IJCAES
ISSN: 2231-4946
Volume III, Special Issue, August 2013

International Journal of Computer Applications in Engineering Sciences
Special Issue on National Conference on Information and Communication (NCIC'13)
www.caesjournals.org

K-Anonymity based Privacy-Preserving location Monitoring Services for Wireless Sensor Networks

Bandana Kumari[1], G. Geetha[2] and L. Bhagyalakshmi[3]

[1,2] Dept. of CSE, Jerusalem College of Engineering, [3] Dept. of IT, Easwari Engineering College, Chennai, India

Email: 1. madhu26nov@gmail.com

***Abstract*- Wireless sensor technologies gave rise to many new applications widely used by general citizens as well as military operations. Numerous cases of these applications are based on the information of personal locations. Observations of these locations with untrusted server cause privacy threats to the individuals being monitored. To deal with such a privacy break, the concept of aggregate location information has been proposed using counting sensors, which would also pretence retreat breaches. To overcome this problem a new method, that is k-anonymity privacy method, wherein every person is indiscernible among k-persons, has been proposed in this paper. The objective this paper is to enable the system to provide high quality location monitoring services for system user, while preserving personal location privacy. The proposed system comprises two algorithms, namely, resource aware algorithm and quality aware algorithm where the former is used to reduce the communication and computational overhead cost while the later maximizes the accuracy of aggregate locations. Having the information of aggregate locations, the distribution of monitored persons is estimated using spatial histogram technique. Finally with the estimated distribution data the range queries can be answered to provide location monitoring services. Through the simulation results it is demonstrated that the proposed system provides a brilliant location monitoring services while ensuring location privacy of the monitored person.**

***Key words*- Aggregate location, Networks Anonymity, Privacy preserving, Quality aware algorithm, Resource aware algorithm, Spatial Histogram, Wireless Sensor.**

I. INTRODUCTION

Monitoring personal location with untrusted server pretences privacy threats to the individuals. Wireless sensor technologies have been widely used in numerous applications including monitoring and surveillance of civilians and military operations. Many cases of these applications are based on the information of personal location, for example, surveillance and location systems are observed by using counting or identity sensor. Using identity sensor each individual carries a signal sender/receiver unit with unique public identifier. With identity sensor, the exact indication of location of each monitored person is possible. Similarly the counting sensor, like photoelectric sensors and thermal sensors are installed to report the number of person. The sensor nodes in the identity sensors report the exact location information to the server. Thus using identity sensor monitoring is high but privacy is low, it immediately cause a significant privacy contravene. To deal with such a privacy break, the concept of cumulative location information has been developed. Aggregate location information in a collection of location data related to a category or a group from which individual identities have been removed. This is an effective approach to protect location privacy. Although the counting sensor by nature provide cumulative location information, they also create privacy breach. Fig. 1 provides an example of a privacy breach in a location monitoring system using counting sensor. Here, it is assumed that there are11 counting sensor nodes installed in 9 different rooms R_1 to R_9, including two hall ways C_1 and C_2. The nonzero number of objects detected by each sensor node is depicted as a number in

parenthesis. Fig. 1a and 1c provide the number respond by the same sensor node at two consecutive times T_1 and T_2 respectively. If R_3 is some object's office room, an adversary knows the particular object is in R_3 at time T_1 and also adversary knows that the object left R_3 at time T_2 and likewise, the adversary can infer that an object left C_2 at time T_2 and left to R_7. Such information leakage may result in several privacy threats knowing that a person has visited certain health rooms may know the entire activities in that location and also knowing that a person has visited particular hotel may reveal confidential information.

