

基于 MCMC-UKF 的直接序列扩频信号盲估计

马 超, 张立民

(海军航空大学 信息融合研究所, 山东 烟台 264001)

摘 要: 针对长码直接序列扩频信号的扩频码及信息序列盲估计问题, 提出基于重叠分段 MCMC-UKF 的扩频码及信息序列联合估计算法。在贝叶斯框架模型下, 结合重叠分段的思想, 利用 UKF 算法求解非线性模型, 估计各参数后验概率的均值和方差, 通过 MCMC 方法迭代抽样得到各分段扩频序列, 进行序列拼接以完成对扩频序列及信息序列的估计。仿真结果表明, 该算法能适应较低的信噪比环境, 且不受扩频序列类型的限制。

关键词: 直接序列扩频信号; 贝叶斯模型; 无迹卡尔曼滤波; 分段; 序列估计

中文引用格式: 马超, 张立民. 基于 MCMC-UKF 的直接序列扩频信号盲估计[J]. 计算机工程, 2019, 45(7): 95-102.

英文引用格式: MA Chao, ZHANG Limin. Blind estimation of direct sequence spread spectrum signal based on MCMC-UKF[J]. Computer Engineering, 2019, 45(7): 95-102.

Blind Estimation of Direct Sequence Spread Spectrum Signal Based on MCMC-UKF

MA Chao, ZHANG Limin

(Institute of Information Fusion, Naval Aeronautical University, Yantai, Shandong 264001, China)

[Abstract] Aiming at the problem of spreading code and information sequence blind estimation for long code Direct Sequence Spread Spectrum (DSSS) signal, this paper proposes a joint estimation algorithm based on overlapping segment Markov Chain Monte Carlo-Unscented Kalman Filtering (MCMC-UKF) for spreading code and information sequence. Under the Bayesian framework model, combined with the idea of overlapping segmentation, the UKF algorithm is used to solve the nonlinear model, and the mean and variance of the posterior probabilities of each parameter are estimated. The segmentation spread spectrum is obtained by MCMC method, sequence splicing to complete the estimation of the spreading sequence and the information sequence. Simulation results show that the algorithm can adapt to lower Signal-to-Noise Ratio (SNR) and is not limited by the type of spread spectrum sequence.

[Key words] Direct Sequence Spread Spectrum (DSSS) signal; Bayesian model; Unscented Kalman Filtering (UKF); segmentation; estimation of sequence

DOI: 10.19678/j.issn.1000-3428.0050464

0 概述

在非合作直接序列扩频 (Direct Sequence Spread Spectrum, DSSS) 信号通信系统中, 接收者无法截获信号的先验知识, 因此, DSSS 信号伪随机码的估计对信息的获取是一个关键的、重要的技术。对于短码 DSSS 信号的 PN 码估计, 已经有较多的研究。长码 DSSS 信号一个扩频周期需调制多位信息符号, 截获长码 DSSS 信号的 PN 码的盲估计变得更加困难和具有挑战性。

2000 年, 文献[1]提出利用特征值分解 (EVD) 方法估计 DSS 信号的扩频码, 该方法已广泛地应用于工程上, 并有许多改进的 EVD 方法^[1-3], 其中的一

些方法已经应用于长码 DSSS 信号的 PN 码估计, 其扩频码长为 m 序列^[4-6]。文献[7]利用爬山算法实现对同步 DSSS 信号的扩频码估计, 通过将扩频波形估计问题视为一个低秩缺失数据矩阵近似问题。文献[8]提出一个适用于长码 DSSS 信号的扩频波形估计方法。文献[9]提出 2 种基于自相关和主分量分析的方法来估计 DSSS 信号的序列长度。三阶统计数据对缺少样本不敏感, 可以公开更多的信号信息, 因此广泛应用于信号检测和参数估计。文献[10]将三阶相关函数 (TCF) 应用于 m 序列估计。文献[11]提出三阶统计过程可以用于非合作环境下 DSSS 信号检测甚至确定码特征。文献[12]提出对于周期长码信号, 可以通过将其等效为周期时变信

基金项目: 国家自然科学基金重大研究计划 (91538201); 泰山学者专项 (ts201511020)。

作者简介: 马超 (1988—), 男, 博士研究生, 主研方向为通信侦察技术; 张立民, 教授、博士。

收稿日期: 2018-02-09 **修回日期:** 2018-05-24 **E-mail:** machao085736@163.com

道来解决。文献[12]提出一种新的方法解决信号方差矩阵。该方法首先对接收信号进行码片延迟相乘,然后分析相关矩阵,得到用于长伪随机码 DSSS 系统的伪随机码转换序列的估计序列。文献[13]提出对于周期长码直扩信号,可以等效为同步多用户短码直扩信号来解决。文献[14]提出将周期长码直扩信号等效为多用户短码直扩 CDMA 系统,该方法减弱了对信号类型的限制,并且可以获得非同步信号的扩频序列估计。文献[15]引入了一种由主分量分析实现自适应特征提取的在线无监督学习神经网络的方法,来实现直接序列码分多址(DS-CDMA)信号的伪码序列盲估计。文献[16]针对批处理方法在实现非等功率同步 DS-CDMA 信号伪码序列盲估计时,存在复杂度高、收敛速度慢的问题,引入了 Sanger NN、LEAP NN 和 APEX NN 3 种多主分量神经网络,实验结果表明,该方法具有较低的复杂度。但是,上述方法同样存在不足,即所处理的信号模型都是线性的,当信号模型为非线性时,上述方法均失效。

无迹卡尔曼滤波(UKF)方法在处理非线性问题时具有明显的优势,因此,本文针对长码直扩信号的扩频码及信息序列盲估计问题,结合重叠分段和 MCMC 算法,提出一种基于重叠分段 MCMC-UKF 的盲估计算法。

1 信号模型

持续时间为 N 个扩频周期的接收信号为:

$$r(t) = s(t) \times g(t) + w(t) = \sum_{i=0}^{N_0-1} b(i) q(t - iT_s - \tau) \times \sum_{m=0}^{N-1} h(t - mT - \tau) e^{j2\pi ft - \varphi} + w(t) \quad (1)$$

其中, $s(t)$ 表示持续时间为 N 个扩频周期的发射信号模型:

$$s(t) = A \sum_{i=0}^{N_0-1} b(i) q(t - iT_s) \sum_{m=0}^{N-1} h(t - mT) e^{j2\pi ft} \quad (2)$$

