# Low Cost and Secure Smart Meter Communications using the TV White Spaces

**Omid Fatemieh**, University of Illinois Urbana-Champaign
**Ranveer Chandra**, Microsoft Research, Redmond, WA
**Carl A. Gunter**, University of Illinois Urbana-Champaign

***Abstract*—We investigate the use of *white spaces* in the TV spectrum for *Advanced Meter Infrastructure* (AMI) communications. We provide a design for using white spaces for AMI and show its benefits in terms of bandwidth, deployment, and cost. We also discuss ongoing work on applying machine learning classification techniques to improve the attack resilience of spectrum data fusion in the proposed architecture.**

## I. Introduction and AMI Background

*Advanced Meter Infrastructure* (AMI) is an integral part of the recent smart grid initiatives in the United States. It refers to systems that measure, collect, and analyze energy usage and interact with smart (advanced) meters through some communication media. The reconfigurable nature and communication capabilities of smart meters allow for deploying a rich set of applications in the smart grid. Prime application instances are automated meter reading, outage management, demand response, electricity theft detection, and support for distributed power generation.

The communication architecture for AMI must meet the needs of current and future applications in a cost-effective, scalable, reliable, and secure way. Of particular interest are two-way communications between the smart meters and service providers such as the utility companies. Figure 1 depicts a common approach to AMI communication in the existing deployments. In this model, hundreds to thousands of meters form a mesh network using proprietary protocols in the public industrial scientific and medical (ISM) frequency bands. The mesh network is used to route the data to an access point (often mounted on a telephone pole). The access point aggregates and relays data between the meters and the utility. This part is often performed using cellular data services such as GPRS or EVDO. This approach suffers from at least three shortcomings. First, the ISM bands are noisy and crowded in urban areas and not well suited to the distances needed in rural areas. Second, cellular links incur the extra expense associated with licensed bands. Moreover, there is considerable competition for this bandwidth in urban areas and limited availability in rural areas. Third, the use of proprietary mesh network technology reduces inter-operability and impedes meter diversity.

In this paper we consider the idea of using *white spaces* as part of AMI. White space communications leverage licensed spectrum opportunistically when it is not being used by incumbent transmitters such as digital TV transmitters. We believe the high bandwidth and long transmission ranges offered by white spaces can provide substantial benefits to the AMI. To that end, we propose a two-layer architecture for AMI communication using a combination of standardized protocols and successful research prototypes. We show that the proposed architecture can address some limitations of the state of the art, particularly in terms of bandwidth, deployability, and cost. In addition, we investigate reliability and security issues associated with the proposed architecture.

The two main goals of the paper are to sketch a strategy for using white spaces for AMI and investigate the use of new techniques for addressing 'data fusion' resilience when combining spectrum sensing data. In white space networks, clients report spectrum sensing data about presence of incumbents to the base station which uses this information to decide the TV channels to use for communication. We consider the problem of dealing with malicious nodes that report such data inaccurately, with specific application to AMI. In the proposed architecture, the white space service provider needs to collect spectrum sensing data from multiple sources such as smart meters, mobile units, or equipment on consumer premises (this process is also known as *crowdsourcing* [9]). These devices have varying or unknown degrees of integrity and risk of compromise. It is important that such uncertainties do not disrupt the AMI communications. We consider the idea of building a *classifier* using *training* data representing natural signal propagation in a region as a means to achieve resilient data fusion. More specifically, we use the classifier to identify attacker-dominated (small) regions. This on-going investigation involves using house density data from the US Census Bureau, digital TV transmitter data from FCC, and terrain data from NASA for a region in Illinois. The early results show that the technique is quite effective against coordinated attacks by malicious nodes.

