Reliability, Scalability and Security in Smart Utility Networks

by

Gopalakrishnan B Iyer

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 4, 2014

Keywords: Wireless mesh networks, Smart utility networks, Performance analysis, Directed acyclic graphs, MMPP, OMNeT++

Approved by

Prathima Agrawal, Chair, Ginn Distinguished Professor, Electrical and Computer Engineering
Shiwen Mao, McWane Associate Professor, Electrical and Computer Engineering
Thaddeus Roppel, Associate Professor, Electrical and Computer Engineering
Alvin Lim, Associate Professor, Computer Science and Software Engineering

Abstract

With the rapid increase in deployment of smart meters across North America, resource consumption data of utilities such as water, gas and electricity is being collected at a much higher granular level. The huge amount of data is being used to make demand-response applications smarter. A large part of this deployment of smart utility networks and Advanced Metering Infrastructure (AMI) is being done using wireless mesh architecture due to the low deployment costs offered by this method. Although wireless environment parameters such as fading and path loss differ widely from home, outdoor to industrial, in-building scenarios, efficient protocols at the Medium Access Control (MAC) and Physical (PHY) layers and lower layer agnostic data routing protocols can be employed to overcome these challenges. For such routing protocols to be applicable to smart utility networks, reliability, scalability and security are vital metrics which will determine their performance in such networks.

Our objective in this dissertation is to present a Reliability, Scalability and Security (RSS-SUN) suite of recommended protocols most applicable for Smart Utility Networks. We present the performance analysis of wireless mesh routing protocols most applicable to smart utility networks. New and emerging routing protocols being proposed as an alternate standard by the Internet Engineering Task Force (IETF) community have been chosen for performance comparison. Herein, we analyze the reliability of three routing protocols, viz. RPL (IPv6 Routing Protocol for Low power and Lossy Networks), LOAD (6LoWPAN Ad Hoc On-Demand Distance Vector Routing) and a proprietary flavor of Geographical Routing developed by Landis+Gyr.

Routing metrics of packet success probability, end-end delay, hop count, link quality and packet delivery ratio have been considered for this performance analysis. Further, realistic network topologies, obtained from actual smart meter network deployments in North America have been modeled in simulations to derive at results pertinent to the applicability of these wireless mesh routing protocols to Advanced Metering Infrastructure comprising of smart utility networks.

With respect to scalability, we address the issue of scalable routing protocols being vital to its successful deployment in smart utility networks. To that effect, we present two approaches to modeling the wireless mesh network with the goal of analyzing the scalability of routing protocols in such networks with respect to large scale deployment.

Approach I models the network as being connected with a Poisson distribution with density and transmission range as parameters. We quantify an upper bound to the network size at a maximum total expected successful packet transmission as determined by the packet success probability of each link. We validate these results with simulation data from a large scale network using the supercomputer infrastructure at the Alabama Supercomputing Authority located in Huntsville, AL.

In Approach II, we model the wireless mesh network traffic arrival process as a Markov Modulated Poisson Process (MMPP) with two distinct modes. Further, an MMPP(2)/M/1/N queuing model is analyzed with the same goal of finding a network size upper bound, such that stability is maintained in the network. Verification of the model with analytical and simulation data is presented with conclusive scalability bounds affecting performance

With respect to Security, we present the urgent need to implement state-of-the-art cryptographic schemes to devices in smart utility networks. Further, we entail the security risks and possible mitigation practices. Two popular industry adopted public-key cryptographic schemes of RSA (Rivest-Shamir-Adleman) and Elliptic Curve Cryptography (ECC) are compared for performance on the Tmote-Sky hardware platform. We present the impact and trade-offs involved in implementing security, an indispensable requirement in smart utility networks.

Acknowledgments

There are several people in Auburn that I owe a debt of gratitude for assisting me in making my doctoral dreams a reality. I would like to express my sincere gratitude to my advisor, Dr. Prathima Agrawal, for convincing me to take the Ph.D. qualifying exams at the end of my Master's program here at Auburn. She has been a constant source of encouragement during the highs and lows of this program. Without her support, this research work and dissertation would not have been possible.

I would like to thank Dr. Shiwen Mao, Dr. Thaddeus Roppel and Dr. Alvin Lim for serving as members on my advisory committee. They have been kind, patient and provided guidance, corrections and suggestions to keep me steering the right way. I would also like to thank Dr. Saad Biaz for agreeing to be the external reader of this dissertation.

I would like to acknowledge the efforts and support of Ms. Shelia Collis, Ms. Jo Ann Loden, Ms. Penny Christopher and Ms. Jennifer Jackson for helping me with paper work related to graduate studies, immigration and travel matters.

My thanks go out to my present and former colleagues in Broun 405 for keeping the office lively and engaging in cerebral discussions. I would like to specially thank the group of Santosh Kulkarni, Vijay Sheshadri, Hui Zhou, Vignesh Sivakumar and Saketh Reddy for their excellent company. My thanks to former colleagues Yogesh Kondareddy, Nida Bano, Nirmal Andrews, Santosh Kulkarni, Pratap Prasad, Suryakant Bhandare and Mithun Raghav for making my time in Auburn filled with fun.

A special thanks to my friends Neeraj Naik, Vikas Iyer, Disha Vira, Roshan Pandey and Gurudutt Telang who have been there for me, both during the most testing times and enjoyable ones.

Above all, I am grateful to my parents Chitra and Balasubramanian, my sister Ranjani Kumar and my brother in-law Kartik Kumar for their encouragement and motivating talks. They have been a constant reminder of my purpose during these long years. Without their moral and emotional support, this research work and dissertation would not have been possible. I lovingly dedicate this work to them.

This work was supported by the NSF Wireless Internet Center for Advanced Technology (WICAT), the NSF Broadband Wireless Access and Applications Center (BWAC) and the Wireless Engineering Research and Education Center (WEREC) at Auburn University. We also acknowledge the support of Landis+Gyr, Alpharetta, GA as industry members of WICAT and BWAC.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Smart meters are devices which measure utility resource consumption, usually over a period of time or in units specified by the utility provider. These devices are comparable with sensors which are deployed in large scale to measure data that is important in analyzing a particular metric. The value in large scale collection of sensor data is greatly impacted by the collection method [1]. In the case of legacy utility meters, historically, manual collection of consumption data has been the norm followed by management and billing. With advancement in radio frequency technology and miniaturization of transceivers, it has been possible to aggregate such sensor data remotely for several decades. Large scale deployment of such remote data aggregating systems is now possible due to rapidly declining costs and mass manufacturing of such devices in accordance with Moore's law for silicon technology. Advanced Metering Infrastructure (AMI) comprising of smart meters, collectors, transmission and distribution devices and automated aggregation systems are now being widely deployed as a part of the Smart Grid initiative [2]. Figure 1.1 shows that over 36 million smart meter devices have been deployed as of 2012 in the United States alone. This is a 33% increase over 2011 deployments and is projected to be at 65 million by 2015 [3]. In a majority of deployments, devices in the AMI have wireless interfaces for communication to collector devices. In addition, smart utility networks are characterized by harsh wireless environment conditions in which they are deployed. Wireless fading and path loss have spatial and temporal components which require complex models to study performance in such environments. The enormous scale of the network also makes scalable solutions an important part of achieving efficiency in such networks.

A typical smart meter network can range from a few thousand metering nodes reporting to a single collector to millions of metering end points reporting to the head-end systems. These networks can be deployed using a plethora of evolved communication technologies that are available today. Physical layer technologies such as wireless cellular, power-line communications, fiber-optical networks and wireless mesh networks are under consideration for deploying AMI. The most popular technology yet has been low power wireless systems used in the wireless mesh architecture for deploying these networks. The wireless mesh architecture provides several advantages over the others including flexibility, minimal infrastructure, extensive network coverage and low deployment cost.

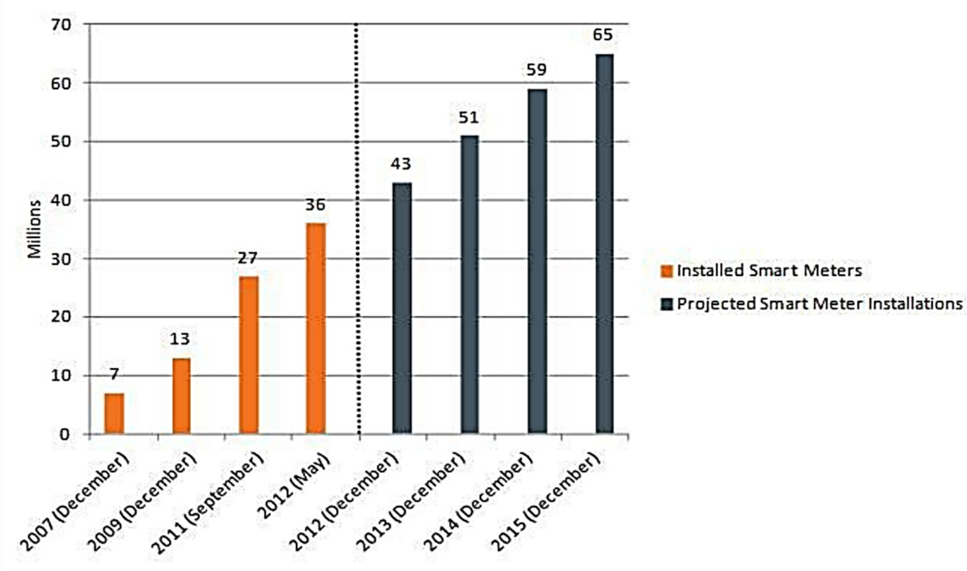


Figure 1.1: Current and Projected Smart Meter deployment data [3].

In addition, the usage of license free spectrum for deploying wireless mesh networks is highly lucrative. It is hence evident that an efficiently deployed smart utility network or AMI is vital to several key applications at the utility hierarchy. An efficient data routing protocol, which is agnostic to lower layer technologies is imperative to the efficiency and reliability of the smart utility network in general. In addition, inter-operability requirement is inherent to all devices in the AMI and must be the same with a routing protocol used in such a deployment. Three such IPv6 based routing protocols developed by the IETF (Internet Engineering Task Force) community viz. RPL [4] [5], LOAD [6] and LOADng [7] have been studied herein. Additionally, a proprietary flavor of Geographical routing protocol developed by Landis+Gyr has been studied for comparison [8]. Previous work done in [9] entails performance analysis using the metrics of hop-count and end to end delay, comparing RPL and Geographical routing for a real deployed 500 node smart utility network. We extend the performance analysis to LOAD routing protocol in smart utility networks. In addition, varying network sizes of up to 7500 nodes with a single collector have been used herein, allowing a high-level scalability analysis of these routing protocols.

Due to the enormous scale of deployment, scalability is a vital requirement of protocols deployed in smart utility networks. Here, we analyze scalability by modeling the smart utility network using two approaches. One approach aims to determine the scalability according to the link packet success probability by modeling the node connectivity using a Poisson distribution. In this analysis, we arrive at an analytic result showing the dependency of scale on packet success probability. Further, we verify the analytic results by using a large scale simulation of such networks. A second approach models the network as a single server queue with a Markov Modulated Poisson arrival process, exponential service and finite buffer. We present an analytic result in which the scale is dependent on the mode of traffic in the network. The scalability relationships thus derived from both the approaches can be used in different smart utility metering scenarios to estimate a viable scale in such networks.

Security is an inherent requirement in smart utility networks. Since these networks are tied to physical systems, a compromise in confidentiality, integrity and availability of control in such networks will be far more consequential than in traditional internetworks. Addition of security in smart utility networks presents an inevitable trade-off in performance and power consumption. The impact of security overhead is analyzed herein. Two popular public-key cryptographic schemes, viz. RSA-1024 bit and ECC-160 bit are implemented on the Tmote-Sky hardware platform and their power consumption in a test mesh network are compared versus a node with no-encryption for their impact on power consumption in smart utility networks.

Chapter 2

Reliability of routing protocols in Smart Utility Networks

Smart utility networks can be deployed using various wireless network architectures. These networks can be implemented as wireless cellular networks or as a wireless mesh networks. The goal is to provide complete connectivity to all metering nodes in the network. High reliability and low end to end latency combined with 100% packet success and delivery ratios are the requirements of an AMI network. In such networks, the quality of service is of utmost importance. Cellular network architectures in general are providers of greater quality of service as more often than not, they use licensed radio spectrum for their service and are fairly immune from the effects of interference. This type of deployment bears a significant infrastructure cost and complexity of deployment. The cellular deployment also requires extensive radio planning in order to optimize installation and use of infrastructure. On the other hand, a wireless mesh network does not involve use of complex infrastructure. An electric utility AMI network consists of the electric meters themselves, equipped with wireless communication capabilities and several collection points. In this mesh network, the electricity meters form an ad-hoc self-organizing mesh and communicate with the collection points known as collectors, thus requiring minimal deployment effort and network management. Wireless mesh network deployment is an optimal choice for deploying an AMI. Figure 2.1 shows the hierarchical architecture of Smart Utility Networks, in which several metering devices with routers and collectors in a mesh network form the Neighborhood area network and report to the head end systems though backhaul wide area networks.

2.1 Radio and Wireless Environment

The wireless environment places fundamental limitations on the performance of communication networks based on radio frequencies. The path between a pair of transmitter and receiver can vary between plain line of sight to tall buildings, forest areas and mountainous terrain. These variations in geographical topology have given rise to the challenging problems of signal fading due to path loss and fitting of complex signal propagation models based on statistical data.

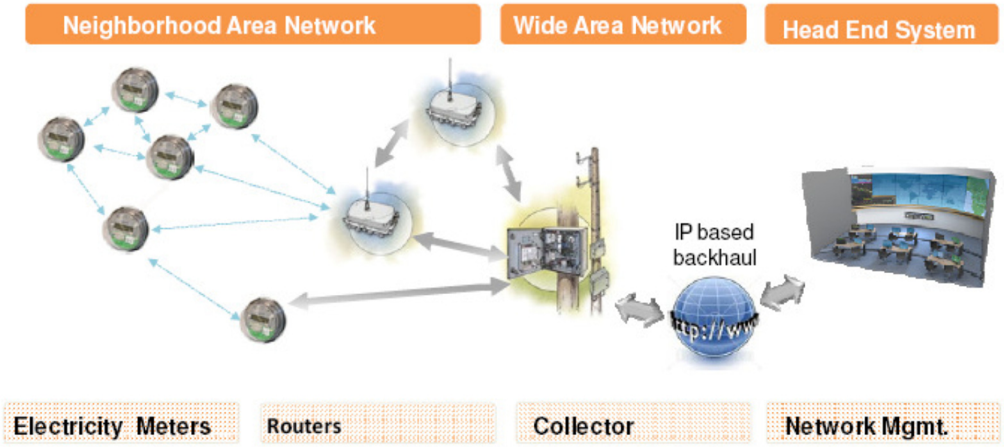


Figure 2.1: Smart Utility Mesh based Architecture [8].

In addition, the wireless channel is plagued by the ubiquitous problem of interference from a variety of sources ranging from cosmic background noise, to the electromagnetic noise generated by household microwaves. Even though there is a big corpus of theoretical results and experimental measurements for wireless communication in general, they tend to practically affect areas with more commercial impact such as cellular communication.

In Smart utility networks and the AMI, the wireless environment varies from indoor for home automation application to outdoor and industrial, highly noisy conditions. One important aspect of this wireless channel modeling is to estimate the average path loss between two nodes, or in general, two points in space. For smart utility networks, where the separation of nodes is from a couple of meters to a hundred meters, the lognormal shadowing model has been shown to give accurate estimates for average path loss. Equation 2.1 returns path loss in dB as a function of the distance between two nodes.

$$PL(d) = PL(d_0) + 10\,\eta \log\left(\frac{d}{d_0}\right) + X_\sigma \tag{2.1}$$

PL (d) is the path loss at distance **d**, ***PL (d0)*** is the known path loss at a reference distance ***d0***, ***η*** is the path loss exponent, and **X$_\sigma$** is a Gaussian zero-mean random variable with standard deviation ***σ.***

Another very important aspect of the wireless channel is the temporal variation of channel quality. This is especially pronounced in rapidly changing environments as those experienced in a smart utility network. Figure 2.2 shows a typical profile of channel variation. Notice the sharp drops (measured up to -50dB) and that most of the time the channel is below the average path loss (0dB). There are several theoretical models to describe temporal variation, taking their names from corresponding complex distributions (e.g., Rayleigh, Weibull, Nakagami, Gamma and lognormal) [10].

The choice of radio for any hardware mainly depends on the factors of power, data-rate and operating frequencies. Proprietary radios generally operate in a licensed spectrum for which the user is billed. On the other hand, IEEE standards based radios operate in the license free spectrum allocated for target applications by the FCC. Wireless Local Area Network (WLAN) and Worldwide Interoperability for Microwave Access (WiMAX) are the most popular examples operating on IEEE 802.11 and IEEE 802.16 standards, respectively.

Smart Utility Networks are comprised of low power devices communicating at a low data rate. The maximum allowed transmission power for such devices is 30dBm (1000 mW) and can communicate at a maximum data rate of 250 kbps. Typically, devices in Smart Utility Network operate at 0dBm and a variable data rate between 25 kbps and 250 kbps. The IEEE 802.15.4 specifications for MAC and PHY are built for such target devices, operating at low power, low data rates and in highly lossy environments. Moreover, these devices operate in the license free Industrial, Scientific and Medical (ISM) radio bands centered around 900 MHz and 2.4 GHz frequency bands. The Zigbee alliance has adopted these standards for all Zigbee enabled devices. The IETF working group 6LoWPAN (IPv6 over Low power Wireless Personal Area Networks) has adopted this standard for building a wireless embedded internet [11], [12], [13].

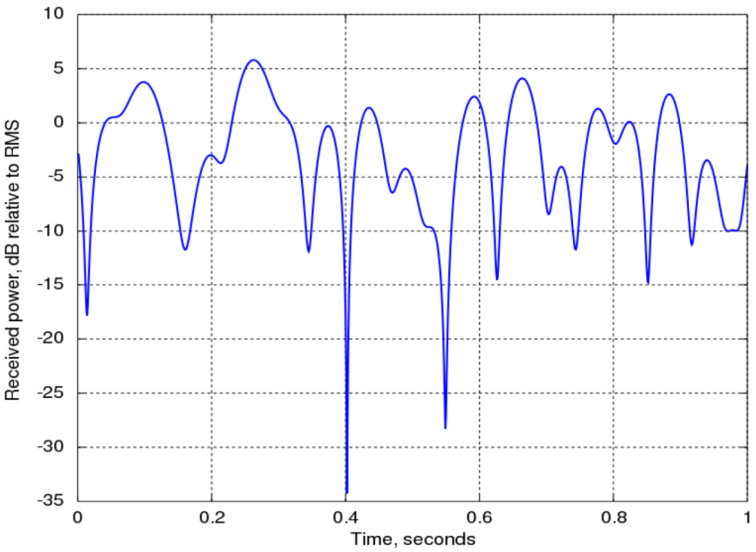


Figure 2.2: Temporal fading in a wireless channel [14].

2.2 Real Network Topology

In the study of wireless mesh networks, a playground grid size is chosen for the simulation. In this playground, nodes are placed at random locations based on a chosen seed or they are distributed according to a uniform distribution with chosen mean and standard deviation. Although these distributions are good for simulations, they do not represent the way in which smart metering nodes are distributed in a real-world network.

In order to study routing protocols for smart utility networks, forming a wireless mesh, incorporating real world topologies are essential. Trace files containing geographical locations of real world deployed smart metering nodes were obtained and used in simulations to arrive at simulation results and conclusion herein. Varying network size ranging from 25 nodes to 7500 metering nodes with the collector located at the center of the deployed area derived from a large AMI deployment were used as network topologies for the simulation of data traffic in smart utility networks. In order to obtain real world comparable performance metrics from the routing protocols, it was necessary to use real deployed network topologies instead of a randomized topology. In general, connectivity and performance are directly related to the network topology in which a routing protocol is used. Figure 2.3 shows the real geographical distribution of a 7500 node network used for the performance analysis.

