Deployment Strategies for Crowdsourced Power Outage Detection

Santiago Correa
STIMA Lab, Department of Electrical and Computer Engineering
University of Massachusetts, Amherst
scorreacardo@umass.edu

Noah Klugman
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
nklugman@berkeley.edu

Jay Taneja
STIMA Lab, Department of Electrical and Computer Engineering
University of Massachusetts, Amherst
jtaneja@umass.edu

Abstract—**Smart grids, typically driven by smart meters, enable the use of information and communication technologies to collect grid status in real-time. However, while smart meters are typically essential to the smart grid vision, many utilities globally have not installed smart meters mainly due to technical or budget constraints. In this paper we evaluate deployment strategies of GridWatch, a novel crowdsourcing system to detect electricity outages using smart phones. Using demographic, user mobility, and outage data relevant for Nairobi, Kenya, we develop an agent-based model (ABM) simulation to understand the factors that optimize the deployment strategy of GridWatch in different sub-regions of the city while maintaining high confidence of outage detection. Our results show that outage detection improves dramatically with increasing density of households per transformer, so a higher penetration of GridWatch devices is needed in areas with sparser grids.**

I. INTRODUCTION

Smart grids enable a transformative improvement in monitoring of the state of the generation, transmission, and distribution tiers of electricity grids. However, while smart meters are typically a canonical component of the smart grid vision, their deployment is uneven worldwide; while some regions have converted entirely to smart meters, deployment has slowed considerably in others – in the U.S., only 50% of endpoints have smart meters, and much of the developing world has installed few if any smart meters. To achieve the promise of more efficient, more responsive, and less costly electricity systems of the smart grid, we study the potential to offer smart grid-like services without smart meters.

In particular, this work examines deployment strategies for a system called GridWatch, which aims to enable commodity smartphones owned by electricity customers to automatically provide notification of electricity outages to utilities [1]. GridWatch proposes an unprecedented paradigm for mobile-driven, crowdsourced sensing approaches for monitoring infrastructure systems. However, little is understood about the density of GridWatch devices needed to accurately measure electricity outages, and even less is understood about the dynamics of those requirements in different settings. Further, as deploying GridWatch at increasing penetrations seemingly requires increasing effort and expense, we seek to understand the minimum deployment penetration needed based on the characteristics of each setting.

To study these topics, we examine potential deployment strategies in the context of a real-world electricity grid in an emerging economy. Using this novel, detailed dataset, we apply an agent-based modeling technique to characterize the theoretical effectiveness of GridWatch across the 17 administrative regions that comprise Nairobi, Kenya, a city of 4 million residents. We consider differences in population density, grid infrastructure, and outage patterns among these regions, and examine their resulting effects on detecting outages using the GridWatch system. We intend for this study to inform deployments of GridWatch and other mobile-based infrastructure monitoring systems.

II. RELATED WORK

The growing penetration of mobile devices and sensors has created unprecedented opportunities to modernize the power system by enabling information acquisition in real-time and from end-users. Smart grids envision a versatile and intelligent electricity infrastructure that allows participation from end-users to achieve better resource optimization, information services, and monitoring [2]. In particular, this work investigates the deployment strategy of a crowdsourcing system, GridWatch, that aims to measure the reliability of electricity grids using smartphones. Like other crowdsourcing systems, the number of users of GridWatch not only impacts the robustness of detection but also the cost associated with the deployment [3]. Clouqueur, *et al.* [4] address the deployment strategy problem for sensor networks that are placed to detect a moving target using path exposure and signal measurements. However, our work addresses this problem for mobile users by understanding their charging patterns and WiFi penetration for outage detection in distribution grids.

