On the Capacity of a Wireless Backhaul for the Distribution Level of the Smart Grid*

Babak Karimi†, Student Member, IEEE, Vinod Namboodiri‡, Member, IEEE

Abstract—Many application scenarios envisioned as part of the ‘Smarter’ Grid depend on an effective underlying communication infrastructure for power distribution. The advanced metering infrastructure (AMI) is one such application scenario where bidirectional communication between electric meters of customers and the utility control center is required for demand response and load control. Several technologies have been suggested to meet communication needs for the backhaul that connects customer data collection points with the utility’s control center. In this work these technologies are compared with a justification of why wireless communication technologies may be most suitable for the backhaul. A linear chain multi-hop wireless communication architecture is then proposed and its ability to meet application requirements of the communication backhaul is evaluated through simulations. Based on capacity limitations first seen in the simulation results, a theoretical analysis is done to understand the data carrying capacity of using linear chain wireless technologies for the communication backhaul. Finally, the AMI application scenario is used as a case study to understand the implications of any limitations imposed by the proposed communication architecture.

Index Terms—Smart Grids, Distribution Level, Advanced Metering Infrastructure (AMI), Communications, Smart Meter, Capacity.

I. INTRODUCTION

The need for improved communications at the power distribution level takes on greater importance with the introduction of the Smart Grid approach. Title XIII of the Energy Independent and Security Act 2007 [1] requires improved operation of distribution systems. This includes development and incorporation of demand-side and energy-efficiency resources, deployment of real-time, automated, interactive technologies that optimize the physical operation of appliances and consumer devices, and provisions for timely information and control options to consumers, to name a few. These developments and deployments require additional capabilities of the grid, especially a better communication infrastructure beyond the supervisory control and data acquisition (SCADA) system.

The critical requirement for advancement in the distribution system is real time information sharing and automation. By improving the communication infrastructure, a vital ingredient for the Smart Grid, a more reliable approach could be taken to better manage assets. In addition to asset and outage management tasks, communication will also aid in better energy management and tariff-related information. Deployment of smart distribution systems necessitates proper identification of power system requirements and integrating suitable communication and control infrastructure. The information communication and control layer of the smart grid brings about numerous advances, including the empowerment of customers to actively participate in the maintenance of a supply-demand balance round the clock and reliability improvements in electricity service.

Advanced Metering Infrastructure (AMI) initiatives are a popular tool to incorporate the changes for modernizing the electricity grid, reduce peak loads, and meet energy-efficiency targets. With the introduction of AMI technology, two-way communication between a smart meter (SM) and the control center, as well as between the smart meter and customer loads would be facilitated for demand response, dynamic pricing, system monitoring, cold load pick-up, and greenhouse gas emission mitigation [2]. AMI uses technology to capture and transmit energy use to a concentration point on an hourly or sub-hourly basis in contrast to standard meters that provide a daily energy usage total and a cumulative monthly bill. This application requires bidirectional communication: control commands from the control center of the utility to smart meters, and load profiles and logs from smart meters to the control center.

There are many communication technology and architecture options in deploying the AMI. As shown in Figure 1, the end-to-end communication infrastructure can be divided into two major phases: (i) the consumer level home area network (HAN) between a smart meter and various electrical equipment/appliances, (ii) the backhaul link that collects information from smart meters and carries it to the utility control center. The backhaul can either be a direct link from individual smart meters to the utility control center, or a link (or series of links) from a concentrator node, which aggregates data from multiple smart meters through some form of a mesh network, to the control center. There is reasonable consensus on using ZigBee-based star topologies in HANs [3], [4], and ZigBee or Wi-Fi based mesh topologies to collect and aggregate data at concentrators [5], but the technologies and topologies to be used for the backhaul is still an open problem.

This work thus looks at the design of backhaul communication for the AMI. An approach of deploying the backhaul
communication infrastructure is to create a network through the existing feeder infrastructure. This approach has the advantage of reducing costs by reusing existing infrastructure of electric poles and feeder lines if needed. Further, it can also integrate well with sensors deployed on feeder poles to improve distribution automation through applications like automated fault location and self-healing feeders. It is not yet clear though what would be the communication capacity of a very long linear chain network deployed for the AMI application between concentrator nodes and the utility control center. The AMI application presents the challenge of collection and management of data from smart meters. By current standards, each smart meter sends a few kilobytes of data every 15 minutes to a smart meter [6], [7]. When this is scaled up to large numbers, many existing communication architectures will find it difficult to handle the data traffic due to limited bandwidth. To this end, this work explores the feasibility of utilizing a wireless linear chain network to meet communication requirements for AMI, analyzes the limitations imposed by such an architecture on the amount of data traffic that can be carried, and explores how grid operators could plan and operate their deployments.

