
Susanna Spinsante(*), Madhyiar Sarayloo(*), Ennio Gambi(*), Chirag Warty(**), Claudio Sacchi(***)

(*) Dipartimento di Ingegneria dell’Informazione, Università Politecnica delle Marche, Ancona (Italy)
(**) Intelligent Communication Lab, Mumbai (India)
(***) Dept. of Information Engineering and Computer Science (DISI), University of Trento, Trento (Italy)
Outline

- Introduction and aims of the paper;
- Efficient generation of large sets of De Bruijn sequences;
- De Bruijn sequences properties;
- Statistical analysis of DS/CDMA system performance;
- Numerical results;
- Conclusion.
Introduction

- **Random (or quasi-random) spreading sequences for DS/CDMA**
  - Direct Sequence Code Division Multiple Access (DS/CDMA) still represents a core technology for the physical layer of commercially remunerative applications and standards (radiolocalization, automotive radar, 3G UMTS);

- A very critical issue of Spread Spectrum and CDMA: keeping the probability of intercept the lowest possible;

- Secure information hiding must be guaranteed at the physical layer level: random spreading sequences should be applied;

- Due to complexity of generating truly random sequences, deterministic sequences (i.e. pseudorandom) are used in real applications;

- Required features: pseudo-noise auto-correlation patterns, quasi-orthogonal cross-correlation.
Introduction

- **Gold and De Bruijn sequence sets**
  - Typical choice: **Gold codes**, generated as logical combination of linear shift register (LSR) sequences (*preferred pairs*) of span $n$ (= number of LSR cells);
  - Gold codes features:
    - favorable statistical properties;
    - small cardinality $= N + 2$, where $N$ (sequence length) $= 2^n - 1$
  - In the literature, the alternative use of **De Bruijn** binary sequences for DS/CDMA has been recently proposed [SPI11]. Their features are:
    - generation by nonlinear shift register;
    - maximal length ($N = 2^n$);
    - very large cardinality $2^{2(n-1)-n}$
    - interesting correlation-related features [AND10, SPI11, SPI13, WAR13, SAR14]
Aims of the paper and advancement with respect to related work

- Propose an efficient sequence generation algorithm based on De Bruijn graphs theory and Eulerian cycles;

- **Formal statistical analysis of De Bruijn sequences in DS/CDMA with explicit computation of 2\textsuperscript{nd} and 4\textsuperscript{th} order statistics (variance and normalized kurtosis) of multi-user interference (MUI), in asynchronous BPSK-modulated DS/CDMA transmission;**

- **Closed form computation of average bit-error-probability (BEP):**
  - by Gaussian approximation [PUR76]
  - by non-Gaussian evaluation [TES99], based on the Generalized Gaussian modeling of the global detection noise affecting the CDMA receiver (Gaussian noise + MUI)

- MUI statistics and BEP performance comparison to Gold codes with and without code selection driven by a formal criterion.

At the end of this analysis we may have more insights about the use of De Bruijn sequences in real DS/CDMA systems.
Efficient generation of large sets of De Bruijn sequences

**DEF:** In a binary De Bruijn sequence viewed cyclically over a period, each binary n-tuple appears exactly once, including the all-zero n-tuple, due to the non-linear nature of the generating register.

- **Generation by Non Linear Feedback Shift Registers (NLFSRs)**
- NLFSR state at time $t$: $s(t) = (s_1(t), s_2(t), ..., s_3(t))$, $s_i(t) \in A = \{0, 1\}, \text{for } i = 1, 2, ..., n$
  
  where

  - At each clock transition:
    - each memory cell content shifted one position to the right
    - leftmost cell $s_n(t)$ updated by the output of a nonlinear feedback function $g(.)$
    - $g(.)$ defines a mapping of $A^n \rightarrow A$

