CHAPTER 2

LITERATURE REVIEW

Sensor network routing is a challenging one due to characteristics that differentiate it from contemporary communication and wireless ad-hoc networks. It is impossible to build a global addressing scheme to deploy the vast number of sensor nodes. So, classic IP-based protocols can’t be used for sensor networks (Akkaya & Younis 2005). Secondly, in contrary to typical communication networks sensor networks applications require sensed data flow from many regions (sources) to a specific sink. Thirdly, the generated data traffic contains redundancy as multiple sensors can generate similar data within near a phenomenon. This must be used by routing protocols to improve energy and bandwidth use. Fourthly, sensor nodes are highly constrained regarding transmission power, on-board energy, process capacity and storage and so need proper resource management.

Data aggregation algorithms aim at gathering data energy efficiently to enhance network life. WSN provides an attractive data gathering method in distributed system architectures having dynamic access through wireless connectivity. Data aggregation, aggregates sensor data via aggregation approaches. General data aggregation algorithms work as seen below, using sensor data from sensor node and aggregating data through aggregation algorithms like LEACH, centralized approach and Tiny Aggregation (TAG). Aggregated data is sent to sink node through selected path (Patil & Patil 2010). Literature survey for the following topics is discussed here:
• Cluster Based Routing
• Energy Aware Routing
• Evolutionary Algorithms Based Routing
• Data Aggregation in Wireless Sensor Networks
• Cluster Head Selection for Wireless Sensor Networks

2.1 DATA AGGREGATION IN WIRELESS SENSOR NETWORKS

Bhasker (2014) proposed a genetically derived secure cluster-based data aggregation in WSN. Initially the CHs were selected based on the node connectivity, which acts as a data aggregator. Then, the clustering process was executed using the genetic algorithm. When a cluster member wants to transmit the data to aggregator, a data encryption technique is utilised that offers authenticity, confidentiality and integrity. By simulation results, it showed that the proposed technique minimised the energy consumption, ensured data security and reduced the transmission overhead.

A sequential decision process model considering sample arrival instants randomness and uncertainty of availability of multi-access communication channel to analyze a problem and determine optimal decision policies with local information was proposed by Zhenzhen et al (2009). It revealed that when statistics of sample arrival and channel availability satisfied some conditions, there were optimal control-limit-type policies easy to implement. If required conditions were not satisfied, performance loss of using the new control-limit-type policies was characterized. Generally, a finite-state approximation was proposed and two on-line algorithms solved it. Practical distributed data aggregation simulations demonstrated the developed policies effectiveness which also achieved desired energy-delay trade-off.
Roy et al (2012) proposed the synopsis diffusion approach, against attacks which compromised nodes to contribute false sub-aggregate values. In large sensor networks, in-network data aggregation reduced communication and energy consumption. Recently, research community proposed a robust aggregation framework called synopsis diffusion combining multipath routing schemes with duplicate-insensitive algorithms to compute aggregates (predicate Count, Sum) despite message losses due to node and transmission failures. It specifically presented a new lightweight verification algorithm through which the base station determines if computed aggregate includes false contributions. Theoretical analysis and simulation proved that the new algorithm outperforming existing approaches.

A new in-network aggregation technique based on Chinese Remainder Theorem (CRT) which exploited the advantages of source coding when using a distributed approach without coordination among network nodes was proposed by Campobello et al (2013). WSNs need low complexity techniques to reduce bits flowing throughout the network, to decrease bandwidth waste and increase network life. For this, spatial correlation caused by deployed devices high density can be exploited. An appropriate CRT based network coding mechanism was combined with source coding to ensure lossless transmission from sources to sink.

A Data Aggregation and Authentication protocol (DAA), integrating false data detection with data aggregation and confidentiality was proposed by Ozdemir & Cam (2010). To support data aggregation with false data detection, every monitoring nodes of data aggregator’s conducted data aggregation and computed corresponding small-size message authentication codes for data verification as pair mates. To support confidential data transmission, sensor nodes between two consecutive data aggregators verify data integrity on encrypted data instead of plain data. Performance analysis
revealed that DAA detected false data injected by up to T compromised nodes, and that such false data are not forwarded beyond next data aggregator on path. Despite false data detection and data confidentiality increasing communication overhead, simulation showed that DAA could reduce transmitted data up to 60% through data aggregation and early false data detection.

Different channels were assigned to each sensor node on data aggregation tree as a non-overlapping channel is limited in wireless networks according to Yen et al (2009). Then, data aggregation routing along with channel assignment minimizing total transmission power under restricted non-overlapping channels was performed. The problem is NP-complete and devised a heuristic algorithm, Iterative Channel Adjustment Data Aggregation Routing algorithm (ICADAR) with other three heuristics to solve it. Simulation showed that ICADAR algorithm outperformed the 3 algorithms under experimental cases.

A cooperative Multiple-Input-Multiple-Output (MIMO) and data-aggregation technique to reduce energy consumption per bit in WSNs by reducing data for transmission and using network resources through cooperative communication was proposed by Gao et al (2010). This was a new energy model that considered correlation between nodes generated data and distance between them for cluster-based sensor network using the combined techniques. Using the model, cluster size effect on average energy consumption per node is analyzed. It is shown that network’s energy efficiency can be greatly improved in cooperative MIMO systems with data aggregation, compared to cooperative MIMO systems without data aggregation or data-aggregation systems without cooperative MIMO, when sensor nodes are properly clusterized. Both centralized and distributed data-aggregation schemes for cooperating nodes to exchange and compress data
were suggested and appraised, leading to diverse data correlation impact on energy performance of integrated cooperative MIMO and data-aggregation systems.

A packet attribute concept, defined as the identifier of data sampled by various sensors or applications was introduced by Ren et al (2013) which proposed an Attribute-aware Data Aggregation (ADA) scheme of a packet-driven timing algorithm and special dynamic routing protocol. Imbued by physics potential and ant colony pheromone concepts, a potential-based dynamic routing was got up to support an ADA strategy. Performance evaluation in varied scenarios verified that the ADA scheme made packets with similar attribute spatially convergent and improved data aggregation efficiency. Further, ADA scheme offered other properties, like scaling regarding network size and being adaptable to track mobile events.