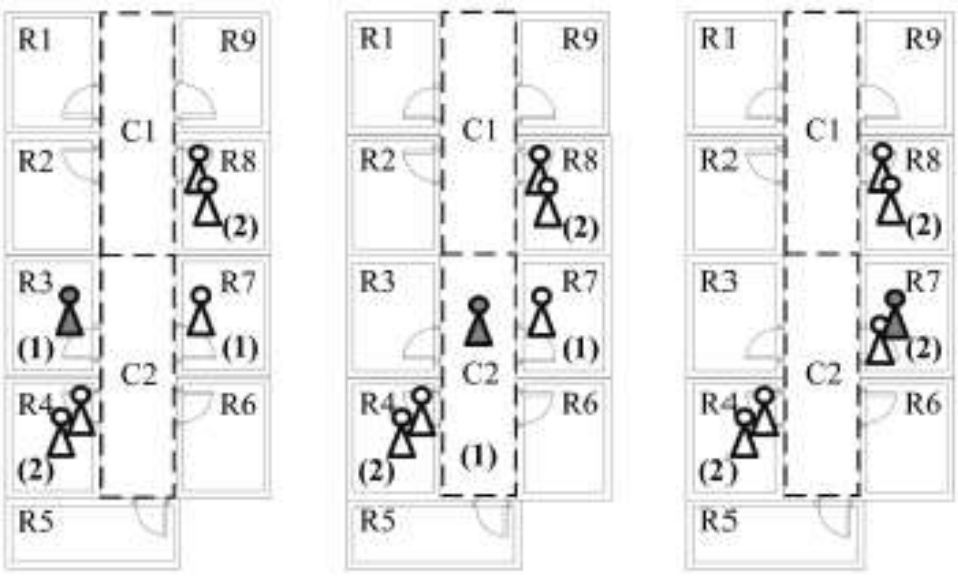

Fig. 1. Location monitoring system using counting sensor.

The organization of the paper is as follows: In section II, related work is presented. The proposed work is developed in section III. Required performance matrices to design the system are given in section IV. Finally the proposed algorithm is simulated in section V followed by conclusion in section VI.

II. RELATED WORK

Culler et al [5] reported environmental monitoring and condition-based maintenance has maintained. The main objective of the paper has described about dense instrumentation with actual time access and in network processing make a qualitative difference in this ability to perceive what is happening throughout large physical structure. It is self-organized network and conserving power and band width. An ideal location sensor for use in indoor environments would possess several important properties. Harter et al [9] provide fine-grain spatial information at a high renew rate, but would it also be low profile, cheap, scalable and robust.

Further, Kaupins G et al [4] show how the location technologies allow the employers to monitor the location of employees. The technologies range from global positioning systems capable to determine outdoor locations worldwide to sensor networks able to determine locations within buildings. Few international laws and no American laws directly address location monitoring.

Gruteser et al [11] reported the concept to hide the location of the message source and make it more difficult for an adversary to trace the messages back to the source location. The adversary observes the wireless communication within a certain detection range and traces towards the message source by moving, in each step, to the node that transmits the detected target information.

Onesystems Technologies **[**10**]** documented most vision-based counting techniques depend on detecting individuals. It has used for the concept of estimating the number of people in a crowded environment consecutively

to count, an idealistic intention in crowded settings. A geometric algorithm is introduced to calculate bounds on the number of persons in each region of the projection.

Sweeney. L [12] gave explanation for formal protection model named k-anonymity and set of supplementary policies for employment. A k-anonymity protection has introduce, if the information for each person contained in leave go of, cannot be distinguished from at least k-1 individuals whose information also appears in the release.

After an exhaustive literature survey on privacy preservation, a k-anonymity based location monitoring services with guaranteed location privacy algorithm is developed in this paper.

