The Consultative Committee for Space Data Systems

Research and Development for Space Data System Standards

CORRELATED DATA GENERATION

EXPERIMENTAL SPECIFICATION

CCSDS 551.1-O-1

ORANGE BOOK

July 2015

The Consultative Committee for Space Data Systems

Research and Development for Space Data System Standards

CORRELATED DATA GENERATION

EXPERIMENTAL SPECIFICATION

CCSDS 551.1-O-1

ORANGE BOOK

July 2015

AUTHORITY

Issue:	Orange Book, Issue 1
Date:	July 2015
Location:	Washington, DC, USA

This document has been approved for publication by the Consultative Committee for Space Data Systems (CCSDS). The procedure for review and authorization of CCSDS documents is detailed in *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4).

This document is published and maintained by:

CCSDS Secretariat
National Aeronautics and Space Administration
Washington, DC, USA
E-mail: secretariat@mailman.ccsds.org

FOREWORD

Through the process of normal evolution, it is expected that expansion, deletion, or modification of this document may occur. This Recommended Standard is therefore subject to CCSDS document management and change control procedures, which are defined in the *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4). Current versions of CCSDS documents are maintained at the CCSDS Web site:

http://www.ccsds.org/

Questions relating to the contents or status of this document should be sent to the CCSDS Secretariat at the e-mail address indicated on page i.

At time of publication, the active Member and Observer Agencies of the CCSDS were:

Member Agencies

- Agenzia Spaziale Italiana (ASI)/Italy.
- Canadian Space Agency (CSA)/Canada.
- Centre National d'Etudes Spatiales (CNES)/France.
- China National Space Administration (CNSA)/People's Republic of China.
- Deutsches Zentrum für Luft- und Raumfahrt (DLR)/Germany.
- European Space Agency (ESA)/Europe.
- Federal Space Agency (FSA)/Russian Federation.
- Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
- Japan Aerospace Exploration Agency (JAXA)/Japan.
- National Aeronautics and Space Administration (NASA)/USA.
- UK Space Agency/United Kingdom.

Observer Agencies

- Austrian Space Agency (ASA)/Austria.
- Belgian Federal Science Policy Office (BFSPO)/Belgium.
- Central Research Institute of Machine Building (TsNIIMash)/Russian Federation.
- China Satellite Launch and Tracking Control General, Beijing Institute of Tracking and Telecommunications Technology (CLTC/BITTT)/China.
- Chinese Academy of Sciences (CAS)/China.
- Chinese Academy of Space Technology (CAST)/China.
- Commonwealth Scientific and Industrial Research Organization (CSIRO)/Australia.
- Danish National Space Center (DNSC)/Denmark.
- Departamento de Ciência e Tecnologia Aeroespacial (DCTA)/Brazil.
- Electronics and Telecommunications Research Institute (ETRI)/Korea.
- European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
- European Telecommunications Satellite Organization (EUTELSAT)/Europe.
- Geo-Informatics and Space Technology Development Agency (GISTDA)/Thailand.
- Hellenic National Space Committee (HNSC)/Greece.
- Indian Space Research Organization (ISRO)/India.
- Institute of Space Research (IKI)/Russian Federation.
- KFKI Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
- Korea Aerospace Research Institute (KARI)/Korea.
- Ministry of Communications (MOC)/Israel.
- National Institute of Information and Communications Technology (NICT)/Japan.
- National Oceanic and Atmospheric Administration (NOAA)/USA.
- National Space Agency of the Republic of Kazakhstan (NSARK)/Kazakhstan.
- National Space Organization (NSPO)/Chinese Taipei.
- Naval Center for Space Technology (NCST)/USA.
- Scientific and Technological Research Council of Turkey (TUBITAK)/Turkey.
- South African National Space Agency (SANSA)/Republic of South Africa.
- Space and Upper Atmosphere Research Commission (SUPARCO)/Pakistan.
- Swedish Space Corporation (SSC)/Sweden.
- Swiss Space Office (SSO)/Switzerland.
- United States Geological Survey (USGS)/USA.

DOCUMENT CONTROL

Document	Title	Date	Status
CCSDS 551.1-O-1	Correlated Data Generation, Experimental Specification, Issue 1	July 2015	Original issue
EC 1	Editorial Change 1	September 2015	Adds corrections/ clarifications to text in 4.2.2.1.4.5, 4.2.2.1.5.2, 4.2.2.1.5.6, 5.1.1, and figure B-6

CONTENTS

Section Page

1 INTRODUCTION 1-1

1.1 PURPOSE 1-1
1.2 SCOPE 1-1
1.3 APPLICABILITY 1-1
1.4 RATIONALE 1-2
1.5 DOCUMENT STRUCTURE 1-2
1.6 CONVENTIONS AND DEFINITIONS 1-3
1.7 NOMENCLATURE 1-3
1.8 ISSUES RELATED TO PATENTS 1-5
1.9 REFERENCES 1-5

2 OVERVIEW 2-1

2.1 METHOD FOR CORRELATED DATA GENERATION 2-1
2.2 CONDITIONS FOR IMPLEMENTING THE PROPOSED CORRELATED DATA GENERATION METHOD 2-2
2.3 CONDITIONS RELATED TO DATA BLOCKS CONTAINING TRANSFER FRAMES 2-3

3 DATA REQUIREMENTS 3-1

3.1 REQUIREMENTS FOR DATA SOURCE SIGNALS (DATA) 3-1
3.2 REQUIREMENTS FOR SIGNALS (DATA) GENERATED ON THE RECEIVING END 3-1

4 TECHNIQUES AND ALGORITHMS MAKING A BASIS OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION 4-1

4.1 TECHNIQUE FOR IMPROVING SYNCHRONIZATION QUALITY OF RECEIVED DATA BLOCKS 4-1
4.2 ALGORITHMS FOR CORRELATED DATA GENERATION 4-3

5 CRITERIA AND TECHNIQUES TO EVALUATE CAPABILITIES OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION 5-1

5.1 CRITERIA AND MODELS TO EVALUATE ALGORITHMS FOR IMPROVING DATA RELIABILITY 5-1
5.2 TECHNIQUES TO EVALUATE CAPABILITIES OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION 5-9

CONTENTS (continued)

Section Page

ANNEX A OPTIMAL WEIGHT CHARACTERISTICS OF RELIABILITY FOR IMPLEMENTING CORRELATED DATA GENERATION ALGORITHMS (NORMATIVE) A-1
ANNEX B REQUIRED EXPLANATIONS (INFORMATIVE) B-1
ANNEX C INFORMATIVE REFERENCES (INFORMATIVE) C-1

Figure

1-1 Bit Numbering Convention.. 1-4
2-1 Correlated Data Generation Process 2-1
2-2 Specified Data Block (Transfer Frame With the Attached Sync Marker and Correcting Bits of the Reed-Solomon Code) 2-3
2-3 TM Transfer Frame Structure 2-4
3-1 Graphs of Signal (Structure of Data) at the Diverse Channel Output (Software-Hardware Output For Correlated Data Generation) 3-2
4-1 Regeneration of Sync Pulses by the Proposed Method 4-3
5-1 Received Data Blocks Exposed to Model-Simulated Interferences 5-5
5-2 Flow-Graph of Simulated Interference Situation States 5-8
B-1 One of the Modulation Coding Techniques...B-1
B-2 Geometrical Representation of Zones Related to Elementary Signals u_{ij} and u_{kr} B-3
B-3 Results of Algorithm A_4 Application B-7
B-4 Results of Algorithm A_4 Evaluation via Simulated Data B-8
B-5 Examples of RRS Complementarity..B-9
B-6 Improvement of Correlated Telemetry Data Reliability.............................B-11

Table

4-1 WCR Values W_{ki} for $n = 5$, $m = 4$ 4-7
4-2 Examples for Estimations of Reliability, W_{ke}, Depending on Values of e_i, $i = 1, 2, \ldots, n$ Acquired via Signal Diverse Channels, with $n = 5$, $m = 4$ 4-7
4-3 Examples of Receiving Correlated Signals e_{cor} from e_i, $i = 1, 2, \ldots, n$, Acquired via Signal Diverse Channels, with $n = 5$, $m = 4$ for $k = 1, 2, \ldots, 166$.......... 4-8
5-1 Received Data Blocks Exposed to Model-Simulated Interferences 5-6
A-1 WCR for Algorithm A_4 ($n = 5$)... A-1
A-2 WCR for Algorithm A_4 ($n = 4$)... A-4
A-3 WCR for Algorithm A_4 ($n = 3$)... A-4
A-4 WCR for Algorithm A_{42} ($n = 5$)... A-5
A-5 WCR for Algorithm A_{42} ($n = 4$)... A-5
A-6 WCR for Algorithm A_{42} ($n = 3$)... A-5

1 INTRODUCTION

1.1 PURPOSE

The purpose of this Orange Book is to serve as a relevant CCSDS Experimental Specification for correlated data generation through using data obtained in compliance with the CCSDS requirements and acquired via diverse channels.

This Experimental Specification presents the diverse reception method that gives the ability to improve data reliability.

1.2 SCOPE

This Experimental Specification defines the correlated data generation method in terms of:

a) services provided to users of this Specification;

b) data formats; and

c) procedures to be performed for correlated data generation, processing, and evaluation.

This Experimental Specification is not extended to:

a) execution or product individuality;

b) methods or technologies needed to perform procedures; or

c) control actions required to control and monitor the correlated data generation facilities.

1.3 APPLICABILITY

This Experimental Specification is the Roscosmos contribution to CCSDS. It describes the method for correlated data generation through using data acquired via diverse channels that is destined for applications in near and outer space.

The Specification is not contemplated for updates or, far less, for the abolition of any provisions of the CCSDS current documents. On the contrary, it is intended to be compatible with them.

Strictly, this Experimental Specification is applicable to operations with data blocks acquired via diverse channels with the aim of improving the data reliability. Its implementation precludes any influence on the facilities previous to outputs of these channels. The Specification aims at improving reliability of received (no matter by what method) data. However, it is clear that the correlated data reliability depends on technical characteristics of the aforesaid facilities and on characteristics of data acquired from the diverse channels. Besides, different algorithms for correlated data generation ensure different reliability. Tools for controlling this dependence (criteria and techniques to evaluate efficiency of correlated data generation) are defined in this Experimental Specification thereby widening its application.

1.4 RATIONALE

Quality of data acquired via different diverse channels at receiving stations can be improved by delivering the data to a single receiving station where the data are correlated by the proposed method. The data receiving stations may include the relative software-hardware facilities of measuring stations located on territories of several states, with a common receiving station appointed MCC. The correlated data of improved reliability generated at this MCC may be transferred to MCCs of other Space Agencies.

Correlated data may be generated not only at a certain, a priori assigned, single receiving station (of the aforesaid MCC type). The correlated data may be generated at a ground station, with, for example, a diversity in carrier frequencies (when data are transferred on two frequencies) and in polarization (vertical and horizontal), or if the ground station is provided with the needed software and hardware facilities.

The proposed method for correlated data generation offers higher reliability against the known analogs based on auto-selection and data majorization.

Clearly, the data recorded at receiving stations shall be identical in structure, or a single receiving station shall be capable to convert the data acquired from different receiving stations into the required structure (for that, the associated software is needed). Herewith, certain requirements for the transfer signals (data) needs to be met so that the initial data blocks (data blocks at outputs of diverse channels) exhibit characteristics required to implement the proposed method for correlated data generation. In addition, the effective scheduling of tasks with the application of this method or alternative techniques requires the criteria and techniques to be developed, with the consideration of the proposed method specifics and without the margin for ambiguous evaluations.

Therefore the proposed method offers a useful tool for improving the received data reliability and is compatible with the current CCSDS documents. However, its full implementation requires certain conditions to be met as related to the cross-support. To create these conditions, as well as conditions required to control the method evolution process, the method itself, and the associated technical solutions must be documented.

1.5 DOCUMENT STRUCTURE

This document is divided into five numbered sections and three annexes:

- Section 1 presents the purpose, capabilities, applicability, and explanation of this Experimental Specification, as well as lists of conventions, definitions, and references used throughout the document.

- Section 2 provides an overview of the proposed method for correlated data generation and discusses conditions required for its implementation.
- Section 3 defines basic requirements concerning data structure.
- Section 4 defines techniques and algorithms making a basis for the proposed correlated data generation method.
- Section 5 defines criteria and techniques to estimate capabilities of the proposed method for correlated data generation.
- Annex A (normative) presents calculated optimal Weight Characteristics of Reliability (WCR) for implementing correlated data generation algorithms.
- Annex B (informative) presents necessary explanations of the content of certain provisions set forth in the document.
- Annex C presents informative references.

1.6 CONVENTIONS AND DEFINITIONS

1.6.1 DEFINITIONS OF TERMS USED IN THIS EXPERIMENTAL SPECIFICATION

This Experimental Specification makes use of the following terms.

mission phase (reference [1])**:** A period of a mission during which specified communication characteristics are fixed. A transition between two consecutive mission phases may cause an interruption of the communication services.

physical channel (reference [1])**:** A stream of bits transferred over a space link in a single direction.

diverse channels: Physical channels for transferring a stream of bits from one data source.

data block: A final number (set) of data having a certain structure (Transfer Frame—see reference [2]), a Transfer Frame with an attached sync marker and correcting bits of the Reed-Solomon code (see reference [2]), etc., specified data blocks).

correlated data: Data obtained by the diverse reception methods.

1.7 NOMENCLATURE

1.7.1 NORMATIVE TEXT

The following conventions apply for the normative specifications in this Recommended Standard:

a) the words ‘shall’ and ‘must’ imply a binding and verifiable specification;

b) the word ‘should’ implies an optional, but desirable, specification;

c) the word ‘may’ implies an optional specification;

d) the words ‘is’, ‘are’, and ‘will’ imply statements of fact.