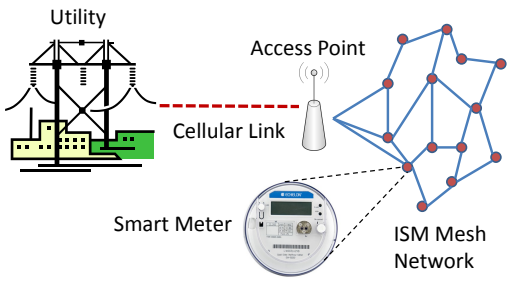


Fig. 1. Current AMI Communication Architecture.

The rest of this paper is organized as follows. In Section II we provide background on white-space networking. In Sec-

tion III we describe the proposed architecture and discuss its strengths and challenges towards a practical deployment. In Section IV we discuss the classification-based solution to the resilient data fusion problem. In Section V we speculate on a possible application of the proposed technique. Section VI concludes the paper.

## II. White-Space Networking

Traditionally, governments have assigned wireless spectrum to interested parties using long-term licenses. This has created significant inefficiencies. For instance, recent measurements in 7 locations across the United States show that, on average, only a fraction (0% to 30%) of the spectrum in the 30-2900 MHz bands is in use [4]. Previous measurements have shown that depending on the region, temporal and geographical variations in the utilization of the assigned spectrum range from $15\%$ to $85\%$ [1].

Dynamic Spectrum Allocation (DSA) is a new paradigm to make spectrum use more efficient. It enables secondary users to opportunistically use frequencies that are not occupied by incumbent (*a.k.a. primary*) users. FCC's recent historical ruling allows unlicensed radio operation in the unused portions of the VHF and UHF spectrum; this is the typical definition of 'white space' [2]. Wireless communications in this spectrum benefit from great signal propagation and penetration properties, which allows for long transmission ranges. *Cognitive Radio* (CR) is an important enabling technology for DSA. A CR can change its transmitter parameters (*e.g.* transmission frequency) based on interaction with the environment in which it operates, and is often able to perform spectrum sensing [10]. The cognitive radio is also referred to as a *secondary* user in this context.

As mandated by FCC, spectrum sensing enables discovering the frequencies that are in use by primary transmitters. All the existing approaches to realizing white-space networks require collecting and combining the sensing data from the cognitive radios in order to build a dynamic database of regions and frequencies that are occupied by incumbents. An alternative is to build the database using information about registered primary transmitters and signal propagation models. This has a variety of limitations including the difficulty of developing a good database of transmitters, inaccuracies in signal propagation models, and factors such as shadow-fading from buildings and trees which are not captured in signal propagation models. Both approaches have been endorsed by FCC and are envisioned to complement each other in the path to the adoption of this technology.

White-space networking is significantly more challenging compared to popular Wi-Fi connections in ISM bands. First, it requires detecting and avoiding interference to incumbents. Second, the network must be able to operate in spectrum bands of varying widths. Third, transmissions in white spaces are subject to temporal variations because primaries such as wireless microphones can become active at any time [6].

White-space networks can be deployed using *infrastructured* or *distributed ad-hoc* architectures. The majority of proposals, however, consider and infrastructured model, which relies on base stations for communication. This is perhaps due to the ease of implementation and conformance to the FCC requirements. The prime example of an infrastructured architecture is the IEEE 802.22 wireless regional area networks (WRAN) standard that is under development at IEEE LAN/MAN standards committee [3], [19]. Another successfully implemented instance from the research community is WhiteFi, which provides connectivity similar to Wi-Fi using the white spaces [6].

*IEEE 802.22 Overview:* IEEE 802.22 (802.22 hereafter) focuses on constructing wireless regional area networks that utilize UHF/VHF TV bands between 54 and 698MHz while ensuring that no harmful interference is caused to the incumbent TV broadcasting and low-power licensed devices such as wireless microphones. It specifies a fixed point-to-multipoint wireless air interface whereby a base station manages its own network and all the associated users. The application for 802.22 is providing wireless broadband access to areas of typically 17-30 km or more in radius (up to 100km) from a base station and serving up to 255 fixed clients with antennas located at about 10m above ground level, similar to a typical UHF/VHF TV receiving antenna. The main focus for 802.22 is providing access to the less populated areas. The minimum throughput delivered to clients is similar to a T1 rate (1.5 Mbps) in the downstream and 384 Kbps in the upstream. It supports both mechanisms for incumbent detection; spectrum sensing and building a database using transmitter data. It also involves detailed coordination mechanisms for co-existence of neighboring networks with overlapping areas of coverage.

