

基于非高斯度量的多用户 DS-CDMA 系统 多干扰消除与用户检测

崔志富 张 杭 张 江 辜方林

(解放军理工大学通信工程学院 南京 210007)

摘 要 非同步的直接序列扩频码分复用(DS-CDMA)系统可能同时存在符号间串扰、多址干扰以及外部干扰。将扩频系统普遍存在的外部窄带干扰作为讨论重点,依据函数空间理论,对干扰进行有限维函数空间上的分解,并将基函数作为混合向量,基函数下的坐标作为待恢复源,通过过采样处理,结合 DS-CDMA 系统多用户检测模型,提出一种基于盲分离的算法。该算法在未知扩频序列以及信道多径的情况下完成盲多用户检测与盲均衡处理,同时实现外部窄带干扰消除,提高系统的干扰容限。仿真结果表明,在不同的信噪比以及信干比条件下,该算法普遍优于带理想均衡器的 RAKE 接收机,尤其在干扰显著的情况下,能获得更高的增益。

关键词 直接序列扩频,盲多用户检测,盲均衡,干扰消除,盲源分离,最大非高斯

中图法分类号

文献标识码 A

External Interference Cancellation and Blind Multi-user Detection of DS-CDMA System Based on Non-gaussianity Measure

CUI Zhi-fu ZHANG Hang ZHANG Jiang GU Fang-lin

(Institute of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China)

Abstract In an asynchronous DS-CDMA system of multiple users, there are inter-symbol interference (ISI), multiple access interference (MAI) and external interference (EI). According to function space theory, decomposing the narrow-band external interference on a function space with finite dimensions, regarding each basis of the function space as a vector of the mixing matrix, and taking the coefficient of corresponding basis as one of sources, we proposed a BSS-based method, requiring the received signal to be over-sampled, which can realize both blind equalization and blind multi-user detection in the case that all these three interferences coexist simultaneously. The simulation result indicates that this method performs better than RAKE with perfect equalization, especially in lower SNRs.

Keywords DS-CDMA, Blind multiple user detection, Blind equalization, External interference cancellation, Blind source separation, Maximum non-gaussianity

非同步直接序列扩频的码分复用(DS-CDMA)系统可能同时存在符号间串扰(Inter Symbol Interference, ISI)、多址干扰(Multiple Access Interference, MAI)以及外部干扰(External Interference, EI),系统的扩频增益是系统用户容量的主要决定因素。然而,存在的各类干扰不仅影响用户的工作性能,而且也将成为系统容量的主要制约因素。符号间串扰一般通过均衡技术来消除,减小多址干扰的影响主要依赖多用户检测,而外部干扰消除往往需要阵列天线等前端设备的支持,这无疑增加了系统代价并导致在很多情况下难以实现。

盲源分离由于能够在源信号和混合系统已知信息较少的情况下实现对源信号和混合系统的估计,因此在 DS-CDMA 系统盲多用户检测以及盲均衡(Blind Equalization)领域获得了广泛的应用。A. D. Baynast 和 H. Chen 等人^[1-3]针对 DS-CDMA 通信系统中通信信号盲检测问题,采用过采样技术将单通道或多路接收信号转化为一个高维多入多出模型,借助用户扩频序列等先验信息,利用盲源分离技术实现了特定用户

信号检测。文献[4,5]等将上述模型推广到了卷积非同步多人多出系统,但与文献[1-3]一样,只考虑了系统中存在多址干扰的情形,没有考虑外部干扰的影响。K. Raju 在其博士论文中考虑了 DS-CDMA 通信系统同时存在多址干扰、区间干扰和外来干扰的情形^[6],提出了 ICA-Rake 接收机,但该方法需要依赖阵列天线来减少外来干扰对接收机性能的影响。T. Huovinen 在其博士论文中考虑了扩频通信系统同时存在多址干扰和外来干扰的情形^[7],但在单天线接收的条件下,该方法只能处理与载波相同频率的单音干扰。文献[8,9]在上述基础上,分别提出基于模型和空时特性的盲多用户检测与盲均衡方法。本文针对多用户非同步 DS-CDMA 系统存在外部窄带干扰的情况,结合信道的多径效应和外部干扰的窄带特征,提出一种联合实现扰信分离与多用户检测的方法,该方法不依赖多天线支持,仅利用一路观测信号便能同时消除上述 3 种干扰。

本文首先分析 DS-CDMA 系统特点以及链路特性,据此

建立基于盲源分离的多用户检测模型,并结合外部干扰的窄带特征,将其分解并将分解后的系数视为源信号,反映在高维的多入多出(MIMO)模型中;以该模型为基础,利用基于最大非高斯的盲分离方法完成多用户检测并实现干扰消除;最后给出仿真验证的结果以及相应的结论。

1 多用户信号模型

DS-CDMA 系统内可能同时存在多个在线用户,每路信号经历不同的信道传输过程,到达接收端时一般是非同步的,对于单天线接收的多用户系统,当存在外部干扰(External Interference)时,接收信号可以表示为

$$r(t) = s(t) + j(t) + n(t) \quad (1)$$

式中, $s(t)$ 表示系统内多个用户信号经过多径信道后相互叠加的结果; $j(t)$ 表示外部干扰; $n(t)$ 表示观测噪声,一般认为是加性高斯信号。

每路扩频信号由于经历不同的信道传输过程,因此具有不同的信道衰减和传输延时,则对任意采样时刻 t , 有

$$s(t) = \sum_{k=1}^K \sum_{m=1}^M b_k(m) \sum_{l=1}^{L_k} a_{kl} c_k(t - mT_s - \tau_{kl}) \quad (2)$$

