IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

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

**Rajani S. Sajjan[1*], Vijay R. Ghorpade[2]**

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

[*]*Corresponding Author: rajanisajjan78@gmail.com, Tel.: 08308602157*

**Available online at: www.ijcseonline.org**

Accepted: 19/Dec/2018, Published: 31/Dec/2018

***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 $\mathbf{2^{n/2}}$ adversarial queries where $\boldsymbol{n}$ 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 $\mathbf{2^{n}\,/\,p}$ adversarial queries where, $\mathbf{p} = (\mathbf{n/m})$, where $\boldsymbol{n}$ is a block size and $\boldsymbol{m}$ 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 1: Different approaches to AEAD scheme

| AEAD Approach | Process 1 | Process 2 | Strongly unforgeable | Application |
|---|---|---|---|---|
| Encrypt then MAC | Plaintext is encrypted | MAC is generated based on ciphertext produced in process 1 | Yes | IPsec |
| Encrypt and MAC | Plaintext is encrypted | MAC is generated as per plaintext | No | SSH |
| MAC then encrypt | MAC is generated based on plaintext | Plaintext and MAC is encrypted | No | SSL/TLS |

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

Authenticated encryption schemes [6] [7] [8] [9] [10] provide birthday bound security.

***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 $2^{rn/(r+1)}$ 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 comapred 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 ε, empty string. For any string y, |y| indicates length of string y, such that y ϵ $\{0,1\}^*$. $|y|_n$ 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 || 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 $y \overset{@}{\leftarrow} Y$, 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 $\mathbf{K_e}$ is the key space, such that **Ke** ϵ $\{0,1\}^n$. Assume K is a key used in authenticated encryption process, where K ϵ $\mathbf{K_e}$. Plaintext M is given for processing where M ϵ $\{0,1\}^n$. The authenticated encryption operation can be written as E: **Ke** $X\,\{0,1\}^n \rightarrow \{0,1\}^n$. For any given K ϵ $\mathbf{K_e}$. , a fixed value, the encryption operation $E_K : \{0,1\}^n \rightarrow \{0,1\}^n$ is n-bit permutation function. Its inverse i.e. $D_K$ is written as $D_K = E_K^{-1}$. Assume that Permutation(n) is a set of all n-bit permutations. Adversary **A** has access to encryption oracle. Let K $\overset{@}{\leftarrow}$ $\mathbf{K_e}$. and ρ $\overset{@}{\leftarrow}$ Permutation(n) then pseudorandom permutation (PRP) advantage of adversary **A** against E is defined as

$$Adv_E^{prp}(\mathbf{A}\ ) = \left|\ \Pr[\mathbf{A}^{\ Ek\,(.)} = 1\ ] - \Pr\,[\mathbf{A}^{\ \rho(.)} = 1\ ]\ \right|$$

Here probabilities are taken randomly over K and ρ and adversary **A** . If $Adv_E^{prp}(\mathbf{A}\ )$ 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** $X\,\{0,1\}^m \rightarrow \{0,1\}^n$. Where it takes input as a key K and plaintext M and returns ciphertext C; K ϵ $\mathbf{K_{fu}}$ ,M ϵ $\{0,1\}^m$ and C ϵ $\{0,1\}^n$.

When any fixed value is taken as a K ϵ $\mathbf{K_{fu}}$ ; $F_{uK}$ is a function from $\{0,1\}^m$ to $\{0,1\}^n$ ; which can be represented as $F_{uK} : \{0,1\}^m \rightarrow \{0,1\}^n$. 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 $\overset{@}{\leftarrow}$ $\mathbf{K_f}$. and η $\overset{@}{\leftarrow}$ Fu(m,n) then pseudorandom function (PRF) advantage of adversary **A** against $F_u$ is defined as

$$Adv_{Fu}^{prf}(\mathbf{A}\ ) = \left|\ \Pr[\mathbf{A}^{\ Fuk\,(.)} = 1\ ] - \Pr\,[\mathbf{A}^{\ \eta\,(.)} = 1\ ]\ \right|$$

