraic, changing S-box completely doesn’t seem appealing. Additionally dynamic generation of key dependent S-box makes the process lengthy. By considering these aspects in AES-VR we have proposed changes for the key schedule process. The construction of AES-VR, its algorithm is given in detail along with security analysis and security measure [12]. AES-VR provides \( 2^{(m+log(2^n, 2^m))} \) variations where \( m \) represents key size; \( n \) represents bit variance induced in data.

B. GCM-AES-VR:
GCM-AES-VR is a proposed nonce based AEAD scheme. It is based on encryption then MAC approach. GCM-AES-VR is a variant of GCM which aims to achieve BBB security. The PRF \( F_1 \) discussed in the previous section is used in the construction of GCM-AES-VR. It has two passes encompassing encryption operation which uses PRF \( F_1 \) and universal hash function and EDM construction is used to generate an authentication tag.
C. Working of GCM-AES-VR:
It has two main operations to perform, encryption and
decryption, described as below. For performing two
encryption operations with one universal hash function we
need three keys viz K₁, K₂, K₃ as input. Instead of getting
these three keys from user as inputs here the proposed
module itself generates them by just taking one key, K, as
input from user. This key derivation algorithm is described
below. This keeps proposed system key management free.
The key Kᵢ is used for first pass of encryption and key K₂
for second pass of encryption process. The third key K₃ is
used by universal hash function to generate tag value. The
underlying block cipher \( E : K \times \{0,1\}^n \rightarrow \{0,1\}^n \) is used
in the model.

D. Key derivation function:
The construction of proposed model involves two
encryption operations, which need two different security
keys. Additionally one more key is used for generating hash
value of data. So in all three secret keys are needed to
complete the process. In the proposed model we have made
user to enter only one key called master key, and this key is
processed to generate three secret keys by adding other
parameters. This operation is depicted as shown in figure 1.
As shown in the figure 1, Kₘ is a master key entered by
user. This key is given to a key derivation function fᵢ
alongwith constant parameter Cᵢ. Kₘ ∈ \{0,1\}ⁿ and n ≥ 1. fᵢ
is a set of functions used to generate kᵢ. The kᵢ where i = 1 to
m, m is chosen as per the function requirement, here it is
taken as 3.

These functions can be written as
\[
\begin{align*}
K₁ &= f₁(Kₘ, C₁, N, P, R, 128) \\
K₂ &= f₂(Kₘ, C₂, N, P, R, 128) \\
K₃ &= f₃(Kₘ, C₃, N, P, R, 128).
\end{align*}
\]
The kᵢ belongs to the key space kᵢ ∈ \{0,1\}ˡ and l = 128bits. The Cᵢ represents constant values those
are already set by proposed model itself, but one can make
them get generated dynamically. Here Cᵢ ∈ \{0,1\}ᵇ and b ≥ 1.
The key generation process is elaborated in the algorithm
1 given below.

E. Encryption operation:
The encryption operation is represented as,
\[
E : K \times N \times M \times H \rightarrow C \times T
\]
where,
Key K ∈ Ke X Kh = \{0,1\}ᵇ is the key space,
Nonce N ∈ \{0,1\}ˡ is the nonce space,
Associated data space H ⊆ \{0,1\}*,
Plaintext space M ⊆ \{0,1\}*,
Ciphertext space C ⊆ \{0,1\}*,
Tag space T ⊆ \{0,1\}*.
The encryption operation takes plaintext M ∈ M, key K ∈
K, nonce N ∈ N, and associated data A ∈ H as inputs and
returns C ∈ C as the ciphertext and tag T ∈ T.

F. Decryption operation:
This operation is inverse encryption operation. Decryption
operation is represented as
\[
D : K \times N \times X \times C \times T \rightarrow M \cup \{ \square \}
\]
The decryption operation takes key K, nonce N, associated
data A, ciphertext C, and tag T as inputs and returns M
plaintext or \( \square \) error as output. Error symbol \( \square \), indicates
failure of decryption operation.