The work in [8] is closest to our work by considering a wireless backhaul architecture. Their focus is however on enhancements to the air interface and network protocols instead of network capacity. The work in [9] considers most last-mile options for telecommunication in a Smart Grid, especially backhaul solutions for the distribution network. The authors of [10] considered PLC in low- and medium-voltage distribution grids to connect network nodes (e.g., meters, actuators, sensors) through multi-hop transmission. The authors in [11] proposed a unified solution for Advanced Metering Infrastructure (AMI) integration with a Distribution Management System (DMS). They found that a challenge of the integration of AMI and DMS is that it entails different communication protocols and requirements for handling various meter information models.

While the prior work above discusses how to design the communication network at the distribution level for AMI, some researchers have looked beyond. For example, [12] discussed one of the key components of a future Smart Grid called load leveling and elaborates how existing techniques from computer networking research could be potentially applied to solve these problems. The authors of [13] described an approach to modeling wireless communications at the link layer of the power grid, with emphasis on employing a medium access control/physical layer model to measure link performance in terms of reliability, delay, and throughput. Some researchers believe that wireless communication is not enough to meet the entire needs of Smart Grid communication. For example, the authors in [14] describe a hybrid Wireless-Broadband over Power Lines (W-BPL) technology. They believe that this combination is suitable for rural and remote areas.

The remainder of the paper is organized as follows. In Section II we justify using wireless communication as the medium of choice for meeting communication requirements at the distribution level of the Smart Grid. We identify communication requirements at the distribution level by considering specific applications and compare the suitability of existing wireless technologies to meet these requirements in Section III. A wireless multi-hop linear chain architecture is proposed as an ideal solution followed by evaluations through simulations on the ability of this architecture to meet some of the requirements. Section IV fundamentally analyzes the data traffic capacity of the proposed linear chain architecture and uses this analysis to determine how many chains might be needed to support a given amount of data traffic. The AMI application scenario is analyzed in Section V as an example to determine what might be the capacity requirements in the future from the communications infrastructure, and how our capacity analysis can be used in setting data gathering parameters and leveraging multiple technologies in current and future deployments. Finally in Section VI we make concluding remarks and discuss the utility of this work in communications infrastructure planning for grid operators and final conclusions.

II. COMMUNICATION OPTIONS AT THE DISTRIBUTION LEVEL

In this section we motivate the choice of using wireless communication technologies over other options, discuss possible limitations, and the types of wireless architectures possible. This sets the stage for further study into the communication performance of such architectures in the later sections of this paper.

A. Choice of Communication Medium

A communication network for the power grid allows power utility companies to access electricity usage data and services remotely, regardless of their geographic position. Real-time monitoring of transmission and distribution lines for protection against natural disasters or even malicious attacks are all reasons to have a secure, reliable, and scalable communication
network for the power grid. Several last-mile options are available for getting a communication network to be operational; Broadband technologies like Digital Subscriber Line (DSL), Fiber Options (FTTX), and Power Line Communication (PLC) are some examples. All these technologies have their limitations, however, due to their fixed nature and lack of flexibility.

When distribution feeders are considered, PLC is well-suited, because it is a no-cost medium for the utility and is spread along the distribution system. PLC has the potential to transmit data at a maximum rate of 11 Kbit/s and the maximum data rate can be achieved only in a narrow frequency range of 9 to 95 kHz [15]. This low rate of communication is not enough for supporting applications where large amounts of data may be transferred, for example when large number of smart meters connected to end-user loads send periodic information using the AMI.

The current developments in the Broadband over Power Line (BPL) could create an impression that this is the best technology, but that is not so. The distribution system is consistently affected by voltage transients and harmonics that are unpredictable, and is thus prone to a high level of disturbance. High-frequency signals involved in BPL need to bypass transformers to avoid high attenuation [16]. The attenuation in a radial distribution feeder is high, and this would increase the number of regenerators needed. It is expected that a typical 20-mile-long rural feeder needs regenerators on the order of 30 to 100 [17]. Thus, even if the medium of communication is free in BPL, there are infrastructure costs involved. In addition, the high-frequency signals may be blocked by voltage regulators, reclosers, and shunt capacitors, which are common in long radial feeders [17] posing problems.

Another option would be to use dedicated wired communication; however, one problem with copper wire connections is interference and attenuation. Fiber optic cables would be a solution to the interference, but would increase the cost. It should be noted that the investment required for a fiber optic network would be on the order of $10-100 million for 100 nodes [18]. Newly developing communities could get to install a fiber optic communication network close to the feeders, enabling infrastructure to be shared for both the power grid and consumer communication needs. One of the advantages of this medium is that the utility would bear only the costs of the terminal equipment and for leasing the line, which would reduce the utility’s overhead and improve communication. On the other hand, the utility would not have control over the medium because, in most cases, it would not own the dedicated wired network, and would require physical connections that reduce flexibility.

All the above technologies further have the disadvantage that when an electric pole goes down, it takes the communication link down as well. This would be a major concern when the communication is used for automatic fault location and system restoration, where communication is expected to help the power grid back to normalcy. For Smart Grid applications, a highly reliable communication network is necessary, with some prior work recommending an availability as high as 99.995% for the communication network [19]. This percentage requirement would result in the per-year unavailability of communication to less than 4.4 hours. All these concerns build the case to explore other options for the communication medium at the feeder level.

