

## COMPARISON OF CONTIGUOUS AND DISTRIBUTED FRAME SYNCHRONIZATION SEQUENCES USING STATISTICAL BIFIX APPROACH

Čedomir Stefanović, Dragana Bajić, Faculty of Technical Sciences, University of Novi Sad

**Abstract** – Short introduction to statistical bifix approach used to evaluate the expected duration of the search for frame synchronization sequence is presented. Contrary to the previous contributions, this is a general case when synchronization symbols are not equiprobable. Based on this, a comparison between some best-known contiguous and distributed sequences is made.

### 1. INTRODUCTION

It is a well-known fact that one the fundamental requirements in digital communication systems is to achieve the synchronization between transmitter and receiver on several levels. In this paper, synchronization on the frame level is considered, which is needed for the proper interpretation of received data.

Frame synchronization is accomplished by regular insertion of so-called synchronization sequence (pattern) in data stream. First methods dealt with the contiguous sequences ([1], [2], [3], and [4]), while recently interesting approach based on distributed sequences was introduced in [5].

Generally, both approaches seek for the synchronization sequences that have optimum properties in terms of aperiodic correlation of the sequence with the shifted version of itself. The basic idea is that a sequence should have as small as possible correlation “sidelobes” (values of correlation for non-zero shifts). It can be easily shown that such condition minimizes the probability of false interpretation of the received data as a synchronization sequence [6].

Another criterion by which the comparison of synchronization sequences can be made is the expected duration of the search for synchronization sequence in received data. From one point of view, this time should be as short as possible, which relates to the situation when the receiver is in the mode of acquisition of synchronization. From another point of view, the longer expected duration of search process means the smaller probability of accident appearance of the same sequence between two periodic insertions of synchronization pattern. In [7] the statistical bifix approach was introduced that dealt with expected duration of search process. Here was shown that, for contiguous sequences with equiprobable symbols, the expected duration of the search process depends on probability of synchronization sequence seen as random data and its structure – the bifices. (The sequence has the bifix of  $i$ -th order, if its first  $i$  and last  $i$  symbols are the same, i.e. they are both prefix and suffix of the sequence). It is interesting to notice that, in contrast to sequences with bifices, bifix-free sequences have better correlation properties, while they also have shorter expected search time.

The organization of this paper is as follows. Section 2 gives a short introduction to the generalization of statistical bifix approach, both for contiguous and distributed sequences

with non-equiprobable symbols. Section 3 briefly recapitulates the best-known contiguous and distributed sequences. In section 4 comparisons between these synchronization sequences are made, based on the expected duration of the search process. Results are commented and concluding remarks are given in section 5.

### 2. GENERALIZED STATISTICAL BIFIX APPROACH

Let us consider (instead of one synchronization sequence) a set of synchronization sequences that are all considered valid by the receiver. For instance, this pertains to the case when receiver tolerates errors in the synchronization sequence due to the transmission. Another example is the usage of distributed synchronization sequences, when the synchronization symbols together with random data symbols comprise the whole set of sequences. We will denote the set of  $M$  allowed synchronization sequences by  $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\}$ . All sequences are of length  $N$ .

Next we introduce the cross-bifices. Sequences  $\mathbf{p}_i$  and  $\mathbf{p}_j$  have cross-bifix of  $m$ -th order, if the last  $m$  symbols of sequence  $\mathbf{p}_i$  and first  $m$  symbols of sequence  $\mathbf{p}_j$  are identical. We denote this by cross bifix indicator  $h_{ij}^{(m)} = 1$ . If previous statement does not hold, then  $h_{ij}^{(m)} = 0$ . By default,  $h_{ij}^{(0)} = 1$ ,  $h_{ij}^{(N)} = 0$  and  $h_{ii}^{(N)} = 1$  for  $1 \leq i, j \leq M$ .

Receiver makes the first test, i.e. determines whether one of the synchronization sequences is received, upon the reception of the first  $N$  symbols. If this is not the case, last  $N-1$  symbols together with the next received symbol are considered, and another test is made. The search process continues in the same fashion until the first successful test.

