Simplified MAP-MUD for Active User CDMA

Pedram Pad, Ali Mousavi, Ali Goli, and Farokh Marvasti, *Senior Member, IEEE*

***Abstract*—In CDMA systems with variable number of users, the decoder consists of two stages. The first stage is the active user identification and the second one is the multi-user detection. Most of the proposed active user identification methods fail to work in overloaded CDMA systems (when the number of users is more than the spreading factor). In this paper we propose a joint active user identification and multi-user detection for binary input CDMA systems using the Viterbi algorithm. We will show that the proposed identification/detection method is Maximum A posteriori Probability (MAP) and outperforms the pervious results. In addition, some suboptimum decoders will be proposed that have lower computational complexities but lower performances.**

***Index Terms*—CDMA, active user identification, multi-user detection, MAP, Viterbi Algorithm.**

I. Introduction

In practical situations, the users in the system are not always known and may vary in different time indicies. Thus, the more realistic model for a CDMA system is

$$Y = \mathbf{A}_{\mathcal{A}} X_{\mathcal{A}} + N \tag{1}$$

where $\mathcal{A} \subseteq \{1, \ldots, n\}$ is the set of the indices of the active users in the system, $\mathbf{A}_{\mathcal{A}}$ is the $m \times |\mathcal{A}|$ signature matrix, $X \in \{\pm 1\}^{\mathcal{A}}$ that each of its entry is data of a user, N is the $m \times 1$ channel noise vector and Y is the $m \times 1$ received vector. Inserting zero as data of the incactive users, we get

$$Y = \mathbf{A}X + N \tag{2}$$

where $\mathbf{A}$ is the $m \times n$ signature matrix.

In such systems, the knowledge of active users is assumed for multi-user detection. Thus, in general, at the receivers ends there are two main modules. The task of the first module is to identify the active users and the second module attempts to extract the user's data that are identified as active. The performance of the overall decoder is highly dependent on the performance of the active user identification section.

Active user detection in multiuser systems has been dealt in several papers [2]–[6]. The subject of identifying an individual user arriving or leaving at system is studied in [3], [4], [5]. In [7], the authors have utilized procedure on the basis of multiple signal classification (MUSIC) algorithm for recognizing the active users. In [8] and [9] authors have implemented MAP system with 3 users and spreading factor of 7 which can be considered as an underloaded and small scale CDMA system.

Manuscript received November 27, 2010. The associate editor coordinating the review of this letter and approving it for publication was T. Tsiftsis.

The authors are with the Advanced Communication Research Institute (ACRI), Electrical Engineering Department, Sharif University of Technology, Tehran, Iran (e-mail: {pedram_pad, ali_mousavi, agoli}@ee.sharif.edu, marvasti@sharif.edu).

Digital Object Identifier 10.1109/LCOMM.2011.040111.102320

However, in this paper we implement a highly overloaded and large scale CDMA system with 96 users and spreading factor of 64 with significantly reduced complexity and slightly improving the traffic model in comparison with [8] and [9].

For the traffic model we assume that the interval of being active or inactive for each user have exponential distributions. An active user remains active in the time interval with probability p and becomes inactive with probability $1-p$. Also, an inactive user remains inactive with probability q and become active with probability $1-q$. Thus, the activation/inactivation process of the users forms a Markov chain.

II. Design of the Optimum Decoder

Suppose that the L user data vectors are

$$\underline{X}_{n\times L} = [X_1, X_2, \ldots, X_L] \tag{3}$$

(each vector X_i contains zeros for the inactive users). The resultant vectors that are sent through the channel are

$$\mathbf{A}\underline{X} = [\mathbf{A}X_1, \mathbf{A}X_2, \ldots, \mathbf{A}X_L] \tag{4}$$

If the matrix $\mathbf{A}$ is such that the mapping $X \to \mathbf{A}X$ is injective, $\mathbf{A}X_i$'s determine X_i uniquely. Also, the corresponding received noisy vectors are

$$\underline{Y}_{m\times L} = [Y_1, Y_2, \ldots, Y_L] \tag{5}$$

The decoder is

$$\hat{\underline{X}} = \text{argmax}_{\tilde{\underline{X}}} \bar{f}\left(\tilde{\underline{X}}|\underline{Y}\right) \tag{6}$$

where $\tilde{\underline{X}} = [\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_L] \in \{0, \pm 1\}^{n\times L}$ and $\bar{f}$ is the $n \times L$-dimensional PDF of $\underline{Y}$. According to the Bayes rule, we have

$$\bar{f}\left(\tilde{\underline{X}}|\underline{Y}\right) = \frac{\mathbb{P}\left(\tilde{\underline{X}}\right)\bar{f}\left(\underline{Y}|\tilde{\underline{X}}\right)}{\bar{f}\left(\underline{Y}\right)} \tag{7}$$