K 表示系统内活跃的用户数, a_{kl} 、 τ_{kl} 分别表示第 k 路信号、第 l 条多径的信道衰减和延时, $b_k(m)$ 表示第 k 个用户的第 m 个信息符号, c_k 表示第 k 个用户的扩频序列, T_s 表示符号周期。对于一般 DS-CDMA 系统,可以提出如下合理假设:

1) $b_k(m)$ 是独立同分布信号,且不同源信号之间相互独立;

2) 扩频序列 $c_k(t) = \sum_{j=1}^{L_c} \beta_{kj} \phi(t - jT_c)$, 其中 L_c 表示扩频增益, β_{kj} 是码片序列,一般认为是二元的, $\phi(t - jT_c)$ 表示宽度为 $T_c = T_s/L_c$ 的脉冲波形;

3) 信道是准稳态的。

以码片速率 $1/T_c$ 对接收信号进行采样,并将码元周期 T_s 作为一个观测周期。令 $d_{kl} = \lfloor \tau_{kl}/T_c \rfloor$, $\lfloor \cdot \rfloor$ 表示下取整运算,且令 $\epsilon_{kl} = (\tau_{kl}/T_c) - d_{kl}$ 。当信道最大的多径延时不超过一个符号周期时, $d_{kl} = L_s - 1$, $\epsilon_{kl} \in [1, 0)$ 。如果以码片速率对观测信号 $r(t)$ 进行采样,将在第 m 个码元周期上获得的全部采样点作为一个向量 $r(m)$ 。多用户信号 $s(t)$ 可以表示为^[2-5]

$$s(m) = \sum_{k=1}^K b_k(m-1) \sum_{l=1}^{L_k} a_{kl} C_{kl}^{d_{kl}} + \sum_{k=1}^K b_k(m) \sum_{l=1}^{L_k} a_{kl} C_{kl}^{\epsilon_{kl}} \quad (3)$$

其中

$$C_{kl}^{d_{kl}} = [0^{1 \times d_{kl}}, c_k(1), \dots, c_k(L_s - d_{kl})]^T \quad (4)$$

$$C_{kl}^{\epsilon_{kl}} = [c_k(L_s - d_{kl} + 1), \dots, c_k(L_s), 0^{1 \times (L_s - d_{kl})}]^T$$

该模型可以扩展至多径延时超过一个符号周期的情况^[5]。在多径延时不超过一个符号周期的模型中,令

$$g_{2k-1} = \sum_{l=1}^{L_k} a_{kl} C_{kl}^{d_{kl}}, g_{2k} = \sum_{l=1}^{L_k} a_{kl} C_{kl}^{\epsilon_{kl}} \quad (5)$$

且令

$$G = [g_1, g_2, \dots, g_{2K}] \quad (6)$$

$$b(m) = [b_1(m-1), b_1(m), b_K(m-1), b_K(m)]^T$$

则第 m 个观测周期上,多用户系统合成信号 $s(m)$ 可以表示为

$$s(m) = Gb(m) \quad (7)$$

2 干扰模型

对于接收信号 $r(t)$ 中的外部干扰 $j(t)$,考虑其为窄带干

扰的情况。任意窄带过程都可以表示为

$$j(t) = a(t) \cos(2\pi\Delta ft + \phi(t)) \quad (8)$$

式中, $a(t)$ 表示干扰的幅度, Δf 表示干扰频率与载波频率差, $2\pi\Delta ft + \phi(t)$ 表示信号的瞬时相位, $\phi(t)$ 表示信号相位的变化过程。不失一般性,假设该 Δf 是非时变的,且在一个观测周期 $[(m-1)T_s, mT_s]$ 上, $a(t)$ 为缓变的, $\phi(t)$ 的变化满足

$$\tau_0 = 0, \tau_{P+1} = T_c,$$

$$\phi(t) = \sum_{p=0}^P \phi_{pm}(u(t - \tau_p - (m-1)T_c) - u(t - \tau_{p+1} - (m-1)T_c))$$

式中, $u(t)$ 为阶跃函数, P 表示一个观测周期上的相位跳变次数, ϕ_{pm} 表示不同间隔上 $\phi(t)$ 的取值。将一个观测间隔分成 P 个固定的子区间,且令 $q_p = \lfloor (\tau_p - \tau_{p-1})/T_c \rfloor$, $\lfloor \cdot \rfloor$ 表示四舍五入运算, $n=1, \dots, P$; 将第 m 个观测间隔上窄带干扰 $j(t)$ 的采样值表示为 $J(m)$,并分别定义相互正交的矢量组为

$$J_c^1(m) = [1, \cos(2\pi\Delta f T_c), \dots, \cos(2\pi\Delta f (q_1 - 1)T_c), 0^{1 \times (L_s - q_1)}]^T$$

$\vdots$

$$J_c^P(m) = [0^{1 \times (L_s - q_P)}, \cos(2\pi\Delta f q_P T_c), \dots, \cos(2\pi\Delta f L_s T_c)]^T \quad (9)$$

$$J_s^1(m) = [0, \sin(2\pi\Delta f T_c), \dots, \sin(2\pi\Delta f (q_1 - 1)T_c), 0^{1 \times (L_s - q_1)}]^T$$