其中, f 为载波频率, N_0 为 N 个扩频周期长度包含的信息符号的数目。信号经过信道传播,其表达式为:

$$g(t) = \delta(t - \tau) e^{j2\pi ft} \quad (3)$$

其中, $\delta(\cdot)$ 为狄拉克函数, τ 为时延。

初相 $\varphi = 2\pi f\tau$ 通过载波同步^[16]估计出信号载频和初相,对接收信号 $r(t)$ 进行下变频,得到等效基带信号为:

$$r(t) = \sum_{i=0}^{N_0-1} b(i) \times \sum_{l=1}^L \tilde{\beta}_l q(t - iT_s - \tau_l) \times \sum_{m=0}^{N-1} h(t - mT - \tau_l) + \tilde{w}(t) \quad (4)$$

其中, $\tilde{w}(t) = w(t) e^{-j2\pi ft + \varphi}$ 为均值为 0、方差为 σ^2 的加性复高斯白噪声。以码片速率对接收信号进行采样,定义 $r(k) \triangleq r(kT_c)$, 则可得接收信号的离散形式为:

$$r(k) = \sum_{i=0}^{N_0-1} b(i) q\left(k - iP - \frac{\tau}{T_c}\right) \times \sum_{m=0}^{N-1} h\left(k - mR - \frac{\tau}{T_c}\right) + \tilde{w}(k) \quad (5)$$

将式(5)以扩频周期进行分段,得到的分段观测数据如图 1 所示。

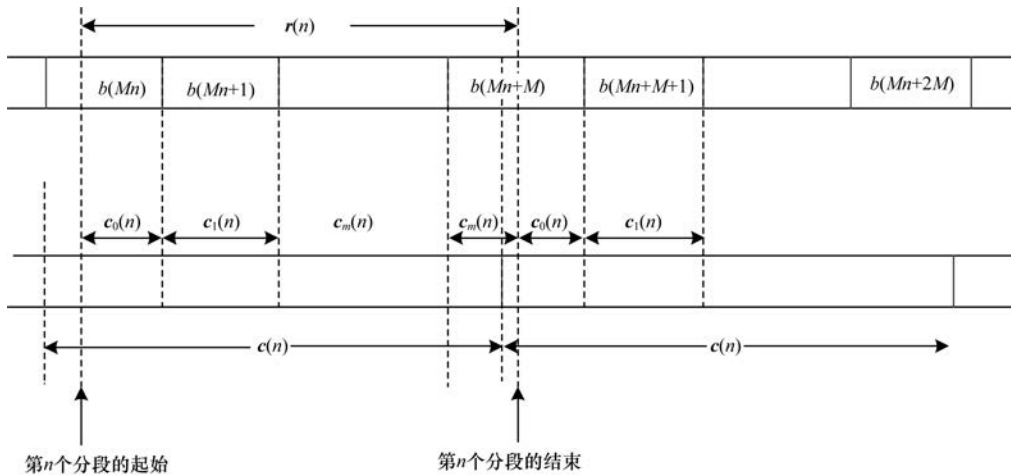


图 1 观测数据的分段示意图

得到的第 n 个分段信号表示为:

$$r(n) = [b_0(n)c_0(n) + b_1(n)c_1(n) + \cdots + b_M(n)c_M(n)] + w(n), 1 \leq n \leq N \quad (6)$$

其中, $M = R/P$ 为扩频比, $r(n) \triangleq [r(nR+1), r(nR+2), \dots, r(nR+R)]$, $b_m(n) \in \mathbb{C}^1$, $c_m(n) \in \mathbb{R}^{1 \times R}$, $0 \leq m \leq M$,

且有:

$$\begin{cases} b_0(n) = b(Mn) \\ b_1(n) = b(Mn+1) \\ \vdots \\ b_M(n) = b(Mn+M) \end{cases} \quad (7)$$

$$\begin{aligned}
c_0(n) &= [c(\tau/T_c + 1) \ c(\tau/T_c + 2) \cdots c(P) \ 0 \ \cdots \ 0]^T \\
c_M(n) &= [0 \ \cdots \ 0c(1) \ c(2) \cdots c(\tau/T_c)]^T \\
c_m(n) &= [0 \ \cdots \ 0c(1) \ c(2) \cdots c(P) \ 0 \ \cdots \ 0]^T \quad (8)
\end{aligned}$$

将式(6)进一步写成下面的向量形式:

$$r(n) = b(n)c(n) + w(n), 1 \leq n \leq N \quad (9)$$

其中, $b(n) = [b_0(n) \ b_1(n) \ \cdots \ b_M(n)]$, $c(n) = [c_0(n) \ c_1(n) \ \cdots \ c_M(n)]^T$ 。结合图1并分析式(6)~式(9)可以发现, 在 n 取不同值时, 各时间窗内第一个符号起始时刻相对于该时间窗起始时刻的偏移不相同, 从而导致 $c(n)$ 为一个时变的向量, 因此无

法通过建立统一的数学模型, 采用 MCMC 方法对扩频码和信息序列进行联合估计。

针对上述问题, 下文将一个完整的扩频周期分成小的片段, 使扩频码在每个片段的持续时间内对应一个符号, 从而使 $c(n)$ 保持恒定, 这样就可以通过 MCMC 方法估计出每个片段的扩频码及信息序列。对于长码直扩信号, 为进一步提高算法的估计性能, 避免每一段估值运算时产生的扩频码极性不一致问题, 将一个扩频周期向量分成 S 个的相互重叠。图2所示为重叠长度为 H_0 、长度为 H 的信号片段, 其中, 信号片段的长度 H 远小于扩频增益 P 。

