Blind Identification of Co-Existing Synchronous and Asynchronous Users for CDMA Systems

Zhengyuan Xu

***Abstract*—In a code-division multiple access (CDMA) system, if all users are either bit-synchronous or bit-asynchronous, the number of active users can be easily estimated from the dimension of the signal subspace, and users can be correspondingly identified. However, identification of co-existing synchronous and asynchronous users has not been addressed in the literature. In this letter, a subspace-based blind identification approach is presented by exploring correlations of relevant data vectors and evaluating projection of possible code vectors onto the noise subspace.**

***Index Terms*—Delay estimation, subspace decomposition, user identification.**

I. Introduction

DIRECT-SEQUENCE (DS) code-division multiple access (CDMA) has demonstrated high capacity and has been shown to be applicable for the new wireless networks [3], [7]. Thus, tremendous efforts have still been focused on multiuser detection [1], [2], [4].

A typical CDMA communication system often has a dynamic load due to the time-varying number of active users and duration of service for each user. Identification of active users will help the system to promptly process requests and efficiently allocate channels. In such a way, system capacity can be increased.

For clarity of presentation in the current context, let us define synchronous users with delays equal to multiples of bit period, and asynchronous users with other delays. We also assume each user is chip-synchronized. The scenario with fractional delay in chip period needs further investigation which is beyond the scope of this letter. Then, if all users are either synchronous or asynchronous, the number of active users can be easily estimated from the dimension of the signal subspace, and users be identified correspondingly [5], [6]. However, since delays are arbitrary, synchronous and asynchronous users may co-exist in the system. Identification of these users has not been addressed in the literature. Accordingly, this letter presents a subspace-based blind identification method. First, any delay can be factored by bit period T_b and chip period T_c as $\tau_k = d_k T_b + j_k T_c (j_k = 0, 1, \ldots, P-1)$ where d_k is an integer and P is the spreading factor. We categorize asynchronous users into three groups based on whether their chip delays j_k are smaller, equal or larger than *a priori*-set value around $P/2$. Then numbers of users in different groups and the number of synchronous users can be uniquely solved from a set of linear equations based on dimensions of corresponding correlation matrices. Secondly, active users can be identified and their delays be estimated by minimizing a projection error. Simulations show that the proposed method can provide reliable estimation.

Manuscript received November 20, 2000. The associate editor coordinating the review of this paper and approving it for publication was Prof. N. D. Sidiropoulos.

The author is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: dxu@ee.ucr.edu).

Publisher Item Identifier S 1070-9908(01)05239-7.

II. Data Model

Consider a CDMA system with k_s synchronous users and k_a asynchronous users. User k is assigned periodic spreading codes, which are put in a code vector $\boldsymbol{c}_k = [c_k(0), \ldots, c_k(P-1)]^T$. Let $K = k_a + k_s$ be the total number of active users in the system. For notational convenience, let us assign indices $1 \sim k_a$ to asynchronous users and $(k_a+1) \sim K$ to synchronous users. Then during the nth bit period, P chip samples can be collected in a vector. This data vector has the following expression:

$$\boldsymbol{y}(n) = \sum_{k=1}^{k_a} [\boldsymbol{c}_{k,1} w_k(n-d_k) + \boldsymbol{c}_{k,2} w_k(n-d_k-1)] + \sum_{k=k_a+1}^{K} \boldsymbol{c}_k w_k(n-d_k) + \boldsymbol{v}(n) \quad (1)$$

where channel gain is absorbed into the input $w_k(n)$. $\boldsymbol{c}_{k,1} = [0, \ldots, 0, c_k(0), \ldots, c_k(P-1-j_k)]^T$ and $\boldsymbol{c}_{k,2} = [c_k(P-j_k), \ldots, c_k(P-1), 0, \ldots, 0]^T$ are signatures corresponding to two consecutive bits from an asynchronous user. They depend on the user's chip delay j_k. $\boldsymbol{v}(n)$ is a white Gaussian noise vector.

In the input/output equation (1), there are two terms in the first summation corresponding to each asynchronous user, while only one term in the second summation corresponding to each synchronous user. In this sense, this observation window is synchronized to the synchronous users. However, if we take another window still with length P but going back l chip samples, then the new vector $\boldsymbol{y}_l(n)$ will have a different representation depending on the scenario of all delays. To clearly express $\boldsymbol{y}_l(n)$, it is necessary to differentiate k_a asynchronous users based on chip delay j_k. Therefore, we divide all k_a asynchronous users into three groups. Group 1 has k_1 users with $0 < j_k < P-l$, group 2 has k_2 users with $j_k = P-l$, and group 3 has k_3 users with $P-l < j_k < P$. Then it can be easily observed that the new window is only synchronized to the users in group 2, while not synchronized to all other users (totally $K-k_2$). This can