$\vdots$

$$J_s^P(m) = [0^{1 \times (L_s - q_P)}, \sin(2\pi\Delta f q_P T_c), \dots, \sin(2\pi\Delta f L_s T_c)]^T$$

将 $J(m)$ 在这组矢量下分解,得到相应的系数为

$$a_c^1(m) = a_1(m) \cos(2\pi f \tau_m + \phi_{1m})$$

$\vdots$

$$a_c^P(m) = a_P(m) \cos(2\pi f \tau_m + \phi_{Pm}) \quad (10)$$

$$a_s^1(m) = -a_1(m) \sin(2\pi f \tau_m + \phi_{1m})$$

$\vdots$

$$a_s^P(m) = -a_P(m) \sin(2\pi f \tau_m + \phi_{Pm})$$

则第 m 个观测间隔上干扰的采样 $J(m)$ 可以表示为

$$J(m) = a_c^1(m) J_c^1(m) + a_s^1(m) J_s^1(m) + \dots + a_c^P(m) J_c^P(m) + a_s^P(m) J_s^P(m) \quad (11)$$

根据以上分析,分别将基矢量和相应的系数作为新的混合向量和源信号,对式(6)加以拓展,令

$$\tilde{G} = [J_c^1, J_s^1, \dots, J_c^P, J_s^P, G]^T \quad (12)$$

$$\tilde{B} = [a_c^1(m), a_s^1(m), \dots, a_c^P(m), a_s^P(m), b(m^T)]^T \quad (13)$$

则第 m 个观测周期上的接收信号 $r(t)$ 的采样序列 $r(m)$ 可以表示为

$$r(m) = \tilde{G}\tilde{B}(m) + n(m) \quad (14)$$

式中, $n(m)$ 表示 $n(t)$ 相应的采样序列。此时,一个多径信道在单通道接收条件下,多用户非同步 DS-CDMA 系统的盲均衡、多用户检测与干扰消除问题转变成为了一个高维的时不变瞬时混合的盲分离问题。一般认为 DS-CDMA 系统中,同时在线的用户数远小于扩频序列长度 L_c ,且窄带过程参数变化较缓慢,即在一个符号周期内 P 值较小,可认为 $L_c \geq 2K + 2P$ 。此时,该系统的信息传递是完备的,即混合系统是适定或者是超定的。

如式(14)所示,一个单通道接受的多用户非同步的 DS-CDMA 系统被转变为了一个多入多出的瞬时混合盲分离问

题,系统模型如图 1 所示。基于上述模型,利用盲源分离的方法同时实现 CDMA 系统的多用户检测以及干扰消除。

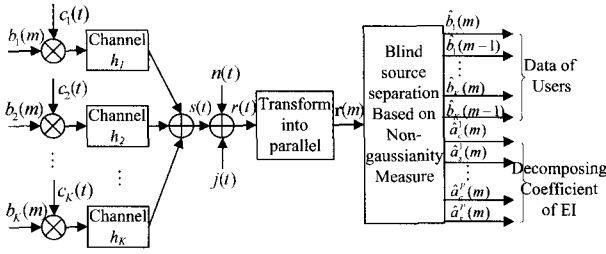


图 1 多用户检测及干扰消除模型

3 基于最大非高斯度量的多用户检测与干扰消除算法

3.1 算法描述

对于图 1 所示的瞬时混合模型,其等效源为 $a(m)$ 和 $b(m)$,即便假设 $b_i(m)$ 、 $b_j(m)$ 和 $j(t)$ 是相互独立的,由于信道编码的存在,同一用户信息序列不同符号之间(如 $b_i(m)$ 和 $b_i(m-1)$)的独立性也是没有保证的;此外,由于 a_{mi} 和 φ_{mi} 分布的不确定性,假设 $a_i^*(m)$ 与 $a_i^*(m)$ 、 $a_i^*(m)$ 是相互独立的也不尽合理,这就意味着基于独立成分分析的分方法不一定奏效,起码不能保证理想的分离效果。

一些前人的工作为我们提供了解决相关源盲分离问题^[10,11]的途径。联合块对角化算法(简称 JBD 算法)是 JADE 算法的扩展^[10],该算法能够将源信号进行归类处理,其分类则是根据源之间的统计独立性,并在此基础上对相互独立的不同类联合块对角化处理,进而完成类与类的分离。这种方法的优势在于即便某些源信号之间存在强相关,仍不影响该算法实现类的分离。然而,由于算法要求属于不同类的源个数一致,当条件不能得到满足时,该算法的有效性将明显下降。除此类方法外,C. F. Cai 还证明基于最大非高斯度量的盲扰信分离方法^[11]在处理相关源时往往比基于最小互信息(MinMI)的方法更有效,在此基础上 Michael Novey 将该方法扩展至复数情况^[12]。

若采用基于最大非高斯度量的方法实现相关源的分离,首先需要建立能够量化源信号估计的非高斯性的代价函数,在此基础上搜索观测信号的某个线性组合,使得该代价函数取得最值。本文将待估源信号概率密度函数与标准高斯分布函数之间的欧几里德距离(l^2 -Euclidean Distance)作为实现分离的代价函数^[13],该距离可以表示为

$$\Gamma(p_y) = \int_{\gamma} [\Phi(y) - p_y(y)]^2 dy \quad (15)$$

式中, γ 表示积分变量 y 的变化范围, $p_y(y)$ 表示其概率密度函数分布,在本文的应用中, $\Gamma(p_y)$ 表示源信号估计的概率分布,是未知的, $\Phi(y)$ 表示高斯分布的概率密度函数,即

$$\Phi(y) = N(0,1) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}y^2) \quad (16)$$