- At time $(t+1)$, the state of the register is given by: $s_i(t + 1) = \begin{cases} s_{i+1}(t), & \text{for } i = 1, 2, ..., n - 1 \\ g(s(t)), & \text{for } i = n \end{cases}$
Efficient generation of large sets of De Bruijn sequences

- Various generation methods proposed in the literature:
  - use of a lower-order and n-bit pattern initial stage [TUR11]
  - prefer-one, prefer-opposite [ALH10], and prefer-same [FRE82] approaches for bit insertion
  - n-stage FSRs [CHA90, ZHA09]

- To increase time efficiency, De Bruijn graphs theory and Eulerian cycles are here exploited to ignore a large number of sequences that do not verify the definition given above:
  - each acceptable sequence starts from an arbitrary vertex and walks through the graph, by crossing each degree no more than once;
  - a rotated sequence to either the right or the left is not a different De Bruijn sequence;
  - the bitwise not of each De Bruijn sequence originates a distinct De Bruijn sequence of the same family: generation of half the set allows to obtain the whole family;
  - decimal representation of the array used to store the sequences.

A De Bruijn graph: every four-digit sequence occurs exactly once if one traverses every edge exactly once and returns to one's starting point (an Eulerian cycle)
Generating algorithm pseudocode:

1. Parameters initialization: \( n \), \( L_{Seq} \) (length of sequence), \( N_{Seq} \) (number of distinct sequences), \( T_{Seq} \) (matrix to store the generated De Bruijn sequences)
2. Set \( Poss_{Seq} \) to \( \{1, \{0\}^n, 1\} \) or \( \{0, \{1\}^n, 0\} \)
3. Calculate \( Dir_{Vectors} \) according to \( Poss_{Seq} \)
4. Calculate \( Next_{Num} \) according to \( Dir_{Vectors} \)
5. Loop (\# generated sequences < \( N_{Seq} \))
   1. Calculate next possible vertex w.r.t. \( Dir_{Vectors} \)
   2. Update \( Poss_{Seq} \) according to the next possible vertex
   3. Update \( Next_{Num} \) according to \( Dir_{Vectors} \)
   4. If (achieved sequence meets De Bruijn definition) then
      - Calculate bitwise NOT of the generated sequence
      - Rotate the generated sequence and its complementary such that it starts with \( 0^n \)
      - Store the decimal value of both the generated sequence and its complementary one, in \( T_{Seq} \)
6. End Loop

Generation time for different span:

<table>
<thead>
<tr>
<th>Span</th>
<th>Time (sec.)</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (sec.)</td>
<td>0.263</td>
<td>0.416</td>
<td>70.764</td>
<td>≈ 4 days</td>
</tr>
<tr>
<td># generated sequences</td>
<td>2</td>
<td>16</td>
<td>2048</td>
<td>4000000</td>
<td></td>
</tr>
<tr>
<td># sequences</td>
<td>2</td>
<td>16</td>
<td>2048</td>
<td>67108864</td>
<td></td>
</tr>
</tbody>
</table>

Sequence sets: length and cardinality comparison:

<table>
<thead>
<tr>
<th>( m )-sequences</th>
<th>Gold</th>
<th>De Bruijn</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>length</td>
<td># seq.</td>
</tr>
<tr>
<td>5</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>63</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>127</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>255</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>511</td>
<td>48</td>
</tr>
<tr>
<td>10</td>
<td>1023</td>
<td>60</td>
</tr>
</tbody>
</table>
De Bruijn sequences properties

- Periodic auto-correlation $C_{aa}^P[k]$ of span n De Bruijn sequence $a$ for a shift $k$:
  - $C_{aa}^P[k] = 2^n$, for $k = 0$
  - $C_{aa}^P[k] = 0$, for $1 \leq |k| \leq n - 1$ (Zero Correlation Zone)
  - $C_{aa}^P[k] \neq 0$, $|k| = n$
  - $C_{aa}^P[k] \equiv 0 \pmod{4}$, $\forall k, n \geq 2$