*WhiteFi Overview:* WhiteFi is an implementation of a Wi-Fi like protocol on top of the UHF white spaces that addresses the key challenges of white space communications. It offers a centralized architecture in which the clients connect to an access point and perform single hop communication. The network adaptively configures itself to operate in the most efficient contiguous chunk of the available spectrum. WhiteFi is designed for, and tested in the UHF TV spectrum. It involves techniques for detecting incumbents, as well as methods for detecting WhiteFi access points and handling disconnections without causing interference to the incumbents. The communication between the access point and clients is done using stock Wi-Fi cards along with software-implemented UHF translators. This has the advantage that Wi-Fi is a mature technology which is inexpensive and easily available. The spectrum sensing is performed using a separate UHF scanner which is a combination of a UHF antenna and a receive-only daughter board on a software-defined radio. The authors show the abundance of spectrum availability in rural and suburban areas and achieve coverage of up to 1 km in UHF bands (this figure would be 2km in VHF bands) with downstream data rate of 7 Mbps for each 3.5 Mhz of spectrum available to the network [6].

## III. White Space Communication for AMI

We believe the bandwidth, range, and cost improvements offered by white spaces can provide substantial benefits to

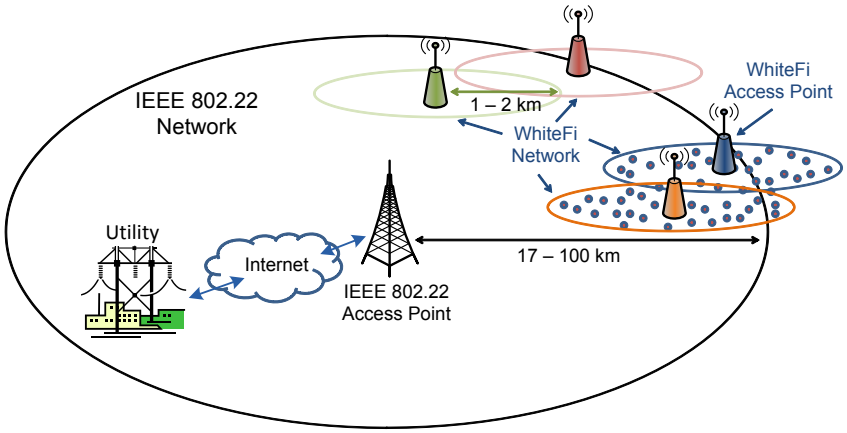


Fig. 2. Proposed Architecture for AMI Communication over White Spaces.

AMI. In this section, we propose an architecture that advances the state of the art in AMI communications and investigate its benefits, limitations, and challenges. To the best of our knowledge this is the first paper that provides a detailed architecture for using white spaces in AMI.

### A. Proposed Architecture

Figure 2 depicts the proposed architecture for AMI communication. This architecture involves two types of wireless networks in a hierarchy. At the lower level, there are small-scale white space networks which are represented as small circles in the picture. A prime candidate for implementing such networks would be WhiteFi. For simplicity we refer to the general class of such networks as WhiteFi in the rest of this paper. Due to the favorable propagation characteristics of the TV spectrum, WhiteFi networks can easily expand in areas with radius of up to 2km, while using commodity Wi-Fi transmitters and conforming to FCC regulations. The WhiteFi networks are envisioned to be established and maintained by the utility companies.

