Towards automatic classification of appliances: Tackling cross talk in EMF sensors using blind source separation techniques

Suman Giri, Mario Bergés
Carnegie Mellon University, USA
sgiri@andrew.cmu.edu

Abstract. Non Intrusive Load Monitoring (NILM) is the method of obtaining information about appliance-level power consumption inside a building from voltage and/or current measurements made at a central location in the electrical system. The system that does this kind of power disaggregation typically relies on some sort of training step to help it recognize what types of changes in the observed signal correspond to what kinds of appliances. To help train NILM systems recognize appliances within the house, auxiliary sensors like Electromagnetic Field (EMF) sensors have been proposed (Rowe et al., 2010). One of the problems with using EMF sensors to automatically recognize an appliance is that of cross talk. Since a typical home constitutes of settings where there are multiple appliances at close vicinity, EMF sensors are prone to picking up unwanted signals from appliances that are not of interest. In this paper, we use blind source separation techniques like Independent Component Analysis to remedy cross talk in EMF sensors.

1. Introduction

George Hart pioneered the technique of disaggregating whole house power consumption into appliance-level detail in the late 1980s (Hart, 2002). The idea was to monitor voltage and/or current at the main circuit that feeds into the house to calculate power, and based on changes in its features -like its magnitude, shape, etc.- decide what appliance caused the change. The ultimate goal was to give consumers feedback on how much each appliance inside the house was contributing towards their electricity bill, so they could take control of their consumption habits and save energy. Multiple studies have looked into the effect of appliance level consumption feedback on energy consumption behavior and reported savings in the range of 10 to 18% (Ehrhardt-Martinez et al., 2010). The knowledge of appliance level consumption is also invaluable to the supply side (electrical utilities) as they can provide personalized recommendations to users, as well as perform demand response and improved load forecasting (Carrie Armel et al., 2013). The information also has the potential to be used in fault detection and diagnosis of appliances (Orji et al., 2010).

Since the inception of the idea in 1986, there have been multiple breakthroughs, in both the hardware and algorithm sides in this field. A comprehensive review can be found in (Zeifman and Roth, 2011) and (Zoha et al., 2012). Lately, the push has been towards sensor-aided NILM: an effort to augment the algorithms in their training and classification phase using additional sensors such as magnetic, acoustic, light, etc. (Kim et al., 2009; Schoofs et al., 2010). In this vein, we explored the possibility of using the electromagnetic field (EMF) sensor proposed by Rowe et al. for automatic classification and training of appliances in a NILM setup (Rowe et al., 2010; Giri and Berges, 2012). Although, the study concluded that the training is feasible, the issue of noise picked up by the EMF sensor from other appliances was not dealt with in these publications.

The cocktail party problem in audio signal processing is very similar to the problem of cross talk in EMF sensors. The cocktail party problem involves a setting in which multiple speakers are speaking simultaneously in a crowded place (say, a cocktail party), and a particular listener is interested in listening to only one of the conversations (Jutten and Herault, 1991). The human brain is adept at solving this problem easily because the ability to identify the source of interest and focus on it while tuning out additional noise is innate. In computers, however, this problem becomes complicated. The problem is inherently a source separation problem where the listener is first trying to disaggregate the combined sound signal into individual components, and then look focus on the ones that are of interest. To tackle this problem algorithmically, a wide variety of methods have been proposed over the years (Jadhav and Bhalchandra, 2008). Independent component analysis (ICA) has been identified as one of the more successful and easy to implement techniques to achieve the goal of signal source separation in the audio domain (Comon, 1994; Hyvärinen and Oja, 2000)

In this paper, we first outline the theory behind using EMF sensors for automatic labeling and training of appliances and the theory of Independent Component Analysis. We then describe our experimental setup and discuss the results from a case study involving three different appliances.

Following this, we discuss the implications of our results in sections that follow, and finally propose ideas for further research.

2. Theory

2.1 Using EMF sensors for appliance detection

The idea behind using EMF sensors to train a NILM system has been explored in (Rowe et al., 2010) and (Rajagopal et al., 2013). The vision was to label EMF sensors by placing them next to particular appliances and then use them as event detectors. In this way, every time the appliance changed state (turned *on* or *off*, for instance), the sensor would detect a change in the amplitude of the magnetic field around it, and convey that information to the central NILM system. Armed with this information, the central NILM system learns to associate the changes it sees at the aggregate level with the appliance that caused it. After a few days of training, the NILM system learns to recognize all the appliances, and hence the EMF sensors can be removed.