Since the states of the users change according to the Markov chain described before, we have

$$\mathbb{P}\left(\tilde{\underline{X}}\right) = \mathbb{P}\left(\tilde{X}_1\right)\prod_{i=1}^{L-1}\mathbb{P}\left(\tilde{X}_{i+1}|\tilde{X}_i\right) \tag{8}$$

Now, notice that since the channel is memoryless, we have

$$\bar{f}\left(\underline{Y}|\tilde{\underline{X}}\right) = \prod_{i=1}^{L} f\left(Y_i|\tilde{X}_i\right) \tag{9}$$

where f is the m-dimensional PDF of Y. For an AWGN channel with noise variance σ^2, we have, $f\left(Y_i|\tilde{X}_i\right) = \left(2\pi\sigma^2\right)^{-m/2}\exp\left(-\|Y_i - \mathbf{A}\tilde{X}_i\|^2/\left(2\sigma^2\right)\right)$.

According to (8) and (9) and the fact that $\bar{f}(\underline{Y})$ is a constant term, (6) can be rewritten as

$$\hat{\underline{X}} = \text{argmax}_{\tilde{\underline{X}}} \mathbb{P}\left(\tilde{X}_1\right) f\left(Y_1|\tilde{X}_1\right) \times \prod_{i=1}^{L-1} \mathbb{P}\left(\tilde{X}_{i+1}|\tilde{X}_i\right) f\left(Y_{i+1}|\tilde{X}_{i+1}\right) \quad (10)$$

Direct implementation of the above maximization needs 3^{nL} operations. By using Viterbi algorithm, there are 3^n states in each level of the algorithm. The weights of the states of the first level are $\mathbb{P}\left(\tilde{X}_1\right) f\left(Y_1|\tilde{X}_1\right)$ and the transition weights from state $\tilde{X}_i$ to $\tilde{X}_{i+1}$ is $\mathbb{P}\left(\tilde{X}_{i+1}|\tilde{X}_i\right) f\left(Y_{i+1}|\tilde{X}_{i+1}\right)$, according to (10). At each state of the $(i+1)^{th}$ level,we multiply the weights of the states of the i^{th} level by the transition weight and pick the path that has the maximum weight. Therefore, going through the trellis level by level, performing the above algorithm and choosing the survivor path at the end of the trellis, we arrive at the solution to maximizing (10). The complexity of this algorithm is $3^n 3^n L$ operations. Now we desire to show that this maximization can also be done by $2^n 3^n L$ operations. Through further simplifications, we will perform the MAP decoder with $2^n 2^m 3^{n-m} L$ operations for a class of signature matrices $\mathbf{A}$.

The transition weights of the trellis discussed above has two terms. The first term is $\mathbb{P}\left(\tilde{X}_{i+1}|\tilde{X}_i\right)$ which, according to (8), depends only on the activeness/inactiveness pattern of the users in $\tilde{X}_i$ and $\tilde{X}_{i+1}$ and does not depend on the data of the active users. The second term $f\left(Y_{i+1}|\tilde{X}_{i+1}\right)$ is not a function of $\tilde{X}_i$. Hence, the 3^n states of level i can be put in 2^n categories (according to their patterns of activeness/inactiveness) and from each category we can save only the one that has the maximum $f\left(Y_{i+1}|\tilde{X}_{i+1}\right)$ as the representative. At each state of level $i+1$, we only need to search among these 2^n representatives instead of all 3^n states. Thus at the end of each level we save only 2^n paths and we are sure that the survivor path will not be omitted at all. This will decrease the complexity of the decoder down to $2^n 3^n L$ which is much less than the previous $3^n 3^n L$ for typical values of the number of users n.

Similar to [1] and [10] we can decrease this complexity even further; we can reduce the complexity from $2^n 3^n$ to $2^n 2^m 3^{n-m}$. Therefore, we can perform the trellis operations with $2^n 2^m 3^{n-m} L$ computations.

A. Low Complexity Decoder for Signature Matrices

Similar to [1] and [10] if the signature matrix of the system is

$$\mathbf{A}_{km\times kn} = \mathbf{P}_{k\times k} \otimes \mathbf{D}_{m\times n}, \quad (11)$$

where $\mathbf{P}$ is invertible and $\otimes$ denotes the Kronecker product, the decoder of the $km \times kn$ system can be decomposed to k decoders of the $m \times n$ systems. It has been shown that if $\mathbf{P}$ is unitary, this decoder is MAP. Hence, we can have a CDMA system with kn users and km chips that exploits MAP active user identification and multiuser detection with the computational complexity of $k2^n 2^m 3^{n-m} L$ instead of 3^{knL} operations.