NOTE – These conventions do not imply constraints on diction in text that is clearly informative in nature.

1.7.2 INFORMATIVE TEXT

In the normative sections of this document, informative text is set off from the normative specifications either in notes or under one of the following subsection headings:

- Overview;
- Background;
- Rationale;
- Discussion.

1.7.3 CONVENTIONS

In this document, the following conventions are used to identify each bit in an *N*-bit field. The first bit in the field to be transmitted (i.e., the most left justified when drawing a figure) is defined to be ‘Bit 0’, the following bit is defined to be ‘Bit 1’, and so on, up to ‘Bit *N*-1’. When the field is used to express a binary value (such as reciprocal), the Most Significant Bit (MSB) is the first transmitted bit of the field, i.e., ‘Bit 0’ (figure 1-1).

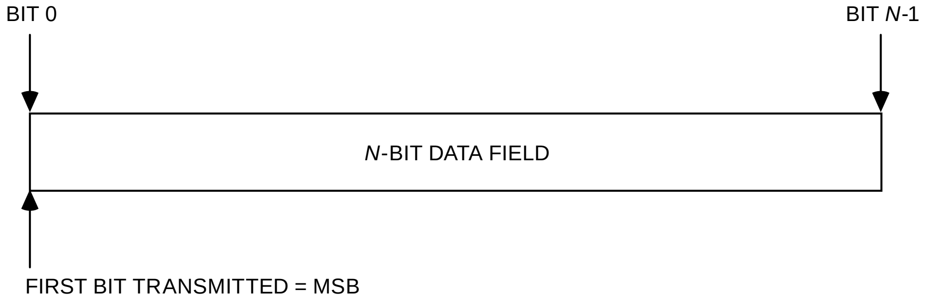


Figure 1-1: Bit Numbering Convention

In accordance with the standard data-communications practice, data fields are often grouped into 8-bit ‘words’ which conform to the above convention. Throughout this Recommended Standard, such an 8-bit word is called an ‘octet’.

The octets within the data structure are numbered starting from ‘0’.

1.8 ISSUES RELATED TO PATENTS

It must be noted that the techniques described in section 4 and related to the method for improving a synchronization quality of received data blocks and the method for developing correlated data generation algorithms (including the developed algorithms A_4 and A_{42} as such) are patented (Vorontsov, V.L., Method for Generation of Sync Pulses During Reception of Digital Signals, Pat. № 2446438, БИ № 9, issued March 27, 2012; and Vorontsov, V.L., Method for Definition of Weight Characteristics of Reliability for Processing the Received Multi-Position Signals, Pat. № 2339164, БИ № 32, issued November 20, 2008, respectively).

The patent holder is prepared to grant a free-of-charge license to an unrestricted number of applicants on a Reasonable and Non-Discriminatory (RAND) basis to make use of the aforesaid patented technologies. However, the patent holder reserves the right to grant a license on reasonable terms and conditions (but not free of charge) in the future.

1.9 REFERENCES

The following publications contain provisions which, through reference in this text, constitute provisions of this Experimental Specification. At the time of publication, the editions indicated were valid. All publications are subject to revision, and users of this Experimental Specification are encouraged to investigate the possibility of applying the most recent editions of the publications indicated below. The CCSDS Secretariat maintains a register of currently valid CCSDS publications.

[1] *TM Synchronization and Channel Coding*. Issue 2. Recommendation for Space Data System Standards (Blue Book), CCSDS 131.0-B-2. Washington, D.C.: CCSDS, August 2011.

[2] *TM Space Data Link Protocol*. Issue 1. Recommendation for Space Data System Standards (Blue Book), CCSDS 132.0-B-1. Washington, D.C.: CCSDS, September 2003.

NOTE – Annex C contains lists of informative references.

2 OVERVIEW

2.1 METHOD FOR CORRELATED DATA GENERATION

The method for correlated data generation is described below.

Data processed for transmission via communication channels are inserted in data blocks (e.g., in transfer frames). To improve the data reliability, diverse channels are used. Data are transferred from a data source to a receiver via different routes (via different diverse channels). If the number of such routes (diverse channels) is *n*, a receiver obtains the *n* number of data blocks corresponding to the same transferred data block. Each of the *n*-received data blocks is to a different degree corrupted with interferences in the diverse channels. From the *n*-received data blocks a single correlated data block can be generated through selecting the most reliable data (figure 2-1).

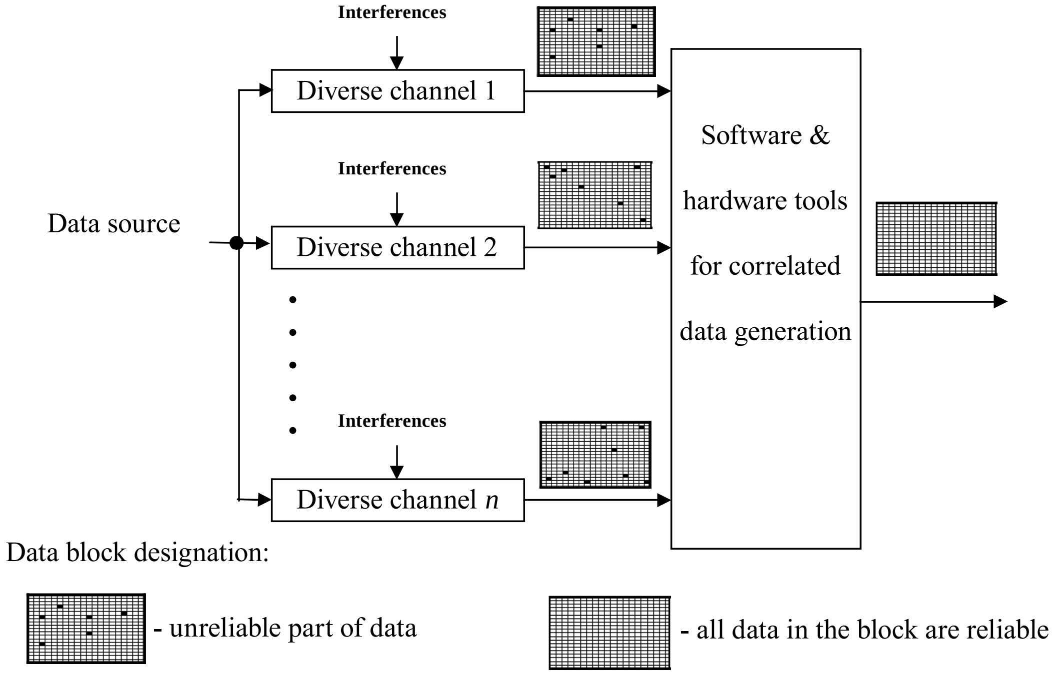


Figure 2-1: Correlated Data Generation Process

Practically, auto-selection and majorization are widely used to obtain correlated data.

With the auto-selection, data are automatically selected for a correlated data stream from a single diverse channel with a gain of 1, and with a gain of 0 for the rest. One of the auto-selection modifications is remarkable in that for a correlated data stream the data are selected one block at a time. Its major constraint is eliminating the possible complementarity of data blocks acquired via diverse channels related to the same transmitted data block and differently distorted with interferences.

In case of majorization, data is selected for a correlated data set by voting, with a gain for each diverse channel being equal to $^1/_n$ (where n is the number of diverse channels). If, for example, {0,1} is the data alphabet, $n = 5$, data values for 3 diverse channels are equal to 0, and for the rest two channels, 1; data whose value is equal to 0 is regarded as transferred (selected for the correlated data set). Its constraint is a low interference resistance when data are distorted in a large number of diverse channels.

The proposed method for correlated data generation offers the improved data reliability against auto-selection and majorization.

The initial data and generated correlated data refer to the second (channel) level of the Reference Model for Open Systems Interconnection.

Semantic properties of data (their semantic load) are of no importance. For example, a data block may contain spacecraft telemetry parameters, panoramic images of any planet, audio data (voice, melody, etc.), textual data, etc. A structure of words inserted in a data block also makes no matter.

To generate a correlated data block, the most reliable data (elementary data) extracted from a digital signal symbol is selected from data blocks acquired via diverse channels. Their size may be 1 or 2 bits (e.g., with two- or four-position signal, respectively). If each initial (received via a diverse channel) m-bit word ($m >> 1$) contains some invalid elementary data, there is a reason to expect that the generated word will not contain invalid correlated elementary data.

To obtain correlated data by this method, conformance of data acquired via different diverse channels to the same transferred data must be known. The required condition for ensuring such conformance is that boundaries of data blocks received via different diverse channels must be correctly defined. To fulfill this condition, the sync signals distorted by interferences in the diverse channels must be recovered.

2.2 CONDITIONS FOR IMPLEMENTING THE PROPOSED CORRELATED DATA GENERATION METHOD

2.2.1 DATA CHARACTERISTICS

Data (signals containing these data) diversity is achieved via radio signal polarization (extraction of vertical and horizontal polarization signals), frequency diversity (some carrier frequencies are used, e.g., VHF and UHF), space diversity (with several spaced receiving antennas), and time diversity (relaying of data blocks, e.g., through the onboard memory).

It should be noted that with two diverse channels the capabilities for selecting the most reliable data for their subsequent insertion in a data stream are rather limited (reduced to auto-selection). With three diverse channels, only auto-selection and majorization are possible. The number of diverse channels should be 4 or 5. Furthermore, when the data reception is stable, the proposed method becomes not relevant (as a matter of fact, this is also the case with other methods for improving data reliability). The proposed method is the most advantageous when data acquired via several diverse channels are equally distorted with interferences (then, in selecting reliable data, complementarity is possible).

2.2.2 INTERFERENCE SITUATION CHARACTERISTICS

The proposed method for correlated data generation is intended for rather unfavorable conditions (particularly, if capabilities of the methods recommended by CCSDS—see reference [1]—are exhausted), in cases when:

– because of high-intensive interferences in diverse channels, the error bit probability in received data blocks may be below 10^{-2};

– because of limited diversion capabilities, interferences are possible in all diverse channels simultaneously.

2.3 CONDITIONS RELATED TO DATA BLOCKS CONTAINING TRANSFER FRAMES

2.3.1 CONDITIONS FOR MEETING DATA STRUCTURE REQUIREMENTS

The aforesaid data blocks may be Transfer Frames intended to transfer telemetry with the attached sync markers and correcting bits of the Reed-Solomon Code shown in figure 2-2 (see figure 4-1 in reference [1]).

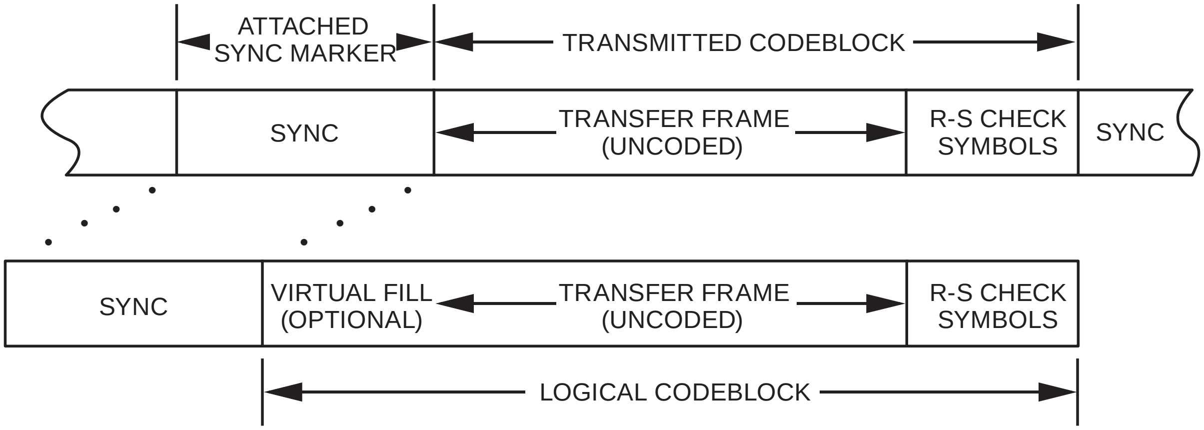
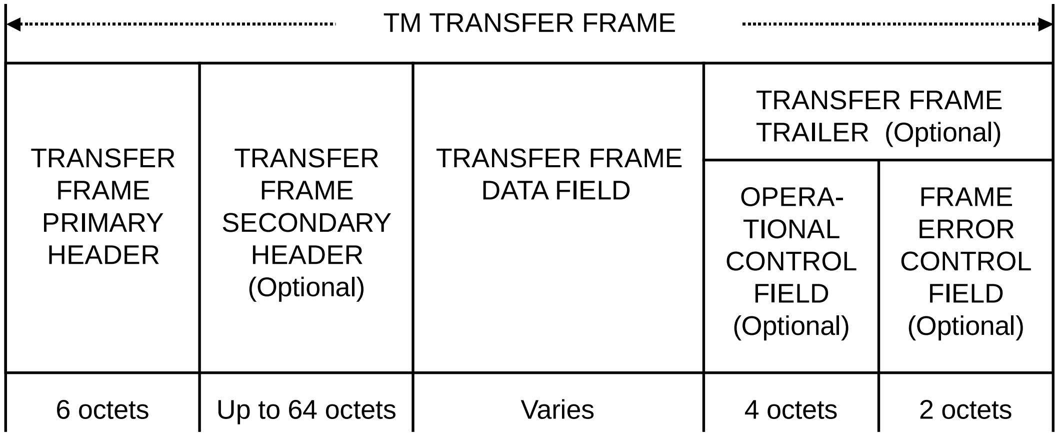


Figure 2-2: Specified Data Block (Transfer Frame With the Attached Sync Marker and Correcting Bits of the Reed-Solomon Code)

NOTE – Contents of the Transfer Frame are described in section 4 of reference [2]. Its structure is shown in figure 2-3 (see figure 4-1 in reference [2]).