Wireless communication is a promising alternative for distribution-level communication. One of the important characteristics of wireless communication is the feasibility of communication without a physical connection between two nodes, thus ensuring continued connectivity even when a few poles are down. In other words, redundant paths for communication are possible without additional cost. Another advantage of using wireless communication is that the utility must own only the terminal units, which are relatively cheap and could be integrated with cost-effective local processors. When multi-hopping is used in wireless communication, the range of communication could be extended, and the nodes located on the feeder would be able to communicate with the control center. A major concern with a wireless medium is easy accessibility, which could result in security issues. This could be avoided by using security mechanisms presented in prior work like [20]. The disadvantages of wireless communication could be interference due to the presence of buildings and trees, which could result in multi-paths. Further, rural feeder sections could be long, and the range of communication could become a concern. These issues could be avoided to some extent with improved receivers and directional antennas. We discuss the issue of interference from power lines to wireless communication next.

B. Impact of Interference Due to Transmission Lines on the Wireless Medium

One of the concerns in using wireless communication along power lines is the interference from high-voltage transmission lines. Electromagnetic noise generated around high-voltage power lines is an undesirable disturbance, which can affect wireless data transmission. This noise can be observed as an additive signal to the original one, and it can interrupt, obstruct, degrade, or limit the performance of communication systems. According to [21] this noise is due to the following: Discharges between line components: This occurs only in power lines less than 70 kV. This type of noise is generated in insulators, in metallic parts, or in faulty or improperly installed equipment. The noise tends to dominate the frequency spectrum between 10 and 20 MHz. Its effects can be controlled by ensuring a correct power line installation and proper maintenance. Corona effect: This affects power lines over 110 kV and tends to dominate the frequency spectrum between 10 and 30 MHz. It is generated due to partial discharges in areas with a very high electric field and causes acoustic noise, radio interference, and mechanical vibrations.

In [21], it was concluded that the radio interference generated by high-voltage lines diminishes logarithmically with the distance to the power line and with increasing frequency. Therefore, it is recommended that communication modules be operated at frequencies greater than 100 MHz. A selection of wireless communication technologies like Wi-Fi, ZigBee, or WiMax, which operate in the GHz range, could be utilized in the distribution system with minimal interference.
C. Types of wireless communication architectures

Wireless communication architectures could be classified based on whether they use single-hop or multi-hop communication between two end-points, and whether they employ only a single technology or a combination of multiple technologies. Single-hop communication is possible only when the distance between the two points of interest is small enough to fit within the communication range of a technology. In many cases, when large geographical distances are involved, such as the case of the distribution system, multi-hop communication will be required where intermediate nodes forward data from the source to the destination. Wireless multi-hop communication provides several benefits like extending coverage due to multi-hop forwarding, greater throughput due to shorter hop distance, and possibly lower costs than long-distance communications. A combination of wireless technologies could be employed in cases where one technology cannot cover a region, or does not have the capacity to support generated data traffic, or due to economic reasons. Later in this paper we will propose and evaluate a multi-hop wireless communication architecture and also look at the capacity limitations of using only a single technology, and how it scales up when using a combination of multiple technologies.

III. A WIRELESS COMMUNICATION NETWORK ARCHITECTURE FOR THE DISTRIBUTION SYSTEM

Having made a case for using wireless technologies in the previous section, we describe some communication requirements for emerging applications at the distribution level. The fault location application is one such example. Existing practices for locating possible faults at the distribution level involve a lot of manual interventions. Only approximate locations of faults are known at the time of an outage, and operators need to spend time to identify the exact location and determine cause of the failure and then fix it. If each feeder pole had a wireless master node, with the help of local sensors on power lines, it could provide useful information on the power lines passing through it in a matter of seconds. Thus, the fault location and recovery could be much faster and lead to much more automation involving significantly less manual intervention. The AMI application would also benefit as a wireless communication architecture would make it easy to collect data from smart meters requiring little infrastructure support like cabling and provide flexible re-configurations. To help realize such applications and make progress towards identifying communication performance we compare multiple candidate wireless technologies, and propose and evaluate the ability of linear chain architecture to meet some of these requirements.

A. Communication Requirements

Using the two applications of AMI and automated fault location mentioned above as a guideline, and based on the general characteristics of the distribution system, the following is a list of requirements we have identified as important for any communication architecture to satisfy:

1) Low-latency Communications: Any wireless technology used should be able to provide low-latency communications from the data generation or collection point to the eventual destination. For an application like fault location, as soon as an abnormal state is sensed, this event should be communicated from the feeders to the control center for possible action. Any control commands from grid operators should similarly reach localized points on the grid with minimal delay. For applications like AMI, a higher latency is tolerable for the data collected from smart meters to the control center, but there are control commands in the other direction (from the control center to smart meters) for controlling loads and remote connect/disconnect which need to be communicated immediately.