Here probabilities are taken randomly over K and η and adversary **A** . If $Adv_{Fu}^{prf}(\mathbf{A}\ )$ 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 $K_h$ ϵ $\mathbf{K_h}$ and input message y ϵ $\{0,1\}^m$ is given and output z ϵ $\{0,1\}^n$ is given by H. The H is said to be almost XOR universal hash function if for y ϵ $\{0,1\}^m$ and z ϵ $\{0,1\}^n$,

$$\Pr[K_h \overset{@}{\leftarrow} \mathbf{K_h}\ : H_{Kh}(y) = z\ ] \le \alpha$$

For y ϵ $\{0,1\}^m$ and y’ ϵ $\{0,1\}^m$ where y and y’ are two distinct values and z ϵ $\{0,1\}^n$

$$\Pr[K_h \overset{@}{\leftarrow} \mathbf{K_h}\ : H_{Kh}(y) \oplus H_{Kh}(y') = z\ ] \le \beta$$

Where H(α,β) 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 $b = b_{n-1}\ldots\ldots b_1 b_0$, string b can be defined with the help of polynomial b(x) ϵ 𝔑|x| by $b(x) = b_{n-1}x^{n-1} + \ldots.. + b_1x + b_0$ where $b_i$ ϵ {0,1} for any i ϵ [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 XOR. For example addition of p and q is denoted as $p \oplus q$ where p,q ϵ $GF(2^n)$.

In GF multiplication operation 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 ϵ $GF(2^n)$; 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.

$$\mathbf{E} : \mathbf{K}\ \ X\,\mathbf{N}\ \ X\,\mathbf{M}\ \ X\,\mathbf{H}\ \ \rightarrow \mathbf{C}\ X\,\mathbf{T}$$

$$\mathbf{D} : \mathbf{K}\ \ X\,\mathbf{N}\ \ X\,\mathbf{H}\ \ X\,\mathbf{C}\ \ X\,\mathbf{T}\ \ \rightarrow \mathbf{M}\ \ U\,\{\ □\}$$

i.e.

$$\mathbf{E}\,(K, N, A, M) = \mathbf{E}_k(N, A, M) \rightarrow (C,T)$$

$$\mathbf{D}(K, N, A, C,T) = \mathbf{D}_k(N, A, C,T) \rightarrow M\ /\ □$$

Where,
Key K ϵ **K**,
Nonce N ϵ **N**,
Associated data A ϵ **H**, where **H** ϵ {0,1}*
Plaintext M ϵ **M**, where **M** ϵ {0,1}*
Tag T ϵ **T,** where **T** ⊆ {0,1}*,
Error symbol □, indicates failure of decryption operation.

$$\mathbf{E}_k(N, A, M) \rightarrow (C,T) \text{ iff } \mathbf{D}_k(N, A, C,T) \rightarrow M.$$

When $\mathbf{D}_k$ returns □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. Psuedorandom 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 $F_{1\ \text{is}}$ defined as

$$F_1 : \{0,1\}^l \rightarrow \{0,1\}^{ns},\ F_1(x) = (y_1,y_2,\ldots.,y_s).\ y_i = Q_2\,(Q_1\,(\text{inc}^i\,(x \parallel [0]_{n-l}\,))\oplus \text{inc}^i\,(x)\,) \in \{0,1\}^n$$

$$\text{For } i \in [1, s\,] \text{ and } s \leq 2^{n-l} - 1,\ l \leq n \text{ and } x \in \{0,1\}^l$$

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 □ = qs blocks, then PRF-advantage of adversary, **A** , against function, $F_1$, is upper bounded by,*

$$Adv_{F_1}^{prf}(\mathbf{A}\ ) \leq \frac{□}{2^n}$$

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 ← fun ( l, ns ) and s ← fun ( l, n ) are the functions. The PRF advantage of adversary **A** against function $F_1$ is given as,

$$Adv_{F_1}^{prf}(\mathbf{A}\ ) = |\ \Pr[\mathbf{A}^{\ F1\,(.)} = 1\,] - \Pr\,[\mathbf{A}^{\ S\,(.)} = 1\,]\ |$$

Consider EDM construction in the reduced form with n - 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

$$Adv_{f}^{prf}(\mathbf{G}\ ) = \left|\ \Pr[\mathbf{G}^{\ f\,(.)} = 1\,] - \Pr\,[\mathbf{G}^{\ s\,(.)} = 1\,]\ \right|$$

$$\leq \frac{q_i}{2^n} + \frac{(\sigma^{\,q_i+1})}{2^{n\sigma}}$$

Then we construct a hybrid function $H_j^i$, where i indicates f functions and j indicates s.

$$H_j^i = (f, \ldots\ldots\ldots, f, s, \ldots\ldots\ldots, s)$$

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

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

If i = j then, $H_i^i = (f, \ldots\ldots\ldots, f)$ indicates $F_1$.