In our previous work [5], we evaluate the deployment strategy of GridWatch using a purely stochastic model where we described the occurrence of power outages as a non-homogeneous stochastic process [6]. In contrast, in this work we use an agent-based model technique that provides more accurate estimate of deployment penetration given that we are able to model individual interactions instead of making assumptions of homogeneity across the regions of study. In addition, one of the signals used by GridWatch to attempt to detect an outage is by monitoring sudden drops of WiFi signals; we account for the ability of this signal to measure outages both at a single residence as well as at nearby residences. We consider the effects of varying penetration of WiFi across the city, and characterize the tradeoff in outage
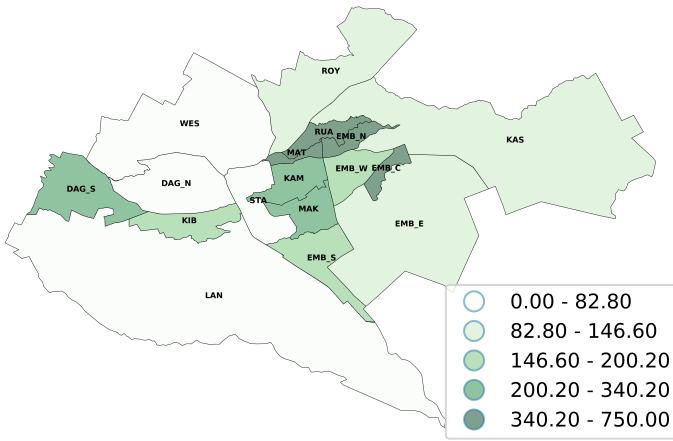
Fig. 1. Map of Nairobi and its 17 administrative and geographic regions denominated constituencies. Darker colors represent higher densities of households per transformer, and labels are provided in Table I.

Fig. 2. Measurement of households versus outages per km^2 in Table I. A line of best fit with R^2 of 0.60 shows a moderately high correlation between household and power outage density.

estimation between higher confidence and better coverage via varying the outage detection threshold.

Agent-based modeling (ABM) is a computational modeling technique that describes the behavior of individual agents to observe the results of their interactions in a complex system [7]. This technique is widely used in engineering, social, and natural sciences but, to our knowledge, has not been used for this application. Existing ABM simulations related to electricity infrastructure aim to understand the sustainability of energy management provided by smart homes [8], market outcomes and consumer behavior in demand response programs [9], effects of intermittent renewable sources in electricity markets [10]–[12], and impacts of traditional and new electric loads such as plug-in electric vehicles in the planning and performance of smart grids [13], [14]. It is worth noting that most of the aforementioned work relies on the availability of smart meter data which is either limited or entirely unavailable in developing countries.

III. BACKGROUND

We aim to understand the features that impact the detection of power outages by the GridWatch system. To do this, consider a large and growing urban setting: Nairobi, Kenya. We identify the feature space based on available data as provided by the largest utility. We characterize the 17 administrative units that constitute the county of Nairobi, Kenya based on ranges of inter-household distance, the population density, number of power outages in the distribution grid, and power infrastructure density.

A. Constituencies of Nairobi

Nairobi, a city of four million residents, lies primarily within Nairobi County which consists of 17 administrative and geographic regions denominated constituencies; these constituencies are listed in Table I. We characterize densities in terms of number of households using information from the 2009 Kenya Census and data from the Kenya National Bureau of Statistics (KNBS) and Society for International Development
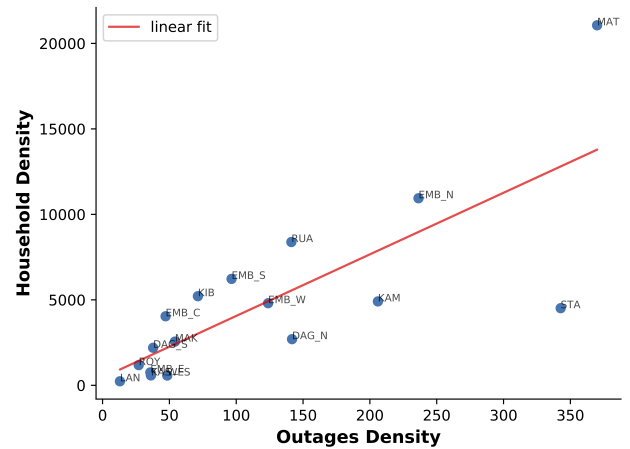
(SID) [15]. Figure 1 shows a map of the constituencies of Nairobi. Most of the population density in Nairobi is clustered in the center of the county corresponding to the constituencies of Mathare, Kamukunji, and Ruaraka. The Nairobi National Park, with a very low population density, is located in the southern part of Langata Constituency.

The densities of power distribution infrastructure and outages provided for each constituency are for the period from September 2014 to October 2015. We measure infrastructure in terms of the number of distribution transformers available in each region and outages using the datasets mentioned in Section IV-A.

B. GridWatch

GridWatch is a mobile-based crowdsourcing system for electricity grid outage and restoration measurement [1]. It uses sensors embedded in smart phones to detect absence of power supply. GridWatch detects an outage by observing the charging patterns of users, detecting sudden drops of WiFi signals and ambient lights, among other sensors.