Attention on amendment strategies was focused on by Feng et al (2011) which assumed that networks adopted connected dominating sets based aggregation scheduling as basic aggregation scheduling strategy. Amendment scheme includes localized aggregation tree repairing algorithms and distributed rescheduling algorithms. The former found a new aggregation tree for a corrupted node’s child, while the latter achieved interference free data aggregation scheduling after amendment. Amendment strategies impacted very limited nodes near corrupted node and amendment process was transparent to all nodes. Theoretical analyses and simulation proved the scheme improving data aggregation operation efficiency by reducing message and time costs compared to rebuilding aggregation tree and rescheduling the network.

Data aggregation scheduling problem was emphasized to reduce delay (or latency) by Xu et al (2011). It suggested an efficient distributed algorithm that resulted in a collision-free schedule for WSN data aggregation.
It proved theoretically that aggregation schedule delay generated by the new algorithm is at most $16R + \Delta - 14$ time slots. Here, $R$ is network radius and $\Delta$ is maximum node degree in the original network communication graph. This algorithm improved earlier known best data aggregation algorithm with upper delay bound $24D + 6\Delta + 16$ time slots, where $D$ is network diameter ($D$ can be as large as $2R$) to conduct simulations to study the proposed data aggregation algorithm’s practical performance. Simulation results corroborated theoretical results revealing that the algorithm performed better in practice. To ensure that overall delay lower bound for data aggregation under an interference model is $\max\{\log n, R\}$, where $n$ is network size. The lower bound of delay was derived through a protocol interference model where $r < rI < 3r$ and $rI \geq 3r$.

A common aggregation model independent of specific application environments proposed by Lin et al (2011) was based on evolutionary game theory called Evolutionary Game-based Data Aggregation Model (EGDAM) in WSNs. EGDAM is made up of formal definition, functional model and its general process was defined to map aggregation procedure competition and cooperation into games avoiding perfect rationality. The authors then applied the theoretical model. Guided by model, an evolutionary game-based adaptive weighting algorithm named EGWDA, guided by model, was provided for pixel-level data aggregation with homogeneous sensors. Sensors reasonable weight distribution is possible during WSNs aggregation. Experiments on self-constructed data and from reference ensured good performances.

Barton & Zheng (2010) proposed the problem of data aggregation at a single sink. The predominate traffic patterns in a WSN were many-to-one and one-to-many communication. Hence, the performance of WSNs is characterized by the rate at which data could be disseminated from or aggregated to a data sink. It was shown that a data aggregation rate of
A new concealed data aggregation scheme extended from homomorphic public encryption system was proposed by Lin et al (2013). It has three contributions. First, it is designed for multi-application environment where base station extracts application-specific data from aggregated cipher texts. Next, it mitigates impacts of compromising attacks in single application environments and lastly, it degrades unauthorized aggregations damage.

Data aggregation, information trust and fault tolerance to improve the correctness of collected information and trustworthiness was considered by Yan et al (2012). Based on WSNs multilayer aggregation architecture, a trust-based framework for data aggregation with fault tolerance which aimed at reducing erroneous data impact and ensuring measurable trustworthiness for aggregated results was designed. By extracting statistical characteristics from various sources and extending Josang's trust model, it is proposed to compute self-data trust opinion, peer node trust opinion and peer data trust opinion. According to trust transfer and trust combination rules designed, sink node trust opinion was from final aggregated result. Specifically, this framework evaluated discrete data and continuous media streams in WMSNs through a mechanism. Results from simulation and experiments on a real WMSN test bed proved the proposed framework’s validity and efficiency, as it greatly improved multimedia information quality and evaluated the trustworthiness of collected information.

A multidimensional privacy preserving data aggregation WSN scheme efficient and ensuring strong security was presented by Yang et al (2011). This ensured a countermeasure against active and passive privacy compromising attacks, coalition attacks from malicious base stations and
captured sensor nodes, in addition to being robust to data loss. Also, the new scheme ensured data aggregation with communication overheads, reducing transmission cost suiting it for large scale WSNs. The new scheme was the first to address privacy and efficiency issues in WSNs simultaneously.

A new graph-based model to aggregate sensors' data at WSN gateways was proposed by Habib & Marimuthu (2011), where it guided and assessed resources for data aggregation. They modelled all sensors' tasks in a graph so that collected data was smoothly aggregated (scheduled) at gateways without losing/overlapping collected data. A typical WSN has hundreds of sensors and a few gateways; also, each sensor executes three tasks periodically and sequentially like sensing, processing and transmitting. The three tasks were modelled as a Directed Acyclic Graph (DAG) per sensor and then all DAGs were grouped into a Super Task-Flow-Graph (STFG). The data aggregation problem was solved by scheduling the tasks within STFG, where the authors have utilized three scheduling algorithms:

1. As soon as possible.
2. As late as possible.

Dan & Wong (2011) proposed using a novel distributed data structure called Distributed Data Cube (DDC). A DDC meticulously maintains a set of special forms of aggregate values (prefix sum, prefix average, prefix max and prefix min) in distributed sensor nodes. First, fast algorithms are presented to build a DDC within a sharp and time bound. Then, the efficient distributed query-processing algorithms were presented to handle aggregate queries by using a DDC. For a query region with n sensor nodes, the algorithms could return within $O(\sqrt{n})$ time. Finally, extensive simulation studies confirmed that a DDC can be built very quickly, which was
consistent with the theoretical and time bound issue. The network traffic injected while constructing a DDC was acceptable and also scalable as the network size grows. Query processing on a DDC was fast and energy efficient in terms of the time units needed and the number of messages incurred.

Data-aggregation techniques are based on statistical information extraction capturing the effects of aggregation over different scales was proposed by Jiang et al (2010). An accurate estimation of sensory data parameters using expectation-maximization algorithm is also designed. The proposed techniques not only greatly reduced the communication cost but also retained valuable statistical information it has lost in many existing data-aggregation approaches for sensor networks. Moreover, simulation results showed that the proposed techniques were robust against link and node failures and performed consistently well in broad scenario with various network configurations.