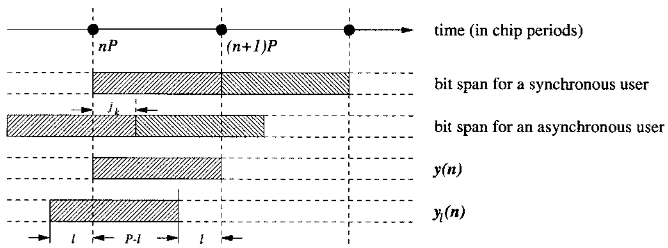

Fig. 1. Illustration of bit spans and corresponding data vectors.

be illustrated by Fig. 1 and is reflected in the expression for the new data vector $\boldsymbol{y}_l(n)$

$$\begin{aligned}\boldsymbol{y}_l(n) = &\sum_{k=1}^{k_1} [\tilde{\boldsymbol{c}}_{k,\,11} w_k(n-d_k) + \tilde{\boldsymbol{c}}_{k,\,12} w_k(n-d_k-1)] \\ &+ \sum_{k=k_1+1}^{k_1+k_2} \boldsymbol{c}_k w_k(n-d_k-1) \\ &+ \sum_{k=k_1+k_2+1}^{k_a} [\tilde{\boldsymbol{c}}_{k,\,31} w_k(n-d_k-1) \\ &\qquad + \tilde{\boldsymbol{c}}_{k,\,32} w_k(n-d_k-2)] \\ &+ \sum_{k=k_a+1}^{K} [\tilde{\boldsymbol{c}}_{k,\,41} w_k(n-d_k) \\ &\qquad + \tilde{\boldsymbol{c}}_{k,\,42} w_k(n-d_k-1)] + \tilde{\boldsymbol{v}}(n) \qquad (2)\end{aligned}$$

where there is only one term for each of the users in group 2 as indicated by the second summation, while all other users contribute two terms to this new vector. Signature $\tilde{\boldsymbol{c}}_{k,\,ij}$ is again a shifted version of the code vector $\boldsymbol{c}_k$, but the amount of shift depends on both j_k and l. It can be concluded from (1) and (2) that the choice of the observation window affects the total number of terms in a data vector. It also determines correlations between data vectors from different windows. These properties will be employed in identifying active users as described next.

III. Subspace-Based Identification of Active Users

We will first derive our method to estimate the number of asynchronous users in each group and the number of synchronous users. Then we identify active users based on these knowledge.

A. Determination of Numbers of Active Users—(k_1, k_2, k_3, k_s)

From (1) and (2), it can be found that numbers (k_1, k_2, k_3, k_s) of corresponding active users are closely related to the total number of terms in the expressions of corresponding data vectors, respectively. The latter can be obtained based on the estimation of dimensions of associated signal subspaces. The singular value decomposition (SVD) of the corresponding correlation matrix will be performed to achieve this goal. Totally at least four equations are required to estimate four unknowns k_1, k_2, k_3, k_s.

Assume that the input sequences are i.i.d. They are also independent of the AWGN. From (1), the autocorrelation matrix $\boldsymbol{R}$ of $\boldsymbol{y}(n)$ can be derived. The SVD on $\boldsymbol{R}$ yields the noise subspace $\boldsymbol{E}_n$ and the signal subspace $\boldsymbol{E}_s$. The dimension of $\boldsymbol{E}_s$ is observed as $2k_a + k_s$ by counting the number of independent[1] signatures in (1). It can be estimated by a serial search (see [6]). If the search result is denoted by ν, then the first equation follows:

$$2(k_1 + k_2 + k_3) + k_s = \nu \qquad (3)$$

where $k_a = k_1 + k_2 + k_3$ is used. It is worth mentioning that if all users are synchronous as discussed in [6], then ν is equal to k_s, which completes the determination process. On the other hand, if all users are asynchronous as discussed in [5], then ν is equal to $2k_a$, which makes determination also easier. However, for the current setup, more information is required to achieve our goal.