As was mentioned in [5], even though the system enables the detection of outages in settings with limited sensing infrastructure in the distribution grid, GridWatch is not yet broadly deployed. In order to account for the main techniques GridWatch would employ to detect an electricity outage, we incorporate the two main detection mechanisms in our model using inter-household distance for WiFi detection and user charging patterns across the day. These mechanisms are able to simultaneously interact to sense outages, improving the confidence of detection. Thus we design scenarios in our model in which they overlap.

IV. METHODOLOGY

In this section we explain the datasets, agent-based model approach, and the parameters of our model.

TABLE I
METRICS PER CONSTITUENCY IN NAIROBI COUNTY, KENYA.

Constituency in Nairobi	Households (per km^2)	Transformers (per km^2)	Outages (per km^2)	Median num. houses within 10 mts	Median num. houses within 30 mts	Median num. houses within 50 mts	% houses in range (within 30 mts)
Mathare (MAT)	21060	31.7	370.0	58	177	371	100
Starehe (STA)	4512	66.2	342.7	32	73	142	99
Embakasi North (EMB-N)	10943	23.6	236.4	18	53	115	99
Kamukunji (KAM)	4908	23.1	206.0	39	90	184	100
Dagoretti North (DAG-N)	2700	35.5	141.8	33	52	88	97
Ruaraka (RUA)	8382	24.6	141.3	37	114	255	99
Embakasi West (EMB-W)	4803	32.3	123.9	22	72	172	100
Embakasi South (EMB-S)	6225	34.4	96.5	94	245	533	100
Kibra (KIB)	5214	30.5	71.3	14	38	74	97
Makadara (MAK)	2558	10.9	54.1	12	28	62	98
Westlands (WES)	573	8.5	48.4	21	34	63	92
Embakasi Central (EMB-C)	4039	5.4	47.1	24	86	209	100
Dagoretti South (DAG-S)	2196	6.5	37.8	26	43	79	98
Kasarani (KAS)	579	4.7	36.2	18	47	103	96
Embakasi East (EMB-E)	753	6.8	35.7	28	57	107	97
Roysambu (ROY)	1182	8.2	27.0	30	78	170	97
Langata (LAN)	236	3.3	13.0	14	26	51	90

A. Datasets

As with our previous work [5], we use datasets for historic outage data collected by the largest electric utility and phone charge events using the publicly-available StudentLife dataset [16] to model the conditions GridWatch requires to sense power outages. Each entry in the outage data contains information on the problematic piece of equipment; these pieces of equipment are defined as the agents in our model. Their probability of failure depends on the number of events of that particular device presented in the dataset at a particular hour of day. Using phone charge events we obtain the proportion of users charging their phones during the course of the day. Even though the StudentLife dataset provides good insight about charging patterns of smartphone users in the United States, we recognize the potential for regional differences in user behavior and intend to collect phone charging data from our target setting to provide a more realistic simulation environment. Unfortunately, this information is not available.

In addition, since GridWatch measures changes in WiFi signals, we leverage household meter locations collected by the utility and calculate the median number of neighbors within 10, 30, and 50 meters. We use these metrics to obtain the proportion of households with at least one neighbor within WiFi range. Different IEEE 802.11 standards explored [17] define an approximate range coverage of 30 meters which is the range we choose for this metric. Table I summarizes the percentage of household with at least one neighbor within a 30-meter range, median number of households within 10, 30, and 50 meters, the demographic, power infrastructure, and outage density for each constituency. Figure 2 illustrates the positive relationship between household versus outage densities across all the constituencies.

B. Agent-Based Model Design

We model the occurrence of power outages in the constituencies mentioned in Section III. We aim to observe the dynamics of those outages from the perspective of each individual problematic piece of equipment that generated

the failure and then we create possible scenarios of outage detection using the GridWatch app.

We identify feeders and transformers in the distribution grid as the infrastructure where the outages occur and we choose them as the agents of the model. From the underlying datasets, we obtain the frequency of failure across hours of day. In order to recreate dynamics of the outages, we use this information to calculate the probability that if an outage occurs, it matches to a specific hour of day and agent.

We assign a set of attributes to the system and each agent called global and local variables respectively. We define the following global variables:

- Number of outages per hour of day. Using the outage dataset from the utility, we separate the number of outages that occurred inside each constituency. In turn, we obtain the respective number of outages that occur across the hours of day. We observe a common pattern of higher occurrence of outages during the peak hours across all constituencies.

