Wireless Network Design for Transmission Line Monitoring in Smart Grid

Benazir Fateh, *Student Member, IEEE*, Manimaran Govindarasu, *Senior Member, IEEE*, and
Venkataramana Ajjarapu, *Fellow, IEEE*

***Abstract*—In this paper, we develop a real-time situational awareness framework for the electrical transmission power grid using Wireless Sensor Network (WSN). While WSNs are capable of cost efficient monitoring over vast geographical areas, several technical challenges exist. The low power, low data rate devices cause bandwidth and latency bottlenecks. In this paper, our objective is to design a wireless network capable of real-time delivery of physical measurements for ideal preventive or corrective control action. For network design, we formulate an optimization problem with the objective of minimizing the installation and operational costs while satisfying the end-to-end latency and bandwidth constraints of the data flows. We study a hybrid hierarchical network architecture composed of a combination of wired, wireless and cellular technologies that can guarantee low cost real-time data monitoring. We formulate a placement problem to find the optimal location of cellular enabled transmission towers. Further, we present evaluation results of the optimization solution for diverse scenarios. Our formulation is generic and addresses real world scenarios with asymmetric sensor data generation, unreliable wireless link behavior, non-uniform cellular coverage, etc. Our analysis shows that a transmission line monitoring framework using WSN is indeed feasible using available technologies. Our results show that wireless link bandwidth can be a limiting factor for cost optimization.**

***Index Terms*—Cyber-physical network design, electric power grid, wireless sensor network.**

I. INTRODUCTION

CURRENTLY, the electric power infrastructure is highly vulnerable against many forms of natural and malicious physical events [1], which can adversely affect the overall performance and stability of the grid. Additionally, there is an impending need to equip the age old transmission line infrastructure with a high-performance data communication network that supports future operational requirements like real-time monitoring and control necessary for smart grid integration [2], [3]. Wireless sensor based monitoring of transmission lines provides a solution for several of these concerns like real-time structural awareness, faster fault localization, accurate fault diagnosis by identification and differentiation of electrical faults from the mechanical faults, cost reduction due to condition based maintenance rather than periodic maintenance, etc. The use of sensor networks has been proposed for several applications like mechanical state processing [4], [5] and dynamic transmission line rating applications [6].

These applications specify stringent requirements such as fast delivery of enormous amount of highly reliable data. The success of these applications depends on the design of a cost-effective and reliable network architecture with a fast response time. The network must be able to transport sensitive data such as current state of the transmission line and control information to and from the transmission grid. This research provides a cost optimized framework to design a real-time data transmission network. Our objective is to formulate a communication framework to transport enormous amount of sensitive data at the time scale of Supervisory Control and Data Acquisition (SCADA) cycle.

To monitor the status of the power system in real-time, sensors are put in various components in the power network [7]–[9]. These sensors are capable of taking fine-grained measurements of a variety of physical or electrical parameters and generate a lot of information. Delivering this information to the control center in a cost efficient and timely manner is a critical challenge to be addressed in order to build an intelligent smart grid [2]. Network design is a critical aspect of sensor based transmission line monitoring due to the large scale, vast terrain, uncommon topology, and critical timing requirements. The goal is to deploy multiple different sensors in critical locations of the transmission line to sense mechanical properties of its various components and transmit the sensed data through a suitable wireless network to the control center. At the control center, it can be combined with existing electrical data in the system to arrive at an ideal preventive or corrective control decision. We design a hybrid hierarchical network that spans wired, wireless and cellular technologies to provide cost optimized delay and bandwidth constrained data transmission. Further, we present the feasibility analysis considering various practical issues concerning the deployment and operation of the sensor network.

This paper is organized as follows. Section II explains the related work followed by sensor network design in Section III. Section IV presents evaluation studies and Section V concludes the paper.

II. RELATED WORK

Managing the communication burden and resulting data latency is essential for efficient analysis and fast control responses and calls for distribution of intelligence throughout the infrastructure [2], [10]. Given the vast geographical expanse of the

Manuscript received April 23, 2012; revised August 28, 2012, November 30, 2012; accepted January 03, 2013. Date of publication April 18, 2013; date of current version May 18, 2013. This research was funded in part by the Electrical Power and Research Center (EPRC), Iowa State University. Paper no. TSG-00228-2012.

The authors are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50010 USA (e-mail: benazir@iastate.edu; gmani@iastate.edu; vajjarapu@iastate.edu).

Digital Object Identifier 10.1109/TSG.2013.2241796

transmission line infrastructure, wireless networking presents a feasible and cost effective solution for transmission of monitoring data [3]. Several works [11]–[15] and [16] propose to improve the state of the art in transmission line monitoring by harnessing the power of wireless sensor networks for real-time monitoring and control. In these works, the goal is to deploy multiple different sensors in critical and vulnerable locations of the transmission line to sense mechanical properties of its various components and transmit the sensed data through a suitable wireless network to the control center. However, most of these works address this theme at a very high level of abstraction. Small scale real world deployments of wireless sensors include tension monitoring using load cells [17], power donut for conductor surface temperature monitoring [18], sagometer [19], etc.

The authors of [13] and [14] were the first to propose a two level model specifically for supporting the overhead transmission line monitoring applications. But considering the topological constraints posed by the transmission lines, the low bandwidth, low data rate wireless nodes would fail to transmit huge amount of data in a multihop manner. The hierarchical model proposed in [11], offers a very expensive solution with the idea of deploying cellular transceivers on every tower. While such a network can provide extremely low latency data transmission, this model is highly cost inefficient as it incurs huge installation and subscription costs. The only work that addresses the problem of finding optimal locations of cellular transceivers is presented in [12].

In [12] authors develop a quadratic equation based solution for finding the optimal locations of cellular transceivers aiming to minimize the delay in information delivery. We contrast this work on the following grounds:

- The formulation presented in [12] relies heavily on symmetry. The underlying network infrastructure and the cellular infrastructure is assumed to be symmetric and available at all times. Further, it is assumed that all transmission towers are identical and transmit the same amount of data. However, several factors bring in asymmetry as enumerated below:
  - Sparse cellular coverage (due to unavailability of cellular towers in the area) or cellular outage.
  - Variation in the amount of data transmitted by the towers in lieu of its location or situation.
  - Irregular terrain in certain regions of the transmission line might prevent the usage of any wireless device forcing the use of only cellular network.
- The analysis presented in [12] considers minimizing delay as an objective. While cost consideration is mentioned in the paper, deployment and maintenance costs are not used as factors restraining the number of cellular transceivers. In the absence of cost constraints, the latency can be minimized by putting a cellular link on each tower which can be highly cost inefficient.
- The method presented in [12] utilizes a quadratic equation to find out the number of cellular enabled towers required and subsequently the location of such towers. Roots of quadratic equation are rounded off to depict the number of cellular enabled towers, which must be an integer. This rounding off can lead to incorrect results. Also, the factors such as latency and bandwidth which affect the placement of cellular transceivers (referred to as representative node in [12]) in a group are not considered simultaneously, rather bandwidth constraints are considered as an afterthought. This leads to suboptimal results.

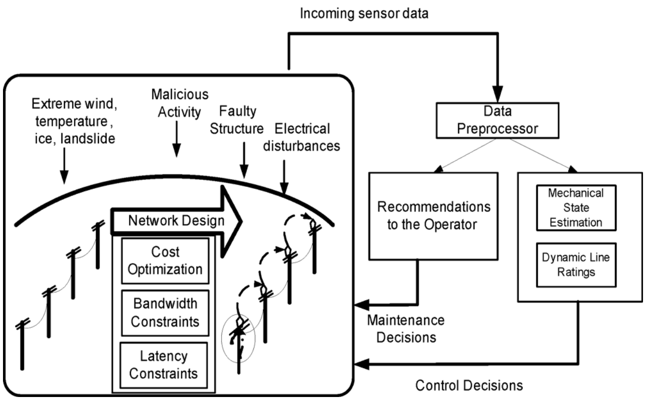

Fig. 1. Sensor network design framework.

