

Packet Arrival Analysis in Wireless Sensor Networks

Krishna Doddapaneni*, Purav Shah*, Enver Ever*, Ali Tasiran*, Fredrick A Omondi*,
Leonardo Mostarda* and Orhan Gemikonakli *

* School of Science and Technology, Middlesex University, London, UK
Email: K.doddapaneni@mdx.ac.uk

t Computer Engineering Program, Middle East Technical University
Northern Cyprus Campus, Turkey

^School of Science and Technology, University of Camerino, Camerino, Italy

Abstract—Distributed sensor networks have been discussed for more than 30 years, but the vision of Wireless Sensor Networks (WSNs) has been brought into reality only by the rapid advancements in the areas of sensor design, information technologies, and wireless networks that have paved the way for the proliferation of WSNs. The unique characteristics of sensor networks introduce new challenges, amongst which prolonging the sensor lifetime is the most important. WSNs have seen a tremendous growth in various application areas including health care, environmental monitoring, security, and military purposes despite prominent performance and availability challenges. Clustering plays an important role in enhancement of the life span and scalability of the network, in such applications. Although researchers continue to address these grand challenges, the type of distributions for arrivals at the cluster head and intermediary routing nodes is still an interesting area of investigation. Modelling the behaviour of the networks becomes essential for estimating the performance metrics and further lead to decisions for improving the network performance, hence highlighting the importance of identifying the type of inter-arrival distributions at the cluster head. In this paper, we present extensive discussions on the assumptions of exponential distributions in WSNs, and present numerical results based on Q-Q plots for estimating the arrival distributions. The work is further extended to understand the impact of end-to-end delay and its effect on inter-arrival time distributions, based on the type of medium access control used in WSNs. Future work is also presented on the grounds that such comparisons based on simple eye checks are insufficient. Since in many cases such plots may lead to incorrect conclusions, demanding the necessity for validating the types of distributions. Statistical analysis is necessary to estimate and validate the empirical distributions of the arrivals in WSNs.

I. INTRODUCTION

WSNs have attracted a wide range of disciplines where close interactions with the physical world are essential. These unique characteristics and intrinsic properties of individual sensor nodes and WSNs separate them from other communication networks, and also present unique challenges for the development of communication protocols in terms of energy consumption as the stringent energy reserves of the sensor nodes make the energy consumption of primary importance. The past several years have seen an interesting increase in the development of WSNs. WSNs, with a wide range of applications are rapidly becoming an integral part of our lives. The application of sensor networks are diverse, ranging from habitat monitoring to surveillance and physical intrusion detection and can be categorised into environment, health, military, home, disaster relief, space exploration and other commercial areas. The flexibility, fault tolerance, low cost,

rapid deployment characteristics and high sensing fidelity of sensor networks create many new and exciting applications in the field of remote sensing. Recent work [1] illustrated tools and methodologies for the modelling, simulation and script generations for simulation tools for various WSN applications and performance evaluation by employing physical environment as well. Performance modelling and evaluation should consider new metrics for WSNs, such as system lifetime and energy efficiency, and the introduction of new traffic attributes. Depending on the area of application, information monitoring and reporting may further be classified as continuous, periodic, or event-based (driven) [2], [3]. In all these cases, data arrival delay is clearly determined by the nature of application and the chosen monitoring scheme. Quality of Service (QoS) provision in relation to the end to end delay of transmitted packets remains a serious concern along with the commonly accepted challenges such as energy consumption, network connectivity, data aggregation, computation power [4]. Characterization of the end-to-end delay distribution is fundamental for real-time communication applications with probabilistic QoS guarantees. In [3], cross layer analysis of the end to end delay distribution in WSNs was studied and the results show that inter-arrival time mostly follow exponential distribution except for low periodic traffic. There are many studies which consider exponential arrivals to sensor nodes [5]-[9]. However, in other quarters there have been mixed opinions on the appropriate distribution for modelling inter arrival delay of WSN data packets [3], [10], [11]. In other works, there has been mixed opinions on the appropriate distribution for modelling interarrival time of WSN data packets [10]. This strongly indicates the need for a study to identify acceptable types of distributions for inter-arrival times used in modelling WSNs. In this paper, an investigation is carried out to establish the most appropriate distribution for the inter-arrival times at Cluster Heads (CH) and relay nodes. Simulation results are presented and analysed in detail to characterize end to end delay between arriving data packets. Regardless of the medium access scheme employed, energy efficiency is of utmost importance in WSNs. A MAC protocol must certainly support the operation of power saving modes for the sensor node. The main motivation must be to minimize the medium access delay that may occur due to high traffic rate. In this paper, the average end-to-end delay for various application rates is also presented, whilst various MAC protocols are considered to save energy.