III. PROPOSED WORK

In this paper propose a K-anonymity based privacy preserving location monitoring system for wireless sensor networks that provide excellent and privacy controlled monitoring services. The system is based on the well-established privacy concept K-anonymity, which states into a cloaked area, where at least K person are resent. Aggregate location information is reported by each sensor node, which is in a cloaked area A, with the no. of object N which is located in the cloaked area. So that each clocked area A contains at least K person, then N, K is reported to the server. A small k indicates less privacy protection as a smaller cloaked region will be reported by the sensor node. These ensure superior monitoring service, on the other hand, a larger K marks in a larger area will reduce the quality of location monitoring services, but it ensures better protection of privacy. The privacy leakage can avoided by the system while providing low-quality location monitoring service especially for small areas may possibly allow adversary to follow users as shown in Fig. 1. hence, it provides high quality services for larger region. The resource aware algorithm goal to minimize both communications cost and quality aware algorithm helps to minimize the cloaked area, to minimize the accuracy of the collective location. Each sensor node in resource aware algorithm finds a sufficient number of an objects and it find the cloaked area. On the other hand quality aware algorithm takes place from a mapped area A, computed by the resource aware algorithm.

In this system each sensor node blurs its sensing area into cloaked area, in which at least K- persons are residing. Each sensor node reports only aggregate location information, which is in a form of cloaked area, along with number of person to the server, where the system can still report the no of person in a certain region, through answering range queries. A Spatial histogram examines the aggregate to evaluate the distribution of monitored objects in the system.

The system overview of this project is shown in Fig. 2. In this it has three main entities sensor node, server and system user. A set of sensor nodes $s_1, s_2, s_3, \dots s_n$ with detecting areas $A_1, A_2, A_3, \dots A_n$ rectangular region consist the sensing region of a set of sensor nodes S and N is the number of items present in that sensing area of the sensor node in S, such that $N_i \geq k$, $N_i = |\, \mathrm{U}_{s_j \in s_i} O_j \mid\ O_j = \{o_l \,|o_l \in a_j\}$, $1 \leq i \leq n$ and $1 \leq i \leq m$; and spatial histogram metod helps to answer the aggregate question Q that questions about the total number of objects in a certain region A. Area is reported by the sensor nodes based on the aggregate locations.

1) *Sensor node*

Each sensor node is responsible for determining the number of objects in its sensing area, blurring its sensing area into a cloaked area A, which includes at least k objects, and reporting A with the number of objects located in A as aggregate location information to the server.

2) *server*

The system only requires a communication n path from each sensor node to the server through a distributed tree each sensor node is also aware of its location and sensing area

3) *System Users*

Authenticated administrator and users can issue range queries to the system through either the server or sensor nodes or spatial histograms used to answer their queries .This proposed system has divided into four phases which is illustrated bellow.

A. Cloaked Area Setup

Each sensor node reports only aggregate location information, which is in a form of a cloaked area. The basic idea of this phase is that each sensor node blurs its sensing area into a cloaked area that includes at least k objects, in order to satisfy the k-anonymity privacy requirement.

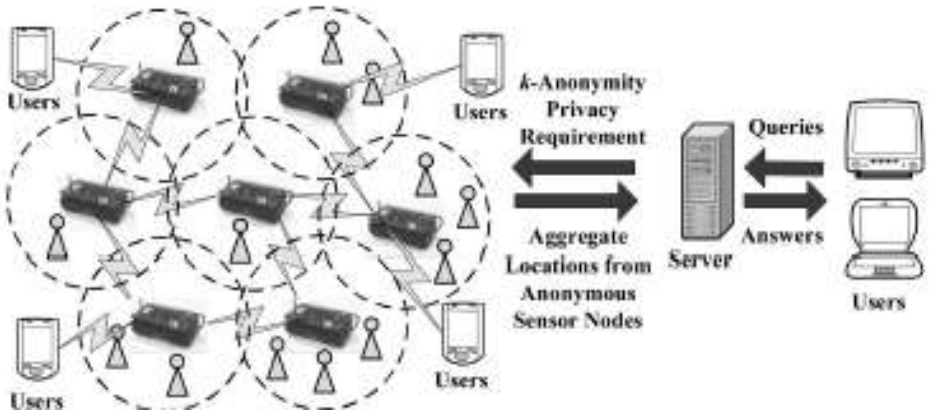

Fig. 2. System Architecture.

