A Probability based Energy Efficient Clustering Protocol in Wireless Sensor Network

Richa Agarwal ; Amit Kumar Gautam

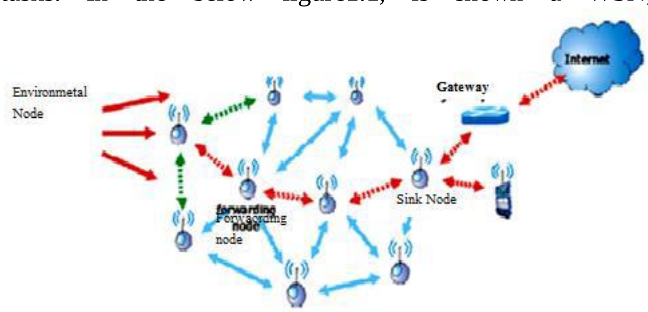
Abstract

The goal of this paper is to devise an approach that can work well in heterogeneous environment of sensor nodes. Heterogeneity in terms of energy is considered. In our scheme, we presumes that nodes varies on the content of their energies. So our algorithm effectively chooses cluster-heads(CH) for the successive round clusters formed in the network. While selecting a CH, parameters under consideration besides overall network's energy are node's initial and left-over energy. On the basis of these energies, the probability of node to be chosen as CH is defined. Thus when decisions are needed to be taken on the option of node for the charge of CH, node's energy play a vital role. As cluster-heads are chosen on the basis of energy left, this prevents the weaker nodes i.e. low energy node to die earlier or charging the single node with the responsibility of CH, resulting in earlier death of that node. In this way, network resources are managed effectively

***Keywords*- Wireless Sensor Network ,Cluster Head (CH) Distributed Unequal Clustering, Multi-hop**.

Introduction

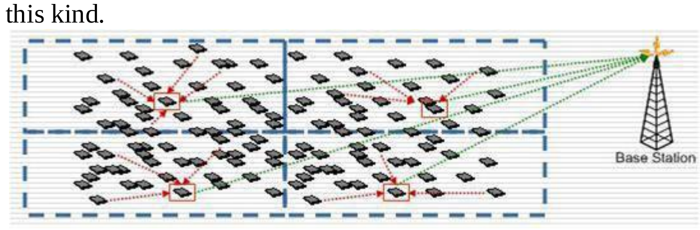
Sensor nodes are deployed for pervasive computing. Each of the sensor node is capable of limited amount of processing, upon coordination with the other node's information, gains the power to accomplish complex functionalities. Thus in Wireless Sensor Network (WSN), a collection of the sensor coordinates well with each other, using wireless communication link, in order to pursue some specific tasks. In the below figure1.1, is shown a WSN,



Upon detecting the event, the source nodes sends the data to the sink node. Upon the receipt of data , the sink node either process it or saves it in its memory. it is possible that sink node is PDA or sensor node in the WSNs. One of the limitations in the large WSNs is that due to multi-hops communication, data cannot be directly transmitted to the sink node from the source nodes .

A. Design Issues of WSN

Energy : Almost every sensor node deployed in the network operates on battery. So often its not practical to replace the dead batteries. Thus one of critical issues of WSN is energy efficiency. The MAC protocol must operate the motes in appropriate fashion to accomplish the energy loss as less as possible. Latency : This factor varies with applications. Like ,in the real time scenarios, whenever a sensor node recognizes an event, it must report immediately, to the sink, for appropriate actions. Fairness : This ensures that every node accessing the same channel get equal chance to get the access to the channel. Throughput : Throughput varies according to the application. In some case , accuracy of result depends upon the large amount of data, whereas in others , only single event is sufficient to report at sink All of the above, energy efficiency is the most crucial factor. Among three functionalities of network i.e. among communication, sensing and computation, major part of energy is consumed during communication. An alternative found to this is that during low traffic load, when node sits idle without any data to transmit, its radio can be switched off, to preserve the energy. This technique targets to stimulate the network efficient and practicable. In this, groups or clusters are formed, such that instead to sink node, the data sensed by the cluster-member is transmitted to CH, which in turn then exercise some computation or unification on data prior sending it to sink node. Thus ceasing the expenditure of energy through data conglomeration and communication in local. LEACH [11] and PEGASIS[14], are popularly known algorithms of this kind.



