Computationally Efficient Iterative Multiuser Detection and Decoding

Suman Das Elza Erkip Behnaam Aazhang
Department of Electrical and Computer Engineering, Rice University
MS 366, Houston, Texas 77005-1892
Phone (713)527-8750x2899 Fax: (713)-524-5237
Email: {suman, elza, aaz}@rice.edu

Abstract

We describe an iterative detection and decoding scheme for the uplink in a convolutionally coded direct sequence code division multiple access (DS-CDMA) system. We consider maximum-a-posteriori (MAP) decoding and reduce the complexity through an iterative interference cancellation scheme combined with a suboptimal channel decoding algorithm. The MAP priors are updated at every iteration step. We investigate the performance of this low complexity scheme and observe that it is close to optimal.

1 Introduction

The uplink in a cellular system consists of multiple transmitters communicating with a common receiver. Direct sequence code division multiple access (DS-CDMA) is an attractive channel accessing scheme in which each user's information bearing signal is modulated on a distinct spreading code. However, due to the inherent asynchronous nature of the system, the signature waveforms cannot be designed to be orthogonal for all possible delays. This causes CDMA systems to be interference limited. In order to mitigate this problem a number of multiuser detection schemes have been proposed (See [8] for an excellent review).

Apart from the multiple access interference (MAI), the second source of error at the receiver is the thermal noise. The use of convolutional codes in conjunction with the DS-CDMA system have been shown ([6, 9]) to provide sufficient protection against errors. The optimal decoding rule for an asynchronous multiuser CDMA system employing convolutional codes consists of combining the trellises of all the users. However, this requires exponential complexity in the number of users, and the feasibility of implementing the joint decoding algorithm is severely limited.

Traditional way of reducing the complexity of the multiuser decoding problem is to first detect the signals of different users using a multiuser detector, and then decode these signals by single user channel decoders. Recently there have been efforts to design the detector and decoder jointly in a low complexity manner and get performance

closer to optimal [3, 7, 2]. In this study, we describe a joint detection and decoding scheme based on the maximum-a-posteriori (MAP) criterion. Reduction in complexity is achieved through iterative interference cancellation and suboptimal channel decoding. The update of priors at every stage of the iteration ensures that loss in performance is not substantial.

In the next section we describe the system model and the basic iterative detection and decoding scheme. In Section 3, we introduce a suboptimal channel decoding algorithm and investigate the performance in conjunction with the iterative detection and decoding scheme.

2 Iterative detection and decoding

We investigate a baseband DS-CDMA system where the signals of K active users, each possibly with a different strength, arrive asynchronously at the receiver. The information bits of user k , denoted by $\mathbf{b}_k$, are encoded by a linear convolutional channel encoder. This coded bit sequence $\mathbf{d}_k$ is then modulated by a distinct signature waveform s_k ($\langle s_k, s_k \rangle = 1$) of length N_c which assumes nonzero values over the symbol period $[0, T]$.

The amplitude of the received signal from user k is A_k , and it accounts for the emitted signal power as well as channel attenuation. The signal of user k arrives with a delay τ_k at the receiver. Thus the cumulative received signal over the period of N coded bits is given by

$$r(t) = \sum_{k=1}^K \sum_{i=1}^N A_k d_k(i) s_k(t - iT - \tau_k) + z(t),$$

where $z(t)$ is the additive white Gaussian noise. The discretized chip matched filter output of this signal is an $N_c(N+1) \times 1$ vector

$$\mathbf{r} = \mathbf{S}\mathbf{A}\mathbf{d} + \mathbf{z}, \quad (1)$$

where $\mathbf{S}$ is an $N_c(N+1) \times NK$ matrix of the signature waveforms, $\mathbf{A}$ is the $NK \times NK$ diagonal matrix of user amplitudes and $\mathbf{d}$ is the coded bits of all users over period

[0,NT]. Further details about this model of the uplink can be found in [6]. If the received signal is passed through a bank of code matched filters, then the output of the filter bank can be expressed as an $NK \times 1$ vector

$$\mathbf{y} = \mathbf{R}_N \mathbf{d} + \boldsymbol{\eta}, \quad (2)$$

where $\mathbf{R}_N = \mathbf{S}'\mathbf{S}$ is the asynchronous code correlation matrix and $\boldsymbol{\eta}$ is the additive noise vector. The output of the bank of code-matched filter provides a sufficient statistic for the estimation of the vector $\mathbf{d}$.