The main drawback of the proposed decoding method is its delay and memory since it must wait for the all the vectors to be received. In the next section, we will propose some sub-optimum decoders that make the delay and the required memory very short and also decrease the computational complexity even further.

III. Suboptimum Decoders

The idea of the first sub-optimum decoder is to truncate the trellis temporally to decrease the delay of the decoding process.

A. Temporally Truncated Trellis

In this method we only consider the last $M(\ll L)$ levels of the trellis. The delay of this decoder is M the memory needed is $M \times 2^n$ which may be much less than the $L \times 2^n$ memory needed for the optimum decoder proposed in the previous section. The following expressions can be stated about this method.

- *High values of* E_b/N_0: Since for the system that has high E_b/N_0 the decoded data at each time index is more certain, it is not needed to observe the samples for a long time before and after for making a decision. In the limit, if we have $E_b/N_0 = \infty$ (noiseless channel), irrelevant to the traffic model of the system, i.e., irrelevant to p and q, we can decode the transmitted vector with no errors without any need of trellis ($M = 0$).
- *Small values of* $|p+q-1|$: The matrix of the transition between activeness and inactiveness of the users in the Markov model described before is

$$\mathbf{Q} = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} \quad (12)$$

  This matrix has two eigen values of 1 and $p+q-1$. If $p+q-1 = 0$, we have that the state of each user in a time index is independent of its state in any other time indices. Thus, $M = 0$ gives the optimum decoder. Therefore, roughly speaking, since the transition probability matrix of activation/inactivation of a user in time indices $i - M$ and i is $\mathbf{Q}^M$, if $|p+q-1|^M$ is very close to 0, we can say that the states of the users in time indices with distance more than M are (almost) independent. Consequently saving the last M levels of the trellis gives almost optimum decoder. In fact, if we have $p+q-1 = 0$, we have the unrealistic system assumed in [10] and our decoder reduces to the decoder proposed in the same reference.

B. Neighboring Search

In this method, at each state of level $i+1$, we do not search among all 2^n states of the previous level for finding the maximum-weight path. Take a non-negative integer d. For each state, we only check the states of the previous level that differs with the current state in only d positions. This decreases the computational complexity of the decoder by a factor of $2^{-n}\sum_{i=0}^{d}\binom{n}{i}$. It can be expressed that the *larger* p *and* q*,the* *smaller the value of* d becomes. Notice that for any $d < n$, there is a positive lower bound for the error even in noiseless channel. But the probability of such alterations are negligible for typical values of p and q.

C. Permuting Signatures

We propose a sub optimum decoder with the complexity of

$$2^n \left(\sum_{i=0}^{m} \binom{m}{i} + \sum_{i=1}^{n-m} \binom{n}{m+i} 2^i \right) L \tag{13}$$

We need to check $2^{|\mathcal{C}'|}$ possibilities where $|\mathcal{C}'|$ is the number of active users among the last $n-m$ columns of $\mathbf{A}$. Therefore, if we can lower $|\mathcal{C}'|$, the search in users becomes easier. This can be done by permuting the order of the users in the system. This decrease the computational complexity of the decoder as given in (13). The condition for this method to work is that every m columns of the signature matrix $\mathbf{A}$ must be linearly independent, which is not very restricting in typical systems. But since all of the $m \times m$ sub-matrices of $\mathbf{A}$ cannot be unitary, this decoder is not optimum. Again, it is easy to prove that for this method if E_b/N_0 increases, the probability of error tends to zero.

IV. Simulation Results

In this section we compare the behavior of the optimum and suboptimum decoders introduced in the previous sections and compare them with the previous works. We simulated a highly overloaded binary CDMA system with 96 users and chip rate 64. We utilize $\mathbf{A}_{64\times 96} = \frac{1}{\sqrt{16}} \mathbf{H}_{16} \otimes \mathbf{D}_{4\times 6}$ as the signature matrix where $\mathbf{H}_{16\times 16}$ is the 16×16 Hadamard matrix and $\mathbf{D}_{4\times 6}$ was constructed randomly with the conditions that its first four columns form a 4×4 unitary matrix, the norm of all of the columns are unity and $\mathbf{D}$ is injective over the set $\{0, \pm 1\}^6$. According to II-A, the decoding problem of this system can be reduced to 16 decoding systems of size 4×6. Thus, we focus and discuss the decoding problem of a system of size 4×6. Also, in the simulated system, we choose $p = q = 0.9$. Three performance curves can be assumed for every decoder, the rate of the active users that are identified as inactive, the rate of error in detecting the data of the active users and the rate of the inactive users that are identified as active. These three performance curves for the optimum decoder are shown in Fig. 1 in comparison with the best decoder that proposed in [10].

