2011 IEEE/ACM International Conference on Green Computing and Communications

The Input-Sensing Problem in Ternary Computing and Its Application in Household Energy-saving

Jingjie Liu[#*1], Lei Nie[#*2], Zhiwei Xu[#3]
Institute of Computing Technology, Chinese Academy of Sciences
*Graduate University of Chinese Academy of Sciences
{[1]liujingjie09, [2]nieleimail}@gmail.com, [3]zxu@ict.ac.cn

Abstract—**This paper formulates and studies the problem of accurately acquiring energy consumption information of physical objects influenced by human behavior into the cyber space. We formulate this *input-sensing problem* within the ternary computing framework, which allows real-world problem instances to be studied, with different constraints on human efforts, sensors needed, error bounds, and computational complexity. We focus on the input-sensing problem commonly found in the energy-efficient design and use of household electric devices (appliances). The main challenge is how to distinguish the electric currents of individual appliances even with only a single sensor. We recast this current disaggregating problem into a computer science problem called the approximate phase space learning problem. We develop a novel technique called the principal component manifold method to solve the learning problem. This method uses a linear manifold, spanned by the principal components, to approximate the phase space of an appliance. Experimental results show that our approach can trace the dynamic currents of the appliances with continuously variable load, and estimate the power consumption values with low errors.**

Keywords-ternary computing; input-sensing problem; current disaggregating; approximate phase space learning;

I. INTRODUCTION

Green computing is a comprehensive and far-reaching new field. It includes at least two objectives: (1) to make computing systems (or IT systems) themselves more energy efficient and (2) to use information technology to make other non-IT systems more energy efficient. This paper relates to the second objective and focuses on the input information sensing problem commonly found in the energy-efficient design and use of household electric devices (also called appliances), such as lamps, air conditional units, washers, refrigerators, computers, TV sets, etc.

To save household electricity consumption, there are two approaches: (1) to optimize household behavior of using electric devices and (2) to produce and deploy better, more energy-efficient electric devices. In both cases, we need to acquire the electricity usage information on each electric device, which originally exists in the physical world.

In studying the long-term and fundamental trends of China's information technology development into 2050, Chinese Academy of Sciences presents two new concepts called *ternary universe* and *ternary computing* [12][19]. The concepts state that future computing systems will utilize resources in the ternary universe of the human society, the cyberspace, and the physical world in problem solving computational processes. Ternary computing calls for enriching the traditional computer science, to solve computational problems in this human-cyber-physical ternary universe.

At a high level, the household electricity saving problem can be described in a ternary computing framework illustrated in Fig. 1. Three computational processes are shown. An *input-sensing process* acquires the electricity usage information of each electric device in a household. This information from many households (e.g., the hundreds of millions of households in China) can then be used as *input data* by two other computational processes, one to optimize the household users' behavior, and the other to upgrade electric devices. This is why the first process is called the *input*-sensing process.

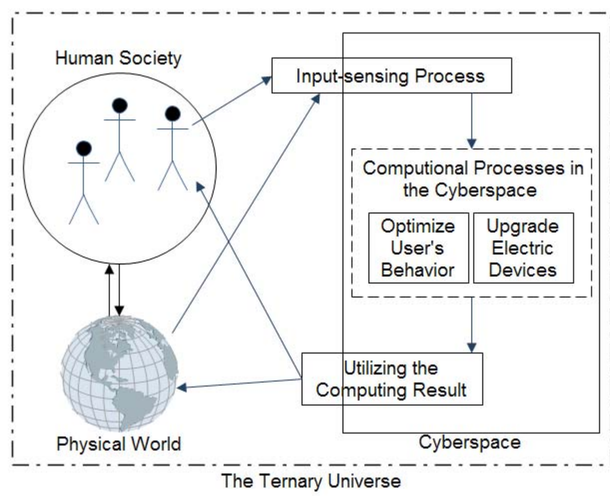


Figure 1. Illustration of ternary universe and ternary computing

The information produced by a good input-sensing process is valuable. For example, American homes consume 38% of the total electricity used in 2009 [20]. Households in Beijing (in China) consume 16.8% (about 13%) of the total electricity used in 2007. A recent study by the China central government reports that Chinese households each in average consumes 87 KWH per month, which is 45% higher than modern but environment-conscious homes in urban Beijing. If the input-sensing information is utilized to change the user behavior and thus reduce electricity consumption by 20%, we could power another IT industry! Accurate electricity usage data of

This work is supported in part by the Frontier Research Project of Chinese Academy of Sciences, the China Basic Research Program and the major national science and technology special projects (2011CB302500, 2011CB302800, 2010ZX03004-003-03).

individual appliances could be crucial information in researching new appliances.

