

基于变分稀疏贝叶斯学习的频谱检测方法

朱翠涛 杨凡

(中南民族大学 电子信息工程学院 武汉 430074)

摘要 为了降低对宽带信号进行压缩频谱感知的复杂度,提出了一种基于变分稀疏贝叶斯学习的频谱检测方法.该算法直接利用压缩测量值对授权用户的位置、个数以及功率传播图进行了估计,在先验知识未知的情况下,利用变分稀疏贝叶斯求解稀疏权值.而且用简单函数因子逼近的方法降低了边缘似然函数的计算难度.实验结果表明:该方法在感知精度和速度上有显著提高.

关键词 认知无线电; 频谱检测; 变分稀疏贝叶斯学习

中图分类号 TN911.76 **文献标识码** A **文章编号** 1672-4321(2013)01-0065-05

Spectrum Detection Based on Variational Sparse Bayesian Learning

Zhu Cuitao, Yang Fan

(College of Electronic and Information Engineering, South-Central University for Nationalities, Wuhan 430074, China)

Abstract A novel spectrum detection algorithm based on variational sparse bayesian learning is proposed for reducing the complexity of the compressed spectrum sensing for wideband. The algorithm directly uses the compressed measurement to estimate the location and number of the primary users and adopts variational sparse learning to obtain the sparse weights although the priori knowledge is unknown. It reduces the computational difficulty of marginal likelihood function by adopting the approximation of simple factorial function. The experimental results show that the algorithm significantly improves the accuracy and speed of sensing.

Keywords cognitive radio; spectrum detection; variational sparse Bayesian learning

在认知无线网络中,由于频谱利用率低而导致认知用户接收的宽带信号在频域具有稀疏性,因此,人们将压缩感知技术用于频谱检测,降低对宽带信号采样的压力^[1].大部分文献^[2-3]利用压缩感知技术,先恢复信号,再进行谱的估计,这样导致检测算法的复杂度增加.文献^[4]提出了稀疏贝叶斯方法,直接利用观测值进行相关参数的估计.基于相关向量机的稀疏贝叶斯模型不仅能够估计出授权基站的位置而且也能测出正在通信授权用户的个数.但是,基于相关向量机^[5]的稀疏贝叶斯模型在计算过程中收敛速度很慢.而且,当权值的先验分布服从高斯分布时,单个的稀疏参数可能会在闭合式中

估计^[6].

针对以上情况,我们提出了基于变分稀疏贝叶斯学习模型的方法.该方法用概率分布估计未知变量.这样就避免了参数闭合式估计的情况.而且,对于难以求解的函数,变分学习方法能够分解为几个简单函数来逼近,降低了计算难度.变分方法在推导图示模型时还能够提供一致的推导框架.

1 系统模型

系统模型如图1所示,设在 $N_p \times N_p$ 的区域内,有 N 个认知用户(简称CR用户)、 M_p 个授权用户

收稿日期 2012-12-17

作者简介 朱翠涛(1967-),男,教授,博士,研究方向:认知无线网络及分布式计算, E-mail: zhucuitao@scuec.edu.cn

基金项目 国家自然科学基金资助项目(61072075)

(简称 PU 用户) 和一个融合中心. CR 用户出现的位置用笛卡尔坐标表示为:

$$X = (x_1 \ x_2 \ x_3 \ \dots \ x_n)^T \quad x_i = (x_{ix} \ x_{iy}) .$$

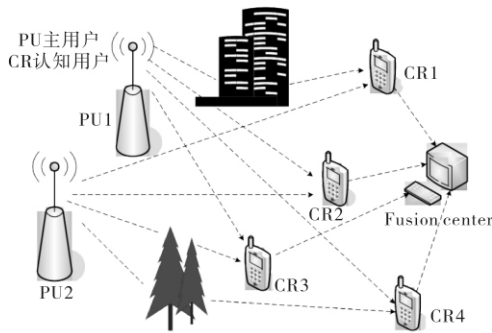


图1 系统模型

Fig.1 System model

该系统模型中,无线电功率传播模型为^[7]:

$$P_r(d) = p_t - (\overline{PL}(d_0) + 10n \ln(\frac{d}{d_0})) + W. \quad (1)$$

式(1)中 n 为路径损耗指数 d_0 为近距离参考距离 d 为发射机与接收机之间的距离 $d = \sqrt{(x_{ix} - u_{jx})^2 - (x_{iy} - u_{jy})^2}$ $\mu_j = (u_{jx} \ u_{jy})$ 描述 PU 用户的位置参数 P_t 为发射功率 $P_r(d)$ 为接收功率 $\overline{PL}(d_0)$ 为在参考距离 d_0 处的路径损耗. W 表示信号在传播过程中受到的影响.