Previous work by the authors attempted to extend this idea further by trying to make the labeling process automatic as well (Giri and Berges, 2012). The idea was to have an EMF sensor recognize what appliances it is placed next to automatically based on the shape of the magnetic field signature that it receives, based on the hypothesis that each appliance tends to have a distinct EMF signature. The sensor then relays that information wirelessly to the central NILM platform that is monitoring the aggregate power changes so that it can learn to associate what changes correspond to what appliances. Figure 1 outlines the general idea.

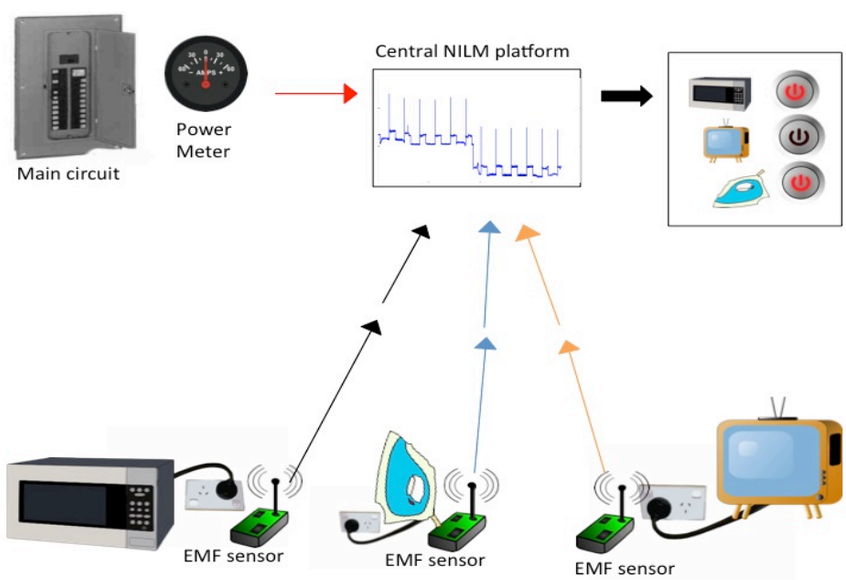

Figure 1. The vision behind EMF sensor aided NILM is to automatically detect magnetic field changes near an appliance and relay that information to the central NILM system so that automated training can occur (Giri and Berges, 2012).

The presence of background magnetic field noise due to other appliances can distort the signature that the EMF sensor receives and hamper its ability to recognize appliances. Imagine a situation like the one depicted in Figure 2, where multiple appliances are in close vicinity of each other. This is a typical situation with many household appliances. Each of the EMF sensors are receiving magnetic field signatures from all the appliances, but the magnitude of the signals coming from each appliance is different in each case depending on how close or far they are from the appliance. Thus, a sensor that is supposed to recognize an appliance based on its isolated magnetic field is not receiving the signature it is trained for. In this case, misclassification can occur from all three sensors.

To remedy this, we propose the use of Independent Component Analysis (ICA). The sensors act as channels that collect data from multiple sources. This lets the problem be framed as a multi-channel blind source separation problem. The only underlying assumption is that the appliance signatures are operating independently of each other, which is a valid assumption in most cases. We run ICA algorithms on the aggregate signal received from each channel and try to disaggregate the source contributions. We then classify the sources and train the NILM algorithm based on that information.

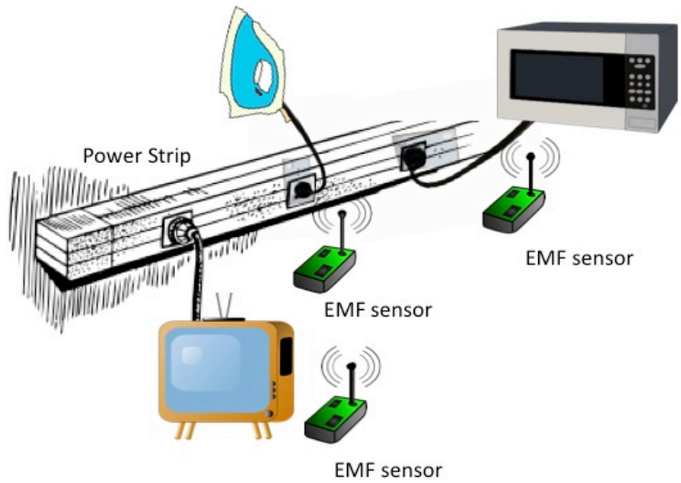

Figure 2 Setup for EMF sensor based training that shows the problem of cross-talk. The sensors place for one appliance can pick up electromagnetic field signals from other appliances in close vicinity, resulting in unwanted noise.