The paper is organised as follows: Section II provides detailed literature survey on performance modelling and existing simulation approaches; Section III briefly discusses the system

considered in this study; Section IV presents the numerical results along with the Q-Q plots for finding theoretical distributions of arrivals at the CH and also the average end-to-end delay incurred for various application rates and MAC protocols employed; Section V summarises the work presented and presents future work on the necessity of validating the types of distributions with the help of other verification approaches.

II. LITERATURE SURVEY

Performance modelling and analysis continues to be of great importance in supporting research as well as in the design, development and optimization of WSN and their applications. The current trend towards the use of WSNs for sensing and control now has the potential for significant advances, not only in science and engineering, but also, on a broad range of applications. This brings the need for performance modelling for the optimization of deployment of WSNs. However, the special design, characteristics of sensors and their applications separate them from the traditional networks. These characteristics pose great challenges for the architecture, protocol design, performance modelling and their implementation. It is essential to consider energy efficiency of WSNs because of their limited energy sources (most of the times batteries). In order to minimise the energy consumption, one of the effective techniques is to place sensors in sleep mode during the idle period [12]. In [13]–[15], a wake-up scheduling scheme at the MAC layer is proposed, which wakes up the sleeping nodes when there is a need to transmit or receive, thus avoiding a degradation in network connectivity or quality of service provisioning. The impact of routing and multi-hop transmission on both energy consumption and latency is presented in [16].

Apart from the common challenges of WSNs including energy consumption, network connectivity, data aggregation, computation power, limited sensor node memory, path-loss [17] the end to end delay of transmitted packets remains a serious concern in relation to Quality of Service (QoS) provision. Characterising delay in distributed systems has been considered in various contexts. Recent work has evaluated the latency performance of WSNs in terms of mean and variance [18]–[20]. However, it can be observed that accurately characterizing end-to-end delay at the CH is still an open problem. Considerable amount of research on sensor networks reported recently has been ranging from network capacity and signal processing techniques, to topology management, algorithms for traffic routing and channel access control. The model presented in [10] is used to investigate system performance in terms of energy consumption, network capacity, delay in data delivery along with the trade-off's that exist between performance metrics and sensor dynamics in active/sleep modes. A Markov model is presented for WSNs, where the nodes may enter into sleep mode. Through standard Markovian techniques, a system model representing the behaviour of a single sensor has been constructed along with the dynamics of the entire network, and the channel contention among interfering sensors. The proposed solution of the system model is then obtained by means of a Fixed Point Approximation (FPA) procedure, and the model has been validated via simulation.

Due to hardware constraints for energy efficiency, optimizing node packet buffer and maximizing the performance is necessary to improve the Quality of Service(QoS) for transmission in WSNs. In [21], a packet buffer evaluation method using queuing network models is proposed, where the blocking probabilities and system performance indicators of each node are calculated using an approximate iterative algorithm. The model considered focuses on a single server model in WSNs and the method used to calculate packet buffer capacity for nodes also indicate that the sink node requires higher performance, when compared to the other nodes in the network. The Markov model of the sensor sleep/active

dynamics is presented in [22], that predicts the sensor energy consumption by acquiring this information for each sensor, while a central controller constructs the network energy map representing the energy reserves available in various parts of the system. Only a single node is represented by a Markov chain, while the network energy status is derived with the help of simulation studies.