2) Low infrastructure development and maintenance costs: Modernizing the grid involves a lot of upgrades. It is imperative that the cost of developing additional infrastructure is minimized by re-using existing infrastructure where possible. Thus any wireless technology used should be inexpensive, possibly unlicensed (to minimize spectrum licensing costs), and be able utilize existing infrastructure (like feeder poles), and be easy to setup and reconfigure. The technology used should also preferably have low maintenance costs, including fees to third party providers/carriers.

3) Scalability: The scale of deployments possible under a single grid operator at the distribution level could be of the order of hundreds or thousands or nodes (even more if smart meters and end-user loads they connect to are included). Any wireless technology and the services it provides should be scalable in terms of the technical features it provides like data carrying capacity and latency.

4) Privacy, Information Security, and Network Availability: With applications like AMI fostering customer participation in energy delivery, ensuring privacy and information security in general is a challenge. With wireless technologies relying on a shared broadcast medium, unlike their wireline counterparts, this issue takes on greater dimensions. Further ensuring that the communication network is available for operation and control can be critical.

In this paper we will mainly focus on scalability of the communication backhaul, and low-latency to some extent. We leave the exploration of other requirements for future work.

B. Selection of Appropriate Wireless Technology and Architecture

The choices of wireless technologies, Wi-Fi, WiMAX, ZigBee and cellular data service are compared in Table I. Based on the information in this table, and the fact that electric poles are typically separated by 100-300 feet\(^1\) (less than 100 meters), Wi-Fi seems to be the wireless technology that meets communication needs at the distribution level. Wi-Fi is based on unlicensed frequency bands and provides cost benefits and

\(^1\)This does vary a lot based on terrain and population density. For example, for some terrain the separation has to be large due to the inability to setup poles any nearer. In low population density, rural, areas electric poles can also be found to have larger separations. In such cases longer range directional antennas might be needed, or long-range technologies like WiMAX will have to be used.
The feasibility of using a collection of cost-effective, fixed wireless nodes relying on Wi-Fi technology forming a linear chain network is evaluated here. We base the feasibility of the architecture on the performance measure of end-to-end delay, and two measures of scalability — packet delivery fraction and node density.

1) Experimental Setup: Simulations are carried out in the open-source network simulator ver. 2, ns2 [25], which allows abstraction of all communication protocols and their performance evaluation for different network topologies and configuration of various network traffic types. The nodes were placed as a linear chain topology (mimicking electric poles on a distribution line) on a 10 km long scenario, each node separated from the next by a distance of 100 meters. This distance between nodes was varied only if the impact of node density was studied, in which case we explicitly mention the density (nodes per unit distance). In the topology, the first node from the left is assumed to be the source while the destination was the right endpoint of the chain to which information was being sent. Continuous bit rate (CBR) data traffic was used from the source with varying data rates in our experiments. For all simulations, a fixed antenna model is used with the two ray propagation channel model. The transmit power is set to be fixed at 0.28 Watt, providing a range of 100 meters in our simulator, which is consistent with practical values for Wi-Fi. The channel data rate was set to 11Mbps. The following performance measures were used for the routing protocol:

Packet delivery fraction (PDF): This is the ratio of packets delivered to packets sent by the traffic generator. End-to-end packet delay: This is the average of the delays encountered by all successfully received packets at the destination from the source node.

Node density: This term is used to describe the constant number of nodes deployed within a fixed geographic length.

2) Selection of Routing Protocol: Ad hoc routing protocols that we can use for our mesh architecture can be divided into two main categories: reactive (on-demand) and proactive (table-driven). Other possible categories include location-based routing (e.g. [26]) and prediction based routing (e.g. [27]). There has been extensive prior work on evaluating routing protocols for ad hoc networks. (e.g. [28]). We picked Ad Hoc On-Demand Distance Vector (AODV) routing protocol [29] as the representative routing protocol in this work due to prior evaluation results and our own comparisons with proactive routing protocols. In on-demand routing protocols, routes are created as and when required. When a source wants to send to a destination, it invokes route-discovery mechanisms to find the path to the destination. These discovered routes time-out after a fixed duration, requiring new routes to be created to replace them. For our evaluations we modified some default parameters of the AODV protocol to work with route lengths expected to be many hundreds of hops. For example, the default network diameter for AODV in ns2 is 30 hops; we modified it to be larger than the number of hops expected in our long chain topology.