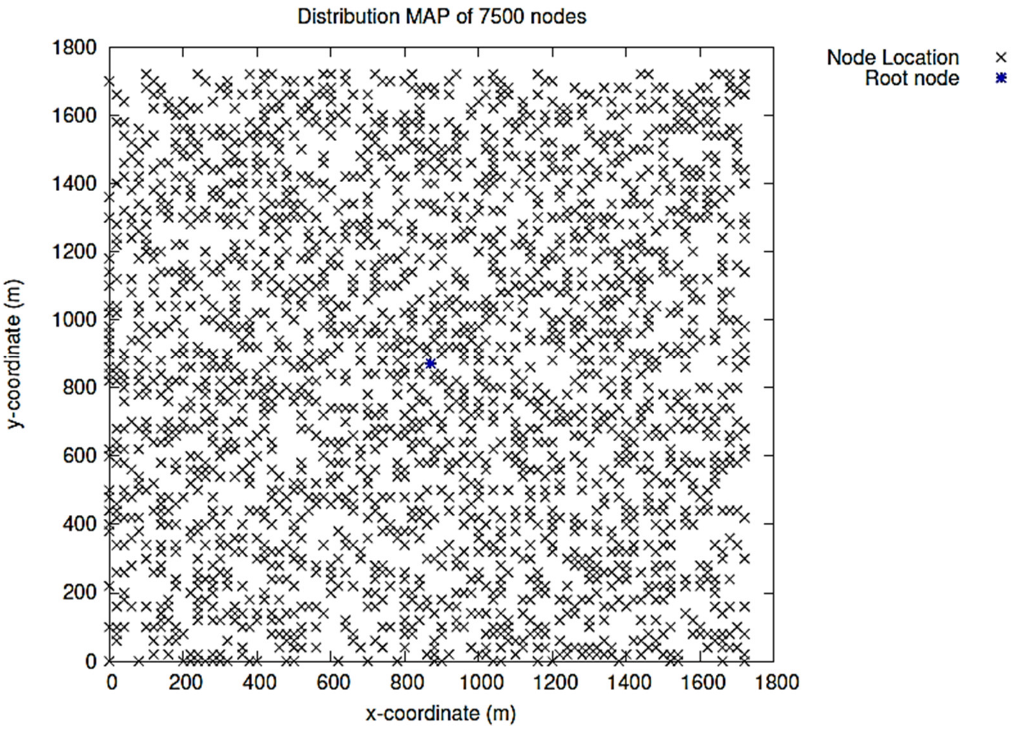


Figure 2.3: Geographical distribution of a 7500 node network

2.3 Routing Protocols

In order for the smart utility network to function as designed for its application, an efficient routing protocol is imperative. A single routing protocol is designed with target applications, according to which the routing algorithm optimizes routes and controls network functions. The requirements for routing in a smart utility network are diverse, as it has to be compatible for a wide range of devices, from home appliances to metering devices to collectors. Routing Protocols designed for the smart utility networks have to adhere to the requirements spelled out in [RFC5826], [RFC5673], [RFC5548] and [RFC5867] which are home automation routing, industrial routing, urban routing environments and in-building routing requirements respectively.

The nodes participating in these networks are low power devices and have severe constraints in terms of CPU power, memory, battery and operating environments. Utility meters are designed to last for approximately 20 years on battery. Due to these constraints, restrictions are placed on routing requirements. For Home automation applications, the duty cycle of devices is less than 1% and the end to end delay requirement for the routing protocol is 500 milliseconds.

The routing protocol has to perform device aware routing, and must avoid battery constrained devices or devices under strain for frequent routing of packets. For industrial automation applications, more focus is on the reliability which has to be as close to 100% as possible. Reliability is measured by the number of successful transmissions over number of attempts in a given period of time. Also minimal network re-configuration is required when new devices enter or leave the network.

Wireless mesh routing protocols in smart utility networks, viz. RPL, LOAD and Geographical routing have been analyzed and compared for performance herein. The following is a brief overview of each of the routing protocols.

RPL is based on construction of a DODAG (Destination Oriented Directed Acyclic Graph) for the network which has the collector or data aggregating device as the root of the directed acyclic graph. It is a distance vector routing protocol for Low Power and Lossy Networks (LLNs). The route is constructed with a goal of optimizing an objective-function. The objective function is constructed using a combination of metrics and constraints to compute the best path for routing data packets to the destination. Thus, there can be multiple DODAGs for the same network depending on the objective-function. Applications such as outage reporting may require hard deadlines on delivery and can use an objective function to minimize latency as an example. Network state information such as connectivity and degree of connectivity of each node to its parent node up to the root is maintained and periodically updated using the RPL control plane. The root node initiates formation of the graph using the DODAG Information Object (DIO) control message. Consider node j receiving a DIO (DODAG Information Object) from node i. Upon receiving the DIO message the j^{th} node will compute its own rank R_j according to eq. (2.2) and determine its position in the DODAG with respect to other nodes. Parameter β defines the granularity of the network topology. It is defined as minimum rank increase and can be adjusted as required during deployment. β varies depending on the network density. For a sparse network, β is small and for a dense network β is large. Figure 2.4 illustrates a RPL instance after formation of the DODAG. DIOs are used to form up-routes for MP2P (Multipoint to Point) traffic.

$$R_j = floor\left({}^{R_i}/_{\beta}\right) \tag{2.2}$$

RPL also creates down-routes for P2MP (Point to Multipoint) traffic. DAO (Destination Advertisement Object) messages are periodically transmitted by the root requesting reverse route to a destination node in the tree. DAO messages are used to advertise prefix reach-ability towards the leaf nodes in support of the down traffic. RPL forms acyclic graphs and inherently does not allow nodes at the same hierarchy level to form an edge. A modification in RPL supports point-to-point (P2P) traffic instead of data packets traversing all the way up to the root and then down-route. RPL makes use of the trickle-timer [15] in order to control flooding at the control plane. Trickle timer governs the DIO conditions and DAO intervals. Flooding of control packets in smart utility networks would be particularly costly where node resources such as energy and computing power are scarce and highly constrained. The adaptive trickle-timer is crucial to the efficiency of RPL. As the routes form and the nodes are connected (a rank is established) the network stabilizes and control plane messages decrease. When an inconsistency arises (such as a loop or a change in the DODAG parameters) as local repair is initiated by the node which detected the inconsistency, the trickle timer is reset and DIOs are exchanged to quickly restore connectivity. In the worst case, a global repair is initiated where the graph is built starting at the root. The metric used to make the rank change decision is Expected Transmission Count (ETX). It is calculated using the link packet success probability. Packet success probability of each link k, denoted p_k is randomly assigned for each link to a parent. The metric ETX of link k, X (k), is computed as the inverse of the packet success probability of each link according to eq. (2.3).

$$X(k) = \frac{1}{p_k} \tag{2.3}$$

The initial value of ETX of any link k is set to be 1.0. The ETX is updated and calculated from Equation 4.1 from a randomized set of packet success probability values when a DIO is received by a node from a parent. Hence, the ETX value of each link keeps changing over time. This introduces the temporal nature of link characteristics observed in Low power and Lossy Networks (LLNs). The rank is computed based on the updated ETX count and is transmitted in the DIO. Figure 2.4 shows an instance of the Directed Acyclic Graph of RPL after DIO control packets exchange. Node 1 is the root of the graph with the least rank and node 5 is the leaf node with the maximum rank.

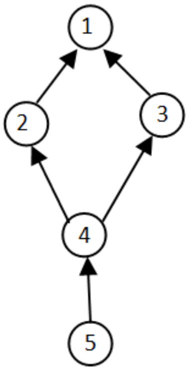


Figure 2.4: Instance of RPL after DODAG formation

LOAD was originally designed to be a light-weight modification of the popular Ad-hoc On-demand Distance Vector routing protocol (AODV) for 6lowpan networks [16]. It has been simplified by not requiring the intermediate nodes to reply a RREQ (Route Request) control packet with the RREP (Route Reply). Instead, only the destination for which the RREQ was intended replies to the route-request. In addition, a table maintaining RREQ for each node has been eliminated. Also, LOAD does not use the precursor list of AODV in order to simplify the routing table structure. LOAD does not use the destination sequence number as gratuitous RREPs are no longer permitted. The accumulated route cost such as Link Quality Indicator (LQI) and the number of hops from the source to the destination are the preferred routing metrics in LOAD. A route is preferred if the number of weak links along the way is smaller and least number of hops from the source to the destination. In LOAD, whenever there is a break in the link, the upstream node of the broken link may try to repair the route locally by using the route discovery mechanism. LOAD uses periodic HELLO messages, governed by a timer to maintain route state information. The process of LOAD control plane message exchange is illustrated in Figure 2.5. Source node S initiates a route request destined for D and is forwarded to D by intermediate nodes N1 and N2. Destination D replies to the RREQ by sending a route reply (RREP) via the reverse route in the RREQ message (i.e. via N2 and N1 to S). After S receives the RREP, data exchange begins. Additionally, if the data packet fails at any node, a route error (RERR) message is issued to the preceding node. In the LOAD control plane, there is no separate mechanism in comparison to RPL for up and down routes.

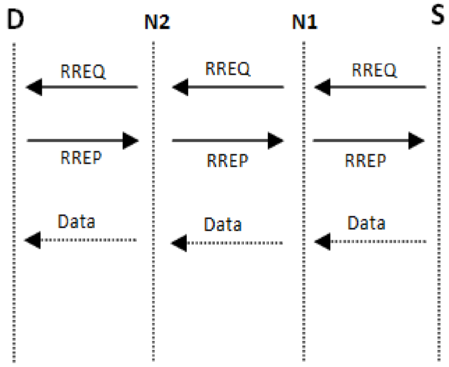


Figure 2.5: Exchange of LOAD control messages

Each source sending RREQ to a particular destination receives a RREP with the reverse route to the destination node contained in the message. This reverse route is used to reach the destination. Thus, a special hierarchy cannot be established in LOAD wherein the root node, participating nodes and leaf nodes, all discover each other by the same mechanism and maintain routes using periodic HELLO messages. Discovery and formation of routes is on-demand. This implies that routes will not be discovered until there is an application packet in the buffer to be sent to a particular destination.

The proprietary flavor of Geographical routing protocol developed by Landis+Gyr is built around P2P traffic exchanges. Nodes poll each other using two mechanisms called TICKLE (similar to request) and ACK (acknowledgement of the TICKLE). These control exchanges contain geographical information in the form of co-ordinates (latitude, longitude and altitude) of the nodes. In addition, they signal if a node is ready to receive a data packet. In order to make routing decisions, the nodes are aware of the geographical location of the collector which is the upstream destination. The collector is also aware of the geographical location of target nodes for downstream communications. The node initiating route discovery builds a scan table to select a neighbor which is geographically proximate to the destination node and computes its distance using eq. (2.4).

$$d_{ji} = \sqrt{\left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2 + \left(z_j - z_i\right)^2} \tag{2.4}$$

Node *j* receiving the geographical information from node *i* computes distance d_{ji} and stores in a table. Sorted according to shortest distance, each node is then polled to transmit data if available at buffer. This selection method also considers the dynamically changing link quality in a wireless environment and combines geographical proximity with a weighted link quality indicator metric in order to select the least cost neighbor. A heuristic approach to routing is employed in which active links are strengthened, ensuring stability in the network topology. Figure 2.6 illustrates a metering source S selecting next hop metering node A to forward data to the collector (destination) D by geographical proximity. The altitude or *z* component is important in the realistic perspective, since collectors or specially designated 'Routers' are typically mounted on top of electric poles or a higher altitude for better line of sight with meters in a dense radio environment. If a designated router is in the scan table, it is selected as a preferred neighbor by default due to better connectivity with the collector. This implementation of geographical routing allows for a minimal state routing approach and can adapt to the highly temporal characteristics wireless medium. It has been deployed in large scale networks while maintaining the same level of complexity. To reflect changing link quality in routing decisions, accumulated packet success rates are considered after choosing the nearest node for the next hop.

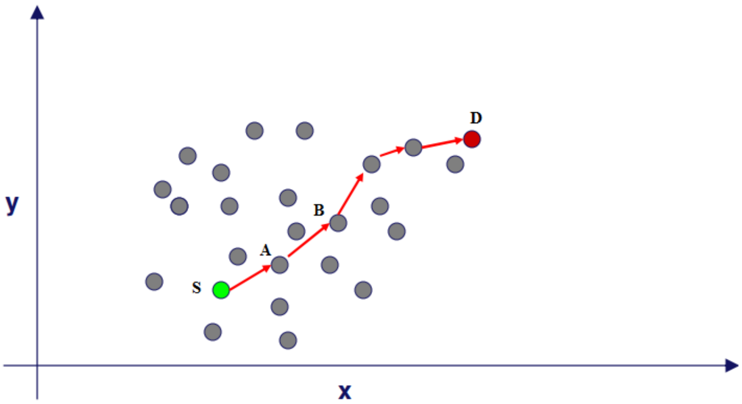


Figure 2.6: Geographical routing mechanism

These values are evaluated continuously in every exchange and the link table keeps track of the metric for routing decisions. Large scale deployment of nodes with geographical routing initially required feeding GPS co-ordinates into the nodes manually. This method has now been deprecated and a geographical co-ordinate system based on the node's link layer address has been developed for rapid deployment.

2.4 Performance Analysis

In order to compare the performance of the wireless mesh routing protocols a simulation setup was developed using the discrete event simulator OMNeT++ [17] and the wireless sensor network framework Castalia [18]. The 802.15.4 MAC/PHY modules provided with the Castalia framework were used as lower layers under the routing layers of RPL and Geographical routing developed at a conceptual level for this performance analysis. As Castalia was developed for BANs (Body Area Networks) it included temporal channel variation and fading models which were close to real world low power and lossy networks. These models made Castalia a right choice over other frameworks available for wireless sensor network simulations. Geographical routing protocol, although proprietary, has been widely deployed in smart utility networks and AMI systems and is currently the protocol running in over 2 million metering end-points.

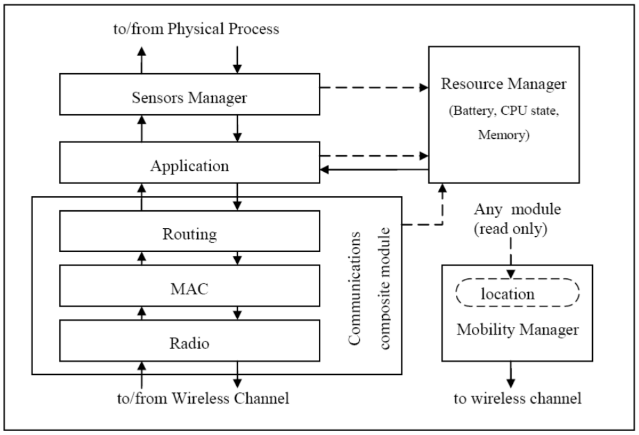


Figure 2.7: Composite node module in Castalia

Figure 2.7 shows the architecture of a node in the Castalia framework comprising of an application module, communication module (Routing, MAC and PHY) and resource management module. Simulation parameters for the IEEE 802.15.4 radio and wireless environment are summarized in tables 2.1 and 2.2 respectively.

Table 2.1: IEEE 802.15.4 PHY parameters

Parameter	**Value**
Radio Data Rate	250 kbps
Modulation Scheme	Frequency Shift Keying (FSK)
Bandwidth	2 MHz
Receiver Sensitivity	-95dBm
Transmission Power	20dBm
Operating Frequency	900 MHz

Table 2.2: Wireless environment parameters

Parameter	**Value**
Path loss at distance of d_0	60dB
Reference distance d_0	1m
Path Loss Exponent (β)	2.7
Standard Deviation of Gaussian distribution (σ)	4.0

In order to analyze multiple routing protocols, the underlying link layer, physical layer and the channel conditions have to be optimized according to the network topology in use and the environment in which such networks are deployed. The choice of IEEE 802.15.4 MAC/PHY standards for such networks is optimal as it is similar to the future standard 802.15.4g for Neighborhood Area Networks. This standard is a likely candidate for deployment in smart utility networks and has been adopted by the Zigbee Alliance for home automation networks. With the aim of quickly validating the link layer model for analysis of the routing protocols, packet success probability rates were gathered from a 400 node real network comprising of smart meters and network devices. The packet success probability distribution in the network was obtained as shown in Figure 2.8. These values represent the real world channel characteristics reflected in the packet success probabilities at the link layer.

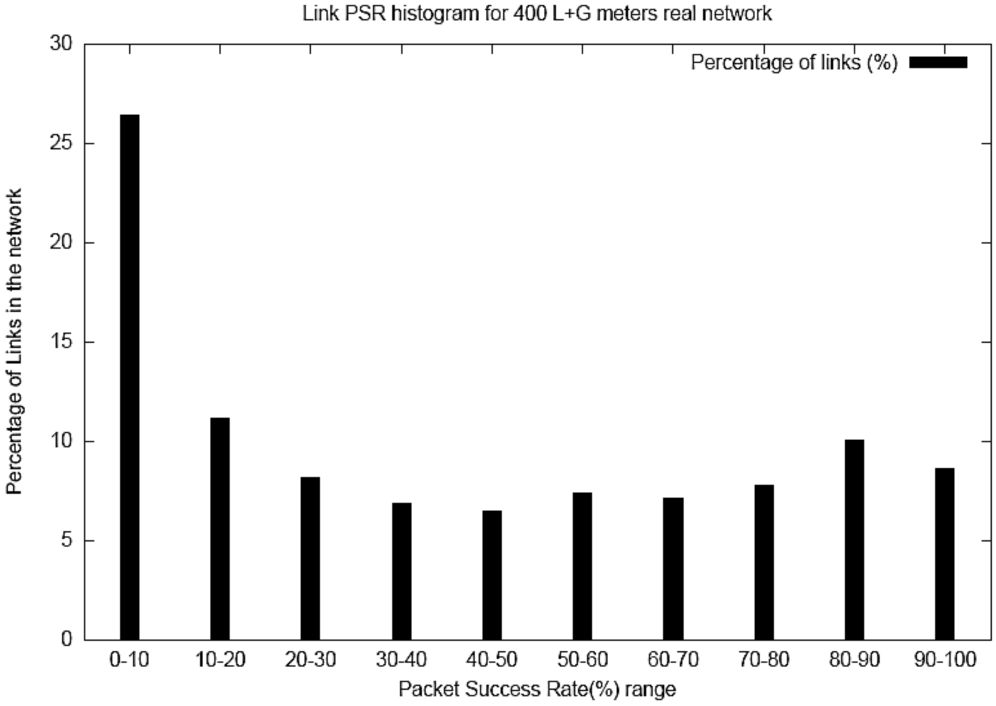


Figure 2.8: Packet Success Rate Histogram for Validation

After having computed the packet success probability distribution for a 400 node real network for comparison, a simulation experiment was conducted to determine the packet success probability yielded by the simulator's link, physical and channel models. This was done for a 100 node subset derived from a large rurally deployed real network. This experiment set each node to broadcast a total of 100 packets in a sequential order. The received packets were counted and the source was recorded. The computed ratio of total transmitted packets to those received by each node was collected as the link packet success probability. The packet success probability values obtained from the experiment were used to plot the connectivity map of the nodes with the collector. Figure 2.9 shows the topology of the 100 node network used for the experiment and all the possible links between nodes. The nodes with links to the collector in figure 2.9, have a packet success rate of more than 60%. The values obtained from the experiment with the 100 node network and the on-field data from the 400 node network showed a fair correlation in packet success probability. This experiment quickly validated the simulation model for IEEE 802.15.4 MAC/PHY to be yielding link characteristics similar to those obtained from field measurements. This validation allowed the use of the IEEE 802.15.4 MAC/PHY simulation models to be used under the two routing layers of RPL and Geographical routing for performance analysis.

In addition to the validation with field data provided by Landis+Gyr, we ran an experiment with 9 nodes in a grid topology in both the simulator and on the Tmote-sky hardware platform. The Tmote-sky radio is a Texas Instrument CC2420 radio conforming to IEEE 802.15.4 standards. The Received Signal Strength Index (RSSI) was obtained over a 24hr period. Figure 2.10 shows that the simulation environment is finely tuned to the radio and wireless characteristics of the real world. This validates the simulators close emulation of IEEE 802.15.4 radio and the wireless channel characteristics used in the simulation parameters.