The optimal decoding rule which minimizes the probability of error at the receiver is the MAP decoding rule. Using MAP decoding, the estimate of the information bits of all users $\hat{\mathbf{b}}$ can be found as

$$\hat{\mathbf{b}} = \arg \max_{\mathbf{b}} \log p(\mathbf{b}|\mathbf{r}).$$

Equivalently, the coded bit estimates are given by

$$\begin{aligned} \hat{\mathbf{d}} &= \arg \max_{\mathbf{d} \in \mathcal{C}} \log p(\mathbf{d}|\mathbf{r}) \\ &= \arg \max_{\mathbf{d} \in \mathcal{C}} [\log p(\mathbf{d}, \mathbf{r}) - \log p(\mathbf{r})] \\ &= \arg \max_{\mathbf{d} \in \mathcal{C}} \log p(\mathbf{d}, \mathbf{r}) \\ &= \arg \max_{\mathbf{d} \in \mathcal{C}} [\log p(\mathbf{r}|\mathbf{d}) + \log p(\mathbf{d})], \end{aligned} \quad (3)$$

where $\mathcal{C}$ denotes the collection of codebooks of all the users. This optimum decoding rule requires joint search through the codebooks of all the users and results in exponential computational complexity in the number of users [4].

A simple suboptimal scheme consists of using y_k , the output of the matched filter for user k , as a statistic to estimate the bits of that user through a single user decoder. However this statistic is corrupted by the interference from other users. In order to improve performance of the single user decoder, we must provide it with an interference-free statistic. One way of reducing the interference is to use an iterative interference cancellation scheme [3]. It was shown in [2] that combining the iterative interference cancellation with prior updates improves the performance. We next describe this scheme briefly.

We define c as the user of concern and the set I as the set of interfering users. The coded bit estimates of the interfering users $\hat{\mathbf{d}}_I$ are used to eliminate interference from the received signal $\mathbf{r}$. The resulting signal $\hat{\mathbf{y}}_c$ is given by

$$\hat{\mathbf{y}}_c = \mathbf{S}_c^\top (\mathbf{r} - \mathbf{S}_I \mathbf{A}_I \hat{\mathbf{d}}_I). \quad (4)$$

The signal $\hat{\mathbf{y}}_c$ is then used to estimate the data transmitted by user c using the MAP decoding rule. Hence, the estimates of the coded bits of user c are given by

$$\begin{aligned} \hat{\mathbf{d}}_c &= \arg \max_{\mathbf{d}_c \in \mathcal{C}_c} \log p(\mathbf{d}_c|\hat{\mathbf{y}}_c) \\ &= \arg \max_{\mathbf{d}_c \in \mathcal{C}_c} [\log p(\hat{\mathbf{y}}_c|\mathbf{d}_c) + \log p(\mathbf{d}_c)], \end{aligned} \quad (5)$$

where $\mathcal{C}_c$ is the codebook of user c . The above maximization requires considering the likelihood $p(\hat{\mathbf{y}}_c|\mathbf{d}_c)$ as well the prior $p(\mathbf{d}_c)$. Iteration among users and update of priors at every iteration step ensures that reliable estimates for all users' bits can be obtained.

Our initial priors are uniform on all the coded bits and initial estimates $\hat{\mathbf{d}}_c^{(0)}$ for all users $c \in \{1, \dots, K\}$ are obtained through the matched filter outputs. Suppose at the end of iteration i , $\hat{\mathbf{d}}_1^{(i)}, \dots, \hat{\mathbf{d}}_K^{(i)}$ denote the estimates of the coded bits. At the next iteration step $i + 1$, $\hat{\mathbf{d}}_I^{(i)}$ is used to cancel interference in (4), and the posterior $p^{(i)}(\mathbf{d}_c|\hat{\mathbf{y}}_c)$ of the previous iteration step i is used as the prior $p(\mathbf{d}_c)$ in (5). The iterative procedure stops when there is no further change in the estimates of the coded bits in successive iterations. A more detailed explanation of this iterative algorithm along with storage saving methods, performance and complexity analysis is given in [2].