With regard to analytical studies, results on the capacity of large stationary ad-hoc networks are presented in [23]. Two network scenarios were considered: one including arbitrarily located nodes and traffic patterns, while the other one with randomly located nodes and traffic patterns. An analytical approach on network coverage and connectivity of sensor grids is presented in [24]. The sensors are considered unreliable and fail with a certain probability leading to random grid networks. Results on coverage and connectivity are derived as functions of key parameters such as the number of nodes and their transmission radius.

Several approaches based on simulations and experiments, have been proposed for performance evaluation of IEEE 802.15.4 networks [25]. In [26], an analytical framework based on a Markov chain characterization of the MAC protocol is proposed for IEEE 802.11 networks in saturation conditions. Based on this pioneering work, several approaches have been proposed for the characterization of the MAC performance in IEEE 802.15.4 networks with a star topology. In this work, a scenario with acknowledgement (ACK) messages is considered and an evaluation of the network performance in both saturation and non-saturation regimes is presented, while trying to characterize the conditions under which the network enters the saturation region [27]. A simple Markov chain theoretical model to characterize the sensors as well as the channel status is proposed in [28]. The model shows good agreement with ns-2 based simulations. This model allows to investigate throughput and energy consumption metrics within WSNs. In [29], an extended framework of the one proposed by [28] is presented for a 2-hop network scenario, i.e., networks where sensors communicate with the coordinator through an intermediate relay node, which forwards data packets from the sources (the sensors) towards the destination (the coordinator). Similar works have been presented in [30], [31], emphasising the use of a relay for interconnecting two different clusters in IEEE 802.15.4 networks and analysing the performance through a queuing theoretical analysis. However, the proposed scenario models the (simpler) cases where the relay does not contain the medium access to the sensors. Hence, it is observed that accurately characterizing arrivals at the cluster head in WSNs is still an open problem. Although it is quite difficult to analyse each possible application in WSNs, it is sufficient to analyse each class of applications classified by data delivery models, as most of these applications in each class have common requirements on the network [32].

Table I: Application Requirements of Data-Delivery Models

Factor	Event-Driven	Query-Driven	Continuous	Hybrid
Interactivity	✓	/	X	✓
End-to-End Performance	X	X	X	X
Delay Tolerance	X	Query-specific	/	X
Criticality	/	/	/	/

A well established simulation tool Castalia which provides realistic node behaviour, wireless channel and radio models, and enables to mimic and analyse the real life scenarios for various types of applications is employed in this study. From the point of view of network QoS, the network is concerned with how to transmit the sensed data from the sensor field to the sink node, fulfilling the corresponding required QoS. The factors that characterize the application requirement are presented in Table I.

TIT. SYSTEM COMMUNICATION PARADIGM

A system of WSN with identical sensor nodes deployed in a cluster tree topology is considered. The sensor nodes used are assumed to self-configure during initial deployment and remain stationary thereafter. All the nodes in a cluster and adjacent CHs are considered directly connected to the CH. The primary focus is to study the inter-arrival distribution of packets at the CH. The total arriving data packets at the CH at any given time is therefore equal to the sum of all the independent arrivals from the cluster nodes and arrivals from adjacent CHs forwarding their data to the sink. For this case continuous monitoring of event driven systems is considered.

In this set up all nodes are considered to be equipped with an omnidirectional antenna and they also have a common maximum radio range r within which they are able to sense event occurrences and also transmit information to the CH based on the 802.15.4/Zigbee standards. The topology of interest is shown in Figure 1. For simplicity, all sensor nodes are shown connected directly to the CHO in Figure 1. CHO can forward data to the sink either through CH1 or CH4, whereas CH2 and CH3 forward their packets to the sink passing through CHO. It is also shown that nodes N1 to N8 are directly connected to the CHO.

Each sensor node is able to independently monitor its habitat and organise the information sensed into fixed data units storable at the sensor buffer before finally forwarding to the CH. The buffers, both at the sensor nodes and at the CH are assumed to have infinite capacity and are following First in First out (FIFO) queuing discipline. The Cluster Head is only able to receive or transmit at one go within the assigned time slots of unit duration. Once Information sensed and aggregated at the nodes are forwarded to the CH, it finalizes cluster aggregation and transmits all the information to the sink either directly or through other intermediary CHs. It is assumed that at least one path always exists towards the sink [10].