要计算式(16)给出的二次欧几里德距离,必须首先获得源信号估计的概率密度函数 $p_y(y)$ 。本文采用 Parzen 窗的方法来估计,这是一种非参数化的方法,其利用核函数建立起变量实际采样值与其概率密度之间的函数关系。本文采用高斯核函数,概率密度的估计方法如式(17)所示,其中函数 $\Phi(\cdot)$ 的定义如式(16)所示。

$$\hat{p}(y) = \frac{1}{Mh} \sum_{i=1}^M \Phi(\frac{y-y(i)}{h}) \quad (17)$$

将式(15)展开,有

$$\Gamma(p_y) = \int \Phi^2(y) dy - 2 \int \Phi(y) p_y(y) dy + \int p_y^2(y) dy \quad (18)$$

将式(18)中的 $p_y(y)$ 用式(17)中的 $\hat{p}(y)$ 做一下替换,则式(18)转变为

$$\Gamma(\hat{p}_y) = \frac{1}{2\sqrt{\pi}} + \Gamma_1(\hat{p}_y) + \Gamma_2(\hat{p}_y) \quad (19)$$

其中

$$\begin{aligned} \Gamma_1(\hat{p}_y) &= -2 \int \Phi(y) \hat{p}_y(y) dy \\ &= -\frac{2}{M\sqrt{h^2+1}} \sum_{m=1}^M \Phi(\frac{y(m)}{\sqrt{h^2+1}}) \end{aligned} \quad (20)$$

$$\begin{aligned} \Gamma_2(\hat{p}_y) &= \int \hat{p}_y^2(y) dy \\ &= \frac{2}{M^2\sqrt{h^2+1}} \sum_{m=1}^M \sum_{n=1}^M \Phi(\frac{y(n)-y(m)}{\sqrt{2}h}) \end{aligned} \quad (21)$$

h 是确定参数,有文献证明当 h 取 $1.06 \times M^{-1/5}$ 时,能够保证概率密度估计的均方误差(Minimum Mean Integrated Square Error, MISE)最小^[11]。

由于假设传输的数据是 BPSK 调制的,因此可以通过对观测信号进行空间白化来确保估计得到的源信号的方差为 1 (unit-variances)。空间白化的好处在于解决了分离结果的比例不确定问题,当然符号的正负变化仍将是不可避免的,好在另有一些方法能够解决符号的不确定问题,比较典型的的就是差分编码方法,本文不对此做更多的讨论。对等效接收信号的空间白化,我们采用基于 KL 变换(Karhunen-Loeve Transformation, KLT)的空间滤波器来完成,以期获得一组空间上相互无关的向量集合,其具体的变换式如下。

$$\tilde{r} = V\Lambda^{-1/2}V^T r \quad (22)$$

式中, V 是由观测信号协方差矩阵 $R_r = E[rr^T]$ 的特征向量所构成的矩阵; Λ 是一个对角阵,其对角线元素由 R_r 的特征值构成,且我们假设这些特征值均不为零。

若定义 $\tilde{D}$ 为基于白化数据的新的理想分离矩阵,则该矩阵的每一行都一定是归一化的。在我们的应用中,问题就简化为在 $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_{L_s}$ 张成的子空间上搜索 $2K+2P$ 个使得非高斯度量函数 $\Gamma(p_y)$ 取得极值的点, y 就是某路源信号的估计,且可以表示为等效观测信号的某种线性组合,如式(23)所示。

$$y(m) = \tilde{d}_1 \tilde{r}_1(m) + \tilde{d}_2 \tilde{r}_2(m) + \dots + \tilde{d}_{L_s} \tilde{r}_{L_s}(m) \quad (23)$$

由于分离矩阵各行是归一化的,即 $\sum_{i=1}^{L_s} \tilde{d}_i^2 = 1$,因此分离矩阵的各行所属的整个向量解空间均可以由 L_s-1 维超球面内的一个角度表示。上述关系可以描述为

$$\tilde{d} = [\cos\theta_1, \sin\theta_1 \cos\theta_2, \dots, \prod_{n=1}^{L_s-2} \sin\theta_n \cos\theta_{L_s-1}, \prod_{n=1}^{L_s-1} \sin\theta_n]^T a \quad (24)$$

若要搜索函数的局部极值点,可采用基于梯度下降的方法,其运算量小,又能保证算法的收敛性。

由于

$$\begin{aligned} \frac{\partial \Gamma(p_y)}{\partial \tilde{d}} &= \frac{\partial \Gamma_1(p_y)}{\partial \tilde{d}} + \frac{\partial \Gamma_2(p_y)}{\partial \tilde{d}} \\ &= \frac{2}{M\sqrt{h^2+1}} \sum_{i=1}^M y(i) \Phi(\frac{y(i)}{\sqrt{h^2+1}}) \tilde{r}(i) + \frac{1}{M^2 h^2} \sum_{i=1}^M \end{aligned}$$

$$\tilde{r}(i) \sum_{j=1}^M \left(\frac{y(j) - y(i)}{\sqrt{2}h} \right) \Phi \left(\frac{y(j) - y(i)}{\sqrt{2}h} \right) \quad (25)$$

因此, θ 的迭代方程为

$$\theta^{(n+1)} = \theta^{(n)} - \frac{\partial \Gamma(p_y)}{\partial \tilde{d}} \frac{\partial \tilde{d}}{\partial \theta} \quad (26)$$