TM TRANSFER FRAME

TRANSFER FRAME PRIMARY HEADER	TRANSFER FRAME SECONDARY HEADER (Optional)	TRANSFER FRAME DATA FIELD	TRANSFER FRAME TRAILER (Optional)	
			OPERA-TIONAL CONTROL FIELD (Optional)	FRAME ERROR CONTROL FIELD (Optional)
6 octets	Up to 64 octets	Varies	4 octets	2 octets

Figure 2-3: TM Transfer Frame Structure

NOTE – Contents of the Attached Sync Marker (ASM) and Reed-Solomon (R-S) codes are described in reference [1], sections 8 and 4, respectively.

2.3.2 CONDITIONS TO MEET TRAINING SAMPLE REQUIREMENTS

2.3.2.1 There are several ways to insert a training sample (test data) in a data block:

a) The training sample data can be can be inserted as the optional secondary header of the Transfer Frame (see figure 2-3).

b) The training sample bits can be uniformly distributed in the Transfer Frame Data Field (their volume, location, and values must be known a priori).

2.3.2.2 With a training sample inserted in the place of the secondary header of the Transfer Frame, regard must be paid to the following:

a) This technique for inserting a training sample gives the ability to use a standard structure of a Transfer Frame (see figure 2-3), given that the data inserted in its secondary header are irrelevant.

b) A training sample volume is limited to the maximum length of the Transfer Frame secondary header (64 octets).

2.3.2.3 With the training sample bits uniformly inserted in the Transfer Frame Data Field, regard must be paid to the following:

a) The test data bits need to be distributed a priori among other bits of the Transfer Frame Data Field (see 3.1 b)) according to the established procedure (rules), with alternation being optional (see section 6 in reference [1]).

b) Upon completion of correlated data generation and data correction by correcting data of the Reed-Solomon Code, prior to de-encapsulating the Transfer Frame, the test data bits need to be removed from the Transfer Frame Data Field to obtain a standard Transfer Frame (see figure 2-3).

c) Constraints related to a training sample (a number of injected test data bits) are associated with the fulfillment of requirements set forth in 3.1 d), e).

3 DATA REQUIREMENTS

3.1 REQUIREMENTS FOR DATA SOURCE SIGNALS (DATA)

To implement the proposed method for correlated data generation, the following requirements must be met:

a) data-block length must be static and known a priori throughout all mission phases;

b) training sample bits (test data) shall be uniformly distributed in the Transfer Frame Data Field, with the indication of their:

  – location;

  – value (alternating '0' and '1', unless other training sample, concurred a priori with a user, is defined);

c) a data block should be transferred within 100 ms;

d) a training sample size should be at least 400 bits;

e) a ratio between volumes of the entire data block and its training sample should be up to 40;

f) throughout all mission phases,

  1) data blocks shall be transferred with the attached sync markers defining boundaries of these data blocks;

  2) contents of these sync markers shall be known a priori to the extent sufficient for generating sync signals by the receiver's ground facilities;

g) the selected carrier modulation methods should support symmetry properties of diverse channels (see B1).

3.2 REQUIREMENTS FOR SIGNALS (DATA) GENERATED ON THE RECEIVING END

3.2.1 To implement the proposed method for correlated data generation, the requirements described hereafter shall be met.

3.2.2 Structures of logged data blocks generated of data acquired via diverse channels and structures of logged correlated data shall be identical and shall be of the form shown in figure 3-1.

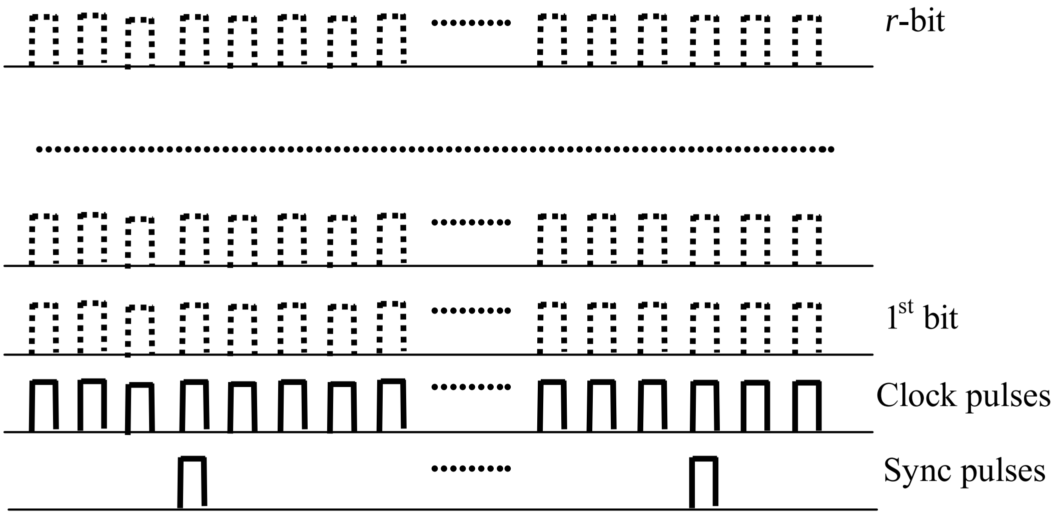


Figure 3-1: Graphs of Signal (Structure of Data) at the Diverse Channel Output (Software-Hardware Output For Correlated Data Generation)

NOTE – Contents of data structure elements are shown in figure 3-1:

3.2.3 Information data may be supplemented with data generated by the ground receiving facilities (particularly, with the station time data); in this case, for the address separation of different purpose data, the respective identifiers shall be generated in the structure (for that, bits r+1, r+2, … shall be used).

NOTE – Situations may occur when bits corresponding to the shown clock pulses are irrelevant (e.g., when this data structure is allocated in the computer memory);

3.2.4 Sync pulses separating the logged data blocks shall be generated through using sync data from diverse channels (specifically, using a sync marker attached to a Transfer Frame on the data source side—see reference [2]).

NOTE – This document does not define a method for generating sync signals (sync pulses) through using the aforesaid sync data.

3.2.5 Reliable sync signals (sync pulses) shall be synchronous with the last word of the logged data block.

4 TECHNIQUES AND ALGORITHMS MAKING A BASIS OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION

4.1 TECHNIQUE FOR IMPROVING SYNCHRONIZATION QUALITY OF RECEIVED DATA BLOCKS

4.1.1 BACKGROUND

Sync signals (sync pulses) separating data blocks (see figure 3-1) are generated through using attached sync markers (ASM) detected in a received signal (see section 8 in reference [1]). Because of interferences in diverse channels, the generated sync signals (sync pulses) may be invalid, that means:

- a sync signal (sync pulse) is not found in the place where it must be;
- a false sync signal (sync pulse) is found.

It should be noted that, with synchronization of a low interference resistance, the application of algorithms for correlated data generation becomes senseless. Therefore the concerted performance of the diverse reception and synchronization techniques is required. Hence, it would be advantageous to make additional efforts for improving synchronization quality of the received data blocks.

4.1.2 METHOD IMPLEMENTATION

4.1.2.1 Implementation Sequence

4.1.2.1.1 Data blocks shall be generated in the form of r-bit signals in the parallel binary code followed by clock pulses and sync pulses M_{rec} corresponding to the position of markers (see figure 3-1).

4.1.2.1.2 Counting shall be done from each generated sync pulse M_{rec} corresponding to the detected marker of data, $-k_1N_0$, $-(k_1-1)N_0$, $-(k_1-2)N_0$, …, $-N_0$, N_0, $2N_0$, …, $(k_2-1)N_0$, k_2N_0, and the counting boundaries shall be marked with corresponding signals $M_{imag}(-k_1)$, $M_{imag}(-k_1+1)$, $M_{imag}(-k_1+2)$, …., $M_{imag}(-1)$, $M_{imag}(1)$, $M_{imag}(2)$, …, $M_{imag}(k_2-1)$, $M_{imag}(k_2)$ of a similar amplitude.

4.1.2.1.3 Signals $M_{imag}(-k_1)$, $M_{imag}(-k_1+1)$, $M_{imag}(-k_1+2)$, …., $M_{imag}(-1)$, $M_{imag}(1)$, $M_{imag}(2)$, …, $M_{imag}(k_2-1)$, $M_{imag}(k_2)$, and M_{rec} shall be amplified in compliance with the specified amplification factors.

4.1.2.1.4 The amplified signals conforming to different sync pulses M_{rec}, but to similar time points, shall be combined so that the total amplitude is equal to a sum of amplitudes of the combined signals.

4.1.2.1.5 Each total signal amplitude shall be compared with an established threshold and, if the threshold is exceeded, a regenerated sync pulse M_{reg} shall be formed.

4.1.2.1.6 The regenerated sync pulse M_{reg} shall be inserted in the place of the corresponding sync pulse M_{imag} or M_{rec}; if no corresponding sync pulse M_{reg} is found, sync pulse M_{rec} shall be removed.

4.1.2.1.7 The r-bit data signals followed by clock pulses shall be grouped into blocks with a fixed number N_0 of data in each uncorrupted data blocks, with the data blocks being separated by regenerated sync pulses M_{reg}.

NOTE – The amplification factors, threshold, and values k_1 and k_2 shall be selected assuming that sync pulses is distorted by interferences in a communication channel.

4.1.2.2 Recommended Settings

For practical application, the following settings should be selected: $k_1 = k_2 = 5$, $U_{imag} = U_{rec} = 1$, $U_{thr} = 1$, where U_{imag} (U_{rec}) is a signal amplitude M_{imag} (M_{rec}) in conventional units; U_{thr} is an established threshold level in conventional units.

4.1.2.3 Discussion—Example of How to Apply This Method

For the purpose of this discussion, the following conditions should be assumed:

- The reference (transferred without distortions) markers correspond to the reference (received without distortions) sync pulses $M_{et}(0)$, $M_{et}(N_0)$, $M_{et}(2N_0)$, …., M_{et} $(15N_0)$, where iN_0 is a number of data from a counting start point corresponding to the i-n sync pulse M_{et}, $i = 0, 1, \ldots 15$.
- Interferences in the communication channel are intensified and attenuated after a while.

As a result of disturbances caused by interferences, sync pulses $M_{et}(6N_0)$, $M_{et}(7N_0)$, $M_{et}(8N_0)$, $M_{et}(9N_0)$, and $M_{et}(10N_0)$ are not found (the respective markers are not found because of the interferences), and a false sync pulse $M_{fal}(7{,}3 \cdot N_0)$ is found because of a false marker (figure 4-1).

If the method is applied under the aforesaid conditions, with $U_\Sigma(6N_0) = U_\Sigma(7N_0) = U_\Sigma(8N_0) = U_\Sigma(9N_0) = U_\Sigma(10N_0) = 6$, where $U_\Sigma(iN_0)$ is a result caused by amplified signals corresponding to the position of sync pulse $M_{imag}(iN_0)$, $i = 6, 7, 8, 9, 10$. As far as $U_\Sigma(iN_0) > U_{thr}$, $i = 6, 7, 8, 9, 10$, sync pulses $M_{reg}(6N_0)$, $M_{reg}(7N_0)$, $M_{reg}(8N_0)$, $M_{reg}(9N_0)$, and $M_{reg}(10N_0)$ will be generated (see figure 4-1). A false sync pulse $M_{fal}(7{,}3 \cdot N_0)$ gives rise to sync pulses M_{imag} which location corresponds to $6{,}3 \cdot N_0$, $5{,}3 \cdot N_0$, $4{,}3 \cdot N_0$, $3{,}3 \cdot N_0$, $2{,}3 \cdot N_0$, $8{,}3 \cdot N_0$, $9{,}3 \cdot N_0$, $10{,}3 \cdot N_0$, $11{,}3 \cdot N_0$, $12{,}3 \cdot N_0$ of current data and to no one of the input (distorted) sync pulses (figure 4-1). Therefore $U_\Sigma(7{,}3 \cdot N_0) = 1$ (i.e., the condition $U_\Sigma(7{,}3 \cdot N_0) > U_{thr}$ is not met). The false sync pulse $M_{fal}(7{,}3 \cdot N_0)$ shall be removed from the output sequence of sync pulses.

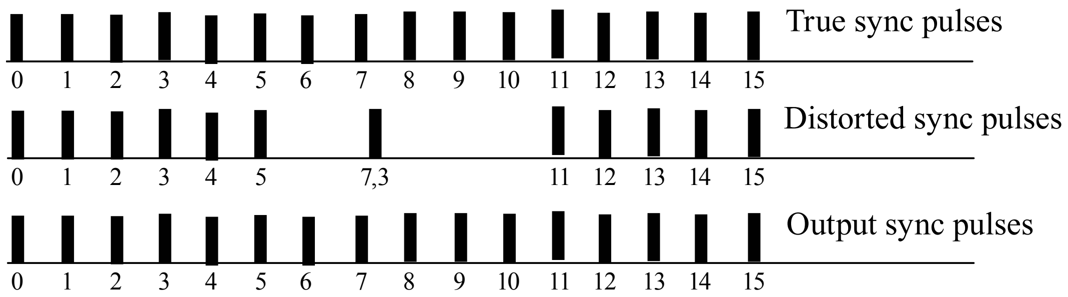


Figure 4-1: Regeneration of Sync Pulses by the Proposed Method

4.2 ALGORITHMS FOR CORRELATED DATA GENERATION

4.2.1 BACKGROUND

The method proposed to develop algorithms for correlated data generation is based on a statement that all possible weight characteristics of reliability as applied to data acquired via diverse channels may be presented as a final (rather small) set W (equation 1) ($\{W_{ki}\}$; $k = 1,\ldots,q$; $i = 1,\ldots,n$) being a matrix of the size of $q \times n$ (where q is a number of combinations of WCR values pertinent to data acquired via diverse channels and n is a number of diverse channels). This gives the ability to select a combination ensuring the best reliability of correlated data among the q combinations by sampling the a priori calculated weight characteristics of reliability $\{W_{ki}\}$.