In this paper, we propose an optimal solution that minimizes the installation and maintenance costs while satisfying all the constraints such as latency and bandwidth. We present a generic formulation that addresses challenges such as asymmetric flow bandwidth requirements, irregular cellular coverage, etc. Further, our proposed method also provides a way to find cost optimized incremental deployment solutions to satisfy newer future specifications.

III. WIRELESS NETWORK DESIGN

The task of designing a robust wireless data communication network involves consideration of various factors such as latency, resiliency, security and bandwidth constraints. While low cost wireless sensor nodes enable large scale deployment and minimal maintenance operation, these low data rate wireless links can prove to be a bandwidth bottleneck when considering the topology of the transmission line network. Transmission towers are deployed in a straight line forming a *linear network* [11] spanning hundreds of miles. An intelligent choice of technologies needs to be made such that the required bandwidth is provided for the intended data to reach its destination in a timely manner.

Fig. 1 shows our proposed framework. It enumerates an array of challenges and constraints associated with monitoring a wide area network like transmission lines. Necessary control or maintenance decisions can be taken once the sensor measurements are validated and the physical structure is critically assessed for the presence of faults.

The linear network topology proves to be a major challenge for wireless network design with respect to latency constraints and bandwidth constraints. Performance evaluation of the linear network model [20] shows that successful delivery ratio of the packets from the nodes far away from the substation is found to be much less than that of nodes near the substation because packets from a farther node have to travel a longer distance and the rate of collision is higher. The effective monitoring of a large transmission line network requires a hybrid communication infrastructure.

TABLE I
TYPES OF TECHNOLOGIES AND THEIR CHARACTERISTICS

Properties	Optical Fiber	Cellular (3GPP LTE)	Wireless (Zigbee IEEE 802.15.4)
Type of Link in the Network	Substations to Control Center (SS, CC)	Transmission towers to cellular towers (k, CC)	Between towers (k, l)or Between tower and substation (k, SS)
Bandwidth	10 Gbps	Uplink 75 Mbps, Downlink 100 Mbps	250 kbps
Delay	$\approx 1\ \mu$sec	$\approx$ 50 msec	$\approx$ 16 msec
Transmission Range	As long as the fiber	100m- 10 km+	10m - 1.5km
Installation Cost	0 since they already exist	$\approx$ 25x-50x	$\approx$ 1x
Operational Cost	$\approx$ 1x	$\approx$ 5x-20x	$\approx$ 2x
Channel Contention	No	No	Yes
Subscription Fee	No	Yes	No

The numerical values presented for installation cost are relative to that of ZigBee and for operational cost are relative to that of Optical fiber.

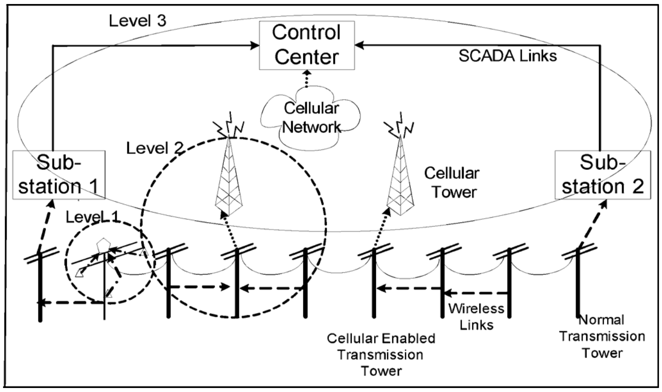

Fig. 2. Hierarchical network structure.

This hybrid infrastructure can be a combination of wired (copper cable/optical fiber) and wireless (cellular/IEEE 802.15.4) standards to enhance the capability of the overall network to meet newer requirements based on emerging smart grid applications.

In this paper, we formulate a hybrid hierarchical network design problem that can provide cost effective data transmission while at the same time respecting the bandwidth, delay, and connectivity constraints. We formulate a placement problem to optimize the number and location of the cellular enabled towers to significantly reduce the operational and installation costs while respecting all the constraints.

A. Three Level Hierarchical Network

We propose a hierarchical three level wireless network model for time critical applications. Each level is equipped with an array of sensors and transceivers with varied capabilities such that together they achieve the required behavior. The design involves the installation of a private WSN of low cost, low data rate links, utilization of the existing SCADA network, and a wide area network such as cellular network comprised of expensive but high data rate links. The proposed network makes use of the existing SCADA links (optical fiber) for communication between substations and control center and strategically utilizes the existing cellular network for data transmission from certain transmission towers directly to the control center. A set of wireless sensors on each tower is installed as part of the private WSN.

Fig. 2 depicts a power transmission corridor with a number of transmission towers, two substations, one at each end of the transmission line, and a control center. Each level of the network forms a cluster supporting *many to one* communication from all the nodes in the cluster to the cluster head.

The first level of the network is responsible for collecting information about the tower. It is composed of sensor nodes installed in each transmission structure forming a sensor array in tower (SAT). This SAT consists of an array of sensor modules such as tension sensors, accelerometers, temperature sensors, tilt sensors, motion sensors, vision-based sensors, and infrared sensors, etc., similar to [13]. Each tower is equipped with a more sophisticated relay node with enhanced computation and communication capabilities. Data from each sensor in the SAT is transmitted to the relay node. The relay node is responsible for compressing the data received from the SAT and transmitting it to the higher level.

The second level of the network is responsible for transmission of data from towers that are far away from the substations. Consider a segment composed of a few towers in the middle of the transmission line network. Data from these towers cannot reach either of the substations due to limited bandwidth of the intermediate wireless links. In such cases, enabling one of these towers with cellular capability can provide a feasible solution as shown in Fig. 2. It is to be noted that it is not required to enable all towers with cellular technology as proposed in [11]. The second level is thus composed of segments of such towers transmitting their aggregated information to the cellular enabled transmission tower which acts as the head of their segment. The cellular enabled tower is a transmission tower equipped with an additional cellular transceiver along with the relay node. This cellular transceiver offers an alternative way to deliver the tower's data directly to the control center through a high bandwidth, low latency cellular network.

The third level of the hierarchical network is composed of a single cluster consisting of two substations and the cellular towers. The control center acts as the cluster head. Thus, level 1 operates at each tower, level 2 operates at the granularity of a group of towers. The size of the group will be dictated by the wireless link bandwidth and the required end to end latency. Level 3 operates at the level of the whole network where substations and cellular towers transmit to the control center. Table I summarizes the characteristics of various communication standards used in this paper.

B. Placement Problem Formulation

In order to provide cost optimized operation in delay constrained and bandwidth constrained linear networks, the

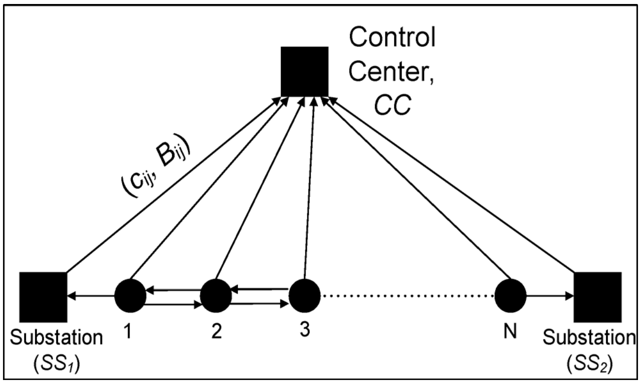

Fig. 3. Placement graph.