In this study continuous monitoring applications where the nodes periodically (deterministic) sense and transmit information are considered for various MAC protocols, in order

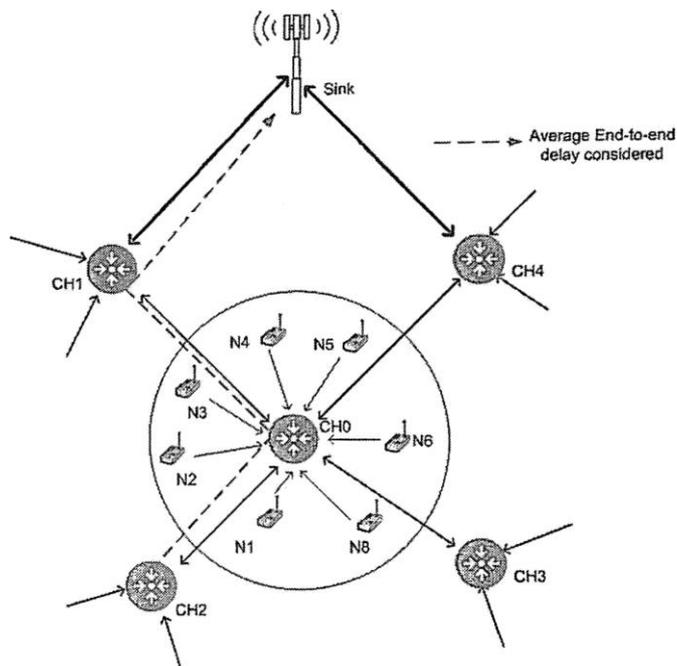


Figure 1: Network topology of the reference scenario

to see the effects of MAC protocols on the distribution of arrival process for the CHs. Castalia simulation environment is employed in order to calculate the arrivals of packets in the system, which are further used to analyse the inter-arrival distribution at the CH with the help of Q-Q plots. For each experiment, packet arrival rate and number of nodes is set at desired values. Desired MAC properties; TMAC, CSMA, and no MAC are then considered for each experiment. The generated inter-arrival distribution time results are then further analysed using Q-Q plots to identify the actual distribution pattern.

IV. NUMERICAL RESULTS AND DISCUSSIONS

Figures 2, 3 present the Q-Q plots of theoretical distributions and the empirical arrival distributions of simulated data series at the CH and the intermediary routing nodes.

The well-known theoretical distributions corresponding properly to the empirical distributions of the simulated data series are Exponential, Gamma, Log-Normal and Mixed Log-Normal distributions. Although we display Q-Q plots to compare empirical distribution to theoretical distributions whether these two population distributions are exactly the same, it is

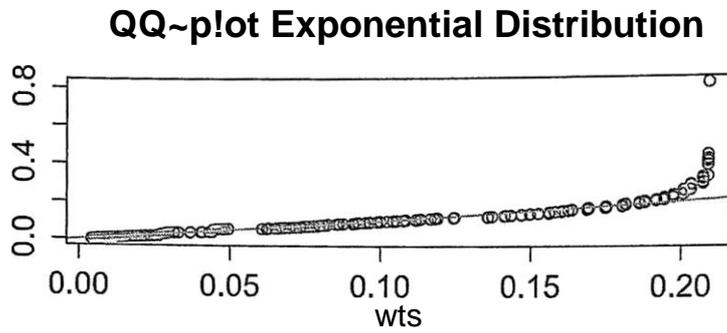


Figure 2. QQ-plo(for Exponential Distribution for 20 nodes with CSMA protocol, sending 1 packet/1 second