A set of WCR combinations will be obtained by solving a combinatory task (including computers aids). Initial data for solving the combinatory task are: a number of diverse channels (n) and a size (m) of the applied signals (or the alphabet volume of data extracted from these signals; e.g., {0, 1, 2, 3} with $m = 4$).

If values n and m are relatively small, the combinatory task solution yields the final optimal set of WCR combinations whose volume is acceptable for practical application in generating correlated data (see annex A). For example, $q = 166$ with $n = 5$ and $m = 4$; i.e., the matrix size is 5×166 (see table A-1).

By the WCR combination the functional dependence is determined between the r-bit data from diverse channels corresponding to the same transmitted data and a correlated r-bit data (see figure 3-1). Each of the WCR combinations obtained via the proposed method determines a new functional dependence.

The optimal volume of WCR combinations means that, with the additional combinations, reliability of the correlated r-bit data is not improved, and, without even one combination, reliability degradation would be preconditioned.

4.2.2 ALGORITHM IMPLEMENTATION

4.2.2.1 Sequence for Implementing the Algorithm for Correlated Data Generation through Use of a Training Sample (Test Data)

4.2.2.1.1 The synchronization quality of received data blocks shall be improved by the method described in 4.1.2.1.

4.2.2.1.2 Numbers $i = 1, \ldots, n$ of the diverse channels shall be ordered in compliance with $N_{test\_rel\_i}$.

NOTE – $N_{test\_rel\_i}$ is the number of valid test elementary data in a data block of the i –n diverse channel.

4.2.2.1.2.1 Each diverse channel shall be assigned a number, depending on the value of $N_{test\_rel\_i}$, $i = 1, \ldots, n$.

4.2.2.1.2.2 The ordering rule shall be $N_{test\_rel\_n} \geq N_{test\_rel\_n-1} \geq \cdots N_{test\_rel\_2} \geq N_{test\_rel\_1}$.

4.2.2.1.3 Data shall be preliminarily sampled from diverse channels for further generation of correlated data in compliance with $N_{test\_rel\_lim}$.

NOTE – $N_{test\_rel\_lim}$ is the specified number of test elementary data in a data block.

4.2.2.1.3.1 If $N_{test\_rel\_i} < N_{test\_rel\_lim}$, the i-n channel data shall be ignored on a condition that at least one validation of $N_{test\_rel\_j}$ exists, with $N_{test\_rel\_j} \geq N_{test\_rel\_lim}$, $i \neq j$.

4.2.2.1.3.2 If $N_{test\_rel\_i} < N_{test\_rel\_lim}$, with $i = 1, \ldots, n$, data shall be sampled from one (j-n) diverse channel where $N_{test\_rel\_j}$ is the maximum value.

4.2.2.1.3.3 If $N_{test\_rel\_i} \geq N_{test\_rel\_lim}$ for one or two diverse channels, data shall be sampled from one (j-n) diverse channel where $N_{test\_rel\_j}$ is the maximum value.

4.2.2.1.3.4 If $N_{test\_rel\_i} \geq N_{test\_rel\_lim}$ for three, four, or five diverse channels, data shall be sampled from these channels.

4.2.2.1.3.5 If $N_{test\_rel\_i} \geq N_{test\_rel\_lim}$ for more than five diverse channels, data shall be selected from five diverse channels by descending values of $N_{test\_rel\_i}$ corresponding to them.

4.2.2.1.4 From a set of WCR combinations $\{W_{ki}\}$ ($k = 1, 2, \ldots, q$; $i = 1, \ldots, n^{/}$) the optimal combination $\{W_{hi}\}$, $k = h$, $i = 1, \ldots, n$ that ensures the largest number of valid correlated elementary test data in a data block shall be selected.

4.2.2.1.4.1 If a number of preliminarily selected channels is $n^{/}= 3$, $n^{/}= 4$, or $n^{/}= 5$, calculations shall be based on the sets of WCR combinations presented in annex A (see A1 and A2, respectively, for algorithms A_4 and A_{42}).

4.2.2.1.4.2 If a single channel is preliminarily selected, auto-selection shall be done from this point on (WCR of annex A shall not be applied).

4.2.2.1.4.3 Reliability of test elementary data shall be evaluated by comparing them with the respective reference data of a training sample, with their values known a priori.

4.2.2.1.4.4 Indices $k^{//}$ in tables of annex A define the application priority of WCR combinations; if, for the k-n, (k+1)-n, ….(k+s)-n WCR combinations, the number of valid correlated test elementary data in a data block is equal, preference shall be given to the k-n WCR combination.

NOTE – When the error probability is low and no errors are found in test data, such a choice of WCR gives the ability to apply more effective majorization instead of auto-selection.

4.2.2.1.4.5 Values W_{ke} for the r-bit data (see figure 3-1) forming elementary data blocks shall be calculated from the following formula:

$$\mathrm{W}_{\mathrm{ke}} = \sum_{\mathrm{i}=\mathrm{i}_1,\mathrm{i}_2,..} \mathrm{W}_{\mathrm{ki}} \;, \tag{1}$$

where:

W_{ke} is the overall reliability estimate e of a correlated elementary data extracted from the m-position signal by using the k-n WCR combination;

W_{ki} is the weight characteristic of reliability for the k-n combination for the i-n diverse channel;

i_1, i_2, …, i_g are numbers of diverse channels where value e of identified elementary data corresponding to the same transmitted elementary data of the m-position signal was repeated;

e is a value of an elementary data extracted from the m-position signal.

4.2.2.1.4.6 Herewith, values of W_{ke} data on the overall estimates of reliability for each value of the received test data shall be calculated, with sets of WCR combinations (presented in annex A) to be used depending on a number of preliminary selected channels $n^{/}$.

NOTE – If, for example, algorithm A_4 is used and $n^{/} = 5$, the WCR values would be selected from table A-1; for A_{42} and $n^{/} = 4$, from table A-5; etc.

4.2.2.1.4.7 If four-position signals are used, the possible values of e (i.e., the elementary data alphabet) shall be equal to 0, 1, 2, and 3, and the respective reliability estimates, W_{k_0}, W_{k_1}, W_{k_2}, W_{k_3}.

NOTE – The calculations are explained by examples. The WCR W_{ki} values required for these calculations (equation 1) are given in table 4-1.

Example 1. If $e_1 = 0$, $e_2 = 1$, $e_3 = 0$, $e_4 = 3$, $e_5 = 0$ (see table 4-2), where e_i is a value of an elementary data transferred via the i-n diverse channel, then:

- $W_{k\_0} = W_{k1} + W_{k3} + W_{k5}$,
- $W_{k\_1} = W_{k2}$,
- $W_{k\_3} = W_{k4}$.

Consequently, for the first WCR combination ($k = 1$, values of combinations are given in table 4-1):

- $W_{1\_0} = 1 + 20 + 62 = 83$, $W_{1\_1} = 10$, $W_{1\_3} = 30$.

Example 2. Similar to Example 1, except for: $k = 2$ (values of combinations are given in table 4-1):

- $W_{2\_0} = 11 + 30 + 92 = 133$, $W_{2\_1} = 20$, $W_{2\_3} = 40$.

Example 3. If $e_1 = 1$, $e_2 = 2$, $e_3 = 1$, $e_4 = 2$, $e_5 = 0$ (see table 4-1), then:

- $Wk\_0 = Wk5,$
- $Wk\_1 = Wk1 + Wk3,$
- $Wk\_2 = Wk2 + Wk4.$

Consequently, for the first WCR combination ($k = 1$, values of combinations are given in table 4-1):

- $W_{1\_0} = 62$, $W_{1\_1} = 1 + 20 = 21$, $W_{1\_2} = 10 + 30 = 40$.

4.2.2.1.5 Among the accepted elementary test data corresponding to the same transmitted elementary test data, data shall be selected whose value $e_{cor\_test\_k}$ ($k = 1, 2, \ldots, q$) corresponds to the maximum overall reliability estimate W_{ke} (where $e_{cor\_test\_k}$ is a correlated elementary test data obtained by using the k-n combination of WCR).

NOTE – Values of W_{k_0}, W_{k_1}, W_{k_2}, W_{k_3} for some e_1, e_2, e_3, e_4, e_5 and WCR combinations have been calculated and are shown in table 4-3 with the respective values of correlated r-bit data e_{cor}.

Example 1 (continued). The maximum value of reliability estimate is 83, and the estimate is related to the position ‘0’ value. This means that $e_{cor\_test\_1} = 0$ ($k = 1$).

Example 2 (continued). The maximum value of reliability estimate is 133, $e_{cor\_test\_2} = 0$ ($k = 2$).

Example 3 (continued). The maximum value of reliability estimate is 62, $e_{cor\_test\_1} = 0$ ($k = 1$).

4.2.2.1.5.1 The reliable correlated test elementary data shall be defined by comparing values $e_{cor\_test\_k}$ with their reference values.

4.2.2.1.5.2 Values of $N_{test\_rel\_cor\_k}$ shall be calculated (where $N_{test\_rel\_cor\_k}$ is a number of reliable correlated elementary test data in a data block generated by using the k-n combination of WCR) for WCR combinations from a set of WCR combinations given in tables of annex A.

4.2.2.1.5.3 Among the obtained values of $N_{test\_rel\_cor\_k}$ the largest value and the respective WCR $\{W_{hi}\}$ combination shall be selected from the set of combinations given in tables of annex A. The selected combination is considered optimal and shall be further used to generate correlated elementary information data.

4.2.2.1.5.4 If $N_{test\_rel\_cor\_k} = N_{test\_rel\_j\_max}$ ($N_{test\_rel\_j\_max}$ is a number of test data in a data block), the further selection of WCR combinations shall be terminated.

4.2.2.1.5.5 A sequence of correlated elementary information and test data shall be generated by using the optimal WCR combination $\{W_{hi}\}$ ($k = h$, $i = 1, \ldots, n'$).

4.2.2.1.5.6 The correlated elementary information data e_{cor_h} (where e_{cor_h} is a correlated elementary data obtained via the optimal WCR h-n combination $\{W_{hi}\}$) shall be generated similarly to $e_{cor\_test\_h}$ (see 4.2.2.1.5). The generated output data shall contain both correlated information data and correlated test elementary data.

Table 4-1: WCR Values W_{ki} for $n = 5$, $m = 4$

k	W_{k1}	W_{k2}	W_{k3}	W_{k4}	W_{k5}
1	1	10	20	30	62
2	11	20	30	40	92
3	11	20	30	40	82
…….	………..	……….	……..	………..	…….
165	51	80	90	100	112
166	41	60	70	80	82

Table 4-2: Examples for Estimations of Reliability, W_{ke}, Depending on Values of e_i, $i = 1, 2, \ldots, n$ Acquired via Signal Diverse Channels, with $n = 5$, $m = 4$

e_i, $i =$					W_{k_0}	W_{k_1}	W_{k_2}	W_{k_3}
1	2	3	4	5				
0	1	0	3	0	$W_{k1} + W_{k3} + W_{k5}$	W_{k2}	-	W_{k4}
1	2	1	2	0	W_{k5}	$W_{k1} + W_{k3}$	$W_{k2} + W_{k4}$	-
3	3	3	3	1	-	W_{k5}	-	$W_{k1} + W_{k2} + W_{k3} + W_{k4}$
…	…	…	…	…	…………..	…………	…………	………………

Table 4-3: Examples of Receiving Correlated Signals e_{cor} from e_i, $i = 1, 2, \ldots, n$, Acquired via Signal Diverse Channels, with $n = 5$, $m = 4$ for $k = 1, 2, \ldots, 166$

e_i, i =					k	W_{k_0}	W_{k_1}	W_{k_2}	W_{k_3}	E_{cor}
1	2	3	4	5						
0	1	0	3	0	1	83	10	-	30	0
					2	133	20	-	40	0
					3	123	20	-	40	0
					….	….	….	….	….	….
					165	253	80	-	100	0
					166	193	60	-	80	0
1	2	1	2	0	1	62	21	40	-	0
					2	92	41	60	-	0
					3	82	41	60	-	0
					….	….	….	….	….	….
					165	112	141	180	-	2
					166	82	111	140	-	2
3	3	3	3	1	1	-	62	-	61	1
					2	-	92	-	101	3
					3	-	82	-	101	3
					….	….	….	….	….	….
					165	-	112	-	321	3
					166	-	82	-	251	3

4.2.2.2 In case of using the noiseless coding methods (e.g., the Reed-Solomon method—see reference [1]), the correlated data should be generated in the first instance and then decoded (see rationale in B2.1).

4.2.2.3 Annex subsection B2.2 presents results of using algorithm A_4, with $N_{test\_rel\_lim}$ = 400, $N_{test\_rel\_j\_max}$ = 640 B, which illustrate its capabilities (where $N_{test\_rel\_j\_max}$ is the maximum number of reliable elementary test 2-bit data in a data block, or a number of elementary test data in a data block).