Algorithm 1: KeyGenProcess( )

Inputs:
Passphrase: Km ; string of characters to be hashed
Salt: for deriving 3 keys here three constants are set
viz. C₁, C₂ and C₃ of any arbitrary length
CostFactor (N): CPU/memory cost parameter
BlockSizeFactor (r): blocksize parameter
ParallelizationFactor (p): Parallelization parameter
DesiredKeyLen: Desired key length in bytes i.e. 128
bits; 16 bytes
Output:
DerivedKey of DesiredLen i.e. 16 bytes

Process:
Get master key i.e. user entered value, Kₘ
\[
\begin{align*}
K₁ &= f₁(Kₘ, C₁, N, P, R, 128) \\
K₂ &= f₂(Kₘ, C₂, N, P, R, 128) \\
K₃ &= f₃(Kₘ, C₃, N, P, R, 128).
\end{align*}
\]
k₁ is used for first encryption process
k₂ is used for second encryption process
k₃ is used in tag value calculation

V. GCM-AES-VR: AEAD SCHEME
The proposed AEAD scheme, GCM-AES-VR, is defined as,
\[
\|J = (Eₖ(............)), Dₖ(............))\]
is a nonce based AEAD scheme. Where K is a key used in
encryption operation and K ∈ \{0,1\}ᵇ. The general
construction of GCM-AES-VR AEAD scheme is
represented in figure 2 as shown. Here two encryption
operations are performed, Eₖ₁ and Eₖ₂, by getting secret
keys k₁ and k₂ generated in key derivation process.
Figure 2 Construction of proposed GCM-AES-VR AEAD scheme

Algorithm 2: Encrypt_GCM_AES_VR()

Input: K₁, K₂, K₃ three keys derived from master key K, associated data Aᵦ, nonce value N, plaintext M
Output: ciphertext C and tag value T

Divide plaintext M into equal blocks i.e. M₁ || M₂ || ....... || Mᵦ such that |Mᵦ| = n for 1 ≤ j ≤ m-1;
P₀ = N || 0^{n/2} 1;
Q₀ = E_K₁( M₁ ⊕ E_K₁(P₀) );
for j = 1 to m do
    Pⱼ = increment(Pⱼ-1);
    Qⱼ = E_K₁( Mⱼ ⊕ E_K₁(Pⱼ) );
for j = 1 to m-1 do
    Cⱼ = Qⱼ;
C = C₁ || C₂ || ....... || Cᵦ;
Q = Q₀ ⊕ Hₖₙ(Aₚ, C);
T = Q₀;
return (C || T);

Algorithm 2: Encryption algorithm of GCM-AES-VR

The encryption and decryption operation performed is described in algorithm 2 and 4 respectively. The hash algorithm used to generate hash value is given in algorithm 3.

Algorithm 3: HashValue()

Input: Kₖ, key derived from master key Kₖ, associated data Aₚ, ciphertext C
Output: Hash value H

Yₐₚ = Aₚ || 0^{Ad[m-n]} || C || 0^{C[n]} || C || 0^{C[n]};
Divide Y into blocks i.e. Y₁ || Y₂ || ....... || Yᵦ such that |Yᵦ| = n for 1 ≤ j ≤ m;
H₀ = 0;
for j = 1 to m
    H = (H ⊕ Yⱼ) · Kₖ
return H

Algorithm 3: Hash algorithm

Algorithm 4: Decrypt_GCM_AES_VR()

Input: K₁, K₂, K₃ three keys derived from master key K, associated data Aₚ, nonce value N, ciphertext C and tag value T
Output: Plaintext M or ⊥ error symbol

P₀ = N || 0^{n/2} 1;
Q₀ = E_K₂( M₁ ⊕ E_K₂(P₀) );
for j = 1 to m do
    Pⱼ = increment(Pⱼ-1);
    Qⱼ = E_K₂( Mⱼ ⊕ E_K₂(Pⱼ) );
for j = 1 to m-1 do
    Mⱼ = Qⱼ ⊕ Cⱼ;
    Mᵦ = Mᵦ || Mᵦ || ....... || Mᵦ;
return (M);
else
return (⊥);

Algorithm 4: Decryption algorithm of GCM-AES-VR

VI. GCM-AES-VR SECURITY ANALYSIS

Security analysis of AEAD scheme considers two aspects viz. privacy and authenticity. In the proposed AEAD scheme these security aspects are discussed below.

Privacy:
Assume that Φ(.............) is a random oracle which takes inputs as (N,A,M) and returns a random string of length |C| + |T|. Assume that there is an adversary A having access to one of the two oracles, Eₖₙ (.............) or random oracle Φ(.............). If all nonces N¹,N₂,...,Nᵦ are always different for all encryption queries i.e. (N¹,A¹,M¹)....(Nᵦ, Aᵦ, Mᵦ) then we say that adversary A is nonce respecting.
We assume that, without loss of generality, adversary A is nonce respecting and it never generates queries with known responses. The privacy advantage of adversary A against Φ = (Eₖₙ(.............),Dₖₙ(.............)) is given as

Advₚₕ(A) = | Pr[A Eₖₙ(.............) = 1 ] − Pr[A Φ(.............) = 1 ] |

Authenticity:
Assume that there is an adversary \( A \) having access to two oracles, \( E_K(\ldots) \) and \( D_K(\ldots) \).