A Data-Centric Storage (DCS) as a system to perform data aggregation on Wireless Sensor and Actor Networks (WSANs), where sensor and actor nodes collaborate in a fully distributed way without a central base station that manages networks or provides connectivity to the outside world was proposed by Cuevas et al (2011). The authors compared various multi-replication DCS proposals and chose the best for application which studying data aggregation. Additionally, the authors ensure mathematical models for production, consumption and overall network traffic for various application profiles based on the ability of a specific application to perform data aggregation and on what traffic type was dominant, either consumption or production. The authors also provided closed formulae for every application profile defining optimal number of replica that reduce overall network traffic. Finally, authors validated proposed models through simulation.
2.2 ENERGY AWARE ROUTING

Ad-hoc on demand distance vector routing was modified by incorporating Signal-to-Noise Ratio (SNR) based dynamic clustering by Ganesh & Amutha (2013). The new scheme, Efficient and Secure Routing Protocol (ESRP) for WSNs through SNR-based dynamic clustering mechanisms partitions nodes into clusters and selects CH from nodes based on energy and non CH nodes join with specific CH based on SNR values. Error recovery was implemented during inter-cluster routing to avoid end-to-end error recovery. Security was achieved by isolating malicious nodes using sink-based routing pattern analysis. Investigation using a global mobile simulator showed that the hybrid ESRP improved energy efficiency and packet reception rate compared to SNR unaware routing algorithms like LEACH and power efficient gathering in sensor information systems.

A new energy-aware WSN routing protocol, Reliable and Energy Efficient Protocol (REEP) which ensued that sensor nodes established reliable and energy-efficient paths for data transmission was proposed by Zabin et al (2008). REEP’s performance was evaluated under various scenarios and was seen to be superior to popular data-centric routing protocol, directed-diffusion which was the benchmark.

Shin & Sun (2011) presented Chain Routing with Even Energy Consumption (CREEC) which solved lifetime issue of WSN by achieving longer average life-times using two strategies:

i) Maximizing the fairness of energy distribution at every sensor node
ii) Running a feedback mechanism that utilizes a preliminary simulation of energy consumption to save energy for depleted sensor nodes.

Simulation results confirmed that CREEC outperformed all previous solutions such as LEACH, PEGASIS, PEDAP and PEDAP-power aware with respect to the first node death and the average lifetime. CREEC performed very well at all WSN sizes, BS distances and battery capacities with an increased converge cast delay.

Zhang et al (2014) proposed an Energy-Balanced Routing Method, based on Forward-Aware Factor (FAF-EBRM). In FAF-EBRM, the next-hop node was selected according to awareness of forward energy density and link weight. In addition, a spontaneous reconstruction mechanism for local topology was designed additionally. In the experiments, FAF-EBRM was compared with LEACH and Energy Efficient Unequal Clustering (EEUC), experimental results showed that FAF-EBRM outperforms LEACH and EEUC, which balanced the energy consumption, prolonged the function lifetime and guaranteed high QoS of WSN.

Yang et al (2013) presented AutoSP-WSN, a novel distributed framework to achieve sustainable data collection while also optimizing end-to-end network performance for SP-WSNs. Initially, the energy-aware support component presented reliable energy monitoring and prediction. This drives the power management component, which is adaptive to time-varying solar power, avoiding battery exhaustion as well as maximizing the per-node utility. Finally, to demonstrate the key design issues of the network protocol component, two self-adaptive network protocols, a routing protocol Solar Powered Backpressure Collection Protocol (SP-BCP) and a rate control scheme PEA-DLEX were proposed. The individual components seamlessly integrate as a whole and the AutoSP-WSN framework exhibited the properties
of context-awareness, distributed operation, self-configuration, self-optimization, self-protection and self-healing. Through extensive experiments on a real SP-WSN platform and hardware-driven simulations, the proposed schemes achieved substantial improvements over previous work, in terms of reliability, sustainable operation and network utility.

A new energy-efficient local metric called Efficient Advancement Metric (EAM), for sensor networks was proposed by Zhang & Zhang (2009). EAM considers maximum forwarding distance and successful transmission probability packet by considering wireless channel condition which enables a forwarding node to choose a most energy-efficient relay node in a geographic-informed routing protocol. Theoretically speaking, the existence of the unique optimal relay node maximizes EAM over a typical Nakagami-m channel of a Code-Division Multiple-Access (CDMA)-based WSN. Furthermore, based on the proposed metric EAM presented a cross-layer packet-forwarding protocol Channel-Aware Geographic-Informed Forwarding (CAGIF) by optimally selecting relay nodes. CAGIF only requires that nodes have knowledge of their own information on location and the location information of source and destination nodes. Numerical examples were presented to show the characteristics of EAM and the optimal distance. Compared with the previous geographic packet-forwarding schemes in WSNs, CAGIF consumed much lower energy and generated a significantly decreased signal overhead.

An Energy-aware Hybrid Data Aggregation Mechanism (EHDAM) designed to adaptively control data transmission based on burst length and time-out based transmissions was proposed by Kim et al (2011). EHDAM adjusted burst length threshold value in reciprocal proportion to a node’s remaining energy state, reducing data transmissions; So, EHDAM extended node life when affected by energy hole issues.
A new online routing scheme, Energy-efficient Beaconless Geographic Routing (EBGR), providing loop-free, fully stateless, energy-efficient sensor-to-sink routing with reduced communication overhead without prior neighbourhood knowledge was proposed by Haibo & Hong (2010). Hop count established lower and upper bounds and upper bound energy consumption under EBGR for sensor-to-sink routing, if there were no packet loss and no failures in greedy forwarding. Also, anticipated total energy consumption on a sink route under EBGR approaches to lower bound with increased node deployment density and extended EBGR to lossy sensor networks to ensure energy-efficient routing in unreliable communication links. Simulation showed that the new scheme outperformed current WSN protocols with dynamic network topologies.