In [8] is shown that probability that  $k$ -th test is successful is:

$$\Pr\{k\} = \sum_{j=1}^M \sum_{i=1}^M \sum_{m=1}^{\min(N,k-j)} (h_{ij}^{(N-m+1)} \cdot r_j^{(m-1)} - h_{ij}^{(N-m)} \cdot r_j^{(m)}) \Pr\{k-m\} \quad (1)$$

where  $\Pr\{k\}$  denotes the probability that  $k$ -th test is successful,  $h_{ij}^{(m)}$  is a cross-bifix of  $m$ -th order, and  $r_j^{(m)}$  is probability of last  $m$  symbols of the sequence  $\mathbf{p}_j$ . It can be shown that  $\Pr\{k\}$  is a probability density function. Based on (1), the expected duration of the search process (expected search time) can be evaluated as:

$$E\{k\} = T = 1 - N + [\Pr\{1\}]^{-1} \sum_{i=1}^M S_i \sum_{j=1}^M C_{ij} \quad (2)$$

where  $\Pr\{1\}$  is the probability that the first test is successful,  $C_{ij}\{k\} = \sum_{j=1}^M h_{ij}^{(N-m+1)} \cdot r_j^{(m-1)}$ , while the parameters  $S_i$  are evaluated using following set of equations:

$$\mathbf{A} \cdot \mathbf{S}^T = \mathbf{0}^T \quad (3)$$

where  $\mathbf{S} = [S_1, S_2, \dots, S_M]$ , and  $\mathbf{A}$  is a  $M \times M$  matrix with coefficients  $A_{ij} = C_{j1} \cdot [r_1^{(N)}]^{-1} - C_{j,j+1} \cdot [r_{i+1}^{(N)}]^{-1}$ ,  $1 \leq i, j \leq M$ .

Variance of the search time is:

$$\sigma^2 = E\{k^2\} - E^2\{k\} \quad (4)$$

therefore, we need second moment, which is:

$$E\{k^2\} = 1 - 2 \cdot N \cdot T - N^2 + [\Pr\{1\}]^{-1} \cdot \sum_{i=1}^M (2 \cdot C_i \cdot T_i + W_i \cdot S_i) \quad (5)$$

where  $T_i = E\{k_i\}$  is the expected search time for  $i$ -th sequence,  $W_i = \sum_{j=1}^M W_{ij}$  and  $W_{ij} = \sum_{m=1}^N (2m-1) \cdot r_j^{(m-1)} \cdot h_{ij}^{(N-m+1)}$ , and  $T_i$  is obtained solving the following set of linear equations:

$$\mathbf{A} \cdot \mathbf{T}^T = \mathbf{B}^T \quad (6)$$

where  $\mathbf{T}$  is vector of expected search times  $\mathbf{T} = [T_1, T_2, \dots, T_M]$ ,  $\mathbf{B}$  is also vector with  $M$  elements,  $\mathbf{B} = [B_1, B_2, \dots, B_M]$  and  $B_i = \sum_{j=1}^M 0.5 \cdot (W_{j,i+1} \cdot [r_{i+1}^{(N)}]^{-1} - W_{ji} \cdot [r_i^{(N)}]^{-1}) \cdot S_j$ ,  $1 \leq i \leq M$ .

For the contiguous sequence case, formulae (2) and (3) are reduced to:

$$E\{k\} = 1 - N + \sum_{m=1}^N h^{(N-m+1)}. \quad (7)$$

$$\sigma^2 = (E\{k\} - N) \cdot (E\{k\} + N - 1) + 2 \sum_{m=1}^N m \cdot h^{(N-m+1)} \cdot \frac{r^{(m-1)}}{r^{(N)}} \quad (8)$$

which is for equiprobable symbols further simplified:

$$E\{k\} = 1 - N + \sum_{m=1}^N h^{(N-m+1)} \cdot L^{N-m+1} \quad (9)$$

$$\sigma^2 = (E\{k\} - N) \cdot (E\{k\} + N - 1) + 2 \cdot \sum_{m=1}^N m \cdot h^{(N-m+1)} \cdot L^{N-m+1} \quad (10)$$

where  $L$  is the cardinality of the symbol-set. If sequence is bifix free, we have:

$$E\{k\} = 1 - N + [\Pr\{1\}]^{-1} \quad (11)$$

$$\sigma^2 = (1 - 2N) \cdot [\Pr\{1\}]^{-1} + [\Pr\{1\}]^{-2} \quad (12)$$

and, for equiprobable symbols, finally:

$$E\{k\} = 1 - N + L^N \quad (13)$$

$$\sigma^2 = (1 - 2N) \cdot L^N + L^{2N} \quad (14)$$

Obtained with different, but similar reasoning, the validity of equations (9) and (13) is also proven in [7].