China State Grid is in the process of installing over 200 million smart meters in Chinese households. But these meters can only acquire information of the total electricity used by a household, not each individual appliance. Their sampling rate is only sufficient for the measurement of macro-information, such as energy consumption and power factor. What is needed is an effective and efficient input-sensing process that can acquire accurate information of each appliance, such as higher harmonics in current signal. This poses scientific and technical challenges. The following questions must be asked:

- How to formulate the input-sensing problem more precisely, to reflect that the input-sensing process is a human-cyber-physical ternary computational process? That is, human behavior, computing systems and physical world elements are all involved.
- How to recast the input-sensing problem into a computer science problem so that we can solve it computationally?
- Answering the above two questions amounts to a methodology for solving the input-sensing problem. The third question is: how to apply the methodology to the household electricity saving use case and to validate the proposed methodology?

This paper makes the following three contributions:

- We formulate the input-sensing problem to capture the ternary computing nature, with four performance evaluation criteria called *effort complexity*, *sensor complexity*, *coding error*, and the traditional *computational complexity*. For the electricity saving application, we also formulate an electric current disaggregating problem as an instance of the input-sensing problem.
- We recast the current disaggregating problem into a computer science problem called the approximate phase space learning problem.
- To solve the learning problem, we develop a novel technique called the principal component manifold method. This method uses a linear manifold, spanned by the principal components, to approximate the phase space of an appliance. Experimental results show that this method can trace the dynamic currents of the appliances with continuously variable load, and estimate the power consumption value with the relative error less than 0.6%.

The rest of this paper is organized as follows. Section II discusses the input-sensing problem in ternary computing, and gives a limited instance for energy-saving application. Section III reduces the current disaggregating problem to the approximate phase space learning problem. Section IV presents the principal component manifold method and its algorithm implementations. Section V presents performance evaluation experiments and results. Section VI discusses related work. Section VII offers concluding remarks.

II. Input-sensing Problem

Considering the input-sensing process, we define the **input-sensing problem** of ternary computing as how to gather information from the physical world or/and the human society, and to transmit it into the cyberspace. To study this problem, we introduce a typical scenario of the input-sensing problem in the ternary computing as Fig. 2(a), and construct a specific scenario for household energy-saving application as Fig. 2(b).

In a typical scenario shown in Fig. 2(a), there are three main elements involved in the whole process: the people in the human society, the things in the physical world and a Turing machine in the cyberspace.

The input-sensing process in Fig. 2(a) starts with the establishment of an application of ternary computing, which determines the format of result data of this input-sensing process. We assume that useful prior knowledge in the cyberspace is already coded on the tape. Next, the machine guides people to operate the devices in physical world, and to input the information from the subjective judgments. Through a computing process, the result of this input-sensing process is provided as the input of subsequent applications.

In contrast to the traditional computing paradigm, people in the ternary universe interact with both the physical world and the cyberspace. The machine does not just process the information on the tape, but also guide people and other devices, to acquire information.

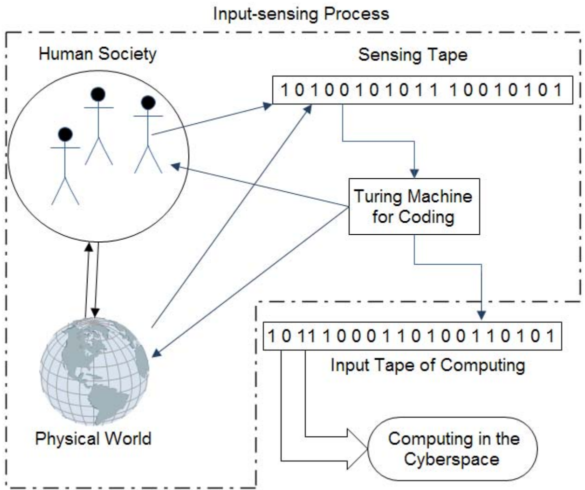
To measure the differences among different input-sensing processes quantitatively, we define four concepts: The *effort complexity* is the total amount of extra effort spent by humans in an input-sensing process. The *sensor complexity* is the maximal number of sensors that are running simultaneously in an input-sensing process. The *coding error* is the difference between the coded information and the true information. The usual *computational complexity* is the total amount of time and space used in an input-sensing process.

The four concepts are important evaluation criteria for the input-sensing process. Together they cover all of three aspects of the ternary universe.

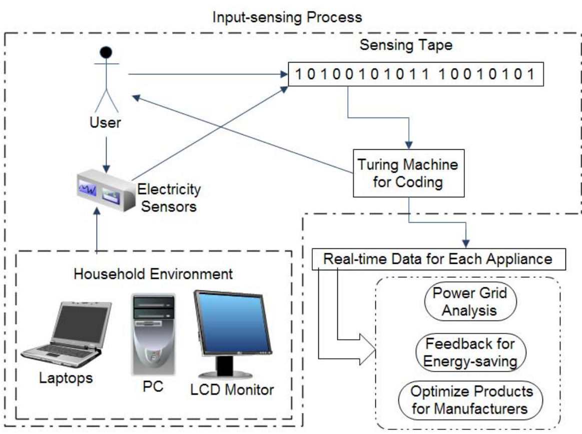
Household energy-saving is an urgent problem nowadays. It leads to a specific scenario of input-sensing problem shown in Fig. 2(b), where the real-time electrical data of each individual appliance is the crucial information [1][14].