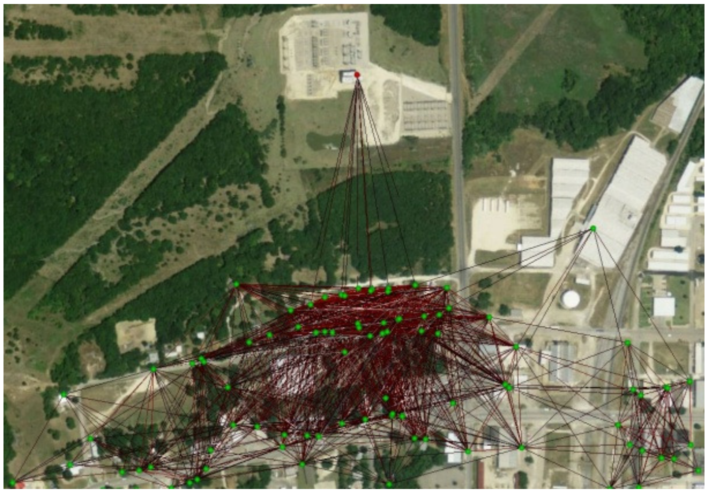


Figure 2.9: Plot of links with PSR greater than 60% in the network

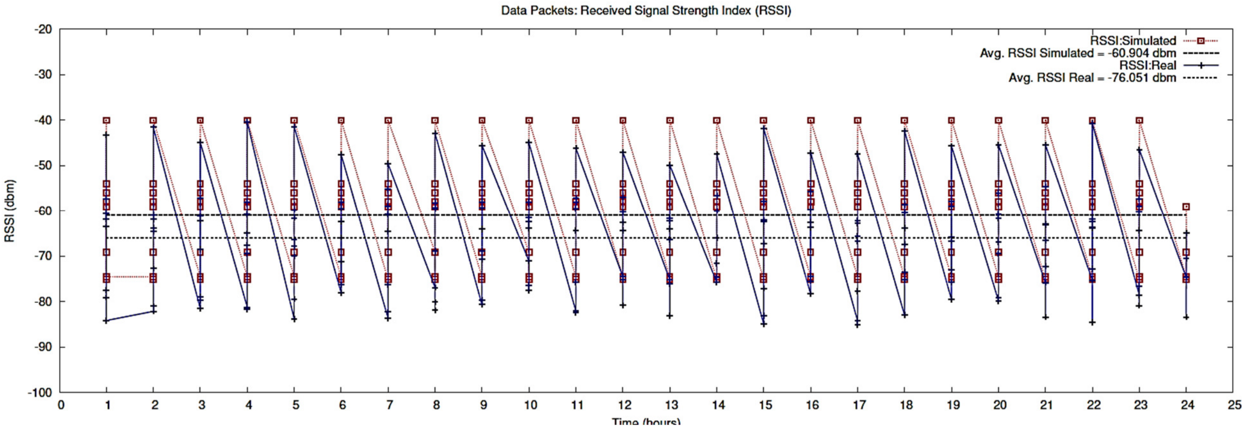
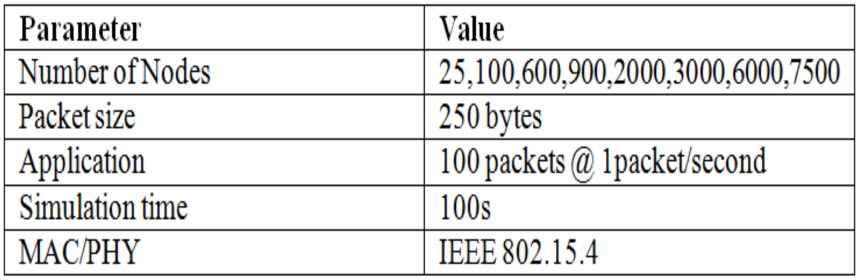


Figure 2.10: RSSI values Real vs. Simulated over 24 hr period.

Table 2.3: Simulation Parameters



Parameter	Value
Number of Nodes	25,100,600,900,2000,3000,6000,7500
Packet size	250 bytes
Application	100 packets @ 1packet/second
Simulation time	100s
MAC/PHY	IEEE 802.15.4

Having validated the simulation model, simulation experiments with varying network size from 25 to 7500 were run. The simulation parameters are listed in Table 2.3. The statistics of End-End delay, hop count and packet delivery ratio were collected from the simulator. The results are presented in figures 2.11, 2.12 and 2.13 respectively. The delay and hop count results are for a network size of 500 nodes. In addition, statistics from the control plane of RPL, LOAD and Geographical routing were collected. The control plane results are presented in figure 2.14.

Simulations of RPL, LOAD and Geographical routing protocols were run with the simulation parameters for varying network sizes. The traffic scenario simulated was MP2P (Multipoint to Point), wherein all meters send periodic data packets to the collector. This traffic scenario is the most prevalent in smart utility networks. The application packet rate was set at 1 packet/second. For a total simulation time of 100 seconds, the application layer generated a total of 100 packets.

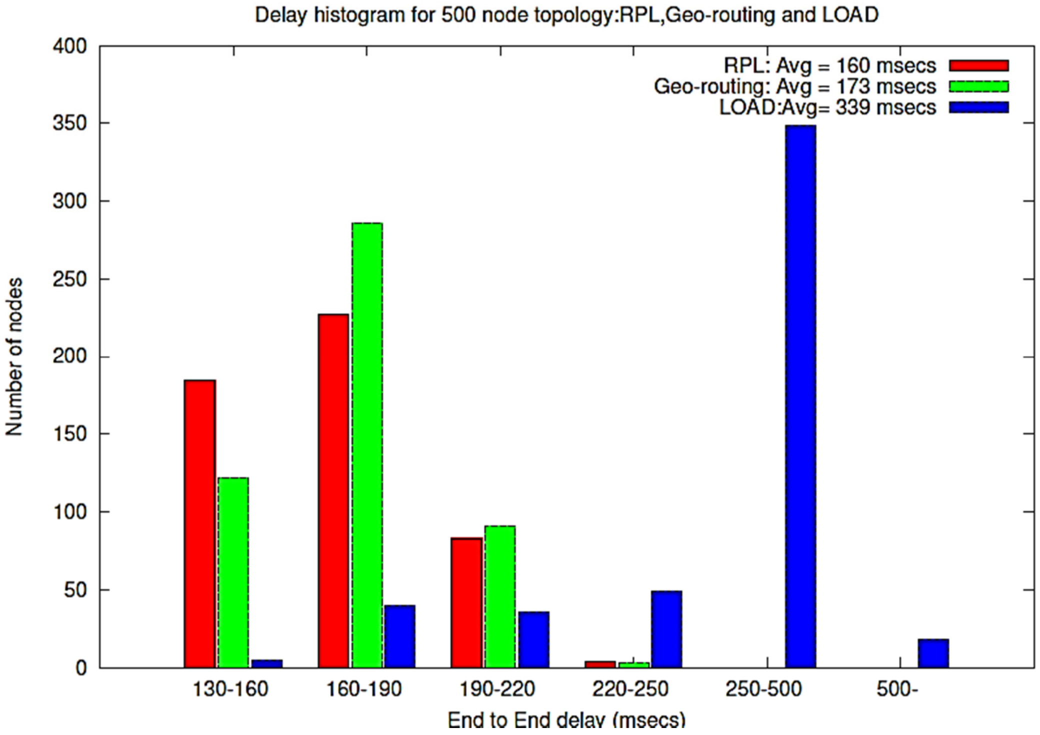


Figure 2.11: Delay histogram

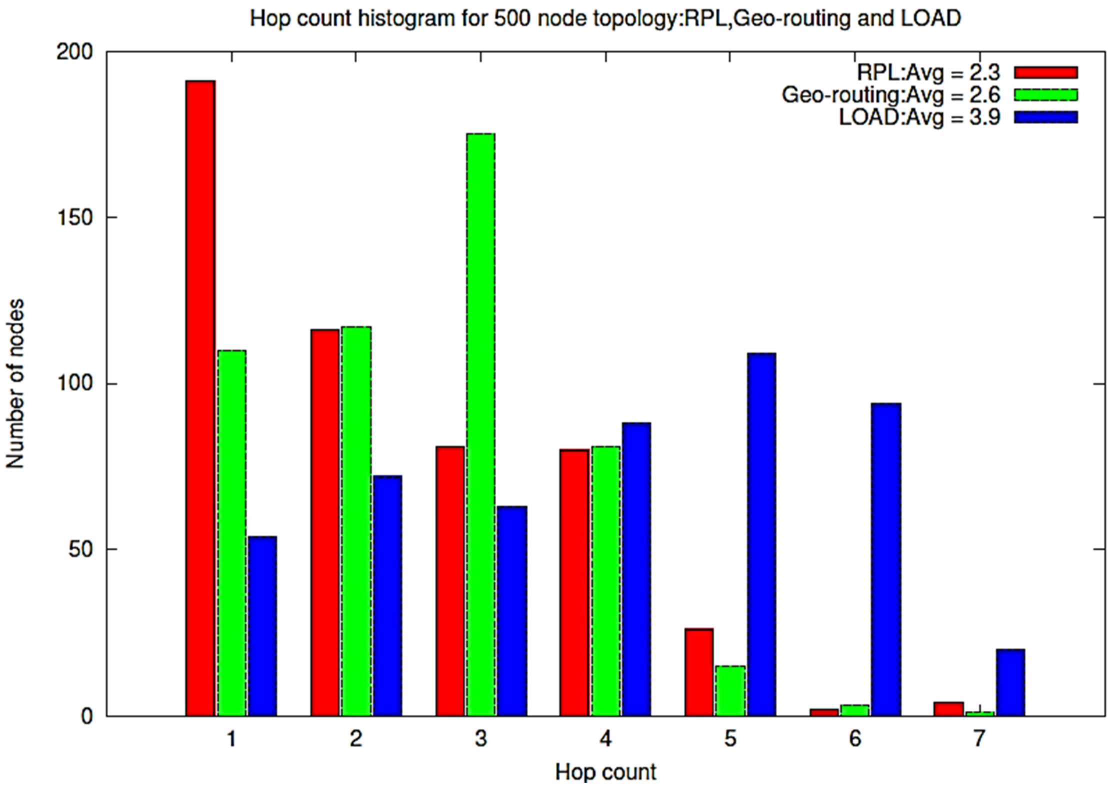


Figure 2.12: Hop Count Histogram

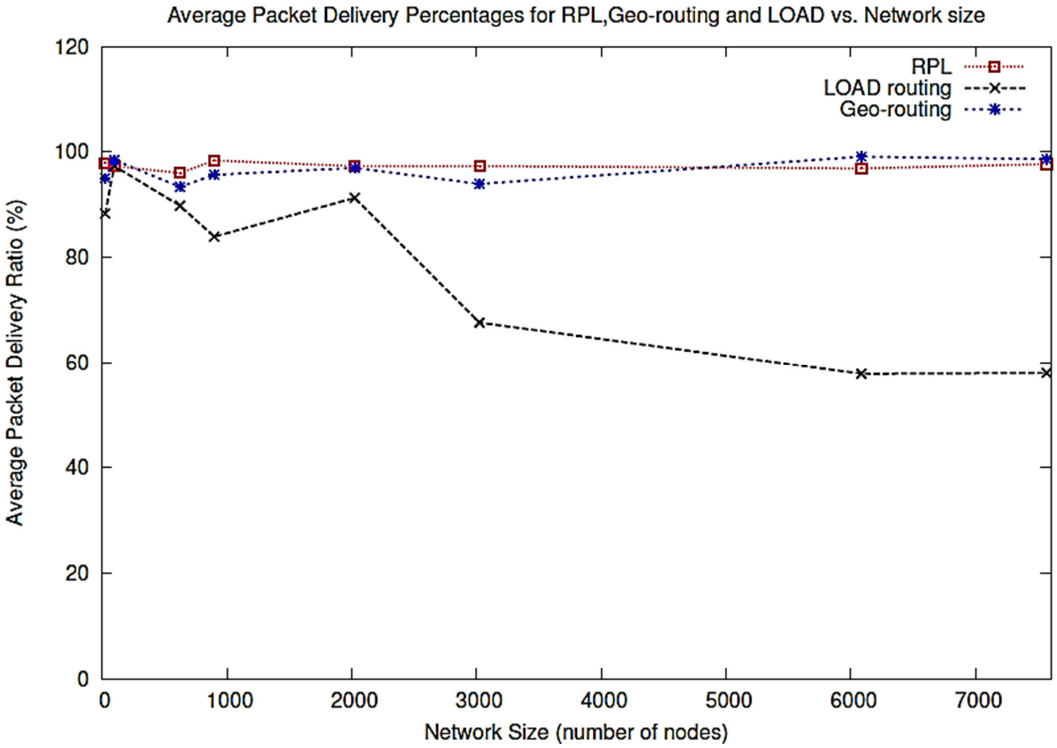


Figure 2.13: Packet Delivery Ratio

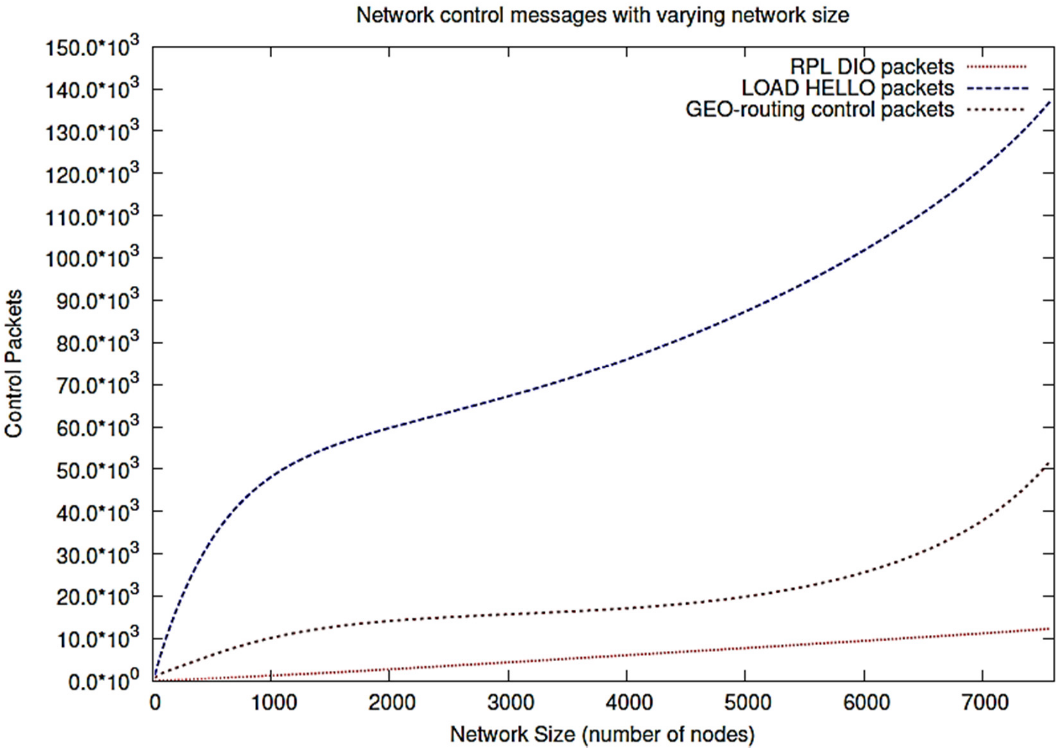


Figure 2.14: Control Messages

Each node transmitted 100 packets with the collector as the destination. The statistics of hop count, delay and packet delivery ratio were gathered for the entire network. At each node, the total number of control packets corresponding to each routing protocol was measured.

Figure 2.11 shows the delay histogram for the 500 node real network for RPL, LOAD and Geographical routing protocols. Overall the packets recorded the least delay with RPL. The data shows that on an average, packets suffer 160msecs, 173msecs and 339msecs of delay for RPL, Geographical routing and LOAD respectively. Reliability, in terms of packet delivery ratio was calculated as the ratio of total packets received over total packets transmitted at the root node.

Figure 2.12 shows the hop count histogram for a 500 node intermediate network size. A maximum of 7 hops was observed for each of the three routing protocols, RPL, LOAD and Geographical routing. RPL constructs a route in which 190 nodes are 1 hop from the root. In contrast 110 nodes are 5 hops away from the collector as constructed by LOAD. 175 nodes are 3 hops away from the collector as constructed by Geographical routing. Among the three RPL places maximum number of nodes 1 hop away from the root which is a good metric of its performance. Overall, RPL has an average hop-count of 2.3 hops, LOAD of 3.9 hops and Geographical routing has an average of 2.6 hops.

Figure 2.13 shows the packet delivery ratio (PDR) comparison of RPL, Geographical routing and LOAD. RPL and Geographical routing perform consistently between 95%-98% for all network sizes. LOAD has the lowest PDR of 58% for the largest network size. It performs reasonably well for smaller networks in terms of reliability. Figure 2.14 shows the control packet comparison of the three routing protocols. RPL shows a linear increase in the control packets with network size while both LOAD and Geographical routing have an exponential increase trend. LOAD generates a significantly more number of control packets throughout all network sizes in comparison to RPL and Geographical routing. These results show that RPL performs better in terms of delay and hop count. It is also the most reliable routing protocol among the three, providing high packet delivery ratios consistently across network sizes. RPL shows a linear increase in control messages with network size while LOAD and Geographical routing show an exponential increase. This is suspected to be the cause of the increased end-end delay and lower reliability statistics for LOAD and Geographical routing.

2.5 Summary

In this chapter, we described the wireless channel conditions in which smart utility networks are typically deployed. We modeled the path loss and fading in the wireless channel and introduced a temporal component to model a dynamic wireless channel. We also described the wireless mesh architecture in which smart meters are typically deployed. Further, we describe the routing protocol mechanisms of RPL, LOAD and Geographical routing and their implementation in the simulation framework of OMNeT++ for performance analysis. The simulation parameters were provided and the results from the simulation of a large scale smart utility network were presented. The results show RPL performing better than Geographical routing and LOAD in terms of hop count, delay, packet delivery ratio and low control overhead. Thus, we choose RPL as the most applicable routing protocol to be deployed in smart utility networks under the conditions simulated.

Chapter 3

Scalability of routing protocols in Smart Utility Networks

Scalability of routing protocols is an important problem in wireless mesh networks [19] [20]. The large scale deployment of smart utility networks poses challenges in determining the maximum capacity and scale which each collector or root node can support [21]. Very little field deployment data is available on the maximum number of nodes supported by each collector. In this work, we attempt to statistically model the smart utility network in order to determine a theoretical limit on the maximum number of nodes supported by each collector. In addition, we verify these results with extensive simulations of large scale networks. Here, we use two distinct approaches to model the smart utility network. One is based on the idea that the network size is limited by the packet success probability of each link in the mesh network. The other analyzes the mesh network as a single server queuing system in which the packet arrival process at the collector is according to a Markov Modulated Poisson Process, service is exponential and the buffer size is limited.

3.1 Approach I (Poisson distributed Connectivity Model)

This approach models the smart utility network as spatially distributed according to a Poisson distribution [22]. Figure 3.1 shows the smart utility network comprised of a collector at the center of a geographical area of radius *R*. Each node in the network has coverage range of radius r. This radius r is derived from eq. (2.1) from Chapter 2 section 2.1. The transmission radius is dependent on the radio transmission power and the wireless channel characteristics. Since the wireless channel model has been validated by comparing with real world experiments in Chapter 2, the r thus calculated provides realistic values for the transmission range of each node in the smart utility network. The total number of nodes in the network is *N*. Typically, the collector node has a higher transmission power than the participating nodes in real deployments. For the sake of simplicity in this analysis, we assume that all the nodes including the collector transmit at the same power, thus having the same transmission range *r*. In addition we define the density parameter ρ calculated by eq. (3.1). The parameter ρ_d is the deployment density and can vary between densely and sparsely deployed smart utility networks. The nodes are connected to each other and the collector at degree *k* by a Poisson distribution whose parameter is n_f. Here, n_f is the number of forwarding neighbors available to each node at the k^{th} hop.