5 CRITERIA AND TECHNIQUES TO EVALUATE CAPABILITIES OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION

5.1 CRITERIA AND MODELS TO EVALUATE ALGORITHMS FOR IMPROVING DATA RELIABILITY

5.1.1 CRITERIA FOR COMPARATIVE EVALUATION OF ALGORITHMS FOR IMPROVING DATA RELIABILITY

5.1.1.1 To evaluate correcting capabilities of the test algorithm A_i, as compared to the selected basic algorithms A_{j1}, A_{j2}, …, A_{jr}, the following criteria should be used:

a) comparative characteristics γ of a number of error in a correlated data block:

$$\gamma(s) = \frac{\left(\min\{N_{j1}(s), N_{j2}(s), \ldots, N_{ir}(s)\} - N_i(s)\right) \cdot 100}{\min\{N_{j1}(s), N_{j2}(s), \ldots, N_{ir}(s), N_i(s)\} + {}^{1}\!/_{8}\min\{N_{j1}(s_M), N_{j2}(s_M), \ldots, N_{ir}(s_M)\}}, \qquad (2)$$

where:

$N_i(s)$, $N_{jh}(s)$ is a number of errors in the analyzed correlated data block generated by using algorithm A_i and A_{jh} ($h = 1, 2, \ldots, r$), respectively;

s is an interference situation state;

s_M is the most unfavorable interference situation of those considered;

$min\{N_1, N_2, \ldots, N_m\}$ is the smallest value among N_1, N_2, …, N_m;

b) rated estimates of reliability E_i calculated via characteristics $\gamma(s)$:

$$\begin{aligned} &E_i(s) = 1, \ \gamma(s) \geq \gamma_{thr}, \\ &E_i(s) = 0, \ -\gamma_{thr} < \gamma(s) < \gamma_{thr}, \\ &E_i(s) = -1, \ \gamma(s) \leq -\gamma_{thr}, \end{aligned} \qquad (3)$$

where:

γ_{thr} is a specified tolerable value of a comparative characteristic of a number of errors in the correlated data stream.

5.1.1.2 Value of $\gamma_{thr} = 15$ (equation 3) should be used.

5.1.1.3 The test algorithm A_i performance in the interference situations, s, should be evaluated via one of three possible values of E_i (equation 3):

a) $E_i = 1$ (‘better that basic algorithms’),

b) $E_i = 0$ (‘much the same as the basic algorithms’), and

c) $E_i = -1$ (‘worse than the basic algorithms’).

5.1.1.4 In case when algorithm A_i is tested in fixed interference situation states, the required condition should be ensured to extend the evaluation results to the overall spectrum of interference situations, i.e., to ensure predictability in transition from one interference situation state to the other. This condition can be fulfilled through using a certain model and selecting interference situation states to be analyzed.

$$N_i(s_q) \le N_i(s_p) \ , \ q < p \ . \tag{4}$$

5.1.1.5 To meet the condition (equation 4), the following criteria (similar to 5.1.1.1) should be applied:

a) the comparative characteristic γ of a number of errors in a correlated data block in transition from one state of interference situation s_q to the other s_p—the transformed formula (equation 2) shall be used:

$$\gamma(s_q, s_p) = \frac{\left(\min\{N_{j1}(s_q), N_{j2}(s_q), \ldots, N_{ir}(s_q)\} - N_i(s_p)\right)\cdot 100}{\min\{N_{j1}(s_q), N_{j2}(s_q), \ldots, N_{ir}(s_q), N_i(s_p)\} + \frac{1}{8}\min\{N_{j1}(s_M), N_{j2}(s_M), \ldots, N_{ir}(s_M)\}}, \tag{5}$$

where s_q, s_p are interference situation states related to the basic (A_{j1}, A_{j2}, …, A_{jr}) and test (A_i) algorithms, respectively;

b) rated estimates of reliability $E_i(s_q, s_p)$ calculated from characteristics $\gamma(s_q, s_p)$ (equation 4):

$$
\begin{aligned}
&- \ E_i(s_q, s_p) = 1 \text{ with } \gamma(s_q, s_p) \ge \gamma_{thr},\\
&- \ E_i(s_q, s_p) = 0 \text{ with } -\gamma_{thr} < \gamma(s_q, s_p) < \gamma_{thr},\\
&- \ E_i(s_q, s_p) = -1 \text{ with } \gamma(s_q, s_p) \le -\gamma_{thr}.
\end{aligned}
\tag{6}
$$

5.1.1.6 The rated estimates of reliability $E_i(s_q, s_p)$ calculated as the estimates $E_i(s)$ (equation 3), with the following possible conclusions:

- if $E(s_q, s_p) = 1$, with $E(s_q) = 1$, $E(s_p) = 1$, in transition from the interference situation state s_q to state s_p, the test algorithm A_i performance is efficient;
- if $E(s_q, s_p) = 0$, with $E(s_q) \ge 0$, $E(s_p) \ge 0$, in transition from the interference situation state s_q to state s_p the test algorithm A_i performs not worse than the basic algorithms;
- if $E(s_q) = -1$ or (and) $E(s_p) = -1$, in transition from the interference situation state s_q to state s_p, the test algorithm A_i performance is inefficient.

NOTE – In the aforesaid cases, the performance quality is predictable, with the worse selected options of its possible variation. If $E(s_q, s_p) = -1$, with $E(s_q) \geq 0$, $E(s_p) \geq 0$, in transition from the interference situation state s_q to state s_p, the quality of the test algorithm A_i is not predicted. Instead, the fact is stated that the efficiency estimations are uncertain because of a rough model of an error source in diverse channels and (or) an improper selection of value γ_{thr}.

5.1.2 MODEL TO EVALUATE ALGORITHMS FOR CORRELATED DATA GENERATION

5.1.2.1 Initial Data Required for the Model Development

5.1.2.1.1 An elementary data volume shall be 2 bits.

5.1.2.1.2 A training sample (a data block) volume shall contain 640 elementary data.

5.1.2.1.3 A number of initial data block formed with consideration for specifics of exposure to interference of four-position signals shall be 30.

5.1.2.2 Model Definition

5.1.2.2.1 To minimize a number of simulated interference situations, a number of interferences to which data blocks in diverse channels are exposed shall be determined.

5.1.2.2.2 A designation ‘interference intensity in the *i*-n diverse channel’ shall be introduced (for short) (u_i), 6 levels in total ($u_{i\,max} = 6$).

NOTE – With $u = 1$, the initial data block is exposed to 16 interferences; with $u = 2$ to 32 interferences; with $u = 3$, to 64 interferences; with $u = 4$, to 128 interferences; with $u = 5$, to 256 interferences; and with $u = 6$ (640), all data are exposed to interferences. Because the simulated channels are asymmetric, a data block exposed to 16 interferences contains 12 errors; with 32 interferences, 24 errors, etc., up to 480 errors when all (640) data are damaged with interferences.

5.1.2.2.3 Each level of interference intensity shall correspond to a specified error probability:

$$\begin{aligned}
&P_{er}(u = 1) = 0{,}01875\ ,\\
&P_{er}(u = 2) = 0{,}03750\ ,\\
&P_{er}(u = 3) = 0{,}07500\ ,\\
&P_{er}(u = 4) = 0{,}15000\ ,\\
&P_{er}(u = 5) = 0{,}30000\ ,\\
&P_{er}(u = 6) = 0{,}75000\ .
\end{aligned} \tag{7}$$

NOTE – Allocation of data damaged with interferences in the received data blocks (figure 5-1) permits simulating a strong dependence of interferences in diverse channels. Herewith, a nature of this dependence is stable that is important for comparing the estimates made for different interference situation states (to justify the procedures described in 5.1.1.4, 5.1.1.5, and 5.1.1.6).

5.1.2.2.4 The developed error source model for diverse channels shall incorporate 30 initial data blocks considering specifics of interferences to which four-position signals are exposed.

5.1.2.3 Discussion

With regard to results, consistency between the interference intensity levels and specific numbers of diverse channels is of no concern (i.e., it does not matter whether $u_1 = 4$ or $u_3 = 4$, with the interference level of 1 in the remaining diverse channels). Therefore the amount of simulated states is reduced to 251 ($s_{max} = 251$) without the analysis quality degradation.

The model is remarkable in that the data blocks are combined (see figure 5-1) and interference situation states are simulated (table 5-1), which permits building a graph (figure 5-2) and gives the ability to implement the condition (equation 4), with the interference situation evidently getting worse. For example, with $s_q = 5$ and $s_p = 6$, it becomes evident that the interference situation (intensity of interference in the 5th channel is getting worse, from 5 to 6, with the interference level unchanged in the rest channels). However, with $s_q = 5$ and $s_p = 7$, the fact that the interference situation is getting worse is not evident (as the interference level in the 4th channel increased from 1 to 2 and decreased from 5 to 2 in the 5th channel). When using a rigorous approach, such a conclusion is not always justified, as no consideration is given to the algorithm performance specifics and interference dependence variations in diverse channels. However, it largely complies with the reality in the analyzed conditions. The graph tops correspond to states, s, and the nodes between the tops suggest that the interference situation is getting worse during the state-to-state transition (equation 4).