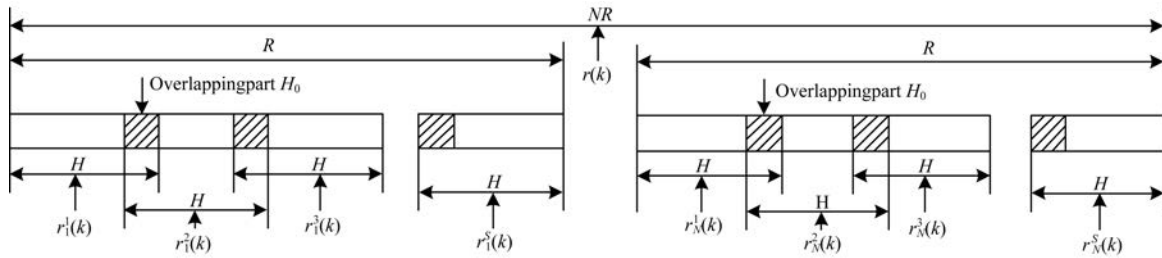


图2 长码 DSSS 信号重叠分段示意图

则式(5)可重新表示成下式的形式:

$$R_s(k) = A(b(s))D(c(s), \tau) + W(k) \quad (10)$$

即:

$$\begin{aligned}
\begin{bmatrix} r_1^s(k) \\ r_2^s(k) \\ \vdots \\ r_N^s(k) \end{bmatrix} &= \begin{bmatrix} b(i_1^s) \\ b(i_2^s) \\ \vdots \\ b(i_N^s) \end{bmatrix} \begin{bmatrix} h_1^s(k) \\ h_2^s(k) \\ \vdots \\ h_N^s(k) \end{bmatrix} + \begin{bmatrix} w_1^s(k) \\ w_2^s(k) \\ \vdots \\ w_N^s(k) \end{bmatrix}, \\
1 \leq k \leq H, 1 \leq s \leq S & \quad (11)
\end{aligned}$$

其中:

$$r_n^s(k) \triangleq \{r[(n-1)R + (s-1)(H-H_0) + k]\}_{1 \leq n \leq N}^{1 \times H}$$

$$b(s) = [b(i_1^s), b(i_2^s), \dots, b(i_N^s)]^T$$

$$i_n^s = \left\lfloor \frac{(n-1)R + (s-1)(H-H_0)}{P} \right\rfloor, 1 \leq n \leq N$$

$$c(s) = \{c[(s-1)(H-H_0) + k]\}_{1 \leq k \leq H}$$

$$h^s(k) \triangleq \left\{h\left[(s-1)(H-H_0) + k - \frac{\tau}{T_c}\right]\right\}_{1 \leq k \leq H}$$

本文研究目标是在知道每一小段数据 R_s 的前提下, 采用 MCMC-UKF 方法对参数空间 $\Theta \triangleq \{b(s), c(s), \tau, \sigma^2\}$ 进行联合估计, 重复进行 S 次独立的运算之后, 进而用所有的样本采样值对扩频码及信息序列进行拼接和恢复。通过分析参数空间 Θ 中的元素可以发现, 对于参数 σ^2 , 可采用 UKF 方法进行估计, 而 $b(s)$ 、 $c(s)$ 和 τ 的全条件后验分布未知且不易抽取, 因此, 采用混合 MH 算法进行抽取。

2 算法原理

2.1 无迹卡尔曼滤波

无迹卡尔曼滤波(UKF)是一种递归式贝叶斯估计方法, 但是无须进行非线性模型的求解(即不需要求解雅可比矩阵), 其基本思想是利用 UT(Unscented Transformation)转换, 用一组确定的样本点近似求解测量条件下系统状态的后验概率密度 $p(x_k | z_k)$ 的均值和方差, 实现系统状态递推均值和方差(一阶矩、二阶矩)的估计。

设具有如下状态空间模型的非线性系统状态方程和量测方程分别为:

$$x_{k+1} = f(x_k) + n_k \quad (12)$$

$$z_{k+1} = h(x_{k+1}) + v_{k+1} \quad (13)$$

其中, n_k 是协方差为 Q_k 的系统噪声, v_{k+1} 是协方差为 R_{k+1} 的测量噪声。基于 UT 变换的 UKF 算法基本过程如下:

1) 初始化 $\hat{x}_0 = E(x_0)$, $P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]$;

2) 确定样本点 Sigma 集合 $s_{i(k)}$ 及相应的权值 $w_{i(k)}$, $i = 0, 1, \dots, 2n$;

3) 根据系统状态方程求取样点传递值:

$$x_{s_{i(k+1),k}} = f(s_{i(k)}) \quad (14)$$

4) 系统状态均值和方差的一步预测, 为:

$$\hat{x}_{k+1,k} = \sum_{i=0}^{2n} w_i^m x_{s_{i(k+1),k}} \quad (15)$$

$$\mathbf{P}_{xx(k+1,k)} = \mathbf{Q}_{k+1} + \sum_{i=0}^{2n} \mathbf{w}_i^c (\mathbf{x}_{s_{i(k+1,k)}} - \hat{\mathbf{x}}_{k+1,k}) \times (\mathbf{x}_{s_{i(k+1,k)}} - \hat{\mathbf{x}}_{k+1,k})^T \quad (16)$$

5) 根据系统量测方程进一步求取状态一步预测的传递值:

$$\mathbf{z}_{s_{i(k+1,k)}} = \mathbf{h}(\mathbf{x}_{s_{i(k+1,k)}}) \quad (17)$$

6) 预测量均值和协方差:

$$\hat{\mathbf{z}}_{k+1,k} = \sum_{i=0}^{2n} \mathbf{w}_i^m \mathbf{z}(s_{i(k+1,k)}) \quad (18)$$

$$\mathbf{P}_{zz} = \mathbf{R}_{k+1} + \sum_{i=0}^{2n} \mathbf{w}_i (\mathbf{z}(s_{i(k+1,k)}) - \hat{\mathbf{z}}_{k+1,k}) \times (\mathbf{z}(s_{i(k+1,k)}) - \hat{\mathbf{z}}_{k+1,k})^T \quad (19)$$

$$\mathbf{P}_{xz} = \sum_{i=0}^{2n} \mathbf{w}_i^c (\mathbf{x}(s_{i(k+1,k)}) - \hat{\mathbf{x}}_{k+1,k}) \times (\mathbf{z}(s_{i(k+1,k)}) - \hat{\mathbf{z}}_{k+1,k})^T \quad (20)$$

其中, \mathbf{P}_{zz} 是量测方差矩阵, \mathbf{P}_{xz} 是状态向量与量测向量的协方差矩阵。

7) 计算 UKF 增益, 更新状态向量与方差, 即:

$$\mathbf{K}_{k+1} = \mathbf{P}_{xz} \mathbf{P}_{zz}^{-1} \quad (21)$$

$$\hat{\mathbf{x}}_{k+1,k+1} = \hat{\mathbf{x}}_{k+1,k} + \mathbf{K}_{k+1} (\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1,k}) \quad (22)$$

$$\mathbf{P}_{xx(k+1,k+1)} = \mathbf{P}_{xx(k+1,k)} - \mathbf{K}_{k+1} \mathbf{P}_{zz} \mathbf{K}_{k+1}^T \quad (23)$$