Various strategies have been devised for prolonging the lifetime of the network. Most of these strategies work on the common principle, in which the node dissipates the energy in the most optimized way. As the node operates on the power supplied by the battery, thus the efficient use of energy is most crucial requirement of WSN. In terms of node's consumption of power, one of the most costly operations is radio communication. Each of the sensor node has a radio on it, which is responsible for the dissipation of energy. Besides transmission, energy is spent in receiving and also in computation or operations such as data fusion being performed at the node

LITRETURE SURVEY

The various existing clustering algorithms can be divided into following categories.

Aggregating Data

Data aggregation basically means that the data is collected from different sensors, some fusion or computation is done on the data, and communicate the computed data to the sink node, thus reducing the possibility of transmission of supererogatory data [6]. Usually, the main purpose of the data aggregation approach is to minimize the consumption of the node's energy as far as possible and thus prolonging the lifetime of the network. There are 2 methods defined of data conglomeration for WSN.

- Cluster-heads gather sensed data from their candidate members, perform operations (if required) before transmitting to the base station [6].
- Gather data over each of the hop passing-by [6].

Rotating the CH's role

As the cluster-head's responsibility is to collect data from its candidate members, and communicate it to the sink node, thus energy expenditure of cluster-head is much more as compared to non cluster-head nodes. Thus the amount of energy consumption per cluster-head can be reduced by circulating the charge from one node to another , time to time and thereby ceasing the prematurely death of the cluster-heads.

Localization of cluster-head at the cluster centre

Besides equilibrating the cluster size [4], circulating the CH responsibility [7][8], another fashion of energy efficiency is emplacing the cluster-heads in the middle of the cluster [6][9]. In n-hop cluster, the candidate node transmits the data to cluster-head via nodes in the path. If cluster-head will be situated in the mediate of the cluster, then sending the data to the head will require less hops. Thus energy preservation can be accomplished.

Designating node with lowest power needed

For ceasing energy dissipation, the fashion that can be adapted is to depute the smallest energy required for communication to the candidate nodes in the cluster, and gateways with the energy required for communication with other clusters' gateways.

Equalizing the size of cluster Energy expenditure can be equilibrated by circulation the charge of cluster-head. Another way of equilibrating the energy consumption is by dividing the load among clusters itself. i.e. attempt is made sectionalize the network into clusters consisting of equal number of nodes.

Optimizing path based on cluster's power level

Among the existing clustering algorithms, most of them presumes uniform deployment of the sensors in WSN. Nonetheless, its not always true. It is possible that concentration of nodes varies in different areas of the network. Within an area, a node can communicate with another node [10] using low power of transmission. However, using low power of transmission, they may not communicate with the other area nodes. Thus the nodes can be unionized in clusters based on power level. This leads to energy preservation.

Designating node with lowest power needed

Eg. : Baras and Manousakis [5], proposed an algorithm which presumes that each candidate has potential to conform its transmission power, due to which lowest energy for communication can be deputed to them. This algorithm has following steps

Step 1 : Creation of clusters.
Step 2 : Deputize the candidate members with smallest transmission power necessary for communication within the cluster.[5]
Step 3 : Choose the gateways for each cluster.
Step 4 : Two or more cluster-heads intercommunicate with each other via gateways.
Step 5 : Deputize these gateways with minimum energy necessary for inter-cluster communication.
Step 6 : By deputing minimum required energy to the nodes, the over-all energy expenditure of the network can be foreshortened.
Step 7: End algorithm

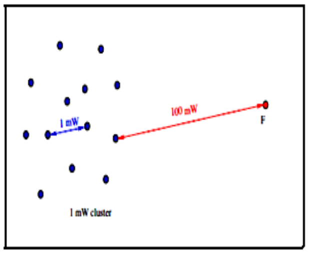


Figure 2.1 A common power level is not appropiate for non-homogenous networks

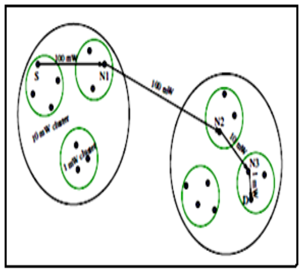


Figure 2.2 Routing by CLUSTERPOW in typical non-homogenous network

LEACH

WSN has LEACH [11] as its first and foremost hierarchical self-organizing algorithm based on clustering. This algorithm virtually sectionalizes the network into clusters or groups. Each cluster has certain number of nodes also known as candidate members. For each of the cluster formed, there is a devoted node known as ClusterHead(CH)

.

