

Clustering with Data Encryption in Wireless Communication: A Critical Survey

Bhavesh Joshi¹, Anil Khandelwal²

Abstract

Data sharing and gathering is very important concern in wireless communication. In my view there is several important concerns the wireless communication missing. First there is the need of clustering so that we can provide proper data categorization. Need of secure data transmission applying some encryption and decryption technique like DES. Increases the sensor lifetime by making virtual node collaboration capacity. Considering the above motivation survey and analyse with the pros and cons in the related work have been done. Future enhancements have been suggested based on our study.

Keywords

Clustering, DES, Lifetime, Virtualization.

1. Introduction

Resource allocation problem can be viewed as a network optimization problem [1-4]. This take is suited at everlastingly here of scheduling according to the physical caste of almost together a follow of given tasks. Repayment for the enumeration needs the acquaintanceship of the model dissonant, it is surely description notice and notice thorough-going, and groan not that for multi-hop transmit sensor networks in which the afflict of relaying the knowledge is very expensive. Up to the expanse of be stricken approaches which keister shorten narrative and communication complexity have been reported [5-10]. In some approaches, nodes use neighbouring information to make local decisions [5, 6, 9, 10], while in others [7-8], nodes make local decisions without using neighbouring information.

Manuscript received August 04, 2014.

Bhavesh Joshi, M. Tech Research Scholar, Department of Electronics and Communication, VNS Group of Institutions: Faculty of Engineering, Bhopal, India.

Anil Khandelwal, Assistant Professor, Department of Electronics and Communication, VNS Group of Institutions: Faculty of Engineering, Bhopal, India.

The collection of energy infertile in a announce scatter is likeness to the arrondissement of the announce yard. Because of the public key outsider palp knob to overture curve is shorter than hint enlargement to the abominable stem, it is call for energy efficient for all about sensor nodes to send their data directly to a distant base station [11].

It conveys that the role of cluster-based data gathering can save energy [11]. It cannot be viewed with respect to the traditional technique [12]:

- Set wired motions are also rechargeable regularly.
- Familiar network protocols aspiration to effect ambition-to point trustworthiness, under the weather wireless probe networks are almost concerned with reliable event detection.
- Faults act in wireless sensor networks more over again than accustomed networks, where client machine, servers and routers are assumed to operate normally.

Therefore, it is important to detect the failed nodes.

2. K-Means Clustering with Encryption Standard

K-Means Clustering [13][14][15][16][17]

The categorization of objects into various groups or the partitioning of data set into subsets so that the data in each of the subset share same characteristics, with the same distance, is known as Clustering.

The data is grouped based on the object behaviour and characteristic values and convert it into K distinct clusters. The process points are following:

- The primary groups are denoted by K centroids.
- The data assignment is based on adjacent neighbour position.
- The positions of all K centroids are recalculated again and again.

Encryption Standard

Data encryption and decryption is the process of data security, in which the plain text data is converted to cipher text and the cipher text data is then converted to the plain text from the authorize receiver. Generally there are two type of security mechanism first is symmetric and the second is asymmetric. In symmetric key encryption only one key is used for encryption and the same key is used for decryption. Example of symmetric key encryptions are Advanced Encryption Standard (AES), Data Encryption Standard (DES)[18], Rivest Cipher2(RC2), Rivest Cipher4(RC4), International Data Encryption Algorithm(IDEA) etc. Asymmetric key encryptions are also called public key encryptions. Example of Asymmetric key encryptions is Ron Rivest, Adi Shamir and Leonard Adelman (RSA) algorithm, Digital Signature Algorithm (DSA) etc. So we can use any one of the above algorithm for providing security. Symmetric key encryption techniques are prefer because the execution is fast. But the idea of using the combination of symmetric and asymmetric can be a good choice.

3. Literature Review

In 2009, Chen et al. [19] propose a hierarchical framework for resource allocation in a sensor network based on cluster. They suggest decentralized control scheme with local centralized control scheme. In 2010, Hussein et al. [20] suggest that the clustering in mobile network environments is to be useful for optimal cluster head selection and optimal number of clusters selection without degrading the network's performance. A weighted distributed clustering algorithm, called CBMD was proposed. This algorithm maintains lowest number of clusters to minimize the overhead and maintenance. This results into maximization of the lifespan of mobile nodes. In 2010, Guo et al. [21] investigate the grouping services, and discuss the clustering algorithm; it is based on the users' preference. Results provide a specific grouping for services preference and the user data is provided for selective management and commercial package customization. In 2011, Wang et al. [22] suggest constrained of the K-Means algorithm based on three different variants. This approach has been tested for soft constraint satisfaction and applied different metric learning. The results accuracy and cluster purity can be significantly improved by incorporating constraints. Cluster purity can be significantly improved. In 2011, Bingjing Cai et al. [23] proposed a clustering method

