Historical Data Processing of Wireless Sensor Networks Data to Reduce Overhead on Data Traversal (WSNRODT)

[1]D. Vendhan [2]Alagumeena A [3]Cithira S [4]Svetha M
[1]Assistant Professor
[1,2,3,4]Department of Information Technology
[1,2,3,4]Kamaraj College of Engineering and Technology, Virudhunagar, India

Abstract

The basic architecture of Wireless Sensor Networks is usually a hybrid type where it is a combination of infrastructure oriented and infrastructure less networks. The Communication from sensor to sensor head takes place through peer-to-peer architecture (infrastructure less) and the communication from Cluster Head to Base Station (BS) involves Broadcast Based (Infrastructure Oriented).This Hybrid architecture is to reduce the energy consumption of sensor nodes as it will be depleted soon when each sensor broadcasts sensed data to base station as and when it senses. Hence a cluster head will be elected for each cluster by considering the battery, memory and processing ability [2] [9]. All the sensors will send their sensed data to the cluster head in a peer-to-peer manner. Cluster head recovers the data and generates the signature using elgammal. Base station sends request to cluster heads of high & heterogeneous clusters and can receive the recovered data from cluster head by verifying signature. The encrypted binary packets are accumulated in BS and it is fed to a database after verification of packets from clusters.
Keyword- Wireless Sensor Networks, Cluster Head, Base Station

I. INTRODUCTION

Wireless sensor networks (WSNs) provide the flexibility of untethered sensing, but pose the challenge of achieving long lifetime with a limited energy budget, often provided by batteries. The proliferation of Wireless Sensor Networks (WSNS) in the past decade has provided the bridge between the physical and digital worlds, enabling the monitoring and study of physical phenomena at a granularity and level of detail that was never possible before [1]. Therefore, there exists a great need and opportunity to increase energy efficiency through distributed, in-network processing [8].

A core challenge in these networks is to minimize energy consumption [7]. Data prediction is proposed in WSNs to extend the system lifetime by enabling the sink to determine the data sampled, within some accuracy bounds, with only minimal communication from source nodes. This data prediction approach is applicable when data is reported periodically - the common case in many pervasive computing applications.

We proposed an efficient hybrid data prediction technique with data aggregation which can drastically reduce energy consumption of sensor nodes during communication [4]. In data prediction the communication can be significantly reduced by avoiding transmission of each raw sample to the sink [3]. This is achieved by using a model to estimate the sensed values, and by communicating with the sink only when there is a change in the sampled data when the aggregation time out is triggered in cluster head. Each node is using a model to predict its own sensor data, and compares the predicted values with those actually observed and generates a confidence value.

A. Problem Statement

Data prediction exploits the fact that many applications can operate with approximate data, as long as it is ensured to be within certain limits of the actual data. This allows huge reductions in communication. Existing results suggest that further reductions in data traffic would have little impact on lifetime, as network costs dominated by control operations. Therefore improvements must directly address the extremely low data rates of DBP (e.g.,) by considering radically different network stacks. Further, data loss in prediction-based systems has the potential to significantly increase application errors. Therefore reliable transport mechanisms must be revisited to ensure application-level quality.

- Energy depletion
- Security risks
- High overhead on transmission
- High overload on database.

II. EXISTING METHOD

Previous studies compute models that aim to reduce the approximation error (Root Mean Squared Error, RMSE) with respect to recent data points, producing models that are accurate with respect to past data. Instead, existing systems computes models capturing the trends in recently-observed data, producing models accurate with respect to future data. Evidently, these trends are best modelled when data exhibit short-term linear behaviour.

Data collection is a fundamental functionality of many WSN applications, and is commonly implemented by nodes periodically taking sensor measurements and reporting the corresponding samples to a data sink. Clustering on various types of sensor nodes and the security concerns for data protection with energy efficiency is not considered.

III. PROPOSED METHOD

We proposed an efficient hybrid data prediction technique with data Aggregation which can drastically reduce energy consumption of sensor nodes during communication. In data prediction the communication can be significantly reduced by avoiding transmission of each raw sample to the sink.

This is achieved by using a model to estimate the sensed values, and by communicating with the sink only when there is a change in the sampled data when the aggregation time out is triggered in cluster head. Each node is using a model to predict its own sensor data, and compares the predicted values with those actually observed and generates a confidence value.