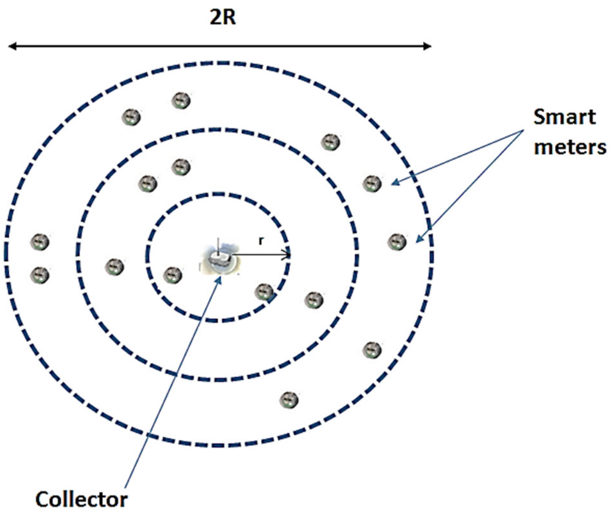


Figure 3.1: Spatial distribution of nodes for Scalability model

Thus the probability of finding a node at k^{th} ring from the collector is given by eq. (3.2). Consequently, the probability that a node is detached from the network is at k=0, given by eq. (3.3).

$$\rho = \frac{E[n_{cell}]}{N} = \frac{\rho_d.\pi .r^2}{\rho_d.\pi .R^2} = \left(\frac{r}{R}\right)^2 \tag{3.1}$$

$$P\left(k; \lambda = n_f\right) = \frac{e^{-n_f}.n_f^k}{k!} \tag{3.2}$$

$$p_{detach} = P\left(k = 0;\ n_f\right) = e^{-n_f} \tag{3.3}$$

Let the packet success probability for each link be *u*. The probability of a successful transmission over *n* hops is given by eq. (3.4). The expectation of successful transmission of packets over *k* hops according to the connectivity model in eq. (3.1) is given by eq. (3.5). Here, the parameter n_{cell} is the number of forwarding neighbors found in the ring of the k^{th} hop.

$$p_f = u^n \tag{3.4}$$

$$E\left[p_f\right] = \textstyle\sum_{k=1}^{k=n} \frac{e^{-n_{cell}}.n_{cell}^k}{k!}\left(u^k\right) \tag{3.5}$$

$$n_{cell} = \rho N$$

Limit of E[p_f] as $n \rightarrow \infty$ yields,

$$E\left[p_f\right] \approx e^{-\rho N}\left(e^{u\rho N} - 1\right) \tag{3.6}$$

Let the differential of *E[p_f]* w.r.t. *N* be zero to find maximum;

$$\frac{dE\left[p_f\right]}{dN} = 0$$

$$\rho N = \frac{1}{u} \, . \, ln\left(\frac{1}{1-u}\right) \tag{3.7}$$

As n → ∞ , the expectation of successful packet forwarding can be approximated to eq. (3.6). To find the number of nodes *N* at which we get the maximum expected successful packet transmission, we let the differential of *E[p_f]* with respect to *N* be zero. It can be easily verified that a maximum exists as the second order differential of *E[p_f]* with respect to *N* is negative. Substituting the result of eq. (3.7) in eq. (3.6) for the maximum expected successful packet transmissions yields eq. (3.8). Here, we have a relationship between the maximum expected successful packet transmissions and the link packet success probability in the network for a network of *N* nodes. Figure 3.2 shows the relationship of *E[p_f]* and *u.*

$$E\left[p_f\right] = e^{-\left[\frac{1}{u}.ln\left(\frac{1}{1-u}\right)\right]}\left(e^{\left[ln\left(\frac{1}{1-u}\right)\right]} - 1\right) \tag{3.8}$$

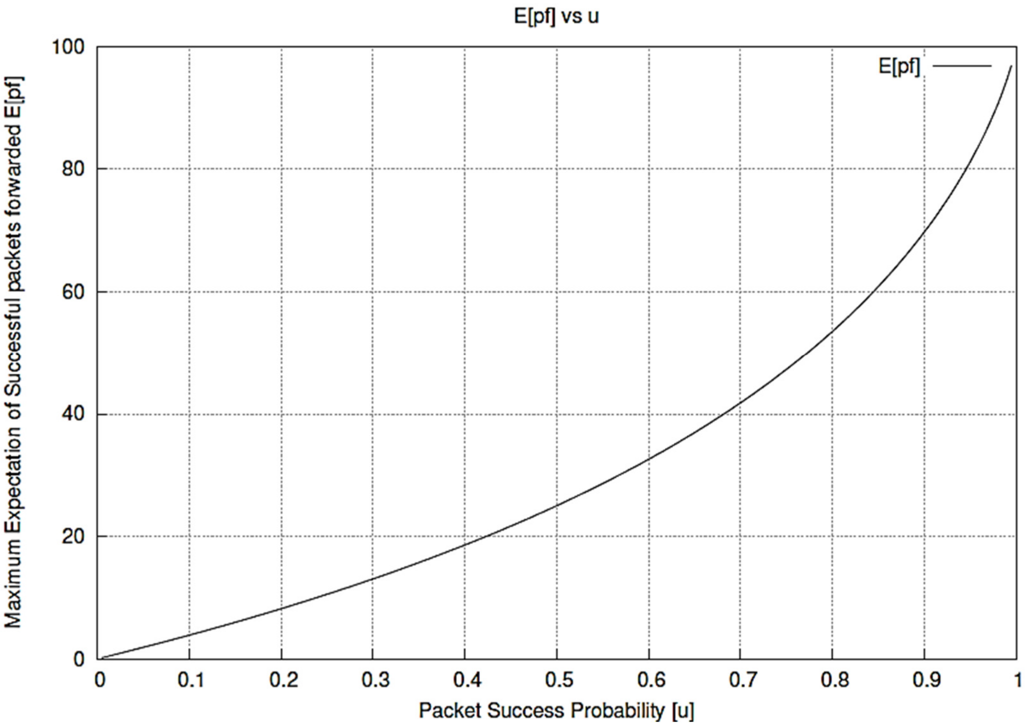


Figure 3.2: E[p_f] vs. u

Figure 3.2 gives a general relationship, allowing for geographical size *R* and transmission range *r* as variable parameters to determine the value of *u* at which we can obtain a maximum expected successful packet transmission for network size *N*. Thus we can determine the scalability in smart utility networks by simulating different network sizes and determine the average link packet success probability yielded by each network size N. Figures 3.3 and 3.4 show the various theoretical values of *N* as a function of the link packet success probability as calculated using eq. (3.7). This is the value of *N* which will yield the maximum *E[p_f]*. In figure 3.3, we vary the transmission range *r* from r=100m to r=450m in steps of 50m and obtain the corresponding values of *N* against the link packet success probability *u*. In practice, this can be done by varying the transmission power of the node. In figure 3.4, we vary the geographical area *R*, keeping the transmission range *r* constant at r=450m. This yields a new set of values of *N* against *u* and by utilizing this data, we can calculate *N* for maximum expected successful packet transmissions as in figure 3.2. In order to verify these theoretical results, we ran simulation scenarios with transmission range set at *r=100m* and geographical area *R=3536m*. Average link success probabilities (*u*) were calculated for varying network size using the same radio and wireless environments described in Chapter 2 section 2.1.

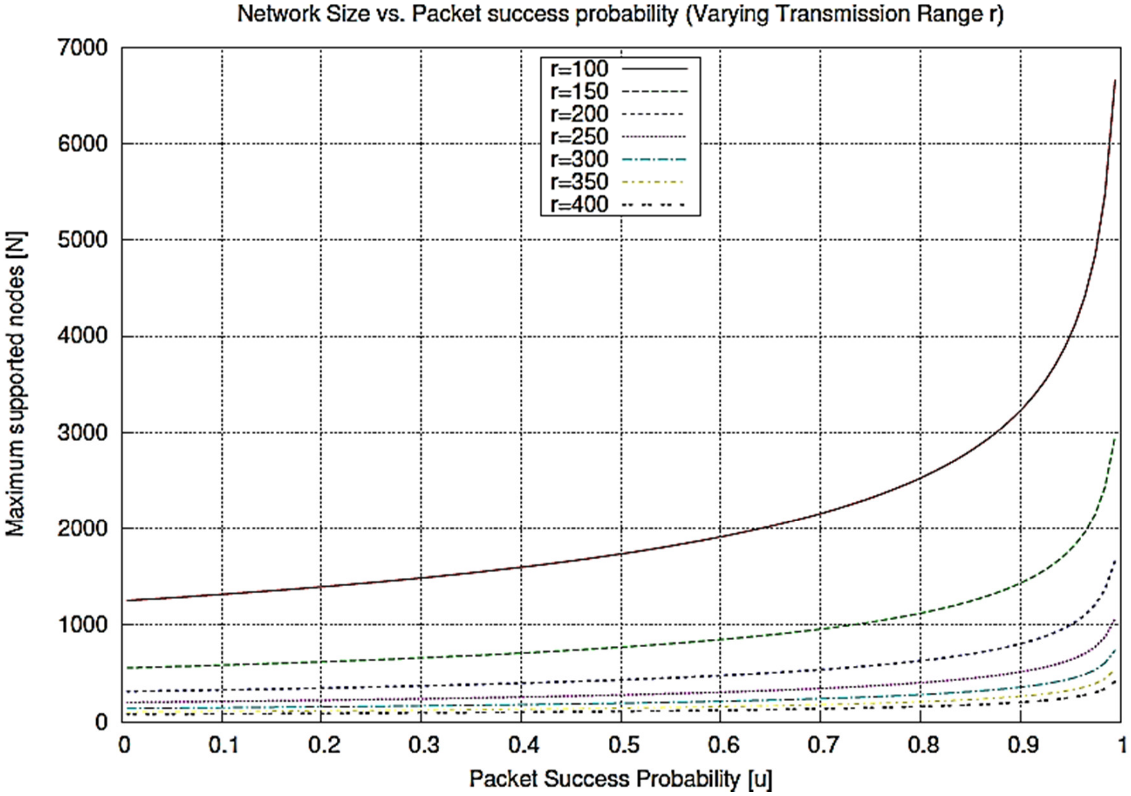


Figure 3.3: N vs. u for varying r

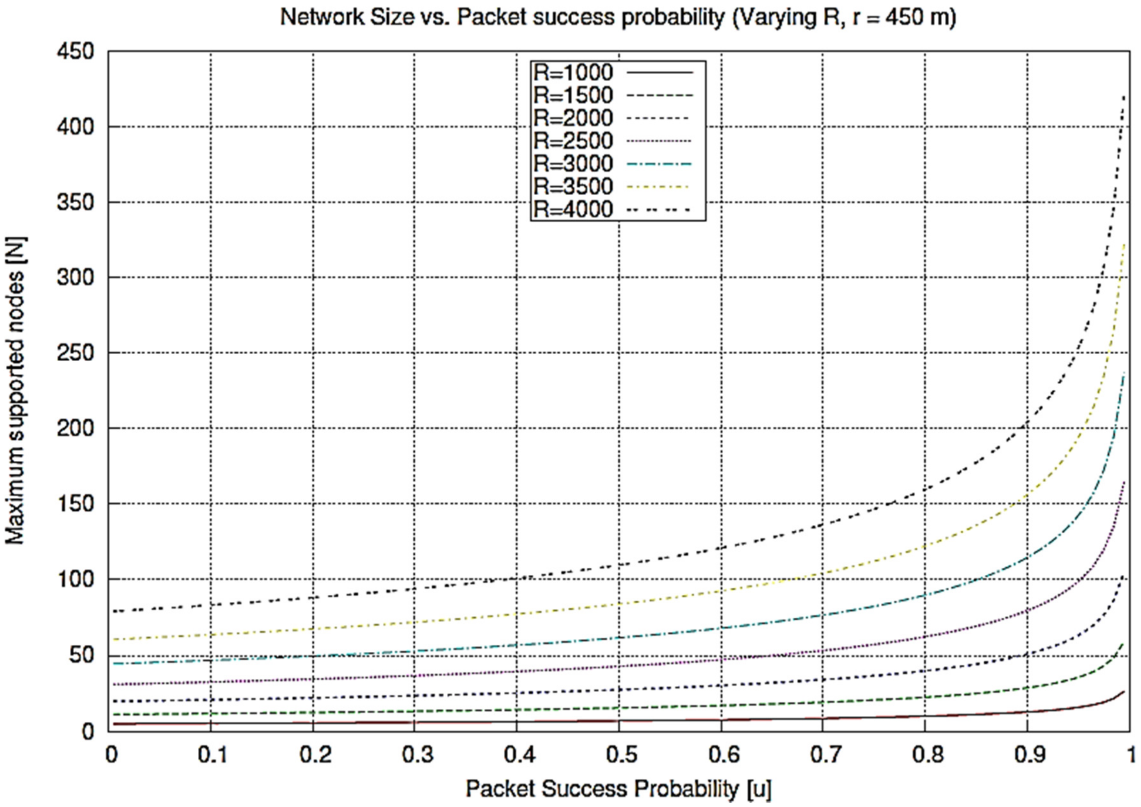


Figure 3.4: N vs. u for varying R

Figure 3.5 shows the maximum *E[p_f]* versus *u* for varying network sizes *N* both calculated and simulated. The values are scaled to fit in a multiple y-axis grid. The graph can be read by choosing a point on the black (calculated) or blue (simulated), projecting it on the x-axis for a value of *u*. Further, we can project this value of *u* onto the red (maximum *E[p_f]*) curve to get a maximum expected successful packet transmission for the network size *N*. The simulated results were obtained for a maximum network size of 5000 nodes before terminating. The difficulties and challenges in simulating a network of size beyond 5000 nodes for this experiment is outlined in chapter 4.

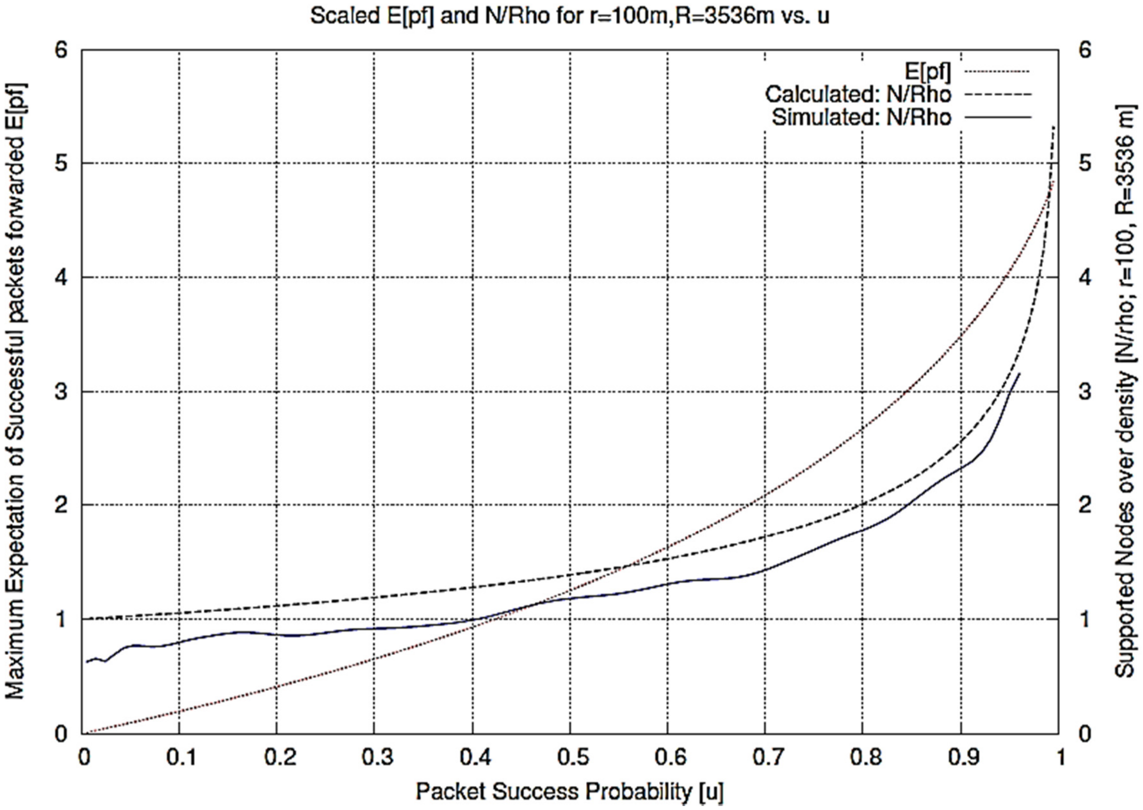


Figure 3.5: Calculated and Simulated N and E[p_f] vs. u (Scaled)

3.2 Approach II (Markov Modulated Poisson Process Traffic Model)

Approach II is based on the idea that, given ideal wireless channel conditions, the scale of the smart utility network is limited by the traffic. In this approach, we model the network as a Single Server Queue (SSQ) [23], the collector node or the root being the server. The Poisson arrival and service process is exponentially distributed. This allows for simplification in calculations due to the fact that we can add the departure rates at each node and generate a cumulative arrival process at the server which is also exponentially distributed [24]. In smart utility networks, the traffic is of two distinct types. One is the metering data which is periodically transmitted to the collector as per the requirements of the utility demand-response applications [25]. This can be modeled as Constant-Bit-Rate (CBR) traffic. The second type of traffic is the emergency data during outages and anomalous power fluctuations. This type of traffic takes precedence over the metering data. Thus we can model the two types of traffic as two modes of a Markov Modulated Poisson arrival Process (MMPP) [26].

In our model the MMPP has two modes m_0 and m_1 with different arrival rates λ_0 and λ_1. The MMPP(2)/M/1/N queue with buffer size *N,* has mode parameters δ_0 and δ_1 for modes m_0 and m_1. The server has an exponentially distributed service time μ. The service times are mutually independent and independent of the arrival process. Since, the inter-arrival times are not independent in this queue, it affects queuing performance, packet loss and utilization. The queue process is a continuous-time Markov chain but its states are two dimensional non-scalar vectors. Each state has two scalar quantities; the mode *m* and the queue size. Let p_{im} for i=0, 1, 2… *N* be the probability that the arrival process is in mode *m* and that there are *i* packets in the system. By obtaining the values of π_{im} the steady-state queue probabilities can be obtained by using eq. (3.9).

$$\pi_{\mathrm{i}} = \pi_{\mathrm{i0}} + \pi_{\mathrm{i1}} \qquad \text{for i} = 0,1,2,\ldots\mathrm{N} \tag{3.9}$$

The probability of arrival in mode *m* can be calculated by solving the set of equations (3.10) and (3.11). Solving these equations gives the steady-state probabilities of arriving in mode *m* as a function of mode parameters δ_0 and δ_1 shown in equations (3.12) and (3.13).

$$P(m=0)\delta_0 = P(m=1)\delta_1 \tag{3.10}$$

$$P(m=0) + P(m=1) = 1 \tag{3.11}$$

$$P(m=0) = \frac{\delta_1}{\delta_0+\delta_1} \tag{3.12}$$

$$P(m=1) = \frac{\delta_0}{\delta_0+\delta_1} \tag{3.13}$$

The probability of the arrival process to be in mode *m,* is equal to eq. (3.14). This follows from equations (3.12) and (3.13).

$$\sum_{i=0}^{i=N} \pi_{\mathrm{im}} = \frac{\delta_{1\text{-m}}}{\delta_{1\text{-m}}+\delta_{\mathrm{m}}} \quad \text{for m=0,1} \tag{3.14}$$

Thus, the average arrival rate of both the modes is λ_{av} given by eq. (3.15). Also, we let utilization of the combined arrival process as ρ^* and is given by eq. (3.16).

$$\lambda_{\mathrm{av}} = P(m = 0)\lambda_0 + P(m = 1)\lambda_1 \tag{3.15}$$

$$= \frac{\delta_1}{\delta_0 + \delta_1}\lambda_0 + \frac{\delta_0}{\delta_0 + \delta_1}\lambda_1$$

$$\rho^* = \frac{\lambda_{\mathrm{av}}}{\mu} \tag{3.16}$$

We also define a parameter ψ called the mode duration parameter. In general, $\delta_{\mathrm{m}} = \psi\, \delta^*_{\mathrm{m.}}$ The mode duration parameter ψ can be varied according to how bursty the traffic is. For example, if we receive frequent outage and emergency information from the meters, the value of ψ can be large. Lower value of ψ indicates that the traffic in each mode is for a long duration or less bursty traffic.

The values of π_{im} can be calculated using the finite set of steady state equations given by eq. (3.17). Q is the infinitesimal generator whose size is 2Nx2N, where N is the size of the buffer. Non-zero entries of Q for [i,j] two-dimensional vectors are given herein.

$$0 = \Pi Q \qquad (3.17)$$

Here, the generator matrices Q and Π are given by equations (3.18) and (3.19).

$$\Pi = \left[\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}, \pi_{20}, \pi_{21}, \dots, \pi_{N-1,1}, \pi_{N0}, \pi_{N1}\right] \qquad (3.18)$$

$\boldsymbol{Q} = \left[Q_{i,j}\right]$; i,j are two-dimensional vectors ; Size of Q is 2Nx2N (3.19)

The non-zero entries of Q for buffer size $N>i>0$ are as follows:

$$Q_{i0,i0} = -\lambda_0 - \delta_0 - \mu\,; \quad Q_{i0,i1} = \delta_0\,; \quad Q_{i0,(i+1,0)} = \lambda_0$$

$$Q_{i1,i0} = \delta_1\,; \quad Q_{01,01} = -\lambda_1 - \delta_1 - \mu\,; \quad Q_{i1,(i+1,1)} = \lambda_1$$

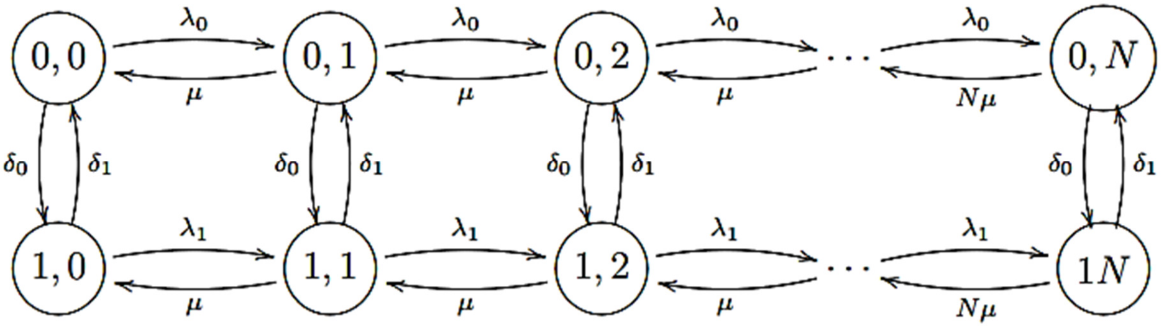


Figure 3.6: State Transition diagram for the modeled MMPP(2)/M/1/N queue

Figure 3.6 gives the state transition diagram for the above modeled MMPP(2)/M/1/N system. The normalizing equation (3.20) and the steady-state equation (3.17) need to be solved in order to obtain steady state probabilities. We use the Gauss-Seidel method of successive approximations or the iterative method to derive these values. The blocking probability, *P*$_b$ can be obtained using equation (3.21).

$$\sum_{i=0}^{i=N} \sum_{m=0}^{m=1} \pi_{im} = 1 \qquad (3.20)$$

$$P_b = \frac{\lambda_0 \pi_{N0} + \lambda_1 \pi_{N1}}{\lambda_{av}} \qquad (3.21)$$

Having constructed the model, we aim to simulate the smart utility network as the modeled queuing system to calculate the blocking probability for different buffer sizes *N* and network sizes. The goal is to find a maximum network size supported by the collector for the smart utility network traffic arrival rate at which blocking probability approaches the value of 1.