Tyndall Heterogeneous Automated Wireless Sensors (THAWS), embedding a reactive routing protocol proposed by Shen et al (2009) was an energy-efficient proactive heuristic protocol and distributed geographical routing mechanism on Tyndall 10 and 25 mm nodes. Mature WSN automation tools needs implementation of various protocols for different topology and application scenarios to optimise overall system performance. Focusing on transport layer energy-aware routing function for Tyndall Heterogeneous protocols, it managed small, medium and large sensor applications adaptively, while THAWS tool end users were transparent from protocol selection.

Mottola & Picco (2011) presented MUSTER, a routing protocol expressly designed for many-to-many communication. To increase network lifetime, MUSTER minimized the number of nodes involved in many-to-many routing and balances their forwarding load. MUSTER was evaluated in emulation and in a real WSN test bed. Results indicated that the proposed protocol built near-optimal routing paths, doubled the WSN lifetime and
overall deliver to the user 2.5 times the amount of raw data with respect to mainstream protocols. Besides, MUSTER is intrinsically amenable to in-network aggregation, making improvements up to a 180 percent increase in lifetime and a four-time increase in data yield.

Huang et al (2011) presented an Energy-aware Interference-sensitive Geographic Routing (EIGR) protocol, which focused on reducing interference and minimising the total network energy consumption. EIGR adaptively used an anchor list to guide data delivery and selected the minimum-interference link from energy-optimal relay region for data delivery. To reduce the energy consumption and interference further, EIGR adjusted the transmission power of each forwarding node so as to just reach the selected next forwarding node. Simulation results demonstrated that the proposed approach exhibited noticeably higher energy efficiency, shorter end-to-end delay and higher packet delivery ratio compared to other geographic routing protocols.

2.3 CLUSTER BASED ROUTING

Jain & Trivedi (2012) proposed an algorithm for energy efficient clustering and multi-hop routing in WSN. The sensor network is constructed in the form of a circular area with the base station. An adaptive cluster selection strategy which selects the Cluster Heads (CHs) not only on the basis of residual energy, but also on the distance from the base station. A key point of the algorithm is that CH selection is done after each round and multiple times within the same round. This is done since the CH near the BS is involved in most of the communications and hence it may get exhausted early.

A type-2 fuzzy based Self-Configurable Cluster Head (SCCH) selection approach to consider CH selection criterion was proposed by Izadi
et al (2013) which presented a cluster backup approach. So, when cluster failed, the system worked efficiently. This protocol’s novelty is the ability to handle communication uncertainty, an inherent operational aspect in sensor networks. Results revealed that SCCH performed better than other developed methods.

A weighted average of CH selection algorithm through learning clustering algorithm’s framework for WSNs based on Back Propagation (BP) Neural Network (NN) was presented by (Guo et al 2010) which directly related node weights to decision-making predictions. The nodes weight distribution is the objective. Simulation revealed the algorithm’s efficiency in eliminating data redundancy, reducing network traffic and extending network life.

A new CH selection framework and sub-carrier allocation to intra-cluster communication in an Orthogonal Frequency Division Multiple Access (OFDMA) based WSN was presented by (Rituraj & Jagannatham (2013) which maximized WSNs overall performance regarding throughput, video quality and network life. The OFDMA Physical Layer (PHY) optimization criterion for rate maximization is thought of as a bi-level CH node selection optimization problem followed by optimal OFDMA sub-carrier allocation. It was demonstrated that optimal resource allocation problem could be reduced to a solution of a weighted bipartite graph matching problem and solved using three different schemes including Hungarian algorithm, game theory based multi-item auction and Gale-Shapley stable matching scheme. Based on the above formulation, also derived an optimal scheme to minimize the energy consumption in the WSN. Further, the performances of the proposed schemes were demonstrated through simulations in the context of scalable video transmission in a video sensor network. Results illustrated that these schemes
significantly outperformed suboptimal WSN cluster-head selection schemes in terms of data rate and video quality.

Liu et al. (2011) proposed a dynamic clustering protocol, focusing on CH selection, cluster member solicitation and cluster reorganization in order to solve the problem of dynamic clustering for WSN triggered by target event. The proposed selection approach in CH selection mechanism considers detected signal strength, a node’s residual energy and distance between cluster-head and sink node. An adaptive method based on task requirement in cluster member selection was suggested regarding cluster range setting. This method constrains cluster size and energy consumption of intra-cluster communication is reduced. Simulation showed that the new protocol conserved energy and prolonged WSN life specially when sink node was far from the network.

Zhang et al. (2010) proposed to tackle the strict limitation of energy supplication in large scale hierarchical structure of WSN. The proposed selection approach in CH selection mechanism considers detected signal strength, a node’s residual energy and distance between cluster-head and sink node. An adaptive method based on task requirement in cluster member selection was suggested regarding cluster range setting. This method constrains cluster size and energy consumption of intra-cluster communication is reduced. Simulation showed that the new protocol conserved energy and prolonged WSN life specially when sink node is far from the network. Considering the character of communication scope of node in cluster, there exists areas of redundant place between joined clusters. To give a precise number of CH, efficient coverage is used, instead of full coverage of cluster to calculate the optimal number of CH per round, according to characters of one and two order partial derivative functions.
Yoon & Chang (2011) proposed a new cluster-based routing protocol using message success rate. To resolve the node concentration problem and design a new CH selection algorithm based on node connectivity and devise cluster maintenance algorithms. Moreover, to guarantee reliability of data communication, message success rate is a popular measure for data communication reliability, is used in order to select a routing path. Finally, to reduce data communication overhead, only information of neighbouring nodes during both cluster-head selection and cluster construction phases were used. Through the performance analysis, it showed that proposed protocol outperformed the existing schemes in terms of communication reliability and energy efficiency.