- Bound on periodic auto-correlation sidelobes values:
  - $0 \leq \max C_{aa}^P[k] \leq 2^n - 4 \left[ \frac{2^n}{2n} \right]^+$, $1 \leq k \leq N - 1, N = 2^n$

Bound on $\max C_{aa}^P[k]$ sidelobe value for $5 \leq n \leq 10$

<table>
<thead>
<tr>
<th>span n</th>
<th>length N</th>
<th>null samples around peak</th>
<th>bound on $\max C_{aa}^P[k]$</th>
<th>ratio $\frac{\max C_{aa}^P[k]}{N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>32</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 4$</td>
</tr>
<tr>
<td>6</td>
<td>64</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 5$</td>
</tr>
<tr>
<td>7</td>
<td>128</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 6$</td>
</tr>
<tr>
<td>8</td>
<td>256</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 7$</td>
</tr>
<tr>
<td>9</td>
<td>512</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 8$</td>
</tr>
<tr>
<td>10</td>
<td>1024</td>
<td>$1 \leq</td>
<td>k</td>
<td>\leq 9$</td>
</tr>
</tbody>
</table>
De Bruijn sequences properties

- Cross-correlation function $C_{a_1a_2}[k]$ for a shift $k$:
  
  $C_{a_1a_2}[k] = C_{a_1a_2}[N - k], 0 \leq k \leq N - 1$
  
  $\sum_{k=0}^{N-1} C_{a_1a_2}[k] = 0$
  
  $C_{a_1a_2}[k] \equiv 0(\text{mod } 4), n \geq 2, \forall k$

- Bound on cross-correlation sidelobes values:
  
  $-2^n \leq C_{a_1a_2}[k] \leq 2^n - 4, 0 \leq k \leq N - 1$

<table>
<thead>
<tr>
<th></th>
<th>Max abs. value</th>
<th>Mean</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$-sequence</td>
<td>11</td>
<td>0.032258</td>
<td>5.65391</td>
</tr>
<tr>
<td>Gold31</td>
<td>9</td>
<td>-0.0447</td>
<td>5.4064</td>
</tr>
<tr>
<td>De Bruijn</td>
<td>32</td>
<td>0</td>
<td>6.0703</td>
</tr>
</tbody>
</table>

Maximum absolute value, mean, and standard deviation of the cross-correlation, for De Bruijn, Gold, and $m$-sequences of span $n = 5$
De Bruijn sequences properties

- Randomness analysis: **Golomb’s postulates [GOL82]**
  - 1st and 2nd postulates (balance and run properties): always verified;
  - 3rd postulate (ideal 2-level auto-correlation): not verified BUT a Zero Correlation Zone is exhibited;

- Linear complexity ($C$) = estimated length of the shortest LFSR which would be able to generate the sequence itself;

- According to Berlekamp [BER68], $C$ provides a numerical description of the amount of information needed to infer the structure of the spreading codes generation algorithm

- De Bruijn sequences: $2^{n-1} + n \leq C \leq 2^n - 1$

- Gold codes: $C = n$
Received multi-user DS/CDMA signal (1):

- From the channel (supposed to be AWGN):
  \[ r(t) = \sqrt{2P} \sum_{k=1}^{K} b_k (t - \tau_k) a_k (t - \tau_k) \cos (2\pi f_0 t + \theta - \phi_k) + z(t) \]

- After coherent demodulation and de-spreading (ref. user 1), sampled at \( t = T \):
  \[ R = \sqrt{\frac{P}{2}} T b_{1,0} + \sqrt{\frac{P}{2}} T \left( \frac{1}{N} \sum_{k=2}^{K} I_{k,1} \right) + \xi \]

\[
I_{k,1} = \left\{ \begin{array}{ll}
\chi_{k,1} (\alpha_k) + \left[ \chi_{k,1} (\alpha_k + 1) - \chi_{k,1} (\alpha_k) \right] v_k \cos (\phi_k) \\
\bar{\chi}_{k,1} (\alpha_k) = \begin{cases}
C_{a_k, a_1} (\alpha_k) & \text{if } b_{k,-1} = b_{k,0} \\
\hat{C}_{a_k, a_1} (\alpha_k) & \text{if } b_{k,-1} \neq b_{k,0}
\end{cases}
\end{array} \right.
\]