3.2.1 Raw data for 2-state MMPP model

Using the validated OMNeT++ simulation model described in Chapter 2, packet arrival rates are collected for varying network sizes from 20 nodes to 6000 nodes at a single collector node. Each network size is simulated for a time of 100 seconds and the number of packets arriving at the collector per second is recorded. This will allow us to approximate the packet arrival rate λ at the collector for each network size. Subsequently we use the LAMBDA algorithm presented in [27] [28] to fit the raw data into the MMPP modeled herein.

In addition to the fitted arrival rates, the LAMBDA algorithm gives the number of states in the MMPP model that the traffic can be classified into. To this effect, we generate two types of traffic while generating data. Periodic metering information sent by each node and random outage information packets. We estimate this type of traffic in smart utility networks to fit into the two-state MMPP model. Figures 3.7, 3.8 and 3.9 show the raw data obtained from varying network sizes of 20 nodes, 200 nodes and 6000 nodes respectively to indicate the scale of data collected form the simulator.

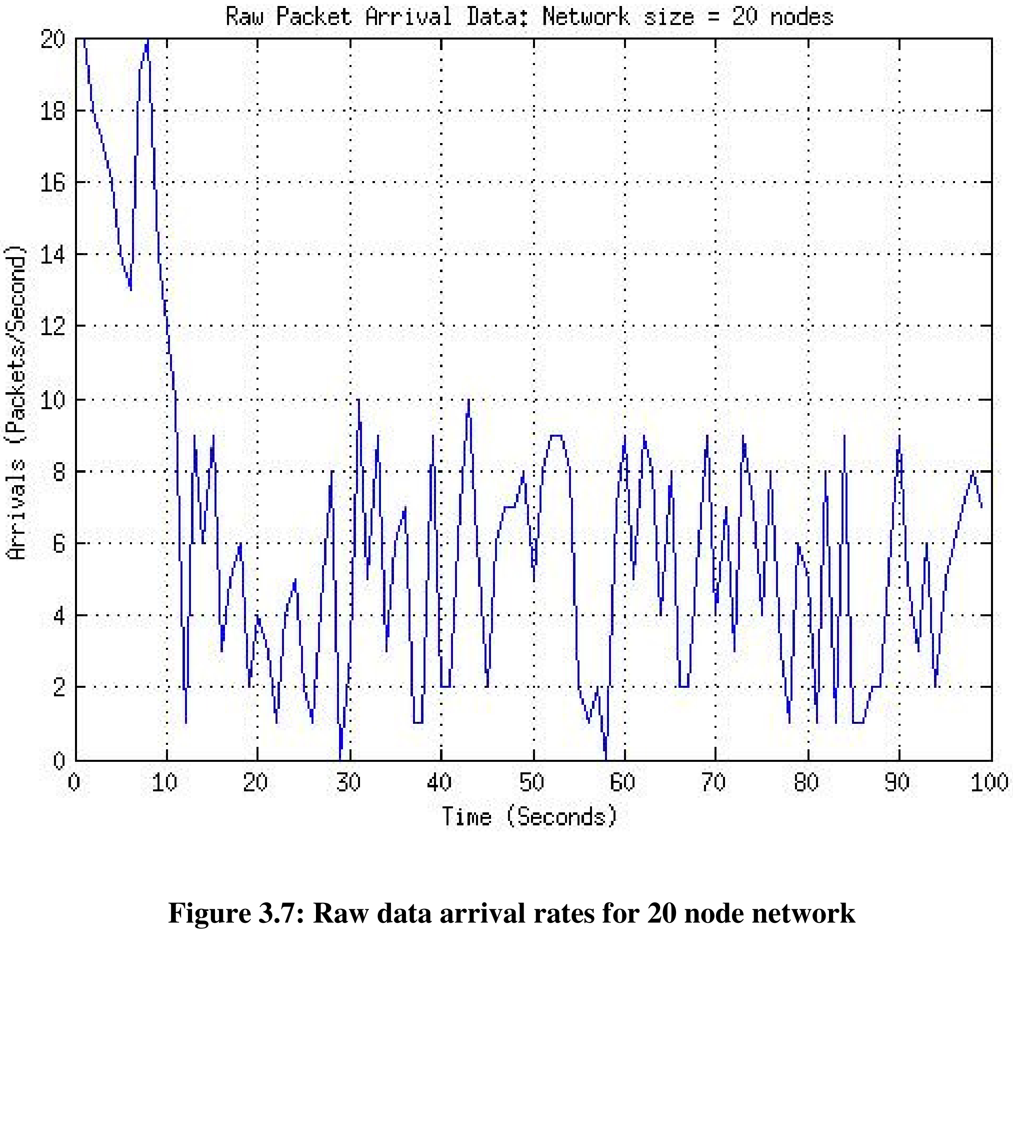


Figure 3.7: Raw data arrival rates for 20 node network

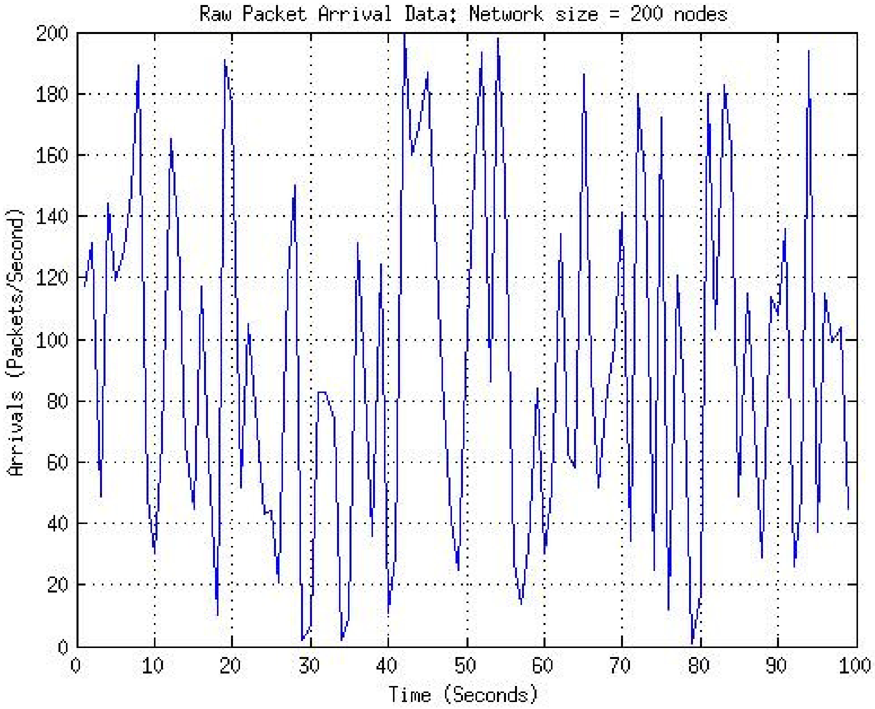


Figure 3.8: Raw data arrival rates for 200 node network

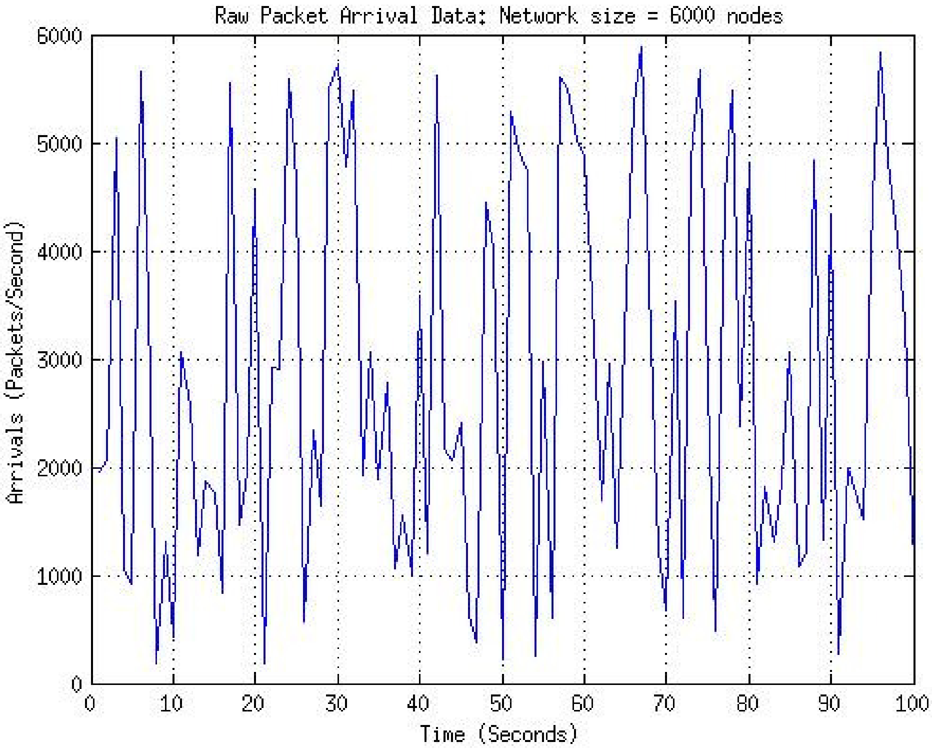


Figure 3.9: Raw data arrival rates for 6000 node network

3.2.2 Fitting raw data to MMPP model

An important application of queuing models in smart utility networks is to size the number of links and buffer sizes at routers. The fundamental part of the queuing model is the arrival process. We have collected raw data from the simulation in order to fit the raw data to the modeled arrival process. Using this data, we present the case that smart meters generate packet arrivals in the form of an MMPP. Moreover, the nature of traffic is closely related to an Interrupted Poisson Process (IPP) in which the arrival rate of one type of traffic is abruptly made 0 when the other type of traffic is active. This accurately models the traffic generated by smart meters.

When an emergency or outage information packet arrives, the collector stop processing the periodic meter reading traffic and prioritizes the outage packet. The arrivals pass through a constant-rate device, or a collector. When the smart meter data generation rate exceeds the rate of the collector, we assume that the excess is either buffered or retransmitted after a short delay. Then the output from the collector is a modification of the original MMPP with a peak rate equal to the rate of the collector. We present a method to capture this effect and develop equations to describe this rate-limited process. Fig. 3.10 shows the different arrival rates and switching of processes by the collector.

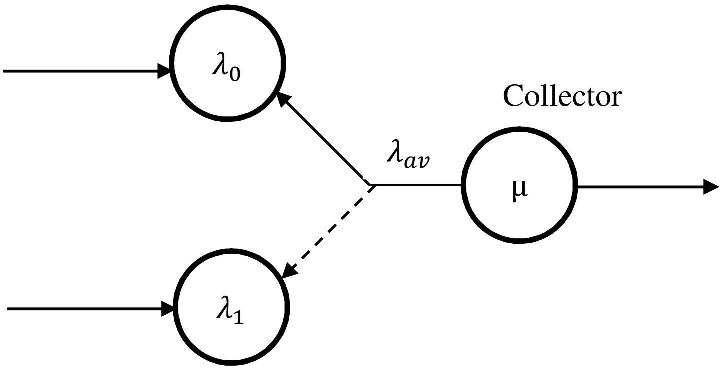


Figure 3.10: Arrival Process at Collector from metering nodes

In [29] the first proposal to fit a traffic stream by an MMPP has been proposed. It used two phases for the Markov chain, gave an estimation procedure for the parameters of the MMPP from count data, and solved several queuing models with exponential service times. Meier-Hellstern [30] provided the first estimation algorithm based on inter-arrival time measurements. It has proven difficult to obtain estimators for MMPPs with more than a few states because of the computational burden. Moreover, traffic data frequently consists of counts of events during fixed length intervals, such as packets per second. Hence, model fitting has to be done with this type of data.

The IPP is the simplest non-degenerate special case of an MMPP. There are two states, and the arrival rate is zero in one of them. In [31], they have developed a maximum likelihood estimation algorithm for MMPPs based on the superposition of IPPs. Their algorithm requires that the arrival times are given in addition to the counts. Lee *et al.* [32] describe the properties of a generalized IPP in which the times between phase transitions need not be exponential. They focus on the autocorrelations in the times between arrival epochs. Since only one positive Poisson rate is used, we do not think this model will fit the data we use.

Since the variance of the Poisson distribution equals the mean (λ), and for large λ the Poisson and Normal distribution functions are close, most of *$P(\lambda)$* will be within $\lambda \pm a\sqrt{\lambda}$ with a=2. When the peak-to-mean ratio of the data set is large, the two rates in states one and two, λ_0 and λ_1, with $\lambda_0 > \lambda_1$,will be very different. Then $\lambda_1 + 2\sqrt{\lambda_1}$ will be much less than $\lambda_1 - 2\sqrt{\lambda_1}$ and traffic with rates between these values will not be described. Hence, there is a need to fit the raw data into two discrete states representing different rates of arrivals to ensure that all the values will be described in the model. Figures 3.11, 3.12, 3.13 and 3.14 show the raw data fitted to the MMPP model for varying network sizes of 20 nodes, 200 nodes and 6000 nodes respectively.