At the upper level in the hierarchy there exist 802.22 networks that provide connectivity between WhiteFi access points and the utility company. As it will be shown below, this provides benefits in terms of cost and broadband penetration in rural areas, while the standardization improves interoperability. The 802.22 networks do not need to be operated by the utility companies. We envision them to be operated by independent broadband service providers that offer service to a utility by admitting the utility's access points in their network. The 802.22 service provider may serve other clients such as residential households and mobile devices as well. This provides broadband access similar to ADSL and cable modems. The two-tier architectures provides a balance between independence and cost savings for the utility, while maintaining high data rate connections to the meters.

The large number and geographical separation of smart meters makes them a valuable resource for distributed spectrum sensing. The smart meters are owned by the utility, however, the 802.22 service provider can obtain the spectrum sensing data from the meters through the WhiteFi base stations. This reduces the number of spectrum sensing units the 802.22 service provider needs to deploy in order to build dynamic spectrum availability maps. This can be an important service that the utility can provide to the 802.22 service provider, and in exchange, receive low service rates and reimbursement for deploying meters with spectrum sensing capability.

In order to diversify its sources for spectrum sensing and avoid relying only on meters that are owned by another entity (the utility), the 802.22 service provider may need to collect spectrum sensing data from other means as well. Such sources could be other clients of the 802.22 service or sensors deployed specifically for this purpose [18]. In addition, in the case of availability of transmitter databases, the list of available channels from both sources (*i.e.* spectrum sensing and transmitter databases) should be intersected to derive the list of available channels. In all these scenarios, the WhiteFi and 802.22 base stations must coordinate their usage of the spectrum using co-existence techniques similar to those proposed in the 802.22 standard draft[1].

### B. Benefits, Limitations, and Challenges

We argue that the proposed architecture provides the following benefits. First, compared to the state of the art, it allows for higher data rates at an economical cost for communication between the meters and the utility. Second, the penetration and long-range transmission properties in white spaces allow for direct communication between the meters and the (WhiteFi) access points. This obviates the need to form complex and unreliable mesh networks that consume considerable power for maintenance and routing. Third, it provides a valuable base of spectrum sensors (the smart meters) for the 802.22 service providers, which may lower their costs and improve their spectrum sensing. This will also provide a leverage to the utilities for discounts from the 802.22 service providers. In addition, this will result in better protection for primary transmitters, which has been the subject of substantial concern by FCC and spectrum license holders. Fourth, since the proposed solution provides cost savings and a revenue stream for 802.22 service providers, it contributes to the cause of

[1]Two months after the initial submission of this paper, service provider Spectrum Bridge announced that it is working with Google to channel data collected by smart residential meters over white spaces, and that they have tested a trial version in Plumas-Sierra County, CA. Technical details were not publicly available at the time of this publication [5].

providing affordable broadband service to rural communities. Fifth, since the approach insists on standardized protocols, it allows for inter-operability between products from different vendors.

One may consider the following limitations for the proposed approach. First, it requires a one-time cost of equipping smart meters with cognitive radios. The cost, however, may be small if the meters are produced at a large scale and could be covered by the spectrum sensing service they provide to the 802.22 provider. Second, there might be times or locations where no white space is available. In this case, the networks can temporarily operate in the ISM bands at lower bit rates. Therefore, in the worst-case scenario the performance would be similar to that of the existing architectures. Alternatively, a narrow band can be purchased at a small cost for emergency backup usage. Either of the above approaches guarantee that the network maintains minimum connectivity. Third, there may exist various security and reliability concerns associated with the proposed architecture. In the rest of this section, we enumerate some of these concerns and we provide solution ideas or references. We defer the detailed treatment of a particularly important problem of *resilient data fusion* and discussion of meter density to Section IV.