We observe that this iterative procedure requires the calculation of $p(\mathbf{d}_c|\hat{\mathbf{y}}_c)$ for users $c \in \{1, \dots, K\}$ across iterations. It was shown in [2] that rather than calculating and keeping $p(\mathbf{d}_c|\hat{\mathbf{y}}_c)$ for all possible $\mathbf{d}_c$ across iterations, keeping probabilities of only the most M probable codewords gives good performance. A Viterbi algorithm, in which we keep the most probable M paths into every state at every stage of the trellis, can be used for this purpose.

The Viterbi algorithm for a convolutional code of constraint length κ and codeword length of N requires $N2^{(\kappa+1)}$ operations for each user, indicating that the complexity becomes prohibitively large for large constraint lengths. In the next section, we describe a suboptimal channel decoding algorithm which has linear complexity in the constraint length of the convolutional code. We show that this channel decoding algorithm used in conjunction with the iterative detection and decoding performs close to single user bound.

3 Suboptimal channel decoding algorithm

In this section we focus on the single user case and describe a suboptimal channel decoding algorithm which picks out M codewords that have large probabilities, but not necessarily the largest, given the channel output. These M large probability codewords can then be used either to find a codeword estimate, or as the list required in the above iterative scheme.

Consider a binary linear code $\mathcal{C}$ of rate R and block length N . This may be a linear block code or a truncated convolutional code. Since the code has rate R , of the N coded bits at most NR of them can be chosen independently. For example in a rate $2/3$ parity block code, any two coded bits exactly determine the third one. We will have more detailed discussion of the choice of NR independent

bits later in this section.

We assume the coded bits $\mathbf{d}$ are transmitted through an additive white Gaussian noise channel using BPSK modulation. We denote the received vector as $\mathbf{r}$. The optimum rule that gives minimum probability of decoding error is the MAP rule explained in (3).

If we ignored the fact that the $\mathbf{d} \in \mathcal{C}$, the MAP decoding rule with uniform prior on $\mathbf{d}$ would result in the sign detector, which gives

$$\hat{\mathbf{d}}_c = \text{sgn}(\mathbf{r}).$$

Our sub-optimum decoding rule is based on using the sign detector for only NR independent bits of the N coded bits. The remaining $N(1 - R)$ bits of $\mathbf{d}$ can be obtained through the code structure of $\mathcal{C}$. The choice of these NR independent bits will be based on the *weight* of the location $i \in \{1, \dots, N\}$ denoted by w_i which is equal to $|r_i|$, the absolute value of the received vector at that location. Since large weight corresponds to lower probability of error in the sign detector, the estimates for these NR bits will be reliable.

In the next subsection, we describe how one can choose NR independent locations that have largest total weight and generate codewords based on these locations. This will give us the first candidate codeword. Extending the results, we will find a total of M candidate codewords. This list of codewords can be used in two ways. The first is to choose the codeword with largest posterior probability from the list as our sub-optimum estimate of the transmitted codeword. The second is to use this list rather than the one obtained through Viterbi decoding in the multiuser iterative detection and decoding algorithm of Section 2. We will also show that the complexity of this suboptimal scheme is much smaller than Viterbi algorithm.

3.1 Algorithm and proof of optimality

In this subsection, we find an algorithm to identify NR independent coded bits that we can estimate with largest reliability out of the length N received vector $\mathbf{r}$. Simply choosing largest weight NR locations will not work as not all NR bit subsets form an independent subset. A set $I \subset \{1, \dots, N\}$ is defined to be an *independent subset* of $\{1, \dots, N\}$ if the bits corresponding to locations in I can be chosen independently without violating any of the conditions for being part of a codeword in $\mathcal{C}$. We call the subset I to be *maximally independent* if I is an independent subset of $\{1, \dots, N\}$ and there does not exist any element e , which is in $\{1, \dots, N\}$ but not in I such that $I \cup \{e\}$ is also an independent subset. We will use three important properties of independent subsets:

1. For a convolutionally coded system of block length N

and rate R there are at most NR locations whose values can be chosen independently.

2. If $J' \subset J$, and J is an independent subset of $\{1, \dots, N\}$ then so is J' .
3. If $J \subset \{1, \dots, N\}$ then the cardinality of all maximally independent subsets of J are equal.