The PRF advantage of adversary **A** against function $F_1$ is upper bounded by,

$$\begin{aligned} Adv_{F_1}^{prf}(\mathbf{A}\ ) &= \left|\ \Pr[\mathbf{A}^{\ F1\,(.)} = 1\,] - \Pr\,[\mathbf{A}^{\ S\,(.)} = 1\,]\ \right| \\ &= \left|\ \Pr[\mathbf{A}^{\ H_i^i} = 1\,] - \Pr\,[\mathbf{A}^{\ H_j^0} = 1\,]\ \right| \\ &= \left|\ \textstyle\sum_i (\Pr[\mathbf{A}^{\ H_j^{i+1}} = 1\,] - \Pr\,[\mathbf{A}^{\ H_j^i} = 1\,]\ \right| \\ &\leq \textstyle\sum_i \left|\ \Pr[\mathbf{G}^{\ f\,(.)} = 1\,] - \Pr\,[\mathbf{G}^{\ s\,(.)} = 1\,]\ \right| \\ &\leq \textstyle\sum_i \frac{q_i}{2^n} + \frac{(\sigma^{\,q_i+1})}{2^{n\sigma}} \\ &\leq \frac{□}{2^n} + \frac{(\sigma^{\,□}+1)}{2^{n\sigma}} \end{aligned}$$

Where inequality is obtained by $\sum_i q_i$ = □. 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 $K_1$, $K_2$, $K_h$ 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 $K_1$ is used for first pass of encryption and key $K_2$ for second pass of encryption process. The third key $K_h$ is used by universal hash function to generate tag value. The underlying block cipher E : $\mathbf{K_e}$ X $\{0,1\}^n \rightarrow \{0,1\}^n$ 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, $K_m$ is a master key entered by user. This key is given to a key derivation funtion $f_i$ alongwith constant parameter $C_i$. $K_m \in \{0,1\}^n$ and n $\geq$ 1. $f_i$ is a set of functions used to generate $k_i$. The $k_i$ where i = 1 to m, m is chosen as per the function requirement, here it is taken as 3.

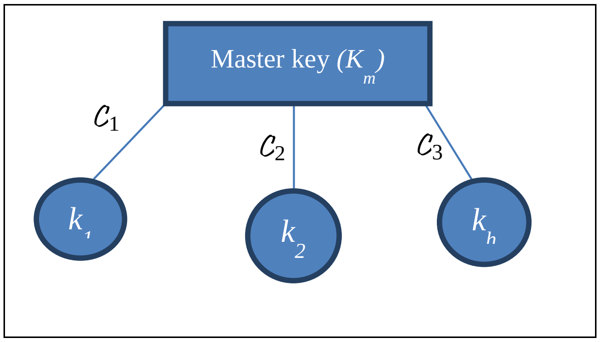

Figure 1 Key derivation process

These functions can be written as

$$k_1 = f_1(K_m, C_1, N, P, R, 128)$$
$$k_2 = f_2(K_m, C_2, N, P, R, 128),$$
$$k_h = f_3(K_m, C_3, N, P, R, 128).$$

The $k_i$ belongs to the key space $k_i \in \{0,1\}^l$ and $|\ k_i\ |$ = l where l = 128bits. The $C_i$ represents constant values those are already set by proposed model itself, but one can make them get generated dynamically. Here $C_i \in \{0,1\}^n$ and n $\geq$ 1. The key generation process is elaborated in the algorithm 1 given below.

*E. Encryption operation:*

The encryption operation is represented as,

$$\mathbf{E} : \mathbf{K} \ \mathrm{X}\, \mathbf{N} \ \mathrm{X}\, \mathbf{M} \ \mathrm{X}\, \mathbf{H} \ \rightarrow \mathbf{C} \ \mathrm{X}\, \mathbf{T}$$

where,

Key K ϵ **Ke** X **Kh** = $\{0,1\}^k$ 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 ϵ **M**, key K ϵ **K**, nonce N ϵ **N**, and associated data A ϵ **H** as inputs and returns C ϵ **C** as the ciphertext and tag T ϵ **T.**