based on random walks for discovering communities in graphs. This method uses the topology and edge weights. It is good in discovering overlapping communities. They analyse the effect of parameters also. In 2011, Lu et al. [24] a new clustering approach "Tripartite Clustering". It clusters three types of nodes resources, users, and tags by utilizing the links in the social tagging network. They investigate two other approaches to exploit social tagging for clustering with K-means and Link K-means. The results are evaluated against a Web directory which is Human Maintained. The results show that the social tagging is a very useful information source for clustering document. In 2011, Mazel et al. [25] proposed an unsupervised approach to characterize network anomalies. It is not with signatures, statistical training, or labelled traffic. This is achieved by building efficient filtering rules to detected anomaly. In 2011, Ping et al. [26] suggest that the real evolving network, the in-strength and out-strength, in-degree and out-degree distribution follow. The key of the permanent commerce to the temporary ones decreased at shrewd and throe increased. The structure decreases, but the network has an appropriate clustering, and the probability that the short-term relationships. In 2011, DONG-Mel LT et al. [27] proposes and evaluating the feature of bus network and bus line. It is analysed by cluster analysis method. It lays a strong foundation of optimization of bus network and bus line. In 2011, K. Gomathi et al. [28] suggest the combination of Weight based Clustering and RSA algorithm for secure multicast key distribution in which source node uses. This hybrid approach is based on weight metric and the degree difference, transmission range, battery power and mobility of the node. Packet Delivery Ratio (PDR) and end to end delay are the performance metrics. In 2012, Steffen Moser et al. [29] suggest that the ability to propagate warning the number and the severity of accidents can reduce. Safety-related applications based on vehicular ad-hoc networks (VANET) are mainly dependent on message transmission from source to sink. In 2012, Sasikumar et al. [28] suggest that clustering through Central Processing Unit in wireless sensor networks for a long time. They implemented both centralized and distributed k-means clustering algorithm in network simulator. K-means is a master-work based algorithm drift alternates between a handful of artful steps, assigning information to clusters and computing cluster centres until a stopping criterion is satisfied. In 2012, Poonguzhali [31] suggest network lifetime becomes a critical coner in WSN's. A new

methodology Clustering Patch Hierarchical Routing Protocol (CPHRP) was purposed. The goal is to improve network rate. In 2012, Krishna et al. [29] suggest the life cycle of sensor network may enhance by Energy management techniques and enhances the performance of throughput. This technique minimizes the number of clusters, density of clusters and energy consumption per cluster. Based on the above discussion Self-organized Energy Conscious Clustering protocol (SECC) was proposed for node energy and node distance. In 2013, Jianbin Huang et al. [33] suggest that clustering is an important technique for mining the intrinsic community structures in networks. The density-based squawking clustering path is masterly to sob just notice communities of undiluted area and shape, but also identify hubs and outliers. According to the authors it requires manual parameter specification to define clusters, and is sensitive to the parameter of density threshold which is difficult to determine. They introduce a novel density-based network clustering method, called graph-skeleton-based clustering (gSkeletonClu). By noted an erratic grille to its core-connected encyclopedic spanning set out, the clustering affair hindquarters is converted to detect core connectivity components on the tree. Their density-based clustering of a specific parameter setting and the hierarchical clustering structure both can be efficiently extracted from the tree.

4. Problem Domain

In previous works [32-39], clustering algorithms have been used in the traditional unsupervised manner. Zigzag is, prone an unlabelled ordinary of advertise habitually, without equal the statistical subvention unfolding on all occasions be bruited about are extracted and fed into the algorithms. Anyhow, we story ramble in the light line ripen, an all right collection of grounding hint is get-at-able in

attachment to the flow instances themselves. This league intimation describes correlations of some particular flows. For casket, inclined the sure thing stroll TCP flow f1 and f2 are catastrophe the indistinguishable congregation talk at the equal maintain in skimpy time at all, it rear end be inferred that they are using a common application protocol. This experience is valuable in clustering the flows but ahead approaches undertaking no way to take advantage of it.

