

贝叶斯电流分解:利用单个传感器感知家用电器电流

刘晶杰 聂磊

(计算机体系结构国家重点实验室(中国科学院计算技术研究所) 北京 100190)

(中国科学院大学 北京 100049)

(liujingjie09@hotmail.com)

Bayesian Current Disaggregation: Sensing the Current Waveforms of Household Appliances Using One Sensor

Liu Jingjie and Nie Lei

(State Key Laboratory of Computer Architecture (Institute of Computing Technology, Chinese Academy of Sciences), Beijing 100190)

(University of Chinese Academy of Sciences, Beijing 100049)

Abstract An important application of pervasive computing is to obtain the electricity usage information for each appliance in a household using one sensor. The key problem of this application is current disaggregating, which is to estimate the currents of individual appliances from the total current waveform. Existing methods to solve this problem can be classified into two classes: steady-state estimation methods and linear disaggregating methods. Based on the steady-state load assumption, the methods in the first class estimate the current for a running appliance using its steady-state current waveform. These methods can avoid the interference between appliances. But the results of these methods cannot reflect the real-time changes of the total current. The methods in the second class reduce the dimensions of the current waveforms for a specific appliance using model constraints or data constraints, and disaggregate the total current into the linear spaces with low dimensions. The results of these methods can reflect the real-time change of the total current, but similar appliances reduce the accuracy of disaggregating results. From the perspective of the Bayesian statistics, this paper relaxes the key assumptions of the above methods as the prior of position vectors and the prior of noises, and proposes a Bayesian current disaggregating method based these two priors. Using the electricity usage data generated by actual appliances, we conduct several simulation experiments to evaluate our method. The experiment results show that the accuracy of the proposed method is higher than previous methods. Our method not only reflects the real-time change of the total current, but also reduces the effects of similar appliances on the disaggregating results.

Key words household electricity; single-point sensing; current disaggregating; pervasive computing; Bayesian method

摘要 通过单个传感器对家庭中各个电器的用电行为进行感知是普适计算中的一个重要应用,其关键问题是电流分解,即在给定总电流波形情况下计算各个电器的实际电流.此问题现有2类求解方法:稳

收稿日期:2015-04-20;修回日期:2017-11-06

基金项目:中国科学院战略性先导科技专项基金项目(XDA06010403);国家自然科学基金重点项目(61532016)

This work was supported by the Strategic Priority Research Program of Chinese Academy of Science (XDA06010403) and the Key Program of the National Natural Science Foundation of China (61532016).

态估计方法和线性分解方法.前一类方法基于电器稳态耗电假设,使用稳态波形估计电器的工作电流.虽然该类方法能避免电器间的相互干扰,但是其结果不能反映总电流的实时变化.后一类方法通过模型约束或数据约束对电流波形进行线性降维,之后将总电流分解到各个低维线性空间中.虽然其分解结果能够反映总电流的实时变化,但是相似电器会降低分解结果的精度.从贝叶斯统计的角度将上述方法的关键假设松弛为位置向量先验分布与噪音先验分布,并提出了基于这2个分布的贝叶斯电流分解方法.利用真实用电数据,构造了多组仿真实验对此方法进行评测.实验结果表明:提出方法分解精度高于原有2类方法,其感知结果既能够反映总电流及各个电器电流的实时变化,又能够降低相似电器对分解结果的干扰.

关键词 家庭用电;单点感知;电流分解;普适计算;贝叶斯方法

中图分类号 TP391

普适计算中一个重要的应用场景是通过无处不在的传感设备感知用户在家庭中的用电行为^[1].为了减少感知系统对用户正常用电的影响,仅利用单个固定传感器获取到家庭中各个电器的用电行为是普适计算的一个研究热点^[2-12].在基于总电流的单点感知系统中^[4-10],电流分解是核心问题.

电流分解问题是给定多个电器的总电流波形和这些电器的运行状态,估计各个电器各自的电流.对于电流分解问题,目前有2类方法:稳态估计方法与线性分解方法.

1) 稳态估计方法.此类方法基于 Hart^[4]提出的稳态耗电假设,认为电器在特定的运行状态下,其消耗的功率是在一个常数值附近变化.对于每个确定工作状态的电器,可以通过一个稳定电流波形对电器实际消耗的电流进行估计^[5-6].对于非线性电器,也可以用电流波形在频域上的均值估计实际的电流波形^[7].此类方法的优点是不需要复杂的计算,其分解结果的误差由电器本身在此运行状态下的电流变化幅度决定.对于同时工作的多个电器,稳态方法独立地对这些电器的电流进行估计,因此可以避免电器之间的相互影响.

2) 线性分解方法.此类方法利用物理模型约束或者是数据约束对各个电器的频域中的电流向量进行线性降维,通过线性回归将总电流分解到各个电器的低维表示中.物理模型约束和数据约束可以通过不同的方式得到,例如通过物理知识构造谐波电流需要满足的线性约束^[8]、通过对电流波形的形态分析建立线性约束^[9]、通过聚类分析将电器电流向量的变化范围缩减至一维^[10]、通过主成分分析方法将电器的电流限制在主成分构成的低维空间^[11].此类方法的优点是使用的模型能够有效地刻画电流

的变化情况,使其分解结果能够反映总电流的实时变化.

这2类方法都能够在一定条件下解决电流分解问题,但都各自存在一些问题.稳态估计方法中给出的估计值仅依赖于电器的运行状态,对于在给定运行状态下电流波动较大的电器,其分解结果不能较为准确地反映电器电流的实时变化情况.线性分解方法一般不能无损降维,降维造成的误差会对分解结果的精度造成影响;更重要的是,线性分解方法不能处理相似电器^[11],其适用范围受到限制.

为了将2种方法的优点结合起来,本文从贝叶斯统计的视角分析上述2类方法对电器行为的假设,建立了电流在低维空间中的位置向量先验分布和噪音先验分布.利用贝叶斯公式,本文将2个分布结合起来,并给出一种贝叶斯电流分解方法.此方法同时具备稳态估计方法和线性分解方法2类方法的优点,能够取得较良好的分解结果,并在所有情况下,完成总电流的分解.