*F. Decryption operation:*

This operation is inverse encryption operation. Decryption operation is represented as

$$\mathbf{D} : \mathbf{K} \ \mathrm{X}\, \mathbf{N} \ \mathrm{X}\, \mathbf{H} \ \mathrm{X}\, \mathbf{C} \ \mathrm{X}\, \mathbf{T} \ \rightarrow \mathbf{M} \ \ \mathrm{U}\, \{ □ \}$$

The decryption operation takes key K, nonce N, associated data A, ciphertext C, and tag T as inputs and returns M plaintext or □error as output. Error symbol □, 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 DesiredKeyLen i.e. 16 bytes
Process:
Get master key i.e. user entered value, $K_m$
        $k_1 = f_1(K_m, C_1, N, P, R, 128)$;
        $k_2 = f_2(K_m, C_2, N, P, R, 128)$;
        $k_h = f_3(K_m, C_3, N, P, R, 128)$;
$k1$ is used for first encryption process
$k2$ is used for second encryption process
$kh$ is used in tag value calculation
```

## V. GCM-AES-VR: AEAD SCHEME

The proposed AEAD scheme, GCM-AES-VR, is defined as,

Ɯ = ($\mathbf{E}_K$(..,..,..,..),$\mathbf{D}_K$(..,..,..,..)),

is a nonce based AEAD scheme. Where K is a key used in encryption operation and K $\overset{\Phi}{\leftarrow}$ $\{0,1\}^k$ . The general construction of GCM-AES-VR AEAD scheme is represented in figure 2 as shown. Here two encryption operations are performed, $E_{k1}$ and $E_{k2}$ , by getting secret keys $k_1$ and $k_2$ generated in key derivation process.

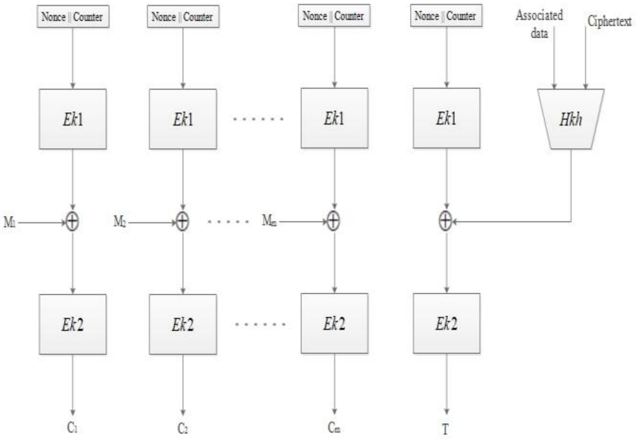

Figure 2 Construction of proposed GCM-AES-VR AEAD scheme

**Algorithm 2: Encrypt_GCM_AES_VR( )**

*Input: $K_1$, $K_2$, $K_h$ three keys derived from master key K, associated data $A_d$, nonce value N, plaintext M*
*Output: ciphertext C and tag value T*

```
Divide plaintext M into equal blocks i.e. $M_1 || M_2 || \dots\dots\dots || M_m$ such that $|M_j| = n$ for $1 \le j \le m-1$;
$P_0 = N || 0^{n-l-1}1$;
$Q_0 = E_{K2}( M_i \oplus E_{K1}(P_0) )$;
for j = 1 to m do
        $P_j$ = increment($P_j$-1);
        $Q_j = E_{K2}( M_i \oplus E_{K1}(P_j) )$;
for j = 1 to m-1 do
        $C_j = Q_j$;
$C = C_1 || C_2 || \dots\dots || C_m$;
$Q = Q_0 \oplus H_{kh}(A_d, C)$;
$T = Q_0$;
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: $K_h$ , key derived from master key $K_m$, associated data $A_d$, ciphertext C*
*Output: Hash value H*

```
$Y = A_d || 0^{|Ad| n.n-|Ad|} || C || 0^{|C| n.n-|C|} || \lceil A_d \rceil_{n/2} || \lceil C \rceil_{n/2}$;
Divide Y into blocks i.e. $Y_1 || Y_2 || \dots\dots\dots || Y_m$ such that $|Y_j| = n$ for $1 \le j \le m$;
$H = 0$;
for j = 1 to m
        $H = ( H \oplus Y_j ) . K_h$
return $H$
```