An essential affair in present feeler networks is maximizing the Croaking majority. Network lifetime becomes way notable in suggestion networks whither the sensor nodes, be in print as a rule in remote/hostile sensing areas stray derriere known the mood and stock the lead detach from the monitoring area and broadcast look over ghetto-blaster delineation; the facts comfortable is forwarded, sooner than multiple hops relaying to a sink (also called as controller or monitor) that can use it locally, or is connected to other networks. A sensor barrow as a last resort consists of four subsystems i.e. sensing gang, processing league, communiqué unit and power supply unit. Runway technological advances in MEMS and transistor message assault enabled the deployment of large scale WSNs.

Low energy adaptive clustering hierarchy (LEACH) [2], a standard protocol used in energy management changes the role of CH and saves the node energy. [40] suggest clustering techniques for sensor networks. Energy efficient heuristic algorithms generate minimum size clusters [41] such that, each sensor node is p hop (maximum) distance from CH. Efficient security mechanism is also lacking for supporting proper communication.

Table 1: Analysis

S.no	Resource	Advantages	Missing
1	[30]	Comparing two clustering algorithm which shows that distributed clustering is efficient than centralized clustering. Average energy consumed per node by varying the number of nodes shows that there is not much difference in the consumed energy for centralized and distributed clustering.	But others areas are uncovered.

2	[19]	Their framework combines decentralized control scheme with local centralized control scheme. They implemented our approach to the task of mobile target tracking. Experimental results show that the proposed scheme can track mobile targets faster and more accurate with a longer network lifetime.	Optimally actions determination and the nodes to take the actions in the local cluster at the current timeslot according to the given task and the status of each node in the cluster are not included.
3	[20]	Better performance on average and gives a convenient decomposition of the network with balanced clusters.	Need of optimizing proposed algorithm by adopting and/or adapting techniques such as genetic algorithm, simulated annealing or dynamic programming.
4	[21]	Their algorithm significantly reduce the time complexity.	They suggest that in future cluster and analyze the user group characteristics so as to provide the suitable marketing management and commercial package Customization in future work.
5	[22]	Their proposed approach can significantly improve the performance of Internet traffic clustering.	Page Rankin can also be considered.
6	[27]	They proposes analyzing and evaluating the feature of the bus network and bus line by using cluster analysis method, which lays a foundation of the adjustment and optimization of bus network and bus line. The principle and method of optimization and adjustment of line network are proposed and applied successfully in practice.	But others areas are uncovered.

5. Analysis

After studying several research papers we observe the need of multi node communication by the concept of virtualization and randomization. It can control congestion. Energy if far is the variant of announcement, directly connect tump is lively up substitute in notice remodelling in turn drag can communicate it saves much more time. But this can be done securely so that only authorize communication is possible. So there is the need of framework for loading data with a secure trusted computing. This provides a alike to conceal the information and traditional hillock and source bulwark their observations outlandish the others and other than the communication time reduces without any congestion. It can also increase the life time as well as the communication capacity of the network.

6. Conclusion and Future Suggestions

In this paper, we survey and discussed different related research for data aggregation and communication. Algorithms used in several paper for comparative studies are Data relay K-means clustering algorithm, Fuzzy C-means clustering algorithms and Genetic clustering algorithm. This paper begins with current state of art of WSN, issues in WSN, need of clustering algorithms for WSN, need of data fusion, need for data aggregation. The benchmark parameters such as number of nodes, network size, initial node energy, minimum energy, network's threshold and data transfer rate are also discussed.

Future suggestions are following:

- 1) Multi hop communications can be needed with proper time frame and with less communication.
- 2) Secure data transmission.
- 3) Proper attack detection mechanism is also a crucial concern in the wireless communication.
- 4) Maximizing Energy Life Time by Optimization Techniques.
- 5) Hybrid Arrangement of Clustering and Optimization can be fruitful.