2.2 Source Separation using Independent Component Analysis (ICA)

To illustrate the idea behind ICA, let us assume a simplified version of the cocktail party problem, where two microphones are in place and two speakers are talking. Say, $x_1(t)$ and $x_2(t)$ are the two signals picked up by the microphones at any time instant t. Let $s_1(t)$ and $s_2(t)$ be the two signals emitted by the speakers. Simple law of superposition dictates that the signal observed at the microphone will be a weighted sum of the two source signals. In mathematical notation:

$$x_1(t) = w_1 s_1(t) + w_2 s_2(t)$$
$$x_2(t) = w_3 s_1(t) + w_4 s_2(t)$$

The goal is to estimate the weights w_i and the signal sources $s_1(t)$ and $s_2(t)$ with the available information about $x_1(t)$ and $x_2(t)$. Since there are four unknown quantities and only two equations, some assumptions need to be made to estimate the unknowns. In ICA, the assumption is that the signal sources $s_1(t)$ and $s_2(t)$ are independent of each other. Independence implies that no correlation exists between the variables, and that knowledge of one variable provides no additional information about the other variable. Probabilistically, for any two variables y_1 and y_2, it amounts to their joint probability density function being equal to the product of their individual probabilities i.e. $p(y_1, y_2) = p(y_1)\, p(y_2)$.

To estimate the weights, it helps to visualize the problem stated above in matrix form. The observed aggregated signals and their decomposition into weighted sources can be written simply as $X = W S$, where W is the mixing matrix, S is the vector of all source signals, and X is the vector of observed signals at each microphone. The task of estimating W and S is done by first initializing W^{-1}, and then iteratively calculating the vectors $S = W^{-1} X$ to minimize some cost function that calculates mutual information between the source signals. This is equivalent to maximizing a cost function that calculates non-Gaussianity. A rigorous treatment of these derivations and variant algorithms for convergence can be found in (Hyvärinen and Oja, 2000)

Although the setup described above was for noiseless and linear case, variations of ICA exist that can tackle cases with noise and non-linearity. The basic idea of independence, however, is central to all these approaches.

3. Case Study

To validate the theoretical assumptions that we made about being able to separate the sources of magnetic fields based on input received at multiple sensors acting as channels, a case study was conducted. The experimental setup consisted of a power strip with three different appliances attached to it. The appliances were – a bulb, a vacuum cleaner and a fan. These appliances were

chosen because they represent different classes of electric loads, namely linear (Bulb), non-linear (Vacuum and Fan), inductive (Fan and Vacuum) and resistive (Bulb). The same EMF sensors used in previous work by the authors (Rowe et al., 2010; Giri and Berges, 2012; Rajagopal et al., 2013) were used here and one sensor was placed closed to each appliance. The sensors were then connected to a National Instruments data acquisition card NI-9215, which converts the analog Magnetic Field information collected by the solenoid in the sensor to a digital form that the computer can understand. To make sure that the results of this finding were general enough that we can attempt translating this to a wireless platform, a sampling rate of 1 KHz was used. The card sampled readings from all three sensors simultaneously. LabVIEW programming language was used to implement a simple program to obtain the sampled waveforms from the NI card. Figure 3 shows a snapshot of the setup.

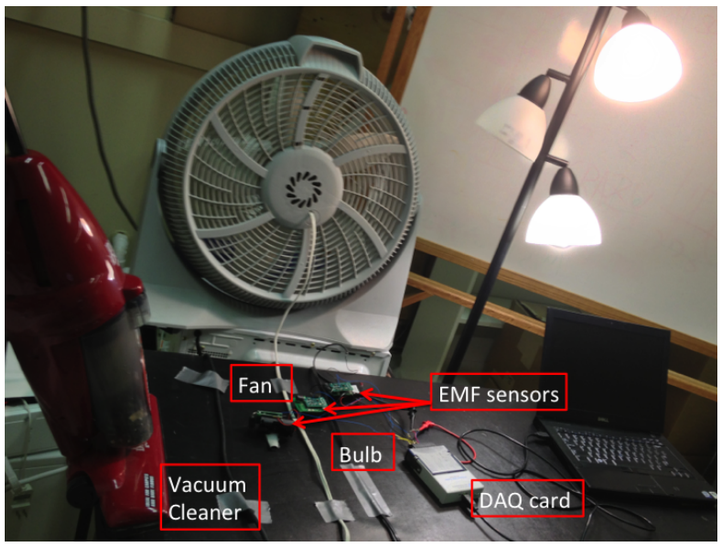

Figure 3. Experimental setup for the case study. The EMF sensors were placed next to appliances and connected to a NI-DAQ card. The DAQ card then was interfaced with a computer using LabVIEW programming language.

