Blind Identification of the Scrambling Code of a Reverse Link CDMA2000 Transmission
Mathieu Des Noes, Valentin Savin, Jean-Marc Brossier, Laurent Ros

To cite this version:

HAL Id: hal-00796351
https://hal.archives-ouvertes.fr/hal-00796351
Submitted on 3 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Blind Identification of the Scrambling Code of a Reverse Link CDMA2000 Transmission

Mathieu des Noes and Valentin Savin  
CEA, LETI, Minatec campus  
Grenoble, France  
Email: {mathieu.desnoes,valentin.savin}@cea.fr

Jean Marc Brossier and Laurent Ros  
GIPSA-Lab, BP46, 38402 Saint-Martin d’Hères, France  
Email: {laurent.ros,jean-marc.brossier}@gipsa-lab.grenoble-inp.fr

Abstract—Interference between macro and femtocells is an important issue for the development of CDMA2000 femtocell networks. More specifically, the reverse link signal of a macro User Equipment may generate an unacceptable level of interference at the femto Base Station. To avoid this situation, interference mitigation techniques could be implemented. All the proposed techniques require to know the state of the scrambling code of the interferer in the reverse link. Unfortunately, it depends on the code mask of the terminal which is unknown by the femto BS. The femto BS has to estimate blindly the state of the scrambling code. An algorithm which performs a blind identification of the scrambling code of a CDMA2000 reverse link transmission is proposed in this article. This gives the possibility to implement interference cancelation algorithm at the femto BS.

INTRODUCTION

The development of femtocell networks is an important perspective for increasing cellular systems capacity. A femtocell is covered by a small Base Station (BS), designed for typically indoor environments (home, business) which does not require a coordinated deployment [1]. Although most of the current research activities on this topic focus on the LTE system [2], current deployed systems use the UMTS-WCDMA or CDMA2000 technologies [3]. In this paper, an algorithm which estimates blindly the state of the scrambling code of a terminal in the reverse link is proposed. It offers the perspective of implementing interference mitigation techniques at a femto Base Station (BS), and hence providing a solution to the problem of macro to femto interference issue in the reverse link of CDMA2000 femtocell networks.

A macro BS provides the overall coverage, while femto BSs offer better indoor coverage to UEs attached to them. A femto BS can be configured to operate in open or closed access mode to visiting UEs [4]. In open access mode, a visiting UE is allowed to handover from the macro BS to a femto BS in order to send its data. This generates additional complexity to route the data packets and also to ensure communication security. In closed access mode, a visiting UE is not allowed to handover. This simplifies the network architecture, but this may lead to an unacceptable interference level at the femto BS. This situation occurs if a visiting UE transmits at high power, while it is located nearby the femto BS. This situation is depicted in Fig. 1. A macro UE transmits at high power because it is either at the cell edge or inside a building. Since the power control procedure concerns only the link with the macro BS, it will be received at the femto BS with a power much larger than the power of a femto UE. This is the well known near-far effect in CDMA systems [5]. If the power of this interferer is too large, the femto BS may not be able to demodulate any communication with its attached UEs. This creates a “dead zone” in the network coverage. In order to avoid this situation, it is required to implement interference mitigation techniques. This subject has been deeply studied in the past two decades for CDMA systems [6]. All the proposed techniques exploit the knowledge of the UE’s scrambling code. More precisely, it is required to know the state of the UE’s scrambling code generator at the beginning of each frame.

The state of the scrambling code depends on its “code mask” which is used to generate a user specific delay of a long scrambling code common to all UEs. Unfortunately, in a closed access mode, there is no signalling link between the macro and femto BSs. Hence, a femto BS has no knowledge of the code mask of an interfering macro UE. It has to estimate the state of the scrambling code of a macro UE blindly. To the authors’ knowledge, [7] is the first and only article addressing this issue. The authors exploit the specificities of the framing, spreading and multiplexing procedures defined by CDMA systems to process the signal so that the resulting signal can be considered to be a linear code depending on the initial state of the scrambling code generator. This processing step can be applied to both WCDMA and CDMA2000 systems [8][9]. The received codeword is then estimated with a max-
log-MAP algorithm [10]. The initial state of the different spreading code generators is eventually obtained by applying a pseudoinverse of the code generator matrix. Concerning the CDMA2000 system, the solution proposed in [7] is adapted to a BPSK modulation which concerns Radio Configurations (RC) 3 to 6. However, it does not work for the 64-ary orthogonal modulation used by RC 1 and 2. These RCs are important because they are used by voice applications.