Algorithm 3: Hash algorithm

**Algorithm 4:Decrypt_GCM_AES_VR( )**

*Input: $K_1$, $K_2$, $K_h$ three keys derived from master key K, associated data $A_d$, nonce value N, ciphertext C and tag value T*
*Output: Plaintext M or ⍰ error symbol*

```
$P_0 = N || 0^{n-l-1}1$;
$Q_0 = E_{K2}( M_i \oplus E_{K1}(P_0) )$;
$Q = Q_0 \oplus H_{kh}(A_d, C)$;
$T' = Q_0$;
if ($T' == T$) then
{
        divide ciphertext C into equal blocks i.e. $C_1 || C_2 || \dots\dots\dots || C_m$ such that
        $|C_j| = n$ for $1 \le j \le m-1$;
        for j = 1 to m do
        {
                $P_j$ = increment($P_j$-1);
                $Q_j = E_{K2}( M_i \oplus E_{K1}(P_j) )$;
        }
        for j = 1 to m-1 do
        {
                $M_j = Q_j \oplus C_i$;
        }
        $M_m = Q_m \oplus C_m$;
        $M = M_1 || M_2 || \dots\dots || M_m$;
        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. The 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, $\mathbf{E}_K$ (..,..,..,..) or random oracle Φ(..,..,..,..). If all nonces $N^1, N^2, \dots\dots, N^q$ are always different for all encryption queries i.e. $(N^1, A^1, M^1) \dots (N^q, A^q, M^q)$ 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 ɰ = ($\mathbf{E}_K$(..,..,..,..),$\mathbf{D}_K$(..,..,..,..)) is given as

$$Adv_{\Phi}^{privacy}(\mathbf{A}) = \left| \Pr[\mathbf{A}^{\mathbf{E}K(..,..,..,..)} = 1] - \Pr[\mathbf{A}^{\Phi(..,..,..,..)} = 1] \right|$$

*Authenticity:*

Assume that there is an adversary **A** having access to two oracles, $\mathbf{E}_K$ (..,..,..,..) and $\mathbf{D}_K$(..,..,..,..).

*Step 1:*
The adversary **A** puts queries ($N^w$, $A^w$, $M^w$) where w ∈ [1,q]. Then **A** returns ($C^w$, $T^w$) = $\mathbf{E}_K$ ($N^w$, $A^w$, $M^w$).

*Step 2:*
The adversary **A** puts a challenge query (N, A, C, T) ∉ {( $N^w$, $A^w$, $C^w$ ,$T^w$)$\}_{w=1}^{q}$ to $\mathbf{D}_K$(..,..,..,..).
If $\mathbf{D}_K$ (N, A, C, T) ≠ □ 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 Ɯ = ($\mathbf{E}_K$(..,..,..,..),$\mathbf{D}_K$(..,..,..,..)) is given as

$$Adv_{\Phi}^{auth}(\mathbf{A}\ ) = \Pr[\mathbf{A}^{\mathbf{E}K\,(..,..,..,..)\mathbf{D}K\,(..,..,..,..)} \text{ forges }].$$

*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 : **Ke** X $\{0,1\}^n \rightarrow \{0,1\}^n$ be a block cipher and H : **Kh** X $\{0,1\}^*$ X $\{0,1\}^* \rightarrow \{0,1\}^n$, and **Ke** and **Kh** are two nonempty sets such that **Ke , Kh** ϵ $\{0,1\}^n$. Let there is a nonce respecting adversary **A** which generates at most $q \le 2^{n/p}$ 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 $\partial = q(m+1)$ oracle queries such that for any adversary **A** ,

$$Adv_{GCM-AES-VR(E)}^{privacy}(\mathbf{A}\ ) \le 2\,Adv_{E}^{prp}(\mathbf{B}\ ) + \frac{\partial}{2^n}$$

The proof takes two steps, as discussed 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 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 ∂ = q(m+1) oracle queries such that

$$Adv_{GCM-AES-VR(E)}^{privacy}(\mathbf{A}\ ) \le 2\,Adv_{E}^{prp}(\mathbf{B}\ ) + Adv_{GCM-AES-VR(Q)}^{privacy}(\mathbf{A}\ )$$

Now we will upper bound $Adv_{GCM-AES-VR(Q)}^{privacy}(\mathbf{A}\ )$. 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 $q \le 2^{n/p}$ adversarial queries to GCM-AES-VR[Q] generating at most ∂ blocks. Then for any adversary* ***A*** *,*

$$Adv_{GCM-AES-VR(Q)}^{privacy}(\mathbf{A}\ ) \le \frac{\partial}{2^n}$$