The real-time electrical data includes currents and voltages. As [20] pointed out, the voltage distortion caused by the appliances is usually less than 5%, whereas the current distortion can be as high as 150%. Furthermore, the voltages of each appliance are equivalent. Consequently the voltages of each appliance can be approximated by the voltage measured at the main electric entrance. The vital task is to deal with the current vectors.

The input-sensing problem for the household energy-saving application is how to gather the real-time current of each individual appliance and transmit it into the cyberspace. In this problem, the active elements are the electricity sensors, the user and a machine; the object to be measured is the circuit of the whole house; and the output of this input-sensing process is the real-time current of each individual appliance.



(a) A typical scenario of the input-sensing problem



(b)The specific scenario for household energy-saving

Figure 2. The scenarios of the inpur-sensing problem

With respect to the above input-sensing process for household energy-saving, we can use the four criteria to evaluate its characteristics. The effort complexity in this scenario contains the extra effort that a household user spends sedulously in the input-sensing process. For example, the effort spent on sensor installations and environment configuration, are part of effort complexity, and the effort in the daily use of electrical appliances are not. The sensor complexity in the scenario is the same as its definition. The coding error is the error on the real-time data for each appliance. The space complexity is the total amount of the used space on all the electronic devices. The time complexity is the time interval between the start time of the input-sensing process and the time when the first outcome is output.

To solve this problem, there are usually two categories of methods to be used: intrusive load monitoring (ILM) and nonintrusive load monitoring (NILM). ILM usually contains plural sensors [14]. For each appliance, an individual sensor needs to be deployed by people. In contrast to ILM, NILM uses only one sensor at the main circuit entry. Every appliance needs a configuration for its learning environment. Because there are some wrong decisions while determining the operating states of the appliances, the result of NILM usually has a relative error around 10% [7].

ILM can get more accurate data than NILM, but has higher sensor complexity. NILM is more attractive than ILM in that it potentially needs only one sensor. But previous NILM has a core weakness: there is no effective approach to address the appliances with continuously variable load (CVL) [5]. The error of the result increases obviously [7] when there are several appliances with CVL.

To deal with multiple appliances with CVL, we limit the condition of the input-sensing problem for household energy-saving, and define a specific instance called the *current disaggregating problem* as follows.

The current disaggregating problem: In the household environment, there are c appliances with CVL. Given the operating states, the training data set for each appliance and the data measured by a fixed sensor at the main electric entrance, calculate the real-time current for each appliance with CVL to minimize the coding error.

In this problem, the sensor complexity is 1; the effort complexity is the total amount of extra effort to obtain the training data set; the usual computational complexity and the coding error are related to the method of solving this problem.

In the rest of this paper, we try to find a method to solve the current disaggregating problem. If this method exists, we can integrate it with other methods of NILM, and then we can solve the input-sensing problem for household energy-saving with low sensor complexity and high accuracy.

III. PROBLEM FORMULATION

In the previous section, we discuss the input-sensing problem and focus on its instance called the current disaggregating problem. In this section, we reduce this problem to a computer science problem called the approximate phase space learning problem.

In the household environment, all of the currents are alternate current. For convenience, we decompose a current waveform into a real trigonometric polynomial. Then the coefficients in this polynomial contain all information of this current waveform, and can be calculated via Fourier analysis. Because of the finite sampling frequency, the number of coefficients is finite. To indicate a cycle of the current waveform, we use a vector of coefficients called *current vector*, which is in the real N-dimensional vector space R^N. The constant N is determined by the sampling frequency.

We define the scenario of the current disaggregating problem as follows:

- There is a training data set L_k for the each of the c appliances with CVL. Each element l_{jk} in L_k is a current vector.

- There is a mixed current set I. Its element I_j is a mixed current vector measured by the fixed sensor at the main electric entrance. Suppose m stands for the size of I.
- The result of this problem is the disaggregated result set R. Each element r_{jk} in set R is a disaggregated current vector of the kth appliance corresponding with I_j in set I.
- There is a validation set $V=\{v_{jk}\}$ to verify the disaggregated result set R. Each element v_{jk} in set V is an actual current vector corresponding with r_{jk} in set R.

According to Kirchhoff's current law, the current at the main electric entrance equals to the sum of the currents of all appliances. This law is also applicable to the current vectors in vector space R^N.

Based on the definition of phase space, every current vector of an appliance must be in the phase space of the appliance's current vectors. The phase space P is usually much smaller than the N-dimensional vector space R^N. Thus, the current vectors must satisfy Equation (1).

$$\begin{cases} I_j = \sum_{k=1}^{c} r_{jk} \\ r_{jk} \in P_k \text{ for any } j,k \end{cases} \tag{1}$$

Here, I_j is a current vector in set I, and r_{jk} is a corresponding current vector of the kth appliance. P_k is the phase space of the kth appliance's current vectors. The first row of Equation (1) is the Kirchhoff's current law; the second row is the definition of phase space.