**Step 1:**
The adversary \( A \) puts queries \((N^w, A^w, M^w)\) where \( w \in [1,q] \). Then \( A \) returns \((C^w, T^w) = E_K(N^w, A^w, M^w)\).

**Step 2:**
The adversary \( A \) puts a challenge query \((N, A, C, T) \not\in \{ (N^w, A^w, C^w, T^w) \}_{w=1}^q \) to \( D_K(\ldots) \).

If \( D_K(N, A, C, T) \not= \square \) then we say that the forgery attempt made by \( A \) is successful.

We assume that, without loss of generality, adversary \( A \) is nonce respecting and it never generates queries with obviously known responses. The authenticity advantage of adversary \( A \) against \( || \) = \( (E_K(\ldots), D_K(\ldots)) \) is given as

\[
Adv_{\text{auth}}^A(\cdot) = \Pr[A \text{Ek}(\ldots)D^\square \text{K}(\ldots) \text{Forges}].
\]

**Security proofs:**

Here we assumed that the underlying block cipher \( E \) is a secure pseudorandom permutation. As per the perspective of information theoretic security, the proposed model achieves close to optimal security i.e. proposed model is provably secure up to \( 2^{n/p} \) adversarial queries, provided underlying block cipher used is secure PRP.

Given below is the privacy proof of GCM-AES-VR.

**Theorem S1:** Assume \( E : K \rightarrow X \{0,1\}^n \rightarrow \{0,1\}^n \) be a block cipher and \( H : K \rightarrow X \{0,1\}^n \rightarrow \{0,1\}^n \), and \( K_e \) and \( K_h \) are two nonempty sets such that \( K_e, K_h \in \{0,1\}^n \). Then there is a nonce respecting adversary \( A \) which generates at most \( q \leq 2^{n/p} \) adversarial queries. It assumes block length of \( m \) bits and execution time as \( t_1 \). Now assume that there is another adversary \( B \) against \( E \) which makes at most \( \delta = q(m + 1) \) oracle queries such that for any adversary \( A \),

\[
Adv_{\text{privacy}}^A(\cdot) \leq 2 Adv_{E}^{pr}(B) + \frac{\delta}{2^n}
\]

The proof takes two steps, as discussed below.

**Step 1:**
Replace encryption operations with random and independent permutations. \( E_{k_1} \) and \( E_{k_2} \) are replaced with \( P_1 \) and \( P_2 \). The keys \( k_1 \) and \( k_2 \) are generated accordingly from \( K \).

**Step 2:**
Now we construct GCM-AES-VR[\( Q \)] as a new module where \( Q = (P_1, P_2) \). By applying hybrid argument we will show that there is another adversary \( B \) against PRP security of \( E \) which makes at most \( \hat{\delta} = q(m+1) \) oracle queries such that

\[
Adv_{\text{privacy}}^{GCM-AES-VR(E)}(A) \leq 2 Adv_{E}^{pr}(B) + Adv_{\text{privacy}}^{GCM-AES-VR(Q)}(A)
\]

Now we will upper bound \( Adv_{GCM-AES-VR(Q)}^{\text{privacy}}(A) \). To do this following Lemma is introduced.

**Lemma 1:**
Assume that from Permutation\( (n) \), \( P = (P_1, P_2) \) these permutations are chosen randomly and independently. And assume that \( A \) is a nonce-respecting adversary which makes at most \( q \leq 2^{n/p} \) adversarial queries to GCM-AES-VR[\( Q \)] generating at most \( \hat{\delta} \) blocks. Then for any adversary \( A \).

\[
Adv_{GCM-AES-VR(Q)}^{\text{privacy}}(A) \leq \frac{\delta}{2^n}
\]

Proof:
We are going to use contradiction argument to prove Lemma 1. Assume that \( E \) [\( Q \)] is the encryption algorithm of GCM-AES-VR[\( Q \)] and \( \exists \), a random function that takes \( (N, A, M) \) as input and gives output as a random string of length \( |C| + |T| \).