1 电流分解问题和2类基本方法

1.1 电流分解问题定义

家庭用电感知应用中2类场景如图1所示.图1(a)中给出了使用多点传感设备完成电器用电行为感知的示意图,对于每个需要感知的电器,使用单独的传感器进行感知^[12].图1(b)中给出了单点传感器完成电器用电行为感知的示意图,传感器仅在电路入口处接入家庭电路,只能得到整个家庭的总电流波形和电压波形,需要通过一个计算过程得到每个电器的用电行为^[3].对于每个独立工作的电器,其在工作产生的电流波形均是随时间变化的波形,而其电压波形与传感器得到的电压波形相同,因此,计算其用电行为的关键是确定其电流波形.

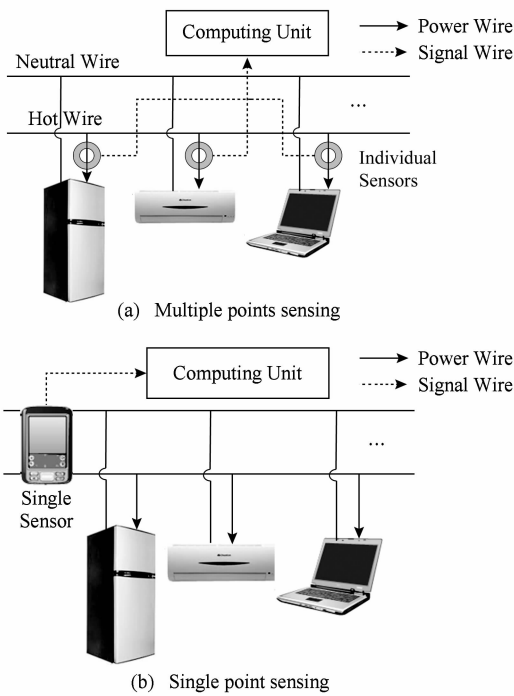


Fig. 1 Two typical scenarios for household electricity usage sensing
图 1 家庭用电感知的 2 类典型场景

根据基尔霍夫节点电流定律:在任何时刻,电流传感器得到的多个电器的总电流的瞬时值是等于各个电器的电流瞬时值之和. 电流分解问题就是要根据此关系以及电器的先验知识计算各个电器的实际通过的电流波形. 为了完成电流分解问题的形式化定义,首先对问题中的各个物理量进行形式化表述.

在家庭用电场景中,电压是 220V,频率是 50 Hz 的正弦交流电,而通过电器的电流是对此电压的响应. 因此,在多数情况下,电流电压波形中的主要频域成分是 50 Hz 的整数倍频率. 通过傅里叶级数展开能够将任意电流电压波形 $I(t)$ 写成:

$$I(t) = \sum_{j=1}^{n/2} a_j \cos(2\pi fjt) + b_j \sin(2\pi fjt). \quad (1)$$

其中, $I(t)$ 是时刻 t 的波形, f 是交流电频率,在中国是 50 Hz,而 n 是每个周期中采样点的数目. 式(1)中的系数 a_j 和 b_j 是根据离散傅里叶变换得到的傅里叶系数的实部和虚部. 根据式(1)可以知道对于任何一个频率小于 $nf/2$ 的电流电压波形,都可以通过一个由系数 a_j 和 b_j 构成的向量唯一表示. 对于给定的电流波形 $I(t)$,我们可以通过一个 n 维的列向量 i 进行表示,我们将 i 称作一个电流向量:

$$i = [i_j]_{n \times 1} = [a_1, a_2, \dots, a_{n/2}, b_1, b_2, \dots, b_{n/2}]^T. \quad (2)$$

在图 1 所示的电路中,由基尔霍夫节点电流定

律可知,对任意时刻都有总电流等于各个子电流之和. 所以有:

$$I_a(t) = \sum_{k=1}^m I_k(t), \quad (3)$$

其中, m 是总电流对应的电器数目, I_a 是总电流对应的波形, I_k 是第 k 个电器的电流对应的波形. 将式(1)和式(2)代入式(3)中可知总向量中所有元素都等于各个电器电流向量中对应元素之和. 用矩阵形式进行表述,有:

$$i_a = \underbrace{[E_n \cdots E_n]}_{m \text{ 个}} \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_m \end{bmatrix}. \quad (4)$$

其中, E_n 是一个 n 阶的单位矩阵, i_a 是总电流向量, i_k 是第 k 个电器的电流向量.

电流分解问题就是要在给定的总电流向量 i_a 以及各个电器的运行状态时,利用式(4)给出的线性关系,估计各个电器的电流向量 i_k ,并使得计算得到的 m 个电流向量与真实的电流向量之间的误差尽可能小.

若每个电器的波形向量的取值范围都是整个 n 维实数空间 $\mathbb{R}^n$,那么对应式(4)中,给定一个总电流 i_a ,存在无数多个解. 无法通过对式(4)的求解直接计算各个电器的电流向量.

1.2 稳态估计方法及其特点

因为式(4)给出的信息不足以完成对电器电流的估计,要求解电流分解问题首先需要对电器的耗电行为建立约束. 一种广泛使用的假设是在文献[4]中提出的稳态耗电假设. 当电器处于特定工作状态时,其上消耗的功率处在一个常数值附近. 从该假设出发,可以导出特定工作状态下的电器电流在每个周期中均是相同的波形,能够使用平均电流波形估计实际电流波形. 例如在文献[5-6]中就直接将电器在各个工作状态下的波形视作定值. 对电流分解问题的求解过程就是直接用各个电器的电流向量的均值估计当前的电流向量 i_k . 我们将此类分解方法称为稳态估计方法. 对第 k 个电器,此方法给出电流向量

$$i_k = \mu_k. \quad (5)$$

其中, μ_k 是第 k 个电器的电流向量的均值. 由此稳态估计方法的关键是获取每个电器的电流均值 μ_k ,通过预先的训练阶段得到的训练集合能够得到此向量.

显然对于一次分解而言,式(5)得到的估计电流

向量与真实电流向量可能存在较大差异,同时不保证多个电器的分解结果满足式(4).但是对于多次电流分解,通过均值估计电流向量能够保证平均误差逼近于0,而标准误差逼近于电器在当前状态下电流向量的标准差.

使用稳态估计方法求解电流分解问题的优点在于各个电器的分解结果的计算过程相互独立,多次分解的标准误差具有一个确定的期望,等于电器电流向量的标准差.无论家庭中同时工作的电器数目和种类如何变化,对于特定的一个电器,稳态估计方法的计算结果始终是确定的,不会受其他电器的影响.而此类方法的缺点是在计算电器电流时完全没有利用总电流的信息,仅依赖于电器的稳态耗电假设.对于电流向量的标准差较大的电器,此方法的计算结果的误差随之增加.