It is easily shown that for set of sequences which are cross-bifix-free, equations (11) and (12) also hold, while for the equiprobable symbols we have:

$$E\{k\} = 1 - N + L^l \quad (15)$$

$$\sigma^2 = (1 - 2N) \cdot L^l + L^{2l} \quad (16)$$

where  $l$  is the number of synchronization symbols in distributed sequence.

### 3. COMPARED SEQUENCES

#### 3.1 CONTIGUOUS SEQUENCES

First discovered contiguous sequences with property that maximum correlation sidelobe is equal to unity were Barker sequences [1]. Here, the influence of the adjacent data symbols when calculating the correlation was neglected. On the other hand, Willard [2] regarded the data symbols. There are several other authors that also contributed synchronization sequences constructed using similar reasoning (refer to Lindner, Jones, Maury and Styles, Turyn).

Tables 1 and 2 list aforementioned sequences up to length 15.

$N$	Barker	Willard	Turyn
5	11101	11010	-
6	-	-	-
7	1110010	0101010	1011000
8	-	00100111	-
9	110110011	000100111	-
10	-	0000111011	-
11	11100010010	00010010111	10110111000
12	-	000101101011	-
13	1111100110101	0000011010111	1111100110101
14	-	00000110100111	11111001100101
15	-	000001011100111	111110011010110

Table 1 *Barker, Willard and Turyn sequences*

$N$	Jones	Maury-Styles
5	00101	-
6	001011	-
7	0001011	1011000
8	00011011	10111000
9	000100111	101110000
10	0000111011	1101110000
11	00010010111	10110111000
12	000011010111	110101100000
13	0000011010111	1110101100000
14	00000101100111	11100110100000
15	000001011100111	111110001010000

Table 2 *Jones and Maury-Styles sequences*

#### 3.2 DISTRIBUTED SEQUENCES

Distributed sequences are presented in [5]. Here, the data symbols are interspersed with synchronization symbols (which are quite naturally placed at fixed positions in frame). Similarly to contiguous sequences, distributed sequences are optimized to have smallest possible correlation sidelobes. An interesting result obtained here is that, for the same number of synchronization symbols, distributed sequences have smaller correlation sidelobes when data symbols are taken into consideration. The drawback is that the length of the distributed sequence (i.e. the length of the frame) depends on the number of synchronization symbols. For instance, for binary sequences, frame length is upper-bounded by:

$$n_1 \cdot n_0 \geq h - 1 \quad (8)$$

where  $n_1$  and  $n_0$  are the number of synchronization symbols equal to 1 and 0 respectively, and  $h$  is the frame length ([5]).

Table 3 gives distributed sequences of maximum length for a given number of synchronization symbols up to 15 ([5]).

$N$	$l$	sequence
7	5	1110xx0
10	6	1110xx0xx0
13	7	111xx0xxx0x10
17	8	111xx0xxx0xxx0x10
20	9	111xx0xxx0xxx0x10xx0
24	10	11xxxx110xx0xx1xxxx0x0x0
28	11	111xx0x0xxxx0xxxx0xxxxxx0110
32	12	111xx0x0x0xxxx0xxxx0xxxxxx0110
37	13	111xx0x0x0xxxx0xxxx0xxxxxx0110
43	14	11xxxx11x0xxxxxxx0xxx0xx0xx1xx10x0x0
46	15	111xxx1xxx1xx0xxxxxx0xxxxxx0xxx0x1011x0xx0

Table 3 Distributed sequences

Finally, it should be noted that each sequence mentioned here actually represents four sequences with identical correlation properties – the first is the original sequence itself, the second is obtained by substituting 1-s with 0-s and inverse in the first one, and remaining two are just time-inversion of the first two.

#### 4. COMPARISON RESULTS

Since the all enlisted sequences are bifix free and cross-bifix-free respectively, comparisons are made using formulae (11) and (12).

First we considered contiguous sequences. Expected search time and search time variance for sequences of equal length behaves identically for all of them, which can be easily observed through formula (11). Namely, since the sequences are bifix free (optimized to minimize correlation sidelobes), number of 1-s and number 0-s are as close as possible in all sequences. Typical graph that plots expected search time ( $T$ ) vs. symbol probability ( $p$  – probability that a synchronization bit takes value 1) is given on figure 1 for Jones sequences of various lengths (parameter  $N$ ).