在估计得到 $\hat{\theta}$ 以后, 利用关系式(24)就可以求得相应的分离向量 $\hat{d}$, 进而利用关系式 $\hat{y} = \hat{\theta} \hat{r}$ 计算相应的源信号估计。

在重新搜索新的极值点之前, 必须移除上一个极值点, 以免搜索结果落在同一个极值区间内。这种极值点移除的方法称为非高斯结构移除, 方法是通过对数据进行处理, 在不影响其在该极值点取值的前提下降低特定位置的非高斯性。本文选择的非高斯结构移除方法在文献[13]中有更详细的介绍。在搜索得到与源个数(L_s 个)相当的 θ 以后, 相应地也就获得了等效源的估计。接下来就需要通过对等效源进一步处理, 以获得每个用户的传输数据。如式(13)所示, 等效源信号中, 必然存在两个等效源之间满足延时相关关系, 即其中一个为 $b_i(m)$ 且另存在一个等效源 $b_i(m-1)$, 即这两个等效源来自相同的用户信息序列的不同延时。利用这一性质, 可以从估计到的等效源中提取每一个用户的信息序列。此外, 由于这种关系的存在, 可以在进行判决的过程中进一步提高判决性能。

利用这种相对时延关系的用户信息序列提取与联合判决方法的实现步骤如下:

首先, 计算每两路信号之间的互延时相关关系, 并取其最大值作为两个等效源之间相关关系的度量, 即

$$\rho_{ij}^1 = E[y_i(m) y_j(m-1)] \approx \frac{\sum_{m=2}^M \hat{y}_i(m) \hat{y}_j(m-1)}{\sqrt{\sum_{m=2}^M \hat{y}_i^2(m)} \sqrt{\sum_{m=1}^{M-1} \hat{y}_j^2(m)}} \quad (27)$$

$$\rho_{ij}^2 = E[y_i(m-1) y_j(m)] \approx \frac{\sum_{m=2}^M \hat{y}_i(m-1) \hat{y}_j(m)}{\sqrt{\sum_{m=1}^{M-1} \hat{y}_i^2(m)} \sqrt{\sum_{m=2}^M \hat{y}_j^2(m)}} \quad (28)$$

并定义 $\tilde{\rho}_{ij} = \max(|\rho_{ij}^1|, |\rho_{ij}^2|)$ 。这样, $\tilde{\rho}_{ij}$ 就可以构成一个矩阵。在此基础上, 每一行的最大值对应的两个等效源认为是来自同一个信息序列。而且, 如果相应的 $\tilde{\rho}_{ij}^k$ (使得 $\tilde{\rho}_{ij} = |\tilde{\rho}_{ij}^k|$ 成立, $k=1$ 或 2) 小于 0, 则意味着两个等效源之间出现了符号相反的情况(这在 BSS 问题中是不可避免的, 因此在计算误码率时, 这种整体的符号变化不识别, 是误码), 需要对其中一路加以纠正。在判决过程中, 如果 $\hat{y}_i(m)$ 和 $\hat{y}_j(m-1)$ 被确认来自同一用户的信息序列, 那么可通过如下方法重构该序列:

$$\begin{aligned} \hat{b}_k(1) &= \hat{y}_i(1) \\ \hat{b}_k(m) &= \Delta_1(m) \hat{y}_i(m) + \text{sign}(\rho_{ij}^2) \Delta_2(m) \hat{y}_j(m-1), m = \\ &2, \dots, M \end{aligned} \quad (29)$$

其中

$$\Delta_1(m) = \begin{cases} 1, & |\hat{y}_i(m) - \text{sign}(\hat{y}_i(m))| < \\ & |\hat{y}_j(m-1) - \text{sign}(\hat{y}_j(m-1))| \\ 0, & \text{else} \end{cases}$$

$$\Delta_2(m) = \begin{cases} 1, & |\hat{y}_i(m) - \text{sign}(\hat{y}_i(m))| \geq \\ & |\hat{y}_j(m-1) - \text{sign}(\hat{y}_j(m-1))| \\ 0, & \text{else} \end{cases} \quad (30)$$

这组新序列即为最终估计到的用户信息序列, 将这种基于最大非高斯判别的分离算法简称为 BSS-NG, 并将相应的基于最小互信息的方法称为 BSS-MI。

3.2 仿真结果

这一节将仿真分析基于最大非高斯度量的非同步 DS-CDMA 系统干扰消除算法的误码率性能(BER), 并将其与其他方法, 尤其是基于最小互信息的方法进行比较, 以说明该算法的优势所在。仿真中, 将本文所提方法简称为 BSS-NG, 比较的算法包括不考虑外部窄带干扰(EI)而直接利用盲分离方法实现多用户检测(BSS-UEI)的方法、基于最小互信息的盲分离方法(BSS-MI)以及几种传统的多用户检测算法 MBER^[3]、IM-P 和 EM-EML^[9]。仿真条件如下:

- 1) 各用户所使用的扩频码均由 GOLD 序列产生, 且相互间具有良好的正交性, 每个用户的扩频均为 $L_s = 31$;
- 2) 外部窄带干扰服从特定区间内的均匀分布, 该区间跨度将决定外部干扰的功率;
- 3) 认为外部干扰在不同处理间隔上的相位 ϕ_{pm} 服从 $[0, 2\pi]$ 的均匀分布, 并认为这样的假设在每个采样点上带来的相位误差 $\Delta\phi(t)$ 是一个零均值随机变量, 且该变量服从 $[-\delta_p, \delta_p]$ 均匀分布, δ_p 决定了近似误差的大小;
- 4) 认为每个观测区间上干扰的相位不发生跳变, 也就是说每个观测区间上不再划分子区间, 即子区间个数 $P=1$;
- 5) 每条信道的多径数目为 10, 其衰减系数 α_{ki} 和相对时延 τ_{ki} 均随机产生, 并假设其最大的多径时延不超过一个观测间隔。