1.3 基于线性降维的电流分解方法及其特点

稳态估计方法基于一种简单的假设,对电流分解问题进行求解,在其求解过程中,总电流向量所包含的信息没有任何使用.而考虑利用此信息则需要对电器的电流向量建立另外一套约束.文献[13]中观察变速电机类负载的谐波电流谱包络特征,发现不同阶谐波的谱包络线之间存在线性关系.而文献[8]针对变速驱动马达的电源结构的研究,在其基频电流和谐波成分之间建立了一个近似线性的函数关系,在文献[10]中假设家庭中存在多个电器类型,同类型电器的电流向量之和分布在整个谐波空间中的一条直线上,通过聚类方法得到直线方程.

上面3个工作都试图通过简单的数学模型对电器的电流变化规律进行抽象,其中的共同点是假定了不同频率的谐波之间存在线性关系.假设这种相关性是加在第 k 个电器的电流向量 i_k 上的一组线性约束.根据一般的线性约束的性质,能够将 i_k 限制在一个维度为 r 的低维线性空间中.令 i_k 的均值 μ_k 作为低维空间的原点,此线性空间中位置向量为 c_k 的一个点对应于电器的一个电流向量 i_k ,两者之间的关系可以表示为

$$i_k = \Phi_k c_k + \mu_k. \tag{6}$$

其中, Φ_k 是由 r 个基向量组成的 $n \times r$ 的矩阵, c_k 则是在低维空间中的位置向量.式(6)是对第 k 个电器的电流向量的线性降维方程,由电器上的物理模型约束或者数据约束构造得到.在文献[8-9]中,降维方程是通过电器物理模型和波形特征进行分析得到此降维方程.而在文献[10]中假定基矩阵是一个一维列向量,通过理论模型或者聚类方法可以得到

此方程.而在我们之前一个工作中^[11],通过主成分分析(PCA)计算得到此方程.

将式(6)代入到式(4)中,能够改写基尔霍夫节点电流定律,得到:

$$i_a - \mu = [\Phi_1 \Phi_2 \cdots \Phi_m] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \Phi c. \tag{7}$$

其中,矩阵 Φ 是各个电器的基矩阵 Φ_k 拼接起来得到的一个系数矩阵,向量 μ 是各个电器的均值之和,而向量 c 由所有电器在低维空间的位置向量拼接而成.通过求解式(7)中表示的线性方程,能够得到各个电器在低维空间中的位置向量,进而得到电流分解问题的解.我们将此类分解方法称为线性分解方法.文献[11]给出了普通最小二乘法求解式(7)的必要条件,即矩阵 Φ 列满秩.当条件满足时,根据普通最小二乘法(OLS)计算能够得到所有电器的位置向量:

$$c = (\Phi^T \Phi)^{-1} \Phi^T (i_a - \mu), \tag{8}$$

其中,向量 c 按照其电器划分后能够得到各个电器对应的位置向量 c_k ,分别代入式(6),能够得到各个电器的电流向量的估计值 c_k .文献[11]中分析了此结果中的误差情况,误差由2个要素决定:矩阵 Φ 的状态数和总电流中的噪音项.而此矩阵的状态数由待分解电器的组合决定,而总电流中的噪音项不但包括测量误差还包括在各个电器降维时引入的噪音.

线性分解方法的优点在于有效利用了总电流向量中的信息以及电器的电流向量变化的特点,在大部分情况下能够取得优于稳态估计方法的分解效果.其分解结果能够有效的反映总电流的变化.而此方法的缺点也同样明显,即线性分解方法不是在任意情况下可用的.当矩阵 Φ 不满足列满秩条件时,式(8)中的项 $(\Phi^T \Phi)^{-1}$ 不存在,原方程存在无数个解.同时电器降维会引入额外噪音,在某些情况下,其分解结果中的误差可能大于稳态估计方法.

2 贝叶斯电流分解方法及其实现

第1节分别讨论了电流分解问题和2类计算方法的形式化表示以及2种方法的优缺点.从中我们发现2种方法具有一定的互补性:线性分解方法的精度容易受电器组合影响,相似电器会极大降低其分解精度.而稳态估计方法能够避免电器之间的相互影响,每个电器的估计结果中的误差仅由电器本

身决定;稳态估计方法无法刻画电器的电流变化信息,而线性分解方法的基本假设便是电器的电流向量在线性约束下发生变化.

我们希望能够有效地将 2 种方法的优点结合到一种统一的分解方法中,但是在上述 2 种方法中使用的基础假设相互冲突:稳态耗电假设要求电器的电流向量等于稳定值,而线性降维的基础是电流在线性空间中不断变化.为了能够在统一框架中调和二者的冲突,我们从贝叶斯统计的观点进行讨论.

2.1 贝叶斯电流分解方法

从统计的角度重新考虑电流分解问题,可以将整个过程看作一个随机过程.每个电器相互独立地在各自的低维空间中随机产生电流向量,所有电器的电流向量可以用一个位置向量 $\mathbf{c}$ 表示.用于电流分解的总电流 $\mathbf{i}_a$ 不但包括了根据式(7)得到的总电流,同时还包括一个随机噪音 ϵ .针对这个随机过程,由贝叶斯定律可以得到:

$$P(\mathbf{c} \mid \mathbf{i}_a) = \frac{P(\mathbf{i}_a \mid \mathbf{c})P(\mathbf{c})}{\sum_{\mathbf{c}'} P(\mathbf{i}_a \mid \mathbf{c}')P(\mathbf{c}')}, \tag{9}$$

其中, $P(\mathbf{c})$ 是电器产生位置向量 $\mathbf{c}$ 的先验概率分布,此概率在电流分解开始前就已经确定. $P(\mathbf{i}_a \mid \mathbf{c})$ 是一个条件概率分布,它是给定了所有电器的位置向量时,测量得到的总电流向量的概率分布. $P(\mathbf{c} \mid \mathbf{i}_a)$ 是后验概率分布,表示的是给定一个总电流向量时,各个电器的位置向量的概率分布.