In Equation (1), $r_{jk}=v_{jk}$ (k is from 1 to c) is a solution. If Equation (1) has a unique solution, it equals to the actual current vectors for each appliance. Then the key point is to learn the approximate phase space S_k for each appliance from the training data set and to ensure Equation (2) with a unique solution:

$$\begin{cases} I_j = \sum_{k=1}^{c} r_{jk} \\ r_{jk} \in S_k \text{ for any } j,k \end{cases} \tag{2}$$

The solutions of Equation (2) for every I_j forms the result set R. To evaluate the difference between the result set R and the validation set V, we define the root mean square error:

$$\mathrm{RMSE}(R,V) = \sqrt{\sum_j \sum_k (r_{jk} - v_{jk})^2 / (c \times m)} \tag{3}$$

Then we can reduce the current disaggregating problem into a learning problem called **the approximate phase space learning problem**. This problem is defined as follows:

Given the training data set L_k for each appliance k and the mixed current set I, find approximate phase spaces S_k for each appliance to minimize the RMSE between the validation set V and the result set R that is the solution according to Equation (2).

IV. PRINCIPAL COMPONENT MANIFOLD METHOD

To solve the approximate phase space learning problem, we need to find a group of proper approximate phase spaces according to the training data set to minimize the RMSE between the validation set and the result set. In this section, we provide a novel method to give the approximate phase spaces. This method is called the principal component manifold method.

A. The Manifold and Its Benefits

In our method, we limit the approximate phase space S_k to be a linear manifold M_k in R^N. A manifold M_k is a linear subspace shifted away from the origin. Then we have the linear formula of the element in S_k:

$$\forall r_{jk} \in S_k,\ r_{jk} = A_k i_{jk} + \mu_k \tag{4}$$

Here, A_k is a transformation matrix from a subspace corresponding with the manifold M_k to R^N. μ_k is the offset vector. i_{jk} is a vector in the subspace and r_{jk} is a vector in R^N. By Equation (4), the elements in S_k can be represented concisely. According to Equation (4), we can simplify Equation (2) as:

$$I_j = \sum_{k=1}^{c} r_{jk} = \sum_{k=1}^{c} (A_k i_{jk} + \mu_k) \tag{5}$$

According to Equation (5), we concatenate all A_k and i_{dk} together. Then we get Equation (6):

$$(A_1 \ldots A_c)(i_{j1} \ldots i_{jc})^T = I_j - \sum_{k=1}^{c} \mu_k \tag{6}$$

Comparing Equation (6) with Equation (2), we see that Equation (2) is an instance of the Subset Sum problem that is NP-Complete. In contrast, Equation (6) is a linear equation, which can be solved in polynomial time.

Furthermore, we can decide that whether there is more than one solution in Equation (6). We can give a necessary and sufficient condition that the coefficient matrix of Equation (6) does not have full column rank.

This condition requires that the height of the coefficient matrix must bigger than its width, and the family of all column vectors in this matrix is linear independent. So we have to find the manifolds with low-dimension and approximate enough with the phase spaces.

B. The Principal Component Analysis to Create a Manifold

In statistics, principal component analysis simplifies variance-covariance structure of samples, and accomplishes the multivariate dimension reduction process without any distribution hypothesis. A group of principal components generated by PCA can map a high-dimensional dataset into a most informative lower-dimensional linear manifold, and the offset vector of this manifold equals to the mean of this dataset. To evaluate the difference between the manifold and the original dataset, we can use a concept called explained variation, which is used to indicate the proportion of the variation contained by the manifold in the total variation.

So we can construct the manifold to approximate the phase space by the principal components and the mean of the training data set. The set of principal components is equivalent to a transformation matrix from a subspace corresponding with the manifold M_k to R^N. Every principal component is a basis vector of this subspace. Thus the number of principal components equals to the dimension of the manifold.

Considering the limitation of appliances physical structure, we make *the manifold sparsity hypothesis*:

The current vectors of an appliance with CVL belong to a linear manifold approximately. The dimension of this manifold is far less than the total dimension N. Meanwhile the explained variation of this manifold is more than 95%.

This hypothesis is verified in Experiment 1. The results show that the principal components can form the manifold with much lower dimension and high explained variation.

These low-dimension linear manifolds can be used to construct Equation (6). If the coefficient matrix satisfies the condition of full column rank, the equation does not have more than one solution. Then we can get the approximate solution of the current disaggregating problem.

C. Errors Analysis of This Method

We have discussed the feasibility of constructing the low dimension manifold by the principal components. Then we analyze the error generated by the approximate phase space.

Observing the derivation of Equation (5), we replaced r_{jk} with a linear formula of i_{ik}. We ignore some the non-principal components. Suppose that the non-principal component of the kth appliance is a_k. From Equation (6), we get Equation (7):

$$\left(A_1 \dots A_c\right)\left(i_{j1} \dots i_{jc}\right)^T = I_j - \sum_{k=1}^{c} \mu_k - \sum_{k=1}^{c} a_k \qquad (7)$$

Comparing Equation (6) and Equation (7), the error acts on the mixed current I_j of Equation (6). To indicate the error magnification of Equation (6), we introduce the concept of condition number from the field of numerical analysis.