2.2 MCMC 算法

MCMC 算法的基本思想是构造一条 Markov 链, 使其平稳分布为待估参数的后验分布。利用这条链上的各个样本值就可以估计参数, 进行各种统计推断, 该算法包括 2 个部分:

1) Monte Carlo 部分利用平均近似一个期望, 如需计算期望 $E[f(\theta)] = \int f(\theta) p(\theta|X) d\theta$, 则一个好的近似为 $E[f(\theta)] \approx \frac{1}{I} \sum_{i=0}^{I-1} f(\theta^{(i)})$, 当 $I \rightarrow \infty$ 时一定收敛于真实期望, 即取等号, 这里 $\theta^{(i)}$ 是依据 $p(\theta|X)$ 产生的一系列样本;

2) Markov 部分产生具有不变分布 $p(\theta|X)$ (一般用 $\pi(\theta)$ 表示不变分布) 的 Markov 链 $\{\theta^{(0)}, \theta^{(1)}, \dots\}$, 一个 Markov 链由 $\theta^{(0)}$ 初始分布和所谓的转移核 $l(\theta^{(i)}, d\theta^{(i+1)})$ 完全定义。经典的 MCMC 方法有 Gibbs 采样和 MH 算法, 它们定义的转移核使得无论 $\theta^{(0)}$ 是何种初始分布, 都一定收敛于代价分布 $p(\theta|X)$ 。

下面给出 2 种算法的具体实现步骤。

1) Gibbs 采样

参数空间 $\theta = \{\theta_1, \theta_2, \dots, \theta_d\}$, 为了从后验分布 $p(\theta|X)$ 中抽取样本值, 利用前一时刻的抽样值, 对各参数的全条件后验分布进行迭代抽取, 得到后一时刻各参数的抽样值。假设在 $n-1$ 次迭代时刻的空间值为 $\theta^{(n-1)} = \{\theta_1^{(n-1)}, \theta_2^{(n-1)}, \dots, \theta_d^{(n-1)}\}$, 于是

n 时刻的空间值通过以下抽取步骤得到:

步骤 1 从 $p(\theta_1|\theta_2^{(n-1)}, \theta_3^{(n-1)}, \dots, \theta_d^{(n-1)}, X)$ 中抽取 $\theta_1^{(n)}$ 。

步骤 2 利用更新值 $\theta_1^{(n)}$, 从 $p(\theta_2|\theta_1^{(n)}, \theta_3^{(n-1)}, \dots, \theta_d^{(n-1)}, X)$ 中抽取 $\theta_2^{(n)}$ 。

步骤 3 利用更新值 $\theta_1^{(n)}, \theta_2^{(n)}$, 从 $p(\theta_3|\theta_1^{(n)}, \theta_2^{(n)}, \theta_4^{(n-1)}, \dots, \theta_d^{(n-1)}, X)$ 中抽取 $\theta_3^{(n)}$ 。

步骤 4 依此类推, 最后利用 $\theta_1^{(n)}, \theta_2^{(n)}, \dots, \theta_{d-1}^{(n)}$, 从 $p(\theta_d|\theta_1^{(n)}, \theta_2^{(n)}, \dots, \theta_{d-1}^{(n)}, X)$ 中抽取 $\theta_d^{(n)}$ 。

2) MH 算法

MH 算法是 Gibbs 采样的扩展, 它主要是针对那些全条件后验分布概率不易抽取的参数, 其具体步骤如下:

步骤 1 根据先验分布选择初始值 $\theta^{(0)}$ 。

步骤 2 在第 n 次迭代时刻, 从建议分布 $q(\theta^*|\theta^{(n-1)})$ 中抽取备择采样值 θ^* 。

步骤 3 计算接收概率:

$$r = \frac{p(\theta^*|X)q(\theta^{(n-1)}|\theta^*)}{p(\theta^{(n-1)}|X)q(\theta^*|\theta^{(n-1)})} \quad (24)$$

步骤 4 以接收概率 $\min(1, r)$ 接收 θ^* 作为 $\theta^{(n)}$ 的更新值, 否则 $\theta^{(n)} = \theta^{(n-1)}$ 。

步骤 5 重复步骤 2 ~ 步骤 4 进行多次迭代, 从而完成对 $p(\theta|X)$ 的抽取。

2.3 贝叶斯模型推导

由第 1 节的信号模型可知, 时延等参数的状态空间模型和观测方程分别为:

$$\tau_n = \tau_{n-1} \quad (25)$$

$$\mathbf{R}_s = \mathbf{A}(\mathbf{b}(s)) \mathbf{D}(\mathbf{c}(s), \tau) + \mathbf{W}(s) \quad (26)$$

1) 参数的先验假设

先验分布的选取一般基于 2 个原则: 简化计算与对后验分布的影响最小 (即无信息先验分布)。在简化计算上一般采用共轭先验分布, 它能保证参数的后验分布与先验分布相同。

由贝叶斯公式可知, 参数空间的后验分布函数等价于:

$$p(\Theta|\mathbf{R}_s(k)) = p(\Theta, \mathbf{R}_s(k))/p(\mathbf{R}_s(k)) \propto p(\mathbf{R}_s(k)|\Theta)p(\Theta) \quad (27)$$

其中, $p(\mathbf{R}_s(k)|\Theta)$ 为似然函数, $p(\Theta)$ 为联合先验概率密度函数。根据式 (10) 可得:

$$p(\mathbf{R}_s(k)|\Theta) = (2\pi\sigma^2)^{-NH/2} \times \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N \|r_{i,:}^s - \mathbf{A}(\mathbf{b}(s))_{i,:} \mathbf{D}(\mathbf{c}(s), \tau)\|^2\right) \quad (28)$$

其中, $(\cdot)_{i,:}$ 表示相应矩阵的第 i 行, $\mathbf{R}_{i,:}^s(k)$ 为矩阵 $\mathbf{R}_s(k)$ 的第 i 行。

$$p(\Theta) = p(\mathbf{b}(s), \mathbf{c}(s), \tau, \sigma^2) = p(\mathbf{b}(s) | \mathbf{c}(s), \tau, \sigma^2) p(\mathbf{c}(s) | \tau, \sigma^2) p(\tau | \sigma^2) p(\sigma^2) = p(\mathbf{b}(s)) p(\mathbf{c}(s)) p(\tau) p(\sigma^2) \quad (29)$$

由于噪声方差 σ^2 的先验分布可认为服从逆伽玛分布 $Ig(\nu_0, \gamma_0)$, 因此概率密度函数表达式为:

$$p(\sigma^2) = \frac{(\gamma_0)^{\nu_0}}{\Gamma(\nu_0)} (\sigma^2)^{-\nu_0-1} \exp\left(-\frac{\gamma_0}{\sigma^2}\right) \quad (30)$$

