Volume 3, Issue 6, June 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Privacy Preservation Against Global Eavesdropper in Sensor Networks Using Network Coding

Pavan Jahagirdar
M.Tech 4th Sem CNE
Dept. Of CSE, CMRIT
Bangalore, India

Mr. Manoj Challa
Associate Professor
Dept. Of CSE, CMRIT
Bangalore, India

Mr. K Sundeep Kumar
Associate Professor & HOD
Dept. Of CSE, CMRIT
Bangalore, India

Abstract— Many security protocols have been developed to provide privacy in sensor networks. Security risk in sensor networks is one of the significant issues to be dealt with. The existing techniques guard the data only against the local eavesdropper who is having limited knowledge of the network topology. A stronger adversary such as global eavesdropper can still analyze the pattern of traffic and launch advanced attacks such as flow tracing and traffic analysis. Due to these advanced attacks the privacy of the users is compromised. This paper proposes a new network coding mechanism to protect the privacy in sensor network against the global eavesdropper. We use the source imitation approach to calculate the candidate traces which can transmit data at the same time and same rate. The proposed scheme also uses the optimal path among the candidate traces for the faster transfer of data. Through the simulation and analysis, we exhibit that the proposed scheme is both energy efficient and successful in providing privacy in sensor networks.

Keywords—Sensor network, network coding, source imitation, eavesdropper, Homomorphic Encryption.

I. Introduction

Wireless sensor networks (WSNs) have been widely employed in many diverse applications because of their ease of installation, cost efficient and portability. A wireless sensor network mainly consists of one or more sensing devices such as acoustic microphones video or still cameras, seismic or magnetic sensors. Each sensor node communicates wirelessly with a few other local nodes within its radio communication range. Even though wireless sensor networks are having these advantages, there are many drawbacks associated with them such as, weak system reliability, shorter coverage range and the most critical drawback is its privacy and security issues. One of the ways to increase the reliability and range of the WSNs is to employ multi-hop routing. The concept of multi-hop routing is to forward a packet to the destination using different path in case of the node failure. But, the critical issue still remains of providing security and privacy in WSNs. Therefore, preserving the privacy of the location of the source node remains critical. Wireless sensor networks are used in many areas such as military supervision where possibility of the eavesdropping the traffic is high to get hold of sensitive information. Exploitation of such information can cause economic losses or cause danger to human lives. To protect such information, researchers are finding out new ways to provide standard security services such as, availability, integrity, confidentiality and authentication. The exchange of information between sensors can disclose sensitive information which can reveal the location information of the critical modules present in the network.

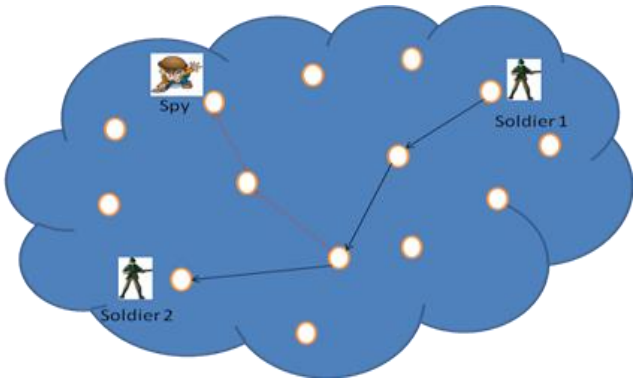

Figure 1(A): Threats in Wireless Sensor Networks

In figure 1(A) shows the WSNs deployed in the military surveillance area, where a soldier 1 is sending confidential information to the soldier 2 i.e. sink via many intermediate nodes. A spy who is present on the same network tries to intercept the information by compromising one of the intermediate nodes. The nodes may reveal sensitive information to the adversary such location of the source or positions of the armed forces in the vicinity.

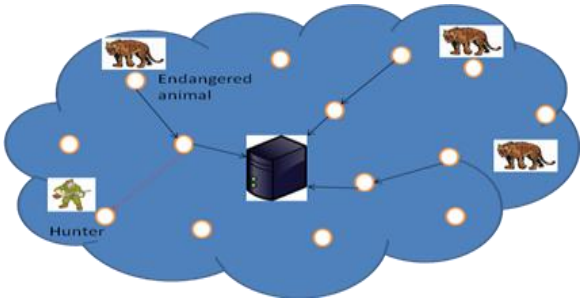

Figure 1(B): Threats in monitoring endangered animals

In figure 1(B) the sensor networks are deployed to monitor the endangered animals in a forest. An event is triggered whenever an animal is spotted in the monitored area. The hunter tries to gather this information and may capture or kill the endangered animal [1]. The above scenario depicts the vulnerability of WSNs is more because of its open wireless medium to transmit the information from source to destination.