The condition number of a linear equation gives a bound on how inaccurate the solution x will be after approximate solution. It is just a property of the coefficient matrix. From the definition of condition number, we know that the relative error of solution is smaller than the condition number times the relative error of the right side of Equation (6), as

$$e / x \le \text{condition number} \times \left(\sum_{k=1}^{c} a_k\right) / \left(I_j - \sum_{k=1}^{c} \mu_k\right) \qquad (8)$$

Here x is the solution of Equation (6); e is the error in x; and $\sum_{k=1}^{c} a_k$ is the error caused by ignoring the non-principal components. The upper bound of error e positively correlates with the condition number.

On the other hand, the condition number of coefficient matrix is also relevant to the rank of coefficient matrix. If the coefficient matrix does not have full column rank, then the condition number is infinite. As a conclusion, we calculate the condition number, and base on it to judge whether the solution is a meaningful solution. Only when the condition number is lower than a threshold came from experiments, the solution is meaningful and with an acceptable error.

D. The Algorithm Implementation[1]

Our algorithm implementation is consisting of two phases. The first one is learning phase for obtaining the principal components for each appliance. The next one is the disaggregating phase for calculating the individual currents from mixed current.

In the learning phase, the first is the respective acquisitions for all appliances. Then transform the currents to the coefficients of real trigonometric polynomial by the short-time Fourier transformation. For each training data set, use *SVD* (singular value decomposition) to calculate its principle components. Because there are polynomial time algorithms for SVD, the running time of this phase is polynomial.

In the disaggregating phase, use traditional NILM system to determine the list Q of running appliances. Use the mixed current vector I_j and coefficient matrix A to construct and solve the linear equation according to Equation (6), and get the current of each appliance in Q. Then we can calculate the real-time waveform by the Inverse Fourier Transform. We also can calculate the real-time power consumption directly, or keep this data for further application.

E. Virtual Combined Appliance

Because of the orthogonality of principal components, the condition number of transformation matrix A_k identically equal to 1. In order to reduce the condition number of Equation (6), we provide a conception of virtual combined appliance.

For a set of appliances, we can construct a virtual combined appliance (VCA) whose current vectors are the mixed current vectors of appliances in this set. Then we replace the appliance set with VCA in Equation (6), and disaggregate the current of VCA from total current. This current is the mixed current of appliances in the set. This replacement can reduce the condition number of Equation (6).

There are many different strategies to construct VCA while disaggregating multiple appliances. Different constructional orders of VCA will have different errors, and a better combining strategy leads to a better result. The effects of different strategies need to verify by the experiments.

V. SIMULATION EXPERIMENTS AND RESULTS

A. The Experiment Environment

We use a single sensor to sampling the data for each appliance respectively. Based on these data we do some simulation experiments.

Since our main issue is to disaggregate currents for CVLs, we make an assumption that we have already known the operating state of every appliance in our experiments.

[1] The details of the algorithms are available online at: http://www.box.net/shared/ld2i8an2zj

Following the Kirchhoff's current law, the mixed currents are generated as the sum of the currents of multiple appliances. Thus we know all the true value of each appliance and the mixed value.

The circuit sensor is divided into two parts. One is a current sensor which measures the alternate current indirectly by a magnetic sensor. The other one uses Resistor network and difference amplifier to get voltage directly. The USB Data Acquisition Card USB5931 accomplishes the AD conversion and data acquisition. The sampling frequency is 12800Hz, and the fundamental frequency in China is 50Hz. According to the Nyquist–Shannon sampling theorem, the maximal frequency is 6400Hz and the total dimension N equals to 128. These devices can be implemented in a smart meter.

All the calculation in our experiments is on the matlab platform. In the following experiments, we use two kinds of error to evaluate the experiment result.

1. The root mean square error (RMSE) between result set and validation set is the evaluation criterion for the waveform accuracy of the disaggregating results.
2. The relative error for power is an evaluation criterion for the computational accuracy of power consumption.

These two errors form different perspectives to analyze the results. The simulation experiments are divided into four parts to explore all aspects of the current disaggregating theory.

We choose four appliances for the experiments. They are a desktop PC (without monitor), a LCD monitor, a laptops and a suit of desktop PC with LCD monitor. The last one is selected for building a bad scenario that there are some similar appliance.

As a reference, we give the relative error for power calculated by NILM in same case. The traditional NILM calculates the power consumption by the average power after the determination of operating states for each appliance.

TABLE I. THE RELATIVE ERROR FOR POWER BY NILM

Desktop PC	LCD monitor	Laptops	PC+LCD
0.778%	-0.377%	-1.355%	0.605%

B. *Experiment 1: The Feasibility of Manifold Construction*

The first phase of the method is to calculate the principal components for the construction of manifold, and all the current vectors in the result set are in the manifold. Then compare it with the original current vector to get RMSE and the relative error for power. The result is in Table II.