本文选取 $\nu_0 = \gamma_0 = 0$, 逆伽玛分布则成为 Jeffrey 无信息先验分布 $p(\sigma^2) \propto 1/\sigma^2$ 。

扩频码向量 $\mathbf{c}(s)$ 的每个元素相互独立, 可认为服从 $\mathbb{S} \triangleq \{-1, +1\}$ 上的离散均匀分布 $DU(\mathbb{S})$, 则其概率密度函数为:

$$p(\mathbf{c}(s)) = (1/2)^R \quad (31)$$

信息序列向量 $\mathbf{b}(s)$ 中的元素相互独立, 可认为服从离散有限字符集 \mathbb{C} 上的均匀分布, 若采用 BPSK 调制, 则 $\mathbb{C} \triangleq \{-1, +1\}$ 。本文中假设调制样式已知, 则其概率密度函数为:

$$p(\mathbf{b}(s)) = \left(\frac{1}{2}\right)^N \quad (32)$$

时延 τ 服从 $[0, T_s]$ 内的均匀分布 $U[0, T_s]$, 则其概率密度函数为:

$$p(\tau) = \frac{1}{T_s} \quad (33)$$

根据上述先验分布, 式(29)可写成:

$$p(\Theta) = p(\mathbf{b}(s)) p(\mathbf{c}(s)) p(\tau) p(\sigma^2) \propto \left(\frac{1}{T_s}\right) \left(\frac{1}{2}\right)^N \left(\frac{1}{2}\right)^R \times \frac{(\gamma_0)^{\nu_0}}{\Gamma(\nu_0)} (\sigma^2)^{-\nu_0-1} \exp\left(-\frac{\gamma_0}{\sigma^2}\right) \quad (34)$$

由此, 可得:

$$p(\Theta | R_s) \propto p(R_s | \Theta) p(\Theta) = (2\pi\sigma^2)^{-NH/2} \times \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2\right) \times \left(\frac{1}{T_s}\right) \left(\frac{1}{2}\right)^N \left(\frac{1}{2}\right)^R \times \frac{(\gamma_0)^{\nu_0}}{\Gamma(\nu_0)} (\sigma^2)^{-\nu_0-1} \exp\left(-\frac{\gamma_0}{\sigma^2}\right) \quad (35)$$

结合式(28)并对 σ^2 积分, 可得:

$$p(\mathbf{b}(s), \mathbf{c}(s), \tau | R_s) \propto \left(\frac{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2}{2}\right)^{\left(-\frac{NH+2\nu_0}{2}\right)} \times DU(\mathbb{C}) DU(\mathbb{S}) U[0, T_s] \quad (36)$$

下文给出各参数的全条件后验分布。

1) σ^2 的全条件后验分布

根据式(26)可知, 在第 n 个观测周期, σ^2 服从逆伽玛分布, 为了方便公式推导, 令 $\beta = \sigma^2$, 即:

$$p(\hat{\beta}_{n+1}^s | \mathbf{b}(s), \mathbf{c}(s), \tau_{n+1}, \Delta f_{n+1}, \sigma_{n+1}^2, R_{n+1}^s) \sim N_c(\hat{\beta}_{n+1}^s, P_{\beta\beta(n+1)}) \quad (37)$$

其中, 均值 $\hat{\beta}_{n+1}^s$ 和方差 $P_{\beta\beta(n+1)}$ 通过无迹卡尔曼滤波方法, 在给出 $\hat{\beta}_{n+1}^s$ 和 $P_{\beta\beta(n)}$ 时进行迭代更新, 其过程如下:

初始化:

$$\hat{\beta}_n^s = E(\beta_n^s) \quad (38)$$

$$P_{\beta\beta(n)} = E[(\beta_n^s - \hat{\beta}_n^s)(\beta_n^s - \hat{\beta}_n^s)^T] \quad (39)$$

根据 UKF 算法, 确定第 n 个周期的样本点 Sigma 集合及相应的权值:

$$\begin{cases} s_{0(n)}^s = \hat{\beta}_n^s \\ s_{1(n)}^s = \hat{\beta}_n^s + \left(\sqrt{(L+\lambda)P_{\beta\beta}}\right)_i, i=1, 2, \dots, L \\ s_{i(n)}^s = \hat{\beta}_n^s - \left(\sqrt{(L+\lambda)P_{\beta\beta}}\right)_{i-L}, i=L+1, \dots, 2L \end{cases} \quad (40)$$

$$\begin{cases} w_0^m = \lambda / (L + \lambda), w_0^c = \lambda / (L + \lambda) + (1 - \alpha^2 + \beta) \\ w_i^m = w_i^c = \lambda / [2(L + \lambda)], i=1, 2, \dots, 2L \end{cases} \quad (41)$$

$$\beta_{s_{i(n+1,n)}^s}^s = F s_{i(n)}^s \quad (42)$$

$$\hat{\beta}_{n+1,n}^s = \sum_{i=0}^{2L} w_i^m \beta_{s_{i(n+1,n)}^s}^s \quad (43)$$

$$P_{\beta\beta(n+1,n)} = \Sigma_V + \sum_{i=0}^{2L} w_i^c (\beta_{s_{i(n+1,n)}^s}^s - \hat{\beta}_{n+1,n}^s) \times (\beta_{s_{i(n+1,n)}^s}^s - \hat{\beta}_{n+1,n}^s)^T \quad (44)$$

$$R_{s_{i(n+1,n)}^s}^s = A(\mathbf{b}(s)) D(\mathbf{c}(s), \tau_n) \quad (45)$$

$$\hat{R}_{n+1,n}^s = \sum_{i=0}^{2L} w_i^m R(s_{i(n+1,n)}^s) \quad (46)$$

$$P_{RR} = \Sigma_W + \sum_{i=0}^{2L} w_i (R(s_{i(n+1,n)}^s) - \hat{R}_{n+1,n}^s) \times (R(s_{i(n+1,n)}^s) - \hat{R}_{n+1,n}^s)^T \quad (47)$$

$$P_{\beta R} = \sum_{i=0}^{2L} w_i^c (\beta(s_{i(n+1,n)}^s) - \hat{\beta}_{n+1,n}^s) \times (R(s_{i(n+1,n)}^s) - \hat{R}_{n+1,n}^s)^T \quad (48)$$

$$K_{n+1} = P_{\beta R} P_{RR}^{-1} \quad (49)$$

$$\hat{\beta}_{n+1,n+1}^s = \hat{\beta}_{n+1,n}^s + K_{n+1} (R_{n+1}^s - \hat{R}_{n+1,n}^s) \quad (50)$$

$$P_{\beta\beta(n+1,n+1)} = P_{\beta\beta(n+1,n)} - K_{n+1} P_{RR} K_{n+1}^T \quad (51)$$

