Keywords: machine learning, time series analysis, non-intrusive load monitoring (NILM), energy disaggregation, rural electrification, electrical grid infrastructure

Background: In Sub-Saharan Africa, over 600 million people – more than two-thirds of the population – live without electricity [1]. Despite governments and private donors making significant financial investments in the rural electrification of this region, spending policy decisions are often made without the support of rigorous evidence. As a result, progress in electrifying this region has been slow [2]. While part of the Africa Electrification Initiative (AEI) is to expand power access to a greater number of people, another important component is to make the access both dependable and affordable. Since inclusive economic growth is the most effective means of reducing poverty and yet many economic activities are impossible without reliable and reasonably-priced energy, it is necessary to focus effort on improving the electrical grid infrastructure of Sub-Saharan African nations [3].

In Kenya, power outages are a common occurrence. This is partially due to the inability of the extant power grid to handle surges in demand from its network of users [4]. These demand spikes cause an overload of the grid which results in a power outage. Since utility companies exhaust resources addressing these outages, the unreliable energy becomes even more expensive. This is counterproductive to the goals of the AEI of effectuating universal access to modern power. In order to understand how to stop these power outages, we need to understand when and how customers utilize their access. Currently, there is limited evidence on even basic patterns of energy demand in Africa [2]. By implementing non-intrusive load monitoring (NILM), or energy disaggregation, we can solve this data deficit.

When a home is connected to an electrical grid, electricity use is monitored by the supplier for billing purposes. Although individual appliances are used, only an aggregate signal for electricity use is reported. To track every appliance used in a household would require the attachment of a power monitor to each device. This is potentially very expensive and therefore impractical for large-scale deployment. Since NILM works by deducing the energy consumption of individual appliances using only the aggregate signal, it presents itself as a very cost-effective method for load disaggregation as just a single sensor is required for monitoring. Using the results from NILM, we can extract energy use tendencies and, in future works, accommodate for them.

Research: Implement a dynamic, innovative approach to non-intrusive load monitoring in the context of rural Sub-Saharan African communities. I will use Kenya as the case study for this research due to availability of TIER lab data for this geographic area and its rapid growth in rural connectivity [1]. I plan to use artificial neural networks (ANN) for individual appliance signal classification and the metaheuristic particle swarm optimization to compute optimal parameters for each network. ANNs have been shown to be cutting edge in neighboring classification fields such as image processing, but their application to NILM using particle swarm optimization has yet to be exhaustively explored [5].

This approach will be compared to the results of classification models for steady-state signature recognition algorithms including support vector machines and K-Nearest Neighbors, aggregate signal extraction techniques like singular spectrum analysis and wavelet decomposition, and single-channel source separation approaches such as discriminative disaggregation sparse coding. By comparing my approach to methods such as these which have been reported to effective for NILM, I will be able to attain a more comprehensive understanding of the energy tendencies of rural Kenyan communities [6].

Unsupervised machine learning algorithms for NILM will also be investigated because they eliminate the need of using richly labeled training data, which is difficult to acquire for rural communities. I will investigate single-channel blind source separation interpreted in the light of signal processing. This is an area of interest because non-negative matrix and tensor factorization techniques have successfully been applied for blind source separation tasks in other research domains such as audio source separation [5, 6].

Data will come from three primary sources: 1) existing data already collected by UC Berkeley's TIER lab in a 2014 field study, 2) data presently being collected by the TIER developed smartphone application Grid Watch, and 3) meters to be deployed in locations throughout Kenya based on the conclusions of my honors in mathematics research thesis and UC Berkeley REU. Existing data comes from diverse and widespread rural Kenyan communities, as will future data.

Since no global performance metric is known by the scientific community for the energy disaggregation problem, I will use a variety of statistics – such as F-measures, Matthew's correlation coefficient, normalized disaggregation error, and root mean square error – to determine the effectiveness of tested models. Results will be visualized wherever appropriate for greater understanding by a broader audience, such as policy makers.

Broader Impacts: By implementing NILM for Kenyan rural electrification, we can gain insight into the behavioral patterns of the electrical grid's users. Using this newfound knowledge, response models for load shedding can be proposed and then incorporated into the Kenyan electrical grid infrastructure to reduce outages due to demand surges. We can also use the observed statistical patterns in collected data to better inform decisions regarding the allocation of electrification investment funds. The conclusion of this thesis has the possibility to greatly further the AEI by advancing research in energy disaggregation for rural African communities. This will, in turn, lead to the potential for cheaper and more reliable energy for this region of the world.