To minimize computational cost, this phase uses a greedy approach to find a cloaked area based on the information stored in Peer List. In a cloaked area setup it contains MBR (Minimum Bounding Rectangle), with the help of location anonymization algorithm namely resource aware algorithm, it covers the sensing area of the sensor nodes.

1) *The Resource Aware Algorithm*

The Resource Aware Algorithm illustrates the concept with 7 sensor nodes assuming from A to G. The required anonymity level is taken as 5 which means, $K = 5$. To prove this, we need a sensing area for the sensor nodes, and a dividing line between the two sensor nodes to indicate that the two sensor nodes can communicate directly with each other. The algorithm is executed in three phases.

Phase 1: *The broadcast Phase*

1. Function RESOURCE-AWARE (Int k, Sensor m, List R)
2. Add in the empty Peer List
3. Send msg by $m's$ identification $m.ID$, sensing region $m.Area$ and object count $m.Count$ to $m's$ neighbor peer
4. *if* Receive a mgs from a peer x, i.e., ($x.ID$, $x.Area$, $x.count$) then
5. Add msg to Peer List
6. *if* m gets required number of objects then
7. Send an alert message to $m's$ neighbors
8. *if* some $m's$ neighbor has not found required number of objects then
9. Transfer the message to $m's$ neighbors
10. *end if*

Phase 2: *The cloaked area phase*

11. $S \leftarrow \{m\}$
12. Calculate score for every peer in Peer List
13. Select peer with highest score from Peer List to S till the total no. of objects in S is at least repeatedly.
14. Area a min. surrounding area of the Senor in S
15. N, the total no. of objects in S

Phase 3: *The validation phase*

16. *if* no relation with Area and $R \in \mathcal{R}$ then
17. Forward $(Area, N)$ to remaining peers within the region & server

18. *else if* $m's$ sensing region is mapped by some $R \in \mathcal{R}$ then
19. Randomly select $R' \in \mathcal{R}$ where $R'.Area$ mapping $m's$ range
20. Forward R' to peers under $R'.Area$ and server
21. *else*
22. Return $Area$ with mapped N to peers in the $Area$ and the server
23. *end if*

B. Computation of Minimal Cloaked Area

A typical sensor network has a large number of sensor nodes; it is too costly for a sensor node m to gather the information of all the sensor nodes to compute its minimal cloaked area. To reduce communication and computational cost, m determines a search space, S, based on the input initial solution, which is the cloaked area, computed by the resource-aware algorithm, such that the sensor nodes outside S cannot be part of the minimal cloaked area. This step takes a set of peers residing in the search space, S, as an input and computes the minimal cloaked area for the sensor node m. In this phase it has propose two optimization techniques to reduce computational cost with the help of quality aware algorithm. The basic idea of the first optimization technique is that it does not need to examine all the combinations of the peers in S; instead, it only needs to consider the combinations of at most four peers. Because at most two sensor nodes define width of MBR and at most two sensor nodes defines height of MBR. Thus this optimization mainly reduces computational cost by reducing the number of MBR computations among the peers in S. The second optimization technique has two properties, lattice structure and monotonicity property. In a lattice structure, a data set that contains n items can generate $2n - 1$ item sets excluding a null set.

2) *The Quality Aware Algorithm*

The quality aware algorithm considers the cloaked area calculated by the Resource Aware algorithm as the initial solution. Now, this quality-aware algorithm refines it until the cloaked area has reached the minimum. Quality aware algorithm by using the input initial solution terminates the existing minimal cloaked region containing the set of sensor nodes that creates the minimal sensing area. This algorithm again has three phases.