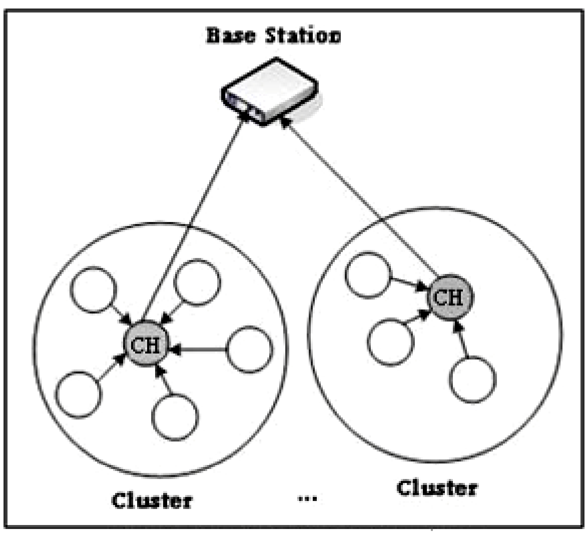


Figure 2.3 clustering in LEACH

Leach keeps on distributing the charge of CH sporadically among nodes because if CHs are kept fastened for the entire lifespan of the network, then these CHs will perish earlier then non CH nodes due to heavy work-load. Thus Leach keeps on deputing the high-energy charge of CH among the nodes, so as not to run out single node's battery. Besides, in order to cease energy expenditure , Leach also executes the fusion on the accumulated data so as to reduce quantity of data being sent to far situated, sink node, thus prolonging the network's life-time. As can be seen from above figure 2.4 this algorithm exercises the idea of rounds. Each round splits the functionality into two steps [11]

A. Setup Phase

In setup phase, selection of CH is done. During this phase, a stochastic number is opted between 0 and 1 by the node. The node took the responsibility of CH, if the *threshold T(i)*

$$T(i) = \begin{cases} \dfrac{P}{1 - P(r \bmod \frac{1}{P})}, & if\ i \in G \\ 0, & otherwise \end{cases}$$

r : current round

P : probability of node to become CH

G : set of nodes those have not been CH since last 1/P rounds

transcends the chosen number. The T(i) is given as

where,

In the advertisement phase, an ADV message is broadcasted using CSMA by the CHs. Rest of the candidate nodes , consider this ADV message, to make decision which CH to join. Whosoever CH is closer to the candidate node, that CH is chosen.

In cluster set-up phase, a join message is send by the candidate node to the chosen CH. This gives an idea to CH that how many candidates members are there in its cluster. The CH then prepares and sends the TDMA schedule to its candidate members [11]. This informs when and which all candidate members are required to send the sensed data and meanwhile the remaining candidate members can turn-off their radio, thus contributing in the conservation of energy. Besides this, scheduling also guarantees the data to be sent without collision.

B. Steady Phase

In this phase, sensed data is transmitted by the candidate members in their scheduled slot, to the CH. Upon gathering of sensed data from candidate member, CH does computation on it and get off it to the sink node. CH suffers from high expenditure of energy, as sink nodes are usually locate very far. At a time, a node can execute either the role of candidate member or the CH. If a node behaves like CH, then it first pass around the information of itself being CH. It then receives the join packet from all non CH nodes who are willing its cluster's member. Upon knowing about all its cluster-members, it creates a TDMA schedule and broadcast it to its candidate members. If a node behaves like cluster member then it receives the ADV packets, and select join the cluster of CH with high intensity packet. Afterwards, it operates in accordance to TDMA Schedule.