TABLE II
SYMBOLS USED IN PLACEMENT PROBLEM FORMULATION

Symbol	Representation
D	End-to-end deadline
l_{ijk}	Latency incurred by the kth flow on link $(i,\ j)$
B_{ij}	Total bandwidth of the link $(i,\ j)$
b_k	Flow bandwidth for node k
c_{ij}	Operational cost incurred by link $(i,\ j)$
IC	Installation cost per cellular transceiver
$S_{i,j}$	Binary Variable. Is 1 if link$(i,\ j)$ is used by any flow. Indicates presence of a flow on link $(i,\ j)$.
Y_i	Binary Variable. Is 1 if tower i is cellular enabled. Indicates presence of a cellular transceiver on tower i.
$X_{i,j,k}$	Binary Variable. Is 1 if node k selects edge $(i,\ j)$ as a link. Indicates presence of kth flow on link $(i,\ j)$.

strategic placement of cellular transceivers becomes very crucial. While the cellular transceivers provide low latency, high bandwidth links, they are costly to install and the subscription charges can dominate the operational cost of the network. On the other hand, the wireless Zigbee devices are relatively inexpensive but provide very low data rates. Thus, there is a tradeoff between cost and delay. In this section, we first explain our network model and state the placement problem. Next, we formulate a mathematical program to find the optimal location of the cellular enabled towers.

1) Network Model: In this paper, the transmission line is modeled as a directed graph, $G = (V, E)$ as shown in Fig. 3. V represents the set of vertices and E represents the set of edges in G. The set of vertices consists of N transmission towers, two substations (SS) at the ends of the transmission line and a control center (CC). Thus, the total number of vertices in the graph is equal to $N + 3$. Edge set E in G represents the set of communication links which can be wired (SS, CC), cellular (k, CC) or wireless (k, l), where $k, l \in N$. Each link (i, j) in the graph can be described by a tuple (c_{ij}, B_{ij}), where c_{ij} is the operational cost incurred by the link and B_{ij} is the total bandwidth of the link. Given that each transmission tower needs to transmit its monitoring data to the control center, there are N flows in this network with each one having a common destination, CC and common delay constraint, D. Each flow can be characterized by the tuple: $F = (\mathrm{Source}, \mathrm{Destination}, b, D)$ where, $\mathrm{Source} \in N$ is the source node of the flow, $\mathrm{Destination} = \mathrm{CC}$ is the destination node of the flow, b is the flow bandwidth requirement, and D is the delay constraint to be satisfied by the flow. l_{ijk} is latency incurred by the kth flow on the link (i, j). The latency incurred is the sum of transmission delay and channel access latency.

2) Placement Problem Statement: Given a directed graph $G = (V, E)$ and a set of N flows, find a feasible path for each flow such that the sum of the cost of all the paths is minimized while respecting the delay and bandwidth constraints of each flow. If the minimum cost path chosen by a tower node k ∈ N includes the edge (k, CC), then a cellular transceiver should be placed on the tower k.

3) Placement Problem Formulation: The input to the algorithm is the transmission line consisting of N transmission towers and the end to end latency constraint, D. The formulation uses the binary variables $S_{i,j}, Y_i$ and $X_{i,j,k}$. $S_{i,j}$ is 1 if link (i, j) is used by at least one flow, showing that operational costs (c_{ij}) are incurred irrespective of if the link is fully utilized or is underutilized. IC denotes the installation cost of a cellular transceiver on a tower, i. Y_i is 1 only if link (i, CC) is used by any flow which means that a cellular transceiver must be installed on node i. $X_{i,j,k}$ represents the choice made for the link (i, j) by the node k. If node k selects edge (i, j) as one of the link in its path, then $X_{i,j,k}$ is equal to 1 otherwise $X_{i,j,k}$ is equal to 0. All the decision variables are binary variables and hence the formulation is an integer linear program (ILP). We used the ILOG CPLEX 12.2 software [21] to solve the ILP. Table II enumerates the different symbols used in the formulation and their meanings. The placement problem can thus be formulated as

$$\textbf{Minimize}:$$

$$f(S_{i,j}, Y_i) = \tau \sum_{(i,j)\in E} c_{ij} S_{i,j} + \sum_{i=1}^{N} IC \cdot Y_i \tag{1}$$

$$\textbf{Subject to}:$$

$$\sum_{(i,j)\in E} l_{i,j,k} X_{i,j,k} \le D \quad \forall k \in N \tag{2}$$

$$-\sum_{(i,j)\in E} X_{i,j,i} = -1 \quad \forall i \in N \tag{3}$$

$$\sum_{k=1}^{N} \sum_{i=1}^{V\backslash \mathrm{CC}} X_{i,\mathrm{CC},k} = N \tag{4}$$

$$\sum_{(j,i)\in E} X_{j,i,k} - \sum_{(i,j)\in E} X_{i,j,k} = 0 \quad \forall k, i \in N, i \ne k \tag{5}$$

$$X_{i,j,k} - X_{j,CC,k} = 0 \quad \forall j \in SS, \forall k \in N \tag{6}$$

$$\sum_{k\in N} b_k X_{i,j,k} \le B_{i,j} \quad \forall (i,j) \in E \tag{7}$$

$$X_{i,\mathrm{CC},k} - Y_i \le 0 \quad \forall i, k \in N \tag{8}$$

$$X_{i,j,k} - S_{i,j} \le 0 \quad \forall (i,j) \in E, \forall k \in N \tag{9}$$

$$X_{i,j,k}, Y_i, S_{i,j} \in \{0,1\} \quad \forall i, j, k \tag{10}$$

Our objective is to minimize the cost function given in (1). Our cost model consists of two types of costs: installation cost and operation cost. Installation cost is a one-time cost of installing cellular transceivers on selected towers. Operation cost is composed of subscription cost and maintenance cost and is recurring in nature. Wireless and SCADA links are assumed to be owned by the transmission company and hence their operational cost is mainly made up of recurring maintenance costs. Cellular links utilize cellular service provided by a third party. Hence the operational costs for cellular links are made up of a recurring subscription cost to be paid to the third party and maintenance cost. As shown in (1), the total cost is the sum of operational cost of all the paths used for data transmission over the operational period τ and one time cost for installing cellular transceivers on selected towers. Equation (2) restricts the end-to-end latency of every flow to less than or equal to the maximum permissible end-to-end deadline, D. Our proposed formulation is capable of addressing multiple latency requirements. Firstly, consider cases where multiple latency requirements are imposed throughout the operational period. In such cases, the constant deadline D in (2), can be modified to flow specific deadline, D_k. Secondly, consider cases where multiple latency requirements are imposed for only a fraction of the operation time. In such cases, a proactive planning approach can be adopted and resources can be reserved for use during emergency situations demanding higher data rate and lower latency data transfer. Specifically, contingency flows, i.e., flows to deal with traffic contingencies can be created during the planning stage, thereby modifying the total number of flows from N to $N^{'}$. The resources reserved for these contingency flows are utilized during emergency situations.

The latency calculations take into account the transmission latency as well as channel access latency experienced by a flow on each link. In order to address fine grained latency calculations, queuing delay can also be taken into consideration, given that sensor measurements may be buffered at one or multiple nodes before transmission. The queuing delay component can be linearly added to the latency component $l_{i,j,k}$ along with transmission latency and channel access latency. Every flow will utilize a set of wireless links and exactly one cellular or SCADA link of type (i, CC) where $i \in N \cup \{\mathrm{SS}_1, \mathrm{SS}_2\}$. On each link, multiple flows can be multiplexed as dictated by the total link bandwidth. The transmission latency of a flow on each link is calculated considering the presence of other flows multiplexed on the same link.