Balaji & Anandamurugan (2013) proposed a scheme to present mobility aware CH selection in the hotspots of WSN. This calculates node reputation to have better CH. Ranking was based on node reputation and mobility rate which increases clustering efficiency regarding cooperativeness and sensor nodes mobility. But, some sensor nodes were seen to have high mobility. By the same token, some nodes are unwilling to serve network. A Mobility Aware Reputation Node Ranking (MARNR) technique to improve clustering efficiency at hotspot regions was suggested. MARNR identifies sensor nodes mobility rate. A node serving as cluster was identified from a cache ensuring a node has minimal mobility threshold and high cluster-head rank probability. Simulation evaluated Evolutionary Computation performance regarding network life, mobility rate, energy consumption, node reputation count, CH ranking, hotspot density and cluster size.

Wang et al (2013) proposed a Fuzzy-based system for CH selection and controlling sensor speed in WSNs. In WSN, cluster formation and CH selection were critical issues. They could drastically affect the network's performance in different environments with different characteristics. The
The proposed system was constructed by two Fuzzy Logic Controllers (FLC). Using four input linguistic parameters for evaluating CH decision probability in FLC1, the output of FLC1 was used and two other linguistic parameters as input parameters of FLC2 to control sensor speed. FLC2 was evaluated by simulations and showed that it achieved good performance.

Zhang et al (2012) proposed the position of CH as an important factor for the network lifetime. Based on this observation, a non-random CH selection scheme based on the concept of the center of mass in physics was proposed. The problem of power consumption in sensor data collection in a WSN was considered. Since sensor nodes operate on batteries, power efficiency is a crucial issue in designing the network. The geographic deployment of sensor nodes is random, with an irregular network topology. In the existing clustering-based protocols for the WSN, the CHs are usually selected at random, which may result in higher power consumption and shorter network lifetime. The purpose of proposed scheme was to use minimal power in the process of data collection. It was shown that proposed scheme could save up to 50% of power consumption.

Hao et al (2009) proposed to draw attention to the randomly deployed WSNs based on the independent homogeneous Poisson process and proposed an Improved Cluster-head Selection (ICS) approach. The selection was subject to two factors:

(i) The number of CHs may differ in different periods of the WSN lifetime.

(ii) This number can determine performance, in terms of WSN lifetime and energy consumption, etc.
First to optimize the dynamic number of CHs. Then the suitable CHs were obtained via improved selection procedures based on three selection phases by introducing the temporary clusters. Meanwhile, Common Energy Consumption Cycle (CECC) -based synchronization was used to coordinate the intra and inter-cluster communication. ICS saved the energy consumption and prolonged the network lifetime. Simulation demonstrated the enhanced network performance.

A new CH selection method for LEACH clustering routing protocol suggested by Tripathi et al (2012) balances network sensor nodes energy consumption. Simulation results showed that sensor network life significantly improved when compared to LEACH routing algorithm for WSNs.

Nazir & Hasbullah (2010) proposed a clustering in terms of energy efficiency in WSN and proposed Energy Balanced Clustering (EBC) in WSN. Algorithms for energy balanced cluster formation, intra cluster and inter cluster communication and CH selection in WSN were proposed. Furthermore, using OMNet-4.0 simulation performance of proposed protocol was compared with LEACH and EEMC using parameters like energy per packet and throughput. Simulation results demonstrated that Energy Balanced Clustering (EBC) was effective in prolonging the network lifetime, than LEACH and EEMC.

Li et al (2011) based on the idea of CH expected frequency appraisal, a LEACH-Head Expected Frequency Appraisal (LEACH-HEFA) algorithm was proposed for improving the performances. Since the CH selection mechanism directly affects the energy consumption and network lifetime, both the cluster selection mechanism and the energy saving strategy play an important role in the design of network routing protocols. Simulation results clearly showed that LEACH-HEFA algorithm could balance energy
consumption of nodes in the WSN, rationalize the clustering process and prolong the network lifetime effectively. It indicated that the proposed algorithm was suitable to water regime monitoring systems.

Hu et al (2011) proposed an energy saving method LEACH-Improved Energy Efficient (LEACH-IMP) based on LEACH. Cluster-head selection algorithm and the inequality of node power consumption were unreasonable disadvantages of LEACH. In the proposed system all nodes were divided into fixed clusters. According to the distance from the CH members of cluster nodes adaptively join cluster. MATrix Laboratory (MATLAB) simulation proved that the improved routing protocol was effective in overcoming non-uniform energy consumption in nodes which was due to random CH selection strategy in LEACH, thereby extending the lifetime of the network.

A LEACH algorithm based energy effective routing protocol to meet key QoS requirements was proposed by Peng et al (2011). Energy efficiency is an important WSN issue. Network layer routing technology is critical to reduce WSN energy consumption. But, reliability and data aggression must also be analyzed. The new protocol focused on traditional LEACH defects and improved energy efficiency and QoS parameters by excluding nodes with improper geographic location to be CHs. Optimum measuring range of head nodes is designed to be a CH selection criterion and every CHs can be selected according to node density threshold in measuring area, confirmed by node distribution situation and communication needs. Simulations evaluated the new protocol in comparison with traditional LEACH algorithm. The performance of new protocol was verified to reduce energy consumption and guarantee communication quality especially in uneven distribution situations.
Nie et al (2010) proposed the idea of common nodes to be prior placed, CHs with more energy and communication capability were deployed according to the special relation of common nodes and designed a CHs deployment based on Genetic Algorithm and Weighed Clustering Algorithm (GA-WCA), in which the position of CHs would be decided by GA and nodes selection within a cluster, by WCA. The simulation result verified the availability of this algorithm, when under the constraint condition of ideal load of CHs and balanced load between CHs, the optimal deployment of CHs could be realized.

Kuila & Jana (2012) presented a distributed clustering and routing algorithm for WSN called Cost-based Energy Balanced Clustering and Routing Algorithm (CEBCRA). The algorithm comprised of three phases, namely CH selection, cluster setup and data routing. The CHs were selected in a distributed manner based on residual energy and the neighbour cardinality. In the setup phase, each non-CH sensor node joined a CH within its communication range based on the cost value of the CHs. In data routing phase, CEBCRA first used single-hop communication within each cluster and then performed multi-hop communication between the clusters. For inter-cluster routing, a CH measured the cost of each path from itself towards base station while selecting other CH as a relay node for data forwarding on those paths. The experimental results showed efficiency of the proposed algorithm in terms of energy consumption and number of live sensor nodes. The results were compared with two existing techniques to show the efficacy of the algorithm.