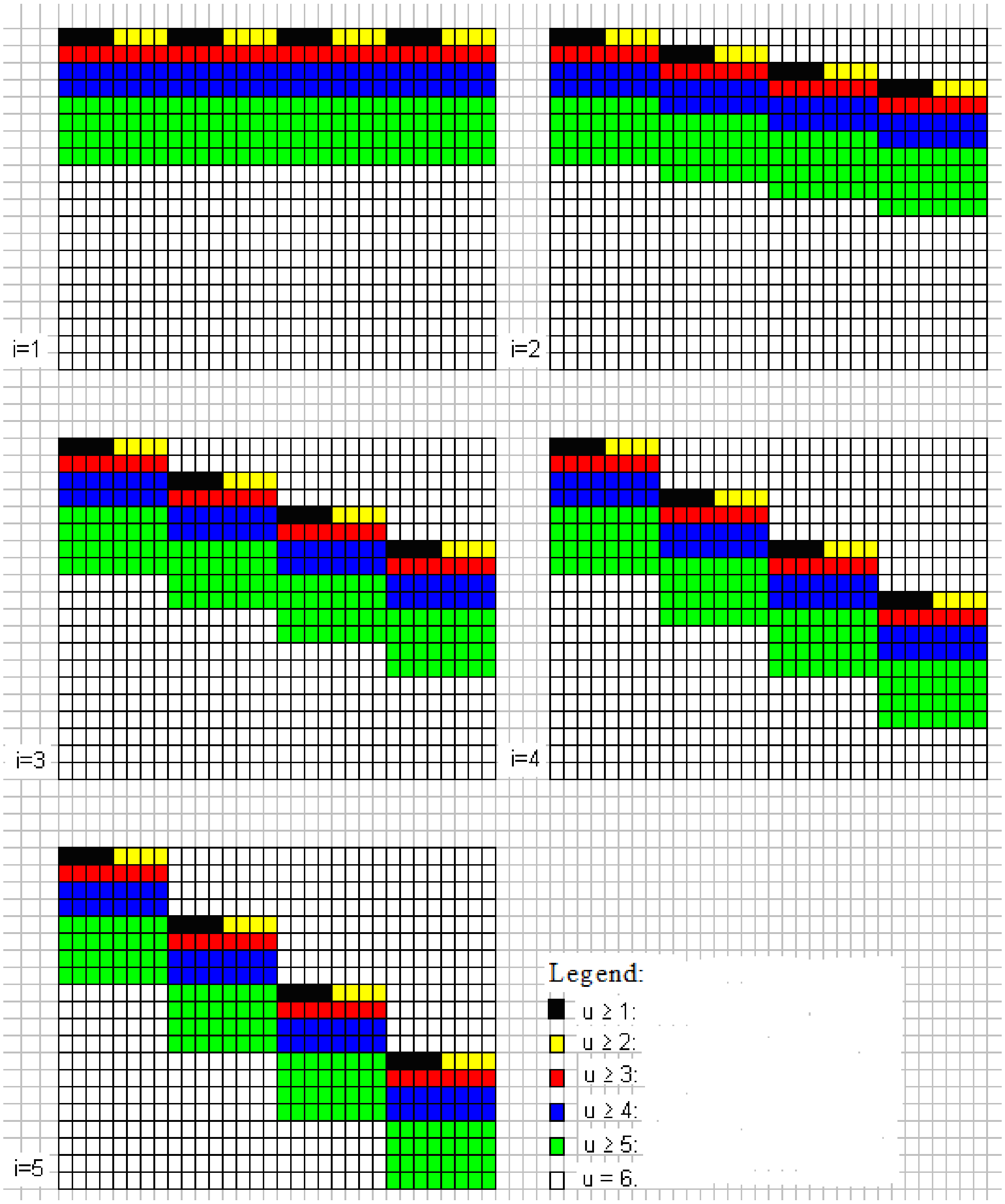


Figure 5-1: Received Data Blocks Exposed to Model-Simulated Interferences

Table 5-1: Received Data Blocks Exposed to Model-Simulated Interferences

s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5
1	1	1	1	1	1	51	1	1	4	5	6	101	1	3	3	6	6	151	2	2	3	6	6	201	3	3	3	4	4
2	1	1	1	1	2	52	1	1	4	6	6	102	1	3	4	4	4	152	2	2	4	4	4	202	3	3	3	4	5
3	1	1	1	1	3	53	1	1	5	5	5	103	1	3	4	4	5	153	2	2	4	4	5	203	3	3	3	4	6
4	1	1	1	1	4	54	1	1	5	5	6	104	1	3	4	4	6	154	2	2	4	4	6	204	3	3	3	5	5
5	1	1	1	1	5	55	1	1	5	6	6	105	1	3	4	5	5	155	2	2	4	5	5	205	3	3	3	5	6
6	1	1	1	1	6	56	1	1	6	6	6	106	1	3	4	5	6	156	2	2	4	5	6	206	3	3	3	6	6
7	1	1	1	2	2	57	1	2	2	2	2	107	1	3	4	6	6	157	2	2	4	6	6	207	3	3	4	4	4
8	1	1	1	2	3	58	1	2	2	2	3	108	1	3	5	5	5	158	2	2	5	5	5	208	3	3	4	4	5
9	1	1	1	2	4	59	1	2	2	2	4	109	1	3	5	5	6	159	2	2	5	5	6	209	3	3	4	4	6
10	1	1	1	2	5	60	1	2	2	2	5	110	1	3	5	6	6	160	2	2	5	6	6	210	3	3	4	5	5
11	1	1	1	2	6	61	1	2	2	2	6	111	1	3	6	6	6	161	2	2	6	6	6	211	3	3	4	5	6
12	1	1	1	3	3	62	1	2	2	3	3	112	1	4	4	4	4	162	2	3	3	3	3	212	3	3	4	6	6
13	1	1	1	3	4	63	1	2	2	3	4	113	1	4	4	4	5	163	2	3	3	3	4	213	3	3	5	5	5
14	1	1	1	3	5	64	1	2	2	3	5	114	1	4	4	4	6	164	2	3	3	3	5	214	3	3	5	5	6
15	1	1	1	3	6	65	1	2	2	3	6	115	1	4	4	5	5	165	2	3	3	3	6	215	3	3	5	6	6
16	1	1	1	4	4	66	1	2	2	4	4	116	1	4	4	5	6	166	2	3	3	4	4	216	3	3	6	6	6
17	1	1	1	4	5	67	1	2	2	4	5	117	1	4	4	6	6	167	2	3	3	4	5	217	3	4	4	4	4
18	1	1	1	4	6	68	1	2	2	4	6	118	1	4	5	5	5	168	2	3	3	4	6	218	3	4	4	4	5
19	1	1	1	5	5	69	1	2	2	5	5	119	1	4	5	5	6	169	2	3	3	5	5	219	3	4	4	4	6
20	1	1	1	5	6	70	1	2	2	5	6	120	1	4	5	6	6	170	2	3	3	5	6	220	3	4	4	5	5
21	1	1	1	6	6	71	1	2	2	6	6	121	1	4	6	6	6	171	2	3	3	6	6	221	3	4	4	5	6
22	1	1	2	2	2	72	1	2	3	3	3	122	1	5	5	5	5	172	2	3	4	4	4	222	3	4	4	6	6
23	1	1	2	2	3	73	1	2	3	3	4	123	1	5	5	5	6	173	2	3	4	4	5	223	3	4	5	5	5
24	1	1	2	2	4	74	1	2	3	3	5	124	1	5	5	6	6	174	2	3	4	4	6	224	3	4	5	5	6
25	1	1	2	2	5	75	1	2	3	3	6	125	1	5	6	6	6	175	2	3	4	5	5	225	3	4	5	6	6
26	1	1	2	2	6	76	1	2	3	4	4	126	1	6	6	6	6	176	2	3	4	5	6	226	3	4	6	6	6
27	1	1	2	3	3	77	1	2	3	4	5	127	2	2	2	2	2	177	2	3	4	6	6	227	3	5	5	5	5
28	1	1	2	3	4	78	1	2	3	4	6	128	2	2	2	2	3	178	2	3	5	5	5	228	3	5	5	5	6
29	1	1	2	3	5	79	1	2	3	5	5	129	2	2	2	2	4	179	2	3	5	5	6	229	3	5	5	6	6
30	1	1	2	3	6	80	1	2	3	5	6	130	2	2	2	2	5	180	2	3	5	6	6	230	3	5	6	6	6
31	1	1	2	4	4	81	1	2	3	6	6	131	2	2	2	2	6	181	2	3	6	6	6	231	3	6	6	6	6
32	1	1	2	4	5	82	1	2	4	4	4	132	2	2	2	3	3	182	2	4	4	4	4	232	4	4	4	4	4
33	1	1	2	4	6	83	1	2	4	4	5	133	2	2	2	3	4	183	2	4	4	4	5	233	4	4	4	4	5
34	1	1	2	5	5	84	1	2	4	4	6	134	2	2	2	3	5	184	2	4	4	4	6	234	4	4	4	4	6
35	1	1	2	5	6	85	1	2	4	5	5	135	2	2	2	3	6	185	2	4	4	5	5	235	4	4	4	5	5
36	1	1	2	6	6	86	1	2	4	5	6	136	2	2	2	4	4	186	2	4	4	5	6	236	4	4	4	5	6
37	1	1	3	3	3	87	1	2	4	6	6	137	2	2	2	4	5	187	2	4	4	6	6	237	4	4	4	6	6
38	1	1	3	3	4	88	1	2	5	5	5	138	2	2	2	4	6	188	2	4	5	5	5	238	4	4	5	5	5
39	1	1	3	3	5	89	1	2	5	5	6	139	2	2	2	5	5	189	2	4	5	5	6	239	4	4	5	5	6
40	1	1	3	3	6	90	1	2	5	6	6	140	2	2	2	5	6	190	2	4	5	6	6	240	4	4	5	6	6
41	1	1	3	4	4	91	1	2	6	6	6	141	2	2	2	6	6	191	2	4	6	6	6	241	4	4	6	6	6

s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5	s	u_1	u_2	u_3	u_5	u_5
42	1	1	3	4	5	92	1	3	3	3	3	142	2	2	3	3	3	192	2	5	5	5	5	242	4	5	5	5	5
43	1	1	3	4	6	93	1	3	3	3	4	143	2	2	3	3	4	193	2	5	5	5	6	243	4	5	5	5	6
44	1	1	3	5	5	94	1	3	3	3	5	144	2	2	3	3	5	194	2	5	5	6	6	244	4	5	5	6	6
45	1	1	3	5	6	95	1	3	3	3	6	145	2	2	3	3	6	195	2	5	6	6	6	245	4	5	6	6	6
46	1	1	3	6	6	96	1	3	3	4	4	146	2	2	3	4	4	196	2	6	6	6	6	246	4	6	6	6	6
47	1	1	4	4	4	97	1	3	3	4	5	147	2	2	3	4	5	197	3	3	3	3	3	247	5	5	5	5	5
48	1	1	4	4	5	98	1	3	3	4	6	148	2	2	3	4	6	198	3	3	3	3	4	248	5	5	5	5	6
49	1	1	4	4	6	99	1	3	3	5	5	149	2	2	3	5	5	199	3	3	3	3	5	249	5	5	5	6	6
50	1	1	4	5	5	100	1	3	3	5	6	150	2	2	3	5	6	200	3	3	3	3	6	250	5	5	6	6	6
																								251	5	6	6	6	6

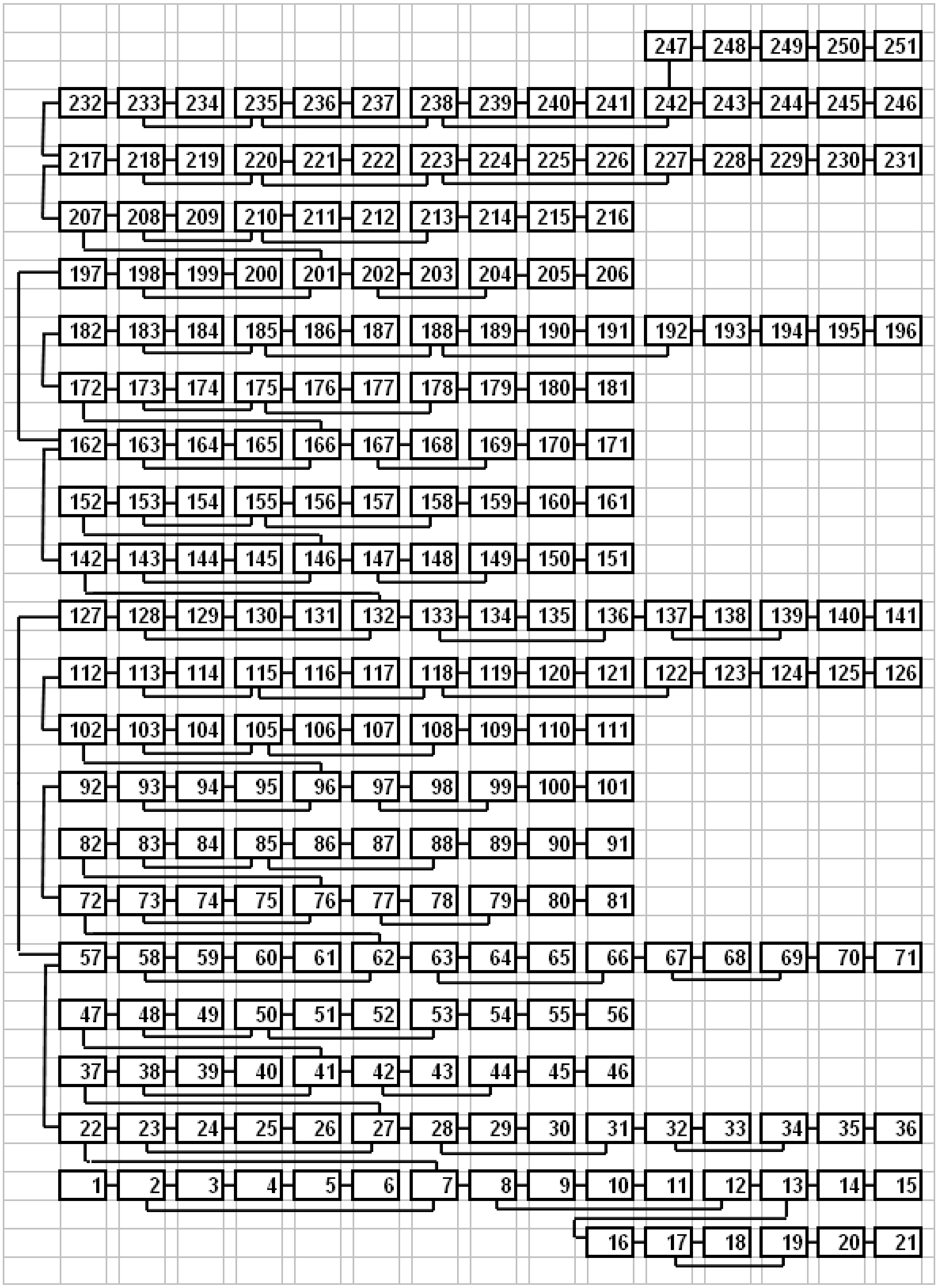


Figure 5-2: Flow-Graph of Simulated Interference Situation States

5.1.2.4 Model Accuracy Criteria

5.1.2.4.1 General

To prove that the test algorithm performance efficiencies estimated in fixed states of an interference situation are applicable to the overall range of varying interference states, the probability P_{rel} of the predicted state-to-state transitions of interference situation should be calculated (see 5.1.1.1, 5.1.1.5, and 5.1.1.6) from the formula:

$$P_{rel} = N_{exp}/N_{tot} \ , \tag{8}$$

where:

N_{exp} is a number of state-to-state transitions of interference situation (worse);

N_{tot} is a total number of the test state-to-state transitions of interference situation (worse).

5.1.2.4.2 Discussion

It is believed that, with P_{rel} being equal to 0,73–0,80, the respective reliability of obtained results can be achieved through selecting u_{max}, s_{max} and γ_{thr}.

A small probability value P_{rel} ($P_{rel} < 0{,}73$) testifies to roughness of the applied model. If $P_{rel} = 1$, the model is excessively accurate (complicated). Hence, it follows that with

$$0{,}73 < P_{rel} < 1 \tag{9}$$

the algorithms' performance efficiency is estimated at a high credibility level and the model optimal accuracy is achieved.

5.2 TECHNIQUES TO EVALUATE CAPABILITIES OF THE PROPOSED METHOD FOR CORRELATED DATA GENERATION

5.2.1 TECHNIQUE FOR COMPARATIVE EVALUATION OF CORRELATED DATA GENERATION ALGORITHMS

5.2.1.1 To compare performance of algorithms for correlated data generation in an attempt to choose the most efficient algorithms, the respective evaluation technique should be developed. The method described hereunder should be applied (selected). The technique implementation steps are described below.

a) The basic algorithms A_{j1}, A_{j2}, …, A_{jr} shall be selected for the test algorithm A_i.

b) A model (data blocks, characteristics of interference situation states, and a graph, such as shown in figure 5-1, table 5-1, and figure 5-2) of an error source in diverse channels shall be developed (selected). The model having parameters shown in figure 5-1, table 5-1, and figure 5-2 shall be applied if its parameters coincide with those of the required model.

c) The interference situation states shall be simulated through combining data blocks (using data blocks and characteristics of interference situation states defined according to b) and shown in figure 5-1 and table 5-1).

d) A number of invalid test data in correlated data blocks generated via basic algorithms A_{j1}, A_{j2}, …, A_{jr} and test algorithm A_i shall be calculated (initial data blocks shall be combined in compliance with characteristics of interference situation states, as shown in table 5-1).

e) Comparative characteristics of errors, (γ), computed according to d), shall be calculated from the formula (equation 5). When calculating comparative characteristics of errors, $\gamma(s_q, s_p)$, the interference situation states s_q and s_p, shall be selected relying on the condition (equation 4) and according to the graph whose analog is shown in figure 5-2.

f) The rated estimates (E) of data reliability shall be calculated from the formula (equation 3) or (equation 6) (depending on the method application purpose). The estimates $E(s_q, s_p)$ (calculated from the equation 6) shall be used to evaluate the accuracy (to justify the application) of the model of errors in diverse channels developed according to b) (equation 9).

g) The rated estimates (E) calculated according to f) shall be analyzed to obtain the overall reliability factor characterizing the test algorithm A_i in terms of the correlated data reliability ensured by this algorithm as follows: 'better', 'worse', or 'much the same', as compared to the basic algorithms A_{j1}, A_{j2}, …, A_{jr}.