The constraints in (3)–(6) explain the flow conservation constraints and ensure that exactly one path is selected for a flow generated at node $k \in N$. Equation (3) restricts each tower to be a source of exactly one flow. Equation (4) depicts that CC serves as destination to exactly N flows. Equations (5) and (6) ensure flow conservation at each tower and substation respectively, between the source and destination. Bandwidth of a link (i, j), is denoted by B_{ij} denoting available link specific bandwidth taking into consideration interference on neighboring links. This bandwidth is assumed to be constant over time. We agree that available bandwidth could change due to varying interference levels. However, the proposed method is an offline planning approach and addressing time varying interference is out of the scope of our work. Equation (7) explains that the total flow on each link must not exceed the available link bandwidth. Equation (8) ensures that whenever any link of type (k, CC) is used by any flow, a cellular transceiver will be placed on the tower k ensuing some installation cost. Equation (9) ensures that cost of link (i, j) is counted at most once for the set of k flows that are multiplexed on this link. The last constraint, (10) ensures that the decision variables are binary variables.

Upon interpretation of the output, the location of cellular enabled towers can be found by the value of the decision variable, Y_i. If it is equal to 1, it means that a cellular transceiver needs to be installed on node i.

The operational cost of links of type (SS, CC) is given a minimal value followed by wireless links (k, l) where $k, l \in N$. The operational cost of cellular links (k, CC) are given highest values modeling the high cellular subscription charges. The installation cost for (SS, CC) is set to zero, thus modeling the already present SCADA links and a constant installation cost can be added to the final output to model the fixed installation cost of SAT on each tower. Installation cost of cellular transceivers is added in the objective function. The edges of type (SS, CC) have the least latency amongst all the edges, followed by the cellular links (k, CC), followed by the Zigbee links (k, l) which have the highest latency.

C. Link Utilization Based Costs

In our proposed formulation, we assume a fixed periodic subscription cost being charged for each active cellular link. This model could be generalized to represent subscription costs proportional to the link utilization. The existing formulation can be easily modified to address the variable cost structure. Consider the modified objective function as shown below:

$$\mathrm{f}(S_{i,j}, Y_i, X_{i,j,k}) = \tau \sum_{(i \in N, j \neq \mathrm{CC})} c_{ij} S_{i,j} + \tau \sum_{i \in N, j = \mathrm{CC}} \left(c_{ij} \sum_{k \in N} \frac{b_k X_{i,j,k}}{B_{i,j}} \right) + \sum_{i=1}^{N} IC \cdot Y_i \quad (11)$$

Equation (11) computes the total cost as the sum of fixed operational costs for wireless and SCADA links, sum of link utilization dependent operational costs for cellular links and the sum of installation costs. Notice that a fixed operational cost is assumed for wireless and SCADA links, because irrespective of the link selected for sensor data transmission, these links are operational for other requirements such as data transmission among the sensors in the SAT. Nonetheless, the formulation can easily be modified if the operational costs for all the links need to be made utilization dependent. We analyze the impact of link utilization based costs in Section IV-F.

D. Link Reliability

In the proposed formulation presented before, we assume perfect wireless link conditions. However, in reality, wireless links can be unreliable due to wireless interference, channel loss, multi path fading, etc. [22]. In the following, we present a method to extend our existing formulation to address link unreliability.

Let Rel denote path reliability that can be specified as part of input requirements. This means that each flow must reach the control center, CC with probability Rel. The path reliability Rel is related to the constituent link reliabilities. Consider a link l with link reliability, ρ_l. For a path with n links, the reliability of the path can be calculated as ρ_l^n. Thus, for every flow the condition $\rho_l^n \geq Rel$ must be satisfied. Given link reliability and path reliability, n can be found offline as the highest integer satisfying $\rho_l^n \geq Rel$. This n represents the maximum number of links that a flow can traverse, in order to satisfy path reliability constraint. Thus, wireless link reliabilities can be easily addressed by our existing formulation with the addition of the following constraint:

$$\sum_{(i,j)\in E} X_{i,j,k} \leq n, \quad \forall k \in N \tag{12}$$

Equation (12) ensures that the total number of links traversed by each flow, k must be less than or equal to n.

Link specific reliabilities can also be addressed in our proposed formulation. However, it introduces a quadratic constraint modifying the linear optimization program to a quadratically constrained program. Specifically, consider that the link reliability of link (i, j) is denoted as ρ_{ij}. Then the equation (12) shown above will change to

$$\prod_{(i,j)\in E} \rho_{ij} X_{i,j,k} \geq Rel, \quad \forall k \in N \tag{13}$$

As evident, addressing link specific reliabilities results in a quadratically constrained optimization problem which is much harder to solve. Thus, link specific reliability can be addressed at the cost of higher complexity.

E. Constrained Cellular Coverage

In the previous section, we assumed that the transmission line considered is uniformly covered by a cellular communication network. This means that cellular transceivers can be placed on any tower. However, owing to the diverse geographical terrains traversed by the long transmission lines, there might be remote areas where cellular coverage is not available. Or there could be prolonged outage on certain cellular towers. In such cases, there are additional constraints on the placement of cellular transceivers. We refer to this version of the problem as *Coverage Constrained* placement problem where relay nodes can only be installed on a subset of the transmission towers which are covered by cellular service.

The *Coverage Constrained* placement problem formulation requires little modification from our original placement problem formulation. Let $V^{'}$ denote the set of nodes not covered by the cellular service. The constrained edge set, $E^{'}$ will be $E - \{(k, \mathrm{CC}),\ \forall k \in V^{'}\}$. In order to model the *Coverage Constrained* placement problem, the input graph G needs to be replaced by $G^{'}(V, E^{'})$. Fig. 4 shows the example where towers 3 and 4 do not have any cellular coverage. The edges (3, CC) and (4, CC) are thus removed. In the formulation, the corresponding binary variables, Y_3 and Y_4 are removed to represent cellular unavailability at these towers. Thus, with a simple modification in the formulation, restricted cellular availability can be easily addressed.

If there is a large segment of towers devoid of cellular coverage, then the responsibility of data delivery falls upon wireless Zigbee links. In such cases, the data delivery might not be able to meet the latency constraints given the limited capacity of wireless Zigbee links. Such cases can be easily identified though our proposed method as shown in Fig. 10.

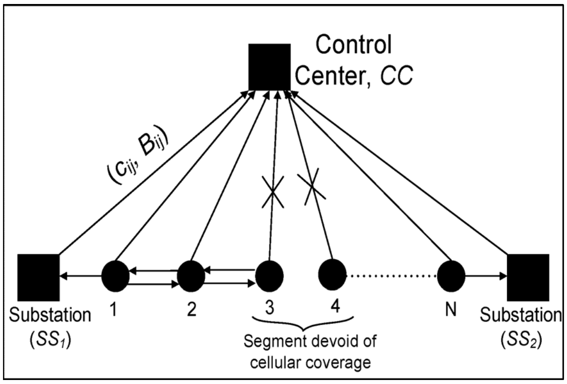

Fig. 4. Placement graph for constrained cellular coverage.

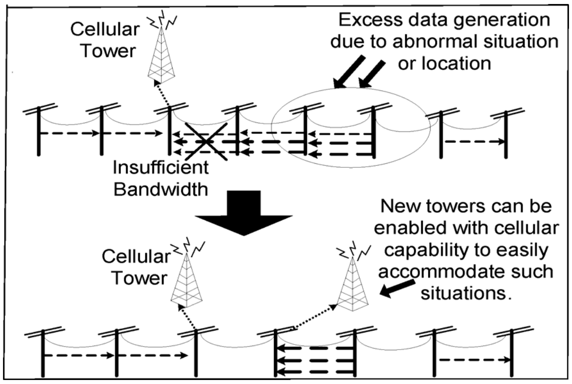

Fig. 5. Accommodating asymmetric flow requirements.

F. Asymmetric Data Generation

The method proposed in [12] relies heavily on symmetry and hence cannot accommodate asymmetric flow bandwidth requirement in the network. There can be several scenarios leading to towers generating sensor data at different rates. This can be due to a requirement of fine grained sensor measurement in order to attain better situational awareness of a particular tower located in a sensitive area. Fig. 5 shows such a scenario. Our proposed formulation can easily accommodate such asymmetric requirements. This is because each flow b_k generated at a tower k is individually formulated. Asymmetric data generation will change the values of b_k and l_{ijk} in (2) and (7). Hence the formulation solves both symmetric and asymmetric cases equally the only difference being in the input file specifying the flow requirements.