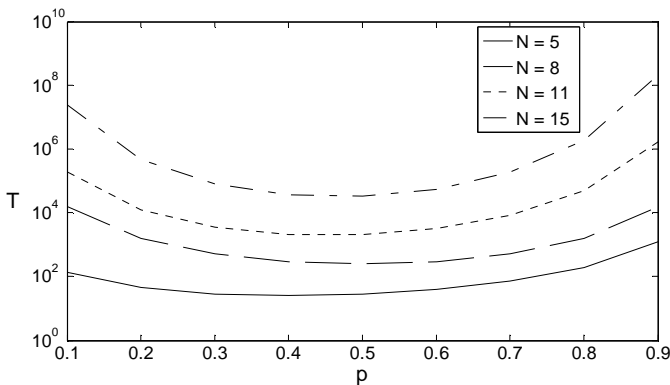


Figure 1 Expected search time for Jones sequences

Figure 2 gives the variance of search time for Jones sequences.

In figure 3 is depicted the expected search time vs. symbol probability for various distributed sequences of maximal

length, where parameter  $l$  is the number of synchronization symbols. Figure 4 gives the variance search time for these sequences.

Finally, figure 5 and 6 show the comparison between Jones sequences (taken as the representative of the contiguous sequences) and distributed sequences of maximum length.