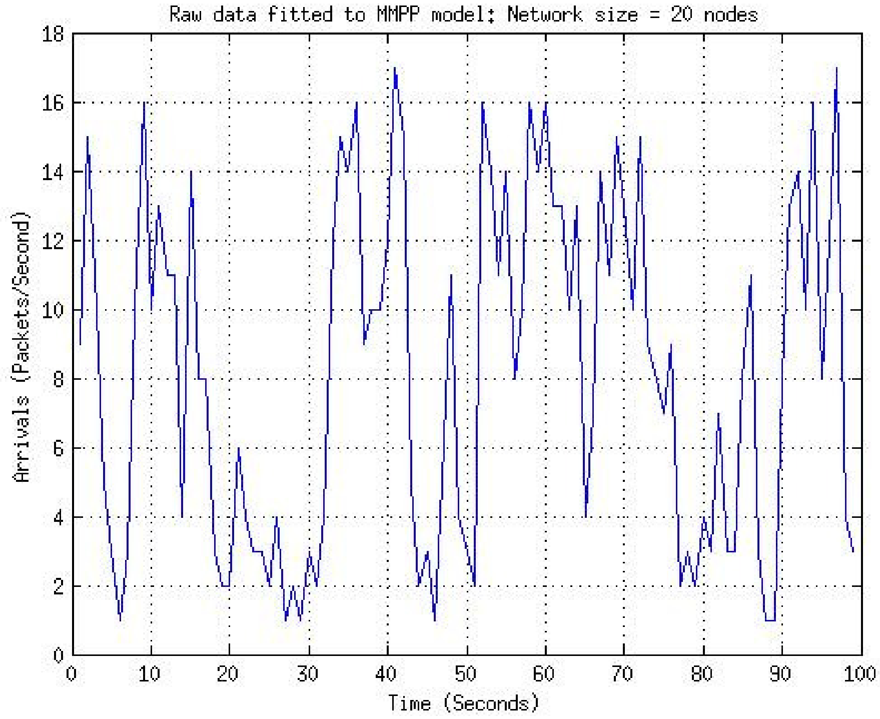


Figure 3.11: Raw data fitted to MMPP process for a 20 node network

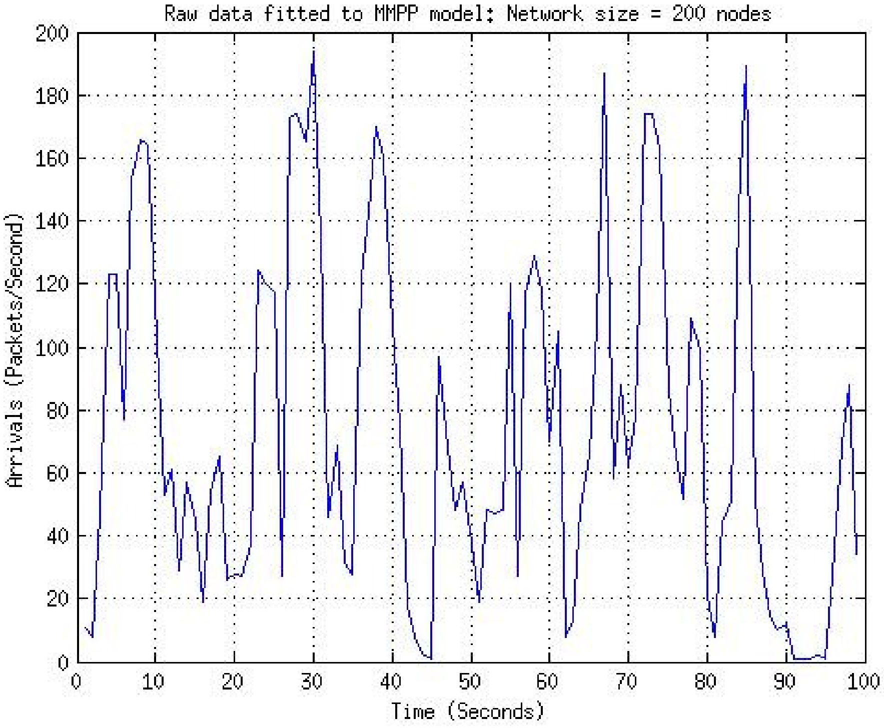


Figure 3.12: Raw data fitted to MMPP process for a 200 node network

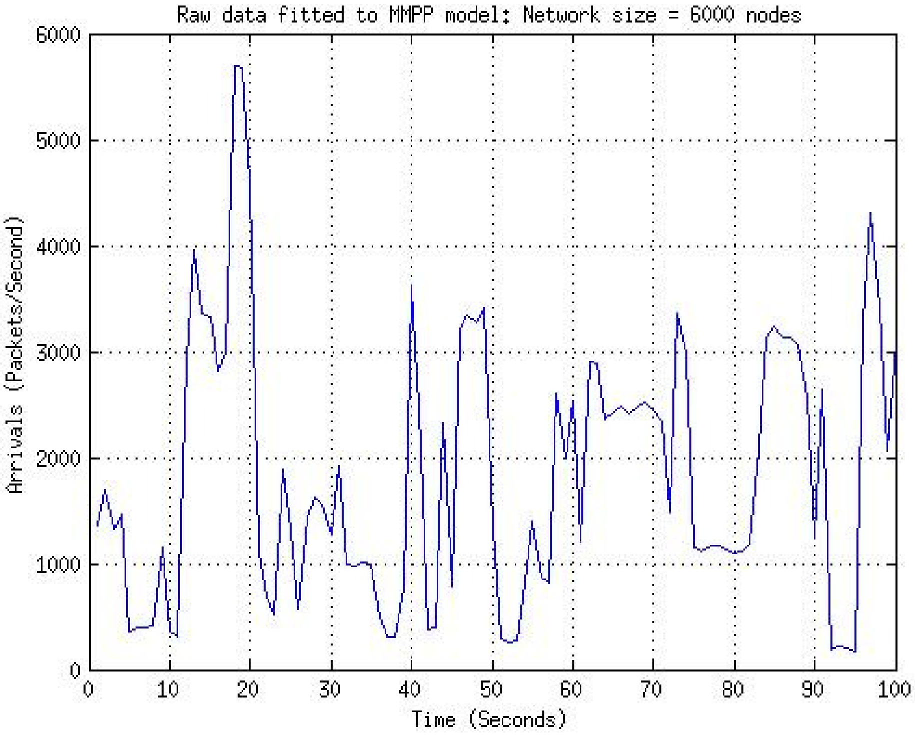


Figure 3.13: Raw data fitted to MMPP process for a 6000 node network

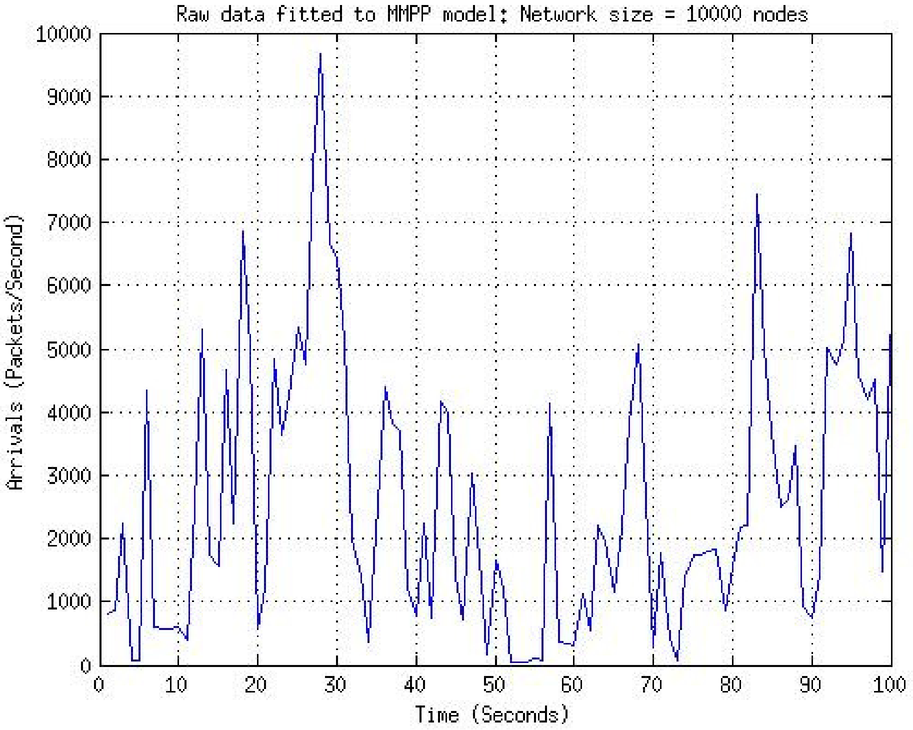


Figure 3.14: Raw data fitted to MMPP process for a 10000 node network

3.2.3 LAMBDA algorithm to determine average rates λ_{av} and states N

The LAMDA algorithm presented in [26] is originally intended for IP traffic at high data rates. The algorithm takes raw data as packet arrivals per second and estimates the average rate λ_{av} and the number of MMPP states that the data fits into. We utilize this algorithm shown in Table 3.4.1 to estimate the number of MMPP states that the data from smart metering simulation trace can be fit into. We expected the algorithm to generate two distinct states, according to the two types of packets at varying arrival rates of λ_0 *and* λ_1. As the starting point of the LAMBDA algorithm, we determine the initial rate to be the rate which covers the largest observations. Equation (3.22) is used to determine this rate at as the upper bound. Observations larger than the initial rate are calculated using Equation (3.23), which is the solution to the quadratic in λ_0. The lower bound of λ_0 is chosen as the upper bound of λ_1 in equation 3.24. The solution to the quadratic in λ_1 is given by equation 3.25. The stopping points are data rates which are 0.

$$\lambda_0 + 2\sqrt{\lambda_0} = Peak\ of\ data \qquad (3.22)$$

$$\lambda_0 = \left(\sqrt{1 + Peak} - 1\right)^2 \qquad (3.23)$$

$$\lambda_1 + 2\sqrt{\lambda_1} = \lambda_0 - 2\sqrt{\lambda_0} \qquad (3.24)$$

$$\lambda_1 = \left(\sqrt{\lambda_0} - 2\right)^2 \qquad (3.25)$$

Table 3.1: LAMBDA algorithm

Algorithm: LAMBDA

```
0: Choose width parameter $a$ [default=2] and ending parameter $\Omega$ [default=0]
1: Choose $\lambda_1$ using Equation (3.22)
    a) Set $N=1$
    b) Set $r = \lambda - a\sqrt{\lambda}$
    c) If $r \leq \Omega$, Stop. Else go to Step 2.
2: Set $j = N+1$
    a) Set $\lambda_j = \left(\sqrt{\lambda_{j-1}} - a\right)^2$
    b) Set $r = \lambda_j - a\sqrt{\lambda_j}$
    c) Set $N = j$
    d) If $r \leq \Omega$, Stop. Else repeat Step 2.
```

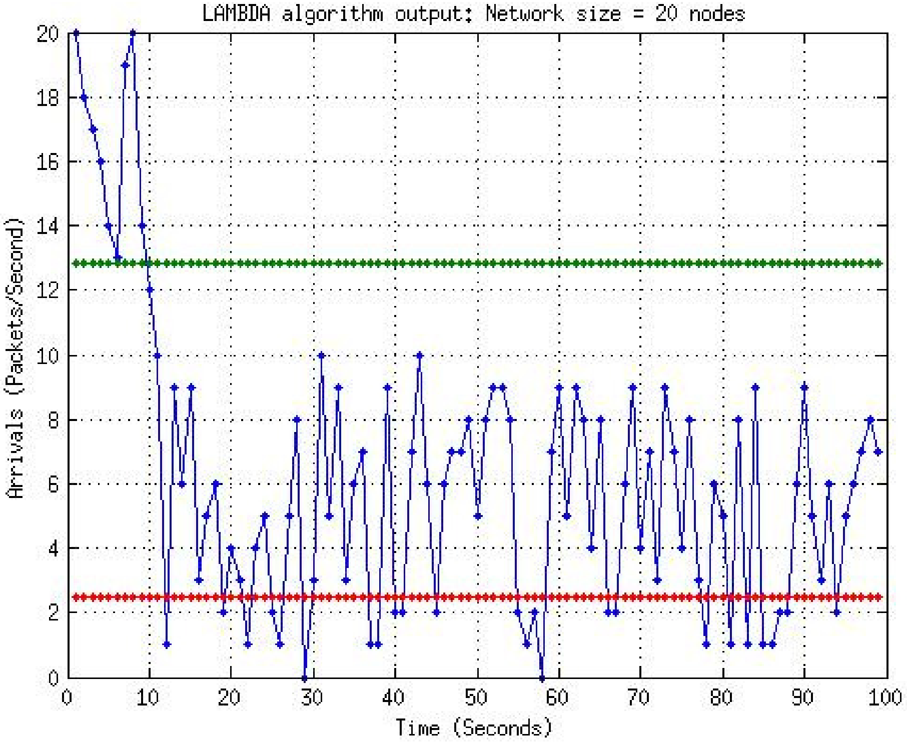


Figure 3.15: Fitted data showing 2 state MMPP process for 20 node network

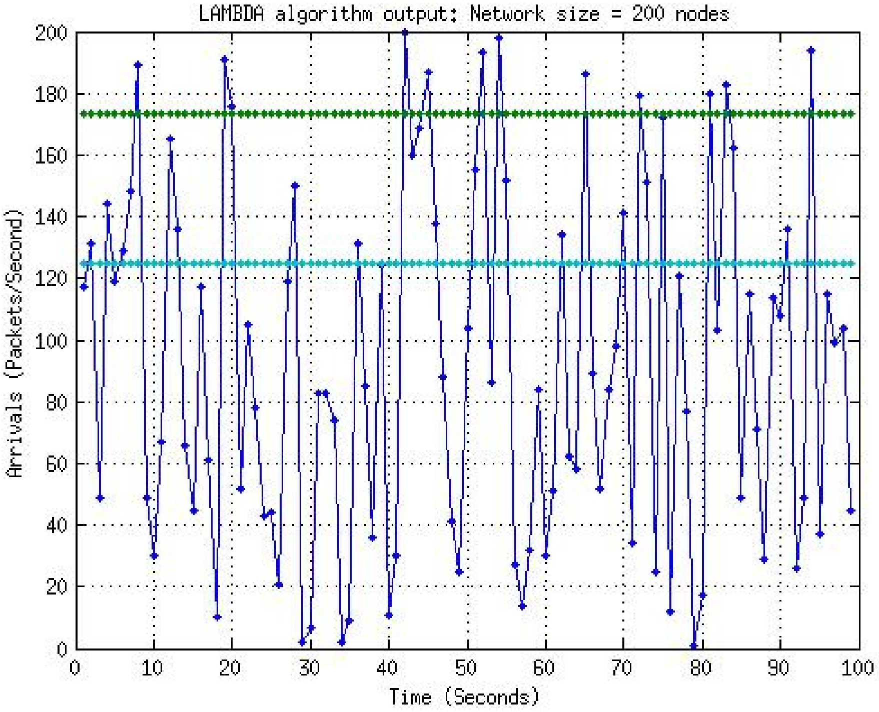


Figure 3.16: Fitted data showing 2 state MMPP process for 200 node network

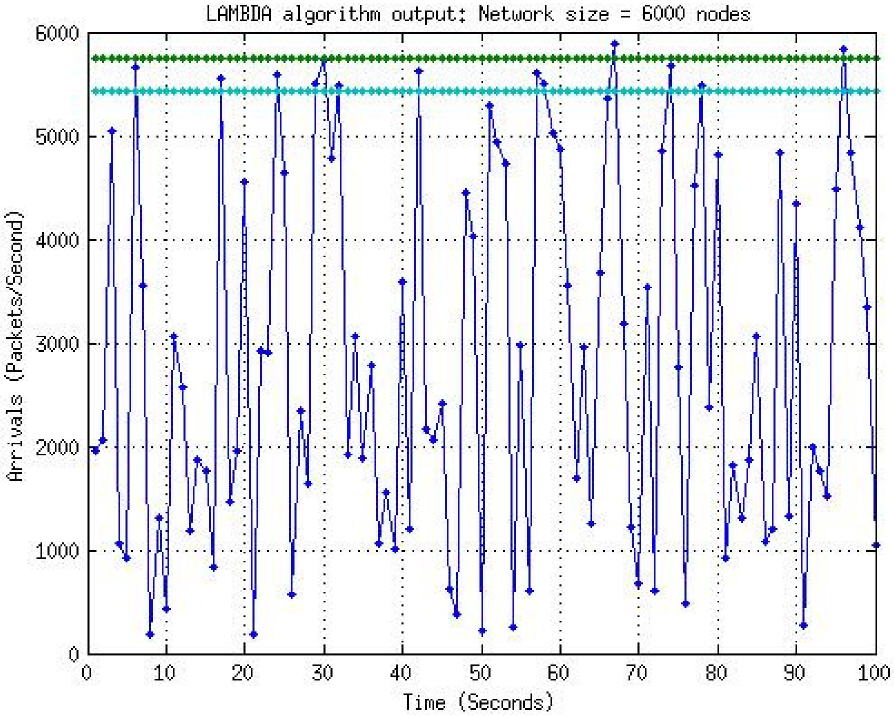


Figure 3.17: Fitted data showing 2 state MMPP process for 6000 node network

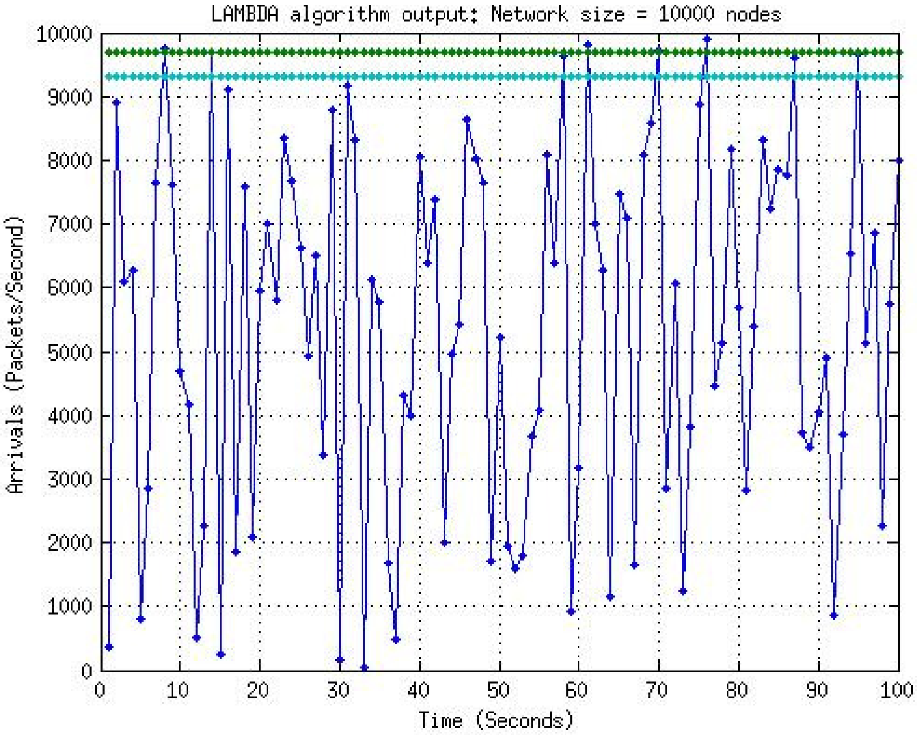


Figure 3.18: Fitted data showing 2 state MMPP process for 10000 node network

Figures 3.14, 3.15, 3.16 and 3.17 show the raw data fitted to the MMPP model with the two distinct states for data rates λ_0 *and* λ_1 as computed by the LAMBDA algorithm.

3.2.4 Fitting data to Markov Chain

Using the LAMBDA algorithm, we have generated the number of states N to fit the data. Here, we have 2 states with rates λ_0 *and* λ_1. We construct the state transition matrix in order to get the steady state probabilities in those states. Let $\{x_i\ ,\ i = 1,2\}$ be the observation states. Let θ_i be the state associated with x_i. The range of x_i for which θ_i is defined, can be given by equation (3.26) as defined by the upper and lower bounds.

$$\lambda_j - 2\sqrt{\lambda_j} < x_i \le \lambda_j + 2\sqrt{\lambda_j} \rightarrow \theta_i = j \tag{3.26}$$

Let $P = (p_{ij})$ be the state transition matrix for the process. p_{ij}. The Maximum Likelihood Estimate (MLE) [33] [34] of p_{ij} is given by equation (3.27).

$$p_{ij} = \frac{Number\ of\ Transitions\ from\ i\ to\ j}{Number\ of\ transitions\ out\ of\ i} \tag{3.27}$$

3.2.5 Mean queue length

Let W_v and W_a be the virtual and waiting time seen by the arrival process respectively. If the mean service time μ is 1, then from the results of [35] [36] we have expected waiting times as equations (3.28) and (3.29), where Q_0 and Q_1 are infinitesimal generator matrices for states with arrival rates λ_0 and λ_1 respectively as defined by equation (3.19).

$$E(W_v) = (3\rho^* - 2bQ_1e)/(2(1 - \rho^*)) \tag{3.28}$$

$$E(W_a) = 1 - \frac{bQ_1e}{\rho^*} + E(W_v) \tag{3.29}$$

$$where, b = \left((1-\rho^*)g + \pi Q_1\right)(e\pi + Q_0 + Q_1)^{-1}$$

Here, g is the stationary probability vector of the irreducible stochastic matrix G, a unique solution to the matrix functional equation (3.30) and π is the stationary vector of the infinitesimal generator function $Q_0 + Q_1$. The starting point for calculating G by successive iteration is $G_0 = e\pi$.

$$G = e^{Q_0 + Q_1 G} \tag{3.30}$$

Using Little's formula [37], the mean virtual queue length at mean queue length seen at arrivals is given by equations (3.31) and (3.32).

$$E(Q_v) = \rho^* E(W_v) \tag{3.31}$$

$$E(Q_a) = \rho^* E(W_a) \tag{3.32}$$

3.3 Packet loss and Blocking

Using the Two-state MMPP fitted data and the model described above, we compute the mean delay and mean packet loss rate for the arrival process. The packet loss rate is given by equation (3.33). Figures 3.17 and 3.18 show the mean packet loss rate and mean delay respectively for the MMPP arrival process. Using thus computed λ_{av} and state transition probabilities for arrival rates of λ_0 and λ_1, we can estimate the blocking probabilities for traffic at varying network sizes. Figure 3.19 shows the blocking probabilities as computed using equation (3.21) for the arrival process at varying network sizes.

$$L_r(i) = (pkts(i) * pkt\_size) - \mu \tag{3.33}$$

Here, the loss rate at a given sample i, L_r is in bps and is obtained by the packets arriving at instant i times the packet size (pkt_size) less the service rate μ of the collector device.