检测期间, CR 用户接收到的信号表示为:

$$t = t_s + n, \quad (2)$$

式(2)中 $t_s = (t_1 \ t_2 \ \dots \ t_N)^T$ $n = (\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n)^T$ 表示信号传输过程中受到的阴影效应 ε_i 表示均值为 0, 方差 σ^2 为的高斯变量.

根据压缩感知理论 t_s 可以表示成转换矩阵 $\Phi(X)$ 与待测未知信号 w 的乘积, 即 $t_s = \Phi(X)w$. 那么观测向量可表示为:

$$t = \Phi(X)w + n, \quad (3)$$

其中 $w = [w_1 \ w_2 \ \dots \ w_M]$,

$$\Phi(X) = \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \dots & \varphi_M(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \dots & \varphi_M(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_1(x_N) & \varphi_2(x_N) & \dots & \varphi_M(x_N) \end{bmatrix} \quad (4)$$

方便起见, 把矩阵(4) 每列简写为 $\varphi_j(X) = (\varphi_j(x_1) \ \varphi_j(x_2) \ \dots \ \varphi_j(x_N))^T$. 根据前面已经给出的无线电功率传播模型(1), 本文算法采用以下函数作为之后的贝叶斯推理的基函数^[4]:

$$\varphi_j(x_i) = \frac{1}{2s_j} \exp\left(-\frac{D}{s_j}\right), \quad (5)$$

式(5)中: $D = \sqrt{(x_{ix} - u_{jx})^2 - (x_{iy} - u_{jy})^2}$, $u_j = (u_{jx} \ u_{jy})$ 描述 PU 用户的位置参数 s_j 为尺度参数.

根据公式(3)、矩阵(4) 可推算出观测向量 t 中的每个元素 t_i 的表达式, 即数学模型如下:

$$t_i = \sum_{j=1}^M w_j \varphi_j(x_i) + \varepsilon_i \quad i = 1 \ 2 \ \dots \ N. \quad (6)$$

式(6) 是一个线性回归模型. 本文采用变分稀疏贝叶斯算法(简称 VSBL 算法) 来求解. VSBL 算法并未对边缘似然函数对数函数直接求解而是通过变分逼近的方法求解出稀疏因子 w .

2 变分稀疏贝叶斯及推理

2.1 变分稀疏贝叶斯模型

频谱检测重构数学模型如公式(6) 所示. 设 CR 用户采样数据是独立同分布的数据集合 $X = \{x_i\}$ $t = \{t_i\}$. 假定 ε_i 为高斯白噪声: $\varepsilon_i \sim N(0, \tau^{-1})$ 那么观测变量的似然函数可表示为^[5]:

$$p(t_i | x_i, w, \sigma) = N(t_i | y(x_i, w), \sigma^{-1}),$$

其中:

$$y(x_i, w) = \sum_{j=1}^M w_j \varphi_j(x_i).$$

那么似然函数可以表示为:

$$p(t | X, w, \sigma) = \prod_{i=1}^N p(t_i | x_i, w, \sigma).$$

如果直接最大化似然函数来估计 w , 会导致模型过学习. 为了克服此问题, 我们假设稀疏参数 w_m 服从均值为 0 方差为 α_m 的高斯分布, 即:

$$p(w_m | \alpha_m) = N(w_m | 0, \alpha_m^{-1}). \quad (7)$$

公式(7) 中 α_m 为超参数, 它决定稀疏因子 w 的稀疏性. 同时为超参数 α_m 引入一个 Gamma 分布:

$$p(\alpha_m) = \Gamma(\alpha_m | a_m, b_m) \equiv b_m^{\alpha_m} \alpha_m^{\alpha_m-1} e^{-b_m \alpha_m} / \Gamma(\alpha_m). \quad (8)$$

式(8) 中 $\Gamma(\alpha_m)$ 是一个 Gamma 函数. 由于 ε_i 噪声协方差为变量, 所以同样为其引入以下分布:

$$p(\tau) = \Gamma(\tau | c, d).$$

由于不存在先验知识, 在初始化时, 我们设定 $a_m = b_m = c = d = 10^{-6}$. 根据以上假设, 变分稀疏贝叶斯概率模型直观图如图 2 所示.