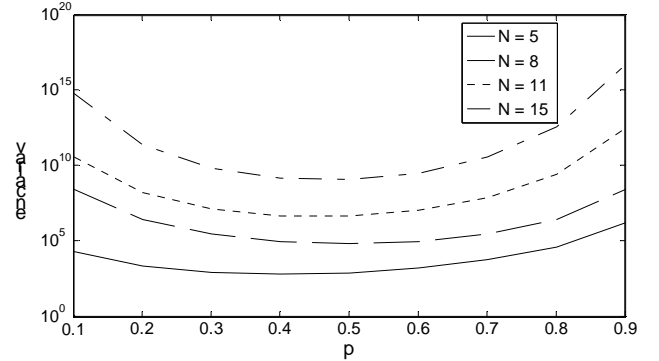


Figure 2 Variance of search time for Jones sequences

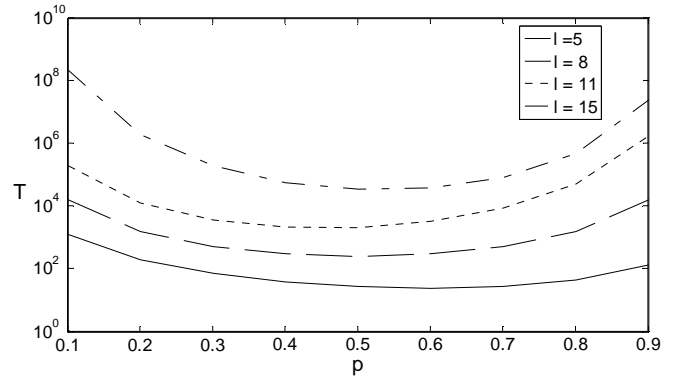


Figure 3 Expected search time for distributed sequences

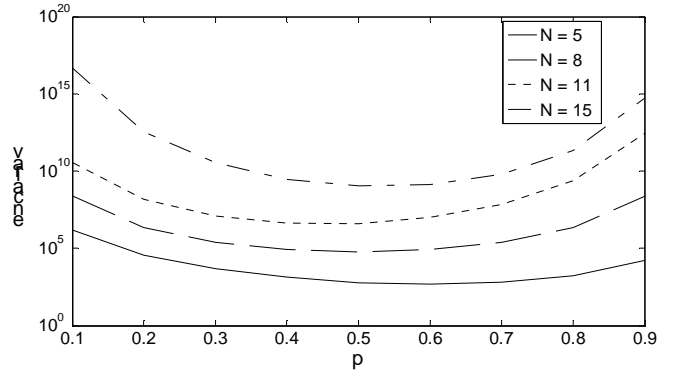


Figure 4 Variance of search time for distributed sequences

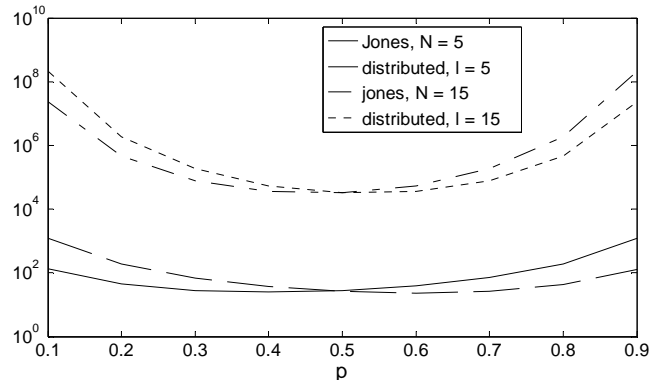


Figure 5 Comparison of expected search times

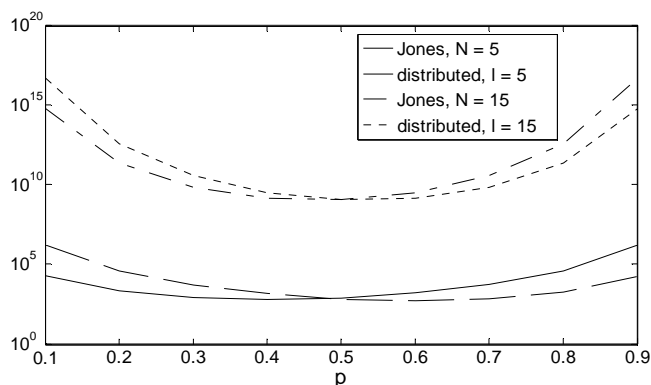


Figure 6 Comparison of search time variance

## 5. CONCLUSION

First it should be noted that, as it was already mentioned in the text above, when in each sequence 1-s and 0-s switch places, another sequence with the same correlation properties is obtained. It means the all graphs can be rotated against point  $p = 0.5$ .

As it can be seen from the given figures, expected search time and its variance rise with sequence length, which was quite predictable.

When contiguous and distributed sequences are compared, it can be observed that the appropriate statistics are almost the same. Distributed sequences have slightly smaller expected search time and its variance, but differences are negligible, especially for sequences with longer number of synchronization symbols.

This drives us to the conclusion that there is no significant difference between contiguous and distributed sequences in the sense of the search process. When the fact that there are limits imposed on the length on distributed sequences (which stem from their very nature) is taken into account, it can be said that contiguous sequences are better choice, since they potentially present less overhead in frame.

Finally, we should mention that the statistical bifix approach, introduced in section 2, could be easily applied to other interesting practical problems that relate to the cases when errors in the received synchronization sequence are tolerated.

## ACKNOWLEDGMENT

The authors are grateful to Slađana Tošić who was the first to check the formula within her diploma thesis and present it as a student's paper at Telfor 2004.

## REFERENCES

- [1] R.H. Barker, "Group Synchronization of Binary Digital Systems", *Communications Theory*, W. Jackson, ed., Academic Press Inc., New York, 1953.
- [2] M.W. Willard, "Optimum Code Patterns for PCM Synchronization", *Proc. Natl. Telem. Conf.*, 1962, paper 5-5
- [3] R.A. Scholtz: "Frame Synchronization Techniques" *IEEE Trans. on Comm.*, pp. 1204-1212, Vol. COM-28, August 1980.
- [4] M.N. Al-Subbagh and E.V. Jones: "Optimum patterns for frame alignment", *IEE Proc. part F - Commun. Radar & Signal processing*, Vol. 135 (6), pp 594-603, December 1988.
- [5] A.J. van Wijngaarden and T.J. Willink, "Frame Synchronization Using Distributed Approach", *IEEE Trans. on Commun.*, vol. 48, no. 12, pp. 2127-2138, December 2000.
- [6] Bernard Sklar, *Digital Communications*, Prentice-Hall International Inc., 1988.
- [7] P.T. Nielsen, "On the Expected Duration of the Search for a fixed pattern in random data", *IEEE Trans. Inform. Theory*, vol IT-19, pp. 702-704, September 1973.
- [8] D. Bajić, Č. Stefanović and D. Vukobratović, "Search Process and Probabilistic Bifix Approach", submitted for ISIT 2005.

## POREĐENJE KONTINUALNIH I DISTRIBURANIH SEKVENCI ZA SINHRONIZACIJU NA NIVOU FREJMA KORIŠĆENJEM STATISTIČKOG PRISTUPA PREKO BIFIKSA

Čedomir Stefanović, Dragana Bajić, *Fakultet tehničkih nauka, Univerzitet u Novom Sadu*

**Sažetak** – U radu je predstavljen kratak uvod u statistički pristup preko bifiksa kojim se može odrediti očekivano vreme pretrage za sinhronizacionom sekvencom u okviru frejma. Za razliku od prethodnih radova na ovu temu, reč je o opštem slučaju kada sinhronizacioni simboli nisu jednakoverovatni. Na osnovu ovoga napravljeno je poređenje između poznatih kontinualnih i distribuiranih sekvenci.