5.2.1.2 The technique application results are presented in B3.1 (the test algorithm, A_4, basic algorithms, auto-selection A_a, and majorization A_m).

5.2.2 TECHNIQUE TO DEFINE THE RATIONAL CONTENT OF THE RECEPTION AND RECORD STATION (DIVERSE CHANNELS)

5.2.2.1 It is often the case when a content of diverse channels (though relatively few in number) ensuring a high data reliability is not evident. Therefore to define a rational content of diverse channels is a current problem (see B3.2). It should be solved by using the technique presented in 5.2.2.2.

5.2.2.2 The technique implementation steps are described below.

a) Options involving the Reception & Record Stations (RRS) shall be defined for their subsequent evaluation. Each RRS set for each specified diversion option (S_{an_1}, …, S_{an_m}) shall be a subset of the basic option (S_b) (i.e., $S_{an\_r} \subset S_b$, $r = 1, \ldots, m$). The basic option S_b shall include all RRS for all diverse channels.

b) Reliabilities of correlated data obtained via RRS of the basic option (S_b) and options defined in a) (S_{an_1}, …, S_{an_m}) shall be compared.

c) Options whose implementation results in a *slight* deterioration of correlated data reliability shall be identified among those defined according to a) (S_{an_1}, …, S_{an_m}). The identified options (S_{rat_1}, …, S_{rat_h}, $h \le m$) shall be referred to as rational.

d) A number of rationale options (g_{rat}) and a corresponding number of basic options (g_b) shall be calculated.

e) Validity of each option defined according to a) (S_{an_1}, …, S_{an_m}) shall be determined from the formula:

$$Q_r = g_{rat\_r} / g_b , \quad r = 1, \ldots, m. \tag{10}$$

f) Relying on the estimates (equation 10), rationality of the options defined according to a) (S_{an_1}, …, S_{an_m}) shall be reasoned.

NOTE – Following from the specifics relative to the choice of a rational content of diverse channels, the proper explanations are presented in B3.3.

ANNEX A

OPTIMAL WEIGHT CHARACTERISTICS OF RELIABILITY FOR IMPLEMENTING CORRELATED DATA GENERATION ALGORITHMS

(NORMATIVE)

A1 WCR FOR ALGORITHM A_4

Table A-1: WCR for Algorithm A_4 (n = 5)

k	$k^{//}$	W_1	W_2	W_3	W_4	W_5	k	$k^{//}$	W_1	W_2	W_3	W_4	W_5
1	166	1	10	20	30	62	84	83	11	30	50	60	72
2	165	11	20	30	40	92	85	82	11	40	50	60	72
3	164	11	20	30	40	82	86	81	11	40	60	70	82
4	163	21	30	40	50	102	87	80	31	40	50	60	82
5	162	31	40	50	60	122	88	79	51	60	80	90	122
6	161	41	50	60	70	142	89	78	21	30	40	50	62
7	160	1	10	20	30	52	90	77	41	50	70	80	102
8	159	11	20	30	40	72	91	76	21	30	50	60	72
9	158	11	20	40	50	82	92	75	31	40	60	70	82
10	157	1	10	20	30	42	93	74	21	40	50	70	82
11	156	21	30	40	50	92	94	73	11	20	30	40	42
12	155	11	20	30	40	62	95	72	21	30	50	60	62
13	154	11	30	50	70	102	96	71	41	60	70	80	112
14	153	21	40	50	80	112	97	70	61	80	100	110	152
15	152	11	30	40	60	82	98	69	31	50	60	70	92
16	151	1	20	30	50	62	99	68	41	60	80	90	112
17	150	31	40	50	60	112	100	67	31	50	80	90	112
18	149	21	30	40	50	82	101	66	41	60	90	100	122
19	148	21	50	80	90	142	102	65	21	40	50	60	72
20	147	21	40	50	70	102	103	64	41	70	90	100	122
21	146	21	50	60	90	122	104	63	21	40	60	70	82
22	145	21	60	70	110	142	105	62	31	50	70	80	92

k	$k^{//}$	W_1	W_2	W_3	W_4	W_5	k	$k^{//}$	W_1	W_2	W_3	W_4	W_5
23	144	31	60	70	80	122	106	61	21	50	60	70	82
24	143	21	50	60	70	102	107	60	41	80	100	110	132
25	142	11	40	50	60	82	108	59	11	30	40	50	52
26	141	1	30	40	50	62	109	58	21	40	60	70	72
27	140	1	10	30	40	52	110	57	41	50	60	130	152
28	139	11	20	60	70	92	111	56	31	40	50	110	122
29	138	11	20	50	70	82	112	55	41	50	70	150	162
30	137	1	10	20	30	32	113	54	31	50	60	130	142
31	136	21	30	70	110	122	114	53	21	30	40	90	92
32	135	11	20	40	70	72	115	52	41	50	60	90	112
33	134	21	30	90	100	132	116	51	31	40	50	80	92
34	133	11	20	50	60	72	117	50	41	50	70	110	122
35	132	11	30	60	80	92	118	49	31	50	60	100	112
36	131	21	30	70	90	102	119	48	21	30	40	70	72
37	130	11	20	40	60	62	120	47	41	50	60	80	102
38	129	21	40	90	100	122	121	46	41	50	60	70	92
39	128	11	30	60	70	82	122	45	51	60	70	80	102
40	127	1	20	30	40	42	123	44	31	40	50	70	82
41	126	21	30	70	80	92	124	43	31	40	50	60	72
42	125	31	40	90	100	112	125	42	41	50	60	70	82
43	124	11	20	40	50	52	126	41	41	50	70	100	112
44	123	21	30	60	70	72	127	40	41	50	70	80	92
45	122	31	50	110	120	142	128	39	51	60	80	90	102
46	121	41	60	130	140	162	129	38	31	50	60	80	92
47	120	21	40	80	90	102	130	37	41	60	70	90	102
48	119	31	50	100	110	122	131	36	51	70	80	100	112
49	118	11	30	50	60	62	132	35	21	30	40	60	62
50	117	21	40	70	80	82	133	34	21	30	40	50	52
51	116	21	30	70	80	132	134	33	31	40	50	60	62
52	115	11	20	50	60	92	135	32	61	90	100	120	162
53	114	11	20	60	70	102	136	31	51	70	80	90	122
54	113	21	30	50	70	102	137	30	61	80	90	100	132
55	112	11	20	40	50	72	138	29	51	80	90	110	142

k	$k^{//}$	W_1	W_2	W_3	W_4	W_5	k	$k^{//}$	W_1	W_2	W_3	W_4	W_5
56	111	11	20	50	60	82	139	28	41	60	70	80	102
57	110	21	30	40	70	92	140	27	51	70	80	90	112
58	109	11	20	30	50	62	141	26	61	80	110	120	142
59	108	11	20	40	60	72	142	25	71	90	120	130	152
60	107	11	30	40	70	82	143	24	31	60	70	90	102
61	106	11	30	50	80	92	144	23	31	50	60	70	82
62	105	31	40	50	100	122	145	22	41	60	70	80	92
63	104	21	30	40	80	92	146	21	41	80	100	130	142
64	103	21	30	50	90	102	147	20	41	60	80	90	102
65	102	21	40	50	100	112	148	19	51	70	90	100	112
66	101	11	20	30	60	62	149	18	31	70	80	100	112
67	100	21	30	50	60	92	150	17	31	60	70	80	92
68	99	21	30	60	70	102	151	16	41	70	80	90	102
69	98	21	30	70	80	112	152	15	21	50	60	80	82
70	97	21	30	40	50	72	153	14	21	40	50	60	62
71	96	11	20	30	40	52	154	13	31	50	60	70	72
72	95	11	20	40	50	62	155	12	61	70	80	90	112
73	94	11	30	40	60	72	156	11	51	60	70	80	92
74	93	11	30	50	70	82	157	10	61	70	90	100	112
75	92	31	40	50	70	92	158	9	71	90	100	120	132
76	91	21	30	40	60	72	159	8	41	50	60	70	72
77	90	21	30	50	70	82	160	7	71	90	100	110	142
78	89	21	40	50	80	92	161	6	61	80	90	100	122
79	88	11	20	30	50	52	162	5	81	100	130	140	162
80	87	31	50	60	70	102	163	4	51	70	80	90	102
81	86	21	40	50	60	82	164	3	61	80	100	110	122
82	85	21	40	70	80	102	165	2	51	80	90	100	112
83	84	11	30	40	50	62	166	1	41	60	70	80	82

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

Table A-2: WCR for Algorithm A_4 (n = 4)

$k^{//}$	W_2	W_3	W_4	W_5
1	20	30	41	42
2	30	40	51	62
3	20	50	61	72
4	10	20	31	32
5	40	50	71	102
6	20	60	71	92
7	20	30	51	62
8	10	30	41	52
9	20	40	51	82
10	10	20	31	42
11	20	30	41	82
12	10	20	31	52
13	20	30	41	92
14	10	20	31	62

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

Table A-3: WCR for Algorithm A_4 (n = 3)

$k^{//}$	W_3	W_4	W_5
1	70	80	82
2	20	30	62

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

A2 WCR FOR ALGORITHM A_{42}

Table A-4: WCR for Algorithm A_{42} ($n = 5$)

k	$k^{//}$	W_1	W_2	W_3	W_4	W_5
1	7	1	2	3	4	11
2	6	1	2	3	4	9
3	5	1	2	3	4	7
4	4	1	2	6	7	8
5	3	1	2	3	4	5
6	2	3	4	5	6	7
7	1	5	6	7	8	9

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

Table A-5: WCR for Algorithm A_{42} ($n = 4$)

k	$k^{//}$	W_2	W_3	W_4	W_5
1	3	2	3	4	11
2	2	2	3	5	9
3	1	2	6	7	8

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

Table A-6: WCR for Algorithm A_{42} ($n = 3$)

k	$k^{//}$	W_3	W_4	W_5
1	2	3	4	11
2	1	7	8	9

NOTE – Value $k^{//}$ is an application priority number for the WCR combination.

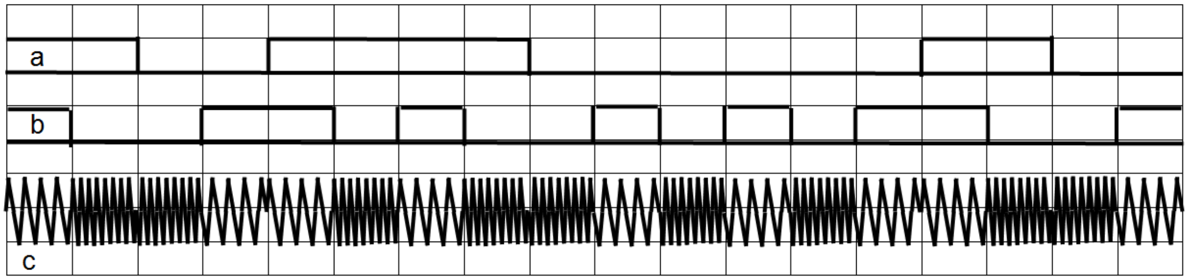
ANNEX B

REQUIRED EXPLANATIONS

(INFORMATIVE)

B1 EXPLANATIONS FOR SECTION 2—IMPROVEMENT OF CORRELATED DATA RELIABILITY VIA NOISELESS MODULATION TECHNIQUES

A set of signals forms a system of signals (reference [C3]). Each signal can be represented as a set of elements or elementary signals (reference [C3]). If $u_{ij\ldots r}$ is an elementary signal (analog implementation of digital signal, symbol), a position of each sequence i, j, …r of subscript characters corresponds to a conventional numeric symbol identifying an information parameter of signal element $u_{ij\ldots r}$, and the character itself contains a quantitative characteristic of this parameter. The amplitude, phase, and frequency can be information parameters. These parameters contain information carried by a signal and have a specified number of values (levels). For signal $u_{ij\ldots r}$, i is a number of levels for the first information parameter; j, for the second; …; r, for n_{par}. For example (figure B-1), for element u_{ij}, the first information parameter is a signal phase, 0°, and 180° are its fixed values; frequency is the second information parameter; and f_1, f_2 are conventions for the applied frequencies. The first information parameter belongs to the first information stream (figure B-1a); the second, to the second information stream (figure B-1b). Each parameter contains information on the transferred data. There are modulation coding techniques with the transferred data splitting into more than two streams.



a) binary code levels for the first stream;
b) binary code levels for the secondary stream;
c) modulation coding of carrier.

Figure B-1: One of the Modulation Coding Techniques

It is demonstrated (reference [C1]) that, as the number of levels of a single information parameter decreases, so does the volume of elementary data transferred via an elementary signal. An elementary data volume can be kept stable by increasing the number of information parameters n_{pa}.

The increased number n_{par} ($n_{par} \geq 2$) is applied to the noiseless modulation techniques. Specifically, the combination of amplitude and phase modulation techniques are widely used (two-dimensional signals are generated, with $n_{par} = 2$). Value n_{par} is increased so that $\Delta F_1 \approx \Delta F_2$, with $n_{par\_1} > n_{par\_2}$ (where ΔF_i is a frequency band of elementary signals in the *i*-n signal system). That is, the noise energy concentrated in the valid signal frequency band is hardly changed with an increased number of information parameters. It is distributed among components of the n_{par}-dimensional space.