It is very challenging to effectively put a stop to these kinds of traffic analysis attacks and provide location privacy in WSNs. Existing schemes such as proxy-based schemes [2], onion routing schemes [3] and Chaum's mix-based schemes [4] [5], may moreover require a series of trusted proxies which can forward the data resulting in the degradation of the performance. These schemes are only capable of dealing with only local eavesdropper and to the limited section of the network. The deployment of network coding in wireless sensor networks can result in the high performance gain and also offer sufficient way to stop traffic analysis and flow tracing attacks. Similar to Chaum's mix-based schemes [4],[5], network coding provides built-in coding/mixing mechanism, which implies that privacy preservation is achieved in distributed manner. There is unlinkability between the both incoming and outgoing packets, which is a very important property for averting the traffic analysis attacks.

In this paper which is based on network coding, we use Homomorphic Encryption (HE) on the Encoding Vector to achieve the efficient privacy in WSNs. With the use of HEs the confidentiality of the Encoding Vector is effectively guaranteed making it almost complicated for an adversary to obtain the plaintext. Instead of Link-to-Link encryption, End-to-End encryption on Encoding Vector is employed to achieve even energy efficiency and avoiding intermediate coding/mixing operations. In this paper we focus on the secure communication methods that preserve the privacy against both global and local eavesdropper. The contributions in this paper are threefold.

- We point out that the assumption of a global eavesdropper who can visualize the entire network and can monitor the traffic traversing the network. We apply the network coding mechanism by tagging the packets with the Encoding Vector.
- We apply the source imitation approach to find out the number of candidate traces present in the network and evaluate the proposed techniques for location privacy in sensor networks.
- We compute the optimal path between the source and destination to maintain the energy reserve in the sensor networks.

II. Existing Approaches

Providing source privacy is one of the most difficult tasks in WSNs. To begin with, an attacker is present in the same wireless medium as that of the defender and with no trouble the attacker can analyze the traffic pattern and intercept the network traffic. He can use the analyzed information to launch active attacks such as node compromising, data modification or data injection. Subsequently, sensors that are deployed are having limited processing speed and energy supplies. The existing system uses the approach of network coding which is carried out in three steps i.e. source encoding, intermediate recoding and sink decoding[6]. In source encoding step the messages belonging to same generation are encoded with the tag and are transmitted to the next hop. Every intermediate sensor nodes that is present between the source and destination perform the process of intermediate recoding. After obtaining the required number of messages from the previous node, random combination of the messages is formed and then forwarded to the next node. This process is repeated until the messages reach the destination. After the messages arrive at the destination, the sink decoding mechanism inverses the received message and obtains the original message.

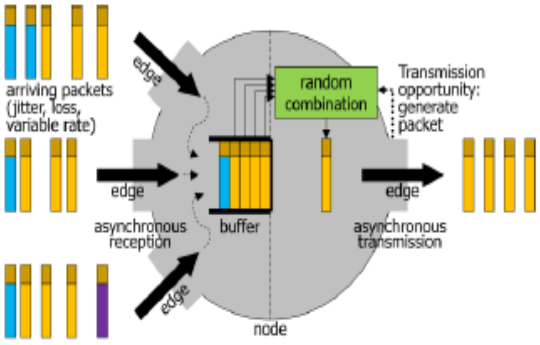

Figure 2: Steps involved in network coding

Figure 2 depicts the scenario of network coding; the messages or packets will be arriving from various sensor nodes asynchronously. All these packets belonging to different generations are stored in the buffer. After the required threshold of the buffer is reached, random mixing or combination is carried out and are passed onto the next nodes. The method of intermediate recoding at the nodes consumes most of the battery power of the sensor nodes and this process is carried out at each node between source to destination. The existing approach only works for local eavesdropper; however a global eavesdropper can still deduce the location of packet. The system works better only when multiple sensors generate data at the same time. If only one sensor node is generating the data then the buffer's maximum threshold condition cannot be met and dummy messages have to be included in this case. This may lead to serious performance degradation of the system as the energy reserve of sensor nodes is compromise in order to provide the security. Delay in packet transmission is observed when the dummy messages are included to meet the maximum threshold of the buffer. To overcome these weaknesses we avoid the process of intermediate recoding in the sensor nodes and focus on strong end-to-end encryption using network coding. To maintain the energy reserve of the sensor nodes we select random number of nodes to act as source nodes using the approach of source imitation and also calculate the optimal path for transmission of packets.