Magnetic Field measurements were collected for 10 seconds first with all appliances off (Figure 4), with the goal of capturing background noise. Then the same data was collected from all the sensors with only appliance turned *on* at a time, for 10 seconds (Figure 5). This was done to have some sort of ground truth as to what the appliance signatures should look like. Then all three appliances were turned *on*, and data was collected for 10 seconds again (Figure 6).

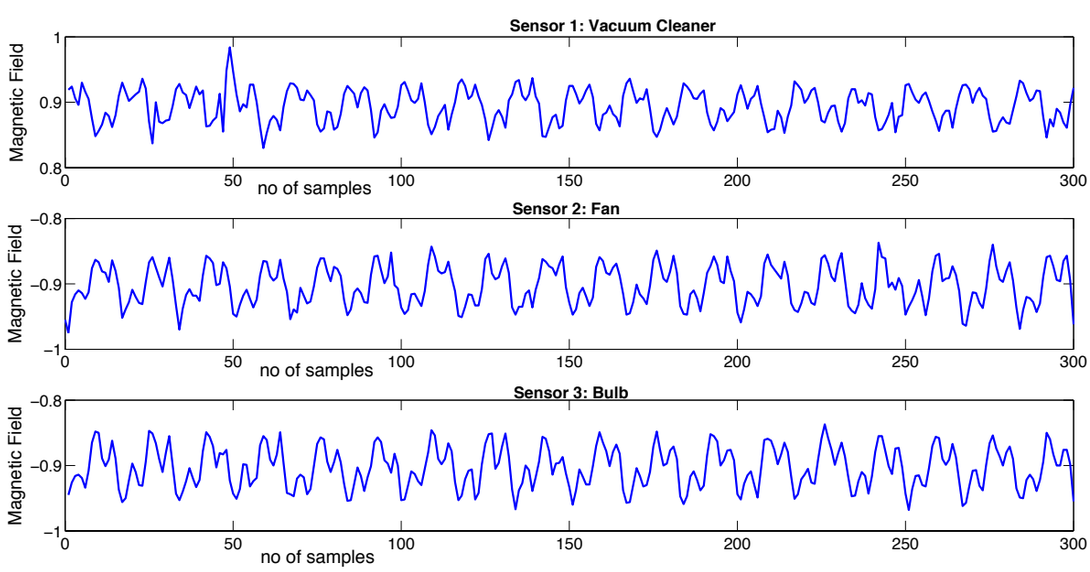

Figure 4. Readings from each of the sensors when none of the appliances were turned *on*. The reading here corresponds to noise. Sensors are numbered according to their proximity to the DAQ card in Figure 3, with 1 being the closest. The EMF sensors were not calibrated, so the magnitude of the magnetic field readings do not correspond to standard units.