Following similar steps, SVD can be performed on the autocorrelation matrix of $\boldsymbol{y}_l(n)$. If we denote the estimate for the dimension of its signal subspace by $\tilde{\nu}$, then from (2), we obtain

$$2k_1 + k_2 + 2k_3 + 2k_s = \tilde{\nu}. \qquad (4)$$

Two more equations can be derived from the crosscorrelation between $\boldsymbol{y}_l(n)$ and $\boldsymbol{y}(n)$ and crosscorrelation between $\boldsymbol{y}_l(n)$ and $\boldsymbol{y}(n-1)$. First, it can be easily verified from (1) and (2) that the dimension of signal subspace associated with SVD of the first crosscorrelation matrix is $2k_1 + k_2 + k_3 + k_s$. Secondly, after obtaining $\boldsymbol{y}(n-1)$ from (1)

$$\begin{aligned}\boldsymbol{y}(n-1) = &\sum_{k=1}^{k_a} [\boldsymbol{c}_{k,\,1} w_k(n-d_k-1) + \boldsymbol{c}_{k,\,2} w_k(n-d_k-2)] \\ &+ \sum_{k=k_a+1}^{K} \boldsymbol{c}_k w_k(n-d_k-1) + \boldsymbol{v}(n-1) \qquad (5)\end{aligned}$$

the dimension of the corresponding signal subspace is found to be $k_1 + k_2 + 2k_3 + k_s$. If we denote their search results by ν_f and ν_b, then we obtain two more equations

$$2k_1 + k_2 + k_3 + k_s = \nu_f \qquad (6)$$

$$k_1 + k_2 + 2k_3 + k_s = \nu_b. \qquad (7)$$

By collecting our unknowns in a vector $\boldsymbol{k} = [k_1 k_2 k_3 k_s]^T$, we can uniquely solve it from (3), (4), (6), and (7) in the least square sense. Since each entry in $\boldsymbol{k}$ should be an integer, a close integer is then taken for each element of $\boldsymbol{k}$ from its estimate. All required correlations are estimated by their sample averages based on received data in N bit periods (or observation windows). In order to obtain a reliable estimate for $\boldsymbol{k}$, l is suggested to be chosen as an integer around $(P/2)$. Furthermore, it is implicitly assumed in SVD operation that the dimension of signal subspace is less than the dimension of the corresponding matrix. This requires that the system should have moderate load.

B. Identification of Active Users

According to our previous analysis, each active synchronous user has delay $j_k = 0$ and contributes only one signature to

[1]We assume signatures are independent. False alarm in identification will occur if this assumption is violated. This topic is still under investigation. For a general discussion, readers can also refer to [5].

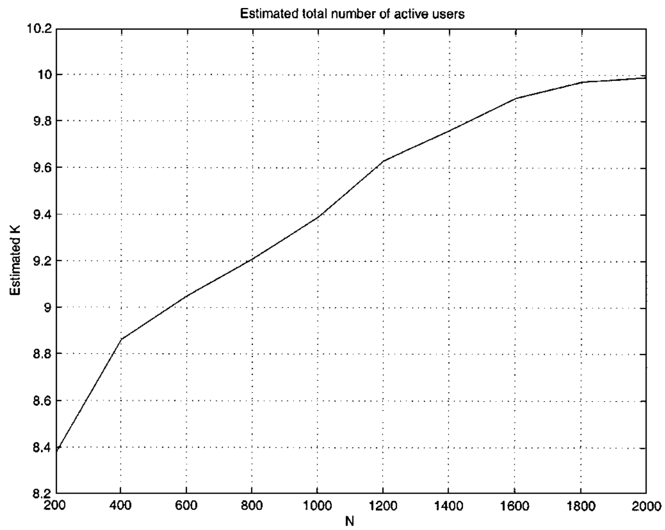

Fig. 2. Estimated total number of active users in the system.

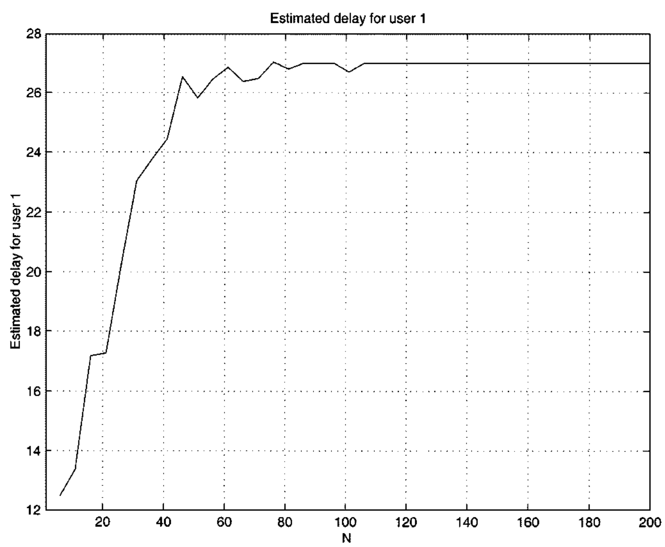

Fig. 3. Estimated delay for a particular active user (user 1).