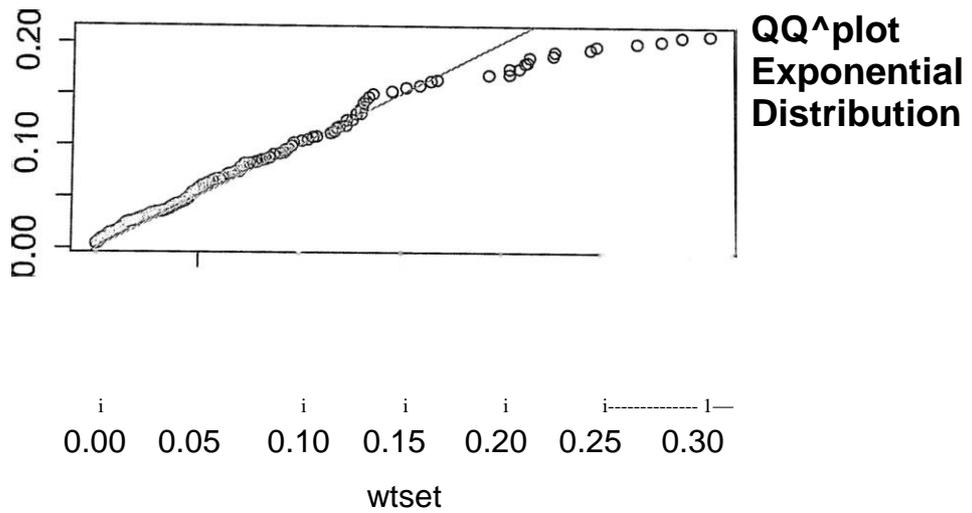


Figure 3: QQ-plot for Exponential Distribution for 10 nodes with CSMA, sending 1 packet/10 minutes

Table II: Average end-to-end delay for various application rates and MAC protocols applied

Nodes	1 packet every 5 min			1 packet every 5 sec			5 packet every sec		
	No MAC	TMAC	CSMA	No MAC	TMAC	CSMA	No MAC	TMAC	CSMA
10	0.03085413	0.03812471	0.03710934	0.036910934	0.0401187	0.039160543	0.040109	0.05791289	0.050281007
20	0.043585577	0.046399014	0.044012909	0.043902188	0.04895094	0.0462243	0.048610932	0.068023776	0.060010211
35	0.053775437	0.057112543	0.05489443	0.0539886	0.0600218	0.058330089	0.059124145	0.081778643	0.070666612
40	0.060177122	0.066062393	0.06276331	0.060289387	0.069779437	0.065319035	0.068324234	0.090668109	0.081224243

necessary to conduct a statistical test to prove it. $C^{\wedge} = \text{eye}$, the quantiles for the **first** distribution versus the quantiles for the second distribution will fall on the 0 Q-Q plots can be insufficient. It can be both difficult and subjective to decide how differences between distributions will yield various kinds of deviations from a straight line. Appendix A presents details about the probability plots or Q-Q plots.

Since the wireless channel is essentially a broadcast medium, only a single transmission is allowed in a transmission area by the MAC protocol. As a result, simultaneous transfers are not possible. Moreover, the MAC layer introduces a non-deterministic delay for channel access because of the activities of other nodes. If a neighbour of a node is transmitting a packet, the MAC protocol delays the transmission for a random amount of time to prevent collisions with the ongoing transmission as well as other neighbours that are trying to access the channel. This may significantly impact the performance of the network. The delay incurred due to various MAC protocols for different application rates is presented in Table II, from which a trade-off can be drawn between energy consumed and average end-to-end delay for the MAC protocol applied. For example, although TMAC reduces energy consumption of the nodes by adaptive duty cycling, there is an extra delay before the node can sense the channel due to node back-off's, therefore causing delays (especially in heavy traffic where it is backing off often). For 40 nodes, and higher packet rate of 5 every second, there is a difference of about 28% in average end-to-end delay when TMAC is employed and no MAC protocol employed, which is quite crucial when considering delay sensitive applications such as military and BANs.