TABLE II. THE RESLUT OF EXPERIMENT WITH SINGLE APPLIANCE

Appliance	Dimension of the Linear Manifold	Explained Variation	RMSE	The Relative Error for power
Desktop PC	7	97.89%	10.339	0.004%
LCD monitor	10	96.23%	10.643	0.009%
Laptops	10	97.11%	12.404	0.014%
PC+LCD	10	96.25%	11.809	0.004%

The total dimension of the four appliances 254

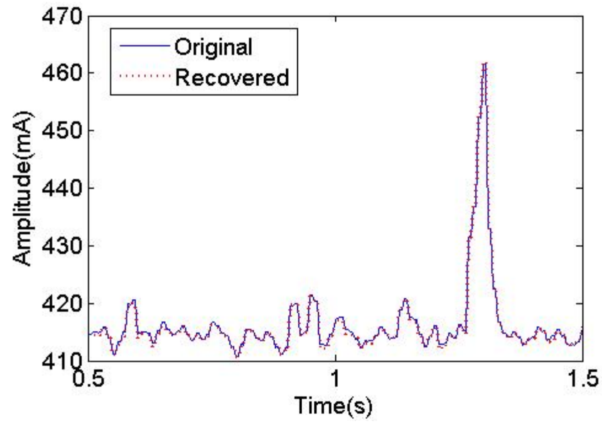


Figure 3. Fundamental harmonic amplitude versus time

To further display this result, we give Fig. 3 to show the curve that the fundamental harmonic amplitude of desktop PC versus time, both original and recovered.

The result in Table II shows that the low-dimension manifolds contain the majority of the information for all appliances. From Fig. 3, the recovered curve is close to the original curve. It means that we can trace the real-time current changes. Meanwhile the relative errors for power calculated by recovered current vectors are much smaller than those calculated by average power in Table I.

C. *Experiment 2: Two Appliances and Condition Number*

Since we have known the error caused by the manifold mapping from Experiment 1, we want to disaggregate the mixed current in Experiment 2. Then we construct several experiments with two appliances. The result is listed in Table III. For comparing the different results in a visible way, we put two groups of results of the desktop PC in Experiment 2 and the original current vectors together, and get Fig. 4.

TABLE III. THE RESLUT OF EXPERIMENT WITH TWO APPLIANCES

Appliance A	Appliance B	Condition Number	RMSE	Errors of A	Errors of B
Desktop PC	LCD monitor	6.857	13.254	12.822	13.675
				-0.000%	0.008%
Desktop PC	Laptops	4.333	13.741	12.977	14.465
				-0.014%	0.038%
LCD monitor	Laptops	3.377	12.675	12.067	13.256
				0.101%	-0.128%
PC+LCD	Desktop PC	21.019	39.172	39.346	38.997
				0.310%	-0.338%
PC+LCD	LCD monitor	31.762	54.997	55.065	54.929
				0.108%	-0.275%
PC+LCD	Laptops	5.124	15.013	14.997	15.029
				0.014%	-0.029%

The first row of the error for each appliance is its RMSE.
The second row of the error for each appliance is its relative error for power.

In Fig. 4, the original curve is the true value of current; the curve with LCD is the results in well condition; the curve with

PC+LCD is the results in bad condition. We can notice that the mean of both results is consistent with the original mean, but there is a greater fluctuation in bad condition. The result in well condition can accurately trace the dynamic evolution of current curve, and the waveform of disaggregated current is reliable.

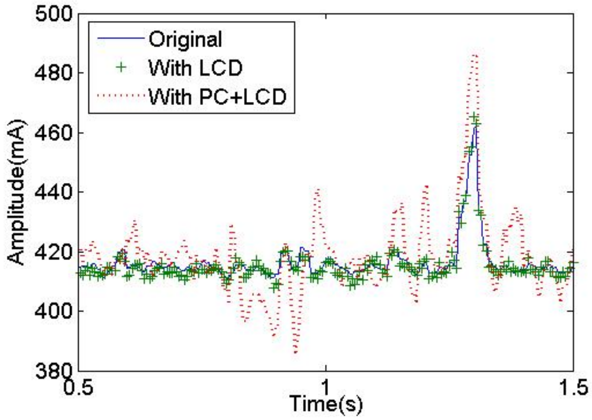


Figure 4. The result of desktop PC in different experiments

D. Experiment 3: Three Appliances and VCA

In the experiments with two appliances, it is not necessary to construct a visual combined appliance (VCA). If there are more than three appliances, the impact of VCA can be demonstrated in the experiment. Our next experiment is to analyze this impact.

In these comparative experiments, we construct three groups of experiments. For each group, there are three appliances. We disaggregate current in two ways. The first is to disaggregate the mixed current directly. The second is to combine the appliance B and C into a VCA, and to disaggregate the mixed current into appliance A and VCA, then to disaggregate the current of VCA. The result is in Table IV.