Phase 1: *The search space phase*

1. Function QUALITY-AWARE (Int k, Sensor m, Set $init\_solution$, List L)
2. $current\_minimum\_cloaked\_area \leftarrow init\_solution$
3. Assume a search area S based on $init\_solution$
4. Get the information of peer located in the area

Phase 2: *The minimum cloaked region phase*

5. Add each node located in the area S to $C[1]$ as an object
6. Add m to each object in $C[1]$ as the initial object
7. *for* $j = 1; j \leq 4; j++$ compute
8. *for* each item set $X = \{a_1, \ldots a_{j+1}\}$ in $C[j]$ compute
9. *if* $Area\ S\ MBR(X)$ $< Area\ (current\_minimum\_cloaked\_area)$ then
10. *if* $N\ (MBR\ (X)) > k$ then
11. $current\_minimum\_cloaked\_Area \leftarrow X$
12. Delete X from $C[j]$
13. *end if*
14. *else*
15. Prune X out of $C[j]$
16. *end if*
17. *end for*
18. $if\ j < 4$ then
19. for each object pair $= \{x_1, \ldots x_{j+1}\}$,

$Y = \{y_1, \ldots y_{j+1}\}$ in $C[j]$ compute
20. if $x_1 = y_1, \ldots x_j = y_j$ and $x_{j+1} \neq y_{j+1}$ then
21. Add an object set $\{x_1 \ \ldots\ldots \ x_{j+1}, y_{j+1}\}$ to $C[j+1]$
22. *end if*
23. $end\ for$
24. *end if*
25. *end for*
26. $Area \leftarrow a$ minimum bounding rectangle of $current\_minimum\_cloaked\_area$
25: $N \leftarrow total\ number\ of\ items$ in $current\_minimum\_cloaked\_area$

Phase 3: *The validation phase*

26. if no relation with Area and $R \in \mathcal{R}$ then
27. Forward $(Area, N)$ to remaining peers within the region and server
28. *else if* $m's$ sensing region is mapped by some $R \in \mathcal{R}$ then
29. Randomly select $R' \in \mathcal{R}$ where $R'.Area$ mapping $m's$ range
30. Forward R' to peers under $R'.Area$ and server
31. *else*
32. Return $Area$ with mapped N to peers in the $Area$ and the server
33. *end if*

C. Construction of Spatial Histogram

A spatial histogram analyses the aggregate locations to evaluate the distribution of monitored objects in the system. A spatial histogram method helps to answer the aggregate question Q that questions about the total number of objects in a certain region A. Area is reported by the sensor nodes based on the aggregate locations. Sensor nodes determine the number of objects in their sensing area by mapping its sensing area into a cloaked region A that includes at least K objects. This finds the number of objects located in the mapped region.

1) *Spatial Histogram*

A spatial histogram that is embedded inside the server to estimate the distribution of the monitored objects based on the aggregate location reported from the sensor nodes. Spatial histogram is represented by a two-dimensional array that models a grid structure G of N_R rows and columns; hence, the system space is divided into $N_R \times N_C$ disjoint equal-sized grid cells. In each grid cell $G(i,j)$, it maintain a float value that acts as an estimator $H[i,j](1 \leq i \leq N_C, 1 \leq j \leq N_R)$ of the number of objects within its area. The accuracy of the spatial histogram indicates the utility of the privacy-preserving location monitoring system.

In the algorithm, initially assume that the objects are evenly distributed in the system, so the estimated number of objects within each grid cell is $H[i,j] = M/(N_R \times N_C)$. The input of the histogram is a set of aggregate locations R reported from the sensor nodes. Each aggregate location R in *C*ontains a cloaked area, $R.Area$, and the number of monitored objects within $R.Area$, $R.N$.

D. System Evolution

System evaluation contains the overall performance metrics, attacker model and setting the privacy preserving location monitoring system in a wireless sensor network.

2) *Attacker Model*

In common, the attacker model is defined as: Given an area A (that corresponds to the monitored area of a sensor node) and a set of aggregate locations $R = \{R_1, R_2, \ldots R_{|\mathcal{R}|}\}$ overlapping with A, the attacker estimates the number of persons within A. Since the validation step in the location anonymization algorithms guarantees that the containment relationship among the aggregate locations reported to the server does not violate the k-anonymity privacy requirement, it do not consider any containment relationship in R.