III. Proposed System

We propose three algorithms to increase the security and reduce the energy consumption of the sensor nodes' battery, (1) Finding the coverage set of sensors that can transmit at same time using Source Simulation method; (2) We propose an algorithm to compute the shortest path between source and destination and (3) We follow the principle of network coding on the messages that will traverse the network. The architectural design for the proposed system is shown in the figure 3 below:-

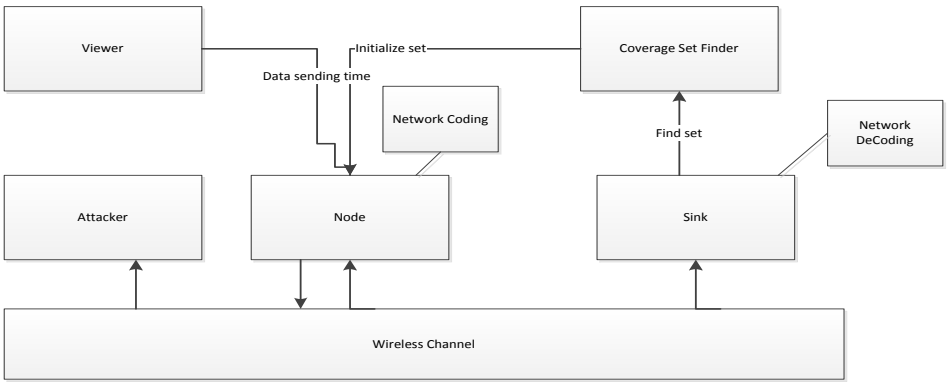

Fig 3: - System architecture of the Privacy Protection against Global Eavesdropper using Network Coding

Viewer is the front end of the system. User configures the system as well as the view to observe the output results using the viewer. User will configure the number of nodes to be present in simulation and the data sending time for the nodes. Using the wireless channel medium nodes communicate with each other and as well as with the sinks. Attacker will listen for all messages on wireless channel and try to identify the source location of packet. Sink use the coverage set finder to identify the set of nodes belonging in one set and using algorithm we propose the optimal path from node to the sink is calculated. Network coding mechanism is implemented in all the nodes present in the network so that privacy of each packet traversing the network is achieved. After sink receives the packet from the nodes, it uses network decoding/source decoding method to obtain the real message.

IV. Implementation

This part mainly concentrates on overcoming the shortcomings in the existing system. In order to overcome those constraints we implement three important algorithms to achieve privacy and energy efficiency and those are i) Coverage set finder algorithm, ii)Shortest path computation and iii) Network coding algorithm

i) Coverage set finder algorithm

The coverage set finder algorithm uses the method of source imitation to reduce the energy consumption in the wireless sensor networks. Figure 4 depicts the source imitation approach i.e. selection of the potential sources in the network.

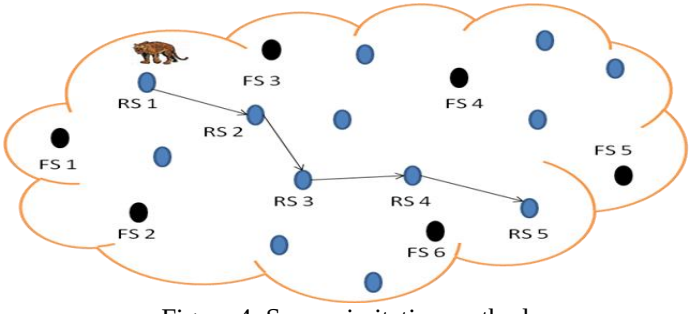

Figure 4: Source imitation method

To reduce the energy consumption in WSNs we prefer to reduce the total number of potential sources in the network [7]. We create multiple candidate traces in the WSN in order to conceal the traffic generated by the real sources. The application in the system determines the number of candidate traces to be generated. In source imitation method, set of fake sources are created in the network field. As shown in figure 3, many sensor nodes are deployed in the field where RS is Real Source and FS is Fake Source. When an event is reported at RS1, the fake sources present in the network generates traffic pattern similar to that of the real source to confuse the attacker. Source imitation works as follows: Before deploying the sensor, we arbitrarily select a set of N sensor nodes and preload each of the randomly selected nodes with a token, each having unique ID. To analyze the behavior of the real source, these tokens are circulated among different sensor nodes. For simplicity, we describe the node having the token as token node. Initially we summarize the behavior of real objects so that we can create candidate traces. After the deployment, every node which is having token will release a signal imitating the signal used by real objects for event detection. This will activate event detection in its local area and generate traffic as if a real event was detected. The circulation of token between different nodes is protected by the secret key exchanged between them.