TABLE IV. THE RESLUT OF EXPERIMENT ABOUT VCA

A	B	C	Condition number	RMSE	Errors of A	Errors of B	Errors of C
LCD monitor	Desktop PC	Laptops	9.673	15.416	15.000	15.763	15.476
					-0.013%	-0.062%	0.187%
			7.256	15.387	14.687	16.047	15.397
					0.080%	-0.043%	0.017%
Laptops	Desktop PC	LCD monitor	9.673	15.416	15.476	15.763	15.000
					0.187%	-0.062%	-0.013%
			5.227	15.552	15.305	16.079	15.257
					0.212%	-0.013%	-0.130%
Laptops	Desktop PC	PC+ LCD	26.111	34.969	15.656	40.961	41.779
					0.008%	-0.363%	0.330%
			5.534	32.162	15.687	37.032	38.545
					0.076%	-0.353%	0.295%

Condition number is in the first disaggregating.
The first row of the error for each appliance is its RMSE.
The second row of the error for each appliance is its relative error for power.

From the result in Table IV, we see that there are two classes of scenarios. First class is no similar appliances, like the first two groups, the VCA reduces the error of appliance A, but does not reduce the total RMSE. The second class is like the third group. Appliance B and C are similar to each other, and the introduction of VCA reduces their RMSE and total RMSE distinctly, but does not reduce the RMSE of appliance A. Due to the cancellation of the relative error, the relative errors for power in Table IV are not following the same pattern as RMSE.

E. Experiment 4: Multiple Appliances and the Strategies

As the increasing number of appliances, the condition number is becoming larger at the same time. In order to get the result with low level error, we suggest combining some appliances into VCA. Due to VCA's impact in the condition number, we believe that different strategies for constructing VCA will lead to different disaggregated results.

The following experiment is to verify this supposition. We use all of the four appliances in this experiment, and calculate the results by using different strategies, then list the errors of those results in Table V, where

Strategy A is to disaggregate mixed current directly, and do not use VCA. Strategy B is to select three appliances to form a VCA, then get the current of laptops, at last disaggregate mixed current for the others directly. Strategy C is to extract the current for one appliance from total in each turn, and we can select the most distinctive one greedy. Strategy D is to extract the current for two appliances from total in each turn, and we follow the greedy selection. Strategy E is to extract the current for every appliance from total and abandon the rest.

TABLE V. THE ERROR OF DIFFERENT STRATEGIES

Strategy	RMSE	Desktop PC	LCD monitor	Laptops	PC+LCD
Strategy A	71.555	63.174	77.792	16.006	100.905
		-0.368%	-0.297%	0.019%	0.439%
Strategy B	58.590	52.291	63.269	15.981	82.088
		-0.324%	-0.438%	0.128%	0.419%
Strategy C	48.726	41.048	54.531	15.981	67.698
		-0.507%	-0.088%	0.128%	0.466%
Strategy D	54.218	50.168	57.334	15.914	75.505
		-0.514%	-0.134%	0.056%	0.508%
Strategy E	49.037	37.740	53.075	15.981	71.567
		-0.266%	0.133%	0.128%	-0.003%

The first row of each appliance is its RMSE.
The second row of each appliance is its relative error for power

From the results in Table V, we see the result of strategy C has the smallest RMSE, and it is a greedy strategy, can be implemented easily. Another greedy strategy E is also much better than other strategies. At the same time, the result in Table V has demonstrated our supposition about using different strategies. On the other hand, our method can obtain more accurate power consumption than the NILM even in a bad condition like Experiment 4.

VI. RELATED WORK

In contrast to the man-machine symbiosis model [13], ternary computing [19] is a new computing paradigm, which emphasizes utilizing resources in the ternary universe. We introduce the input-sensing problem to describe the information coding process of ternary computing.

For electricity load monitoring on device level, there are two methods usually used. One is intrusive load monitoring (ILM). ILM systems are relatively easy to deploy and several products exist [9][17]. Many researchers spend their effort to reduce the hardware complexity of ILM system [4][8][10].

The other method is nonintrusive load monitoring. This concept was first proposed in the 1980's at MIT. George W. Hart wrote a summary for NILM [5], and divide appliance into three models: ON/OFF, FSM and CVL. In 1999 [3], Steven Drenker described this approach as a five-step process.

A research team in the EPRI pays their attention to current harmonics [15], and introduced transient event detection for appliances which are similar in the real and reactive power signature space [11]. In 2010, Scientists in Hong Kong worked out an integrated NILM system [6][7] based on several load signatures and algorithms, and analyzed the results of different methods. From another perspective, some other researchers try to construct the end-user-deployable sensors for household energy consumption [2][16].

There are some methods to deal with the CVL, like [18]. However, they are limited by the assumption of operating states, and their effort focused on transient event detection or switch models of appliances with CVLs. We try to find a new methodology, in which we trace the current instead of the operating state. In our simulation experiments, our method provides positive results to support our theory.

VII. CONCLUSIONS

Accurately acquiring energy consumption information of physical objects influenced by human behavior into the cyber space is a fundamental research problem of green computing. This paper formulates this problem as a general *input-sensing problem* within the ternary computing framework. Four metrics are defined as evaluation criteria, allowing real-world problem instances to be defined and studied, with different constraints on human efforts, sensors needed, error bounds, and computational complexity.