Wang et al (2009) proposed a refined protocol named LEACH-H (hybrid cluster-head selection LEACH) in order to prolong the WSN's lifetime. In the first round of LEACH-H, the base-station selected a CH set through adopting Simulated Annealing (SA) algorithm; in the rounds, the
CHs subsequently would select new CHs in their own cluster. This will not only solve the problem that the CHs were unevenly distributed in LEACH, but also maintain the characteristics of distribution. The energy consumption of the network is cut down and the live time of WSN is extended in LEACH-H.

Enam et al (2012) proposed the consumption of energy during the transmission of data from sensor nodes to the sink. Simulations have shown that 20% of the network's total consumed energy is exhausted in the random CH selection phase itself. This energy consumption could be taken as an overhead energy. To overcome the overhead energy, an energy efficient “Round Rotation” (RR) mechanism has been proposed. By implementing the RR mechanism with random CH selection phase, the overhead energy has been reduced to 4% of the total network energy consumed in data transmissions. This has consequently affected the network lifetime which has been augmented by 20%.

2.4 CLUSTER HEAD SELECTION FOR WIRELESS SENSOR NETWORKS

An efficient CH selection method using K-means algorithm to maximize WSN energy efficiency was proposed by Park et al (2013) based on finding the CH and lowering the sum of Euclidean distances between head and member nodes. Computer simulation revealed that the new approach ensured improved performance than current hierarchical routing protocols like LEACH and HEED as regards network life.

An improved LEACH protocol version which aimed to reduce energy consumption in WSN and prolong network life was proposed by Ray & De (2012). Simulation showed that the new protocol reduced energy consumption and prolonged WSN life with respect to parameters First Node
Dies (FND), Half Node Dies (HND) and Last Node Dies (LND). The new protocol was 41.7% better than LEACH regarding FND, 31.6 % better than LEACH regarding HND and 25.5% better than LEACH regarding LND.

A distributed CH selection algorithm that considers distance from sensors to base station that optimally balances energy consumption among sensors were proposed by Kang & Nguyen (2012). Simulations revealed that the new scheme outperformed current algorithms regarding average node life and time for first node death.

Junping et al (2008) presented a Time-based CH selection algorithm for LEACH (TB-LEACH). A comparison between the proposed protocol and LEACH protocol was provided. The implementation of this protocol was figured out by simulation. Simulation results showed that proposed algorithm outperformed original LEACH by about 20% to 30% in terms of system lifetime.

Ando et al (2011) proposed a new fuzzy-based CH selection system to improve the performance of the network. The proposed system is inherently different from the previous system used by the four input linguistic parameter for CH decision. The proposed system was evaluated by simulations and showed that the proposed system has a good CH selection.

A sleep-wake up based decentralized MAC protocol was applied to LEACH to reduce energy consumption and avoid strict TDMA synchronization requirements by Yang & Sikdar (2007) who presented an analytic framework to ensure optimal probability how a node becomes a CH to minimize network energy consumption. The analysis was first for small networks, assuming that identical expected distance of cluster heads from sink. It was then extended to large networks to consider cases when distances
of various network sections from the sink were different, as nodes farther away require more energy to reach the sink. Simulation showed that using this optimal probability lead to more efficient energy consumption. Compared to current LEACH, the proposal consumed less power.

Li et al (2012) developed and analysed a protocol which forms clusters like LEACH wherein the first round or after one invalid node or more appear. During other rounds previous CHs can transfer the role of CH to the most energy node in corresponding clusters, instead of random reselection of CHs. The information of the residual energy of each node was reported to its CH in data package at the end of the rounds before invalid node appears. Above all, number of CHs could change with number of nodes alive. Simulation result showed that the protocol can maximize the time of the network without nodes using up energy, called ‘no deaths period’.

Li et al (2011) proposed a new improved method called LEACH-N based on LEACH. According to the new protocol, the problem of how to select nodes as the cluster head node depends on the residual energy of nodes in the cluster. This strategy could guarantee the rationality during selecting head nodes. Moreover, the network robustness can be enhanced and the life cycle for the network can be prolonged. The simulation results showed that the algorithm proposed better performance than LEACH.

Han (2010) proposed LEACH-High Performance Routing (LEACH-HPR) an energy efficient cluster head election protocol and using the minimum spanning tree algorithm to construct an inter-cluster routing. Simulation results showed the method is more efficient in reducing and balancing energy consumption and hence prolongs the lifetime of WSN.

Zhao et al (2012) proposed a new method of choosing CHs based on the original LEACH protocol which decreased unnecessary consumption
of energy spent on computing of each node during each round. Based upon the traditional method of choosing formula, neglecting change of nodes' energy will make the nodes acting as CHs die early owing to consuming too much energy. In order to make the energy distribution more even in the network, the consideration of the dynamic change of sensor nodes' energy is introduced during the selection of CHs. Simulations showed that proposed improved protocol performed better than the original LEACH.

A revised cluster routing algorithm named E-LEACH to enhance LEACH hierarchical routing protocol was proposed by Xu et al (2012). The original method to select cluster-heads in LEACH algorithm is random and round time for selection is fixed. In the E-LEACH algorithm, the remnant power of the sensor nodes is used in order to balance network loads and changed the round time depending on the optimal cluster size. The simulation results showed that the proposed protocol increased network lifetime at least by 40% when compared with the LEACH algorithm.

A Multi-hop Routing with LEACH (MR-LEACH) protocol was proposed by Farooq et al (2010) to prolong WSN life. MR-LEACH partitions networks into different cluster layers. Thus, MR-LEACH followed Multi-hop routing from cluster-heads to a base station to conserve energy, unlike the LEACH protocol. Performance evaluation has shown that MR-LEACH achieved significant improvement in the LEACH protocol and provided energy efficient routing for WSN.