Confidentiality

The person who is defending the network the WSN has better knowledge about the topology than the attacker. Therefore, the defender builds a model or profile for the behavior of real objects that are used to create candidate traces in network with probability P. the candidate traces created by the defender will always be considered as valid candidate trace. Let |Ro| be the number of real objects in the network. R_T determines the set of candidate locations which includes an average of |Ro| + |N|*P node IDs. Therefore, the privacy provided by the source imitation approach can be estimated by

$$b = \frac{\log_2 |Ro| + |N|*P}{|Ro|} \quad (1)$$

$$b = \log 2 \quad (1+ |L| * P \ / |Ro|) \quad (2)$$

where b is level of privacy in terms of bits, L is the number of fake sinks that are created.

Total Energy Consumption

We assume that a sensor node present will forward the packet or data whenever the channel is free. Hence Δ is negligible and real packets need not wait in the buffer and there is no need to add multiple fake packets to meet the buffer threshold. Considering Ae be the average number of packets to send an event from source to destination node and we get the formula $W_T = Ae * (|N| + |Ro|) * T / \alpha * \Delta$. This shows that the average communication overhead is increased by a factor of |N| + |Ro| ÷ |Ro| to protect location privacy.

ii) Algorithm to calculate Shortest Path

Input: Source node

Output: Shortest path from source node to base station

```
ConstructPathToSink()
{
        Stack st = new Stack()
        String nodeid = "2"
        push(nodeid);
        pathtosink[2] = "2"
        while (st.empty() == false)
        {
                String cur = (String) st.pop()
                int n = Integer.parseInt(cur);
                for (int i = 0; i < noofNodes; i++)
                {
                if (neighbourmat[i][n] != inf)
                 {
                        if (nodes[i].nextnodetosink == -1 && pathtosink[i] == null)
                        {
                        nodes[i].nextnodetosink = n
                        nodes[i].counttosink = nodes[n].counttosink + 1
                        pathtosink[i] = i + "," + pathtosink[n]
                        push(Integer.toString(i))
                        }
                        else
                        {
                        int hop = nodes[n].counttosink + 1
                        if (nodes[i].counttosink > hop)
                        {
                        nodes[i].counttosink = hop;
                        nodes[i].nextnodetosink = n;
                        }
                        }
```


iii) Network coding method

The network coding approach on packets mainly consists of three phases: source coding, intermediate recoding and sink decoding. In intermediate recoding stage, every sensor node present between source and destination performs transformation of each packet that is received and the node will not forward the packet until the buffer is full. Hence we skip the stage of intermediate recoding to maintain the energy reserve of the sensor nodes. Each sink is assumed to be having two keys, the encryption key *ekey* and decryption key from an trusted authority and encryption key *ekey* is public key and is distribute among other nodes. We mainly concentrate on two stages in network coding and those are source encoding and sink decoding.

Source encoding

Considering that a source has *m* messages say $x_1 \ldots x_m$ to be forwarded. Initially the source prefixes *m* unit vectors to *m* messages respectively as shown in the figure 5.

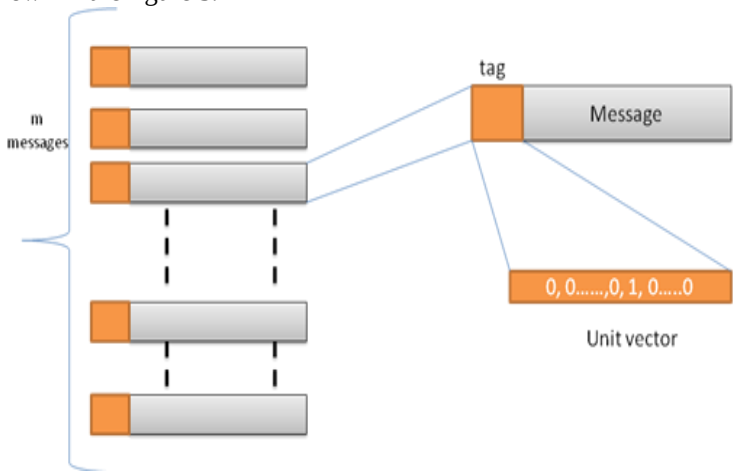

Figure 5: Source encoding operation

After prefixing the tag, the source performs linear encoding on these messages and they are network coded using encoding vector. To increase the confidentiality for the messages the Homomorphic Encryption is applied as follows:

$$C_i(e) = Eekey\ (g_i(e)),\ (1=i=h) \qquad (3)$$

$$C(e) = [C_1(e), C_2(e) \ldots .C_m(e)] \qquad (4)$$

Where ekey is the encryption key.

Sink Decoding

When sink receives the packet, it first decrypts the packet tag using the subsequent decryption key *dkey*.

$$G_i(e) = D_{dkey}(C_1(e),\ (1=i=h) \qquad (5)$$

$$G(e) = [G_1(e), G_2(e) \ldots .G_h(e)] \qquad (6)$$

Once sufficient numbers of packets are received, a sink decodes the packets to obtain the original messages. Then, the sink obtains the decoding vector, which is the inverse of the Encoding Matrix as shown in the following equations.

$$G^{-1} * G = U \qquad (7)$$

$$G=[G(e_1),G(e_2) \ldots .G(e_h)]^T \qquad (8)$$

To end with, the sink can use the inverse to recover the original data, shown as follows

$$\begin{pmatrix} x_1 \\ \ldots \\ \ldots \\ x_m \end{pmatrix} = G^{-1} \begin{pmatrix} y(e_1) \\ \ldots . \\ \ldots . \\ y(e_h) \end{pmatrix} \qquad (9)$$

V. Simulation Results

The simulation of the proposed scheme is carried out in the JProwler simulator. The proposed method can provide privacy preservation by restricting traffic analysis such as size correspondence, time correspondence and message content correspondence. Size correspondence can be avoided by trimming the messages to be of same length. Time correspondence attacks cannot be carried in the proposed scheme because of the absence of buffer in the network. Message content correspondence is avoided with the help of Homomorphic Encryption on Encoding Vectors. The first

step towards implementation is to create the required number of sensor nodes in the network area. Then we simulate the behaviour of sensor nodes with respect to source imitation and network coding.

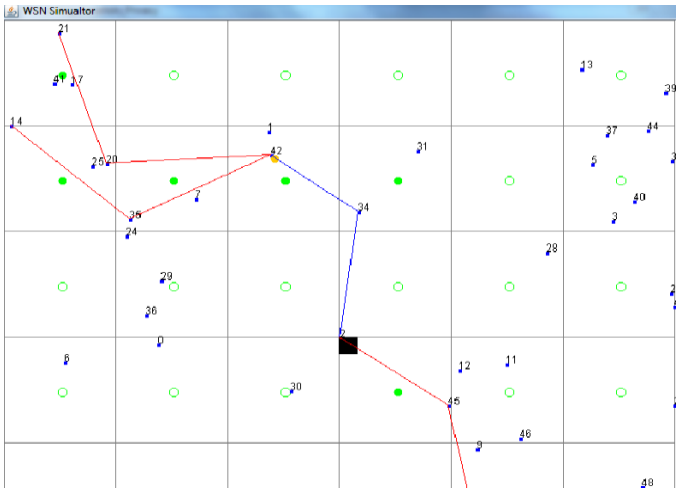

Figure 6: Nodes in the network area with base station as 2 and 42 as source node

The source node is chosen as 42 and base station is fixed at node 2 as shown in figure 6. The source simulation approach is used in this case and original packets is sent from node 42 to base station 2, whereas the other nodes present in the candidate trace will generate fake packets to confuse the adversary. The packets traversing between the nodes are network coded to preserve the confidentiality of the messages. The communication overhead is calculated for source imitation scheme is decided by the number of fake sources created in the network. Thus, we focus of our simulation is on how much cost is paid in order to achieve location privacy. During the simulation process, we assume that only one event is generated in the network. Multiple fake events are created and simulated in the network. The position of the event in the network is randomly selected. We assume that sensor network is handling real-time applications and nodes will forward packets as soon as they receive them. Thus we set the time interval for source imitation as $\Delta = T/10$ and for periodic collection [7] is $\Delta = T$. Thus source imitation approach will forward packets ten times faster than the periodic collection method. Figure 7 shows the communication overhead involved in our source imitation method to achieve the location privacy. We can observe that communication cost increases as the requirement for location privacy increases.