Multi-User Interference (MUI) term

\[
\alpha_k T_c \leq \tau_k < (\alpha_k + 1) T_c \quad v_k = (\tau_k - \alpha_k T_c) / T_c
\]

Even and odd PN cross correlations
Statistical analysis of DS/CDMA system performance

- **Received multi-user DS/CDMA signal (2):**

  - More in details [PUR76, TES99]:
    
    \[
    C_{a_k,a_l}(\alpha_k) = \Psi_{a_k,a_l}(\alpha_k) + \Psi_{a_k,a_l}(\alpha_k - N)
    \]
    
    \[
    \hat{C}_{a_k,a_l}(\alpha_k) = \Psi_{a_k,a_l}(\alpha_k) - \Psi_{a_k,a_l}(\alpha_k - N)
    \]

  - Considering a BPSK modulation and deterministic (known) spreading sequences, the BEP computation is as follows:

    \[
    P_{be} = \Pr \left\{ \frac{error}{b_{k,0}} = -1 \right\} = \Pr \left\{ R > 0 \right\} = \Pr \left\{ \left[ \frac{P}{2} \left( \frac{1}{N} \sum_{k=2}^{K} I_{k,1} \right) + \xi \right] > \sqrt{\frac{P}{2}} \right\}
    \]

    Practically: \[
    P_{be} = \int_{\sqrt{P/2T}}^{+\infty} f_{Z_G}(z) dz \quad \text{where:} \quad Z_G = \sqrt{\frac{P}{2T}} \left( \frac{1}{N} \sum_{k=2}^{K} I_{k,1} \right) + \xi
    \]
How can we NUMERICALLY compute DS/CDMA BEP?

- In other words: can we express in closed form the probability density function of the random variable $Z_G$?
- The answer is NO, therefore, we should resort to some approximation:
  - **Gaussian Approximation (GA)**: it simply considers a Gaussian distribution for $Z_G$. It is reasonable when the number of users is large [PUR76];
  - **Generalized Gaussian Approximation (GG)**: as the pdf of MUI for real-valued binary sequence has an impulsive pseudo-Laplace distribution (*leptokurtic*), we can suppose that the pdf of $Z_G$ fits well with the Generalized Gaussian pdf model [TES99], expressed in terms of its normalized kurtosis:

$$f_{Z_G}(z) = \frac{c\gamma}{\Gamma(1/c)} \exp\left(-\left|\gamma z\right|^c\right)$$

$$\kappa(Z_G) \approx \frac{E(Z_G^4)}{E(Z_G^2)} = \frac{3}{4} \left(\frac{E_b}{\eta}\right)^{-2} + \frac{E(I^4)}{N^4} + 3\left(\frac{E_b}{\eta}\right)^{-1} \frac{E(I^2)}{N^2}$$

$$\gamma = \sqrt{\frac{\Gamma(3/c)}{\Gamma(1/c) \text{var}(Z_G)}}$$

$$c = F(\kappa_{Z_G}) \approx \sqrt{\frac{5}{\kappa_{Z_G}-1.865}} - 0.12$$

$$2 < \kappa_{Z_G} < 10$$

(an alternative, more precise expression of $F$, valid for a wider range of values of the normalized kurtosis is in eq.30 of the paper)

$\Gamma$ = Euler’s Gamma function

$\kappa = 3$ and $c=2$ for Gaussian-distributed r.v.
Approximated analytical expressions for DS/CDMA BEP:

- Using **GA approximation**, BEP is given as follows:

\[
P_{be} \approx Q\left(\sqrt{\text{SINR}}\right)
\text{SINR} \triangleq \left(\frac{E(I^2)}{N^2} + \frac{\eta}{2E_b}\right)^{-1}
\]