### C. Primary Emulation - Unauthorized Spectrum Usage

Primary emulation attacks can disrupt AMI communications over white spaces. In a primary emulation attack, an attacker may modify the air interface of a CR to mimic a primary transmitter signal's characteristics, thereby causing legitimate secondary users to erroneously identify the attacker as a primary user, and abandon the channel. Of the body of existing work, LocDef utilizes both signal characteristics and location of the transmitter to verify a primary signal transmitters' location [8]. If it does not match the known locations for primary transmitters, the signal is from an attacker. This approach, however, requires knowledge of the location of the primary transmitter, and thus may not be practical in some circumstances.

An alternative is using cryptographic and wireless link signatures to authenticate primary users' signal in presence of attackers that may mimic the same signal [15]. This is achieved by using a helper node close to a primary user to enable a secondary user to verify cryptographic signatures carried by the helper node's signals and then obtain the helper node's authentic link signatures to verify the primary users signals. Liu *et al.* [14] also study the problem of detecting unauthorized spectrum usage, where the authorized transmitter may be mobile. They propose two analytical methods and a solution based on machine learning to detect anomalous transmission by using the characteristics of radio propagation.

## IV. Ongoing Work: Resilient Data Fusion using Classification

As mandated by FCC, spectrum sensing is an integral part of white-space networking. The base stations, or fusion centers, collect spectrum sensing reports from the cognitive radios (mainly smart meters in this case) and combine them to determine the unused channels. The process of combining spectrum sensing results from radios by the base station is referred to as (data or decision) fusion (also known as *collaborative sensing*). To effectively perform fusion, the base station divides its area of interest to a grid of small *cells* such that each would contain a few to tens of CRs. Each cell would be the unit of fusion, in which spectrum sensing results are combined to determine primary presence.

In *decision fusion*, each CR reports a 0 or 1 decision to the base station, representing the presence or absence of primaries in a channel, whereas in *data fusion*, raw measurements from CRs are reported to the base station. These reports are often measurements of the signal power on a target frequency channel. In the rest of this paper we focus on data fusion since it is a superset of decision fusion and provides more data to the base station for analysis. A common technique for data fusion in this context is Equal Gain Combining (EGC) which compares the average of power measurements to a detection threshold $\lambda$ to determine whether the primary is present [16].

Given the diverse and potentially unreliable set of nodes used for spectrum sensing in this context, it is possible that some of the nodes be malicious or compromised. Such nodes may work together and seek to *exploit* a spectrum in a given region by falsely reporting that a primary signal is present, or *vandalize* a primary by reporting that its signal is not present thereby encouraging interference from secondaries. Detecting such attacks is challenging due to spatial variations of primary signal, natural differences due to shadow-fading and noise, and temporal variations of primary's presence. These factors makes it easier for compromised nodes (that may collude) to hide their false reports under the legitimate variations.

### A. Detecting Attacker-Dominated Cells

When a base station divides its service area to small cells, nodes within a cell are expected to provide similar readings. Much of the existing work aims to detect malicious nodes by identifying their measurements as abnormal or outlying [9], [11], [16]; if a majority of nodes in a given cell provide a reading in a common range, the other nodes may be identified as outliers. Such techniques only work if a minority of nodes in a cell are compromised. There has also been efforts to detect cells that are dominated by attackers [9]. The authors consider the averages of neighboring cells in a hierarchy and detect cells with outlying averages as attacker-dominated. The above approaches suffer from the following limitations. First, they unrealistically assume complete knowledge about the models and parameters of signal propagation. Second, their performance is dependent on threshold parameters which are either tuned by hand, or depend on the parameters of the signal propagation model.

In the context of AMI, we propose to identify attacker-dominated cells by using real signal propagation data. The data is used to build a *classifier* that is *trained* to differentiate between natural and un-natural signal propagation patterns in a region. The idea is to learn the normal propagation behavior

of the signal from the reliable signal propagation data and use it to spot unnatural propagation of signal, which may be caused by malicious false reports.