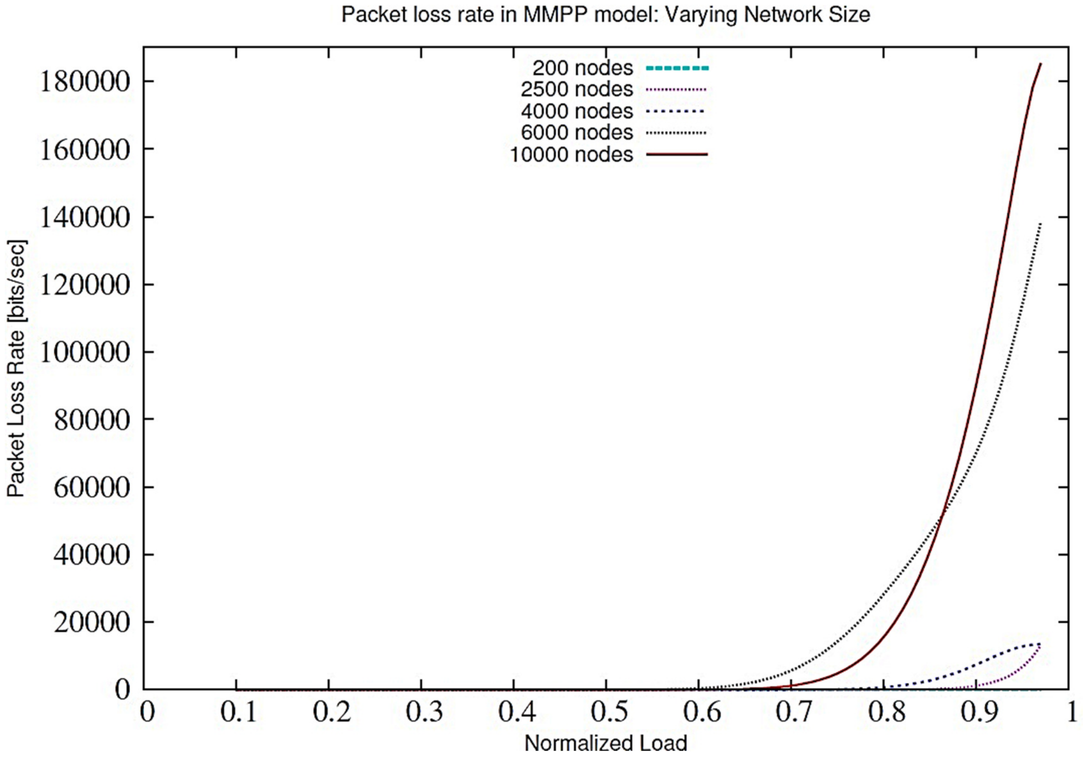


Figure 3.19: Mean packet loss rate for two-state MMPP process

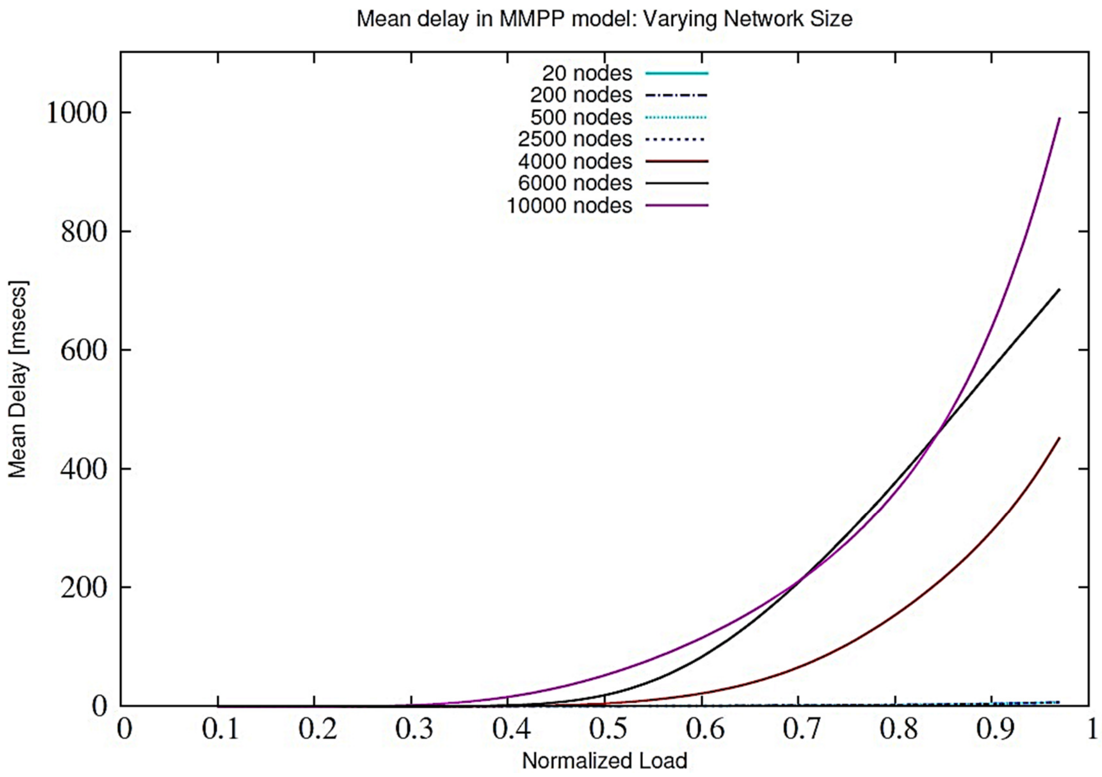


Figure 3.20: Mean delay for two-state MMPP process

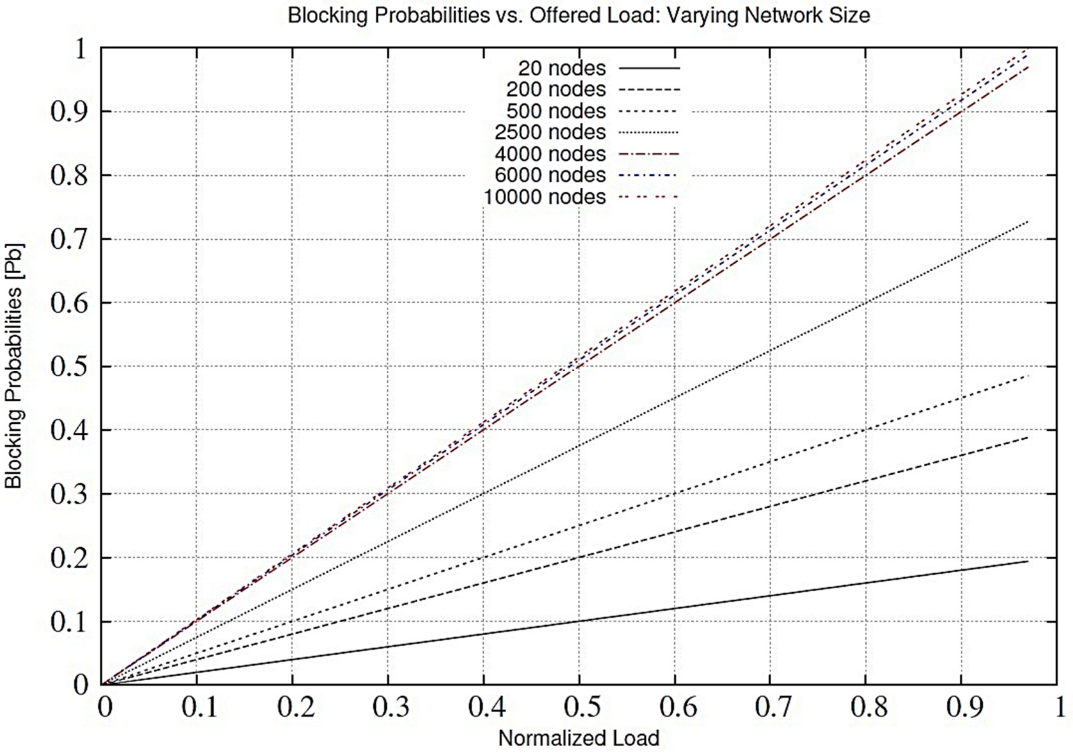


Figure 3.21: Blocking Probability $P_b$with varying network size

3.4 Interpreting MMPP model results

We can conclude the following from the results of fitting the raw data obtained from the smart utility network simulation data of varying network sizes from 20 nodes to 6000 nodes in terms of mean delay, mean packet loss rates and blocking probabilities:

- With offered loads more than 60%, the packet loss rate increases exponentially for network sizes above 2500 nodes.
- Mean delay increases exponentially with offered loads more than 50% for network sizes greater than 2500 nodes.
- The blocking probability is 0.5 for offered loads of 50% and increases linearly thereafter. For network sizes greater than 4000 nodes, the blocking probability approaches 1 at loads greater than 80%.
- In terms of scalability, we can draw a conclusion that deploying smart utility networks of size 4000 nodes per collector in any traffic conditions (offered load) is detrimental to the data acquisition performance of the entire advanced metering infrastructure.

3.5 Summary

In this chapter, we presented two models to address network scalability in smart utility networks. Both the models find an upper bound on the number of nodes that can be supported by each collector device. Approach I uses the link packet success probability and a Poisson connectivity model to measure scalability. We back this approach with simulated data from smart utility networks of up to 7500 nodes per collector. Approach II uses the two-state MMPP to model traffic in smart utility networks. The motivation stems from the two distinct types of traffic with distinct arrival rates in smart utility networks. The regular metering data and the random bursty outage information data are modeled well by the two-state MMPP process. We have observations which are conclusive that as the network size increases, critical performance metrics such as delay and packet loss rate increase exponentially and from the data, we can estimate the upper bound on the network size at which scalability fails, using the raw data arriving at the collector from the smart utility network.

Chapter 4

Cost of Security in Smart Utility Networks

A 2009 report stated that cyber-spies had penetrated into the North American power grid system and left behind malicious software, that could be triggered by them later-on during a time of crisis or war sounded alarm bells for utility companies whose power infrastructure has been proven to be vulnerable. Basic facilities such as electricity grid, water, gas, waste water management systems, traffic signaling systems etc. will be directly affected by a disruption caused to the power grids. Hence the vulnerabilities would imply that the purpose of the utilities serving households, commercial establishments and industry will be severely compromised [38]. At the heart of these vulnerabilities is the supervisory control and data acquisition (SCADA) system. These systems control, collect and analyze data from equipment which are employed largely in power generation stations, transmission and distribution sub-stations and utilities which deliver the power to households, commercial establishments and the industry [39].

4.1 NIST data on Cyber-Physical Security

The National Institute of Standards and Technology (NIST) have specified threats to computer related commerce in North America. Over 66% of these threats relate to SCADA systems, directly connected to power substations [39]. The risks involved in SCADA controllers are not very different from those that personal computers and commercial computer networks encounter. As manual labor is being replaced by automated control, equipment are being networked for control from a single SCADA system. To avoid the cost of private lines which carry control data, the internet is an attractive option to network such devices. This provides an opportunity for intrusion into these systems which control critical infrastructure [40]. Security in these systems can be built around well-established security standards for commercial networks. NIST has established a Smart Grid Cyber Security Coordination Task Group (CSCTG) which addresses and evaluates processes leading to comprehensive cyber security policies for smart grids. The types of risks assessed by the CSCTG are not limited to intentional attacks from disgruntled employees, terrorists and industrial espionage but also include;

- Complexity of the grid, leading to weak points, unintentional errors, openings to attackers.
- Inter-connection of networks, leading to cascading errors.
- Denial of service attacks resulting from complex communication networks.
- Breach of consumer privacy.

The group CSCTG defines cyber infrastructure and cyber security with respect to smart grids as

Cyber Infrastructure:

- Electronic information and communications systems and services and the information contained in these systems and services.
- Information and communications systems and services are composed of all hardware and software that process, store, and communicate information, or any combination of all of these elements. Processing includes the creation, access, modification, and destruction of information. Storage includes paper, magnetic, electronic, and all other media types.
- Communications include sharing and distribution of information. These include the following:
    - o computer systems; control systems (SCADA);

o Networks, such as the Internet; and cyber services (managed security services).

Cyber Security is defined as “The protection required to ensure confidentiality, integrity and availability of the electronic information communication system.” [41]

4.1.1 Risks to the Smart Grid

- Intrusion risks: This form of risk is not new to information networks. Since the smart grid will be based on an information network, hacking into AMI or utility databases is a huge threat. Intrusion vulnerability points include SCADA systems, modems used by components and other networking components such as smart meters and any device participating as a routing device in the Advanced Metering Infrastructure (AMI).
- Malware: Malware or malicious software can be found across the public internet in large numbers. The inter-connection of the smart grid network with the public internet will cause a potential spread of these malware into grid software, causing disruption in power services or false data being exchanged in communication links [42].
- Espionage and information warfare from state sponsored sources.

4.1.2 Mitigating risks in the Smart Grid

A number of steps can be taken to mitigate many of these vulnerabilities. A combination of many of them will provide defense-in-depth solutions that will be much superior to any single strategy or mitigation. We propose the following mitigation strategies:

- Restricting Smart Meter data: The AMI and the smart meters generate a significantly large amount of data even with reporting frequencies as low as 1 packet/4 hours. With over 50 million smart meters currently in deployment, the AMI generates data on par with some internet applications. By restricting the data frequency, we can not only improve the reliability of smart utility networks as evident from the results in Chapter 3, but also mitigate the risk of simple replay attacks and the more complex man-in-the-middle type attacks popular in IP networks.

- Strong Encryption: Using state-of-the-art cryptographic standards popularly deployed in e-commerce applications for smart metering data will mitigate simplistic attacks on intercepting data. Further in this chapter, we will describe two public-key cryptographic schemes and the trade-offs in adding security in smart utility networks.
- Network Isolation: Although this mitigation strategy defeats the purpose of having an Internet Of Things (IoT) architecture, in which all smart metering devices will be accessible on the IPv6 world wide web [43], it is prudent to isolate the communication network of the AMI with the global internet until proven and effective security mechanisms are in place to protect the smart grid infrastructure.

4.2 Public-Key Cryptographic Schemes

Public key cryptographic schemes have been widely deployed in the Internet for a variety of applications ranging from e-commerce to social networks. The two popular public-key cryptographic schemes of RSA (Rivest-Shamir-Adleman) [44] [45] and ECC (Elliptic Curve Cryptography) [46] [47] have been widely used. In this chapter we will describe these schemes in brief and analyze the performance of these cryptographic schemes in smart utility networks and low-power wireless devices in general. We executed simulation runs of a small network comprising of smart metering nodes using the IEEE 802.15.4 low power wireless interface. The simulation framework has been validated and the validation is described in chapters 2 and 3.

4.2.1 RSA Public-Key Cryptography

RSA was proposed in 1977 by Rivest, Shamir and Adleman. Its security is based on the presumed intractability of the integer factorization problem. The RSA function is defined as follows [48]:

1: Let $p \neq q$ be large prime numbers and define $n = pq$ and $\Phi = (p-1)(q-1)$

2: Choose random encryption exponent $e, 1 < e < \Phi$, so that $gcd(e, \Phi) = 1$

3: Use the Euclidean algorithm to find decryption exponent d, $1 < d < \Phi$, so that $ed \equiv 1 \ (mod\ \Phi)$

4. Define $f : Z_n \rightarrow Z_n$ by $f(m) = m^e \ mod\ n$. $f(m)$is the function for encrypting message m.

Encryption of a message m is done by following above steps to choose the random encryption exponent e, and computing the ciphertext $f(m) = c$ by using equation (4.1).

$$c = m^e \ mod\ n \tag{4.1}$$

Decryption of ciphertext c is computed using the decryption exponent d using equation (4.2). The original message $m\ mod\ n$ is obtained as the decrypted plaintext.

$$c^d mod\ n \equiv (m^e)^d \ mod\ n \equiv m\ mod\ n \tag{4.2}$$

In the RSA cryptographic scheme, (n, e) is the public key which can be distributed to the communication entity, while d is the private key which must not be compromised. If public key component n can be factored, then d may be computed and m recovered from c efficiently. However, factoring n for correctly chosen p and q remains a hard problem and the premise for the robustness of the RSA cryptographic scheme.

RSA has a signature scheme with message recovery, which can be described as follows. If two entities, Alice and Bob wish to exchange a signed message m, Alice's signature on m is $s = m^d \ mod\ n$, where d is Alice's private key. On receiving the signature s, Bob computes

$s^e \ mod \ n = m$ using Alice's public key pair (n, e). Thus, Bob will obtain a signed message (m, s) , ensuring the message has not been tampered with. This is an effective defense against man-in-the middle attacks on communication links.

4.2.2 ECC Public-Key Cryptography

The use of elliptic curves over finite fields for data encryption and decryption was first suggested by Neal Koblitz and Victor S. Miller. Elliptic Curve Cryptography (ECC) is an alternate approach to public-key cryptography vis-à-vis RSA. It is based on the algebraic structure of elliptic curves over finite fields. In RSA, the large integer factorization problem is the intractability used to provide security. Similarly, in ECC, it is assumed that finding the Discrete Logarithm of a random elliptic curve element with respect to a publicly known base point is infeasible. This is known as the Discrete Log Problem (DLP) [49].

In current ECC practices, the elliptic curve is a plane curve which consists of points satisfying equation (4.3), along with a distinguished point at infinity. The ECC principle is deployed in the real-world in combination with the Diffie-Hellman (ECDH) key agreement protocol.

$$y^2 = x^3 + ax + b \tag{4.3}$$

In this scheme, suppose Alice wants to establish a shared key with Bob, on an un-secured channel. Initially, the domain parameters $(p, f(x), a, b, G)$, which are the large prime, cruve function, constants and finite field, must be agreed upon. Also, each party must have a key pair suitable for elliptic curve cryptography, consisting of a private key d, randomly selected in the interval $[1, n - 1]$ and a public key $Q, Q = dG$. So, Alice's key pair is (d_A, Q_A) and Bob's key pair is(d_B, Q_B). Each of them must have the other's pubic key using an exchange mechanism. Alice computes a co-ordinate point on the curve $(x_k, y_k) = d_A Q_B$. Bob computes, $k = d_B Q_A$. The shared key is the x co-ordinate point of the curve, x_k. Both Alice and Bob have computed the same number and have a shared key according to equation (4.4).

$$d_A Q_B = d_A d_B G = d_B d_A G = d_B Q_A \tag{4.4}$$

Since only Alice and Bob can compute the shared key, they can encrypt their messages with the shared key without risking decryption by an eavesdropper.

4.3 Power consumption and Processing delay of RSA and ECC

In this section we present the results of power consumed by each of the cryptographic schemes RSA and ECC. It has been shown that a key size of 1024 bits in the RSA scheme offers equivalent security to a 160 bit key size in the ECC scheme [50]. We have implemented each of these schemes on the low power wireless hardware platform Tmote Sky [51]. In addition we use the validated simulation framework to simulate a 100 node smart utility network to analyze processing delays imposed by deploying these cryptographic schemes. We also apply RSA and ECC to encrypt routing protocol packets presented in Chapter 2 in order to quantify the processing delays incurred due to RSA and ECC.

For a 9 node real network on the Tmote-Sky hardware platform, power consumption comparison of these two encryption schemes versus nodes with no-encryption is shown in figure 4.1. This data shows that there is a trade-off in implementing security in terms of increased power consumption in highly constrained nodes in the low power and lossy smart utility networks. While RSA-1024 bit performs the worst consuming maximum power of the three schemes, ECC-160 bit performs better.

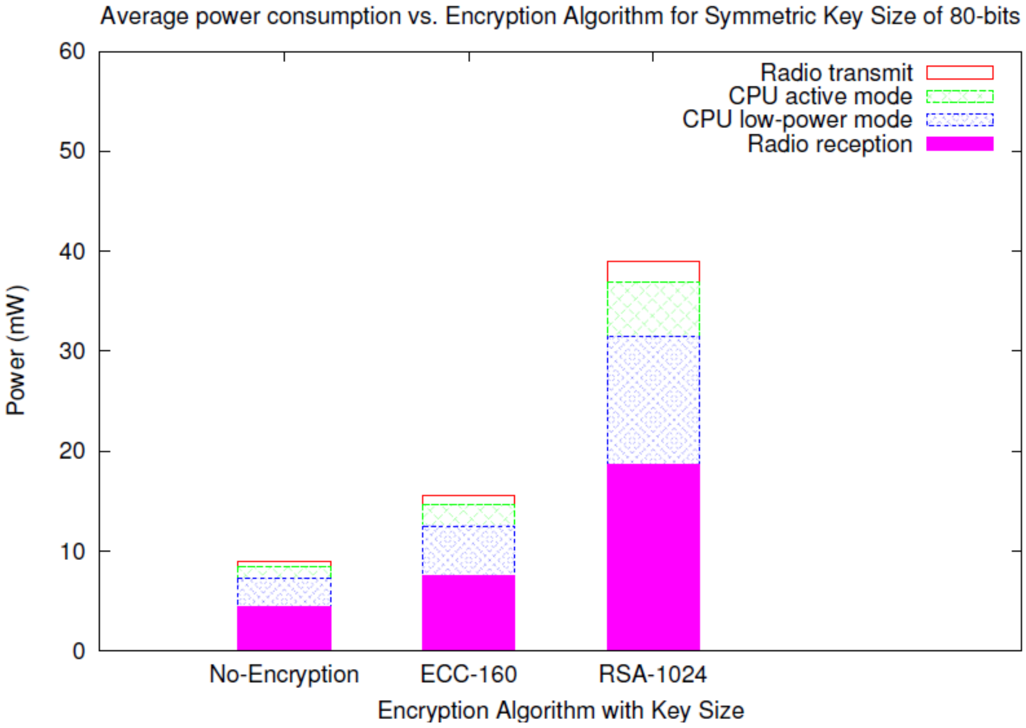


Figure 4.1: Power consumption vs. encryption scheme

Figure 4.2 shows the additional processing delay imposed by implementing RSA and ECC over routing protocols in smart utility networks. RSA imposes 10 times more delay per node than ECC in this analysis. In signature generation, RSA-1024 is 10 times slower than ECC-160 per node. Figure 4.3 shows the power consumed by the hardware during signature generation and verification operations for each of the cryptographic schemes. The order of magnitude faster ECC-160 consumes 11 times less power than the slower RSA-1024 per node during signature generation. Due to the larger number of computations involved in ECC signature verification it is approximately 5 times slower than RSA and similarly consumes approximately 5 times more power than RSA during signature verification operations.