G. Incremental Deployment

With changing data traffic requirements and geographical expansion of the transmission line, the possibility of incremental deployments is always present. If at each such future requirement, an entirely new input is given to the optimization formulation without taking into account the currently existing cellular enabled towers, the solution may be a costlier deployment. We term this method as *memoryless deployment* because it discards any memory of existing cellular enabled towers. This method can result in installation on an entirely new set of towers, thus risking loss of any investment made in installing existing cellular transceivers in the first place.

We present a method to add new cellular links on top of existing network to satisfy newer requirements while minimizing installation costs. In order to do so, only the input to the optimization program needs to be modified. The installation costs for currently deployed towers is made equal to 0 and flow requirements are increased as per new specifications leading to modified latency calculations. Since the installation costs of currently deployed towers is now 0, they will be automatically considered as cellular enabled towers in the optimized answer. The optimization formulation will make as much use of such existing cellular enabled towers as possible in minimizing total costs. Thus, *incremental deployment* reuses the existing cellular enabled transceivers as much as possible.

IV. PERFORMANCE EVALUATION

We consider a transmission line network with 75 towers with an average span length of 800 ft [23]. In order to reflect real-world scenarios, the bandwidth of the optical fiber links (SS, CC) is taken as 10 Gbps. Bandwidth of the cellular links is taken as 75 Mbps and latency incurred in the cellular link due to state transition delay, access delay, and handover, etc., is taken as 50 ms [25]. The bandwidth of the IEEE 802.15.4 wireless links is 250 kbps and latency incurred due to these links is 16 ms [24]. The length of the data packet generated by each tower is 32 kbits [11]. The performance metric of interest is the total cost of the network including the installation and operational costs. We consider three pricing schemes named C_1, C_2, and C_3 and analyze the effect of different costs associated with each type of link present in our network. They vary in their ratio of operational costs attached to each type of link. A pricing scheme can be described as a ratio of operational costs of optical fiber to cellular to Zigbee. Thus, a scheme (1:10:2) would mean that operational cost of the three types of links: optical fiber, cellular and Zigbee are in the ratio 1:10:2.

We study several different scenarios including variation in flow bandwidth, end to end deadline and network size. We compare the results of our proposed formulation called Integer Linear Program (ILP) with the method proposed in [12] referred to here as the Quadratic Equation method (QE). We also evaluate the cost of the network in cases of constrained cellular coverage and incremental deployment.

The proposed formulation provides an optimal solution to the wireless network design problem at the granularity of a transmission corridor as shown in Fig. 2. This solution can be independently computed for each corridor. While computationally expensive, these calculations are required to be done only during the offline centralized network planning stage. Thus, the fact that an ILP formulation is expensive to compute is mitigated by the small number of times such an optimization needs to be performed. However, for cases where the transmission corridor comprises of several hundreds of transmission towers, heuristic methods may be better suited. This is because even though they compromise on the accuracy of the result, they are capable of providing a solution with much less computing resource requirements. Although the ILP model is harder to scale and computationally-intensive, it is helpful in determining a lower bound on the costs. We use the ILOG CPLEX 12.2 software [21] to solve the proposed Integer Linear Program (ILP). The simulations were run on Intel(R) Xeon(R) X5650@ 2.67 GHz machines. The simulations took a minimum of 0.16 seconds and a maximum time of 4.25 hours.

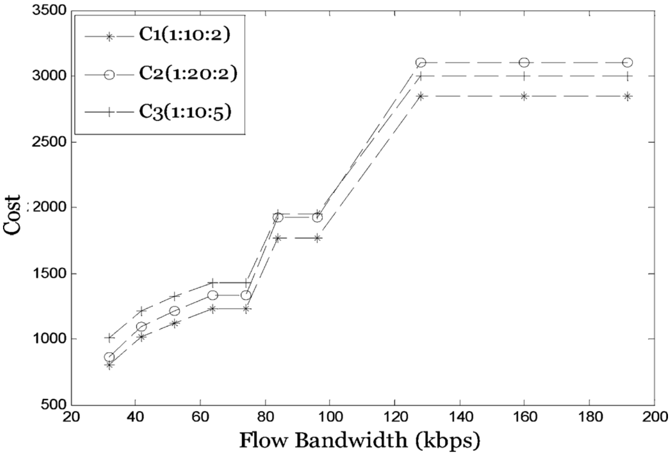

Fig. 6. ILP: Effect of variation in flow bandwidth.

A. Effect of Variation in Flow Bandwidth

Fig. 6 shows the effect of the amount of data generated by each tower and its effect on the feasible operation of the transmission line. In this simulation, we consider a 75 node network with a deadline constraint of 3 s. The packet size of the sensor data generated at each tower is the same. Given constant bandwidth of the wireless links, only a certain number of flows can be multiplexed on any link. Thus smaller the flow bandwidth requirement, more flows can be multiplexed on each link. Given a large deadline requirement, this results in reducing the number of cellular links to be used and hence the cost. The performance graph echoes this observation. Notice in the graph that for values of flow bandwidth greater than or equal to 128 kbps (specifically 128 kbps, 160 kbps, and 192 kbps), the cost becomes constant. This is because given the wireless link bandwidth of 250 kbps, at most one flow can be multiplexed on each link. Thus the network design remains same for each of these three flow bandwidth requirements. Also in cases where, flow bandwidth is greater than wireless link bandwidth, then the remaining options are either deploying an all cellular or all wired solution.

Fig. 7 compares the results given by our proposed algorithm, ILP and the proposed method (QE) in [12] with respect to variation in flow bandwidth. As mentioned earlier, the QE method [12] utilizes a quadratic equation to obtain the number of cellular enabled towers. Roots of quadratic equation are rounded off to the nearest integer to depict the number of cellular enabled towers which must be an integer. This rounding off leads to incorrect results as can be seen in the plotted curves. Once flow bandwidth is greater than 128 kbps, the results should be the same for the three cases (128 kbps, 160 kbps, and 192 kbps) as explained previously. However, the QE method tends to give incorrect results owing to errors encountered due to rounding of roots. Similarly for flow bandwidths of 84 kbps and 96 kbps, the cost incurred by QE method is less, but that is because number of towers selected by the QE method are insufficient leading to constraint violation.

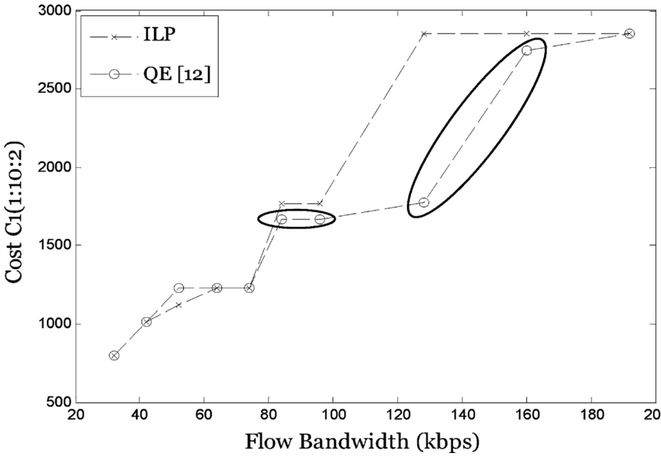

Fig. 7. Comparison between ILP and QE [12] with respect to variation in flow bandwidth.