Given my research experience in machine learning and scientific visualization, involvement in interdisciplinary projects with academics from different fields from my own, time living and working abroad with non-English speaking peers and colleagues, and connection to the TIER lab at UC Berkeley, I feel confident in my ability to complete this research proposal over the course of a Ph.D. program.

References:

[1] IEA (International Energy Agency). 2015. World Energy Outlook. http://www.worldenergyoutlook.org/resources/energydevelopment/energyaccessdatabase/.

[2] Lee, Kenneth et al. "Electrification For "Under Grid" Households In Rural Kenya". Development Engineering 1 (2016): 26-35. Web.

[3] Akinlo, Anthony E. "Energy consumption and economic growth: evidence from 11 Sub-Saharan African countries." *Energy Economics* 30.5 (2008): 2391-2400.

[4] "Milestones In Kenya's Electricity Access | Kplc.Co.Ke". Kplc.co.ke. N.p., 2016. Web.

[5] Dong, Roy, et al. "A dynamical systems approach to energy disaggregation." *52nd IEEE Conference on Decision and Control.* IEEE, 2013.

[6] Figueiredo, Marisa. "Contributions To Electrical Energy Disaggregation In A Smart Home". Ph.D. thesis. University of Coimbra, 2013. Print.

Intellectual Merit Criterion

Overall Assessment of Intellectual Merit
Excellent

Explanation to Applicant
The candidate has an impressive background, which includes multiple awards and scholarships. He had several internship opportunities in highly competitive programs and generated results that are well above the expectations from an undergraduate student, including the publication of one paper in a peer-reviewed IEEE conference. The candidate is clearly a driven individual who will thrive in a graduate program. All these accomplishments are significantly more impressive considering that the candidate is a first-generation college student, an issue that, although underplayed by the candidate, drastically decreases his chances of academic success. Although his research theme needs some refinement, the candidate does have the potential to advance knowledge significantly in the proposed topics.

Broader Impacts Criterion

Overall Assessment of Broader Impacts
Excellent

Explanation to Applicant
The candidate is clearly concerned with his impact on society. From his previous experiences in a congressional committee to his research topic focused on electrification in Kenya and his goal to serve in the United Nations, the candidate demonstrates that his pursuit of a graduate degree is not simply an opportunity to advance his career.

Summary Comments
Mr. O'Brien is a focused and hard-working student who has been dedicated to building a strong CV while addressing issues of societal importance. His research topic has potential and with proper mentoring may have significant impact.

Intellectual Merit Criterion

Overall Assessment of Intellectual Merit
Excellent

Explanation to Applicant
The applicant has strong research record in the proposed area and has also had international research experience.

Broader Impacts Criterion

Overall Assessment of Broader Impacts
Excellent

Explanation to Applicant
The applicant's work on using ML and NLP to determine power meter deployment in rural regions has the potential of benefitting the rural community.

Summary Comments
The applicant has demonstrated strong motivation and expertise on research and has proven to engage in extracurricular activities that advance wellness of international and rural communities. Letter writers have given high marks on the applicant's

research and academic preparation.

Intellectual Merit Criterion

Overall Assessment of Intellectual Merit

Excellent

Explanation to Applicant

Mr. O'Brien has been the recipient of numerous scholarships and awards, and has attended workshops and conferences such as the NSF-sponsored XSEDE 2015. He has participated in REUs, including one at Berkeley, and has actively sought out opportunities to supplement the opportunities available to him. His projects demonstrate creativity and resourcefulness. His letter writers speak highly of him and attest to his ability to pursue graduate studies.

Broader Impacts Criterion

Overall Assessment of Broader Impacts

Excellent

Explanation to Applicant

Mr. O'Brien has diverse experience beyond the lab, with a semester abroad in Peru, where he volunteered as a tech consultant, and participation in a peer-to-peer tutoring program for students from underrepresented groups. He served as a project analyst for the U.S House of Representatives. His overall goal is to use ML to understand and address the needs of emerging regions (using electrification in Kenya as a case study for his dissertation).

Summary Comments

Mr. O'Brien's academic and volunteer activities suggest that he will have a promising career carrying out work that will make both intellectual and societal contributions.