We have also simulated the sub-optimum decoders of Section III. The *Temporally Truncated Trellis* for $M = 10$, the *Neighboring Search* for $d = 3$ and the *Permuting Signatures* are simulated and their BER versus E_b/N_0 are depicted in Fig. 2.

According to equations and algorithms stated in previous sections, it is noteworthy to compare the computational complexity when 50000 vectors are going to be decoded. For a MAP decodes, we need $3^{96\times 50000}$ operations. Utilizing Viterbi algorithm directly, we need $3^{96} \times 3^{96} \times 50000$ operations. Taking advantage of Kronecker product for reducing complexity, we need $16 \times 3^6 \times 3^6 \times 50000$ operations. Using Viterbi algorithm with categorized states, we require $16 \times 2^6 \times 3^6 \times 50000$ operations. Additionaly according to [1], we only demand $16 \times 2^6 \times 2^4 \times 3^2 \times 50000$ operations. Suboptimum decoders reveal outstanding numerical results for computational complexity. Through implementing *Neighboring Search* for $d = 3$, we require $16 \times 42 \times 3^6 \times 50000$ operations. By executing *Permuting Signatures*, we need $16 \times 80 \times 2^4 \times 50000$ operations, which shows notable reduction in comlexity.

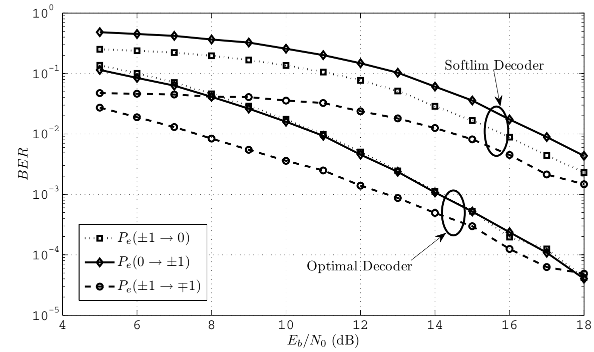

Fig. 1. Probability of error for the three possible kinds of errors for the proposed MAP decoder and the previous best decoder (Softlim).

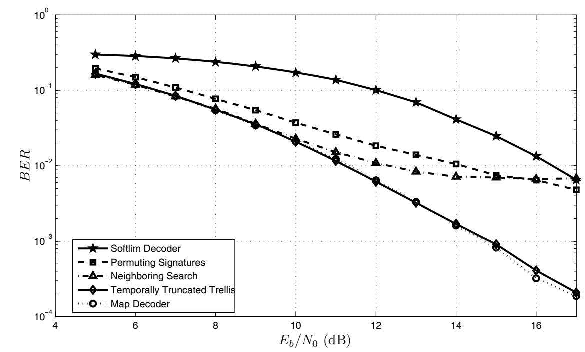

Fig. 2. BER for the proposed sub-optimum decoders.

References

[1] S. Buzzi, A. D. Maio, and M. Lops, "Code-aided blind adaptive new user detection in DS/CDMA systems with fading time-dispersive channels," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2637–2649, Oct. 2003.

[2] K. W. Halford and M. Brandt-Pearce, "New-user identification in a CDMA system," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 144–155, Jan. 1998.

[3] U. Mitra and H. V. Poor, "Activity detection in a multi-user environment," *Wireless Personal Commun.*, vol. 3, no. 12, pp. 149–174, Jan. 1996.

[4] –, "Analysis of an adaptive decorrelating detector for synchronous cdma channels," *IEEE Trans. Commun.*, vol. 44, pp. 257-268, 1996.

[5] T. Oskiper and H. V. Poor, "Online activity detection in a multiuser environment using the matrix cusum algorithm," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 477–493, Feb. 2002.

[6] W.-C. Wu and K.-C. Chen, "Identification of active users in synchronous CDMA multiuser detection," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1723–1735, Dec. 1998.

[7] D. Angelosante, E. Biglieri, and M. Lops, "Multiuser detection in a dynamic environment: joint user identification and parameter estimation," *ISIT07*, June 2007.

[8] –, "Low-complexity receivers for multiuser detection with an unknown number of active users," *Signal Process.*, vol. 9, no. 6, May 2010.

[9] P. Pad, M. Soltanolkotabi, S. Hadikhanlou, A. Enayati, and F. Marvasti, "Errorless codes for over-loaded CDMA with active user detection," *ICC'09*, June 2009.

[10] P. Pad, F. Marvasti, K. Alishahi, and S. Akbari, "A class of errorless codes for over-loaded synchronous wireless and optical CDMA systems," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2705-2715, June 2009.