3) Feasible Operating Conditions: As a first step, it was decided to determine the appropriate node density when deploying the proposed wireless communication architecture. Data rates of 0.01 and 0.5 Mbps were used. The end-to-end network delay for varying node density is shown in Figure 2. As node density increases initially, there are multiple nodes that act as intermediate nodes and interfere with each other. The drop seen for node densities of 20 to 21 in Figure 2 is due to a sudden decrease in the number of hops taken by the AODV routing protocol from the source to the destination. This protocol considers all possible paths from the source to the destination and picks the one that can reach the destination with the least delay, which typically is the route composed of the shortest number of hops. As node density increases, there

### Table I: Comparison of Possible Wireless Technologies

<table>
<thead>
<tr>
<th>Attribute</th>
<th>WiMAX</th>
<th>Wi-Fi</th>
<th>ZigBee</th>
<th>GSM/UMTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>High</td>
<td>Medium</td>
<td>Low/Medium</td>
<td>High</td>
</tr>
<tr>
<td>Range (Single-hop)</td>
<td>500-900 m</td>
<td>200-400 m</td>
<td>10-75 m, 1500 m</td>
<td>1-2 mi</td>
</tr>
<tr>
<td>Max. Data Rate</td>
<td>50 Mbps</td>
<td>54 Mbps</td>
<td>250 kbps</td>
<td>20-800 kbps</td>
</tr>
<tr>
<td>Frequency Band</td>
<td>2.1 GHz, 10-66 GHz</td>
<td>2.4 GHz, 5 GHz</td>
<td>2.4 GHz, 915 MHz, 868 MHz</td>
<td>700 MHz, 2.1 GHz</td>
</tr>
<tr>
<td>Band License</td>
<td>Free and Licensed</td>
<td>Free</td>
<td>Free and Licensed</td>
<td>Licensed</td>
</tr>
<tr>
<td>Flexibility</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Robustness</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>
is a certain point at which the routes taken can skip over some nodes on the path to the destination. From Figure 2, it can be seen that node density of 10 nodes per km, which is also the minimum needed for end-to-end connectivity for the transmit power level used, performs better for both 0.01 Mbps and 0.5 Mbps cases. The lesser the node density, the cheaper the cost of deploying the communication architecture as it reduces the number of nodes that need to be deployed. Therefore, this work suggests using the minimum possible node density for a given communication range of a technology.

Further analysis was done for delay and packet delivery for data rates between 0.01 and 1 Mbps. To compare the effect of node densities, both 10 and 15 nodes per km were simulated. Figure 3 shows the simulation results for these cases. From Figure 3(a) and (b), it can be seen that when the data rate exceeds 0.15 Mbps, the delay increases significantly, and the packet delivery fraction decreases from 100%.

From our simulations, it can be seen that even with the ideal communication scenario (no interference and only one node sending information), there is a limit on the capacity. In subsequent sections we explore these limits of a wireless communication architecture and offer insights on planning deployments with possibly a combination of technologies. Readers interested in additional simulation based evaluation for the Wi-Fi linear chain topology are referred to [30].

IV. CAPACITY ANALYSIS OF FEEDER LEVEL LINEAR CHAIN NETWORK TOPOLOGY

In this section we theoretically analyze the capacity of a linear chain topology likely to be used for feeder level communications at the distribution level. Though a Wi-Fi architecture was proposed as the technology of choice for communications at the distribution level in our previous work [30], we keep the treatment in this section of the paper and following sections to be more general allowing for the use of other wireless technologies as well in any communication architecture that is used.

A. Problem Definition and Assumptions

Consider a linear chain communication network that is multi-hop in nature with \( n + 1 \) nodes \((n \) hops from source to destination) overall separated by a constant inter-node distance. Let \( C \) bps be the maximum single hop throughput or capacity, and \( r \) meters be the communication range of the technology used. The source node generates data at a rate of \( \lambda \) bits per second. Let the sink node be \( L \) meters away from the data concentrator. We would like to determine what is the end-to-end data capacity, \( X_c \), of this linear chain topology.

This problem is representative of the type of linear chain network that will be used at the distribution level of the power grid as described in our proposed Wi-Fi architecture in the previous section. The source node could represent a concentrator or data aggregator where all data from many distributed sources (e.g. smart meters of residences in the AMI application) could be collected and sent over the network. The sink node could represent a control center of the grid operator where all data is collected. The scenario could also be the other way around, where the source node is the control center sending commands to a sink node that could be an end-user’s smart meter. The smart meter to control center direction is more interesting in terms of network capacity analysis as it is expected to have more data; the other direction is mainly envisioned for control commands. Figure 4 shows an illustration of the specific application scenario at the feeder-level. The assumption of constant inter-node distance is based on the assumption that nodes are co-located with existing utility infrastructure (e.g. distribution feeders) that is deployed at some fixed density. Determining the maximum data traffic rate that can be supported by such a linear chain network for any given communication technology that is deployed is useful in network capacity and operational planning, and reliability analysis. Prior research on the capacity of ad hoc wireless networks have studied similar problems, which also include single chain capacity analysis [31], [32]. However, their focus was solely on Wi-Fi technology and considered many other topologies (apart from linear chain), including random networks; our goal is to analyze and study the capacity of linear chain wireless topologies and how it applies to the distribution level, and without being restricted to only one technology.
and the chain. We will use a recursive analysis to calculate $f_n$ sufficiently large. Let $+ 1$ be active simultaneously if they have data to forward (link links that are active in the chain given that the first link $L_1$ is active). Further exactly one of $L_m$ can be active. If $p_i$, $1 \leq i \leq n$ is the probability that any link $L_i$ is active, then we can compute the average number of active links given $L_1$ is active as