- Proportion of users charging their phones (P_{ch}). Based on the StudentLife Dataset, we use this parameter to calculate the probability that a GridWatch user reports an outage when the phone is plugged in and the number of households that are not charging their phones at a specific hour of the day.

- Proportion of GridWatch users (P_{gw}). We vary this parameter to obtain the minimum optimal proportion of users that allow us to detect most of the outages. Note that a user here represents their household; in this work, we do not account for the potential that an individual household can have multiple phones with GridWatch installed. We vary this parameter from $0 - 100$ percent.

- Threshold of reports. When an outage occurs, several GridWatch users can report the event. This parameter is defined as the number of reports above which our model considers an outage as detected. This allows a tradeoff between detection confidence and outage coverage. To observe its impact we vary this threshold from 1 to 5.

- WiFi detection accuracy(P_{acc}). GridWatch aims to detect outages by identifying changes in surrounding WiFi signals. We incorporate a measure of accuracy to observe the fluctuations in outage detection when we vary this parameter in our model.
- Proportion of households with WiFi access points (P_w). The availability of WiFi signals around GridWatch users affects the ability of the system to observe outages indicated by changes in available WiFi networks. This parameter depends on the demographics, income level, and geographic location of households.
- Proportion of households that are located between each other at a distance greater than the WiFi coverage range (P_{out}). Households that are located at a distance from other households greater than the WiFi coverage range can only observe signals from their own access points or charging patterns to detect outages using GridWatch.

The local variables are:

- The average (μ) and standard deviation (σ) of the number of households affected by the outage across every unique agent reported in the dataset. These variables allow us to generate the random variable N which refers to the number of households affected for each simulated outage using a normal distribution.

$$N \sim Normal(\mu, \sigma) \tag{1}$$

- Probability of faults across each hour of day (P_f). This is obtained from the frequency of outages experienced per agent. Given that an outage occurs, we calculate the probability of having a fault at that specific time of day and caused by the given agent.
- Number of households affected having the GridWatch app installed. This is a binomial random variable with parameters of the number of affected households and the proportion of users with GridWatch installed (P_{gw}).

$$n_{gw} \sim Binomial(N, P_{gw}) \tag{2}$$

- Number of GridWatch users that can report only monitoring charging states (n_{ch}).

$$n_{ch} = n_{gw} \cdot P_{out} \cdot (1 - P_w) \tag{3}$$

- Number of GridWatch users that can report only monitoring WiFi signals (n_w).

$$n_w = n_{gw} \cdot (1 - P_{out}) \cdot (1 - P_{ch}) \tag{4}$$

- Number of GridWatch users that can report using both mechanisms simultaneously (n_{cw}).

$$n_{cw} = n_{gw} - n_{ch} - n_w \tag{5}$$

- Number of GridWatch reports when an outage occurs. This local variable is obtained adding the number of reports in n_{ch}, n_w, n_{cw} using binomial random variables:

$$r_i \sim Binomial(n_i, P_d(n_i)) \tag{6}$$

TABLE II
PERCENTAGE OF DEPLOYMENT FOR 80% OUTAGE DETECTION

Constituency	WiFi penetration		
in Nairobi	10%	30%	70%
Kamukunji (KAM)	36	18	13
Westlands (WES)	>100	70	45
Dagoretti North (DAG-N)	87	46	31
Roysambu (ROY)	30	15	10
Starehe (STA)	100	55	35
Langata (LAN)	>100	76	48
Dagoretti South (DAG-S)	21	12	8
Kasarani (KAS)	40	20	13
Ruaraka (RUA)	77	41	26
Kibra (KIB)	87	46	30
Mathare (MAT)	28	15	9
Embakasi West (EMB-W)	>100	83	60
Embakasi East (EMB-E)	48	25	16
Embakasi North (EMB-N)	37	19	13
Embakasi Central (EMB-C)	11	7	4
Embakasi South (EMB-S)	>100	92	57
Makadara (MAK)	100	56	37

Where i denotes each category and their respective probability of detection P_d is:

$$P_d(n_{ch}) = P_{ch} \tag{7}$$

$$P_d(n_w) = P_{acc} \cdot P_w \tag{8}$$

$$P_d(n_{cw}) = P_{ch} + (P_{acc} \cdot P_w) - P_{ch} \cdot (P_{acc} \cdot P_w) \tag{9}$$

V. RESULTS

In this section, we examine the response of our ABM to variations in parameters including the threshold of reports, the proportion of users with GridWatch installed, and the proportion of households with WiFi access points. We also evaluate the deployment strategy required to maintain an outage detection rate of 80% across each of the constituencies and identify the dynamics of detection in each area.