2) *Simulation Settings*

In all experiments, 30×30 sensor nodes that are simulated in uniformly distributed in a 600×600 system space. Each sensor node is responsible for monitoring a 20×20 space. It generates a set of moving objects that freely roam around the system space. The experiments consider $5{,}000$ moving objects that move at a random speed within a range of $[0, 5]$ space unit(s) per time unit, and the required anonymity level is $k = 20$. The spatial histogram contains $N_R \times N_C = 200 \times 200$ grid cells, and it issue 1,000 range queries.

IV. PERFORMANCE MATRICS

System is evaluated in terms of five performance factors

1) *Communication cost*

The measure of communication cost of the location anonymization algorithms is the average number of bytes sent by the sensor nodes during the reporting period. This measure also displays the network traffic and the power consumption by the sensor nodes.

2) *Cloaked region size*

It computes the quality of the average locations reported by the sensor nodes. The lesser the area, the better the correctness of the aggregate location.

3) *Computational cost*

This metric calculates the computational cost of the location algorithms in terms of the avg. number of the MBR computations.

4) *Query error*

This metric process the usage of the system, in terms of the relative error between the query reply, which is the estimated no. of items in the query area based on a spatial histogram, and actual reply M, respectively.

V. EXPERIMENTAL RESULTS & ANALYSIS

Here, analyzing the experimental results w. r. to the privacy protection & k-anonymity based privacy preserving location monitoring system for wireless sensor network. Here we represent a cloaked area where objects move between two different cloaked areas that are monitored by a sensor each. The sensor keeps collecting information about the objects moving in their respective cloaked areas and lose track of objects once they leave their cloaked area (figure. 3).

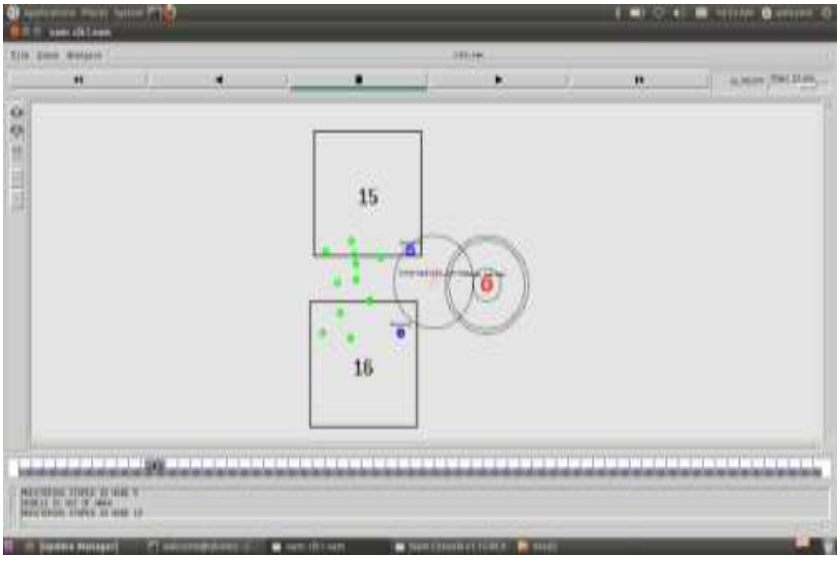

Fig. 3. Cloaked area monitored by a sensor node

(a) *Anonymization Strength*

Fig. 3. is the resilience of the system to the attacker model w.r. to the cloaked area and maximum mobility speed. The performance of the algorithms is shown in Fig. 4a to Fig. 4c. When the anonymity level gets restricted, the algorithms generate larger cloaked regions, which minimize the accuracy of the aggregate location reported to the server.