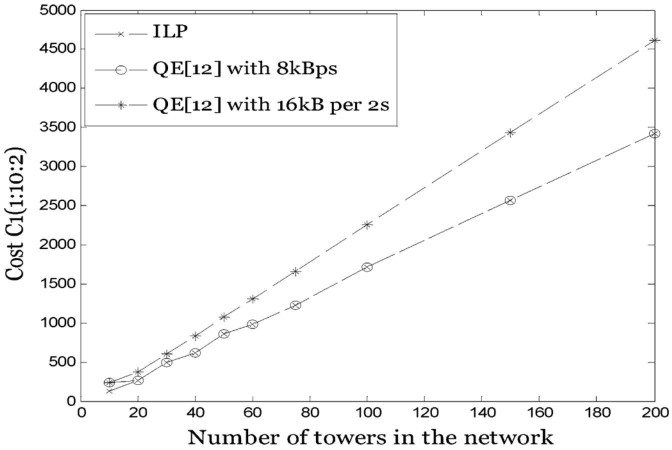

Fig. 9. Comparison between ILP and QE [12] with respect to variation in the number of towers in the network.

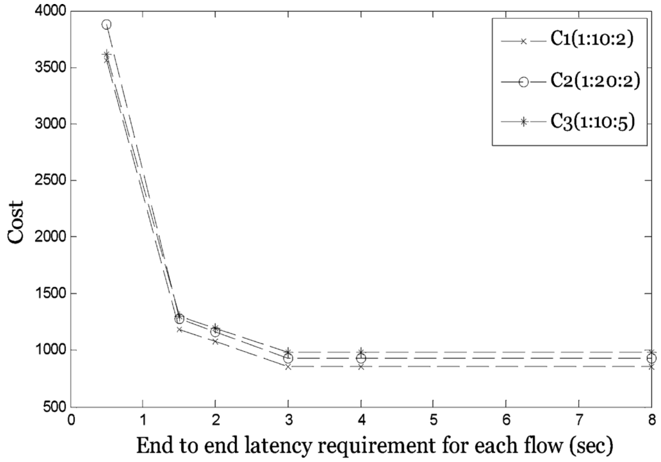

Fig. 8. ILP: Effect of variation in end to end flow latency.

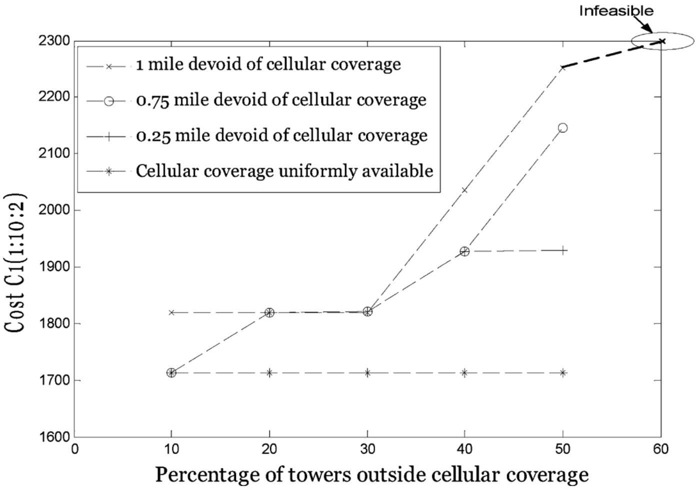

Fig. 10. ILP: Capable of addressing variations in cellular coverage.

B. Effect of Variation in Flow Latency

Fig. 8 shows the effect of variation in end to end flow latency with respect to cost. We consider the time scale of one SCADA cycle which is 4–8 s [13]. The results show that in cases of very stringent deadline requirement (≈ 0.1 s), a cellular transceiver should be installed on each tower. Thus every tower uses the cellular link to avoid any deadline miss hence ensuing a huge cost. In the given 50 node network, for relatively relaxed deadline requirements ($\approx$2–4 s) the lowest cost is attained by fully utilizing the wireless network. At the time scale (≥ 3 s) the cost becomes constant because now the system is more bandwidth limited than latency limited.

C. Effect of Network Size

Fig. 9 shows the effect of variation in the number of transmission towers with respect to the cost. Given the linear structure of the transmission line, the cost increases approximately linearly with respect to the number of towers in the network. We assume all towers generate 64 kbits of data per monitoring cycle [11] and the end to end deadline is 8 s. We compare the results of ILP with two data generation rates in QE one at 8 kBytes per second and one at 16 kBytes per 2 seconds both being equivalent to 64 kbps. The QE solution results in higher cost for the case of 16 kBytes per 2 seconds. For any given network, this graph can be used to find the most cost effective solution by plotting the various pricing curves.

D. Effect of Variation in Cellular Coverage

Fig. 10 shows the effect of variation in cellular coverage and its effect on the installation and operational cost of the network. We consider a 100 node network with constant flow bandwidth of 64 kbps and a deadline of 8 s. We vary the percentage of nodes which can be devoid of any cellular coverage from 10% to 50% of the network size. Further, we vary the length of the segment of the transmission line devoid of any cellular coverage from 0.25 mile to 1 mile. This is done to mirror real world situation where it is more likely that adjacent towers are devoid of cellular coverage rather than random single towers. For comparison, the results of cellular constrained scenarios are compared with the symmetric case where coverage is available throughout the network. In cases where a large number of towers are devoid of cellular coverage, there might not be any feasible solution as shown in the figure. An analysis of this nature helps in finding the feasibility of the wireless option in the cellular constrained areas.

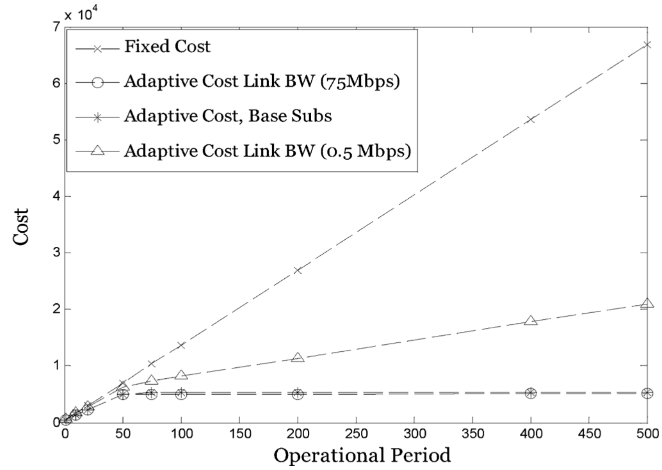

Fig. 11. Effect of variation in operational period.

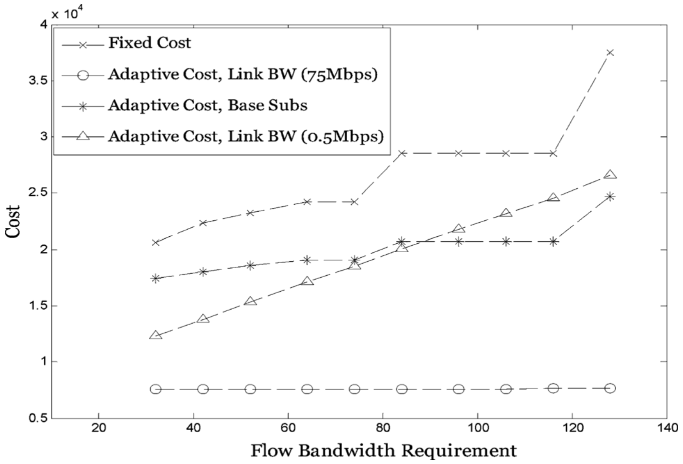

Fig. 12. Effect of link utilization.

E. Effect of Operational Period

Fig. 11 shows the effect of operational period on the total costs (initial installation costs and operational costs over the operational period). In case of fixed operational costs, total costs increase rapidly with increasing operational period. In case of adaptive costs depending upon link utilization the total costs reduce dramatically. Also in this case, if the available link bandwidth is much higher as compared to bandwidth utilization, then operational costs turn out be very less and over time the initial investment on installation is recovered. Thus it can be profitable to deploy a cellular transceiver on a bigger subset of towers. However, if cellular bandwidth is relatively less, making the link utilization based operational costs a bigger part of the total costs, then the number of cellular towers needs to be optimized. This graph illustrates how different factors such as available link bandwidth, operational period and cost ratios affect minimum cost network design.