Now assume that there is a nonce respecting adversary \( A \) against GCM-AES-VR[\( Q \)] such that,

\[
Adv_{GCM-AES-VR(Q)}^{\text{privacy}}(A) = \left| \Pr[A \text{Ek}(\ldots) = 1] - \Pr[A \exists (\ldots) = 1] \right| \leq \frac{\delta}{2^n}
\]

where adversary \( A \) makes \( q \) queries for \( m \) block length to GCM-AES-VR[\( Q \)]. This generates \( \delta = q(m + 1) \) blocks.

Let \( b_1 \) is a random function such that, \( b_1 \in \text{Function}(1, ns), s = m + 1 \). Now let consider that there exists another adversary \( B \) which generates \( q \) queries to an oracle \( OR \), either \( F_1 \) or \( F_2 \), generating \( \delta = qs \).

If \( OR \) is a \( F_1 \) then encryption operation at \( B \) is same as encryption operation at \( A \). This means that

\[
\Pr[B \text{Ek}(\ldots) = 1] = \Pr[A \text{Ek}(\ldots) = 1]
\]

In the same way if \( OR \) is \( b_2 \), then \( B \) gives same functioning for the random function \( \exists \) for \( A \). Hence

\[
\Pr[B \exists (\ldots) = 1] = \Pr[A \exists (\ldots) = 1]
\]

which says that

\[
Adv_{F_2}^{pr}(B) = \left| \Pr[B \text{Ek}(\ldots) = 1] - \Pr[B \exists (\ldots) = 1] \right| \leq \frac{\delta}{2^n}
\]

this contradicts with theorem 1. Hence contradiction hypothesis does not hold. So we say that,

\[
Adv_{GCM-AES-VR(Q)}^{\text{privacy}}(A) > \frac{\delta}{2^n}
\]

At this point Lemma 1 is over. So we conclude that privacy of GCM-AES-VR is secure up to \( q = 2^{n/p} \) adversarial queries by considering that underlying block cipher is a secure PRP in the nonce respecting scenario. In the following section authenticity of GCM-AES-VR is discussed.

**Theorem S2:**
Assume \( E : K_e X \{0,1\}^n \rightarrow \{0,1\}^n \) be a block cipher and \( H : K_h X \{0,1\}^n \rightarrow \{0,1\}^n \) is an universal hash function, and \( K_e \) and \( K_h \) are two nonempty sets such that \( K_e, K_h \in \{0,1\}^n \). Let there is a nonce respecting adversary
A which generates at most $q \leq 2^{n/p}$ adversarial queries. It assumes block length of $m$ bits and execution time as $t_1$. Now assume that there is another adversary $B$ against PRP security of $E$, which makes at most 
$$\tilde{\theta} = (q + 1)(m + 1)$$
oracle queries such that for any adversary $A$:

$$\text{Adv}^{\text{auth}}_{\text{GCM-AES-VR}(E)}(A) \leq 2\text{Adv}^{\text{prp}}_{E}(B) + \frac{d}{2n} + 6 + \frac{n}{2n} + \frac{1}{2^{p-q}}$$

The proof of above theorem is given below.

**Step 1:**
Replace encryption operations with random and independent permutations. $E_{k1}$ and $E_{k2}$ are replaced with P1 and P2. The keys $k_1$ and $k_2$ are generated accordingly from $K$.

**Step 2:**
Now we construct $\text{GCM-AES-VR}(Q)$ as a new module where $Q = (P1, P2)$. By applying hybrid argument we will show that there is another adversary $B$ against PRP security of $E$ which makes at most $\tilde{\theta} = (q + 1)(m + 1)$ oracle queries such that

$$\text{Adv}^{\text{auth}}_{\text{GCM-AES-VR}(E)}(A) \leq 2\text{Adv}^{\text{prp}}_{E}(B)$$

Now we will upper bound $\text{Adv}^{\text{auth}}_{\text{GCM-AES-VR}(Q)}(A)$. To do this following Lemma is introduced.

**Lemma 2:** Assume $\rho > 0$ is an integer. Let $Q = (P1, P2)$ and $P1, P2$ are two permutations chosen randomly and independently from Permuation$(n)$. Additionally there is $H$ an universal hash function, $\delta$. There is a nonce respecting adversary $A$ which makes $q \leq 2^{n/p}$ adversarial queries and one forgery attempt to $\text{GCM-AES-VR}(Q)$ generating at most $\tilde{\theta}$ blocks. Then for any adversary $A$:

$$\text{Adv}^{\text{auth}}_{\text{GCM-AES-VR}(Q)}(A) \leq \frac{\tilde{\theta}}{2n} + 6 + \frac{n}{2n} + \frac{1}{2^{p-q}}$$