This paper addresses the 64-ary orthogonal modulation case and also proposes a different decoding strategy. Exploiting the unique properties of m-sequences, it is shown that an iterative message-passing algorithm [11] can be implemented to decode the received signal after an initial processing step. Then the initial state of the scrambling code can be obtained with a proper used of the “Shift and add” and decimation properties of m-sequences.

The paper is organized as follows. Section I presents the proper used of the “Shift and add” and decimation properties of m-sequences. Then, the method standardized for constructing sequences are built from m-sequences of period $2^{15} - 1 = 32767$ chips [12]. Their characteristic polynomials are:

$$g_I(D) = D^{15} + D^{13} + D^9 + D^8 + D^7 + D^5 + 1$$

$$g_Q(D) = D^{15} + D^{12} + D^{11} + D^{10} + D^9 + D^8 + D^4 + D^3 + 1$$

The sequences are complemented with a ‘0’ in order to create a frame of $N = 32768$ chips. According to the standard specifications, the initial state of each sequence is known by the receiver at the beginning of each frame.

In this article, the gating procedure associated with power control is not considered. This corresponds to a gating rate equals to 1 in the standard [9]. A smaller gating rate ($1/2$ or $1/4$) would require the implementation of parallel instances of the algorithm proposed in this article.

The real part of the complex signal is filtered by the pulse shaping filter $P(t)$, while the imaginary part is delayed by half a chip before pulse shaping. This implements an offset QPSK modulation. The specifications of $P(t)$ are detailed in [9]. The transmitted signal can eventually be modeled as follows:

$$X(t) = \sum_k V(k)G_k(t - kT_c)$$

with $G_k(t) = C_I(k)P(t) + jC_Q(k)P(t - T_c/2)$. $T_c$ is the chip period and is equal to 814 ns. $G_k(t)$ could be considered as a time-varying pulse shaping filter applied to sequence $V(k)$.

The 64-ary orthogonal modulation consists in mapping 6 consecutive binary data $d_i, d_{i+1}, \ldots, d_{i+5}$ onto a Walsh-Hadamard (WH) sequence of length 64 chips [9]. The WH sequence index is selected as follows:

$$m = d_i + 2d_{i+1} + 4d_{i+2} + 8d_{i+3} + 16d_{i+4} + 32d_{i+5}$$

Then each chip of the WH sequence $W_m$ is repeated 4 times and scrambled by the long code $S_n$. The subscript $n$ indicates that the long scrambling code is specific to user $n$. The scrambled signal corresponding to sequence $W_m$ is thus:

$$V(k) = S_n(k)W_m([k/4])$$

$[x]$ is the largest integer smaller than $x$. The chips of sequences $S_n$ and $W_m$ are anti-modal representation ($+1$,−1) of binary sequences $s_n$ and $w_m$.

The content of the $i$th shift register at time $k$ is noted $a_i(k)$. The scrambling code is generated by the modulo-2 inner
product of a 42-bit mask and the 42-bit state vector of the sequence generator:

\[ s_n(k) = \bigoplus_{i=0}^{r-1} a_i(k)m_n(i) \]

where \( \oplus \) denotes the XOR operation. \( m_n(i) \) is the code mask of the \( n \)th user. The vector \( a_0(k), a_1(k), \ldots, a_{r-1}(k) \) is known by the receiver at some specified time since it is signalled periodically by the network. On the other hand, the code mask is unknown to the receiver and thus it cannot generate sequence \( s_n \) with the schematic of Fig. 3.

Due to the “Shift and add” property of m-sequences, \( s_n \) is also a m-sequence with characteristic polynomial \( g_2(D) \) [12]. It is a shifted version of sequence \( a: s_n(k) = a(k + \tau_n) \). \( \tau_n \) is specific to the \( n \)th user. It depends on the code mask and the characteristic polynomial. Using the Fibonacci representation, \( s_n \) can be generated with the generator depicted in Fig. 4[12]. Instead of estimating the code mask, knowing the initial state of sequence \( a \), it is simpler to estimate the content of the shift registers of sequence \( s_n \) with the Fibonacci representation. This is the objective of the algorithm proposed in this paper.

![Code mask](image1)

**Fig. 3.** Long scrambling code generation

![LFSR sequence with the Fibonacci representation](image2)

**Fig. 4.** LFSR sequence with the Fibonacci representation

### II. BLIND IDENTIFICATION OF THE SCRAMBLING CODE

The algorithm is split in 3 steps:

- **step 1**: signal processing which aims at allowing the direct observation of a modified version of the sequence \( s_n \), denoted by \( \tilde{s} \).
- **step 2**: estimation of the initial state of sequence \( \tilde{s} \) with an iterative message-passing decoder.
- **step 3**: determination of the initial state of the sequence \( s_n \) with the use of a transposition matrix.