W. Mardini et al [13], turns up with another approach of betterment over Leach. In the LEACH,as compared to candidate node, the devoted node , i.e. the CH, has extra responsibility of data accumulation, computation and transmission of conglomerated data. Therefore, it is possible that before the periodic selection of next CH, the present CH drops dead, leaving the cluster useless. V-Leach, comes up to address this problem by introducing the Virtual CH for each of the cluster. Virtual CH has all the privileges as CH. Thus in case, if CH drops dead due to heavy duty, then Virtual CH took over the charge of CH, making cluster always associated to sink node as well as network. W. B. Heinzelman et al [3], turns up with idea with claims of betterment over Leach, debating the fact that Leach doesn't provide assurance about Ch's spot. In this protocol termed as Leach-C, each node, using GPS, informs the sink node about its positioning as well as its left-over energy. Using these info, sink node make decision of which nodes to appoint as CHs for next round so that the over-all energy of the motes is less dissipated and load is properly staggered among the clusters. A. Bestavros, I. Matta, G. Smaragdakis, [15], comes up with the approach for network with energy heterogeneity. In this , network consist of kinds of nodes, varying on the content of their initial energy. Normal nodes consist of energy E_0 and advanced node with $E_0(1+a)$. The CH probability and the rotating epoch directly counts on initial energy. However this algorithm, goes well for heterogeneity of two-level, but it fails for further level of heterogeneity.

PROPOSED APPROACH

One of the major drawbacks of leach is that it consider network to be homogeneous, nevertheless this is not the event in pragmatic scenario. It may be possible that initially i.e. at the time of deployment all nodes may have the same energy Eo, but after a certain time period, when some computation has been performed, it may be possible that only some of the motes energy is consumed while others have almost same energy as initial energy because no computation has been performed on them. This leads to heterogeneous network, where nodes differ on the energy level. Another case may be that after certain time-period of deployment, again some new nodes are deployed with initial energy Eo, in the network, and the already existing nodes has lesser energy due to computation being performed on them, thus making the network to be heterogeneous. Thus unlike leach, we will consider the heterogeneous network. Another thing is that as leach consider network to be homogeneous, thus each mote in the network has the equal probability to become the cluster-head. However in our approach, as network is heterogeneous, each mote has different probability to become cluster-head. This probability depends on the left-over energy of the mote at that time and also on the average energy of the network. In the successive subsection, we will look at on all the above mentioned factors one by one.

A. Model of Heterogeneous Network

This section will gives the brief information about the model of the heterogeneous network, we have considered in our approach. Assume the network to be M x M square region, as shown in the figure below. N number of sensors are haphazardly positioned in the network. The nodes are necessitated to transmit data to the base station. Base stations are often located very far from the network. Thus the network is virtually divided into clusters, each clusters consisting of one cluster-head. The cluster nodes transmit their data to their cluster-head. These cluster-heads in turn, performs fusion of data and transmit the conglomerated data to the base station. First we will find out the energy for the network with two-level heterogeneity, and then will generalize the formula for the network with n-level heterogeneity.

B. Two-level heterogeneity

Two types of motes are randomly deployed. Normal node and Advanced node. Let the normal node's initial energy is E_0 .The initial energy of advanced node is a $E_o(1+ma)$ times more then other nodes i.e. If there are m advanced nodes, then there is am times extra initial energy in the network. The total initial energy for such network is

$$E_{Total} = N(1-m)E_o + NmE_o(1+a)$$
$$E_{Total} = NE_o(1+ma) \quad \ldots\ldots\ldots\ldots(1)$$

C. N-Level Heterogeneity

- N Motes are deployed randomly in the network.
- Each mote has its initial energy in the closed interval of . $[E_0, E_0(1+a_{\max})]$
- The maximum value of energy is limited by a_{max} .
- Thus the initial energy of node s_i in the network, is , where $0 \le a_i \le a_{max}$
- Therefore the total initial energy of the network is

$$E_{Total} = \Sigma_{Ni=1} E_o(1+a_i)$$
$$E_{Total} = E_o(N + {}_{i=1}\Sigma_N a_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

D. Computing the Energies

The energy model , we are going to use is already proposed. Suppose the node has to send the message of l bit of length, to another node at the distance d, then the energy spent by radio can be given as,

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\varepsilon_{mp}d^4, & d \geqslant d_0 \end{cases}, \quad \ldots\ldots\ldots\ldots(3)$$