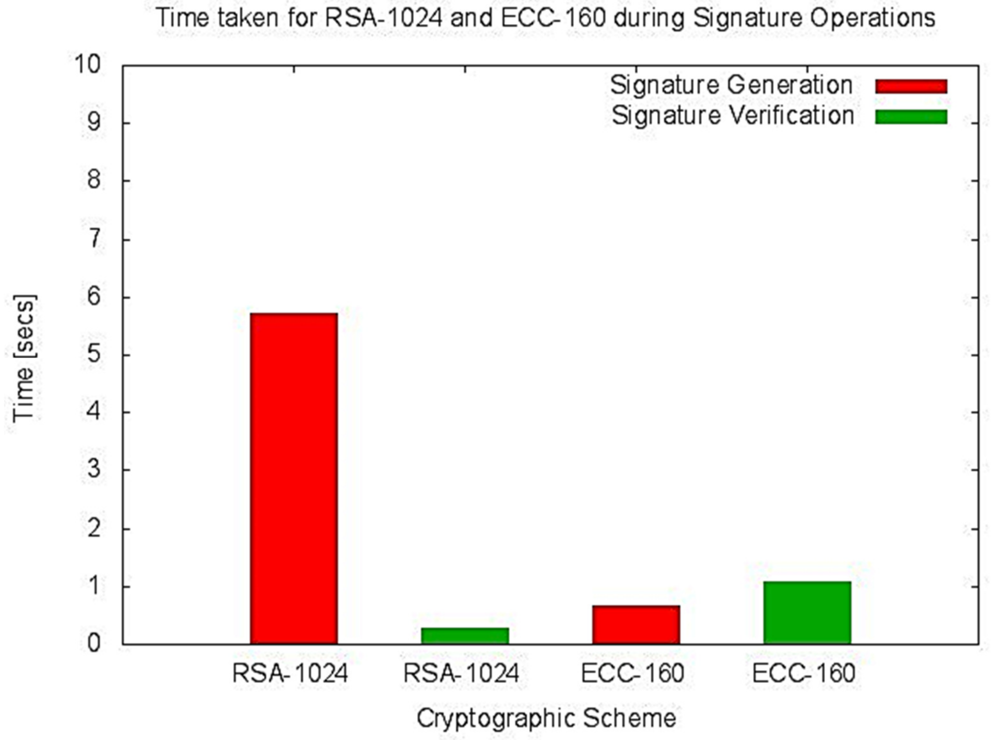


Figure 4.2: Time taken by RSA and ECC for signature generation and verification

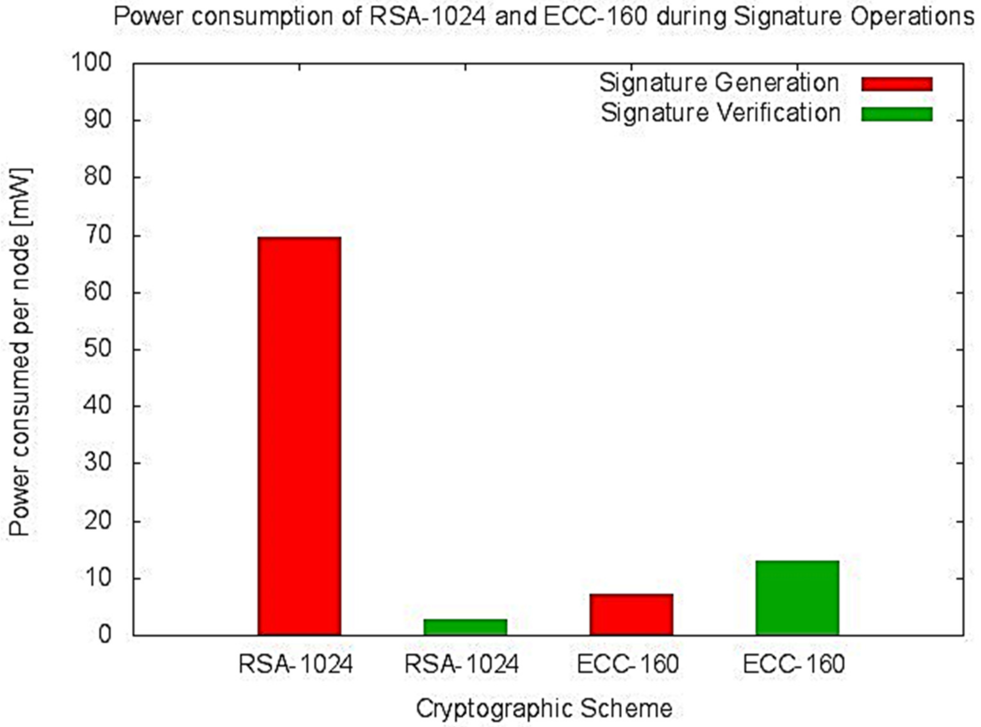


Figure 4.3: Power consumed by RSA and ECC for signature operations

4.4 Summary

In this chapter, we present the urgent need for securing the communication architecture for the smart power grids. We detail the threats to the communication infrastructure and propose appropriate mitigation schemes. Further, we describe two public key cryptographic schemes of RSA and ECC. We compare the performance of the two cryptographic schemes in terms of power consumed and processing delay imposed by implementing them on highly resource constrained nodes in the smart utility network. We find that ECC performs better than RSA in terms of power consumption. Although RSA performs better in signature verification, the gains are moderate in comparison to the power and processing time saved by implementing ECC. We choose ECC has the applicable cryptographic scheme to be deployed in smart utility networks.

Chapter 5

Conclusions and Future Work

In this chapter, we summarize the research work presented in this dissertation and follow it up with possible future directions to this work.

5.1 Summary of Research Work

In this dissertation, we have introduced and studied the aspects of Reliability, Scalability and Security in smart utility networks in detail. We have developed a validated simulation framework using OMNeT++ in order to execute large scale simulations of smart utility networks. We have presented the wireless mesh architecture as the favored method of deploying the Advanced Metering Infrastructure (AMI). Further, we have presented an analysis of emerging wireless mesh routing protocols that are most applicable to these networks. In addition, we have presented two approaches to determine the upper bound on scalability at the collector in smart utility networks, backed by simulation results using data from large scale simulations conducted at the high performance computing facility at Alabama Supercomputing Authority in Huntsville, AL.

In Chapter 1, we introduced the various components of the Smart Grid and emphasized on the importance of efficient data routing protocols to enable the future power grid. We discussed the disadvantage of existing communication infrastructure for the power grid and the urgent need to move to a synergy of internet communication technologies and the power grid. Further, we presented the smart meter and smart utility network as being at the heart of a successful smart grid implementation.

In Chapter 2, we presented the wireless mesh architecture of deployment as the most favored method. We discussed the wireless environment and channel conditions that are prevalent in the deployment zones of smart meters. We presented the simulation framework that we developed using OMNeT++ and Castalia to closely model the wireless channel conditions, low power radio based on IEEE 802.15.4 standards and drafts of routing protocols. Further, we have shown the validation method by comparing corresponding results with real world experiments. We then discuss the three routing protocols of RPL, LOAD and Geographical routing used to simulate network sizes of up to 7500 nodes in the thus developed simulation framework. We presented the performance comparison of these routing protocols and their applicability to smart utility networks. RPL is the most applicable of the three routing protocols analyzed, outperforming LOAD and Geographical routing in terms of end to end delay, hop count and packet delivery ratio. It is more reliable that geographical routing and LOAD in the order of reliability.

In Chapter 3, we studied the scalability in smart utility networks with the aim of determining an upper bound on the number of metering nodes that can be supported at each collector device. We presented two approaches, one dependent on link packet success probabilities and network topology and the other based on network traffic arrival rates and their effect on scalability. We backed this model with large scale simulations, the results of which closely match the theoretical model proposed. The relationship between average link packet success probabilities and maximum expected successful transmissions in the network was established, allowing to estimate network scale. In the second approach we modeled the smart utility network traffic as a two-state Markov Modulated Poisson Process. Following collection of simulation traces of packet arrivals at the collector for varying network sizes, we fit the raw data to the MMPP model and presented the results of how scale affects performance of the queue at the collector. In our findings, network performance degraded exponentially in terms of packet loss rates, delay and blocking probabilities beyond 4000 nodes per collector for more than 50% offered loads. This concluded the scalability analysis in smart utility networks.

In Chapter 4, we presented the urgent need for securing communication protocols for the smart grid. We presented some of the security risks and their mitigation in smart grid communication networks. Further, we detailed two popular cryptographic schemes of RSA and ECC. We presented the results comparing performance of the two cryptographic protocols in terms of power consumption and computation delay. We have implemented both these cryptographic schemes on the Tmote-Sky hardware platform and collected the data from a 9-node real network, representing the low power wireless mesh smart utility networks. We concluded in this chapter, that trade-offs come with adding security to communication protocols in networks with low power radios and highly resource constrained nodes. Under these conditions, the ECC cryptosystem performs better than the RSA cryptosystem in terms of power consumption and time taken for various cryptographic operations.

5.2 Future Work

Several lines of future work can be pursued towards meaningful conclusion following this work.

- *A more efficient Simulation Framework:* Simulating large scale networks poses challenging problems in terms of computation time. The available memory is often insufficient to load and initialize the starting conditions and states for a large number of nodes. In addition, during runtime, maintaining states of the routing layer, MAC layer and Radio states poses severe limitations on the scale. One approach to solving these problems is to move to a distributed computing environment, in which each simulated node's states are loaded onto different machines. Although this solves the memory constraints problem, inter-process and inter-machine communication to keep data consistent and exchange simulation messages significantly slows down the simulation. At scale, the run-time increases dramatically. Figure 5.1 shows the simulation time versus network deployment. The speed advantages we get from parallelization and the use of OpenMPI library for implementing parallelization for OMNeT++ and Castalia simulation frameworks will make for an interesting study.

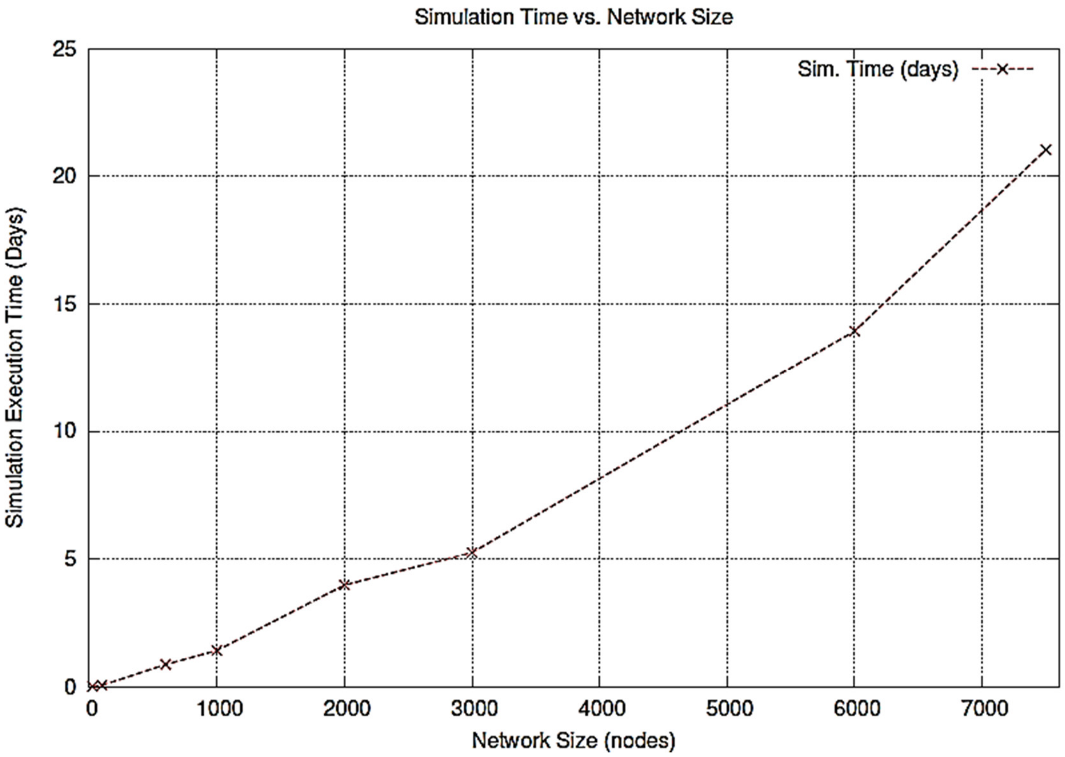


Figure 5.1: Simulated Network size versus Simulation Execution Time

- *Wireless Channel Models* with higher accuracy: Although we have validated the simulation framework with real-world results, there are spatio-temporal aspects of wireless environment that can be modeled more closely and accurately. The real challenge is in making the model scalable to large scale simulations and efficiently implemented on computing systems.
- *Rapid Prototyping:* Although we have tried to cover all the aspects of the model and routing protocol drafts in our code implementations, there are always patches and modifications emerging from IETF standards bodies and other groups. A suitable framework to rapidly prototype such changes on both simulation frameworks and suitable hardware devices will greatly benefit the community of researchers in the wireless networks domain.

Bibliography

[1] I.F. Akyildiz and X. Wang, “A Survey on Wireless Mesh Networks," *IEEE Communications Magazine*, vol. 43, Sept. 2005.

[2] “NIST Framework and Roadmap for Smart Grid Interoperability Standards,” *US Dept. of Commerce*, Release 1.0.

[3] “Utility Scale Smart Meter Deployments, Plans and Proposals,” *Institute for Electric Efficiency*, May 2012.

[4] T. Winter, P. Thubert, “RPL: IPv6 Routing Protocol for Low Power and Lossy Networks,” *draft-ietf-roll-rpl-19.*

[5] T. Watteyne, K. Pister, D. Barthel, M. Dohler, I. Auge-Blum, “Implementation of Gradient Routing in Wireless Sensor Networks,” *IEEE GLOBECOM*, Hawaii, December 2009.

[6] K. Kim, S. Daniel, et.al. , “6lowpan Ad-hoc On-demand Distance Vector Routing Protocol (LOAD),” *draft-daniel-6lowpan-load-adhoc-routing-03*

[7] T. Clausen, et.al. , “The Lightweight On-demand Ad hoc Distance-vector Routing Protocol – Next Generation (LOADng),” *draft-clausen-lln-loadng-08.*

[8] B. Lichtensteiger, B. Bjelajac, C. Müller and C. Wietfeld, “RF Mesh Systems for Smart Metering: System Architecture and Performance,” *IEEE SmartGridComm*, Maryland, Oct. 2010.

[9] G. Iyer, P. Agrawal, E. Monnerie, R. Cardozo, “Performance Analysis of Wireless Mesh Routing Protocols for Smart Utility Networks,” *IEEE SmartGridComm*, Brussels, Oct 2011.

[10] T.S. Rappaport, “Wireless Communications: Principles and Practice,” *IEEE Press*, Piscataway, NJ, 1996.

[11] Zigbee Alliance, “Zigbee Specification,” October. 2010 *http://www.zigbee.org/Specifications.aspx*

[12] Zigbee Alliance, “Understanding Zigbee Gateway,”, September 2010 *http://www.zigbee.org/LearnMore/WhitePapers.aspx*

[13] M. McGranaghan, D. Von Dollen, P. Myrda, E. Gunther, “Utility experience with developing a smart grid roadmap,” *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, 2008

[14] A. Boulis, “Castalia Wireless Channel characteristics,” *Castalia User Manual*, v. 3.2, NICTA, March 2011.

[15] P. Levis, et.al. , “The Trickle Algorithm,” *IETF RFC 6206.*

[16] I.C. Perkins, et.al. , “Ad hoc On-Demand Distance Vector (AODV) Routing,” *IETF RFC 3561.*

[17] A.F. Varga, R. Hornig, “An overview of the OMNeT++ simulation environment,” *ACM SIMUTOOLS 2008,* no. 60.

[18] Castalia Framework for OMNEST simulator, *http://castalia.npc.nicta.com.au/*

[19] J. Jun and M.L. Sichitiu, “The nominal capacity of wireless mesh networks,” *Wireless Communications Magazine*, Vol. 10, no. 5, pp. 8-14, Oct. 2003.

[20] P. Gupta and P.R. Kumar, “The capacity of wireless networks,” *IEEE Trans. On Information Theory*, vol. 46, no. 2, pp. 388-404, Mar’ 2000.

[21] J. Li, C. Blake, D.S.J. De Couto, H.I. Lee, R. Morris, “Capacity of Ad-hoc wireless networks,” *Proceedings of ACM MobiCom*, July, 2001.

[22] G. Iyer, P. Agrawal, R. Salazar, “Analytic model and simulation study for network scalability in smart utility networks,” *IEEE ISGT Asia 2013*, pp. 1-6, Bangalore, India. November 2013.

[23] Y. Feng, X. Shen, Z. Gao, G. Dai, “Queuing Based Traffic Model for Wireless Mesh Networks,” *15th International Conference on Parallel and Distributed Systems*, 2009.

[24] X. Wu, J. Liu, G. Chen, “Analysis of Bottleneck Delay and Throughput in Wireless Mesh Networks,” *IEEE Conference on Mobile Adhoc and Sensor Systems*, 2006.

[25] C. Hauser, E.D. Bakken, A. Bose, “A Failure to Communicate,” *IEEE Power and Energy Magazine*, pp. 47-55, Mar-Apr, 2005.

[26] W. Fischer, K. Meier-Hellstern, “The Markov-modulated Poisson process (MMPP) cookbook,” *Performance evaluation*, Elsevier, 1993.

[27] P. Heyman, D. Lucantoni, “Modeling Multiple IP Traffic Streams with Rate Limits,” *IEEE/ACM Trans. On Networking*, vol. 11, no. 6, December, 2003.

[28] S. Bali, V.S. Frost, “An algorithm for fitting MMPP to IP traffic traces,” *IEEE Communication Letters*, vol. 11, no. 2, pp. 207-209, February 2007.

[29] H. Heffes, “A class of data traffic processes—Covariance function characterization and relating queuing results,” *Bell Syst. Tech. J.*, vol. 59, pp. 437–488, 1980.

[30] K. Meier-Hellstern, “A fitting algorithm for Markov-modulated Poisson processes having two arrival rates,” *Eur. J. Oper. Res.*, vol. 29, pp. 370–377, 1987.

[31] S. Andersson and T. Ryden, “Maximum likelihood estimation of a structured MMPP with applications to traffic modeling,” *13th ITC Specialist Seminar*, Monterey, CA, 2000.

[32] Y. D. Lee, A. van de Liefvoort and V. L.Wallace, “Modeling Correlated Traffic with a Generalized IPP,” *Perform. Eval.* , vol. 40, no. 1–3, pp. 99–114, 2000.

[33] I. L. MacDonald and W. Zucchini, “Hidden Markov and Other Models for Discrete-Valued Time Series,” London, U.K.: Chapman-Hall, 1997.

[34] C. Blondia and T. Theimer, “A discrete-time model for ATM traffic,” *RACE*, Doc. PRLB-123-0018-CD-CC/UST-123-0022-CD-CC, 1989.

[35] D. Lucantoni, “The BMAP/G/1 Queue: A Tutorial,” *Models and Techniques for Performance Evaluation of Computer and Communications Systems*, L. Donatiello and R. Nelson, Eds. New York: Springer-Verlag, 1993, pp. 330–58.

[36] D. Lucantoni, “New results for the single server queue with a batch Markovian arrival process,” *Stochastic Models*, vol. 7, pp. 1–46, 1991.

[37] J. D. C. Little, “A proof for the formula L = λW,” *Oper. Res.*, vol. 14, pp. 723–27, 1961.

[38] W. Dong, L. Yan, S. Paul, J. Mohsen, R. Kenneth, M. Michael, “Power Infrastructure Security: Fundamental Insights of Potential Cyber Attacks and Their Impacts on Power Grid,” *Part of project “Protecting Intelligent Distributed Power Grids against Cyber Attacks” for DOE.*

[39] O. Paul, O.S. Edmund, F. Deborah, “Concerns About Intrusions into Remotely Accessible Substation Controllers and Scada Systems.” *Schweitzer Engineering Laboratories, Inc.*, 2000.

[40] W. David, “Security and Vulnerability in Electric Power Systems,” *NAPS 2003, 35th North American Power Symposium*, University of Missouri-Rolla in Rolla Missouri, October 20-21, 2003. pp.559-566.

[41] National Institute of Standards and Technology; “NIST Framework and Roadmap for Smart Grid Interoperability Standards Release 2.0,” *U.S. Department of Commerce, September 2009.*

[42] G. Iyer, P. Agrawal, “Smart Power Grids,” *42nd IEEE SSST 2010*, Tyler, TX. pp. 152-155, March 2010.

[43] C. Bennett, D. Highfill, “Networking AMI Smart Meters,” *IEEE Energy 2008, Energy 2030 conference*, pp. 1-8, November 2008.

[44] R.L. Rivest, A. Shamir, L. Adleman, “A method for obtaining digital signatures and public-key cryptosystems,” *Communications of the ACM magazine*, vol. 21, no. 2, pp. 120-126, Feb. 1978.

[45] H. Williams, “A modification of RSA public-key encryption procedure,” *IEEE Trans. On Information Theory*, vol. 26, no. 6, pp. 726-729, November 1980.

[46] N. Koblitz, “Elliptic curve cryptosystems,” *Mathematics of Computations*, no. 48, pp. 203-209, 1987.

[47] S. Kumar, M. Girimondo, A. Weimerskirch, C. Paar, “Embedded end-to-end wireless security with ECDH key exchange,” *46th IEEE Midwest Symposium on Circuits and Systems*, vol. 2, pp. 786-789, December 2003.

[48] D.R. Hankerson, D.G. Hoffman, D.A. Leonard, C.C. Lindner, K.T. Phelps, C.A. Rodger, J.R. Wall, “Coding Theory and Cryptography: The Essentials,” New York, Marcel Dekker, 2000.

[49] R. Balasubramanian, N. Koblitz, “The Improbability That an Elliptic Curve Has Subexponential Discrete Log Problem under the Menezes—Okamoto—Vanstone Algorithm,” *Journal of Cryptology*, Springer-Verlag, vol. 11, no. 2, pp. 141-145, March 1998.

[50] K. Piotrowski, L. Peter, S. Peter, “How public key cryptography influences wireless sensor node lifetime,” *SASN 2006 4th ACM workshop on Security of ad hoc and sensor networks*, pp. 169-176, 2006.

[51] Tmote sky IEEE 802.15.4 development kit, http://www.ti.*com*/tool/msp430-3p-motei-tmotesky-dsgkt