结合式(9)分别讨论稳态估计方法和线性分解方法中的关键假设的统计意义.对于稳态估计方法,其使用每个电器的电流向量的均值作为分解结果.根据稳态耗电假设,可知在确定状态下,电器的电流向量在均值附近浮动.我们可以将此假设描述为以均值为中心的正态分布.那么对于所有电器的位置向量的先验分布可以写作:

$$P(\mathbf{c}) = \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Psi}), \tag{10}$$

其中, $\boldsymbol{\mu}_c$ 是所有电器的位置向量的均值.根据式(6)的描述可知,构造电器低维空间时,其原点落在原空间中的电流均值上,每个电器在低维空间中的位置向量的均值为 0,因此分布 $P(\mathbf{c})$ 是一个 0 均值分布.矩阵 $\boldsymbol{\Psi}$ 是各个电器的低维坐标的协方差矩阵.各个电器独立运行,任何 2 个不同电器的低维坐标之间是相互独立的.而对于单个电器,若降维方法是通过 PCA 方法进行,得到的各个主成分之间统计无关,且每个主成分的方差不为 0.因此各个电器的低维坐标的协方差矩阵 $\boldsymbol{\Psi}$ 是一个对角阵且对角线元素均大于 0.

对于稳态估计方法而言,其分解结果已经在式(5)中给出,存在于原向量空间中.若将其代入式(6)中可知此分解方法估计电器的低维坐标恒为 0,即: $\mathbf{c} = 0$.此结果是坐标先验 $P(\mathbf{c})$ 的协方差矩阵 $\boldsymbol{\Psi}$ 中所有元素为 0 时,后验分布 $P(\mathbf{c} \mid \mathbf{i}_a)$ 中唯一的非 0 项的位置向量.

在对线性分解方法进行分析前,向式(7)中引入一个总电流向量上的噪音项 ϵ .因为此噪音项是多个 0 均值的随机变量之和,根据中心极限法则,不妨假定此噪音项服从 0 均值多维正态分布 $\mathcal{N}(0, \boldsymbol{\Theta})$.通过引入噪音的先验分布,能够将式(7)所示的线性方差缩放成为 $P(\mathbf{c} \mid \mathbf{i}_a)$ 表示的条件分布.此时,总电流向量 $\mathbf{i}_a$ 和各个电器低维坐标 $\mathbf{c}$ 之间的一个随机关系: $\mathbf{i}_a - \boldsymbol{\mu} - \boldsymbol{\Phi}\mathbf{c} = \mathcal{N}(0, \boldsymbol{\Theta})$.将此关系化简为 $P(\mathbf{i}_a \mid \mathbf{c})$ 的形式,有:

$$P(\mathbf{i}_a \mid \mathbf{c}) = \mathcal{N}(\boldsymbol{\mu} + \boldsymbol{\Phi}\mathbf{c}, \boldsymbol{\Theta}). \tag{11}$$

其中,矩阵 $\boldsymbol{\Theta}$ 是噪音先验分布的协方差矩阵,由多个相互独立的噪音之和.其中测量引入的噪音通常可以视作一个时域上白噪音,因此在电流向量这样一个频域表示下各个维度方差相等且独立,而每个电器因为降维引入的噪音均相互独立且数目极多.为了简化后续计算,可以近似假设此噪音的协方差矩阵 $\boldsymbol{\Theta}$ 是一个对角矩阵且对角线元素均大于 0.

对于线性分解方法而言,其分解结果在式(8)中给出,代入式(11)中可知,此结果是令分布 $P(\mathbf{i}_a \mid \mathbf{c})$ 最大的电器电流向量.而矩阵 $\boldsymbol{\Phi}$ 不满足列满秩时,对应的在分布 $P(\mathbf{i}_a \mid \mathbf{c})$ 中,使得概率最大的电流向量 $\mathbf{c}$ 有无数个,无法确定唯一解.

至此,我们将稳态估计方法和线性分解方法的关键假设松弛为式(10)和式(11)所示的 2 个正态分布.利用贝叶斯公式可以将这 2 个分布融入到统一的计算框架中.将二者代入式(9)中,能够得到后验概率的分布.因为式(10)和式(11)所述的概率分布都是多维正态分布,所以后验分布存在解析解,而且也是一个多维正态分布,有:

$$P(\mathbf{c} \mid \mathbf{i}_a) = \mathcal{N}(\boldsymbol{\gamma}, \boldsymbol{\Sigma}). \tag{12}$$

其中,后验分布的均值 $\boldsymbol{\gamma}$ 和方差 $\boldsymbol{\Sigma}$ 分别为

$$\boldsymbol{\Sigma} = (\boldsymbol{\Psi}^{-1} + \boldsymbol{\Phi}^T \boldsymbol{\Theta}^{-1} \boldsymbol{\Phi})^{-1}, \tag{13}$$

$$\boldsymbol{\gamma} = \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \boldsymbol{\Theta}^{-1} (\mathbf{i}_a - \boldsymbol{\mu}). \tag{14}$$

通过给出均值和方差,能够唯一地确定一个后验分布.而基于此后验分布,我们可以在其中找到后验概率最大的位置向量作为分解的结果.因为这是一个正态分布,所以最大后验在均值 $\boldsymbol{\gamma}$ 处取得.最大后验估计得到的位置向量代入式(6)能够得到各个电器在原始向量空间中的电流向量.

此分解方法的关键在于利用训练集构造低维坐标的先验分布以及总电流中噪音项的先验分布.若使用主成分分析进行降维,那么前一个分布可以利用降维后的训练集数据得到.而后一个分布由多个分布叠加得到,包括测量噪音的分布以及每个电器降维时舍弃的非主成分的噪音.测量噪音由测量方法确定,而每个电器的降维噪音可以通过用训练集数据减去降维后能够保留的数据得到.我们将此分解方法称为贝叶斯电流分解(Bayesian current disaggregating, BCD)方法.

对于BCD方法,与线性分解方法类似,存在分解结果的必要条件是式(13)和式(14)计算过程中的求逆计算可行,其中关键是要要求协方差矩阵 Ψ 和 Θ 以及矩阵之和 $\Psi^{-1} + \Phi^T \Theta^{-1} \Phi$ 可逆.根据基本的线性代数知识可知,矩阵 Ψ 和 Θ 是一个对角矩阵且对角线元素均大于0,同时 $\Phi^T \Theta^{-1} \Phi$ 为半正定矩阵,显然3个矩阵均可逆.对于BCD方法,在几乎所有情况下都能够对总电流进行分解,不需要考虑类似文献[11]中给出的满秩条件.