As a concrete instance of the general input-sensing problem, we define the electric *current disaggregating problem* for accurately sensing the electricity consumption behavior of individual appliances in a household, limiting the sensor complexity and the effort complexity. Utilizing Kirchhoff's current law, we simplify the current disaggregating problem into a learning problem called the *approximate phase space learning problem*, amiable to existing computer science tools.

The learning problem is solved by a principal component manifold method that constructs a low-dimension manifold for each appliance as its approximate phase space. This is a polynomial time method. The upper bound of error positively correlates with the condition number of Equation (6).

We conduct experiments using synthetic and real data to validate our approach. The highest relative error for power is 0.514%. The RMSEs for electric waveforms of appliances are less than 16 mA, if the condition number of Equation (6) is bounded by 10. In ill-conditions, we offer a virtual combined appliance method to divide one disaggregating process with high condition number into several disaggregating processes with lower condition numbers.

Theoretical analysis and experimental results indicate that our approach offers a promising research direction for solving the input-sensing problem in disaggregating currents of household appliances.

REFERENCES

[1] M.E. Berges, H.S. Matthews, and L. Soibelman. A system for disaggregating residential electricity consumption by appliance, In Proc. 2010 IEEE International Symposium on Sustainable Systems and Technology, Arlington, USA, 2010.

[2] J.J. Cooley, D. Vickery, A. Avestruz, Z. Clifford, and S.B. Leeb. A retrofit 60 Hz current sensor for power monitoring at the circuit breaker panel. In Proc. 12th IEEE Workshop on Control and Modeling for Power Electronics, Boulder, USA, 2010.

[3] S. Drenker and A. Kader, Nonintrusive monitoring of electric loads. IEEE Computer Applications in Power, 1999, 12(4): 47-51.

[4] D. Guinard, M. Weiss, and V. Trifa. Are you Energy-Efficient? Sense it on the Web! In Adjunct Proc. 7th International Conference on Pervasive, Nara, Japan, 2009.

[5] G. Hart. Nonintrusive appliance load monitoring. Proceedings of the IEEE, 1992, 80(12): 1870-1891.

[6] L. Jian, S. Ng, G. Kendall, and J. Cheng. Load Signature Study—Part I: Basic Concept, Structure, and Methodology, IEEE Transactions on Power Delivery, 2010, 25(2): 551-560.

[7] L. Jian, S.K.K. Ng, G. Kendall, J.W.M. Cheng. Load Signature Study—Part II: Disaggregation Framework, Simulation, and Applications, IEEE Transactions on Power Delivery, 2010, 25(2):561-569.

[8] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. Culler. Design and Implementation of a High-Fidelity AC Metering Network. In Proc. 8th ACM/IEEE International Conference on Information Processing in Sensor Networks, San Francisco, USA, 2009.

[9] Kill a Watt, http://www.p3international.com/products.

[10] Y. Kim, T. Schmid, Z. Charbiwala, and M. Srivastava. ViridiScope: Design and Implementation of a Fine Grained Power Monitoring System for Homes. In Proc. 11th International Conference on Ubiquitous Computing, Orlando, USA, 2009.

[11] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong. Power signature analysis, IEEE Power Energy Magazine, 2003, 1(2): 56–63.

[12] G. Li, editor. Information Science and Technology in China: A Roadmap to 2050. Science Press Beijing and Springer-Verlag Berlin, 2010.

[13] J. Licklider. Man-computer symbiosis. IRE Transaction on Human Factors in Electronics, 1960, HFE-1(1): 4-11.

[14] Friedemann Mattern, Thorsten Staake, and Markus Weiss. ICT for green: how computers can help us to conserve energy. In Proc. 1st International Conference on Energy-Efficient Computing and Networking, Passau, Germany, 2010.

[15] L.K. Norford and S.B. Leeb. Non-Intrusive Electrical Load Monitoring in Commercial Buildings Based on Steady State and Transient Load-Detection Algorithms. Energy and Buildings, 1996, 24(1): 51-64.

[16] S. N. Patel, S. Gupta, and M. S. Reynolds. The design and evaluation of an end-user-deployable, whole house, contactless power consumption sensor. In Proc. 28th ACM Conference on Human Factors in Computing Systems 2010, Atlanta, USA, 2010.

[17] Plogg, http://www.plogginternational.com.

[18] W. Wichakool, A.T. Avestruz, R. W. Cox and S. B. Leeb, Resolving power consumption of variable power electronic loads using nonintrusive monitoring. In Proc. 38th IEEE Power Electronics Specialists Conference, Orlando, USA, 2007.

[19] Z.Xu and G. Li. Computing for the Masses, Communications of the ACM, to appear.

[20] IEEE Working Group on Nonsinusoidal Situations. A Survey of North American Electric Utility Concerns Regarding Nonsinusoidal Waveforms, IEEE Transactions on Power Delivery, 1996, 11(1): 73-78.