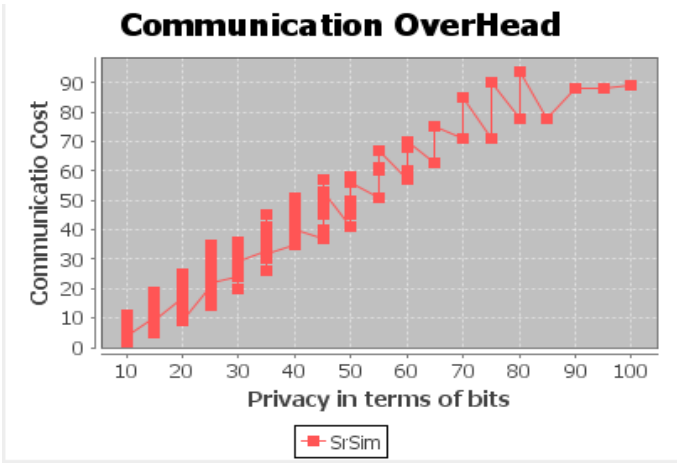

Figure 7: Graph to depict Communication Overhead

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed as efficient source imitation approach with the combination of network coding for preserving the location privacy in sensor networks. With the use of Homomorphic Encryption on Encoding Vectors, the proposed idea offers protection against traffic analysis attacks and also preserves the confidentiality of the messages. Because of the shortest path calculation, the data travels faster between sensor nodes and no computation is carried out in the intermediate nodes maintaining the energy reserve of the sensor nodes. The simulation evaluation demonstrates that the communication cost is increased with requirement of location privacy and becomes stable after reaching certain number of bits. In our future work we can further increase the location privacy by sink imitation approach to protect the location of destination node.

REFERENCES

[1] M.Shao, Y. Yang, S. Zhu, and G. Cao, "Towards statistically strong source anonymity for sensor networks," in

Proc. IEEE INFOCOM'08, pp. 51-55, 2008.
[2] Wensheng Zhang · Guohong Cao · Tom La Porta." Dynamic proxy tree-based data dissemination schemes for wireless sensor networks" May 2006
[3] http://en.wikipedia.org/wiki/Onion_routing.
[4] M. Rennhard and B. Plattner, "Introducing MorphMix: peer-to-peer based anonymous Internet usage with collusion detection," in *Proc. ACM Workshop on Privacy in the Electronic Society*, pp. 91-102, 2002.
[5] Sumit Jaiswal , Jaydeep Howlader, Prasenjit Choudhury" A Review Of Anonymous Communications– Mix.
[6] Yanfei Fan, Yixin Jiang, Haojin Zhu, Jiming Chen, Xuemin Shen, *"Network Coding Based Privacy Preservation against Traffic Analysis in Multi-Hop Wireless Networks,"* IEEE Transaction on Wireless Communication, Vol. 10, no. 3, 2011
[7] Kiran Mehta, Donggang Liu, Matthew Wright, *"Protecting Location Privacy in Sensor Networks against a Global Eavesdropper"* IEEE Transactions on Mobile Computing, Vol. 11, No.2, 2012

Authors

Mr. Pavan Jahagirdar is presently doing Master of Technology in Computer Networks and Engineering at CMR Institute of Technology, Bangalore, Karnataka. He obtained is Bachelor of Engineering degree in Information Science and Engineering from SKSVMA College of Engineering and Technology, Laxmeshwar, Karnataka in the year 2011.

Mr. Manoj Challa is pursuing Ph.D(CSE) in S.V.University, Tirupati, India. He completed his M.E(CSE) from Hindustan College of Engineering, Tamil Nadu in 2003. He is presently working as Associate Professor, CMR Institute of Technology, Bangalore. He presented nearly 18 papers in national and international conferences. His research areas include Artificial intelligence and computer networks.

Mr. K Sundeep Kumar received the M.Tech (IT) from Punjabi University in 2003, ME (CSE) from Anna University in 2009 and pursuing Ph. D (CSE) from JNTUA. He is with the department of Computer Science & Engineering and as an Associate Professor, CMR Institute of Technology, Bangalore. He presented more than 15 papers in International and national Conferences. His research interests include Image Processing, OOMD, Software Engineering and Data Warehousing. He is a life member in ISTE.