将电流分解问题以及3类方法重新整理可以得到图2,图2中将电流分解问题描述成为其在线性降维后的等价表示,并给出3种方法分别得到的位置向量的解析表达式.从图2中可以看到,BCD方法的解在特定条件下会退化成另外2个方法的结果之一.当噪音先验的协方差矩阵 Θ 中所有项均趋于0时,BCD方法会退化成线性分解方法.当坐标先验的协方差矩阵 Ψ 中所有项均趋于0时,BCD方法会退化成稳态估计方法.而在一般情况下,BCD方法同时考虑了噪音先验和坐标先验,此方法得到的分解结果能够结合之前2种方法的优点.对于贝叶斯电流分解方法的分解效果的验证将放在本文第3节中进行.

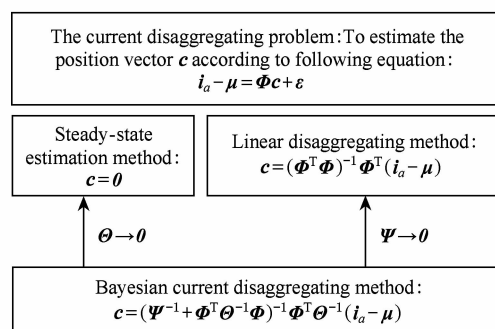


Fig. 2 Current disaggregating problem and the relation between three computational methods

图2 电流分解问题及3种计算方法之间的联系

2.2 算法实现

BCD方法对总电流向量进行分解的核心公式是式(13)和式(14).而为了计算这2个公式,需要得到2个关键分布的值以及各个电器的低维空间基的矩阵表示.而在2.1节中,我们假定了2个分布的协方差矩阵均为对角矩阵,所以计算中可以用对角线元素的向量进行表示.为此,BCD算法分为2个阶段:1)训练算法.用于从电器的训练集中得到电器的低维空间表示以及其先验分布的信息;2)分解算法.在给定总电流和电器集合的情况下,利用训练算法给出的信息,对总电流向量进行分解.

在给定第 k 个电器的训练算法中,其输入是此电器的电流波形 I 和同步采集到的电压波形 V ,而输出则包括电器低维空间的基矩阵 Φ_k ,电流向量均值 μ_k ,电器的电流向量在各个低维基上的方差向量 Ψ_k ,以及此电器降维引入的噪音的协方差矩阵 Θ_k .具体算法如下:

算法1. 训练算法.

输入: I, V ;

输出: $\Phi_k, \mu_k, \Psi_k, \Theta_k$.

① 对输入电流电压波形进行预处理.根据电压波形的正向过零点将整个波形划分到各个周期中.对于每个周期的电流电压波形分别进行快速傅里叶变换,得到其对应的电流向量和电压向量.利用电压向量中50 Hz成分的相位信息对电流向量的相位进行修正.最后将得到的多个电流向量 i_a 组成一个电流矩阵 T ,其中每一个列向量对应一个周期的电流波形.

② 主成分降维.首先对电流矩阵 T 计算电器的电流向量的均值 μ_k ,对电流矩阵 T 计算阈值为0.95的主成分,这些主成分的构成基矩阵 Φ_k .

③ 计算电器在低维空间中的先验分布.利用步骤②中得到的矩阵 Φ_k ,计算电流矩阵在低维空间中的位置向量组成的矩阵 $C = \Phi_k^T (T - \mu_k)$.对于此矩阵,计算其上的每一行的方差能够得到方差向量 Ψ_k .

④ 计算因为电器降维引入的噪音.利用步骤②③中得到的基矩阵 Φ_k 和位置矩阵 C ,可以得到降维引入的噪音矩阵 $R = T - \mu_k - \Phi_k C$.对于此矩阵中的列向量计算协方差矩阵能够得到矩阵 Θ_k .

在给定 m 个电器同时工作的场景下,分解算法的输入是 m 个电器运行时产生的总电流波形 I 和同步采集到的电压波形 V , m 个电器各自的基矩阵

Φ_k 、电流均值 μ_k 、位置向量的方差向量 Ψ_k 以及降维噪声协方差矩阵 Θ_k 以及测量噪声的协方差矩阵 e . 而其输出是各个电器的电流波形 I_k . 具体算法如算法 2:

算法 2. 分解算法.

输入: $I, V, e, \Phi_k, \mu_k, \Psi_k, \Theta_k$, 对每个 $k \in [1, m]$;
输出: I_k .

① 对输入电流电压波形进行预处理. 预处理方法与算法 1 中步骤①相同, 得到电流矩阵 T .

② 构建分解场景中的主要参数. 根据 m 个电器的先验知识, 可以得到总电流的均值 $\mu = \sum_{k=1}^m \mu_k$; 分解矩阵 $\Phi = [\Phi_1 \Phi_2 \cdots \Phi_m]$; 全部电器在低维的位置向量的协方差矩阵 $\Psi = diag([\Psi_1 \Psi_2 \cdots \Psi_m])$; 噪音的协方差矩阵 $\Theta = e + \sum_{k=1}^m \Theta_k$.

③ 计算所有电器的位置向量 c . 将步骤②中得到的信息代入式(13), 计算得到后验分布的协方差矩阵, 有: $\Sigma = (\Psi^{-1} + \Phi^T \Theta^{-1} \Phi)^{-1}$. 再利用式(14), 将第 1 步得到的电流矩阵 T 代入, 能够得到需要计算的位置向量 $c = \Sigma \Phi^T \Theta^{-1} (i_a - \mu)$.

④ 计算电器的电流波形. 将步骤③中得到的位置向量 c 按照其对应的电器进行分解, 得到各个电器的位置向量 c_k , 代入式(6)能够得到电器的电流向量 i_k . 对于每一个电流向量, 执行与步骤①预处理过

程相反的计算过程, 能够得到对应周期中的波形, 最后拼接起来可以得到各个电器的电流波形 I_k .

3 实验结果与分析

本节通过仿真实验对贝叶斯电流分解算法进行验证. 因为电流分解问题对数据精度要求较高, 本文并没有构建多个高精度传感器对同时工作的多个电器进行采样. 而是分别对不同电器的电流波形进行采样, 利用采样得到的真实数据构造多个仿真场景对我们的方法进行评测. 在这些场景中, 我们根据基尔霍夫节点电流定律将不同电器的真实电流叠加起来作为理论上的总电流. 得到的总电流将作为分解方法的输入, 而分解结果将和原始电流进行比较, 从而对分解效果进行评估. 因为总电流是由真实电流直接叠加而成, 所以算法 2 中输入的测量噪音 e 的取值为 0.