2.2 变分稀疏贝叶斯推理

若已知模型参数的先验概率是 $p(w, \alpha, \tau)$ 那么后验概率可表示为:

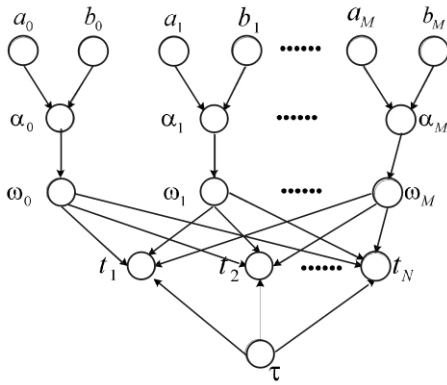


图 2 概率模型直观图

Fig.2 Directed graph representing the probabilistic model

$$p(w \alpha \sigma | t) = \frac{p(t | w \alpha \sigma) p(w \alpha \sigma)}{p(t)},$$

若用稀疏贝叶斯求解后验概率,则先将后验概率分解为:

$$p(w \alpha \sigma | t) = p(w | t \alpha \sigma) p(\alpha \sigma | t),$$

然后再用积分求解 $p(w | t \alpha \sigma)$,算法实现比较复杂. 本文采用变分稀疏贝叶斯的方法对后验概率进行逼近,即:

$$p(w \alpha \sigma | t) \approx Q(w \alpha \sigma). \quad (9)$$

由于变量参数 $w \alpha \sigma$ 对于 Q 分布是相互独立的,那么:

$$Q(w \alpha \sigma) = Q_w(w) Q_\alpha(\alpha) Q_\tau(\tau).$$

从文献[8]可知各个参数服从以下分布:

$$Q_w = N(w | m_w, \sum_w),$$

$$Q_\tau(\tau) = \Gamma(\tau | \bar{c}, \bar{d}),$$

$$Q_\alpha(\alpha) = \prod_{m=0}^M \Gamma(\alpha_m | \bar{a}_m, \bar{b}_m),$$

其中:

$$\sum_w = (\text{diag}(\alpha_m) + \bar{\tau} \sum_{n=0}^N \varphi_n \varphi_n^T)^{-1},$$

$$m_w = \bar{\tau} \sum_{n=0}^N \sum_w \varphi_n^T t_n, \quad (10)$$

$$\bar{a}_m = a + 1/2, \bar{b}_m = b + \frac{w_m^2}{2}, \quad (11)$$

$$\bar{c} = c + (N + 1) / 2, \quad (12)$$

$$\bar{d} = d + \frac{1}{2} \sum_{n=1}^N t_n^2 - \bar{w}^T \sum_{n=1}^N \varphi_n t_n + \frac{1}{2} \sum_{n=1}^N \varphi_n^T \bar{w} \bar{w}^T \varphi_n, \quad (13)$$

$$\bar{w} = m_w \bar{w} w^T = \sum_w + m_w m_w^T, \quad (14)$$

$$\bar{\alpha}_m = \bar{a}_m / \bar{b}_m, \ln \bar{\alpha}_m = \Psi(\bar{a}_m) - \ln \bar{b}_m, \quad (15)$$

$$\bar{\tau} = \bar{c} / \bar{d}, \ln \bar{\tau} = \Psi(\bar{c}) - \ln \bar{d}. \quad (16)$$

公式(15)、(16)中 Ψ 函数定义为:

$$\Psi(a) = \frac{d}{da} \ln \Gamma(a).$$

公式(9)的证明如下:

根据贝叶斯公式,观测数据的边缘概率表示为公式:

$$p(t) = \frac{p(t | w \alpha \sigma)}{p(w \alpha \sigma)}. \quad (17)$$

依据变分理论,引入一个关于变量 $w \alpha \sigma$ 的 Q 分布. 为了保证公式(17)两边等式成立,公式(17)可变换为:

$$p(t) = \frac{p(t | w \alpha \sigma)}{Q(w \alpha \sigma)} \bigg/ \frac{p(w \alpha \sigma | t)}{Q(w \alpha \sigma)}. \quad (18)$$