In order to ease the comprehension of the algorithm, the description is restricted to an Additive White Gaussian Noise (AWGN) channel model. The robustness to multipath channel will be discussed in section II-E.

#### A. Signal processing

The received signal is down-converted to baseband and sampled at rate \( T = T_e/E \), where \( E \) is the oversampling factor.

It is modeled as follows (see Eq. (2)):

\[ R(q) = e^{j\theta} \sum_k V(k)G_k(qT - kT_e) + N(q) \] (3)

where \( \theta \) is the phase rotation introduced by the channel and \( N(q) \) the additional noise modeling thermal noise as well as other sources of interference. The objective is to obtain an observation of a modified version of the sequence \( s_n \), and thus get rid off \( G_k(t) \), \( \theta \) and sequence \( W_n \). The pulse shaping filter defined by the standard generates inter-chip interference. As a consequence, a simple matched filter is not sufficient and an equalizer shall be implemented. Hence, the first step consists in implementing an equalizer to the time varying pulse shaping function \( G_k(t) \). Let us define this equalizer as a FIR filter with \( L \) coefficients: \( F_k(0), \ldots, F_k(L - 1) \). It can be implemented either with a zero-forcing (ZF) or a minimum mean square error (MMSE) criterion [14].

If the receiver is synchronized with the transmit signal, after equalization and downsampling, the signal is well approximated by:

\[ Y(k) = \sum_{l=0}^{L-1} R(kE + L/2 - l)F_k(l) = e^{j\theta}V(k) + N_{eq}(k) \] (4)

\( N_{eq}(k) \) is the filtered noise.

If the receiver is not synchronized, the equalizer output can be regarded as a random noise sequence. This is due to the correlation properties of m-sequences \( c_l \) and \( c_q \) [12].

The second step consists in removing the dependency upon sequence \( W_n \) and phase \( \theta \). Since, the chips of sequence \( W_m \) are repeated 4 times, a differential multiplication is performed within each window of 4 repeated chips. For instance, the differential multiplication using only the first two chips of the 4 chips length window will give:

\[ \begin{align*}
Y(4k+1)Y(4k)^* & = V(4k+1)V(4k) + N_{diff}(k) \\
& = S_n(4k+1)S_n(4k) + N_{diff}(k)
\end{align*} \]

where \( N_{diff}(k) \) contains all the noise cross-product terms.

In fact there are 7 possible differential products:

\[ U_{i,j}(k) = Y(4k+i+j)Y(4k+i)^* = S_n(4k+i+j)S_n(4k+i) + N_{i,j}(k) \] (5)

\[ i = 0, 1, 2 \text{ and } j = i + 1, \ldots, 3. \] Let us define the sequence \( \hat{S}_{i,j}(k) = S_n(4k+i+j)S_n(4k+i) \), which is rewritten in \( GF(2) \) by: \( \hat{S}_{i,j}(k) = s_n(4k+i+j) \oplus s_n(4k+i) \).

\( U_{i,j} \) is thus a noisy observation of sequence \( \hat{S}_{i,j} \). The objective is now to find the relationship between sequences \( \hat{S}_{i,j} \) and \( s_n \).

To do so, two interesting properties of m-sequences will be exploited [12]:

---

\( N_{eq}(k) \) is the filtered noise.

If the receiver is not synchronized, the equalizer output can be regarded as a random noise sequence. This is due to the correlation properties of m-sequences \( c_l \) and \( c_q \) [12].

The second step consists in removing the dependency upon sequence \( W_n \) and phase \( \theta \). Since, the chips of sequence \( W_m \) are repeated 4 times, a differential multiplication is performed within each window of 4 repeated chips. For instance, the differential multiplication using only the first two chips of the 4 chips length window will give:

\[ \begin{align*}
Y(4k+1)Y(4k)^* & = V(4k+1)V(4k) + N_{diff}(k) \\
& = S_n(4k+1)S_n(4k) + N_{diff}(k)
\end{align*} \]

where \( N_{diff}(k) \) contains all the noise cross-product terms.