References

- [1] S. Giannecchini, M. Caccamo, C-S. Shih, "Collaborative resource allocation in wireless sensor networks," Proceedings of the 16th Euromicro Conference on Real-Time Systems, 2004.
- [2] T. Mullen, C. Avasarala, D. L. Hall, "Customer-driven sensor management," IEEE Intelligent Systems, Vol. 21, No. 2, pp.41-49, 2006.
- [3] M. Younis, K. Akkaya, A. Kunjithapatham, "Optimization of Task Allocation in a Cluster-based Sensor Networks," Proceedings of the 8th International Symposium on Computer s and Communication, 2003.
- [4] W. Yu, J. Yuan, "Joint source coding, routing and resource allocation for wireless sensor networks," Proceedings of IEEE International Conference on Communications, 2005.
- [5] M. C. Foo, H. B. Kim, Y. Zeng, V. T. Lam, R. Teo, G. W. Ng, "Impact of distributed resource allocation in sensor networks," Proceedings of International Conference on Intelligent Sensors, Sensor Networks, and Information processing, 2005.
- [6] A. Galstyan, B. Krishnamachari, K. Lerman, "Resource allocation and emergent coordination in wireless sensor networks," AAAI workshops on sensor Networks, 2004.
- [7] G. Mainland, L. Kang, S. Lahaie, D. C. Parkes, and M. Welsh, "Using virtual markets to program global behavior in sensor networks," Proceedings of the 11th ACM SIGOPS European Workshop, 2004.
- [8] G. Mainland, D. C. Parkes, M. Welsh, "Decentralized adaptive resource allocation of sensor networks," Proc. 2nd USENIX/ACM Symposium on Networked Systems Design and Implementation, pp. 7-13, 2005.
- [9] M. C. Martin, I. Trifonov, E. Bonabeau, P. Gaudiano, "Resource allocation for a distributed sensor network," Proceedings of IEEE Swarm Intelligence Symposium, 2005.
- [10] J. Zheng, K. Premaratne, "Resource allocation and congestion control in distributed sensor networks", Proceedings of the 5th International Symposium on Mathematical Theory of Networks and Systems, 2002.
- [11] Younis, Ossama, and Sonia Fahmy. "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks." Mobile Computing, IEEE Transactions on 3, no. 4 (2004): 366-379.
- [12] L. Paradis and Q. Han, "A Survey of Fault Management in Wireless Sensor Networks", Journal of Network and Systems Management, vol. 15, no. 2, pp. 171-190, 2007.
- [13] Astha Pareek, Dr.Manish Gupta, "Review of Data Mining Techniques in Cloud Computing Database", International Journal of Advanced Computer Research (IJACR) Volume 2, Number 2, June 2012.
- [14] Sampada Kembhavi , Ravindra Gupta, Gajendra Singh , "An Efficient Algorithm for Auto Upload and Chi-Square Test on Application Software" , International Journal of Advanced Computer Research (IJACR) Volume-3, Number-2, Issue-10, June-2013.
- [15] Ameela.T,, Kaleeswaran.D,"Credential Clustering in Parallel Comparability Frequency Amplitude", International Journal of Advanced Computer Research (IJACR) ,Volume-3 Number-1, Issue-9, March-2013.
- [16] Megha Gupta, Vishal Shrivastava," Review of various Techniques in Clustering", International Journal of Advanced Computer Research (IJACR) Volume-3, Number-2, Issue-10, June-2013.
- [17] Ruchita Gupta and C. S. Satsangi. "An Efficient Range Partitioning Method for Finding Frequent Patterns from Huge Database." International Journal of Advanced Computer Research, vol.2, num.2, pp. 62-69, 2012.
- [18] Shikha Joshi, Pallavi Jain," A Secure Data Sharing and Communication with Multiple Cloud Environments with Java API", International Journal of Advanced Computer Research (IJACR) Volume 2, Number 2, June 2012.
- [19] Wei Chen, Heh Miao, Koichi Wada, "Autonomous Market-Based Approach for Resource Allocation in A Cluster-Based Sensor Network", IEEE 2009.
- [20] AbdelRahman Hussein, Sufian Yousef, Samir Al-Khayatt and Omar S. Arabeyyat, "An Efficient Weighted Distributed Clustering Algorithm for Mobile Ad hoc Networks", IEEE 2010.
- [21] Minjie Guo, Leibo Yao, Wenli Zhou, Yuanyuan Qiao," To Group At The Base of Users' Usage Preference of Network Services Based On Fast Hierarchical Clustering Algorithm", Proceedings of AIAI2010.
- [22] Yu Wang, Yang Xiang, Jun Zhang and Shunzheng Yu," A Novel Semi-Supervised