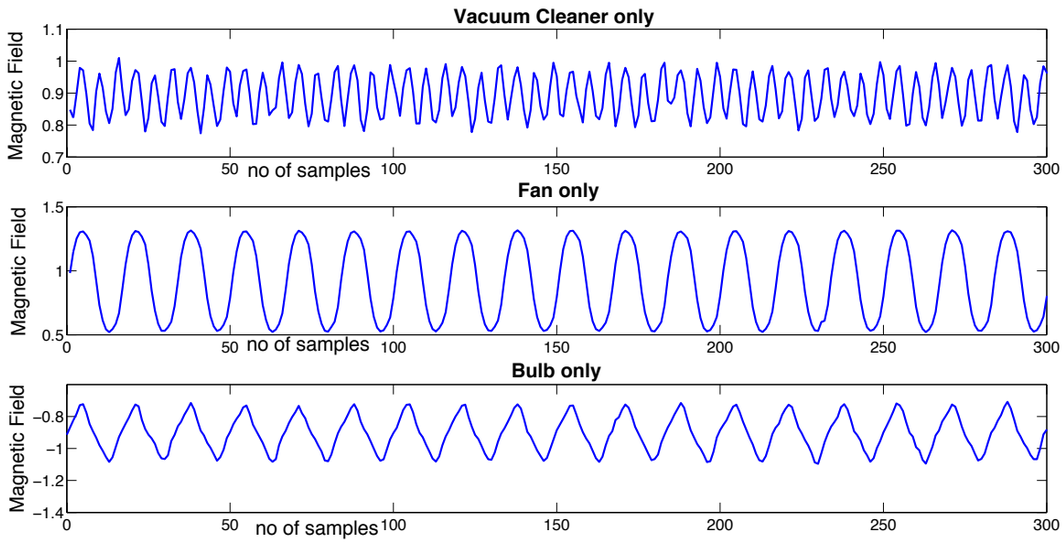

Figure 5. Magnetic Field signatures collected from each appliance for ground-truth purposes. Rows 1, 2 and 3 respectively have information from sensors 1, 2 and 3 when only the Vacuum cleaner, Fan and Bulb were operated- one at a time.

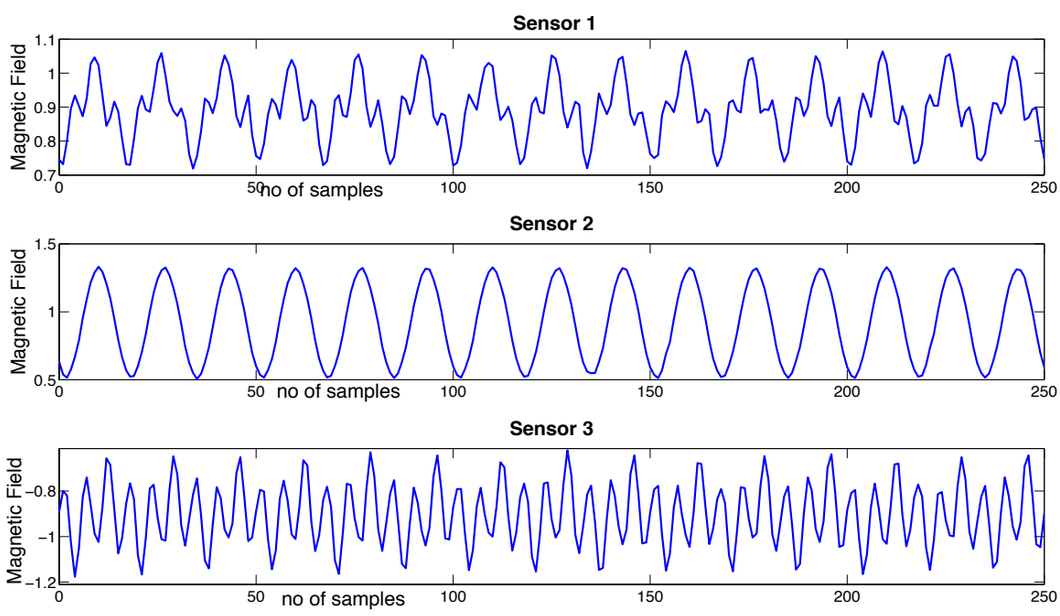

Figure 6. Rows 1, 2 and 3 correspond to readings from sensors 1, 2 and 3 when all the appliances were turned *on*.

4. Results

4.1 Source Separation

Data collected from all three channels when all appliances were *on* (Figure 6) was then used to perform ICA. The fastICA package in MATLAB was used for this purpose (Gävert et al., 2005). A variant of the algorithm that estimates the independent component one-by-one using hierarchical decorrelation (deflation) was used first (Hyvarinen, 1999). Upon visual inspection, it is easy to conclude that the three independent components extracted corresponded to noise, vacuum and fan. This is tested rigorously in section 4.2. Figure 7 details the results.