The first property, as we already observed, is due to the dimension of the code space. The second property results from definition of an independent subset. The third property follows because the cardinality of the maximally independent subset is equal to the rank of the linear system that describes it.

We will choose an independent subset I such that given the received vector $\mathbf{r}$ the *total weight* associated with I is maximum. That is, we need to find an independent subset such that $\sum_{i \in I} w_i$ is maximized. It is clear that I must be a maximally independent subset; otherwise we could add elements to I without violating the independence condition and thereby form another independent subset whose weight is larger than that of I .

The algorithm below obtains a maximally independent subset I of $\{1, \dots, N\}$ which has largest total weight.

Algorithm: Maximally independent subset

- . Set $I = \emptyset$
- . Given the received vector $\mathbf{r}$, sort the *weights* w_i of elements in $\{1, \dots, N\}$
- . While $|I| < NR$
 - Chose the location $e \notin I$ from $\{1, \dots, N\}$ with largest weight such that $I \cup \{e\}$ is still an independent subset of $\{1, \dots, N\}$
 - $I = I \cup \{e\}$

Lemma 1 *The above algorithm gives a maximally independent subset of $\{1, \dots, N\}$ with largest total weight.*

Proof: Suppose the above lemma is false. Let $I = \{e_1, e_2, \dots, e_k\}$ be the maximal subset given by the above algorithm and let $I' = \{q_1, q_2, \dots, q_k\}$ be a maximally independent subset with total weight larger than I . Since I and I' are both maximally independent subsets of $\{1, \dots, N\}$ their cardinality k are the same.

We assume $w_{e_1} \geq w_{e_2} \geq \dots \geq w_{e_k}$ and $w_{q_1} \geq w_{q_2} \geq \dots \geq w_{q_k}$. Of course there might be some elements common in I and I' . Chose the least index m such that $w_{q_m} > w_{e_m}$. Thus the element added at the m^{th} step is not one of $q_1, q_2, \dots, q_m$ and for $1 \leq i \leq m$, either

$q_i \in \{e_1, \dots, e_{m-1}\}$ or $\{e_1, \dots, e_{m-1}, q_i\}$ is not an independent subset of $\{1, \dots, N\}$.

In other words the set $\{e_1, \dots, e_{m-1}\}$ of cardinality $(m - 1)$ is a maximally independent subset of $\{e_1, \dots, e_{m-1}, q_1, \dots, q_m\}$. But the set $\{q_1, \dots, q_m\}$ of cardinality m is an independent subset of $\{e_1, \dots, e_{m-1}, q_1, \dots, q_m\}$. This contradicts the third property that the cardinality of all maximally independent subsets are equal. This proves that our assumption is wrong and our algorithm gives the desired maximally independent subset. $\square$

In a convolutional code, deciding whether a particular subset is independent or not and generating codewords from $\mathcal{C}$ based on an independent subset requires some extra observations. A convolutional code can be represented by a system of linear equations relating the coded bits to the information bits. For any equation involving n variables, as soon as $n - 1$ of these variables have been determined, the last variable can be found uniquely. Therefore, the locations corresponding to these n variables cannot form an independent subset.

Based on these observations, we next present an algorithm that selects the maximum weight independent subset along with the associated codeword.

Algorithm: Independent subset/codeword

- Set $I = \emptyset$
- Given the received vector $\mathbf{r}$, sort the *weights* w_i of elements in $\{1, \dots, N\}$
- While $|I| < NR$
 - Chose the location $e \notin I$ from $\{1, \dots, N\}$ with largest weight such that $I \cup \{e\}$ is still an independent subset of $\{1, \dots, N\}$. The received signal corresponding to the location e is r_e .
 - Set $\hat{d}_e = \text{sgn}(r_e)$.
 - Consider the convolutional code equations. In all the equations that location e appears, reduce the number of unknowns. If any of the equations has only one unknown left, solve for that unknown.
 - $I = I \cup \{e\}$

The independent subset selection together with codeword generation algorithm can easily be extended to choose M largest weight independent subsets. The suboptimum channel decoder then selects the codeword that has largest probability given the channel output $\mathbf{r}$ among these M candidate codewords. In a multiuser scenario, when we use the

iterative detector and decoder described in Section 3, we will use these M codewords together with their posterior probabilities not only to find a codeword estimate, but also to update priors in the next iteration step.