In fact there are 7 possible differential products:

\[ U_{i,j}(k) = Y(4k+i+j)Y(4k+i)^* = S_n(4k+i+j)S_n(4k+i) + N_{i,j}(k) \] (5)

\[ i = 0, 1, 2 \text{ and } j = i + 1, \ldots, 3. \] Let us define the sequence \( \hat{S}_{i,j}(k) = S_n(4k+i+j)S_n(4k+i) \), which is rewritten in \( GF(2) \) by: \( \hat{S}_{i,j}(k) = s_n(4k+i+j) \oplus s_n(4k+i) \).

\( U_{i,j} \) is thus a noisy observation of sequence \( \hat{S}_{i,j} \). The objective is now to find the relationship between sequences \( \hat{S}_{i,j} \) and \( s_n \).

To do so, two interesting properties of m-sequences will be exploited [12]:

---

\( N_{eq}(k) \) is the filtered noise.

If the receiver is not synchronized, the equalizer output can be regarded as a random noise sequence. This is due to the correlation properties of m-sequences \( c_l \) and \( c_q \) [12].

The second step consists in removing the dependency upon sequence \( W_n \) and phase \( \theta \). Since, the chips of sequence \( W_m \) are repeated 4 times, a differential multiplication is performed within each window of 4 repeated chips. For instance, the differential multiplication using only the first two chips of the 4 chips length window will give:

\[ \begin{align*}
Y(4k+1)Y(4k)^* & = V(4k+1)V(4k) + N_{diff}(k) \\
& = S_n(4k+1)S_n(4k) + N_{diff}(k)
\end{align*} \]

where \( N_{diff}(k) \) contains all the noise cross-product terms.

In fact there are 7 possible differential products:

\[ U_{i,j}(k) = Y(4k+i+j)Y(4k+i)^* = S_n(4k+i+j)S_n(4k+i) + N_{i,j}(k) \] (5)

\[ i = 0, 1, 2 \text{ and } j = i + 1, \ldots, 3. \] Let us define the sequence \( \hat{S}_{i,j}(k) = S_n(4k+i+j)S_n(4k+i) \), which is rewritten in \( GF(2) \) by: \( \hat{S}_{i,j}(k) = s_n(4k+i+j) \oplus s_n(4k+i) \).

\( U_{i,j} \) is thus a noisy observation of sequence \( \hat{S}_{i,j} \). The objective is now to find the relationship between sequences \( \hat{S}_{i,j} \) and \( s_n \).

To do so, two interesting properties of m-sequences will be exploited [12]:
The successive steps of the algorithm are the followings. At section II-B). The vector finds a valid codeword, the initial state of sequence is restarted when the next sample is received. If the decoder or a missed detection happened. In both cases, the procedure is matched to the characteristic polynomial of sequence according to Eq. (4). Then a differential multiplication is with the beginning of the frame and equalizes the input signal by a simple matrix multiplication (see section II-C).

A m-sequence is thus a cyclic linear code with rate \( n \). It is thus possible to estimate the initial state of sequence \( s_n \) using its characteristic polynomial \( g(D) \). Then, the receiver assumes that it is synchronized \( \tau \) such that:

\[
\tau \times D = \tau \quad \text{such that:}
\]

\[
(\tau)_k \equiv g_{\tau} g_{\tau-1} \cdots g_0 \equiv 0
\]

These 2 properties allow us to conclude that \( \tilde{s}_{i,j} \) is a shifted version of sequence \( s_n \). As a consequence, \( \tilde{s}_{i,j} \) has the same characteristic polynomial as \( s_n \): \( g_n(D) \).

The algorithm’s principle is thus to decode the initial state of sequence \( \tilde{s}_{i,j} \) and to recover the state of sequence \( s_n \) using a standard iterative message-passing algorithm exploiting the sequence’s characteristic polynomial \([11][15][16]\). It is thus possible to estimate the initial state of sequence \( \tilde{s}_{i,j} \) using its characteristic polynomial \( g_s(D) \).

The second step eventually provides vector \( A_{s_n} \). The decoder provides a vector of \( r \) bits representing the initial state of sequence \( \tilde{s}_{i,j} \), denoted by \( A_5 \). We want to find the initial state of sequence \( s_n \) at the beginning of the frame: \( A_{s_n} = (s_n(0), \ldots, s_n(41)) \).