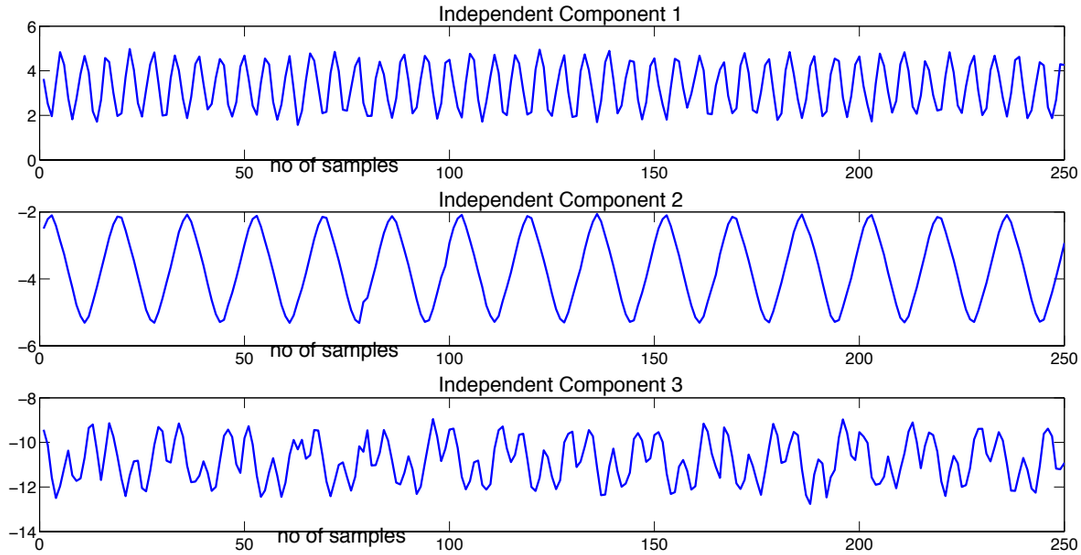

Figure 7. Independent Components (IC) obtained from ICA on data from all three sensors using hierarchical decorrelation. On comparison with Figure 6, the first IC is clearly a Vacuum Cleaner, the second is a Bulb, and the third is noise. The y-axis doesn't necessarily correspond to any quantifiable unit.

Then another method that estimates all the independent components in parallel, with symmetric decorrelation was used (Hyvärinen and Oja, 1997). This disaggregated the signals into contributions from noise, vacuum and bulb. Figure 8 details the results for this case. In conclusion, the contributions of all the individual appliances were retrieved based on the aggregate signals at the three different channels.

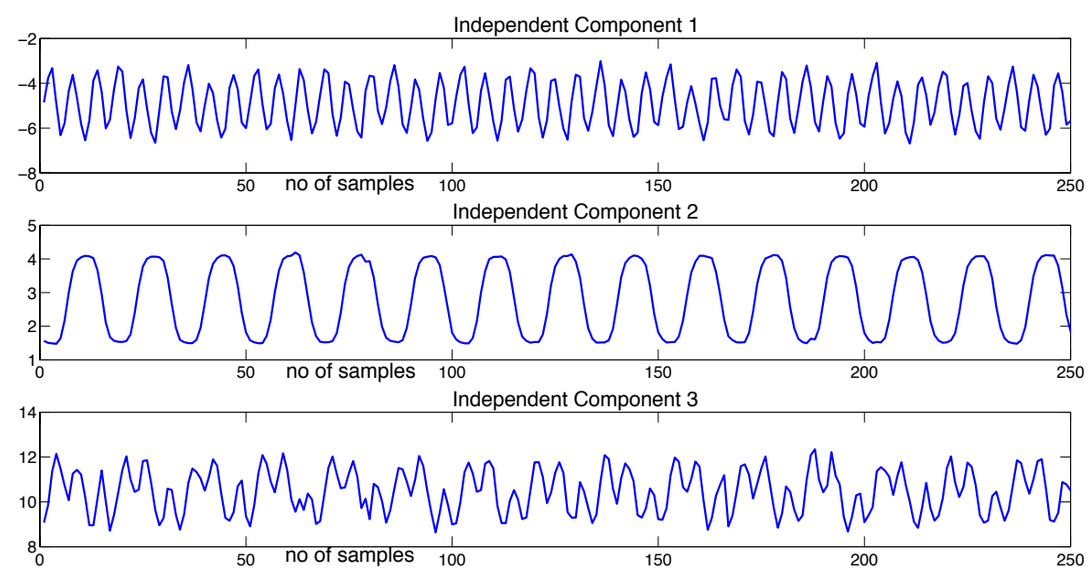

Figure 8. Independent Components (IC) obtained from ICA on data from all three sensors using symmetric decorrelation. On comparison with Figure 6, the first IC is clearly a Vacuum Cleaner, the second is a Fan, and the third is noise. The y-axis doesn't necessarily correspond to any quantifiable unit.

4.2 Classification

To check how close the disaggregated signal was to the ground truth we used a simple k-nearest neighbor classifier, with k = 1. The algorithm works by assigning the class label of the closest training data (in terms of its Euclidean distance) to the test data. Upon real world implementation the EMF sensor will have a database of signatures that tells it what each appliance's magnetic field signature looks like. To simulate that, we used the signatures collected in isolation (Figure 5) as training data. We collected 50 sample points (3 periods of a 60 Hz signal) of the magnetic

field of each appliance and centered it by correlating it with three periods of a 60 Hz sine wave. That was used as the feature (or signature) for training and classification. Figure 9 shows what the features looked like after processing. Fifty sample points were also extracted from the Independent Components, and were scaled and centered appropriately (Figure 9). The process was repeated for 10 different segments of the data, with perfect classification results.