$$f(n) = 1 + \sum_{j=1}^{m} (p_j * f(n - (m + j) - 1))$$

Next we compute $g(n)$ letting $q_i$, $1 \leq i \leq n$ denote the probability that $L_i$ is active. Then,

$$g(n) = \sum_{i=1}^{m} (q_i * f(n - i + 1))$$

where we assume any link $i$ from the first $m$ links can be active removing the conditionality imposed by $f(n)$ on $L_1$ being active. Continuing,

$$g(n) = \sum_{i=1}^{m} q_i \sum_{j=1}^{m} (p_j * (1 + f(n - i + 1 - (m + j) + 1)))$$

$$= \sum_{i=1}^{m} q_i \sum_{j=1}^{m} p_j + \sum_{j=1}^{m} q_j * f(n - i + 1 - (m + j) + 1)$$

$$= 1 + \sum_{j=1}^{m} p_j * g(n - i + 1 - (m + j) + 1))$$

Assuming that $g(n)$ is a linear function of $n$ and using $an + b$ as a possible solution with $b = 0$, and $p_i = 1/m$, $1 \leq i \leq m$, we can solve and obtain $a = \frac{1}{2} \frac{2}{3m-1}$. Thus, $g(n) = an = \frac{2}{3m-1}$ is the average number of links active in linear chain at steady state when $n$ is large. This results in an overall chain capacity, $X_c$, of

$$X_c = \frac{g(n)}{n} C = \frac{2C}{3m-1}$$

where $C$ is the single hop capacity. For smaller values of $n$ we can utilize the recursion in Equation 1 to compute different values of $g(n)$ and compute end-to-end chain capacity.

Validation of analysis through simulations

To verify our assumptions we conducted a series of simulations and compared the results of our theoretical result above with simulations of the Wi-Fi technology. Note that other wireless technologies could be compared similarly on introducing correct parameter values for those technologies. The simulations were done with the ns2 simulator [25] whose parameters were set to a 2 Mbps channel data rate. The average end-to-end chain throughput is shown along with 90% confidence intervals. Additional details of our simulation setup were described in III-C1. For the theoretical results we set the value of $C$ in the Equation 1 to 1.7 Mbps as overhead reduces useful throughput from the 2 Mbps channel rate. The communication range $r$ was set constant to 100 m and the interfering range of each node was set such that $m = 4$ in the above analysis. The distance $L$ was varied to vary the chain length. The corresponding values of $g(n)$ were computed using Equation 1 and the resulting capacity values for each chain length plotted.

From Figure 5 the results of our theoretical analysis and simulations can be compared. For the largest data packet size of 1500 bytes, which is expected to have the greatest throughput, it can be observed that the difference between the two results is very small ($< 6\%$) validating the analysis. We can further improve on our analytical result above (and get

\[\text{Fig. 4. The backhaul linear chain architecture considered in this work for AMI.}\]
closer to the simulation results) by adding a correction for the way the IEEE 802.11 standard medium access control (MAC) protocol works for Wi-Fi. Under this protocol, there could be instances where none of the \( m \) consecutive links could be transmitting as they might be counting down backoff slots in accordance with the binary exponential backoff algorithm used [33]. We do not do account for this to keep our analytical result presented in this section more general.

![Graph](image)

(a) Theoretical chain capacity
(b) End-to-end chain throughput through simulations

Fig. 5. Comparison of theoretical capacity and end-to-end throughput through simulations

**Practical Considerations**

Another step to confirm our results would be to perform experiments on a large-scale testbed. However, this approach would have the limitation that we would not be able to scale the network up to gather any meaningful results. For large networks (or long chains), simulations is the best way to gauge performance. The NS2 simulator has been used for such large scale simulations many times and provides a useful check against our analytical results. In our opinion, comparing both simulation and analytical results provides a good validation check and is common in prior research with large scale networks. Having said that, there has been related work that has considered the capacity of multi-hop wireless networks like [34] and [35], but the most relevant practical scenario to our architecture is presented in [36] where authors have implemented a testbed to validate the path capacity on multi-hop fixed rate wireless networks. The testbed included several 802.11b laptops kept about 70 ~ 80 meters (which is similar to the distance between power distribution poles) apart in a chain topology. For comparison, we have shown their result in Figure 6. Remarkably, their experimental results are very similar to ours in this work, adding confidence to the practicality of our results.

![Graph](image)

Fig. 6. Practical experiment results of Ad Hoc Probe in [36] on wireless multihop testbed.