然后对公式(18)两边取对数:

$$\ln(p(t)) = \ln \frac{p(t | w \alpha \sigma)}{Q(w \alpha \sigma)} - \ln \frac{p(w \alpha \sigma | t)}{Q(w \alpha \sigma)}, \quad (19)$$

设定 $\iiint Q(w \alpha \sigma) dw d\alpha d\tau$ 积分等于 1. 公式(19)可以转化为:

$$\begin{aligned} \ln(p(t)) &= \iiint Q(w \alpha \sigma) \ln \frac{p(t | w \alpha \sigma)}{Q(w \alpha \sigma)} dw d\alpha d\tau = \\ &= - \iiint Q(w \alpha \sigma) \ln \frac{p(w \alpha \sigma | t)}{Q(w \alpha \sigma)} dw d\alpha d\tau. \end{aligned}$$

简记:

$$\ln(p(t)) = L(Q(w \alpha \sigma)) + KL(Q(w \alpha \sigma) \| p(w, \alpha, \sigma | t)). \quad (20)$$

式(20)中 $KL(Q(w \alpha \sigma) \| p(w \alpha \sigma | t)) \geq 0$, 那么对于 $\ln(p(t))$ 来说 $L(Q(w \alpha \sigma))$ 是下界值. $\ln(p(t))$ 数值一定时 $KL(Q(w \alpha \sigma) \| p(w \alpha \sigma | t))$ 值越小 $L(Q(w \alpha \sigma))$ 值越大,函数越收敛. 因此 $Q(w \alpha \sigma)$ 与 $p(w \alpha \sigma | t)$ 越接近越好,即:

$$p(w \alpha \sigma | t) \approx Q(w \alpha \sigma),$$

同时用 $L(Q(w \alpha \sigma))$ 作为算法的收敛条件其公式如下^[8]:

$$\begin{aligned} L(Q(w \alpha \sigma)) &= \langle \ln p(T | X, w, \sigma) \rangle + \\ &= \langle \ln p(w | \alpha) \rangle + \langle \ln p(\alpha) \rangle + \\ &= \langle \ln p(\tau) \rangle - \langle \ln Q_w(w) \rangle - \\ &= \langle \ln Q_\alpha(\alpha) \rangle - \langle \ln Q_\tau(\tau) \rangle. \end{aligned} \quad (21)$$

3 问题求解及算法实现过程

在整个算法实现过程中,由于 u_j 与 s_j 更新速度慢. 所以,超参数 α 与噪声方差 τ 迭代更新一次 μ_j 与 s_j 将更新 P 次. 在处理大量数据时,为了减轻计算负担,本文先计算稀疏参数 w , 然后根据设定的阈值 η ,

当稀疏因子 $w_m < \eta$ 时,删除 w_m 及其所对应的基函数.最后在 w, α, τ 都给定的前提下,我们用梯度下降法对 PU 用户定位^[4].即:

$$\begin{bmatrix} u_{j,x}(k) \\ u_{j,y}(k) \\ s_j(k) \end{bmatrix} = \begin{bmatrix} u_{j,x}(k-1) \\ u_{j,y}(k-1) \\ s_j(k-1) \end{bmatrix} - \delta \begin{bmatrix} \frac{\partial Q}{\partial u_{j,x}} |_{u_{j,x}(k-1)} \\ \frac{\partial Q}{\partial u_{j,y}} |_{u_{j,y}(k-1)} \\ \frac{\partial Q}{\partial s_j} |_{s_j(k-1)} \end{bmatrix} \quad (22)$$

式(22)中 k 为迭代次数, $Q = -L, \delta$ 为学习步长.其中:

$$\frac{\partial Q}{\partial u_{j,x}} |_{u_{j,x}(k-1)} = \frac{\tau}{2} w^T \frac{\partial \Phi^T \Phi}{\partial u_{j,x}} w - \tau w^T \frac{\partial \Phi^T}{\partial u_{j,x}} t, \quad (23)$$

$\frac{\partial Q}{\partial u_{j,y}} |_{u_{j,y}(k-1)}, \frac{\partial Q}{\partial s_j} |_{s_j(k-1)}$ 与公式(23)类似,只需

把 $\frac{\partial \Phi^T \Phi}{\partial u_{j,x}}, \frac{\partial \Phi^T}{\partial u_{j,x}}$ 相应更换为 $\frac{\partial \Phi^T \Phi}{\partial u_{j,y}}, \frac{\partial \Phi^T}{\partial u_{j,y}}, \frac{\partial \Phi^T \Phi}{\partial s_j}, \frac{\partial \Phi^T}{\partial s_j}$.