- Using **GG approximation**, we obtain after some mathematical manipulations:

\[
P_{be} \approx \frac{1}{2} - \frac{1}{2} \Gamma_{inc} \left(\frac{1}{c} \Gamma \left(\frac{3}{c}\right) \left(\text{SINR} \right)^{c/2}, \frac{1}{c}\right)
\]

\[
\Gamma_{inc} (x, s) = \frac{1}{\Gamma(s)} \int_0^x t^{(s-1)} e^{-t} dt
\]

«Incomplete» Gamma function
Numerical results

- **Random sequence selection**
  - This really means: no selection criterion applied, random indices of the De Bruijn matrix have been used to select the sequences;
  - GA and GG approximations have been compared with tight upper and lower bounds on DS/CDMA BEP computed as in [LEH89];
  - BPSK modulation with Reed-Solomon coding (RS(31,23)) have been considered in deriving numerical results (an analytical lower bound on BER is available for RS coding);
  - $N=32$ and $K=4$ users have been considered in BEP computations.

**MUI Statistics**

- $E(I^2)/N^2$ and $\kappa(I)$

<table>
<thead>
<tr>
<th></th>
<th>De Bruijn</th>
<th>Gold</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(I^2)/N^2$</td>
<td>0.0349</td>
<td>0.0306</td>
</tr>
<tr>
<td>$\kappa(I)$</td>
<td>3.56</td>
<td>3.43</td>
</tr>
</tbody>
</table>

Gold sequences performs better than De Bruijn ones thanks to their superior “Gaussianity” (the GA curve is closer to GG and upper and lower bounds)
Numerical results

- **Making things smarter: sequence selection criterion** ($K$ sequences selected, for each value of $N$)
  
  i. all the sequences in each set are assessed for their minimum aperiodic auto-correlation sidelobe $C_{a_k,a_1}(\alpha_k)$
  
  ii. looking at the lowest minimum aperiodic auto-correlation sidelobe values found in i), the subset featuring the lowest sidelobe value joint a number of sequences $K$ is selected;
  
  iii. $K$ sequences are extracted from the subset obtained in ii), by looking at sequence pairs featuring the most favorable aperiodic cross-correlation;
  
  iv) if it is not possible to find a close subset of $K$ sequences as per iii), they are selected randomly over the subset obtained in ii);

Number of groups of non-duplicated $K$ sequences out of $M$ (cardinality of the set):

\[
G_{M,K} = \binom{M}{K} = \frac{M!}{K!(M-K)!}
\]

Check the whole family for sequences featuring the **MINIMUM** auto-correlation sidelobe

select the subset $S$ for which:

(lowest auto-correlation sidelobe) **AND**

(# seq. $\geq K$)

select $K$ sequences from $S$ with best aperiodic cross-correlation

**OR**

randomly select $K$ sequences in $S$
Numerical results

- **Sequence selection**
  - Minimum aperiodic cross-correlation sidelobe criterion;
  - Things are changing: selected De Bruijn sequences decreases both variance and normalized kurtosis of MUI;
  - As result, BEP is noticeable decreased with respect to Gold sequences (the selection criterion is not effective for small Gold sets).
Conclusion

- Performance of binary De Bruijn sequences assessed, as spreading codes in multiple users DS/CDMA systems, through a formal statistical analysis of link performance, in comparison with traditionally used Gold codes;

- The formal statistical analysis shows that De Bruijn codes exhibit performance comparable to Gold codes and even worse if no selection criterion is applied;

- On the other hand, the selection criterion based on the minimization the pairwise aperiodic cross-correlation among the sequences associated to different users may lead to remarkably improved performance of De Bruijn sequences;

- The much greater cardinality, and better randomness-related properties of De Bruijn sequences, could anyway improve the robustness of the communication system against interception or security attacks.
References