An improved LEACH algorithm called partition-based LEACH (pLEACH), which partitions a network into optimal sectors and then selects a node with highest energy as head for each sector, using centralized calculations was proposed by Gou & Yoo (2010). Simulation and analysis showed that pLEACH achieved better performance in WSN regarding energy dissipation, network life and quality of communication.
2.5 MODIFIED LEACH ROUTING IN WSN

Multi-hop transmissions and Single-hop transmissions LEACH (MS-LEACH) to enhance Secure-LEACH (S-LEACH) security by providing data confidentiality were proposed by El-Saadawy & Shaaban (2012). They also used node to CH authentication using pairwise keys shared between CHs and cluster members. A secure clustering protocol that achieves security goals while ensuring acceptable energy consumption levels is a WSN challenge. LEACH is a basic clustering-based WSN routing protocol. S-LEACH is modified LEACH version with cryptographic protection against outsider attacks. The proposed MS-LEACH’s security analysis revealed that it had efficient security properties, achieving all WSN security goals compared to present secured LEACH protocol solutions. MS-LEACH performance based evaluation proved its effectiveness and that it achieved desired security goals outperforming other protocols regarding energy consumption, network life, network throughput and normal routing load.

Ahmed et al (2013) proposed a basic LEACH protocol modified with the LEACH-Mobile Average Energy (LEACH-MAE) based protocol to overcome its shortcomings to support mobility along with the new average energy based CH selection technique. The simulation showed that proposed algorithm improved network life time up to 25% as well as helped to maintain the equal distribution of energy resource among the sensor nodes.

Mahmood et al (2013) modified one of the most prominent WSNs routing protocol LEACH as modified LEACH (MODLEACH) by introducing efficient CH replacement scheme and dual transmitting power levels. The modified LEACH, in comparison with LEACH outperformed it using metrics of CH formation, throughput and network life. Afterwards, hard and soft thresholds were implemented on MODLEACH that boosted the performance even more. Finally a brief performance analysis of LEACH, MODLEACH,
MODLEACH with Hard Threshold (MODLEACHHT) and MODLEACH with Soft Threshold (MODLEACHST) was undertaken considering metrics of throughput, network life and CH replacements.

Distance Aware Intelligent Clustering protocol (DAIC), a new hierarchical routing protocol was proposed by Gautam et al (2010) to minimize energy consumption and increase network life with the concept of dividing networks into tiers and selecting high energy CHs at nearest distance from BS. It was seen that much energy could be conserved by selecting CHs at nearest distance from BS. Also, CHs number was computed dynamically to avoid choosing unnecessarily many CHs in a network. Simulation showed that the new DAIC outperformed LEACH and LEACH-C (Centralized LEACH) by 63.28% and 36.27% as regards energy conservation respectively. Distance aware CH selection method adopted by the new DAIC protocol could be adapted to other hierarchical clustering protocols for improved energy efficiency.

A Location-aware and Fault tolerant Clustering Protocol for Mobile WSN (LFCP-MWSN) that is energy efficient and reliable was proposed by Karim & Nasser (2012). LFCP-MWSN incorporated a range free approach to localise sensor nodes during cluster formation whenever a sensor moved into another cluster. Existing cluster-based mobile routing protocols like LEACH-Mobile, LEACH-Mobile-Enhanced and Constant Bit Rate (CBR)-Mobile consider only sensor nodes energy efficiency. Most existing mobile routing protocols are not designed as fault tolerant. These protocols allocate extra timeslots using Time Division Multiple Access (TDMA) scheme to accommodate nodes that enter a cluster because of mobility and thus, increases end-to-end delay. Moreover, existing mobile routing protocols were not aware of location and assumed that sensor nodes verily knew their coordinates. Simulation results showed that LFCP-MWSN protocol has about 25-30% less network energy consumptions and slightly less end-to-end data
transmission delay than the existing LEACH-Mobile and LEACH-Mobile-Enhanced protocols.

A new vehicular clustering scheme integrating hierarchical clustering based on classical routing algorithm was proposed by Liu et al (2010). Clustering algorithm significantly impacted WSN operations. Effective clustering algorithm ensures efficient WSN operation. Hierarchical clustering is a new WSN clustering scheme. Simulation revealed that the new scheme mitigated WSN hotspot problem efficiently achieving improved network life and load balance compared to the old algorithms like Direct, LEACH and DCHS.

A framework enabling practical development of centralized cluster-based protocols supported by optimization methods for WSNs was proposed by Hoang et al (2014). Based on this, a protocol using a music-based meta-heuristic optimization method, Harmony Search Algorithm (HSA) is designed and implemented in real time for WSNs. It anticipated reducing intra-cluster distances between cluster members and CHs and optimized WSN’s energy distribution. A study of HSA cluster-based protocol was carried out where WSNs equipped with the new protocol were deployed indoors to monitor fire detection ambient temperature. Well-known cluster-based protocols developed for WSNs like LEACH-C and Fuzzy C-Means (FCM) clustering algorithm were compared. Experiments proved that the new protocol with HSA could be realized in centralized cluster-based WSNs for building environments safety and surveillance applications. Test results proved that it extended WSN life compared to LEACH-C and FCM protocols.

Gao & Li (2011) proposed the Link Estimation and Parent Selection (LEPS) multi-hop routing protocol and the tree topology with a sink node as the root node by using nesC language in the TinyOS operating system. Adopting TOSSIM simulation tools a graphical debugging program
Tinyviz was provided for the simulation of this system. Finally, the routing module was replaced by the modified LEACH protocol. Experimental tests showed that the automatic routing function could be realized and the cluster was also determined by the election based on the remaining energy and spatial distribution of the region.

### 2.6 EVOLUTIONARY ALGORITHMS BASED ROUTING

Sector Based Clustered Routing (SBCR) protocol for distributed Data-Centric Storage (DCS) in WSNs was proposed by Ahmed & Gregory (2012) which identified an optimally synchronized routing algorithm that achieved high power efficiency and simultaneously reducing end-to-end delay, updates and query traffic and collisions. SBCR was cross pollinated with MAC layer (Sensor-MAC and/or Telecommunication-MAC) to achieve this objective, building a cross-layer optimization model. Based on network-wide cluster flow distribution, a sector member was assigned a non-overlapping TDMA slot by sector head for member-head collision free communication while a stateless location based multi-hop routing scheme handled inter-sector communications. A description of SBCR implementation was ensured in open-source Castalia simulator. Simulation proved that SBCR yielded improved efficiency for DCS and data queries when compared to current approaches like LEACH and Greedy Perimeter Stateless Routing (GPSR).