根据以上假设, 外部干扰 EI 可以表示为

$$j(t) = a(m) \cos(2\pi\Delta f t + \phi(m) + \Delta\phi(t)) \quad (31)$$

为检验所提算法的抗干扰性能, 首先针对单用户在线情况进行了仿真, 并与其他方法进行了比较。作为比较对象的算法包括基于最小互信息的方法(简称为 BSS-MI)、将不考虑外部干扰的存在仅将盲分离方法用于 DS-CDMA 系统解扩的方法(简称为 BSS-UEI)以及带有理想均衡器和无均衡机制的匹配滤波接收机, 分别简称为 MFR-PE 和 MFR-WE。在不存在外部干扰和多址干扰的情况下, 若已知扩频序列, 则带有理想均衡器的匹配滤波接收机是最理想的接收机。由于此时不存在多址干扰, 因此影响系统性能的决定性因素是外部干扰和码间串扰(仿真中干扰是窄带的), 图 2 和图 3 分别给出了不同算法分别随信噪比(Signal to Noise Ratio, SNR)和信干比(Signal to Interference Ratio, SIR)的变化曲线。仿真中, 将最大的相位近似误差 δ_p 取为 $\pi/20$, 这就意味着各采样点上的相位近似误差服从 $[-\pi/20, \pi/20]$ 的均匀分布。图 2 表明, 各种算法(BSS-NG, BSS-MI, BSS-UEI 和 MFR)在信噪比逐渐增加的情况下, 性能均有所改善, 且本文所提算法 BSS-NG 性能的改善尤其明显。观察图 3 容易发现, 若排除随机因素的影响, 算法 BSS-NG 和 BSS-MI 的 BER 性能几乎不随 SIR 变化。这是由于盲分离方法所固有的尺度不确定性(Scaling Uncertainty)造成的, 简单地说, 盲分离方法在实现过程中可能将源信号功率的差异作为混合矩阵相应元素的差异来处理

理,因此不可避免地使分离结果的功率比不同于真实源信号之间的功率比,换言之,只要源信号之间的功率没有极端的差异,盲分离方法均对这种差异不敏感。为此,我们通常进行空间白化处理,使得分离结果是方差归一化的。后续的仿真也表现出这种对源信号功率的不敏感性。

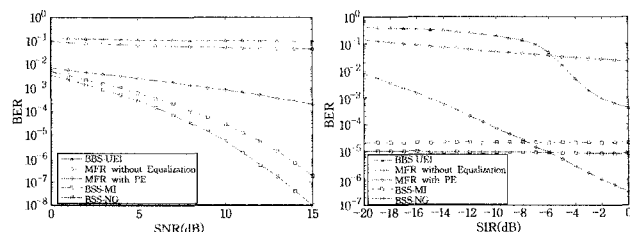


图2 单用户条件下 BER 随信噪比的变化曲线(信干比 $SIR = -15\text{dB}$, 最大相位误差 $\delta_P = \pi/20$)

图3 单用户条件下 BER 随信干比的变化曲线(信干比 $SNR = 10\text{dB}$, 最大相位误差 $\delta_P = \pi/20$)

结合图2和图3,可以认为在当前的仿真环境下,无论随SNR还是SIR变化,基于非高斯度量的算法BSS-NG都表现出了比基于最小互信息的算法更优越的性能。在干扰功率较大($SIR = -6 \sim -20\text{dB}$)时,该算法的性能优于带有理想均衡器的匹配滤波接收机,随着干扰功率的下降,匹配滤波接收机的性能逐步改善,这是由于单用户的CDMA系统本身就具有一定的抗干扰容限,多用户CDMA系统之所以不能对抗外部干扰,是因为多址干扰已经达到了该容限。当干扰功率与信号功率相当时,带有理想均衡器的MFR具有最好的BER性能,而BSS-UEI以及无均衡器的MFR在较低的信干比条件下几乎是不能工作的。此外不难发现,同基于最小互信息的方法一样,与SIR相比,SNR是影响该算法性能的主要因素。由于信道滤波的存在,码间串扰显然是不可避免的,该仿真也证明该算法在消除码间串扰方面同样是有效的。

图4为不同算法随SNR变化的BER性能曲线。该仿真中,令 $K=2$,且令信干比 $SIR = -10\text{dB}$,相位误差仍选择为 $[-\pi/20, \pi/20]$ 的均匀分布。可以看出,在当前条件下BSS-NG获得的性能最优,BSS-MI、EM-EML和带有理想均衡器的IM-P依次次之,但性能比较接近。MBER算法^[3]和带有理想均衡器的ACMA算法^[15]在高信噪比条件下仍能工作,但性能较差,而在SNR较低的情况下两种算法是无效的。