$$\frac{\partial \Phi^T \Phi}{\partial u_{j,x}} = \begin{bmatrix} 0 & \varphi_1^T \varphi'_{j,\mu_x} & 0 \\ \varphi_1^T \varphi'_{j,\mu_x} & \dots & \varphi_M^T \varphi'_{j,\mu_x} \\ 0 & \varphi_M^T \varphi'_{j,\mu_x} & 0 \end{bmatrix}_{M \times M}, \quad (24)$$

$$\frac{\partial \Phi^T}{\partial u_{j,x}} = \begin{bmatrix} 0 \\ \varphi^T |_{u_{j,x}} \\ 0 \end{bmatrix}_{M \times N}. \quad (25)$$

同理 $\frac{\partial \Phi^T \Phi}{\partial u_{j,y}}, \frac{\partial \Phi^T}{\partial u_{j,y}}, \frac{\partial \Phi^T \Phi}{\partial s_j}, \frac{\partial \Phi^T}{\partial s_j}$ 与公式(24),

(25)类似,只需要把 $u_{j,y}, s_j$ 相应地代替 $u_{j,x}$.

算法的实现过程如下:

(1) 根据公式(5)初始化 M 个基函数,创建基函数矩阵 Φ .

(2) 频谱检测不存在先验知识,初始化 $a = b = c = d = 10^{-6}, k = 0$.根据 CR 用户采集数据及公式(11)、(12)、(13)、(15)、(16)计算 $\alpha, \tau, \bar{a}_m, \bar{b}_m, \bar{c}, \bar{d}$.

(3) 其次由步骤2得到的结果及根据公式(10)计算并 m_w, Σ_w 由公式(21)计算收敛条件 $L(K)$ 的值.

(4) $k = k + 1$ 根据公式(15)、(16)重新计算 α, τ 并根据公式(10)更新 m_w, Σ_w .

(5) 为了减轻计算负担,删除 $w_j < \eta$ 的稀疏参数 w_j 以及所对应的基函数.根据公式(15)、(16)更新 α, τ, Φ .

(6) 然后再用梯度下降法对 PU 用户进行定位.根据公式(22)更新 u_j, s_j ,并且根据公式(10)估计 m_w, Σ_w .再次计算本次迭代之后收敛条件 $L(K)$ 的数值.若 $L(K) < L(K-1)$,表示算法已经收敛.结束循环,否则迭代步骤(6) P 次之后返回到步骤(4).

(7) 根据公式(15)、(16)、(14)、(10),重新估计 $\alpha, \tau, w, \Sigma_w$ 输出 $M_p = M, u_j, s_j$.

4 试验结果及分析

假定被检测区域为 30×30 .在该区域内有3个 PU 用户正在通信(即 $M_p = 3$).PU 用户的发射功率为 30dB(即 $p_t = 30$),在 $PL(d_0)$ 为在 $d_0 = 1$ 处的平均功率为 5dB.功率传播的路径损耗 $n = 1.3$.稀疏因子 w 的选择阈值定义为 $\eta = 2$.

实验(1) 采样率 = 0.9 时本算法对频谱功率传播图进行重构.采样率计算公式如下:

$$\text{Measurement rate} = N/N_p \times N_p.$$

重构效果如图3所示.

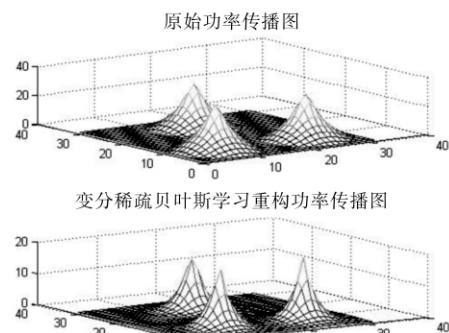


图3 变分稀疏贝叶斯重构图与原图比较

Fig.3 Comparison of image of variational sparse bayesian learning algorithm vs original

图3可看出,采用变分稀疏贝叶斯学习方法能实现很好的重构.

实验(2) 分别设定噪声功率为 10dB、20dB、30dB,比较不同噪声功率对功率图重构的影响.性能比较图如图4所示.