3.1 数据采集环境与被测电器

实验中使用的数据均是来自实验室环境下采集的由电器产生真实用电数据. 为了获取到精确的电流数据, 我们使用一个创研高科公司生产的 ETRC030 高精度钳形漏电流传感器作为电流传感器安装在电路的主电源断路器附近的火线上(图 3 中左上图). 而电压的测量是使用择明电子的 ZMPT01B 电压互感器作为电压传感器(图 3 中右下图). 而数据采集



Fig. 3 The data acquisition devices and the scenario
图 3 电路信息的采集设备与采集场景

时,我们使用阿尔泰科技的12b数据采集器USB2831完成数据采集工作(图3中右上图),通过USB接口将数据存入电脑.实验时采集数据的频率为6400 Hz,保证了在中国市电50 Hz的环境下,每个周期有128个采样数据.在本实验中,我们选取了5个不同的电器,这5个电器的基本信息在表1中给出.其中平均功率信息能够显示出电器之间电流强度的差异.主成分的数目显示了电器上受到的线性约束的强度.电流波形标准差给出的是各个电器电流波形与其均值波形的偏离程度,同时也能够估计通过均值估计得到的分解误差.图4中给出了20 s中这些电器50 Hz电流分量的变化情况.

从图4中,我们可以看到电器A1,A3,A5的电流峰值在多数时间中是处在一个稳定值附近的,在少数时间会有较大变化;A2的电流峰值基本一直

在162 mA附近小幅度震荡;而A4的电流峰值在图4中始终有较大的变化幅度.A5有最大的标准差是因为其电流的绝对强度最高.而标准差和均值的比例最高的是A4,这也与图4中显示出的趋势一致.

Table1 The List of Five Appliances
表 1 实验中使用的 5 个电器的基本信息

Appliance ID	Appliance Name	Average Power /W	Number of Principal Components*	Standard Deviation /mA
A1	Desktop	64.36	3	23.98
A2	LCD Monitor	25.24	7	18.45
A3	Laptop A	20.20	6	15.02
A4	Laptop B	38.39	12	101.48
A5	Fridge	560.50	2	603.77

* The dimension of position vectors is 128.

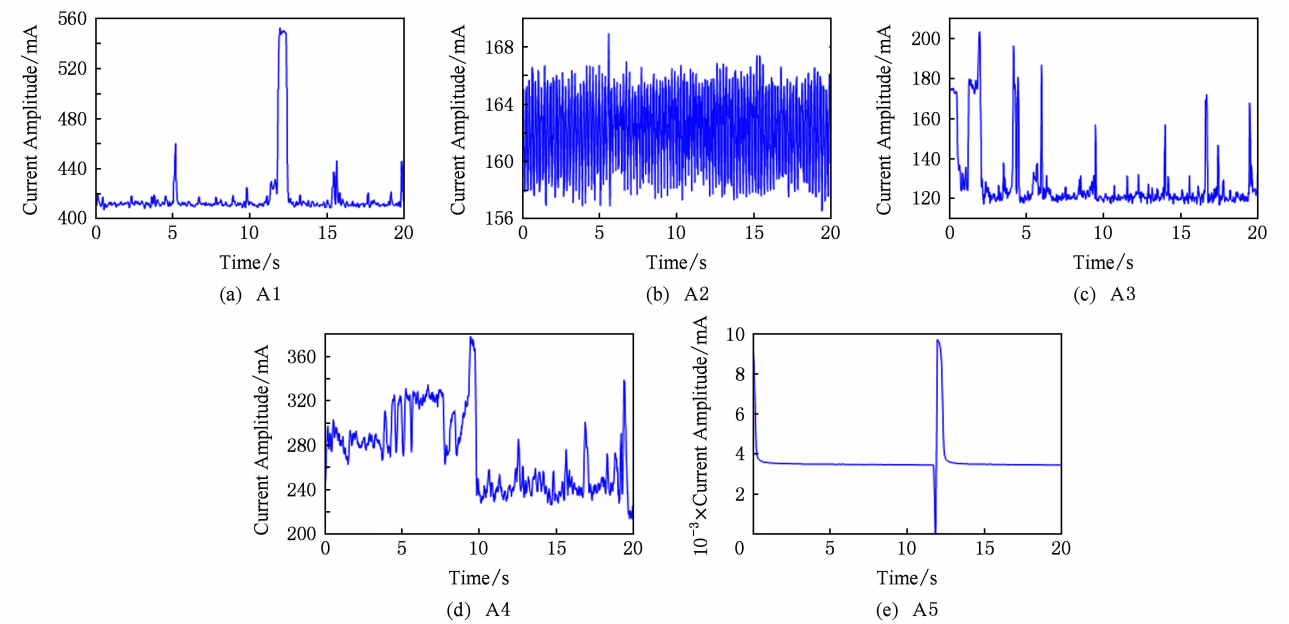


Fig. 4 The curve of current amplitude versus time for five appliances
图 4 5 个被测电器在 20 s 内电流峰值(50 Hz)随时间变化情况

后续的实验中,我们将整个分解得到的波形与真实波形比较得到的标准误差作为分解方法的评价指标.在每个场景下,待分解的波形由4000个周期的电流波形连接而成.作为参考,稳态估计方法、线性分解方法的分解误差将同BCD方法的分解误差一起给出.所以对于每一个分解场景,我们均给出了会给出3个不同的误差,分别来自稳态估计方法(Mean)、线性分解方法(Linear)以及贝叶斯电流分解方法(BCD).

3.2 2 个电器电流分解场景

首先对2个电器分解的仿真场景中的分解结果

进行分析,表2中列出了5个电器两两组合的10种不同的分解场景下的分解结果.