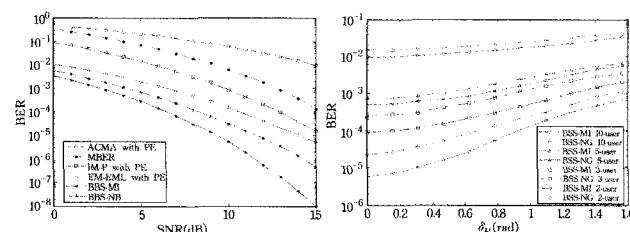


图4 不同算法 BER 性能随 SNR 的变化曲线($K=2$, $SIR = -10\text{dB}$, $\delta_P = \pi/20$)

图5 BSS-CEI 算法 BER 性能随相位近似误差 δ_P 的变化曲线($SNR = 10\text{dB}$, $SIR = -10\text{dB}$)

考察CDMA系统干扰消除算法的工作性能知,有较大影响的参数之一是相位的近似误差。在仿真条件中交代过, $\Delta\phi(t)$ 服从 $[-\delta_P, \delta_P]$ 上的均匀分布。图5给出了不同用户条件下,随着相位近似误差的变化,BSS-NG和BSS-MI两种算法

的BER性能曲线。仿真中,令 δ_P 从0变化到 $\pi/2$,用户数量分别为 $K=2,3,5,10$ 。不难发现,在相同的相位误差水平下,BSS-NG算法性能比BSS-MI更优,且当用户数量固定时两种算法的性能随着相位误差的增加逐步下降。显然,减小相位近似误差有助于改善算法性能,而若要减小相位近似误差代价,将需进一步增加用于分解的基函数的数量,这就意味着等效源数量的增加,源信号数量增加也同样会引起分离性能下降的问题,同时带来运算复杂度提高的问题,为此必须寻求两者的合理折中。

此外,我们仿真检验了本文算法在不同用户数量条件下的BER性能。在不同外部干扰以及噪声的条件下,不同用户数量对应的BER曲线分别如图6和图7所示,其中SIR和SNR分别为 -10dB 和 10dB 。图6显示,BSS-NG算法以及BSS-MI算法随SIR的变化几乎保持不变,这进一步证明了我们得到的结论。结合图6和图7知,在不同在线用户数量的条件下BSS-NG算法的性能均优于BSS-MI算法,当系统在线用户数量较少时,这种优越性尤其明显;且随着系统在线用户数量的增加,两种算法的性能都会下降,这也是盲分离方法固有的缺陷之一。一般来讲,当混合矩阵的估计存在误差时,即便对矩阵各元素的估计误差并没有增加,源恢复的性能也会随着源的数量增多而变得越差。

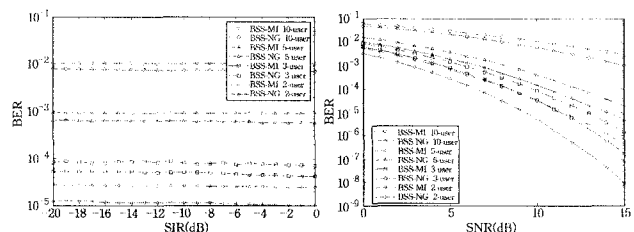


图6 不同用户数量条件下算法 BER 性能随 SIRs 的变化曲线($SNR = 10\text{dB}$, $\delta_P = \pi/20$)

图7 不同用户数量条件下算法 BER 性能随 SIRs 的变化曲线($SIR = -10\text{dB}$, $\delta_P = \pi/20$)

结束语 本文针对单通道接受的、信道存在滤波效应的数字通信系统,提出了一种通过过采样处理后经适当变化并利用适定盲分离提取多用户信号的方法。在此基础上,以DS-CDMA系统作为立足点,针对其可能存在的窄带干扰,提出了一种基于三角函数分解的干扰近似方法,将其与前述的用于多用户检测的盲分离模型相结合,建立一个高维的多输入多输出的瞬时混合系统,并分别采用基于最小互信息和最大非高斯度量的方法实现用户信号以及干扰等效分量的分离,在同步实现多址干扰和码间串扰的消除同时,实现了外部窄带干扰的消除。

参考文献

- [1] Baynast A D, Fijalkow I. Blind source extraction in a multiuser context[C]// IEEE WCNC, Chicago USA: IEEE Press, 2000: 1172-1177
- [2] Leong W Y, Homer J. Blind multiuser receiver for DS-CDMA wireless system[J]. IEE Proc. -Commun, 2006, 153(5): 733-739
- [3] Lathauwer L D, Baynast A D. Blind deconvolution of DS-CDMA signals by means of decomposition in rank-(1, L, L) terms[J]. IEEE Trans. on Signal Processing, 2008, 56(4): 1562-1571

(下转第84页)

时,每个上传文件约 60M。运行几小时后,在浏览器上通过 50070 端口查看文件分布情况,如图 7 所示。

从图 7 中可看到,文件被均匀地上传到集群各个节点,集群运行状况良好。查看文件系统里上传成功的文件,没有出现数据丢包现象,存储过程成功完成。

3.3 数据查询性能分析

调节数据发送端的发送频率,产生文件大小和数量不同的几组数据。启动查询客户端,查询各组数据。将同样的数据放到 Oracle 系统(四核 CPU,10 内存,硬盘 438G)里查询并统计时间。为减少偶然因素的影响,每个数据都是经 3 次以上的测试所得的平均值^[11,12]。表 3 将两种情况进行对比。

表 3 查询时间对比情况

查询条数	Oracle(S)	Hadoop 集群(S)
5.4 万	0.8	10.4
9.7 万	5.3	12.5
31.7 万	14.3	20.2
63.5 万	15.6	25.5
127 万	20.4	29.3
253.8 万	56.3	55.1
507.7 万	113.5	85.1
1015.3 万	201.4	150.2