1) In our proposed design we promise to give high authenticity of each sensing data and integrity of the same in a recoverable environment for Concealed Data Aggregation (CDA) by privacy homomorphism encryption Scheme using Ecc-Elgammal signature in a binary transmission for three completely different network clusters.
2) Base station can recover each sensing data as well as can compute on it. Overhead is greatly reduced as cluster heads of high and heterogeneous sensors can respond for base station requests.
3) So communication cost is drastically reduced that a low cluster network can also be deployed to a WSN [5].
4) Binary transmission of data, encryption using public key, private key, cluster key and thus creating a signature for each data and (two layer authenticity) verifying authenticity at both cluster heads and in base station helps to securely send data in huge WSN containing different clusters.
5) Query processing on historical data enables to visualize the data in database in a light weighted way both graphically and analytically [10].

A. Network Formation & Electing Cluster Head

Three completely different clusters are formed in a wireless sensor network named,

1) Homogenous Low Sensors Cluster
2) Homogenous High Sensors Cluster
3) Heterogeneous Sensor Cluster

Cluster heads are elected based on the battery, memory and processing ability for each cluster and only the cluster head aggregate the received data from its cluster sensors and sends to the base station [6].

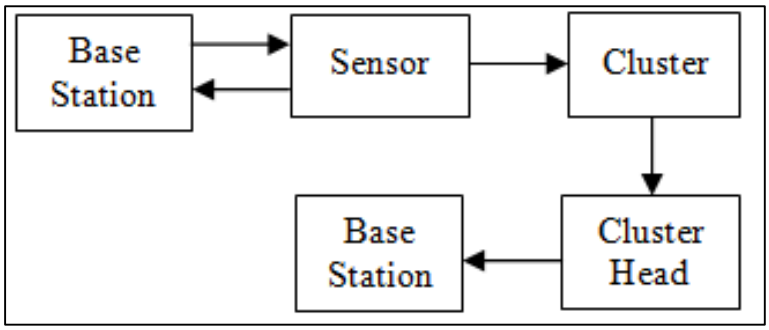

Fig. 1: Network Formation & Electing Cluster Head

The homogeneous high and heterogeneous clusters have the high capability sensors and hence can do computations on aggregated data and send the corresponding result and also respond to base station request thereby reducing overhead drastically. So homogeneous low clusters can also be incorporated into WSN since it contains low sensors which just aggregate and sends the cipher text to BS. Each node is equipped with 3 keys - in which bit length varies depending on type of cluster.

1) Cluster key
2) Private key
3) Public key

B. Data transmission and Aggregation

Sensors send its own sensing data to its cluster head as each and every sensor knows its own cluster head and generates a shortest path to reach it and transmits through it. Each sensed data is converted into a packet and is encrypted and the cipher is subjected to signature generation process.

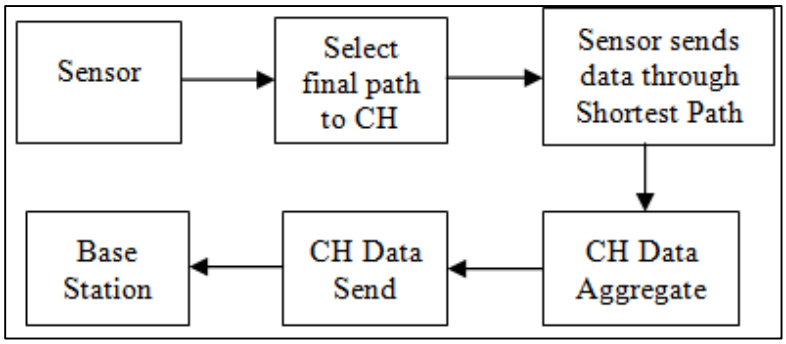

Fig. 2: Data transmission and Aggregation

The cluster head receives the encrypted cipher text and signature is verified and the data is aggregated. Cluster head recovers the data and in homogeneous high and heterogeneous and generated the signature using elgammal. For homogenous low and heterogeneous clusters only aggregation process takes place as homogenous low cluster is memory constrained.

C. Recovering Data using Signature

The Aggregated data are converted into a single packet when aggregation time out is triggered in cluster head. Now a cipher text is created by encryption and a signature is also generated for the same and the cipher is concealed in signature by compressing and converting the whole compressed content to binary. In homogeneous low cluster the head just verifies signature sends the data to BS without aggregating packets.