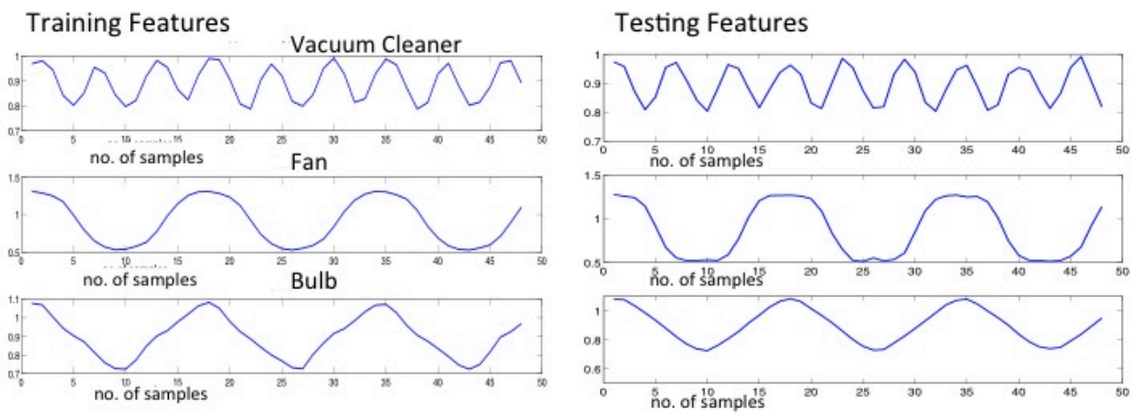

Figure 9. Training and testing features. 50 samples (Three 60 Hz periods) were extracted from the signal to act as training features and centered using a 60 Hz sine wave. Same was done for the Independent Components extracted from sensor data, which was used for testing. A nearest neighbor classifier was used for classification.

5. Discussion and Limitations

From Figure 6 it is obvious that any algorithm trained for signals from Figure 5 would fail to recognize the signatures. Hence cross talk and unwanted noise can seriously hamper the ability of an EMF based NILM training system to recognize appliance and consequently train the system. ICA provides a convenient way to get around the problem. The basic assumption in ICA is that the signals that the sensors receive are independent of each other. Given that it was possible to extract the source signals from noisy data, it can be said that the magnetic field of the three appliances were operating independently of each other. That we were also able to extract the noise, which is independent of the operation of the appliances, adds to the credibility of our model.

Implementing this concept of first separating the signals in a real time environment will need several practical considerations. Computational power and battery life are things that need to be considered when using EMF sensors for recognizing appliances that they are placed next to. A practical work around might be to have the sensors send any information they gather from appliances to a central gateway (which could be a computer) where the processing, including source separation- if necessary, can be done. To make sure that the sensors are not operating and wasting battery life during times when they are not needed, an event based waking system can be adopted i.e. every time an event (considerable change in power consumption according to a predefined threshold) is noted at the central main circuit level, all the nodes should be woken up and asked to send magnetic field signals. Based on the received signatures and knowledge of all other appliances that are already operating, the system can decide what appliance just changed states.

Knowing when to separate the signal into sources and when not is an integral part of the setup. Currently our proposed method is unable to automatically decide which sensors are placed close to each other and are thus receiving data from multiple appliances. So, the user has to label beforehand what sensors are at a risk of getting corrupted signal. Similar problem arises in cases when a single sensor is used but there are multiple appliances that might be operating simultaneously in its vicinity. So, multiple channel source separation techniques like ICA will not be useful there. Single channel source separation techniques like complex matrix factorization or non-negative matrix factorization can be clever ways to tackle this.

Another issue worth discussing is the varied convergence of the two ICA solutions. Depending on how many of the required appliances have been identified, both the algorithms can be used in tandem. However, the topic of which algorithm is better suited for what kind of appliances needs more research.

6. Conclusions and Future Work

In this paper we outlined the current state of Non-intrusive Load Monitoring (NILM) techniques and the motivation for using EMF sensors in NILM platforms. Then we focused on the issue of cross talk in EMF sensors and proposed Independent Component Analysis (ICA) as the way of tackling this problem when there exist measurements taken from multiple EMF sensors in close proximity. We presented an experimental setup to test these ideas and provided results from a case study that involved three appliances. In this case study we determined that using a simple combination of ICA and k-Nearest Neighbors could perfectly identify all three appliances even when the signals measured from each sensors included noise from different appliances.