A. Detection with varying WiFi penetration

We begin setting an outage detection goal for each of the constituencies of 80% and evaluate the deployment strategy required at three different levels of WiFi penetration (low, medium, and high): 10%, 30%, and 70%. Table II summarizes the results for each constituency and shows that for low WiFi penetration, in certain constituencies we never met the detection goal even having a deployment of 100% of users with GridWatch installed. The regions with this high deployment requirement match the ones with a low number of households per transformer. In this experiment we set the threshold of detection to three and the WiFi detection accuracy to 50%.

Figure 3 shows the results of varying WiFi penetration at different proportions of users with GridWatch in Kibra Constituency, which has a large population under the poverty
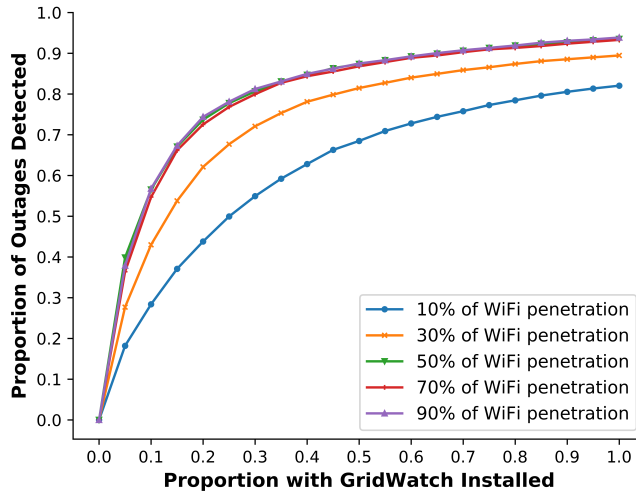
Fig. 3. Outage detection in Kibra Constituency, which has a large population under poverty line, at different levels of WiFi penetration. Outage detection starts to saturate after 50% penetration.
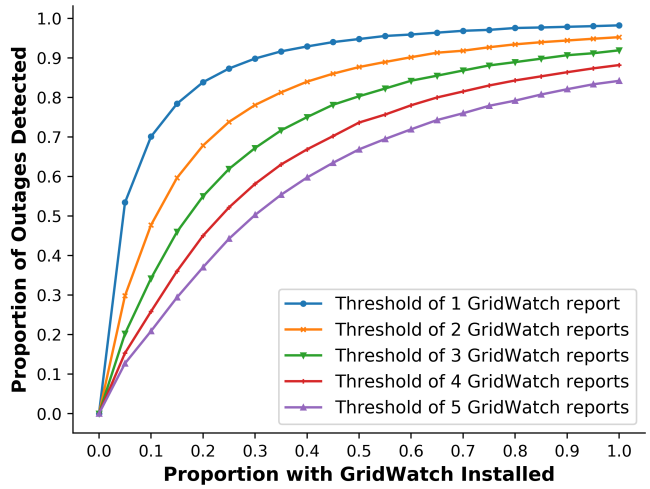

Fig. 4. Outage detection in Dagoretti North Constituency at different thresholds of outage reporting. This constituency has one of the largest densities of outages per km^2.

line. We observe a substantial improvement in detection (up to 35%) in changes from 10% to 50% WiFi penetration. However, increments beyond 50% WiFi penetration yield diminishing returns. As expected, increments in the proportion of GridWatch users increases the proportion of outages detected. In the following experiments we approximate and fix the WiFi penetration to 30% based on the number of internet broadband subscriptions reported by the Communications Authority of Kenya in the second quarter of financial year 2016-2017 [18], which reports a total penetration level of 28.7% with a growth rate of 6.7%. Even though the penetration in Nairobi might be higher and not all of these connections represent WiFi access points, we consider 30% a suitable estimate.