其中, L 为多样本点数, $\hat{\beta}_{n+1,n}^s$ 为状态预测均值, $P_{\beta\beta(n+1,n)}$ 为状态预测值的协方差矩阵, $\hat{R}_{n+1,n}^s$ 为预测的观测均值, P_{RR} 是观测方差矩阵, $P_{\beta R}$ 是状态向量与观测向量的协方差矩阵, K_{n+1} 为卡尔曼增益。由此,

可得 $\tilde{\beta}_{n+1}^s$ 的全条件后验分布为:

$$p(\tilde{\beta}_{n+1}^s | \mathbf{b}(s), \mathbf{c}(s), \tau, R_s) \sim \text{Ig}\left(\nu_0 + \frac{NH}{2}, \frac{1}{2} \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2 + \gamma_0\right)$$

即 σ^2 的全条件后验分布为:

$$p(\sigma^2 | \mathbf{b}(s), \mathbf{c}(s), \tau, R_s) \sim \text{Ig}\left(\nu_0 + \frac{NH}{2}, \frac{1}{2} \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2 + \gamma_0\right) \quad (52)$$

2) $\mathbf{b}(s)$ 、 $\mathbf{c}(s)$ 和 τ 的全条件后验分布

由式(36)可知, $\mathbf{b}(s)$ 、 $\mathbf{c}(s)$ 和 τ 的全条件后验分布表达式均为:

$$\left(\frac{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2}{2} \right)^{\left(\frac{NH+2\nu_0}{2} \right)} \quad (53)$$

$$A(\mathbf{c}_n(s), \mathbf{c}_n^*(s)) = \min \left\{ 1, \left(\frac{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2}{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}_n^*(s), \tau)\|^2} \right)^{\left(\frac{NH+2\nu_0}{2} \right)} \right\} \quad (54)$$

$$A(\mathbf{b}(i_n^s), \mathbf{b}^*(i_n^s)) = \min \left\{ 1, \left(\frac{2\gamma_0 + \|R_{i,:}^s - A(\mathbf{b}(i_n^s))_{i,:} D(\mathbf{c}(s), \tau)\|^2 + \Sigma}{2\gamma_0 + \|R_{i,:}^s - A(\mathbf{b}^*(i_n^s))_{i,:} D(\mathbf{c}(s), \tau)\|^2 + \Sigma} \right)^{\left(\frac{NH+2\nu_0}{2} \right)} \right\} \quad (55)$$

$$A(\tau, \tau^*) = \min \left\{ 1, \left(\frac{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2}{2\gamma_0 + \sum_{i=1}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau^*)\|^2} \right)^{\left(\frac{NH+2\nu_0}{2} \right)} \right\} \quad (56)$$

其中, $\Sigma = \sum_{l=1, l \neq m}^N \|R_{i,:}^s - A(\mathbf{b}(s))_{i,:} D(\mathbf{c}(s), \tau)\|^2$

为常数, $\mathbf{c}^*(s)$ 、 $\mathbf{b}^*(s)$ 和 τ^* 与更新前的 $\mathbf{c}(s)$ 、 $\mathbf{b}(s)$ 和 τ 完全相似, 只是用 $\mathbf{c}_k^*(s)$ 代替扩频码向量 $\mathbf{c}(s)$ 的第 k 个值 $\mathbf{c}_k(s)$, $\mathbf{b}_n^*(s)$ 代替信息序列向量 $\mathbf{b}(s)$ 的第 n 个值 $\mathbf{b}(i_n^s)$, τ^* 代替时延 τ 。

2.4 序列拼接算法

由于算法是在对 S 个分段数据做独立抽取后获得的结果, 因此需要将获得的 $\mathbf{b}(s)$ 和 $\mathbf{c}(s)$ 进行拼接和重构 ($1 \leq s \leq S$), 得到完整的扩频码及信息序列。针对不同的片段, 在算法的执行过程中, 首先根据估计到的时延 τ 对基带信号进行码片同步, 得到用户的扩频向量 $\mathbf{c}(s)$ 。在扩频码的拼接过程中, 利用片段间重叠部分的相似性来提高序列的估计性能。

对于信息序列 $\mathbf{b}(s)$, 在 s 为不同值时, 由于在一个完整的扩频周期内包含多个信息序列, 致使每个向量对应的信息符号不是完整的信息序列。可以通过对抽取得到的 $\mathbf{b}(s)$ 按照下标进行排序, 得到完整的信息序列。在算法的实现过程中, 同样可以对重复抽取得到的样本值做均值处理来提高估计性能。

由上文的推导过程可以看出, 本节的算法对于短码和长码直扩信号都能适应, 其具体实现步骤

在计算后验分布时内部的已知条件取值不同。

下文给出采用混合 MH 算法抽样 $\mathbf{b}(s)$ 、 $\mathbf{c}(s)$ 和 τ 的过程, 式(53)已经给出扩频码向量 $\mathbf{c}(s)$ 、信息序列向量 $\mathbf{b}(s)$ 以及时延 τ 中元素的全条件后验概率分布, 并将其分别用作所需抽样参数的不变分布。针对扩频码向量 $\mathbf{c}(s)$, 首先取 $[0, 1]$ 之间的均匀分布值 μ , 如果 $\mu < \lambda$ (λ 为仿真参数, 文中取 $\lambda = 0.5$), 则建议概率分布 $q_1(\mathbf{c}^*(s) | \mathbf{c}(s)) \sim DU(\mathbb{S})$ 随机更新扩频码片, 其中 $\mathbb{S} \triangleq \{-1, +1\}$, 否则以 $q_2(\mathbf{c}^*(s) | \mathbf{c}(s)) \sim N(\mathbf{c}(s), \sigma_s^2)$ 更新扩频码片; 同理对于信息向量 $\mathbf{b}(s)$ 分别取建议概率分布 $q_1(\mathbf{b}^*(s) | \mathbf{b}(s)) \sim DU(\mathbb{C})$ 和 $q_2(\mathbf{b}^*(s) | \mathbf{b}(s)) \sim N(\mathbf{b}(s), \sigma_s^2)$; 时延 τ 的建议分布取 $q_1(\tau^* | \tau) \sim U[0, T_s]$ 和 $q_2(\tau_s^* | \tau_s) \sim N(\tau_s, \sigma_s^2)$, 且 3 种建议概率分布的接受概率分别为:

如下:

步骤 1 根据已知的扩频周期, 将接收到的基带信号按照图 2 所示进行重叠分段, 得到 S 个分段数据 R_s ($1 \leq s \leq S$)。

步骤 2 初始化数据。根据各参数的先验分布, 得到 $\Theta \triangleq \{\mathbf{b}(s), \mathbf{c}(s), \tau, \sigma^2\}$ 的初始值 $\Theta^{(0)}$, 设置 σ^2 的初始方差, 并令迭代变量 $n = 1$ 。

步骤 3 针对每一个片段数据, 在第 n 次迭代过程中, 首先利用混合 MH 算法对 $\mathbf{b}(s)$ 、 $\mathbf{c}(s)$ 和 τ 进行抽样更新; 然后采用 UKF 方法对 σ^2 进行估计。

步骤 4 第 n 次迭代结束, 进入 $n+1$ 次迭代, 重复步骤 3, 直至完成所有迭代。

步骤 5 重复步骤 3 ~ 步骤 4 直到获得所有的序列片段。

最后根据给出的序列拼接方法, 估计出完整的扩频码和信息序列。

3 仿真实验结果与分析

本文算法是在长码直扩信号条件下推导出的, 但由分段模型和推导过程可知, 该算法对短码的直扩信号同样适用。下文通过 2 个仿真实验分别对短码和长码直扩信号的参数估计性能进行仿真验证。

算法共进行500次迭代运算,由于算法收敛需要一个过程,为了增加仿真结果的可信度和准确性,取后400次的迭代结果累计进行相应的运算。

仿真实验1 对短码直扩信号的参数估计性能验证

仿真条件:扩频码为 $R=63$ 位随机序列,调制样式为BPSK,码片速率10 MHz,符号速率为 $10\text{ MHz}/63=158.7\text{ kHz}$,扩频增益 $P=63$,仿真数据为 $N=300$ 个扩频周期,分段长度 $H=8$,重叠长度 $H_0=H-1$,则每个扩频周期分为 $S=56$ 个片段,时延 τ 从 $[0, T_s]$ 上随机选取,并以码片速率进行采样,噪声为零均值加性复高斯白噪声,输入信噪比变化范围为 -15 dB 到 15 dB 。

图3直接给出SNR为 -9 dB 时,由MCMC算法所得的扩频码。由图3可以看出,该算法能在低信噪比条件下较好地适应短码直扩信号。

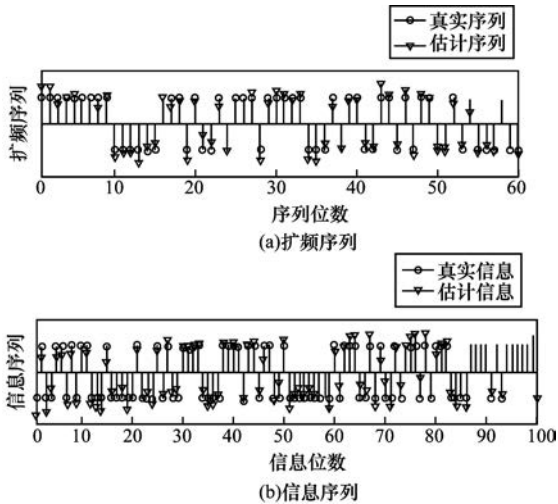


图3 扩频序列和信息序列估计结果

图4给出在不同分段长度 H 和不同周期数目 N 时,算法得出的扩频码及信息序列估计性能随输入信噪比变化的曲线。参数设置为: N 分别取100、200、300;分段长度 H 分别取8、16,且重叠部分均为 $H_0=H-1$ 。从图4可以看出,对于短码直扩信号的扩频码及信息序列盲估计,重叠分段的MCMC-UKF算法利用300个扩频周期的数据和8个码片周期的分段长度,就能在 -12 dB 的信噪比条件下,使得估计扩频码与真实扩频码的相似度超过0.95。该算法性能与使用的周期数目 N 成正比,在具有足够数据窗的情况下,即使具有更低的信噪比,该算法依然适应,但随之而来的是成倍增加的计算量。另外,由图4还可以看出,随着分段长度的减少,算法性能会随之提高,但算法的计算量也同样会随着增加。因此,在实际的应用中,需要考虑性能与计算复杂度的折中。

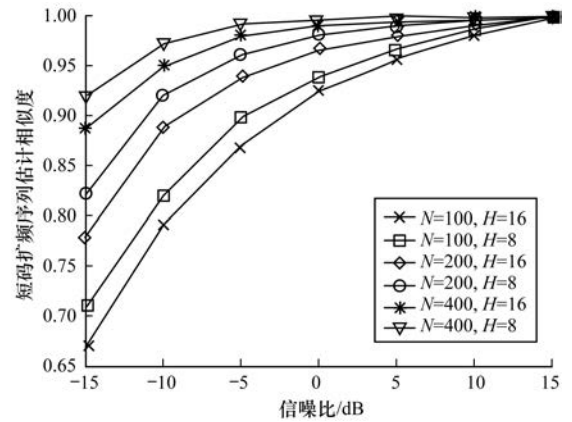


图4 短码直扩信号的估计性能随输入信噪比变化曲线

仿真实验2 对长码直扩信号的参数估计性能验证

仿真条件:扩频码为 $R=63$ 位随机序列,调制样式为BPSK,码片速率10 MHz,符号速率为 $10\text{ MHz}/30=333.3\text{ kHz}$,扩频增益 $P=30$,仿真数据为 $N=300$ 个扩频周期,分段长度 $H=8$,重叠长度 $H_0=H-1$,则每个扩频周期分为 $S=56$ 个片段,其他参数与仿真实验1相同。

图5给出了SNR为 -9 dB 时,采用后400次迭代值进行的时延参数 τ 的估计,并以直方图的形式表示。从图5可以看出,时延估计在真值的位置上具有最大的后验概率,因此本文算法能够较好地在低信噪比条件下对多径时延参数进行估计,仿真时 τ 的真值为 $\{0.6T_s\}$ 。