**C. Multiple Chain Analysis**

If the rate of data generation is higher than what a linear chain communication topology can support, multiple chains could be possibly deployed. However, it is important that these chains do not interfere with each others flows. It is useful at this point to study the impact of interfering chains (or interference in general) on end-to-end throughput achieved. We simulate varying number of interfering chains of length 10 hops, each hop spanning 100 m as before. The source node of each chain was configured to send at a rate of 100 kbps. The average end-to-end chain throughput is shown along with 90% confidence intervals. As Figure 7 shows, each interfering flow results in a rapid decrease in end-to-end throughput. Thus, interfering flows need to be carefully planned. Interfering flow chain could be co-located using the same physical path and the same (or different) nodes, but possibly different non-interfering frequency channels, or use a different non-interfering path along another set of feeders.

![Graph](image)

Fig. 7. Impact of interfering flows on end-to-end chain throughput.

After ensuring flows are non-interfering, it is thus useful to extend our single chain analysis to how many chains would be needed for some given data generation rate. The eventual wireless communication architecture could thus be composed of multiple independent linear chains from many concentrators to the control center. The presence of multiple such chains, if planned and deployed properly, can add some reliability to the communication network by providing the option of alternate paths on re-configuration (from a non-overlapping non-interfering configuration).

Let \( X_c \) be the capacity of single chain as found in Equation 2. Let \( N \) be the ratio of data to be sent from the concentrator, \( \lambda \), to the capacity of chain \( X_c \). In order to estimate the number
of chains that we need to support the required $\lambda$, there is a relation which is always true and is:

$$\lambda \leq NX_c$$  \hspace{1cm} (3)

Thus to get the number of required chains, we define the function $f$ as:

$$f\left(N = \frac{\lambda}{X_c}\right) = \begin{cases} 
1 & \text{if } N \leq 1 \\
\lceil N \rceil & \text{if } N > 1
\end{cases}$$  \hspace{1cm} (4)

V. DATA RATE REQUIREMENTS ON COMMUNICATION INFRASTRUCTURE

The analysis in the previous section looked at how much data rate, $X_c$, a single linear chain communication network can support from a source to sink and how many chains, $N$, will be needed to carry a specified amount of data traffic $\lambda$. In this section an analysis of data communication requirements when using the AMI application will be done so as to better understand the magnitude of $\lambda$ that will have to be supported. The application scenario under consideration was shown in Figure 4 where the smart meter aggregator or concentrator point gathers data from smart meters and sends it to the control center.\(^5\) This section will also give insights on the amount of data and interval at which smart meters should send this data for a given communication infrastructure.

A. AMI data output

Suppose there are $z$ smart meters in a given area as part of the AMI application, all connected to a concentrator. As the rate of data generation by smart meters are not necessarily synchronized, an average rate of data generation per second is a more useful value. Let each meter generate a data packet $P$ bytes long at an interval of $t$ minutes. Thus, the average data traffic rate reaching the concentrator from $z$ smart meters will be:

$$\lambda = \frac{z \cdot P}{60 \cdot t} \text{ bytes/sec} = \frac{z \cdot P}{7.5 \cdot t} \text{ bits/sec}$$  \hspace{1cm} (5)

Current standards specify that each smart meter sends a 512 byte packet every 5, 15, 30, or 60 minutes\(^6\), but 15 minutes is most common \(^7\). Thus, henceforth we use the interval of 15 minutes for evaluations. This translates to a data rate of $\lambda = 4.55z$ bps arriving at the concentrator.

This simple analysis can be further extended to include the concept of smart meter density. For example, assume each house in a residential neighborhood has a smart meter, and the smart meter density (equivalent to housing density) per square meter, say $\rho_H$, is known. So, given the area of the location of interest, $A$, and the housing density, $\rho_H$, the expected data traffic arriving at the concentrator would be:

$$\lambda = \frac{\rho_H A \cdot P}{60 \cdot t} \text{ bytes/sec} = \frac{\rho_H A \cdot P}{7.5 \cdot t} \text{ bits/sec}$$  \hspace{1cm} (6)

Figure 8(a) plots the data traffic at the concentrator for various values of $\rho_H$ for an area of 1000 sq. m and packet size $P = 512$ bytes. Three intervals at which meters could send data are considered. It can be seen that huge amounts of data can be easily generated by applications in dense AMI deployments.

B. Data transfer interval for a given infrastructure

If the communication infrastructure is already in place, the parameter that decides the data rate that will need to be supported by this infrastructure is the interval $t$ at which each smart meter sends data given $\rho_H$, $A$, $P$, $N$, and $X_c$. We can use equations 3 and 6 to get the following inequality

$$\rho_H \cdot A \cdot \frac{P}{7.5 \cdot t} < N \cdot X_c$$

which simplifies to provide a lower bound for the sending interval as

$$t > \frac{\rho_H \cdot A \cdot P}{7.5 \cdot N \cdot X_c}$$  \hspace{1cm} (7)

Figure 8(b) provides an idea of how the sending interval for each smart meter increases with a decrease in end-to-end capacity of the communication infrastructure for $m = 4$, $N = 1$, $A = 1000$ sq. m, and $P = 512$ bytes.