V. CONCLUSIONS AND FUTURE WORK

To the best of our knowledge, this is the first work that provides a trade-off between energy consumption and average end-to-end delay incurred in the network, along with the necessity of validating the types of distributions and limitations of Q-Q plots for estimating the distribution of arrivals. A clustered model is considered characterised by its sending rate, inter-arrival distribution and the service process. The empirical distributions of inter-arrival times of the packets considering such physical events that do not occur frequently are generally assumed by Poisson processes, and the inter-arrival times by exponential distributions. The general practice in published works is thus to compare empirical exponential arrival distributions of wireless sensor networks with theoretical exponential distributions in Q-Q plot diagrams. In this paper, we show that such comparisons based on simple eye checks are not sufficient since in many cases incorrect conclusions may be drawn from such plots. This work is extended to estimate the inter-arrival distributions by estimating Maximum Likelihood parameters of empirical distributions, generate theoretical distributions based on the estimated parameters. By conducting Kolmogorov-Smirnov Test Statistics for each generated data series, we would like to find out, if it is possible, a corresponding theoretical distribution. The effects caused by MAC properties are also analysed by experimenting with well known MAC protocols. Therefore, these results confirm that the assumption of exponential inter-arrival distributions does not hold in all the cases. Exponential arrival distribution assumption of wireless sensor networks holds only when a fewer nodes (10-15),

market every 5-10 minutes with no MAC properties as when CSMA properties are considered. Also, end-to-end delay for various application rates is also presented whilst various MAC

protocols are considered to energy, presenting a trade-off between energy consumed and delay incurred in the network.

ACKNOWLEDGMENT

The authors would like to thank the members of SEN-COT for their valuable feedback.

APPENDIX A

PROBABILITY PLOTS OR QUANTILE-QUANTILE PLOTS

A probability plot or quantile-quantile (Q-Q) plot is a graphical display invented by Wilk and Gnanadesikan [33], to compare a data set to a particular probability distribution or to compare it to another data set. The idea is that if two population distributions are exactly the same, then they have the same quantiles (percentiles), so a plot of the quantiles for the first distribution versus the quantiles for the second distribution will fall on the 0-1 line (i.e., the straight line $y = x$ with intercept 0 and slope 1). If the two distributions have the same shape and spread but different locations, then the plot of the quantiles will fall on the line $y = a + x$ (parallel to the 0 - 1 line) where a denotes the difference in locations. If the distributions have different locations and differ by a multiplicative constant b , then the plot of the quantiles will fall on the line $y = a + bx$ [34], [35]. Various kinds of differences between distributions will yield various kinds of deviations from a straight line.

REFERENCES

- [1] K. Doddapaneni, E. Ever, O. Gemikonakli, I. Malavolta, L. Mostarda, and H. Muccini, "A model-driven engineering framework for architecting and analysing wireless sensor networks" in *Software Engineering for Sensor Network Applications (SESENA)*, 2012 Third International Workshop on, pp. 1-7, 2012.
- [2] N. Bouahdallali, M. E. Rivero-Angeles, and B. Sericola, "Continuous monitoring using event-driven reporting for cluster-based wireless sensor networks," *IEEE Transactions on Vehicular Technology* vol. 58 no. 7, pp. 3460-3479, 2009.
- [3] Y. Wang, M. C. Vuran, and S. Goddard, "Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 20, no. 1, pp. 305-318, 2012.
- [4] I. E. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer networks*, vol. 38 no. 4 pp. 393-422, 2002.
- [5] I. Dbibih, O. Zytoune, and D. Aboutajdine, "On/off markov model based energy-delay aware mac protocol for wireless sensor network," *Wireless Personal Communications*, vol. 78, no. 2 pp. 1143-1155 2014.
- [6] F. Omondi, E. Ever, P. Shah, and O. Gemikonakli, "Modelling wireless sensor networks for performance evaluation," in *Ad-hoc, Mobile, and Wireless Network (J. CichoAD, M. GEL'bala, and M. Klonowski, eds.)*, vol. 7960 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2013.
- [7] Y. Zhang and W. W. Li, "Modeling and energy consumption evaluation of a stochastic wireless sensor network," *EURASIP J. Wireless Comm, and Networking*, p. 282, 2012.
- [8] Y. Zhang and W. Li, "An energy-based stochastic model for wireless sensor networks." *Wireless Sensor Network*, vol. 3, no. 9, pp. 322-328, 2011.