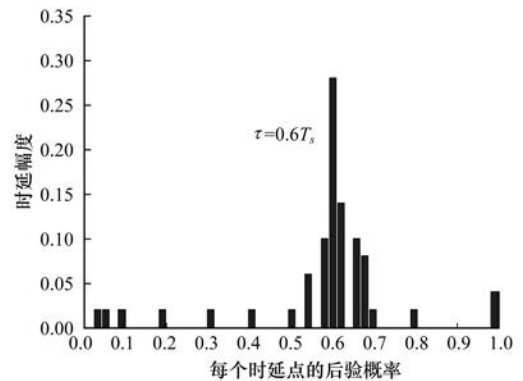


图5 时延的后验分布直方图

图6给出在不同分段长度 H 和不同周期数目 N 时,长码信号扩频码及信息序列的估计性能随输入信噪比变化的曲线。参数设置为: N 分别取100、200、300;分段长度 H 分别取8、16,且重叠部分均为 $H_0=H-1$ 。由图6可以看出,对于长码直扩信号,信息符号在每段持续时间 H 内都有一定发生跳变的可能,这样就会形成干扰,所以当参数条件相同时,长码信号的扩频码及信息序列估计性能明显低于短

码信号。随着分段长度的增加,信息符号发生跳变的概率也随之增大,扩频码及信息序列的估计性能就会相应降低。利用 300 个扩频周期的数据和 8 个码片周期的分段长度,只能在 -10 dB 的信噪比条件下使得估计扩频码与真实扩频码的相似度超过 0.95。因此,针对长码直扩信号中的信息调制影响,如要得到与短码相当的估计性能,就需要更高的信噪比或更多的时间窗。

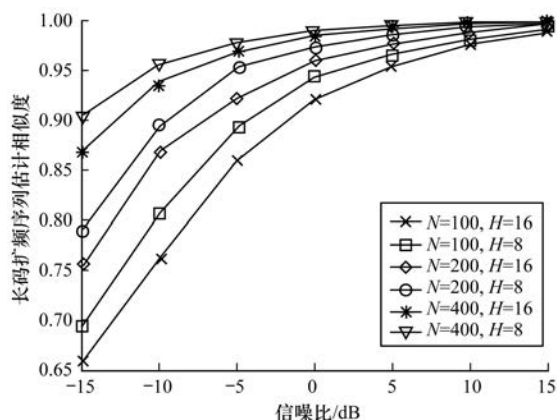


图 6 长码直扩信号的估计性能随输入信噪比变化曲线

4 结束语

本文针对长码直扩信号的扩频码及信息序列盲估计问题,结合重叠分段和 MCMC 算法的思想,提出基于重叠分段 MCMC-UKF 的盲估计算法。该算法能够实现对短码和长码信号的有效估计,并且不受扩频序列类型的限制。仿真实验结果表明,该算法具有适应较低信噪比的能力,且能获得较好的估计性能。

参考文献

- [1] BOUDER C, AZOU S, BUREL G. Blind estimation of the pseudo random sequence of a direct sequence spread spectrum signal [C]//Proceedings of IEEE Military Communications Conference. Los Angeles, USA: [s. n.], 2000: 967-970.
- [2] DONG Zhanqi, HU Nianying. A method for the detection of long pseudo random code DSSS signals based on the processing of delay-multiply (II)—the estimation of the information symbol period and the pseudo-random code sequence [C]//Proceedings of International Conference on Communication Technology. Hangzhou, China: [s. n.], 2008: 233-236.
- [3] QUI P, HUANG Z, JIANG W, et al. Improved blind-spreading sequence estimation algorithm for direct sequence spread spectrum signals [J]. IET Signal Process, 2008, 2(2): 139-146.
- [4] VLOK J, OLIVIER J. Blind sequence-length estimation of low-SNR cyclostationary sequences [J]. IET Communications, 2014, 8(9): 1578-1588.
- [5] ZHANG Tianqi, MU Aiping. A modified eigen-structure analyzer to lower SNR DSSS signals under narrow band interferences [J]. Digital Signal Process, 2008, 18(4): 526-533.
- [6] WANG Qian, GE Qian. Blind estimation algorithm of parameters in PN sequence for DSSS-BPSK signals [C]//Proceedings of 2012 International Conference on Wavelet Active Media Technology and Information Processing. Chengdu, China: [s. n.], 2012: 371-376.
- [7] SHA Zhizhao, HUANG Zhitao, ZHOU Yiyu, et al. Blind spreading sequence estimation based on hill-climbing algorithm [C]//Proceedings of International Conference on Signal Processing. Beijing, China: [s. n.], 2012: 1299-1302.
- [8] ZHANG H, GAN L, LIAO H, et al. Estimating spreading waveform of long-code direct sequence spread spectrum signals at a low signal-to-noise ratio [J]. IET Signal Process, 2012, 6(4): 358-363.
- [9] VLOK J, OLIVAER J. Blind sequence-length estimation of low-SNR cyclostationary sequences [J]. IET Communications, 2014, 8(9): 1578-1588.
- [10] WARNER S, MULGREW B, GRANT M. Triple correlation analysis of m sequences [J]. Electronics Letters, 1993, 29(20): 1755-1756.
- [11] HILL P, COMLEY V, ADAMS E. Techniques for detecting and characterizing covert communication signals [C]//Proceedings of IEEE Military Communications Conference. Monterey, USA: [s. n.], 1997: 1361-1365.
- [12] TSATSANIS M K, GIANNAKIS G B. Blind estimation of direct sequence spread spectrum signals in multipath. Signal Processing [J]. IEEE Transactions on Signal Processing, 1997, 45(5): 1241-1252.
- [13] DANESHMAND S, AGHAEINIA H, TOHIDIAN M, et al. Blind estimation of signal in periodic long-code DSSS communication [C]//Proceedings of IEEE Sarnoff Symposium. Washington D. C., USA: IEEE Press, 2009: 1-6.
- [14] BAI Juan, ZHANG Tianqi, ZHAO Defang, et al. PN sequence estimation of long-code DSSS signal based on virtual multiuser model [J]. Telecommunication Engineering, 2011, 51(8): 29-35.
- [15] 赵军桃, 张天骐, 江晓磊, 等. 基于 LEAP 神经网络的同步 DS-CDMA 伪码序列盲估计 [J]. 计算机应用研究, 2016, 34(2): 552-556.
- [16] 张天骐, 赵军桃, 江晓磊. 基于多主分量神经网络的同步 DS-CDMA 伪码盲估计 [J]. 系统工程与电子技术, 2016, 38(11): 2638-2647.

编辑 索书志