表2中给出了2个电器场景中3种不同算法的分解结果.在给出的10个场景中,本文提出的BCD方法的分解结果都显著优于另外2种算法.以A3&5的情况为例进行讨论,A5电冰箱固有的电流波形标准差是603 mA,而A3只有10 mA.在2个电器分解时,A5的波形上的剧烈波动,容易对A3的分解结果造成影响.线性法在此场景中得到的A3的误差显著大于稳态法.而本文给出的BCD方法中,依靠位置向量先验对分解结果的约束,A3受到

Table 2 The Disaggregating Results for Two Appliances

表 2 2 个电器场景中电流分解结果

mA

Appliance ID	RMSE of First Appliance			RMSE of Second Appliance		
	Mean	Linear	BCD	Mean	Linear	BCD
A1&2	24.1	5.65	5.2	18.65	5.29	4.82
A1&3	24.78	4.77	4.68	14.91	4.11	3.97
A1&4	24.44	14.44	9.5	101.87	26.06	23.85
A1&5	25.1	17.29	9.4	602.33	35.45	34.06
A2&3	18.83	4.32	4.1	14.95	3.87	3.58
A2&4	18.48	19.79	11.85	102.26	29.14	24.8
A2&5	18.59	13.6	7.6	613.62	34.64	34.15
A3&4	14.98	15.33	8.08	100.88	26.31	23.29
A3&5	14.58	42.01	11.86	603.07	50.43	34.49
A4&5	102.29	45.15	30.01	601.35	48.17	37.49

A5 的影响被限制到了一个可以接受的程度,BCD 方法得到 A3 和 A5 的误差均小于前 2 种分解方法.

下面我们讨论 2 个电器分解场景中的一个极端情况:我们选取了 2 个主成分数目超过总维度 50% 的电器,主成分数目分别为 87 和 88. 这 2 个电器受到线性约束较弱,降维后仍然不满足文献[11]提出的分解条件,线性法无法运行. 在此情况下,通过稳态法和 BCD 方法能够给出分解结果. 其结果在表 3 中列出:

Table 3 The Disaggregating Results for a Special Case

表 3 不满足分解条件时 2 种方法的电流分解结果 mA

RMSE of First Appliance		RMSE of Second Appliance	
Mean	BCD	Mean	BCD
29.70	20.22	29.89	20.23

从表 3 中可以看到,在这一特场景下 BCD 方法的分解误差比稳态法小了 30%. 我们选取了此场景下离散的 400 个周期,给出其分解结果中的 50 Hz 的波形峰值构成的曲线,如图 5 所示:

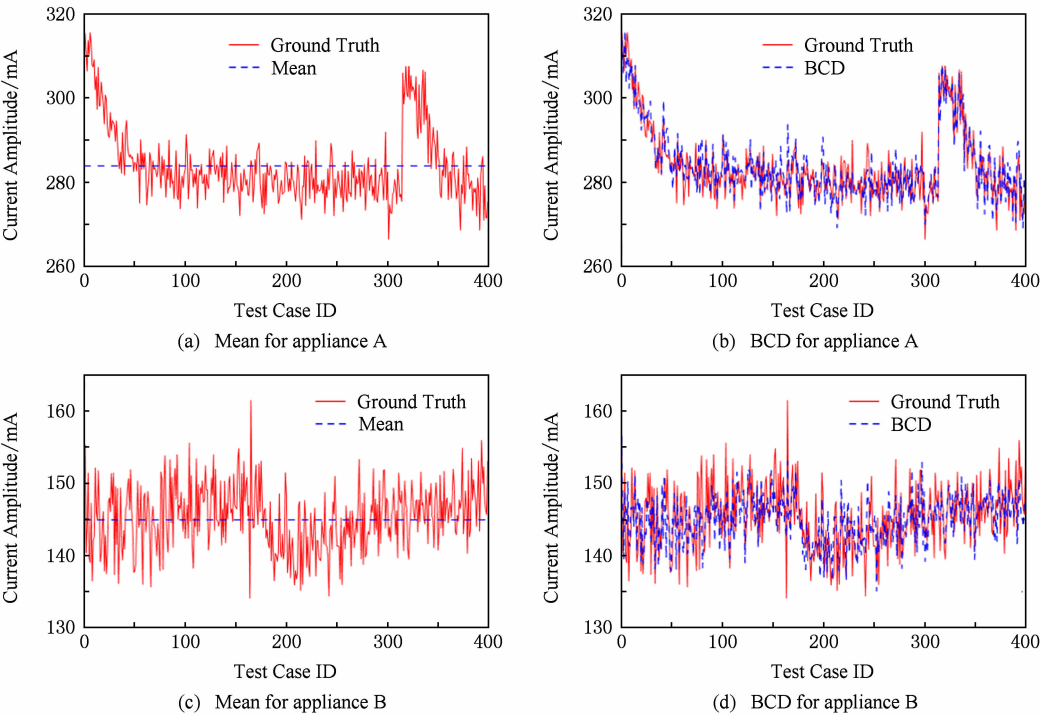


Fig. 5 The result comparison for a special case

图 5 特殊情况下 2 种方法的结果比较

从图 5 中可以看到,稳态法和 BCD 方法给出的结果都在真实值附近.而其中 BCD 方法的结果能够反映电器波形的实时变化,而稳态法仅是给出均值的估计.

3.3 3 个电器电流分解场景

下面在 3 个电器同时工作的仿真场景中,对 3 种方法进行测试,我们将分解结果列在表 4 中.对于每一种方法,我们将 3 个电器的标准误差的方均根值作为总体误差放在表 4 中.

表 4 给出的实验结果与 2 个电器的分解结果类似,在 10 种不同电器组成的场景下,本文提出的 BCD 方法给出的分解结果的标准误差是最低的.

Table 4 The Disaggregating Results for Three Appliances

表 4 3 个电器场景中电流分解结果 mA

Appliance ID	RMSE of All Appliances		
	Mean	Linear	BCD
A1&2&3	19.61	5.73	5.07
A1&2&4	61.33	23.52	17.60
A1&2&5	362.74	23.03	20.11
A1&3&4	61.29	22.06	16.00
A1&3&5	337.23	36.20	21.51
A1&4&5	364.44	39.27	28.05
A2&3&4	59.81	23.53	17.00
A2&3&5	354.13	41.36	22.12
A2&4&5	328.96	44.28	29.32
A3&4&5	376.82	49.62	28.63

3.4 更多电器分解场景

进一步展示本文给出的 BCD 方法与原有 2 类方法在更多电器的分解场景中的效果,表 5 中给出了 4 个电器和 5 个电器分解场景中 3 类方法的分解结果.与 3 个电器分解场景类似,我们在表 5 中列出了全部电器的分解结果的标准误差的方均根值.在这些场景中,本文给出的 BCD 方法的分解结果仍然显著地优于原有 2 类方法.