In order to compare the complexity of this suboptimum scheme with Viterbi algorithm, we note that the independent subset/associated codeword selection method involves sorting N positive numbers which can be done in linear time using the radix sort [1] algorithm. In a convolutional code with constraint length κ , a coded bit can be associated with at most $\kappa + 1$ equations. Thus for each step of the suboptimal channel decoder algorithm, we check at most $\kappa + 1$ equations. This ensures that for a single user the complexity of the suboptimum channel decoding algorithm is upper bounded by $O(MN\kappa)$ where M is the number of codewords in the list. Recall that the original Viterbi algorithm has a complexity of $O(N2^{2(\kappa+1)})$. Thus we have reduced the complexity from exponential in the constraint length to linear. We will also observe in the next subsection that performance loss due to this suboptimal algorithm is not significant.

3.2 Simulation results

We first compare the performance of the suboptimal channel decoding algorithm with Viterbi algorithm for one user. The convolutional code has rate $1/2$ and constraint length 3. We decode 100 information bits at a time. Our simulations have $M = 2$, that is we only have kept a codeword list of size 2 in the suboptimal algorithm. Our simulation results show there is very little performance loss in a single user case. We will observe that the difference becomes even less significant when we consider a multiuser system.

Figure 1: Suboptimal decoding algorithm versus Viterbi decoding

For the multiuser scenario, we have an asynchronous system with 4 users. The users are modulated by Gold-codes [5] of length 7. The interfering users have MAI 1-8 dB higher than the weakest user. We have used the suboptimal decoding in conjunction with the iterative detector and decoder of Section 2. After iteration 1, the difference between the Viterbi algorithm and suboptimal channel decoding becomes insignificant, thus we only show simulation results with the suboptimal decoder. We observe that the low complexity scheme manages to bridge the gap between matched filter decoding and single user bound.

Figure 2: Iterative detection and decoding algorithm with suboptimal channel decoders

4 Conclusion

In this study, we have proposed a computationally efficient multi-user joint detection and decoding scheme for a DS-CDMA system employing convolutional channel coding. The conventional Viterbi decoder has exponential complexity in the constraint length, and this prohibits a strong convolutional coding scheme from being implemented for real time applications. We have proposed a suboptimal decoding algorithm that has linear complexity in the constraint length. Through simulation, we have showed that the performance degradation in a single-user system is minimal. In a multi-user environment the suboptimal channel decoding used together with an iterative detector and decoder gives performance close to single user bound.

Acknowledgements

This work was supported by Nokia Inc., by the Texas Advanced Technology Program under grants 1995-#003604-049 and 1997-#003604-044, and by NSF under grants NCR 9506681 and CDA-9617383.

References

- [1] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. MIT Press, 1995.
- [2] S. Das, E. Erkip, J.R. Cavallaro, and B. Aazhang. Iterative Multiuser Detection and Decoding. In *Seventh Communication Theory Mini-Conference in conjunction with GLOBECOM'98*, Sydney, 1998. To Appear.
- [3] U. Fawer and B. Aazhang. Multiuser Receivers for Code-Division Multiple-Access Systems with Trellis-Based Modulation. *IEEE Journal on Selected Area in Communications*, 14(8):1602–1609, 1996.
- [4] T.R. Giallorenzi and S.G. Wilson. Multiuser ML Sequence Estimator for Convolutional Coded Asynchronous System. *IEEE Transaction on Communication*, 44(8):997–1008, August 1996.
- [5] R. Gold. Optimal Binary Sequences for Spread-spectrum Multiplexing. *IEEE Transactions on Information Theory*, pages 619–621, 1967.
- [6] J. G. Proakis. *Digital Communications*. McGraw-Hill, 1995.
- [7] C.B. Schlegel, M.C. Reed, P.D. Alexander, and J.A. Asentorfer. Reduced Complexity Iterative Multiuser Detection for DS/CDMA with FEC. In *Proceedings of ICUPC*, 1997.
- [8] S. Verdú. *Multiuser detection*. Cambridge University Press, 1998.
- [9] A. J. Viterbi. *CDMA Principles of Spread Spectrum Communication*. Addison-Wesley Publishing Company, 1995.