F. Effect of Link Utilization Based Cost

Fig. 12 shows the impact of link utilization based *adaptive* cost models on total network design cost. In this simulation, we consider a 75 node network with an operational period of 100. We consider 4 cost models explained as follows.

The *Fixed Cost* model considers fixed subscription cost ignoring adaptive link utilization. As evident, it incurs highest costs because the number of deployments increase with increasing flow bandwidth requirement.

The *Adaptive Cost, Link BW(75 Mbps)* model is the link utilization based cost model where cellular bandwidth is 75 Mbps incurring the lowest costs. The optimal solution here is to deploy all towers with cellular capability. This solution is optimal due to two factors. One, the subscription costs become negligible as compared to wireless operation costs given the high cellular bandwidth availability as compared to utilization. Secondly, a huge initial investment can be easily recovered given a large operational period. The negligible subscription cost acts as an incentive to initially deploy all towers with cellular capability.

In order to consider the effect of available cellular bandwidth, we consider an *Adaptive Cost, Link BW(0.5 Mbps)* model with a smaller available link bandwidth of 0.5 Mbps. In this case link utilization based subscription costs are non negligible resulting in increased cost.

We consider another cost model, *Adaptive Cost, Base Subs* where in addition to the link utilization based adaptive costs, a *fixed base subscription fee* is charged per active cellular link per period. The available cellular bandwidth is 75 Mbps. Thus subscription costs are negligible but base subscription fee acts as a major factor in cost calculations. When compared with *Adaptive Cost, Link BW(0.5 Mbps)*, the tradeoff can be observed. At lower flow bandwidth requirements, *Adaptive Cost, Base Subs* incurs more cost due to overpowering base subscription fee. But at higher requirement, link utilization based costs result in more costs by *Adaptive Cost, Link BW(0.5 Mbps)*.

G. Effect of Link Unreliability

Fig. 13 shows the impact of link unreliability on network design cost. In this simulation, we consider a constant flow bandwidth of 32 kbps and a deadline constraint of 3 s. At lower link reliabilities, a path should consist of lesser number of links to maintain path reliability constraint. This leads to more cellular towers being deployed resulting in higher costs. As link reliability increases, more wireless links can be utilized resulting in cost reduction. After a certain point (maximum links $\geq$6), any further increase in link reliability does not affect cost reduction. This is because other optimization constraints such as limited bandwidth and latency limit the number of flows per link and hence the number of links per path.

H. Effect of Incremental Deployment

We perform experiments on a network of 50 towers with each tower generating sensor data at the rate of 32 kbps initially. The data generation rates are then gradually increased upto 128 kbps to mirror the increasing bandwidth demands in the future. Fig. 14 shows the cumulative costs incurred by the two methods: incremental deployment and memoryless deployment when the operational period is equal to one. Memoryless deployment starts with a clean slate each time a new requirement comes in. Due to this, memoryless deployment ends up installing a cellular transceiver on a much bigger subset of towers. Incremental deployment modifies the input to the optimization and adds new links on top of existing network such that the existing design is utilized as much as possible incurring lesser deployment costs.

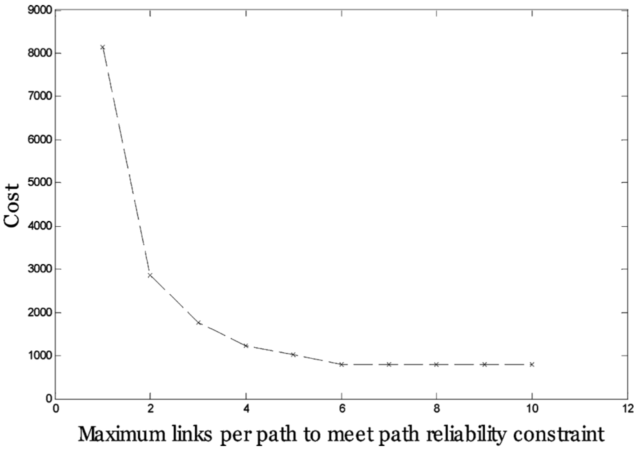

Fig. 13. Effect of link unreliability.

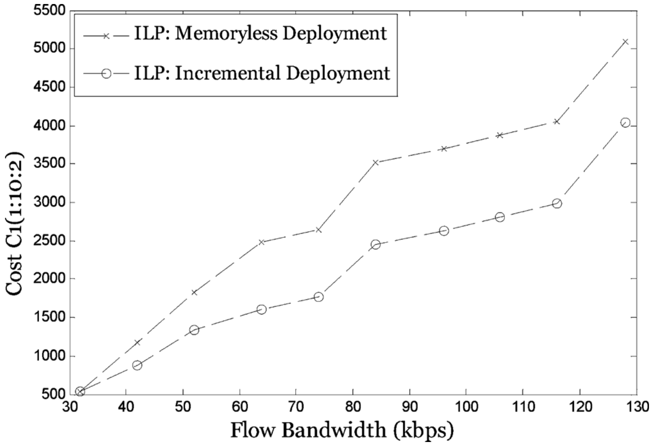

Fig. 14. ILP: Capable of cost optimized incremental deployment.

Please note that both deployment methods are variations of our proposed formulation producing optimal solutions depending on the input being provided to them. Our formulation offers the flexibility to perform different types of analyses. Factors such as operational period and link utilization based costs affect the performance of the two methods. Consider Fig. 15, where operational period is varied from 20 to 350. We consider a cost model [C1(1:10:2)] with the variation that each active cellular link is charged a base subscription fee and link utilization based subscription cost instead of fixed subscription cost. Each point on the graph, represents the total costs of an instance where a set of data requirements (starting from 32 kbps to 128 kbps) was operational for a period τ. Each value on the curve is the total cost of memoryless deployment normalized with respect to the respective value for incremental deployment to show their relatively close but different performance.

In Fig. 15, we observed that at low operational periods, memoryless deployment incurs more cost than incremental deployment. This is because memoryless deployment ignores any investment made in the earlier deployment and a small operational period does not allow enough time to recover from that investment. Incremental deployment on the other hand, avoids extra installations by making the most use of the already deployed towers. At higher operational periods, the installation costs can be recovered and hence memoryless and incremental deployments produce solutions with equal costs.

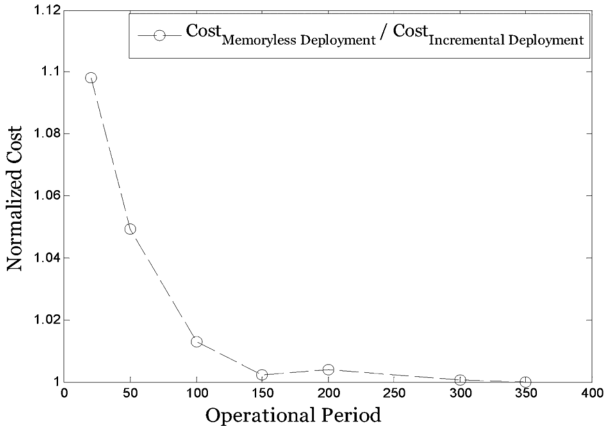

Fig. 15. ILP: Comparing Memoryless deployment with Incremental deployment.