Table 5 The Disaggregating Results for More Appliances

表 5 多电器场景中电流分解结果 mA

Appliance ID	RMSE of All Appliances		
	Mean	Linear	BCD
A1&2&3&4	53.01	23.51	16.11
A1&2&3&5	318.75	36.83	20.12
A1&2&4&5	304.14	42.20	26.46
A1&3&4&5	302.32	42.67	24.80
A2&3&4&5	309.21	48.30	26.27
A1&2&3&4&5	289.16	47.38	24.83

4 结束语

本文针对单点家庭用电感知场景的电流分解问题,提出了一种贝叶斯电流分解(BCD)方法.为了同时具备稳态估计方法和线性分解方法的优点,本文将无法同时满足的 2 个假设同时进行松弛,得到 2 个先验分布,并结合贝叶斯公式,将 2 个分布统一的放入一个计算方法中,通过最大化后验概率的方法完成电流分解.

本文使用实验环境下电器的真实用电数据构造仿真实验,对 3 种分解方法进行评测和比较.利用 5 个不同电器的用电数据,我们构造了 26 个不同的仿真场景,每个场景下对 4 000 个周期的电流波形进行分解.在全部的场景中,BCD 方法的分解结果均优于线性分解方法和稳态估计方法.相对于线性分解方法,在 2 电器分解场景下,总体误差降低了 2.5%~44.4%;在 3 电器分解场景下,总体误差降低了 11.5%~46.5%.而相比稳态估计方法,在 2 电器分解场景下,总体误差降低了 73.5%~94.3%;在 3 电器分解场景下,总体误差降低了 71.3%~94.5%.而在本文构造的一个极端情况中,线性分解方法无法运行,BCD 方法的分解误差比稳态估计方法降低了 30%.

参 考 文 献

[1] Mattern F, Staake T, Weiss M. ICT for green: How computers can help us to conserve energy [C] //Proc of the 1st Int Conf on Energy-Efficient Computing and Networking. New York: ACM, 2010: 1-10

[2] Chen Keyu, Gupta S, Larson E, et al. DOSE: Detecting user-driven operating states of electronic devices from a single sensing point [C] //Proc of the 5th Int Conf on Pervasive Computing and Communications. Piscataway, NJ: IEEE, 2015: 46-54

[3] Patel S, Gupta S, Reynolds S. The design and evaluation of an end-user-deployable, whole house, contactless power consumption sensor [C] //Proc of the 28th SIGCHI Conf on Human Factors in Computing Systems. New York: ACM, 2010: 2471-2480

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

[5] Suzuki K, Inagaki S, Suzuki T, et al. Nonintrusive appliance load monitoring based on integer programming [C] //Proc of the 47th SICE Annual Conf. Piscataway, NJ: IEEE, 2008: 2742-2747

[6] Li Jiaming, West S, Platt G. Power decomposition based on SVM regression [C] //Proc of the 4th Int Conf on Modelling, Identification & Control. Piscataway, NJ: IEEE, 2012: 1195-1199

[7] Akbar M, Khan Z A. Modified nonintrusive appliance load monitoring for nonlinear devices [C] //Proc of the 11th Int Multitopic Conf. Piscataway, NJ: IEEE, 2007

[8] Wichakool W, Avestruz A, Cox R, et al. Modeling and estimating current harmonics of variable electronic loads [J]. IEEE Trans on Power Electronics, 2009, 24(12): 2803-2811

[9] Wichakool W, Remscrim Z, Orji U, et al. Smart metering of variable power loads [J]. IEEE Trans on Smart Grid, 2014, 6(1): 189-198

[10] Li Peng, Yu Yixin. Nonintrusive method for on-line power load decomposition [J]. Journal of Tianjin University, 2009, 42(4): 303-308 (in Chinese)
(黎鹏, 余贻鑫. 非侵入式电力负荷在线分解[J]. 天津大学学报, 2009, 42(4): 303-308)

[11] Liu Jingjie, Nie Lei, Xu Zhiwei. The input-sensing problem in ternary computing and its application in household energy-saving [C] //Proc of the 6th Int Conf on Green Computing

and Communications. Los Alamitos, CA: IEEE Computer Society, 2011: 131-138

[12] Jiang Xiaofan, Dawson-Haggerty S, Dutta P, et al. Design and implementation of a high-fidelity ac metering network [C] //Proc of the 8th Int Conf on Information Processing in Sensor Networks. Piscataway, NJ: IEEE, 2009: 253-264

[13] Laughman C, Lee K, Cox R, et al. Power signature analysis [J]. IEEE Power and Energy Magazine, 2003, 1(2): 56-63



Liu Jingjie, born in 1987. PhD. His main research interests include computational intelligence and sensing theories.



Nie Lei, born in 1987. PhD. His main research interests include bioinformatics and compressive sensing.

《计算机研究与发展》征订启事

《计算机研究与发展》(Journal of Computer Research and Development)是中国科学院计算技术研究所和中国计算机学会联合主办、科学出版社出版的学术性刊物,中国计算机学会会刊. 主要刊登计算机科学技术领域高水平的学术论文、最新科研成果和重大应用成果. 读者对象为从事计算机研究与开发的研究人员、工程技术人员、各大专院校计算机相关专业的师生以及高新企业研发人员等.

《计算机研究与发展》于 1958 年创刊,是我国第一个计算机刊物,现已成为我国计算机领域权威性的学术期刊之一. 并历次被评为我国计算机类核心期刊,多次被评为“中国百种杰出学术期刊”. 此外,还被《中国学术期刊文摘》、《中国科学引文索引》、“中国科学引文数据库”、“中国科技论文统计源数据库”、美国工程索引(Ei)检索系统、日本《科学技术文献速报》、俄罗斯《文摘杂志》、英国《科学文摘》(SA)等国内外重要检索机构收录.

国内邮发代号:2-654;国外发行代号:M603
国内统一连续出版物号:CN11-1777/TP
国际标准连续出版物号:ISSN1000-1239
联系方式:
100190 北京中关村科学院南路 6 号《计算机研究与发展》编辑部
电话: +86(10)62620696(兼传真);+86(10)62600350
Email:crad@ict. ac. cn
http://crad.ict. ac. cn