the data vector $\boldsymbol{y}(n)$. This signature is the code vector $\boldsymbol{c}_k$. As described in [6], these signatures $\boldsymbol{c}_k (k = k_a + 1, \ldots, K)$ lie in the signal subspace $\boldsymbol{E}_s$, thus their projection onto the noise subspace $\boldsymbol{E}_n$ is zero

$$f_k = (\boldsymbol{c}_k^H \boldsymbol{E}_n)(\boldsymbol{c}_k^H \boldsymbol{E}_n)^H = ||\boldsymbol{c}_k^H \boldsymbol{E}_n||^2 = 0.$$

If totally M potential users are provided services in the system, then their code vectors are *priori* known. Check all possible values of f_k for $k = 1, \ldots, M$. If f_k is found to exceed zero, then it can be concluded that user k is not a synchronous user. It could be an asynchronous user or inactive user. After all k_s synchronous users are identified, there are possibly only $M - k_s$ users to be identified. However, for each of all active asynchronous users, two signatures $\boldsymbol{c}_{k,1}$ and $\boldsymbol{c}_{k,2}$ are in the signal subspace. They depend on the user's chip delay j_k. We can construct two code vectors $\boldsymbol{c}_{k,1}$ and $\boldsymbol{c}_{k,2}$ for each of $M - k_s$ possible remaining users with all possible delays $j_k = 1, \ldots, P - 1$. Then we project them onto the noise subspace $\boldsymbol{E}_n$. If

$$\tilde{f}_{k,1} = ||\boldsymbol{c}_{k,1}^H \boldsymbol{E}_n||^2 = 0, \quad \tilde{f}_{k,2} = ||\boldsymbol{c}_{k,2}^H \boldsymbol{E}_n||^2 = 0$$

then this user is identified as an asynchronous user with delay j_k. This process needs $P-$ 1 searches for each user due to the delay j_k. Thus, in total, $(M - k_s)(P - 1)$ searches are required to identify k_a asynchronous users and corresponding delays.

This method is also similar to the MUSIC-type algorithm presented in [5]. However, method [5] is based on the *a priori* knowledge of dimension of the noise subspace or signal subspace. In the current context, the dimension of signal subspace equals $2k_a + k_s$, which must be determined first based on our approach.

IV. Simulations

We simulate a DS-CDMA system with ten equally-powered users and 15 dB AWGN in a flat fading environment. Each user is assigned a Gold sequence of length $P = 31$ as a code sequence. Delays for these users are set to be [27 5 22 0 16 0 7 10 16 21], which indicates two synchronous users and eight asynchronous users. We choose $l = 15$, then theoretically $\boldsymbol{k} = [3\ 2\ 3\ 2]^T$. First, we estimate $\boldsymbol{k}$ based on our proposed method and obtain the total number of active users by adding all elements up. The average result versus the number of bit periods (N) is obtained from 50 realizations and shown in Fig. 2. In each realization, noninteger elements are maintained in the estimated vector without rounding operation. It can be seen that the estimated value is expected to converge to its true value ($K = 10$) as $N \rightarrow \infty$, since the dimension of corresponding signal subspaces can be reliably estimated with more data samples. We further identify these active users in terms of code sequences and delays. For each code sequence in Gold sequences with $P = 31$, we construct the code vectors $\boldsymbol{c}_{k,1}$ and $\boldsymbol{c}_{k,2}$ for possible delay between 0 and $P-$ 1 chip periods. According to their projections onto the noise subspace, the sum of $\tilde{f}_{k,1}$ and $\tilde{f}_{k,2}$ can be obtained. If for a particular code sequence this value has a sharp peak with respect to different delays, then the user with this code sequence is identified to be active, and the delay corresponds to the peak point. Fig. 3 shows the average delay estimate for user one after it has been detected. It is observed that the estimate converges to its true value (27 chip periods) after 120 bit periods. This indicates that even a short data record can provide reliable estimate.

References

[1] H. C. Huang and S. Verdu, "Linearly differentially coherent multiuser detector for multipath channels," *Wireless Personal Commun.*, vol. 6, pp. 113–136, Jan. 1998.

[2] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[3] E. Nikula, A. Toskala, E. Dahlman, L. Girard, and A. Klein, "FRAMES multiple access for UMTS and IMT-2000," *IEEE Personal Commun.*, pp. 16–24, Apr. 1998.

[4] M. Saquib, R. Yates, and N. Mandayam, "Decorrelating detectors for a dual rate synchronous DS/CDMA channel," *Wireless Personal Commun.*, vol. 9, pp. 197–216, May 1999.

[5] E. Strom, S. Parkvall, S. Miller, and B. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84–93, Jan. 1996.

[6] W. Wu and K. Chen, "Identification of active users in synchronous CDMA multiuser detection," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1723–1735, Dec. 1998.

[7] M. Zeng and V. Bhargava, "Recent advances in cellular wireless communications," *IEEE Commun. Mag.*, pp. 128–138, Sept. 1999.