I. Modeling Multiple Constraints

In the proposed formulation, we presented a generalized formulation in order to encompass varied application scenarios. However, it can be difficult to model every operational constraint. Mainly, the most important constraints that need to be modeled for a feasible network design are flow latency and flow bandwidth constraints. This is because Zigbee links can provide the least expensive communication, but their limited bandwidth proves to be a major bottleneck. Also the topology of transmission lines presents critical challenges towards low latency communication being achieved solely by Zigbee links. These are followed by the cellular coverage constraint since extreme unavailability of cellular coverage might render the network design infeasible. Other constraints such as link reliability, link utilization based cost structure, asymmetric data generation, etc., are less critical for a feasible cost optimized network design. The results we present in the paper individually analyze the effect of variation in each of the constraint on the cost of the network while other constraint are assumed to take up average values. Specifically, we find that network design is feasible at very low costs at higher latency; and at very high costs at higher flow bandwidth. Further, we find that if link utilization based cost structure is adopted then flow bandwidth requirement might have no effect on the total costs as shown in Fig. 12.

V. CONCLUSION

In this paper, we presented an optimal formulation for a cost optimized wireless network capable of transmission of time sensitive sensor data through the transmission line network in the presence of delay and bandwidth constraints. Our analysis shows that a transmission line monitoring framework using WSN is indeed feasible using available technologies. We compare behavior of our proposed method with the method proposed in [12]. Our proposed formulation is generic and encompasses variation in several factors such as asymmetric data generation at towers, wireless link reliabilities, link utilization dependent costs, non-uniform cellular coverage characteristics and requirements for cost optimized incremental deployment.

Our evaluation studies show that the main bottleneck in cost minimization is wireless link bandwidth. Further, in cases of increasing flow bandwidth, the limited wireless link bandwidth leads to a feasible but expensive design due to increased dependence on cellular network to satisfy constraints. As part of future work, we plan to study cost efficient fault tolerant network design.

REFERENCES

[1] H. Li, G. W. Rosenwald, J. Jung, and C. Liu, "Strategic power infrastructure defense," *Proc. IEEE*, vol. 93, no. 5, pp. 918–933, May 2005.

[2] P. Zhang, F. Li, and N. Bhatt, "Next-generation monitoring, analysis, and control for the future smart control center," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 186–192, Sep. 2010.

[3] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," *Comput. Netw.*, vol. 50, no. 7, pp. 877–897, May 2006.

[4] P. Ramachandran, V. Vittal, and G. T. Heydt, "Mechanical state estimation for overhead transmission lines with level spans," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 908–915, Aug. 2008.

[5] S. Malhara and V. Vittal, "Mechanical state estimation of overhead transmission lines using tilt sensors," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1282–1290, Aug. 2010.

[6] J. Ausen, B. F. Fitzgerald, E. A. Gust, D. C. Lawry, J. P. Lazar, and R. L. Oye, "Dynamic thermal rating system relieves transmission constraint," in *Proc. IEEE 11th Int. Conf. Transm. Distrib. Construction, Oper., Live-Line Maintenance (ESMO)*, Oct. 15–19, 2006.

[7] V. C. Gungor, L. Bin, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3557–3564, Oct. 2010.

[8] S. Ullo, A. Vaccaro, and G. Velotto, "The role of pervasive and cooperative sensor networks in smart grids communication," in *Proc. 15th IEEE Mediterranean Electrotech. Conf. (MELECON)*, Apr. 26–28, 2010, pp. 443–447.

[9] A. Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 11–19, Jun. 2010.

[10] K. Moslehi and R. Kumar, "A reliability perspective of the smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 57–64, Jun. 2010.

[11] K. S. Hung, W. K. Lee, V. O. K. Li, K. S. Lui, P. W. T. Pong, K. K. Y. Wong, G. H. Yang, and J. Zhong, "On wireless sensors communication for overhead transmission line monitoring in power delivery systems," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun. (SmartGridComm.)*, Oct. 4–6, 2010, pp. 309–314.

[12] Y. C. Wu, L. F. Cheung, K. S. Lui, and P. W. T. Pong, "Efficient communication of sensors monitoring overhead transmission lines," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1130–1136, Sep. 2012.

[13] R. A. Leon, V. Vittal, and G. Manimaran, "Application of sensor network for secure electric energy infrastructure," *IEEE Trans. Power Del.*, vol. 22, no. 2, pp. 1021–1028, Apr. 2007.

[14] J. Chen, S. Kher, and A. K. Somani, "Energy efficient model for data gathering in structured multiclustered wireless sensor network," in *Proc. 25th IEEE Int. Perform., Comput., Commun. Conf. (IPCCC)*, Apr. 10–12, 2006, pp. 8–388.

[15] Y. Yang, F. Lambert, and D. Divan, "A survey on technologies for implementing sensor networks for power delivery systems," in *Proc. IEEE Power Eng. Soc. Gen. Meet.*, Jun. 24–28, 2007, pp. 1–8, vol., no.,.

[16] Y. Yang, D. Divan, R. G. Harley, and T. G. Habetler, "Design and implementation of power line sensornet for overhead transmission lines," in *Proc. IEEE Power Energy Soc. Gen. Meet. (PES)*, Jul. 26–30, 2009, pp. 1–8.

[17] "Load cell technology in practice, Revere Transducers", Mar. 2009, http://www.societyofrobots.com/robottheory/loadcellprimer.pdf.

[18] R. G. Olsen and K. S. Edwards, "A new method for real-time monitoring of high-voltage transmission-line conductor sag," *IEEE Trans. Power Del.*, vol. 17, no. 4, pp. 1142–1152, Oct. 2002.

[19] An EPRI Sponsored Technology, EDM International, Inc., Sagometer.

[20] S. Gumbo and H. N. Muyingi, "Performance investigation of wireless sensor network for long distance overhead power lines; mica2 motes, a case study," in *Proc. 3rd Int. Conf. Broadband Commun., Inf. Technol. Biomed. Appl.*, Nov. 23–26, 2008, pp. 443–450.

[21] IBM ILOG CPLEX Mathematical Optimization Technology [Online]. Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimization-studio/cplex-optimizer/cplex-performance

[22] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[23] American Transmission Company, Tech. Rep. [Online]. Available: www.atcllc.com/pdf/MorganWernerWestapplication.pdf

[24] *IEEE Standard for Information Technology-Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LRWPANS)*, IEEE 802.15.4-2006.

[25] Overview of the 3GPP Long Term Evolution Physical Layer [Online]. Available: http://www.freescale.com/files/wirelesscomm/doc/whitepaper/3GPPEVOLUTIONWP.pdf

Benazir Fateh (S'10) received the B.E. degree in electrical engineering from the University of Mumbai, India. She is currently a Ph.D. candidate in the Department of Electrical and Computer Engineering at Iowa State University, Ames, IA, USA. Her Ph.D. thesis focuses on resource management algorithms for wireless sensor networks with applications in cyber-physical systems.

Manimaran Govindarasu (SM'10) received the Ph.D. degree in computer science and engineering from the Indian Institute of Technology (IIT) Madras, India, in 1998. He is currently a Professor and Associate Chair in the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. He received the Young Engineering Research Faculty Award from ISU in 2003. His research expertise is in the areas of cyber-physical systems security for smart gird, networked real-time systems, and network security. He has published more than 100 research publications in international journals and conference proceedings. He is a co-author of the book *Resource Management in Real-Time Systems and Networks* (MIT Press, 2001). He has given tutorials in reputed conferences, including IEEE INFOCOM 2004 and IEEE ComSoc TutorialsNow (2004), and IEEE ISGT (2012). He serves on the editorial board of the IEEE TRANSACTIONS ON SMART GRID and served as guest co-editor for several special publications, including *IEEE Network* (Jan. 2003) and *IEEE Power & Energy* (Jan. 2012). He currently serves as the Chair of the Cyber Security Task Force at IEEE Power & Energy Society, CAMS Subcommittee.

Venkataramana Ajjarapu (SM'91–F'08) received the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1986. Currently, he is David C. Nicholas Professor in the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. His research interests are in the area of reactive power planning, voltage stability analysis, power system security, and wind and solar energy integration.