The results presented here were based on a limited case study, and future studies should be aimed at adding the number of appliances and hence, complexity to the system so as to validate the findings. That we were able to effectively extract the signals and classify is promising and encouraging. To simulate a more realistic setting, the nearest neighbor classifier could be modified so that if the nearest neighbor were farther than a certain threshold, the signature in question would be classified as noise. This would be crucial in settings where the system has no idea of what the noise should look like. More advanced blind signal separation techniques like complex matrix factorization should also be tried to see if it produces better results.

Acknowledgement
The authors would like to thank Niranjini Rajagopal of Carnegie Mellon University for her invaluable help in the hardware setup and data collection.

References

Carrie Armel, K., Gupta, A., Shrimali, G., Albert, A., 2013. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 52, 213–234.

Comon, P., 1994. Independent component analysis, A new concept? Signal Processing 36, 287–314.

Ehrhardt-Martinez, K., Donnelly, K., Laitner, J.A. "Skip," York, D., Talbot, J., Friedrich, K., 2010. Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Economy-wide Electricity Savings. (No. E105). American Council for an Energy-Efficient Economy, Washington, D.C.

Gävert, H., Hurri, J., Särelä, J., Hyvärinen, A., 2005. FastICA package for Matlab. Helsinki, Finland.

Giri, S., Berges, M., 2012. A study on the feasibility of automated data labeling and training using an EMF sensor in NILM platforms, in: Proceedings of the 2012 International {EG-ICE} Workshop on Intelligent Computing. Herrsching, Germany.

Hart, G., 2002. Nonintrusive appliance load monitoring. Proceedings of the IEEE 80, 1870–1891.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks 10, 626–634.

Hyvärinen, A., Oja, E., 1997. A Fast Fixed-Point Algorithm for Independent Component Analysis. Neural Computation 9, 1483–1492.

Hyvärinen, A., Oja, E., 2000. Independent component analysis: algorithms and applications. Neural Networks 13, 411–430.

Jadhav, S.D., Bhalchandra, A.S., 2008. Blind source separation: trends of new age - a review, in: IET International Conference on Wireless, Mobile and Multimedia Networks, 2008. Presented at the IET International Conference on Wireless, Mobile and Multimedia Networks, 2008, pp. 251–254.

Jutten, C., Herault, J., 1991. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. Signal Processing 24, 1–10.

Kim, Y., Schmid, T., Charbiwala, Z.M., Srivastava, M.B., 2009. ViridiScope: Design and Implementation of a Fine Grained Power Monitoring System for Homes, in: Proc. Ubicomp09, 245–254.

Orji, U.A., Remscrim, Laughman, C., Leeb, S.B., Wichakool, W., Schantz, C., Cox, R., Paris, J., Kirtley, J.J., Norford, L.K., 2010. Fault detection and diagnostics for non-intrusive monitoring using motor harmonics, in: Applied Power Electronics Conference and Exposition (APEC), 2010 Twenty-Fifth Annual IEEE. Presented at the Applied Power Electronics Conference and Exposition (APEC), 2010 Twenty-Fifth Annual IEEE, pp. 1547–1554.

Rajagopal, N., Giri, S., Rowe, A., Berges, M., 2013. A Magnetic Field-based Appliance Metering System. Presented at the Third International Conference on Cyber-Physical Systems, Philadelphia, PA.

Rowe, A., Berges, M., Rajkumar, R., 2010. Contactless sensing of appliance state transitions through variations in electromagnetic fields, in: Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, BuildSys '10. ACM, New York, NY, USA, pp. 19–24.

Schoofs, A., Guerrieri, A., Delaney, D.T., O'Hare, G., Ruzzelli, A.G., 2010. ANNOT: Automated Electricity Data Annotation Using Wireless Sensor Networks, in: 2010 7th Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON). Presented at the 2010 7th Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON), pp. 1 –9.

Zeifman, M., Roth, K., 2011. Nonintrusive appliance load monitoring: Review and outlook, in: 2011 IEEE International Conference on Consumer Electronics (ICCE). Presented at the 2011 IEEE International Conference on Consumer Electronics (ICCE), pp. 239 –240.

Zoha, A., Gluhak, A., Imran, M., Rajasegarar, S., 2012. Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. Sensors 12, 16838–16866.