Consider the *local neighborhood* $N_A$ of a cell $A$ to contain $A$ and its 8 neighboring cells. Using this definition, we represent $A$ by a 9-element tuple containing the 'average' reported powers from the CRs in each of the cells in $N_A$ (in a pre-specified order). We call this the *neighborhood representation* of $A$. Assume we have access to reliable power measurements in (all or a subset of) the region of interest. This data can be used to create one neighborhood representation for each cell in the area. We refer to each of such representations as an 'example.' Therefore, we can assume access to a large number of such examples representing the 'natural' propagation of signal in local neighborhoods. Also, as we will elaborate later, assume we have access to the neighborhood representation for a sufficiently large and diverse set of 'un-natural' (attacker-dominated) cells.

We are now ready to reduce our problem to a binary *classification* problem. Classification is a machine learning technique that is widely used in domains ranging from spam email detection and unauthorized spectrum usage to fraud detection and speech recognition [12], [14]. In a binary classification problem we are given a set of *training* examples with their corresponding labels, $(\overrightarrow{x_i}, y_i)$, where $\overrightarrow{x_i}$ is the representation of the $i^{th}$ example and $y_i \in \{1, -1\}$ ('yes' or 'no') is the corresponding binary label. Each example is described by a vector of its attributes which is often called the feature vector. For example, the neighborhood representation of a cell can serve as its feature vector. The goal is to predict a binary label for a *test* example for which we do not know the label, using the classifier built from training examples [7].

A classifier tries to partition the input feature space into regions where positive examples lie versus regions where negative examples lie. The boundary between regions for positive and negative examples is called the *decision boundary*. Training involves learning the decision boundary and classification involves determining on which side of the decision boundary a test example lies. In our experiments, we opt for using Support Vector Machines (SVM) to construct our classifier. SVMs are one of the most widely used techniques for building classifiers. An SVM constructs a hyperplane or a set of hyperplanes in a high or infinite dimensional space that constitute the decision boundaries. Further details about SVM can be found in [7].

Now we turn to the problem of obtaining training examples. For AMI, there exists a large base of smart meters. At the time of deployment, or right after a physical firmware update, these devices can be trusted to be un-compromised. *Natural* instances can be obtained in a practical *one-time* process based on this trusted sensor grid. An alternative is war-driving where a sensor is moved though the region to collect training data. Having obtained such natural (normal) examples, we modify them to inject *un-natural* training instances to represent attacker-dominated cells. The instances have to be general enough to train the classifier in such a way that it is able to detect vandalism and exploitation attacks mounted by coordinating attackers inside the cell. To that end, we create multiple un-natural examples from each natural example. More specifically, we replace the actual average power in a cell with a value that is a random amount higher (lower) than the primary detection threshold $\lambda$ to represent exploitation (vandalism) attacks.

### B. Evaluation

We perform an early evaluation of the attacker detection system on data from parts of East-Central Illinois that offer a flat terrain and medium to sparsely populated suburban and rural areas. We rely on the house density data from the US Census Bureau, Digital DTV transmitter database from FCC, and terrain database from NASA. We use the Longley-Rice empirical outdoor signal propagation model [17] to generate training and testing signal propagation data. We treat the signal propagation data obtained from Longley-Rice as the ground truth provided by sensors and use this to perform an early evaluation. Our method, however, does not rely on any specific choice of a model. Therefore, if these models have some inaccuracies then we believe that accurate training data and the application of our method will achieve the necessary foundation for integrity protections.

An important factor in our evaluation is the size and sensor density of each cell. We take the approximation of one sensor per house, which is supported by the fact that each house would have at least one meter. Our choice of cell size and density is guided by the following considerations. First, the assumed densities should match the real house densities in the region. Second, at least a few to tens of nodes must exist in a cell to add the required diversity gains promised by collaborative sensing. Third, the variation of average signal power in a cell must not be significant in order for collaborative sensing to be meaningful. Using this criteria, the maximum radius of 5.6 km for a cell is recommended by related work [13]. Fourth, as a low priority recommendation (not a requirement), limiting the density to 3.2 sensors/km$^2$ improves the performance of collaborative sensing by ensuring independence between individual reports [13].