B. Varying threshold of outage detection

An important parameter is based on the definition of an outage using GridWatch. In this approach, we vary the threshold of reports from 1 to 5 apps reporting. Ideally multiple reports need to be considered in order to flag an event as a detected outage so we can reduce the probability of a false positive detection. We want to observe how varying this threshold affects the likelihood of detection across different proportions of GridWatch users. Figure 4 shows the results of this experiment in Dagoretti North Constituency. We can observe how a threshold of 1 can reach a high proportion of outage detection: 70% of outages were detected with only 10% of households having GridWatch installed. However, this scenario is not practical since this detection threshold is not enough to mitigate false positive events. We believe that a reasonable threshold is 3 to balance between detection confidence and coverage at a lower cost. Even though larger thresholds can reduce false positive events, the strategy requires high deployment penetration to capture the desired number of outages.
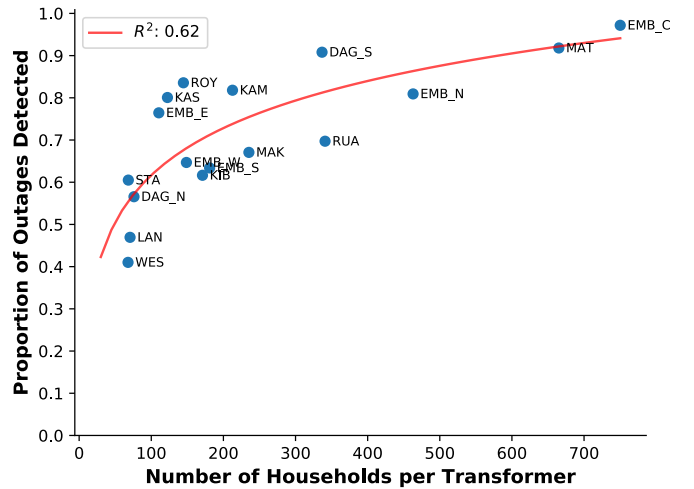

Fig. 5. Proportion of outages detected when 20% of users have GridWatch installed versus number of households per transformer in each constituency.

C. Dynamics between densities and detection

For this experiment, we defined a fixed proportion of GridWatch users at 20%, a threshold of reports equal to 3, WiFi accuracy of 50%, and WiFi penetration of 30% across all the constituencies. Figure 5 shows the proportion of outages detected versus the number of households per distribution transformer. We found that a larger number of households per transformer can detect a higher proportion of outages, with a positive correlation of 0.62. It is worth noting that the correlation is not linear but logarithmic. Thus, to ensure equal outage detection across areas, the ideal deployment penetration of GridWatch should substantially vary by constituency.

Figure 6 shows the deployment penetration required to reach 80% outage detection in each constituency. As with previous experiments, we set the detection threshold to 3 reports, 30% WiFi penetration, and 50% WiFi accuracy. We observe that the best fit shows an exponentially decreasing function with
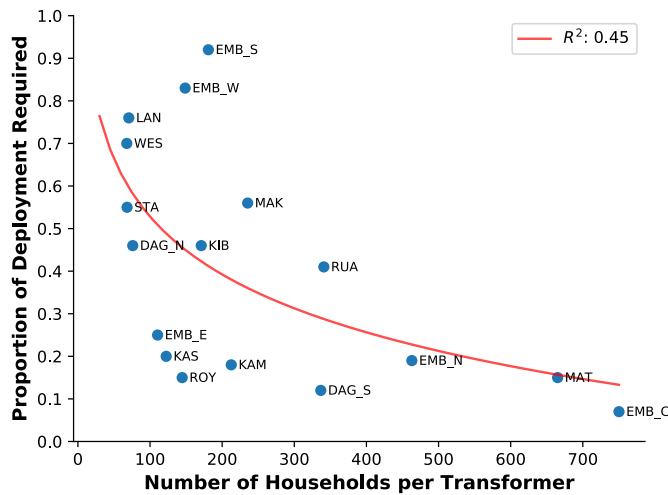
Fig. 6. Proportion of users with GridWatch installed needed for 80% outage detection versus number of households per transformer in each constituency.

R^2 of 0.45; while this is not perfect correlation, we believe this characterizes a fundamental pattern. Even though there is a higher variance between the deployment required and households per transformer to meet the detection goal, it is still possible to observe a trend of decreasing deployment for higher household density per transformer. We note the especially large range in deployment penetrations needed in each constituency to achieve similar detection rates.