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 ,ℑ , 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,

$$Adv_{\text{GCM−AES−VR[Q]}}^{privacy}(\mathbf{A}\ ) = \left|\ \Pr[\mathbf{A}^{\mathbf{E[Q]}\,(..,..,..,..)} = 1\,] - \Pr[\mathbf{A}^{\mathfrak{I}\,(..,..,..,..)} = 1\,]\ \right| \ \ge \frac{\partial}{2^n}$$

where adversary **A** makes q queries for m block length to GCM-AES-VR[Q]. This generates $\partial = q(m+1)$ blocks.
Let $\mathfrak{h}$ is a random function such that, $\mathfrak{h} \in Function(l, ns)$, where $s = m + 1$. Now let consider that there exists another adversary **B** which generates q queries to an oracle **OR ,** either $F_1$ or $\mathfrak{h}$ , generating $\partial = qs$.
If **OR** is a $F_1$ then encryption operation at **B** is same as encryption operation at **A** . This means that

$$\Pr[\mathbf{B}^{\mathbf{F1(.)}} = 1\,] = \Pr[\mathbf{A}^{\text{E[Q]}\,(..,..,..,..)} = 1\,]$$

In the same way if **OR** is $\mathfrak{h}$, then **B** gives same functioning for the random function ℑ for **A .** Hence

$$\Pr[\mathbf{B}^{\mathfrak{h}(.)} = 1\,] = \Pr[\mathbf{A}^{\mathfrak{I}\,(..,..,..,..)} = 1\,]$$

which says that

$$\begin{aligned} Adv_{F_1}^{prf}(\mathbf{B}\ ) &= \left|\ \Pr[\mathbf{B}^{\mathbf{F1(.)}} = 1\,] - \Pr[\mathbf{B}^{\mathfrak{h}(.)} = 1\,]\ \right| \\ &= \left|\ \Pr[\mathbf{A}^{\text{E[Q]}\,(..,..,..,..)} = 1\,] - \Pr[\mathbf{A}^{\mathfrak{I}\,(..,..,..,..)} = 1\,]\ \right| \\ &= Adv_{\text{GCM−AES−VR[Q]}}^{privacy}(\mathbf{A}\ ) > \frac{\partial}{2^n} \end{aligned}$$

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

$$Adv_{\text{GCM−AES−VR[Q]}}^{privacy}(\mathbf{A}\ ) > \frac{\partial}{2^n}$$

At this point Lemma 1 is over. So we conclude that privacy of GCM-AES-VR is secure up to $q \approx 2^{n/p}$ 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 : **Ke** X $\{0,1\}^n \rightarrow \{0,1\}^n$ be a block cipher and H : **Kh** X $\{0,1\}^*$ X $\{0,1\}^* \rightarrow \{0,1\}^n$ is an universal hash function, and **Ke** and **Kh** are two nonempty sets such that **Ke , Kh** ϵ $\{0,1\}^n$. 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 $\partial = (q + 1)\,(m + 1)$ oracle queries such that for any adversary **A** ,

$$Adv^{auth}_{GCM-AES-VR(E)}(\mathbf{A}) \leq 2\,Adv^{prp}_{E}(\mathbf{B}) + \frac{\partial}{2^n} + \mathfrak{G} + \frac{n}{2^n} + \frac{1}{2^{\rho}-q}$$

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 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 $\partial = (q + 1)\,(m + 1)$ oracle queries such that

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

Now we will upper bound $Adv^{auth}_{GCM-AES-VR(Q)}(\mathbf{A})$. 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 Permutation(n). Additionally there is H an universal hash function,* $\mathfrak{G}$*. There is a nonce respecting adversary* **A** *which makes* $q \leq 2^{n/p}$ *adversarial queries and one forgery attempt to GCM-AES-VR(Q) generating at most* $\partial$ *blocks. Then for any adversary* **A,**

$$Adv^{auth}_{GCM-AES-VR(Q)}(\mathbf{A}) \leq \frac{\partial}{2^n} + \mathfrak{G} + \frac{n}{2^n} + \frac{1}{2^{\rho}-q}$$

Proof:

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

$$Adv^{auth}_{GCM-AES-VR(Q)}(\mathbf{A}) = \Pr[\mathbf{A}^{\mathbf{E}K(Q)\mathbf{D}K(Q)} \text{ forges}]$$
$$\leq \left| \Pr[\mathbf{A}^{\mathbf{E}K(Q)\mathbf{D}K(Q)} \text{ forges}] - \Pr[\mathbf{A}^{@1,@2} \text{ forges}] \right| + \Pr[\mathbf{A}^{@1,@2} \text{ forges}]$$