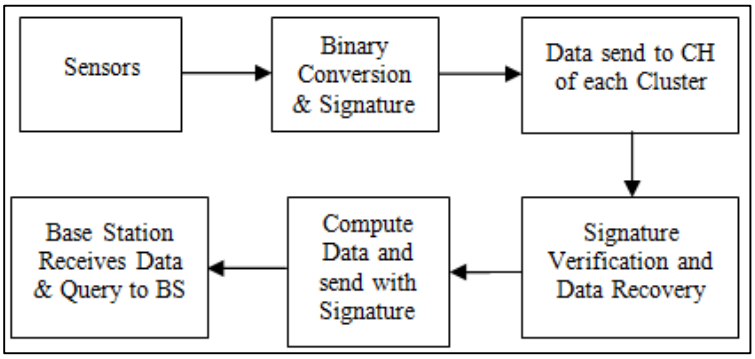

Fig. 3: Recovering Data using Signature

BS once again verifies each and every data by verifying signature thereby ensuring data integrity and authenticity. Base station can also send request to cluster heads of high & heterogeneous clusters and can receive the recovered data from cluster head by verifying signature.

D. Data Prediction and Query Processing

The Low Confidence data from the sensor nodes are dropped in cluster head by data prediction strategy on each cluster head. Hence the redundant data's are free from communication reducing overhead. The dropped packets are shown in a graph for redundancy evaluation by time versus drop count. It is applied to every sensor node for efficient data transmission towards the sink node.

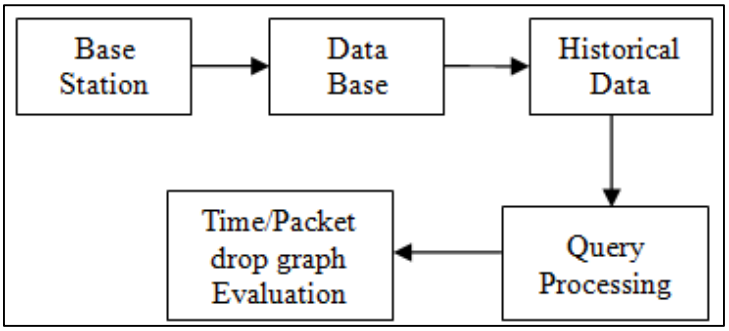

Fig. 4: Data Prediction and Query Processing

The Broadcasted compressed encrypted binary packets are being accumulated on base station and it is fed to a database after verification of packets from various clusters. The historical thus formed by time are subjected to four kinds of query processing.

1) Top-K based query processing (Top ranked values on each cluster).
2) Necessary set based query processing (Values should be present).
3) Sufficient set based query processing (Values that are more than enough).
4) Boundary based query processing (Ranked values in a range).

E. Equations

Minimum range of sensor = Distance of sensor - Range of sensor
Maximum range of sensor = Distance of sensor + Range of sensor
Coverage of sensor = Maximum range of sensor - Minimum range of sensor