The first step consists in finding the initial state of the sequence \( s_{\text{decim}} \), which gives \( \tilde{s}_{i,j} \) by decimation by a factor 4:

\[
\tilde{s}_{i,j}(k) = s_{\text{decim}}(4k)
\]

[19] shows that there is a fixed transposition matrix \( B \) between the state vectors of a m-sequence and its decimated by 2 version, and detailed the procedure to compute this matrix \( B \). As a consequence:

\[
A_{s_{\text{decim}}} = B^2 A_5
\]

The second step eventually provides vector \( A_{s_n} \). It exploits the “shift and add” relation between the sequences \( s_{\text{decim}} \) and \( s_n \):

\[
s_{\text{decim}}(k) = s_n(k+i+j) \oplus s_n(k+i)
\]

Using the state transition matrix \( G \), defined by [13]:

\[
G = \begin{bmatrix}
0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1 \\
1 & g_{r-1} & g_{r-2} & g_{r-3} & \cdots & g_1
\end{bmatrix}
\]
we have:

\[ A_{x_n} = (G^{i+j} + G^j)^{-1} A_{\text{decim}} \]

These two steps are finally combined in a single transposition matrix \( T \):

\[ A_{x_n} = T A_s \]

\[ T = (G^{i+j} + G^j)^{-1} B^2 \]

It is important to note that matrix \( T \) could be computed off-line and stored in memory.

**D. Flow chart of the algorithm**

A flow chart of the algorithm is presented in Fig. 5. The decoding operation must be implemented at the chip rate, which requires a very high data rate decoding capability.

---

\[ \hat{x}_n = \{ 1 \text{ if all parity check equations are satisfied} \}
\]

\[ 0 \text{ otherwise} \]

\( \hat{x}_n \) is the indication function of the decoder.

\( A_{x_n} \) is the estimated initial state of sequence \( x_n \), given by the decoder output.

Performances are measured with the following simulation configurations:

- When measuring \( P_n \), the receiver is synchronized with the beginning of the frame (i.e. hypothesis \( H_1 \) is satisfied), while its is not synchronized when \( P_{fa} \) is evaluated.
- The decoder implements a Self-Corrected Min-Sum (SCMS) message-passing algorithm [20], which provides a very high data rates decoding capability, while performing close to the “optimal” Belief Propagation (Sum-Product) decoding. The decoder stops when either all the parity check equations are satisfied or the maximum number of iteration \( N_{iter} \) is reached.
- The number of variable is \( M = 1500 \) at the decoder input. This value is selected so that it is smaller than the size of a Power Control Group (PCG) which contains 1536 chips [9]. A PCG is equivalent to a slot in the WCDMA system.
- The channel is AWGN.

Simulations over \( 10^7 \) frames did not produce any false alarm. This means that \( P_{fa} < 10^{-6} \) with a good confidence level. Fig. 6 shows the probability of missed detection for a chip level MMSE, ZF or MRC equalizer. The number of RGMs is set to its largest possible value \( n_{RGM} = 6 \) for \( M = 1500 \) variables and \( N_{iter} = 100 \). \( P_m \) is measured as a function of the Signal to Noise Ratio (SNR) at the input of the receiver. At \( P_m = 0.1 \), the MMSE equalizer gives a gain of almost 1 dB and 3 dB with respect to ZF and MRC. Fig. 7 shows the probability of missed detection as a function of the number of RGMs for a receiver configuration with chip level MMSE equalizer. It is not worth increasing the number of RGMs above 4.

According to [21], the target Eb/No at a base station is around 7 dB for the 9.6 kbps service which corresponds to our simulation configuration (all the PCGs are transmitted). Since the combined coding-spreading gain equals 128, the target SNR at a base station is around \(-14\) dB. According to Fig. 7, the algorithm implementing a chip level MMSE equalizer detects the scrambling code of an interfering user at a SNR of 2.5 dB. This corresponds to a very strong interferer since its received power is 16.5 dB larger than a desired user attached to the femtocell. However, the performance of the algorithm is not good enough to detect certain lower but still strong interferes. For instance, an interferer which power is 15 dB larger than the desired user will not be detected even it will have very damaging effects. The bad detection performance of the algorithm is due to the weight of the characteristic polynomial of the scrambling code generator. This effect has been well studied in the framework of fast...
correlation attacks [16]. The weight equals 20 while it should be below 4 or 5 to give good decoding performance [22]. One way to improve performance is to generate a parity check matrix which row weight is below 4 or 5. This could be achieved by implementing the technique proposed in [23].

false alarm probability is below $10^{-6}$ for all the tested SNR values.

**REFERENCES**


**IV. CONCLUSION**

An algorithm which performs a blind identification of the scrambling code of a reverse link CDMA2000 transmission has been presented. It is dedicated to the 64-ary orthogonal modulation used by the radio configurations 1 and 2 of the standard. The simulation results show that with a BP decoding algorithm it is possible to obtain a reliable estimation at a SNR of 2.5 dB in an AWGN channel. It is also noticeable that the