噪声功率越大,对信号干扰越强,重构信号时造成的误差越大.图4可看出,在相同采样率下,噪声功率越大均方误差越大.

实验(3) 比较 VSBL 算法与 SBL 算法对功率传播图重构时造成的均方误差,如图5所示.

从图5可以看出 VSBL 重构均方误差比 SBL 的

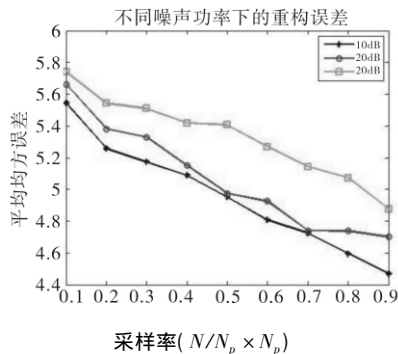


图 4 同噪声功率下功率重构误差比较
Fig. 4 Comparison of mean square error under different noise power

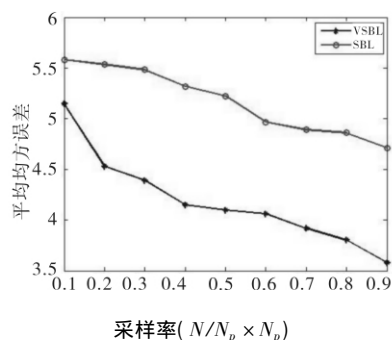


图 5 VSBL 与 SBL 均方误差比较
Fig. 5 Comparison of mean square error of VSBL vs SBL

均方误差要低.

实验(4) 在不同的采样率下,比较 PU 用户真实位置与预测位置之间测均方误差. 如图 6 所示.

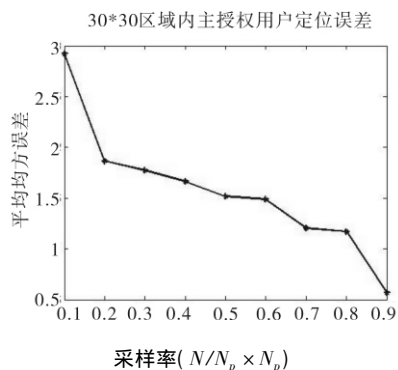


图 6 PU 用户真实位置与预测位置之间均方误差
Fig. 6 Comparison of mean square error of real location vs predicted location

从图 6 可看出,随着采样率的提高,PU 用户真实位置与预测位置之间的均方误差变小.

5 结束语

本文提出的基于变分稀疏贝叶斯频谱感知算法直接利用压缩测量数据对授权用户的位置、个数以及功率传播图进行估计和定位,克服由于采用压缩感知带来的算法复杂性. 且在先验知识未知的情况下,利用变分稀疏贝叶斯求解稀疏权值,用简单函数因子逼近复杂函数的方法降低边缘似然函数的计算难度. 下一步重点研究分布式合作压缩频谱感知算法的一般框架,分析其收敛性质与通信负担.

参 考 文 献

- [1] 石光明,刘丹华,高大化 等. 压缩感知理论及其研究进展[J]. 电子学报,2009,37(5): 1070-1081.
- [2] Tian Z, Giannakis G. Compressed sensing for wideband cognitive radios [C]//IEEE. IEEE International Conference. Honolulu: IEEE, 2007: 1357-1360.
- [3] Ji S, Xue Y, Carin L. Bayesian compressive sensing [J]. IEEE Trans on Signal Processing, 2008, 56(6): 2346-2356.
- [4] Tina Huang D-H, Wu S-H, Wang P-H. Cooperative spectrum sensing and locationing: a sparse bayesian learning approach [C]//GTC. Proc Global Telecommunications Conference. New York: GTC, 2010: 1-5.
- [5] Tipping M E. Sparse Bayesian learning and the relevance vector machine [J]. Journal of Machine Learning Research, 2001, 1: 211-244.
- [6] Shutin D, Buchgraber T, Kulkarni SR, et al. Fast variational sparse Bayesian learning with automatic relevance determination [J]. IEEE Trans on Signal Process, 2011, 59(12): 6257-6261.
- [7] Rappaport T S. Wireless communications: principles and practice [M]. 2nd ed. PTR: Prentice Hall, 2002.
- [8] Bishop C M, Tipping M E. Variational relevance vector machines [C]//CUAI. Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence. New York: CUA I, 2000: 46-53.