Where, E_{elec} :energy required to run the circuitry

d : separation between receiver and transmitte

l : depends on model of transmitter amplifier

Thus the total energy exhausted in the network per round is

$$E_{round} = L(2NE_{elec} + NE_{DA} + k\varepsilon_{mp}d^4_{toBS} + N\varepsilon_{fs}d^2_{toCH}) \quad \ldots(4$$

E_{DA} : cost incurred for data aggregation on the cluster-head

d_{toCH} : average distance between cluster-head and base station

d_{toBS} : average distance between cluster-head and its cluster-members

assuming the uniform distribution of nodes , we get,

$$d_{toCH} = \frac{M}{\sqrt{2\pi k}}, \quad d_{toBS} = 0.765\frac{M}{2}. \qquad \ldots\ldots\ldots\ldots\ldots\ldots(5)$$

As in our approach, cluster-head probability depends on the average energy of the network, we give an estimate to calculate it

$$\overline{E}(r) = \frac{1}{N}E_{total}\left(1-\frac{r}{R}\right) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(6)$$

Where, ***r*** current round
R max no. of rounds

E. Selection of Cluster-Head

In Leach, each node has same probability p_{opt}, to become the cluster-head. Thus, the node get one chance to behave like cluster-head, in every **1/p_{opt}** rounds. Our approach, on the contrary to this, consider that each mote has probability dissimilar to other node's probability, to become the cluster-head. This probability for each mote depends on its left-over energy as well as average energy of the network. At round ***r***,

Ei(r) denotes the left-over energy for node si

pi denotes the average probability to become cluster-head in ni rounds where pi = 1/ni

$\overline{\text{E}}$(r) denotes the average energy of the network at round r

Now, the probability of a mote to be cluster-head can be given as,

$$p_i = p_{opt}\left[1 - \frac{\overline{E}(r) - E_i(r)}{\overline{E}(r)}\right]$$

$$p_i = p_{opt}\frac{E_i(r)}{\overline{E}(r)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

In every round, using threshold probability, mote finds out, whether it can become the cluster-head for that round or not. When the mote has sufficient energy to become cluster-head, it randomly chooses a number from the closed interval ranging from 0 to 1. If the threshold probability is greater than the chosen number of the mote, then that mote becomes the cluster-head for that round. The threshold probability is defined as ,

$$T(s_i) = \begin{cases} \frac{p_i}{1-p_i(r \bmod \frac{1}{p_i})} & \text{if } s_i \in G \\ 0 & \text{otherwise} \end{cases}, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

where ***G*** contains node s_i, which have sufficient energy and has not become cluster-head for last n_i rounds.

F. Calculating the Probabilities.

As we know that the nodes have different energy in the heterogeneous network, thus probability to behave like cluster-head of these nodes also varies. Considering the network with two-level heterogeneity, we calculate the probability of nodes as [13],

$$p(s_i) = \frac{p_{opt}}{1+am} \qquad \text{if } s_i \text{ is the normal node}$$

$$p(s_i) = \frac{p_{opt}(1+a)}{(1+am)} \qquad \text{if } s_i \text{ is the advanced node}$$

Deputizing these values in eq (7), we get,

$$p_i = \begin{cases} \frac{p_{opt}E_i(r)}{(1+am)\overline{E}(r)} & \text{if } s_i \text{ is the normal node} \\ \frac{p_{opt}(1+a)E_i(r)}{(1+am)\overline{E}(r)} & \text{if } s_i \text{ is the advanced node} \end{cases}$$

Extending this derivation for n-level heterogeneity, we get

$$p(s_i) = \frac{p_{opt}N(1+a_i)}{\left(N+\sum_{i=1}^{N}a_i\right)}$$

putting this value in eq 7, we probability of being cluster-head as

$$p_i = \frac{p_{opt}N(1+a)E_i(r)}{(N+\sum_{i=1}^{N}a_i)\overline{E}(r)}.$$

Thus as threshold determines the chances of becoming the cluster-head on the basis of probability and this probability depends on left-over energy of the node as well as average energy of the network, threshold controls the energy to be spent in each round.

V. CONCLUSION AND FUTURE SCOPE

In this dissertation, I have devised a new approach based on clustering technique for efficiently managing the energy of the heterogeneous environment. In our approach, cluster-head selection depends on the left over and initial energy of nodes. By considering the factor over all network's energy, the expenditure of energy is efficiently curbed. Global information about the energy at every round is not necessitated by our approach and thus executes well in N- Level heterogeneity.