IV. EXPERIMENTAL RESULTS AND ANALYSIS

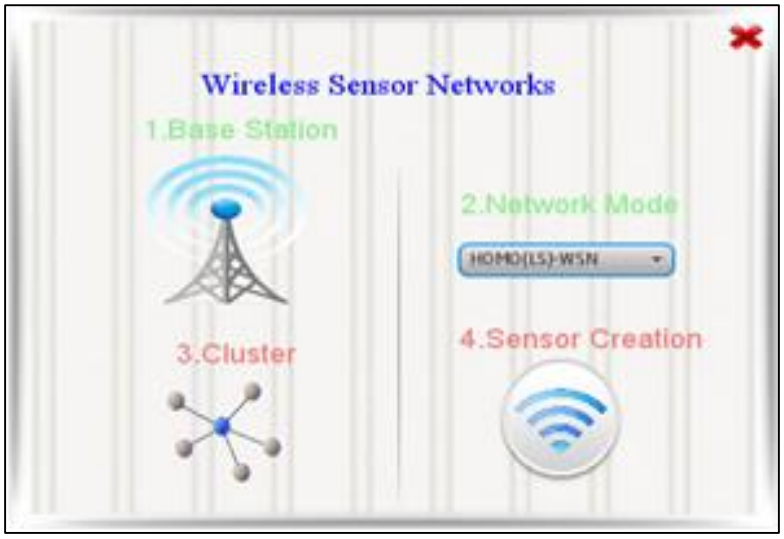

Fig. 5: Wireless Sensor Networks Home Page

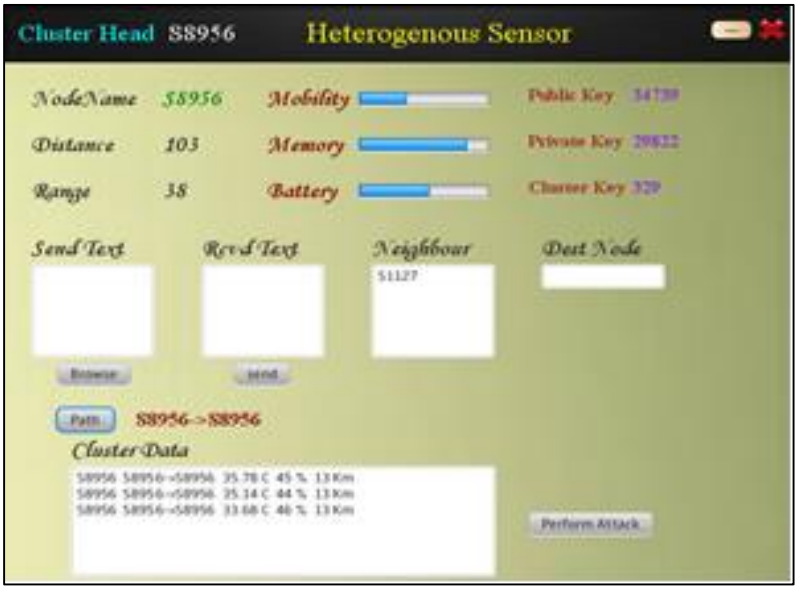

Fig. 6: Shortest Path Identification

Fig. 7: Base Station

V. CONCLUSION

In this paper we designed, developed and evaluated energy efficient, light weighted, secure and reliable system for hybrid WSN to transmit and process the historical data on different types of clusters using Elgammal signature scheme, aggregation methods. We applied data mining approach to remove redundancy and efficient data transmission from sensor node to sink node. We applied data prediction strategy on every sensor node for efficient data transmission towards the sink node. Dealing with huge amount of data transmission is a serious issue of WSN and thus we to introduced data prediction based reduction in wireless sensor network because of energy consumption is main limitation of WSN, in transmitting a data from source node to sink node.

VI. FUTURE ENHANCEMENTS

In this novel approach more work needed because each and every organization wants suitable technique for every kind of wireless sensor network. We can use the Naïve Bayes Prediction (NBP) model for wireless sensor network to reduce the data as well as saving the energy of network for better transmission of data. In future, we will take various WSN applications based on prediction. For query processing a distributed database allows faster local queries and can reduce network traffic.

REFERENCES

[1] T. Palpanas, "Real-time data analytics in sensor networks", I Managing and Mining Sensor Data, Aggarwal Ed. New York, NY, USA: Springer, 2012.
[2] H. Jiang, S. Jin, and C.Wang "Prediction or not? An energy efficient framework for clustering-based data collection in wireless sensor networks" IEEE Trans. Parallel Diatribe, Syst., vol. 22, no. 6, pp. 1064–1071, Jun 2011.
[3] D. Tulane and S. Madden, "PAQ: Time series forecasting for approximate query answering in sensor networks" in Proc. Eur. Workshop Wireless Sen. Newt. , 2006, pp. 21–37.
[4] R. A. F. Mini, M. D. V. Machado, A. A. F. Loureiro, and B. Nath, "Prediction-based energy map for wireless sensor networks" Ad Hoc Netw., vol. 3, pp. 235–253, 2005.
[5] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering" IEEE Trans. Wireless Common., vol. 9, no. 12, pp. 3728–3738, Dec. 2010.
[6] J. Rousselot, A. El-Hoiydi, and J.-D. Decotignie, "Low power medium access control protocols for wireless sensor networks," in Proc. Eur. Wireless Conf., Jun. 2008, pp. 1–5.
[7] D. Chu, A. Deshpande, J. M. Hellerstein, and W. Hong, "Approximate data collection in sensor networks using probabilistic models" in 22nd Int. Conf. Data Eng., 2006, p. 48.
[8] Silberstein, G. Filpus, K. Munagala, and J. Yang, "Data-driven processing in sensor networks" in Proc. Conf. Innovative Data Syst. Res., 2007, pp. 10–21.
[9] M. Hassani, E. Mller, P. Spaus, A. Faqolli, T. Palpanas, and T.Seidl, "Self-organizing energy aware clustering of nodes in sensor networks using relevant attributes," in Proc. 4th Int. Workshop Knowl. Discovery Sen. Data, 2010, pp. 87–96.
[10] D. Tulone and S. Madden, "An energy-efficient querying framework in sensor networks for detecting node similarities," in Proc.Int. Conf. Modeling, Anal. Simul. Wireless Mobile Syst., 2006, pp. 191–300.