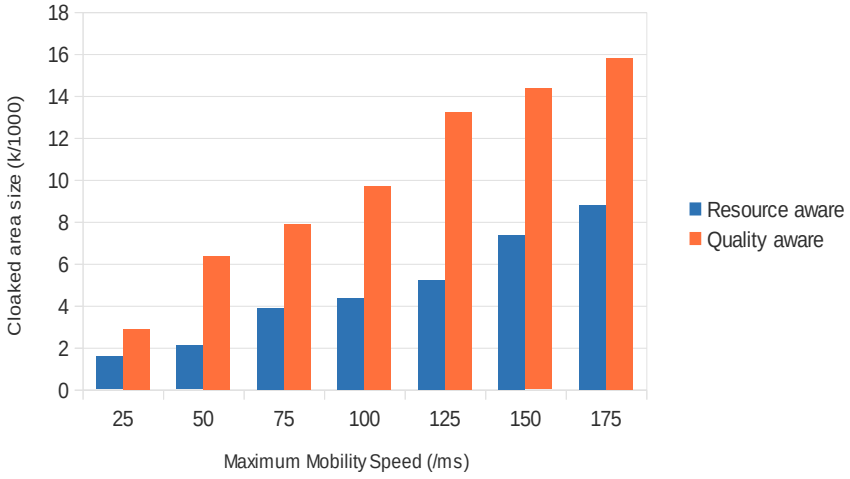

Fig. 4a: Cloaked Area Size

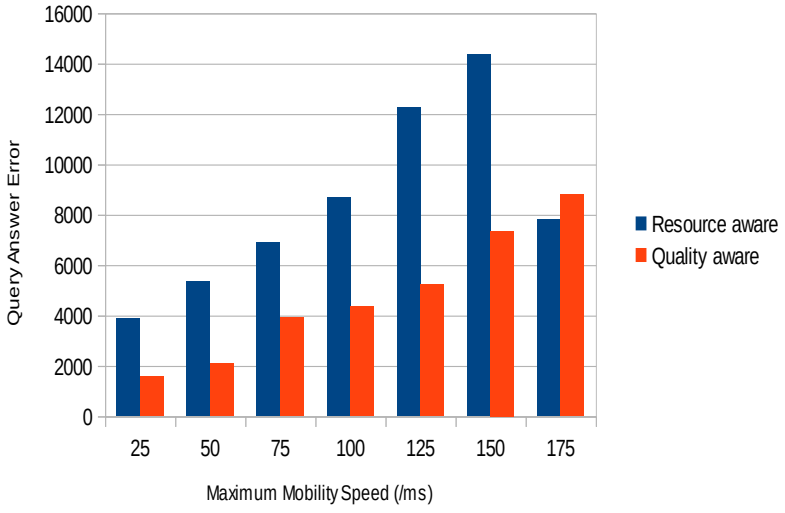

Fig. 4b: Communication Cost

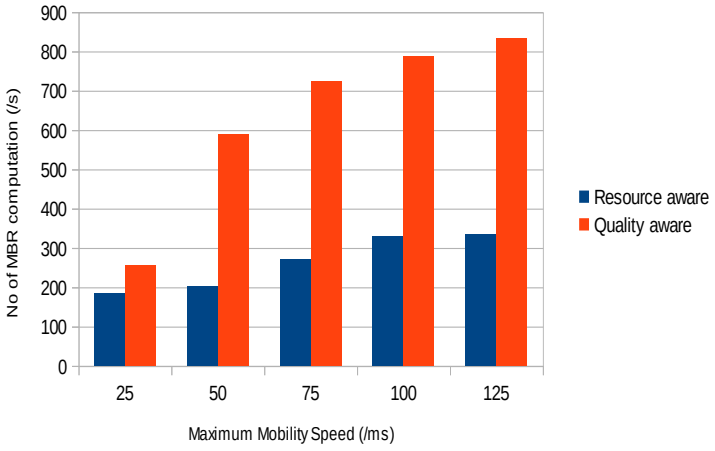

Fig. 4c: Computation cost

(b) *Effect of Query Area Size*

Quality aware and Privacy guard of the location monitoring system w. r. to increase in the query region size ratio range from 0.0010 to 0.1, where query region size ratio is the ratio of the cloaked region to the system area and the query region size ratio0.0010 that corresponds to the size of a sensor node's sensing area.