$\mathbf{E}_K$ (Q) and $\mathbf{D}_K$ (Q) are encryption and decryption operations of GCM-AES-VR(Q). $@_1$ and $@_2$ are the random oracles where $@_1$ is a random oracle which always returns a random string (C,T) $\overset{@}{\leftarrow}$ **C X T** and $@_2$ is a random oracle which returns a random string or a reject symbol, i.e., M / □ $\overset{@}{\leftarrow}$ **M** □ { □ }.

$$\left| \Pr[\mathbf{A}^{\mathbf{E}K(Q)\mathbf{D}K(Q)} \text{ forges}] - \Pr[\mathbf{A}^{@1,@2} \text{ forges}] \right| = Adv^{privacy}_{GCM-AES-VR(Q)}(q + 1, \partial) \leq \frac{\partial}{2^n}$$

This is shown in privacy proof and $\partial = (q + 1)\,(m + 1)$. So we conclude that authenticity of 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, GCM-AES-VR, which has achieved optimal security, $2^n$, where n is length of block. The proposed scheme provides $2^{n/p}$ ; $p = (\frac{n}{m})$ where m indicates bit variance. Table 2 shows the comparative analysis of 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 2: Comparative analysis

| **Parameter** | **GCM-AES** | [13] **OGCM-1** | [13] **OGCM-2** | **Proposed algorithm (GCM-AES-VR)** |
|---|---|---|---|---|
| Key size | 128 | 128 | 128 | 128 |
| Block size *n* | 128 | 128 | 128 | 128 |
| Keys used | 1 | 3 | 3 | 1 |
| Key management needed | Yes | Yes | Yes | No |
| Block cipher calls | 1 | 2a + 2 | 2a + 2 | 2a |
| Security bits | 64 | ≅107 | ≅121 | ≅124 |

*a indicates plaintext block length*

## VIII. CONCLUSION

In this paper a novel approach for authenticated encryption with associated data is proposed. The proposed scheme, 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

[1] Chanathip Namprempre Mihir Bellare, "Authenticated Encryption: Relations among notions and analysis of the generic composition paradigm," *Lecture Notes in Computer Science,Springer-Verlag*, vol. 1976, pp. 531–545, July 2007.

[2] C. Jutla, "Encryption modes with almost free message integrity".

[3] V. Gligor and P. Donescu, "Fast encryption and authentication: XCBC encryption and XECB authentication modes.".

[4] M. Bellare, J. Black, andT. Krovetz P. Rogaway, "OCB: A block-cipher mode of operation for efficient authenticated encryption," 2001.

[5] Rogaway P., "Authenticated-Encryption with Associated-Data," in *9th ACM Conference on Computer and Communications Security*, Washington,USA, 2002, pp. 98-107.

[6] A. Bogdanov, A. Luykx, B. Mennink, E. Tischhauser, and K. Yasuda E. Andreeva, "Parallelizable and authenticated online ciphers".

[7] S. Fluhrer, C. Forler F. Abed, "Pipelineable on-line encryption".

[8] C. Forler, and S. Lucks E. Fleischmann, "McOE: a family of almost foolproof on-line authenticated encryption schemes".

[9] M. Bellare and C. Namprempre, "Authenticated encryption: relations among notions and analysis of the generic composition paradigm".

[10] P.Jovanovic,B.Mennink,and S.Neves R.Granger, "Improved masking for tweakable blockciphers with applications to authenticated encryption".

[11] J.Viega D.A.McGrewand, "The security and performance of the Galois/counter mode (GCM) of operation".

[12] Dr. Vijay R. Ghorpade Rajani S. Sajjan, "AES-VR:A New Approach for Cloud Data Confidentiality," *International Journal of Computer Technology and Applications*, Accepted 2018.

[13] Hong-Gang Hu, Qian Yuan Ping Zhang, "Close to optimally secure variants of GCM," *Hindawi*, vol. 2018, March 2018.

## 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.