We studied the house density of the 102 counties in Illinois using data from the US Census Bureau. The results show that the least dense county contains 2.5 houses/km$^2$, with the 5th, and 50th percentile at 3.5 and 8.5 respectively. In view of this data and the discussion above, we consider base cells of size 2km$\times$2km with the average density of 3.2 sensors per km$^2$. In the rare cases (given the conservative choice of density) that some cells contain less number of sensors, or sensors are closely clustered, the service provider may deploy additional sensing units. We consider a 160km $\times$ 160km square area in East-Central Illinois with southwest and northeast coordinates of $(39.56, -89.4)$ and $(41, -87.5)$ in (*latitude, longitude*) format. The area consists mainly of rural farmland and a few small and mid-size towns.

We hypothesize that given the limited frequency range for TV signals, the signal propagation patterns remain mostly

unchanged across different frequency channels (and therefore transmitters), and are mainly a function of the terrain. To verify this in the Illinois terrain, we build a classifier that is trained by pooling data from multiple transmitters in such a way that there exist sufficient number of examples in any power level from 90 dBm to -130 dBm. Using this criteria, of the tens of transmitters in the region, we pick the following three to build the classifier: WEIU-TV (PBS), WICS (ABC), and KTVI (Fox). We test the classifier on the data from another three randomly selected transmitters: WAOE (MyN), WCIA (CBS), and WQAD-TV (ABC). Table I summarizes the performance of the classifier. It can be seen that in all the power ranges our classifier can successfully detect attacker dominated cells, with low false positive rates.

TABLE I
DETECTION ACCURACY (D.A.) AND FALSE POSITIVE (F.P.) PERCENTAGES FOR THREE DTV TRANSMITTERS.

| | **WAOE (MyN)** | | **WCIA (CBS)** | | **WQAD-TV (ABC)** | |
|---|---|---|---|---|---|---|
| | D.A. | F.P. | D.A. | F.P. | D.A. | F.P. |
| $P > -65$ | 100 | 0 | 99.8 | 0 | - | - |
| $-65 \geq P > -85$ | 100 | 0 | 100 | 0 | 100 | 0 |
| $-85 \geq P > -105$ | 100 | 0 | 100 | 0 | 100 | 0 |
| $-105 \geq P > -114$ | 99.1 | .9 | - | - | 99.6 | .8 |
| $-114 \geq P$ | 97.3 | 3.2 | - | - | 95.1 | 7.6 |
| **Overall** | **99.3** | **.8** | **99.9** | **0** | **99.3** | **1.3** |

## V. SAMPLE APPLICATION

AMI provides a number of general benefits, some of which we enumerated earlier, but are there any benefits of white space AMI for rural regions beyond these general benefits? We speculate here on at least one such possibility. Power grids in rural regions typically have the characteristic that loads are sparsely distributed along the power lines, often only at farm houses and machine sheds. Such loads must be metered and meters are expensive to monitor. Thus a 200 acre farm might have a half a mile of adjacent power line but have power only from a meter at the farm house on a corner of the property. If, for example, there is an electric fence on a remote part of the property then power must be supplied with a battery since it is impractical to run a power line from the meter. The value of power harvesting for military purposes is well recognized (see http://www.ndep.us/Power-Harvesting-Induction-Magic and http://www.ndep.us/Power-Harvesting-The-Bat-Hook for instance) but civilian uses will require metering. If it were feasible to add meters and power links easily along the utility power lines then this problem could be significantly diminished. Generally power companies do not wish to add new meters for such low loads, but when the cost of collecting billing data is diminished by suitable wireless communications infrastructure (based on white space for instance), then costs can be contained and a valuable service becomes feasible.