(c) *Effect of Increase in Number of Objects*

The performance of our system w. r. to increase in the no. of items from 3,000 to10,000.When the no. of objects increases, the communication cost of the resource aware algorithm is effected slightly, whereas the quality aware algorithm significantly lowers the communication cost

(d) *Effect of Movements*

Results show that increase in the object movement speed only slightly affects the communication cost and the cloaked region size. Since the resource-aware cloaked regions are less affected by the mobility speed, the object mobility speed has a very low effect on the required search area computed by the quality aware algorithm.

VI. CONCLUSIONS

In this paper, we propose a K-anonymity based privacy preserving location monitoring services system for wireless sensor networks. In this system two in-network location anonymization approaches namely, resource-aware and quality-aware algorithms has implemented, that not only provides location monitoring services but also preserves personal location details of persons who are going to be monitored. Both proposed algorithms rely on the well defined k-anonymity privacy theory which requires a person is uniquely identified among k persons. The resource aware algorithm helps to minimize communication cost and computational cost, while the quality aware algorithm helps to provide aggregate data with maximum accuracy.

REFERENCES

[1] Chi-Yin Chow, Mokbel, Tian, "A Privacy-Preserving Location Monitoring System for Wireless Sensor Networks," IEEE Transaction on mobile computing, vol.10, no 1, 2011.
[2] Traf-SysInc., "People Counting Systems," http://www.trafsys.com / products / people - counters / thermal - sensor.aspx, 2009.
[3] A.J. Stankovic M.F. Mokbel, T.F. Abdelzaher, S. Guo and T. He, "On Accurate and Efficient Statistical Counting in Sensor- Based Surveillance Systems," Proc. Fifth IEEE Int'l Conf. Mobile Ad Hoc and Sensor Systems (MASS), 2008.
[4] G. Kaupins and R. Minch, "Legal and Ethical Implications of Employee Location Monitoring," Proc. 38th Ann. Hawaii Int'l Conf. System Sciences (HICSS), 2005.
[5] D. Culler and M.S. Deborah Estrin, "Overview of Sensor Networks," Computer, vol. 37, no. 8, pp. 41-49, 2004.
[6] A. Jain, R. Han, and D. Grunwald, G. Schelle M. Gruteser, "Privacy-Aware Location Sensor Networks," Proc. Ninth Conf. Hot Topics in Operating Systems (HotOS), 2003.

[7] J. Kong and X. Hong, (2003), "ANODR: Anonymous on Demand Routing with Untraceable Routes for Mobile Ad-Hoc Networks," Proc.ACM MobiHoc, 2003.
[8] E. Schonberg, K. Bohrer, S. Levy, and X. Liu, "Individualized Privacy Policy Based Access Control," Proc. Sixth Int'l Conf. Electronic Commerce Research (ICECR), 2003.
[9] A. Harter, A. Hopper, A. Ward, P. Steggles, and P. Webster, "The Anatomy of a Context-Aware Application," Proc. ACM MobiCom, 1999.
[10] Onesystem Technology, "Counting People in Building,"http://www.onesystemtech.com.sg/index.php?option= com_ content&task=view&id=10, 2008.
[11] M. Gruteser, G. Schelle, A. Jain, R. Han, and D. Grunwald,"Privacy-Aware Location Sensor Networks," Proc. Ninth Conf. HotTopics in Operating Systems (HotOS), 2003.
[12] L. Sweeney, "Achieving k-Anonymity Privacy Protection UsingGeneralization and Suppression," Int'l J. Uncertainty, Fuzziness andKnowledge-Based Systems, vol. 10, no. 5, pp. 571-588, 2002.