Main application of our approach is vehicle detection, health monitoring via machine, Environmental monitoring, Greenhouse monitoring, Monitoring the level of *etc*. However, besides energy heterogeneity, there are other prospects which also need to be handled. These include security of transmitted data, localization of cluster-head, topology consideration, QOS, etc. When sending data to the sink node which is usually situated far off from the network, security od transmitted data also becomes aspect of concern. It may be possible that intruder may manage to become cluster-head or pretends to original cluster-head to be sink node, thus imposing danger to the data. topology of the network also becomes fact of consideration , when it comes to localization of cluster-head. Sumhow knowing about the topology can help to choose CH in better way leading to cease in expenditure of energy.

References

[1] CHEN Xuhui, YANG Zhiming, CHENG Huiyan, '*Unequal Clustering Mechanism of LEACH Protocol for Wireless Sensor Networks'*, World Congress on Computer Science and Information Engineering, 2009.

[2] A.Allirani, M.Suganthi, '*An Energy Sorting Protocol (ESP) With Low Energy and Low Latency in Wireless Sensor Networks'*, IJCSNS International Journal of Computer Science and Network Security, VOL.208 8 No.11, November 2008.

[3] W. B. Heinzelman et al., '*An Application-Specific Protocol Architecture for wireless Microsensor Networks'*.

[4] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin and A. O. Fapojuwo, 'A Centralized *Energy-Efficient Routing Protocol for Wireless Sensor Networks*', IEEE Radio Communications, pp S8-S13, March 2005

[5] K. S. Manousakis and J. S. Baras, '*Clustering for Transmission Range Contrl and Connectivity Assurance for Self Configured Ad Hoc Networks'*, Proc. of IEEE MILCOM, pp. 1042-1047, 2003..

[6] H.Chen and S.Megerian, 'Cluster Sizing and Head Selection for Efficient Data Aggregation and Routing in Sensor Networks' ,Proc. Of IEEE WCNC 2006.

[7] W. R. Heinzelman, A. Chandrakasan and H. Balakrishnan, 'Energy efficient Communication Protocol for Wireless Microsensor Networks', Proc. of the 33rd IEEE Hawaii International Conference on System Science, Jan. 2000.

[8] J.-S.Liu and C.-H.R Lin, 'Energy-Efficiency Clustering Protocol in Wireless Sensor Networks' Ad Hoc Networks, vol. 3, pp. 371-388, May 2005.

[9] S. Ghiasi, A. Srivastava, X. Yang and M. Sarrafzadeh, 'Optimal Energy Aware Clustering in Sensor Networks', Sensors, pp. 258-269, Jul. 2002.

[10] V. Kawadia and P. R. Kumar, 'Power Control and Clustering in Ad Hoc Networks' , Proc. of IEEE INFOCOM, pp. 459-469, 2003

[11] M.J. Handy, M. Haas, D. Timmermann ,'Low Energy Adaptive Clusterin Hierarchy with Deterministic Cluster-Head Selection' , 2002.

[12] Hang Zhou, Zhe Jiang and Mo Xiaoyan, 'Study and Design on Cluster Routing Protocols of Wireless Sensor Networks', Dissertation,2006,

[13] National Bani Yassein, A. Al-zou'bi, Y. Khamayseh, W. Mardini, 'Improvement on LEACH Protocol of Wireless Sensor Network (VLEACH)', M. International Journal of Digital Content Technology and its Applications Volume 3, Number 2, June 2009.

[14] S. Lindsey, C.S. Raghavenda, 'PEGASIS: power efficient gathering in sensor information systems', Proceeding of the IEEE Aerospace Conference, Big Sky, Montana, March 2002,

Biographies

RICHA AGARWAL received the B.Tech degree in Computer Science and engineering from the Uttar Pradesh Technical University, the M.Tech. degree in Information Technology specialization of Software engineering.
agarwal.r3@gmail.com

AMIT KUMAR GAUTAM received the B.Tech degree in Computer Science and engineering from the Uttar Pradesh Technical University, the M.Tech. degree in Information Technology specialization of Software engineering. Currently working as a Assistant Professor in IMS Engg college ,Ghaziabad.
gautam.biet@gmail.com