## VI. CONCLUSIONS

In this paper we provided an architecture for AMI communications using the white spaces in the TV spectrum. We argued that the proposed architecture offers improvements in terms of bandwidth, deployment, and cost compared to state of the art. We discussed various security and reliability issues associated with the proposed architecture and identified resilient data fusion as a particularly important and challenging one. To address this problem, we considered using classification techniques from machine learning to identify attackers. We performed an early evaluation of this technique on data from suburban/rural areas in Illinois. We also introduced a novel potential application of the infrastructure to provide a new type of rural power service. In future, we will investigate in more detail the effect of attacker strategies, frequency, household density, and most importantly terrain on the classifier-based defense in rural, suburban, and urban areas.

## ACKNOWLEDGEMENTS

This work was supported in part by HHS 90TR0003-01, NSF CNS 09-64392, NSF CNS 09-17218, NSF CNS 07-16626, NSF CNS 07-16421, NSF CNS 05-24695, and grants from the MacArthur Foundation, Boeing Corporation, and Lockheed Martin Corporation. The views expressed are those of the authors only.

## REFERENCES

[1] FCC, ET Docket No 03-222, December 2003.
[2] FCC, ET Docket No 08-260, November 2008.
[3] IEEE 802.22 WRAN WG on Broadband Wireless Access Standards. *www.ieee802.org/22*.
[4] Shared Spectrum Company. *http://www.sharedspectrum.com/*.
[5] Spectrum Bridge Company. *http://www.spectrumbridge.com*.
[6] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, and M. Welsh. White space networking with wi-fi like connectivity. *SIGCOMM Comput. Commun. Rev.*, 39(4):27–38, 2009.
[7] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, 2006.
[8] R. Chen, J.-M. Park, and J. Reed. Defense against primary user emulation attacks in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 26(1):25–37, Jan. 2008.
[9] O. Fatemieh, R. Chandra, and C. A. Gunter. Secure collaborative sensing for crowdsourcing spectrum data in white space networks. *DySPAN '10: IEEE Symposium on Dynamic Spectrum Access Networks*, April. 2010.
[10] S. Haykin. Cognitive radio: brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–220, Feb. 2005.
[11] P. Kaligineedi, M. Khabbazian, and V. Bhargava. Secure cooperative sensing techniques for cognitive radio systems. *ICC '08: IEEE International Conference on Communications*, pages 3406–3410, May 2008.
[12] H. Kim, S. Pang, H. Je, D. Kim, and S. Bang. Pattern classification using support vector machine ensemble. pages II: 160–163, 2002.
[13] H. Kim and K. G. Shin. In-band spectrum sensing in cognitive radio networks: energy detection or feature detection? In *ACM MobiCom '08*.
[14] S. Liu, Y. Chen, W. Trappe, and L. Greenstein. Aldo: An anomaly detection framework for dynamic spectrum access networks. In *INFOCOM 2009, IEEE*, pages 675 –683, april 2009.
[15] Y. Liu, P. Ning, and H. Dai. Authenticating primary users' signals in cognitive radio networks via integrated cryptographic and wireless link signatures. In *IEEE Symposium on Security and Privacy*, 2010.
[16] A. Min, K. Shin, and X. Hu. Attack-tolerant distributed sensing for dynamic spectrum access networks. In *ICNP '09: 7th IEEE International Conference on Network Protocols*, pages 294–303, Oct. 2009.
[17] T. Rappaport. *Wireless Communications: Principles and Practice*. IEEE Press, New York, 1996.
[18] N. Shankar, C. Cordeiro, and K. Challapali. Spectrum agile radios: utilization and sensing architectures. *IEEE DySPAN '05*, Nov. 2005.
[19] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell. Ieee 802.22: the first cognitive radio wireless regional area network standard. *Comm. Mag.*, 47(1):130–138, 2009.