**Proof:**
Assume that there is a nonce respecting adversary $A$ having access to two oracles, $E_K(............)$ and $D_K(............)$. The adversary $A$ puts queries $(N^w, A^w, M^w)$ where $w \in [1,q]$. Then $A$ returns $(C^w, T^w) = E_K(N^w, A^w, M^w)$. The adversary $A$ makes a forgery attempt $(N, A, C, T) \notin [(N^w, A^w, C^w, T^w)]_{w=1}^q$ to $D_K(............)$. As per the definition of authenticity discussed earlier,

$$\text{Adv}^{\text{auth}}_{\text{GCM-AES-VR}(E)}(A) = Pr[A^{EK(Q,DK(Q))} \text{ forges}]$$

$$\leq Pr[A^{EK(Q,DK(Q))} \text{ forges}] - Pr[A^{\oplus1,\oplus2} \text{ forges}] + Pr[A^{\oplus1,\oplus2} \text{ forges}]$$

$E_K(Q)$ and $D_K(Q)$ are encryption and decryption operations of $\text{GCM-AES-VR}(Q)$. $\oplus_1$ and $\oplus_2$ are the random oracles where $\oplus_1$ is a random oracle which always returns a random string $(C,T) \leftarrow C \times T$ and $\oplus_2$ is a random oracle which returns a random string or a reject symbol, i.e., $M/\square \leftarrow M$.

$$\leq Pr[A^{EK(Q,DK(Q))} \text{ forges}] - Pr[A^{\oplus1,\oplus2} \text{ forges}] = \text{Adv}^{\text{privacy}}_{\text{GCM-AES-VR}(Q)}(q + 1, \tilde{\theta}) \leq \frac{d}{2^\theta}$$

This is shown in privacy proof and $\tilde{\theta} = (q + 1)(m + 1)$. So we conclude that authenticity of $\text{GCM-AES-VR}$ is secure up to $q \approx 2^{n/p}$, where $P = n/m$, $n$ indicates block size and $m$ shows bit variance, adversarial queries by considering that underlying block cipher is a secure PRP in the nonce respecting scenario.

## VII. COMPARATIVE ANALYSIS

In this paper we have proposed an improved AEAD scheme, $\text{GCM-AES-VR}$, which has achieved optimal security, $2^n$, where $n$ is length of block. The proposed scheme provides $2^{n/p} \approx 2^{n/p}$, where $m$ indicates bit variance. Table 2 shows the comparative analysis of $\text{GCM-AES-VR}$ with other schemes proposed by researchers with respect to mentioned parameters. The proposed AEAD scheme takes only one key from user and computes three keys required for two encryption steps and hash tag calculation. In addition to this, key management task can be avoided by implying the partial key outsourcing method. The proposed scheme has achieved BBB security.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Key size</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Block size $n$</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Keys used</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Key management</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Block cipher calls</td>
<td>1</td>
<td>$2a + 2$</td>
<td>$2a + 2$</td>
<td>$2a$</td>
</tr>
<tr>
<td>Security bits</td>
<td>64</td>
<td>$\equiv 107$</td>
<td>$\equiv 121$</td>
<td>$\equiv 124$</td>
</tr>
</tbody>
</table>

$a$ indicates plaintext block length

## VIII. CONCLUSION

In this paper a novel approach for authenticated encryption with associated data is proposed. The proposed scheme, $\text{GCM-AES-VR}$, provides cloud data confidentiality as well as authentication. It is based on encrypt then MAC approach. It has achieved beyond birthday bound security. It provides security up to $\approx 2^{n/p}$ adversarial queries by considering that underlying block cipher is a secure PRP in the nonce respecting scenario. Where $n$ indicates block size and $P = n/m$; $m$ is a bit variance that is induced in data. In this scheme we have kept user free from entering three
different keys required for three different operations. The key derivation function is used to produce three keys for processing.

REFERENCES


Authors Profile

Vijay Ram Ghorpade is working as Principal at Bharati Vidyapeeth’s College of Engineering, Kolhapur, MH-India. His research areas are network, ad hoc networks and network security. He has published various research papers in national, international conferences, peer reviewed journals. He has completed his Ph. D. from SRTM University, Nanded in 2008.

Rajani Sangappa Sajjan has completed her post graduation from VTU Belgaon in the year 2008. She is now pursuing her Ph D from Shivaji University, Kolhapur. Her research domain is cloud data security, IoT, big data. She has published various research papers in conferences and journals.