C. Data Handling Requirements for Large, Hybrid Topologies

The analysis so far was based on a homogeneous architecture where only one type of technology was used to transport data from concentrators to the control center. Here we extend our analysis to study expected data traffic that needs to be handled over a large, wide-area network where possibly multiple technologies might be needed from concentrators to reach the control center. For example one technology could be used within metropolitan areas due to a higher population density, while another technology is used for suburban or rural users. We begin by developing tree-based model for the communication infrastructure and subsequently use this model for assessing data traffic handling requirements.

Figure 9 shows how a tree architecture can be used to model $h$ different levels of technologies used end-to-end. There is a root node, which is the control center and the final destination of all data collected from smart meters. All leaf nodes at the

\(^5\)The technologies used in gathering data from each individual customer to the smart meter and the capacity analysis of this network is not addressed in this work and would be part of future work.
lowest level of the tree are smart meters generating data. All other intermediate nodes of the tree that are neither leaves or the root node, are the concentrators. We are interested in the data traffic at each concentrator, and the overall data traffic received at the control center.

Let the height of the tree be \( h \), starting from level 0 to level \( h \). Let there be \( n_c(i) \) cluster(s) at each level \( i \). Then there are \( N(i, k) \) aggregators (concentrators) in \( k \)th cluster of level \( i \), where \( 0 < k < n_c(i) \) and \( i > 0 \). This helps to identify the lower level data traffic senders to which a concentrator \( j \) of cluster \( k \) at level \( i \) is connected to. Thus \( \lambda_{ijk} \) is the rate of data collected at concentrator number \( j \) in cluster \( k \) at level \( i \) and defined as:

\[
\lambda_{ijk} = \sum_{j'=1}^{N(i',k')} \lambda_{i'j'k'}
\]

where \( i' = i - 1 \) (lower level), \( j' \) is the aggregator index in cluster \( k' \), and \( k' \) is the cluster identifier in level \( i' \) that sends its data to aggregator with identifiers \( i, j, \) and \( k \).

To represent the total data traffic that is being collected at any specific level \( i \), we use \( \lambda_{i..} \) and define it as:

\[
\lambda_{i..} = \sum_{j=1}^{n_c(i-1)} \lambda_{ijk}
\]

\( \lambda_{111} \) is a base case and is the data traffic by an aggregator at level 1 directly from smart meters (leaves at level 0) connected to them and equals

\[
\lambda_{111} = \sum_{j=1}^{N(0,1)} P_k \frac{1}{7.5 \cdot t} \text{ bps}
\]

based on the data generation rate for each smart meter from Equation 5 (without the multiplier \( z \) for all concentrators \( j \) at level 1. \( P_k \) is the packet size used for data sent in cluster \( k \).

Finally, the overall data traffic that is being transmitted from smart meters to the root control center over the distribution grid equals to:

\[
\lambda_{\text{Total}} = \sum_{j=1}^{N(0,k)} n_c(0) + \sum_{j=1}^{N(1,k)} n_c(1) \lambda_{1jk} + \cdots + \sum_{j=1}^{N(h,k)} n_c(h) \lambda_{hjk}
\]

\[
= \sum_{i=0}^{h} \sum_{j=1}^{N(i,k)} n_c(i) \lambda_{ijk}
\]

VI. CONCLUSION

Several technologies have been suggested to meet communication needs at the distribution level. This work compared these technologies and provided a justification of why wireless communication technologies may be most suitable. A linear chain wireless communication architecture was proposed and evaluated for its ability to meet communication requirements of scalability and latency. Subsequently, a theoretical study was done to understand the capacity limitations of using linear chain wireless technologies at the distribution level. The application of advanced metering infrastructure (AMI) that encourages customer participation in energy delivery was used as a case study to understand the implications of any limitations imposed by the proposed communication architecture.

Though one specific communication architecture for the distribution level of the power grid was recommended in this paper, an optimal architecture may vary widely. For example, at the distribution level of the power grid, engineers could choose to utilize a configuration in which short-range wireless communication technology is used to send data from sensors installed on electric poles to a master node on the pole (so that no wire is needed between the sensors and router) and use wired technologies like DSL, Fiber Options (FTTX), or Power Line Communication (PLC) to feed the data back to the control center. Alternative configurations may allow the smart meters to send the data directly back to collection point without involving any wired technology. In response to such a wide range of possible network and system configurations, the best practice begins with thorough and careful system architecture and configuration design. After this process, several distinct configurations that involve wired or wireless technologies at varying degrees may be derived. Therefore, apart from proposing a specific architecture based on some common applications in this paper, this paper also contributes towards a step in the direction of providing a suite of communication topologies and an analysis of their technical feasibility, optimal configuration, and deployment considerations.

ACKNOWLEDGEMENT

This work was supported in part by the Power Systems Engineering Research Center (PSERC), as well as contributions from the industrial and academic members of the PSERC S-54 Project, Towards a Privacy-Aware Information-Sharing Framework for Advanced Metering Infrastructures. We would also like to thank Dr. Visvakumar Aravinthan at Wichita State University for his help with our prior work in this area that laid the foundation for this work.
REFERENCES