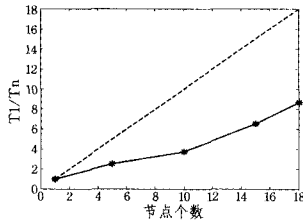


图 8 加速比性能测试结果

从表 3 中可以看出,当数据量低于 250 万条时,单机版系统有很大的优势;随着数据量增大,集群的查询优势越来越明显。这是由于与单机版相比,集群有额外的任务分配及调度时间,当查询数据量较少时,这个过程所用的时间是不可忽略的;当数据量非常庞大时,这部分所用的时间跟总体用时相比是非常微小的。

为测试查询算法的并行性,测试 100 万条数据集的加速比(speedup),实验结果如图 8 所示。从图中可看出,查询算法的加速比是接近线性的,考虑到 100 万条数据对于集群来说并不是很大的数据量,集群性能没有得到全面的体现,因此

图中曲线并没有靠近对角线。这再次说明平台更适合大文件大数据量的存储和查询。

结束语 针对智能电网下状态监测数据的海量、分布式等特点,设计并实现了智能电网状态监测数据的 Hadoop 存储系统。搭建具有 19 个普通 PC 机的 Hadoop 集群,对集群进行基准测试,表明集群适合处理大规模大文件的数据;应用 HDFS 设计实现了智能电网状态监测数据的分布式存储,文件均匀分布于集群各个节点,集群运行状况良好,没有出现数据丢失现象;应用 MapReduce 模型实现了状态监测数据的并行查询以及与单机关系数据查询性能的比较。随着数据量的增大,集群的查询优势更加明显,因此它可以有效地应用于智能电网信息平台建设中。

参考文献

- [1] 毕天姝,素梅智. 智能电网含义及其共性技术探讨[J]. 华北电力学报,2011,38(2):1-9
- [2] 张怀宇,朱松林,张杨,等. 输变电设备状态检修技术体系研究与实践[J]. 电网技术,2009,33(13):70-73
- [3] Newman A, Li Y F, Hunter J. Scalable Semantics-the Silver Lining of Cloud[C]// IEEE Forth International Conference on eScience, 2008
- [4] 于戈,谷峪,鲍玉斌,等. 云计算环境下的大规模图数据处理技术[J]. 计算机学报,2011,10(34):1753-1767
- [5] Gaggero M, Leo S, Mance S. Parallelizing bioinformatics applications with MapReduce[OL]. www.cca08.org/papers/Poster10-Simone-Leo.pdf
- [6] 潘巍,李战怀,伍赛,等. 基于消息传递机制的 MapReduce 图算法研究[J]. 计算机学报,2011,10(34):1768-1784
- [7] 王德文,宋亚奇,朱永利. 基于云计算的智能电网信息平台[J]. 电力系统自动化,2010,34(22):7-12
- [8] 吕跃刚,关晓慧,刘俊承. 风力发电机组状态监测系统研究[J]. 自动化与仪表,2012,27(1):6-10
- [9] White T. Hadoop 权威指南(第二版)[M]. 北京:清华大学出版社,2011:43-44
- [10] 赵卫中,马慧芳,傅燕翔,等. 基于云计算平台 Hadoop 的并行 k-means 聚类算法设计研究[J]. 计算机科学,2011,38(10):166-169
- [11] 刘鹏,黄宜华,陈卫卫. 实战 Hadoop[M]. 北京:电子工业出版社,2011:60-64
- [12] 陈勇. 基于 Hadoop 平台的通信数据分布式查询算法的设计与实现[D]. 北京:北京交通大学,2009
- [13] Friedman J H. Exploratory Projection Pursuit[J]. J. Amer. Statist. Assoc., 1987,82:249-266
- [4] Cheng H, Guo W, Yu J D. Multichannel convolution blind separation algorithm for MIMO DSSS/CDMA system[J]. IEEE Trans. on Circuits Systems Signal Processing, 2007,26(2):249-262
- [5] Zhang Yang, Kassam S A. Blind source separation and equalization in asynchronous MIMO DS-CDMA[C]// IEEE CISS. Princeton NJ: IEEE Press, 2008:316-319
- [6] Raju K. Blind Source Separation For Interference Cancellation In CDMA Systems[D]. Finland: Helsinki University of Technology, 2006
- [7] Huovinen T. Independent component analysis in DS-CDMA multiuser detection and interference cancellation[D]. Finland: Tampere University of Technology, 2008
- [8] Nion D, Lathauwer L D. A block component model-based blind DS-CDMA receiver[J]. IEEE Trans. on Signal Processing, 2008,56(11):5567-5579
- [9] Liu X, Xu Z Z, Lei L. Space-time blind multiuser detection for multiuser DS-CDMA and oversampled systems[R]. German: Springer Wireless Pers Commun, 2010
- [10] Fevotte C, Theis F J. Orthonormal approximate joint block-diagonalization[R]. Technical report GET/Telecom. Paris, 2007
- [11] Caiafa C F, Proto A N. Separation of statistically dependent sources using an L2-distance non-Gaussianity measure[J]. Signal Processing, 2006,86:3404-3420
- [12] Novey M, Adali T. Complex ICA by Negentropy Maximization[J]. IEEE Trans. on Neural Networks, 2008,19(4):596-609
- [13] Friedman J H. Exploratory Projection Pursuit[J]. J. Amer. Statist. Assoc., 1987,82:249-266