VI. Future Work and Conclusions

In this work we developed an ABM simulation to explore the deployment strategy of GridWatch, a novel crowdsourcing system to detect electricity outages using the sensors embedded in commodity smartphones, in real settings. We evaluated the size of deployment required in different sub-regions of Nairobi, Kenya, and observed the changes in the confidence of detection when there are differences in population density, grid infrastructure, and outage patterns among these regions. Outage detection improves with increasing density of households per transformer, so a higher penetration of GridWatch devices is needed in areas with sparser grids. We also considered the variation of parameters that enable the sensing of outages: the availability of WiFi signals exposed by the WiFi penetration and the threshold of reports to consider an outage detected. We found that improvements in outage detection from increased WiFi access saturate at 50% WiFi penetration and varying the threshold number of devices for outage detection provides a tradeoff between coverage and confidence. As we continue this work, we will collect charging patterns specifically from the area of studies and test our deployment strategies in similar settings. With further study and deployment experience, we believe that nontraditional and sidechannel sensing methods can improve monitoring and management of electricity grids.

VII. Acknowledgments

References

[1] N. Klugman, J. Rosa, P. Pannuto, M. Podolsky, W. Huang, and P. Dutta, "Grid Watch: Mapping Blackouts with Smart Phones," in *15th Workshop on Mobile Computing Systems and Applications (HotMobile '14)*, 2014.

[2] S. Karnouskos, "Crowdsourcing information via mobile devices as a migration enabler towards the SmartGrid," in *International Conference on Smart Grid Communications (SmartGridComm))*. IEEE, 2011.

[3] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing," in *Proceedings of the 18th annual international conference on Mobile computing and networking (MOBICOM)*. ACM, 2012.

[4] T. Clouqueur, V. Phipatanasuphorn, P. Ramanathan, and K. K. Saluja, "Sensor deployment strategy for target detection," in *WSNA '02 Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, Sep 2002.

[5] S. Correa, N. Klugman, and J. Taneja, "How Many Smartphones Does It Take To Detect A Power Outage?" in *The ninth ACM International Conference on Future Energy Systems (ACM e-Energy)*, 2018.

[6] H. Pishro-Nik, *Introduction to probability, statistics, and random processes*. Kappa Research LLC, 2014. [Online]. Available: https://www.probabilitycourse.com

[7] U. Wilensky and W. Rand, *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. MIT Press, 2015. [Online]. Available: http://www.jstor.org/stable/j.ctt17kk851

[8] J. Walzberg, T. Dandres, R. Samson, N. Merveille, and M. Cheriet, "An Agent-Based Model to Evaluate Smart Homes Sustainability Potential," in *28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017.

[9] E. Karangelos and F. Bouffard, "Integrating Demand Response into agent-based models of electricity markets," in *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2012.

[10] V. H. Pakka, B. M. Ardestani, and R. M. Rylatt, "Agent-based modelling of the UK short term electricity market: Effects of intermittent wind power," in *9th International Conference on the European Energy Market*. IEEE, 2012.

[11] A. Adepetu and S. Keshav, "Understanding solar PV and battery adoption in Ontario: an agent-based approach," in *e-Energy '16: Proceedings of the Seventh International Conference on Future Energy Systems*. ACM, 2016.

[12] A. Alyousef, A. Adepetu, and H. de Meer, "Analysis and model-based predictions of solar PV and battery adoption in Germany: an agent-based approach," *Computer Science - Research and Development, vol 32 num. 1*, 2017. [Online]. Available: https://doi.org/10.1007/s00450-016-0304-9

[13] S. Hijgenaar, Z. Erkin, T. Keviczky, J. Siemons, R. Bisschops, and A. Verbraeck, "A decentralised energy trading architecture for future smart grid load balancing," in *International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2017.

[14] J. Kays, A. Seack, and C. Rehtanz, "Consideration of smart-meter measurements in a multi-agent simulation environment for improving distribution grid planning," in *Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2016.

[15] KNBS and SID, "Exploring kenyas inequality: Pulling apart or pooling together?" https://www.knbs.or.ke/exploring-kenya-s-inequality-pulling-apart-or-pooling-together/.

[16] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Sep 2014.

[17] S. Sendra, M. Garcia, C. Turro, and J. Lloret, "WLAN IEEE 802.11a/b/g/n Indoor Coverage and Interference Performance Study," *International Journal on Advances in Networks and Services, vol 4 num. 1 and 2*, 2011.

[18] C. A. of Kenya, "Second quarter sector statistics report for the financial year 2016/2017 (october-december 2016)," http://www.ca.go.ke/.