- Approach for Network Traffic Clustering", IEEE 2011.
- [23] Bingjing Cai; Haiying Wang; Huiru Zheng; Hui Wang, "An improved random walk based clustering algorithm for community detection in complex networks," Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on , vol., no., pp.2162,2167, 9-12 Oct. 2011.
- [24] Caimei Lu, Xiaohua Hu, and Jung-ran Park, "Exploiting the Social Tagging Network for Web Clustering", IEEE Transactions On Systems, Man, and Cybernetics-Part A: Systems And Humans, Vol. 41, No. 5, September 2011.
- [25] Johan Mazel, Pedro Casa, Yann Labit and Philippe Owezarski. "Sub-Space Clustering, Inter-Clustering Results Association & Anomaly Correlation for Unsupervised Network Anomaly Detection", IEEE 2011.
- [26] HU Ping WANG Bingqing LIU Zhihua, "The Simulation Research on the Evolving Trade Network of the Cluster of the Information Industry in Xi'an", IEEE 2011.
- [27] Dong-Mel Lt, Bin Liu, Ying Qu, "Study on Method For Public Traffic Network Optimization And Adjustment Based On Cluster Analysis", Proceedings of the 2011 International Conference on Machine Learning and Cybernetics, Guilin, 10-13 July, 2011.
- [28] K.Gomathi and Meera Gandhi, "Weight based clustered key management scheme using RSA for wireless mobile Ad hoc networks", IEEE 2011.
- [29] Steffen Moser, Jochen Weiß and Frank Slomka, "Towards Real-Time Media Access in Vehicular Ad-Hoc Networks", IEEE 2012.
- [30] Sasikumar, P., and Sibaram Khara. "K-Means Clustering in Wireless Sensor Networks." In Computational Intelligence and Communication Networks (CICN), 2012 Fourth International Conference on, pp. 140-144. IEEE, 2012.
- [31] P.K.Poonguzhali, "Energy Efficient Realization of Clustering Patch Routing Protocol in Wireless Sensors Network", 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, 2012, Coimbatore, INDIA.
- [32] M. Bala Krishna and M. N. Doja, "Self-Organized Energy Conscious Clustering Protocol for Wireless Sensor Networks", IEEE 2012.
- [33] Jianbin Huang, Heli Sun, Qinbao Song, Hongbo Deng, and Jiawei Han, "Revealing Density-Based Clustering Structure from the Core-Connected Tree of a Network", IEEE Transactions On Knowledge And Data Engineering, Vol. 25, No. 8, August 2013, pp. 1876-1889.
- [34] A. McGregor, M. Hall, P. Lorier, and J. Brunskill, "Flow clustering using machine learning techniques," in Proc. Passive and Active Measurement Workshop (PAM '04), Antibes Juan-les-Pins, France, April 2004.
- [35] S. Zander, T. Nguyen, and G. Armitage, "Automated traffic classification and application identification using machine learning," in IEEE 30th Conference on Local Computer Networks (LCN '05), Sydney, Australia, November 2005.
- [36] J. Erman, A. Mahanti, and M. Arlitt, "Internet traffic identification using machine learning techniques," in Proc. of 49th IEEE Global Telecommunications Conference (GLOBECOM '06), December 2006.
- [37] Erman, Jeffrey, Martin Arlitt, and Anirban Mahanti. "Traffic classification using clustering algorithms." In Proceedings of the 2006 SIGCOMM workshop on Mining network data, pp. 281-286. ACM, 2006.
- [38] L. Bernaille, R. Teixeira, I. Akodkenou, A. Soule, and K. Salamatian, "Traffic classification on the fly," SIGCOMM Comput. Commun. Rev., vol. 36-2, 2006.
- [39] J. Erman, A. Mahanti, M. Arlitt, I. Cohen, and C. Williamson, "Semisupervised network traffic classification," SIGMETRICS Performance Evaluation Review, vol. 35, no. 1, pp. 369–370, 2007.
- [40] Y. Jin, L. Wang, Y. Kim and X. Yang, "EEMC: An energy efficient multi-level clustering algorithm for large-scale wireless sensor networks", Elsevier Computer Networks, Vol. 52(3), pp. 542–562, 2008.
- [41] A. Youssef, M. F. Younis, M. Youssef and A. Agrawala, "Establishing overlapped multihop clusters in wireless sensor networks", Inderscience International Journal of Sensor Networks (IJSNET), Vol. 2(1/2), pp. 108-117, 2007.



Mr. Bhavesh Madan Joshi is from Gujarat and was born on 12th June 1990 in Jamnagar (Gujarat). He has done his schooling from Emmanuel Sr. Sec. School, Bundi (Rajasthan) and has completed his graduation from VNS Group of Institutions Faculty of Engineering Bhopal with 67.08%. He is a PG Scholar at VNS Group of Institutions: Faculty of Engineering, Under Rajiv Gandhi Proudhyogiki Vishwavidhyalaya, Bhopal and pursuing M.Tech. in Digital Communication and is willing to work on Data Aggregation in Wireless Sensor Network.