Mao et al (2011) proposed, for large scale WSNs, an energy efficient unequal clustering scheme which aims at balancing the node power consumption and prolonging the network lifetime as long as possible. The approach focuses on energy efficient clustering scheme and inter-cluster routing protocol. On one hand, considering each node's local information such as energy level, distance to base station and local density, to use fuzzy logic system to determine one node's chance of becoming CH and estimate the cluster-head radius. On the other hand, Ant Colony Optimization (ACO) was
applied for inter-cluster routing to construct energy-aware routing between CHs and base station. The proposed method solves hot spot problem. The confirmation of experiment results have indicated that proposed clustering scheme has better performance than the other methods such as LEACH and EEUC.

Song et al (2010) proposed a novel hierarchical routing protocol based on Artificial Fish Swarm Optimization (AFSO). AFSO algorithm in cluster formation phase’s main object was to solve the NP-hard problem of finding k optimal clusters according to the given rules. These networks require robust wireless communication protocols for the purpose of balancing load and prolonging the network lifetime. The performance of the novel protocol was compared to the well-known cluster-based protocol LEACH and LEACH-C. As the experiment results show, the protocol not only improves system lifetime but also prevents the networks from serious energy consumption.

Bakr & Lilien (2011) proposed the LEACH-Spare Management (LEACH-SM) protocol - a modification of LEACH. It adds the spare selection phase to LEACH. After a WSN is deployed, its operational lifetime depends on its energy resources. Available results reveal that significant improvement in WSN lifetime could be achieved by making WSNs redundant, i.e., by adding to WSNs spare nodes that are initially asleep but are ready to be switched on when any primary node (i.e., a node that is not a spare) uses up its energy.

Cui & Liu (2009) proposed a Balanced-Clustering, Energy-Efficient Hierarchical Routing Protocol (BCEE), which operates in two phases. During Phase 1, balanced cluster formation was possible through application of k-means clustering strategy. This does not require no exact position of every node but uses Received Signal Strength Indicator (RSSI)
instead. Then, ACO algorithm in Phase 2 established the route with optimal or sub-optimal power consumption between cluster heads and sink node. Simulation showed that the new BCEE protocol ensured better load-balance, significant energy consumption reduction and prolonged network life better than LEACH.

Fareed et al (2012) compared six varied protocols of different scenario that presented energy minimizing, clustering and route selection schemes to ensure effective communication. A good protocol design should scale well in energy heterogeneous and homogeneous environments meet demands of various application scenarios and guarantee reliability. The research was motivated to get an insight in which protocol suited which application and could be a guide-line to design a robust and efficient protocol. MATLAB simulations analyzed and compared performances of LEACH, multi-level hierarchal LEACH and Multi-hop LEACH (M-LEACH).

Badi et al (2010) used simulation modelling to investigate the dependency of the CH percentage on the network node density. LEACH designers observed that there were an ideal percentage of nodes that needed to be CHs to achieve optimal energy performance. For the work and performance analysis they selected 5% as the number of CHs in the network. Most WSN routing protocols followed LEACH and use clustering for energy performance optimization. Most protocols had taken 5% of CH as ideal working setting, without qualifying this assumption, independently. Results reveal that the percentage is not universal for all network settings and is dependent on density. These pave the way for the challenge of WSNs hierarchical networking protocol design.

Nivetha (2012) surveyed different WSN clustering algorithms. A goal in WSN design is its lifetime maximization, constrained by batteries energy capacity. Routing protocols discover and maintain network routes.
Many routing, power management and data dissemination protocols are specifically designed for WSNs where energy awareness is a design issue. WSN routing protocols might differ depending on application and network architecture. Overall, routing techniques are trifurcated based on underlying network structure: flat, hierarchical and location-based routing. Clustering/hierarchical routing is introduced in WSNs due to its network scalability, energy-saving attributes and network topology stabilities. Cluster routing, sensors are divided into groups called clusters in cluster routing, with a cluster electing a node as cluster head, so that sensors communicate information to cluster heads after which cluster heads communicate aggregated information to processing centers to save energy. Thus, cluster creation and assigning special tasks to cluster heads greatly contributes to overall system scalability, life and energy efficiency.

Energy constrained WSN deployed over a region were considered by Sarangi & Thankchan (2012). The task of such networks is to gathering information from nodes and transmitting it to base stations for more processing. Usually a fixed amount of energy is required to receive one bit of information and additional energy to transmit it. The additional amount solely depends on transmission range. A Particle Swarm Optimization based Routing protocol (PSOR) where energy efficiency is a major criterion to perform routing and derive optimized path to forward and process data to base node. PSOR generated a new path of routing by taking energy as fitness value to judge different paths and choose best optimized path whose energy consumption is lesser than that of other routing paths. The experiments performed were done using simulator. The result obtained by performing experiment on the proposed algorithm PSOR and Genetic Algorithm showed that PSOR have better results when compared to Genetic Algorithm.
2.7 CONCLUSION

In this chapter, literature survey related to data aggregation in WSNs, cluster based routing, cluster-head selection for WSNs energy aware routing and evolutionary algorithms based routing was presented. General methods for data aggregation and clustering in WSN were reviewed. Issues on energy conservation were addressed in energy aware routing. From the reviews, it is concluded that Hierarchical structure is a useful mechanism used to prolong lifetime of WSN.

Cluster creation and assigning of cluster-heads greatly contribute to overall system scalability, life and energy efficiency. It is observed that the LEACH routing is widely used for WSN routing. Many modifications to improve the performance of LEACH are proposed in the Literature. Evolutionary algorithms are used to optimize the routing. In this work, it is proposed to investigate popular EA techniques and propose improvements over the existing techniques.