If elementary signals $u_{ij}(t)$ ($n_{par} = 2$) are mutually orthogonal, then (reference [C3]):

$$\begin{aligned}
&\int_{-\infty}^{\infty} u_{ij}(t)u_{ik}(t)dt = 0 \,, j \neq k \quad , \\
&\int_{-\infty}^{\infty} u_{ij}(t)u_{rj}(t)dt = 0 \,, i \neq r \quad , \\
&\int_{-\infty}^{\infty} u_{ij}(t)u_{rk}(t)dt = 0 \;\;, i \neq r \,, j \neq k \,, \\
&\int_{-\infty}^{\infty} u_{ij}(t)u_{ij}(t)dt = E_0 \,,
\end{aligned} \tag{11}$$

where:

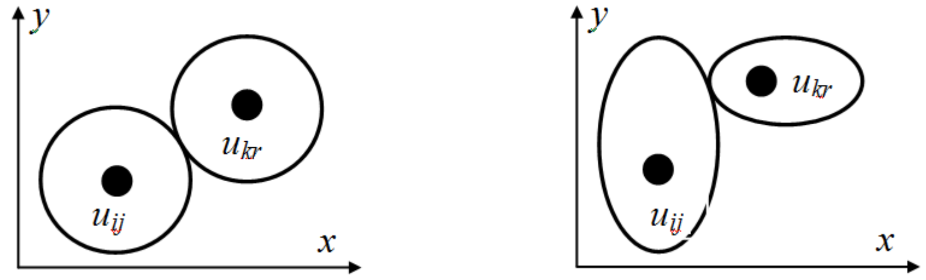
E_0 is signal $u_{ij}(t)$ energy, with different values of i and j ($i \in I, j \in J$);

The required conditions shall be created for the optimal processing of signals (reference [C3]).

Signals $u_{ij}(t)$ ($i \in I, j \in J$) shall be equally spaced (reference [C3]). Distances between any two points identifying ideal (undistorted with interferences) elementary signals in a two-dimensional space shall be equal;

The equidistance property constitutes evidence that the n_{par} –dimensional space is rationally used to assure a good resistance to interference.

NOTE – Figure B-2 shows a geometrical representation of equally spaced elementary signals.



a) symmetrical channel b) asymmetrical channel

Figure B-2: Geometrical Representation of Zones Related to Elementary Signals u_{ij} and u_{kr}

If signals are equally spaced and energy impacting the signals is also uniformly distributed in the n_{par}-dimensional space, a zone of the probable position of each signal is restricted to the n_{par}-dimensional sphere with a point in its center identifying an ideal elementary signal. In this case, the communication channel is symmetrical (reference [C5]) (figure B-2a), and the probability of misidentifying an elementary data is independent of its transferred value. As the alphabet of elementary signals increases through increasing n_{par}, the distance between the center of the sphere and the point on it remains unchanged. Herewith, a number of spheres is not increasing, and stable noise energy concentrated in the valid signal frequency band is distributed in space extended through increasing $n_{par.}$ A signal-to-noise ratio is not reduced. If the points corresponding to ideal elementary signals, $u_{ij}(t)$ and $u_{rk}(t)$, are matched (e.g., through biasing $u_{ij}(t)$ to $u_{rk}(t)$), their zones of the most probable signal occurrence will coincide. With such a uniform distribution of noise energy over the n_{par}-dimensional sphere (in this case $n_{par} = 2$), the optimum signal reception (reference [C3]) (according to equation 11) and the equally probable transfer of the alphabet of elementary signals, the following expressions will be justified:

$$
\begin{aligned}
&P(e_{i\_er}|u_{ij\_tran}) = P(e_{i\_er}|u_{ik\_tran}),\ j \neq k\ , \\
&P(e_{j\_er}|u_{ij\_tran}) = P(e_{j\_er}|u_{rj\_tran}),\ i \neq r, \\
&P(e_{i\_er}|u_{ij\_tran}) = P(e_{j\_er}|u_{ij\_tran}),
\end{aligned}
\tag{12}
$$

where:

$P(e_{i\_er}|u_{ij\_tran})$ is the misidentification probability of a transferred data e_i contained in a transferred elementary signal $u_{ij}(t)$ that corresponds to the i-n level of the first parameter of this signal (misidentification is designated as ‘*er*’);

$P(e_{j\_er}|u_{ij\_tran})$ is the misidentification probability of a transferred data e_j contained in a transferred elementary signal $u_{ij}(t)$ corresponding to the j-n level of the second parameter of this signal.

As follows from equation 12, $P(e_{i\_er}|u_{ij\_tran}) = P(e_{i\_er}|u_{ik\_tran}) = P(e_{j\_er}|u_{ij\_tran}) = P(e_{j\_er}|u_{rj\_tran})$, or

$$P(e_{i\_er}|u_{ik\_tran}) = P(e_{j\_er}|u_{rj\_tran}) \ . \tag{13}$$

That is, the probability of misidentifying data that are selected from elementary signals and correspond to any levels, both of the first and the second information parameter of this signal, is independent of values of data contained in the respective transferred elementary signals.

In the given case, it would be reasonable to expect that the nature of error dependencies is similar in diverse channels for data corresponding to both the first and the second information parameters (the first and the second information streams—see figure B-1). Thus correlated data can be obtained via one-bit data corresponding to the first and the second information parameter and acquired from diverse channels (data from the first and the second streams, respectively—see figure B-1).

To implement the aforesaid algorithms, A_4 and $A_{42,}$ (see 4.2.2), the optimum set of WCR combinations (q) shall be calculated, with the WCR volume equal to $n \times q$ (see annex A). It depends on the alphabet of elementary data being used (i.e., a volume of a data selected from an analog implementation of a digital signal). Algorithm A_4 is intended for two-bit elementary data handling, with $n = 5$, $q = 166$; algorithm A_{42}, for one-bit elementary data handling, with $n = 5$, $q = 7$. It is obvious that the amount of calculations required to implement algorithm A_{42} is far less compared to algorithm A_4. When the n_{par}-dimensional (multi-position) signals are used, even with a relatively small number of information parameters (even with a small volume of an elementary data), a number of WCR combinations q may turn out unacceptably large in terms of the algorithm's practical use for correlated data generation. Therefore the rule presented hereinbefore, based on the equidistance of signals, is of current concern. As evidenced by this rule, a volume of the WCR optimum set can be drastically reduced using algorithm A_{42}. When multi-position (m-position) signals are used, with $m > 2$, the rational use of algorithms for correlated data generation is possible only via algorithm A_{42}. It should be also remarked that the condition required to ensure the equidistance of signals entails a choice of a modulation technique.

B2 EXPLANATION TO SECTION 4

B2.1 RATIONALE FOR APPLICATION PRIORITY OF TECHNIQUES IMPROVING DATA RELIABILITY

The next approach to selecting a better strategy is presented below (reference [C1]).

From strategies u_1 and u_2 the most efficient one is selected (ensuring the highest reliability), with a number of errors $N_{er\_\Sigma\_1}$ in a data block after using strategy u_1 and $N_{er\_\Sigma\_2}$ (after using strategy u_2), with

$$u_1 = \langle Met_{\mathrm{CD}}, Met_{\mathrm{DR}}\rangle,\ \ u_2 = \langle Met_{\mathrm{DR}}, Met_{\mathrm{CD}}\rangle\ , \tag{14}$$

where:

Met_{CD} is a noiseless coding-decoding of data;

Met_{DR} is a diverse reception method.

When using the Met_{CD} methods, a number of errors N_{er_det} is detected, and a number of errors N_{er_rem} in a data block is detected and corrected, with $N_{er\_det} \geq N_{er\_rem}$. It is known (reference [C2]) that either the incomplete or complete decoder is used for decoding. If $N_{er\_i} > N_{er\_rem}$ in a data block acquired from the i-n diverse channel, no decoding is done with an incomplete decoder; for decoding, a complete decoder is required. To make the case simple and certain, it can be assumed that the decoder is incomplete; hence:

$$\begin{aligned} &N_{er\_\mathrm{CD}\_i} = 0, \text{ with } N_{er\_i} \leq N_{er\_rem}\ , \\ &N_{er\_\mathrm{CD}\_i} = N_{er\_i}\ , \text{ with } N_{er\_i} > N_{er\_rem}\ , \end{aligned} \tag{15}$$

where:

N_{er_i} ($N_{er\_\mathrm{CD}\_i}$) is a number of errors in a data block of the i-n diverse channel before (after) data decoding by the Met_{CD} method.

When the Met_{DR} methods are used, the number of errors in data blocks of the first, the second, …, n diverse channels is N_{er_1}, N_{er_2}, …., N_{er_n}, respectively, and N_{er_cor} in a correlated data block.

$$N_{er\_1} \geq N_{er\_2} \geq \cdots N_{er\_n}\ . \tag{16}$$

When applying rational methods $Met_{\mathrm{DR,}}$

$$0 \leq N_{er\_cor} \leq N_{er\_n}\ . \tag{17}$$

Further, all considered methods Met_{DR} will be assumed rational.

Efficiency of strategies u_1 and u_2 depends on ratios between values N_{er_rem}, N_{er_cor}, and N_{er_n}:

$$N_{er\_rem} < N_{er\_cor} < N_{er\_n} \, , \tag{18}$$

$$N_{er\_rem} < N_{er\_n} < N_{er\_cor} \, , \tag{19}$$

$$N_{er\_cor} < N_{er\_rem} < N_{er\_n} \, , \tag{20}$$

$$N_{er\_cor} < N_{er\_n} < N_{er\_rem} \, , \tag{21}$$

$$N_{er\_n} < N_{er\_rem} < N_{er\_cor} \, , \tag{22}$$

$$N_{er\_n} < N_{er\_cor} < N_{er\_rem} \, . \tag{23}$$

Ratios between equations 19, 22, and 23 are impossible as inconsistent with equation 17.

For the conditions corresponding to ratio (equation 21), $N_{er\_\Sigma\_1} = N_{er\_\Sigma\_2} = 0$. When strategy u_1 is applied, with equation 15 considered, $N_{er\_\mathrm{CD}\_n} = 0$, and, with $N_{er\_\mathrm{CD}\_n} = 0$ and equation 17 considered, $N_{er\_cor} = 0$. When strategy u_2 is applied, it appears (equation 21) that $N_{er\_cor} < N_{er\_rem}$, therefore all errors are corrected by applying the Met_{CD} method after the Met_{DR} method (according to equation 15).

For the conditions corresponding to ratio (equation 20) $N_{er\_\Sigma\_2} = 0$, as, with $N_{er\_cor} < N_{er\_rem}$ (equation 20), correlated data decoded by the Met_{CD} method will not contain errors (according to equation 15). If the incomplete decoder is used (the condition accepted earlier), $N_{er\_\Sigma\_1} \leq N_{er\_n}$, as $N_{er\_rem} < N_{er\_n}$ (equation 20) and no decoding is done (equation 15). Strategy u_2 is more efficient.

For the conditions corresponding to ratio (equation 18) $N_{er\_\Sigma\_1} = N_{er\_\Sigma\_2} = N_{er\_cor}$ (equation 18), as $N_{er\_n} > N_{er\_rem}$, $N_{er\_cor} > N_{er\_rem}$ (equation 18); in this case, no decoding is done (see equation 15), no matter what strategy, u_1 or $u_{2,}$ is applied.

It is clear that, if the condition (equation 20) is fulfilled, strategy u_2 is more efficient. As for all other cases, strategies u_1 and u_2 are equal. As a result, strategy u_2 is the best suited for the incomplete decoder.

B2.2 EXAMPLE FOR PRACTICAL APPLICATION OF ALGORITHM A_4

N1 Функц СЕК: 123 Изм: 64000 ММ: 12 МБ: 8000 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199__.994 Dt: 0____.000

Y: 512__ Двоичные единицы = 116 , время (сек) = 130.947 из файла: E:\EMN\8к82км от 5.11.08\1 ступень\Поток А\11А_отр_ЛС_У

N2 Функц СЕК: 123 Изм: 64000 ММ: 12 МБ: 8000 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199 .994 Dt: 0 .000

Y: 512__ Двоичные единицы = 508 , время (сек) = 127.448 из файла: E:\EMN\8к82км от 5.11.08\1 ступень\Поток А\15А_отр_ЛС_У

N3 Функц СЕК: 123 Изм: 64000 ММ: 12 МБ: 8000 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199 .994 Dt: 0 .000

Y: 512__ Двоичные единицы = 474 , время (сек) = 126.806 из файла: E:\EMN\8к82км от 5.11.08\1 ступень\Поток А\26А_отр_ЛС_У

N2 Функц СЕК: 123 Изм: 64000 ММ: 12 МБ: 8000 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199 .994 Dt: 0 .000

Y: 512__ Двоичные единицы = , время (сек) = из файла: E:\EMN\8к82км от 5.11.08\1 ступень\Поток А\28А_отр_ЛС_У

N3 Функц СЕК: 123 Изм: 64008 ММ: 12 МБ: 8001 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199 .994 Dt: 0 .000

Y: 512__ Двоичные единицы = , время (сек) = из файла: E:\EMN\8к82км от 5.11.08\1 ступень\Поток А\32А_отр_ЛС_У

N4 Функц СЕК: 123 Изм: 64048 ММ: 12 МБ: 8006 Поток: 1 ОК: 1 КЗ: 4 ЛК: 64 [ст][мл]--> УК-64 Ткп: 199 .994 Dt: 0 .000

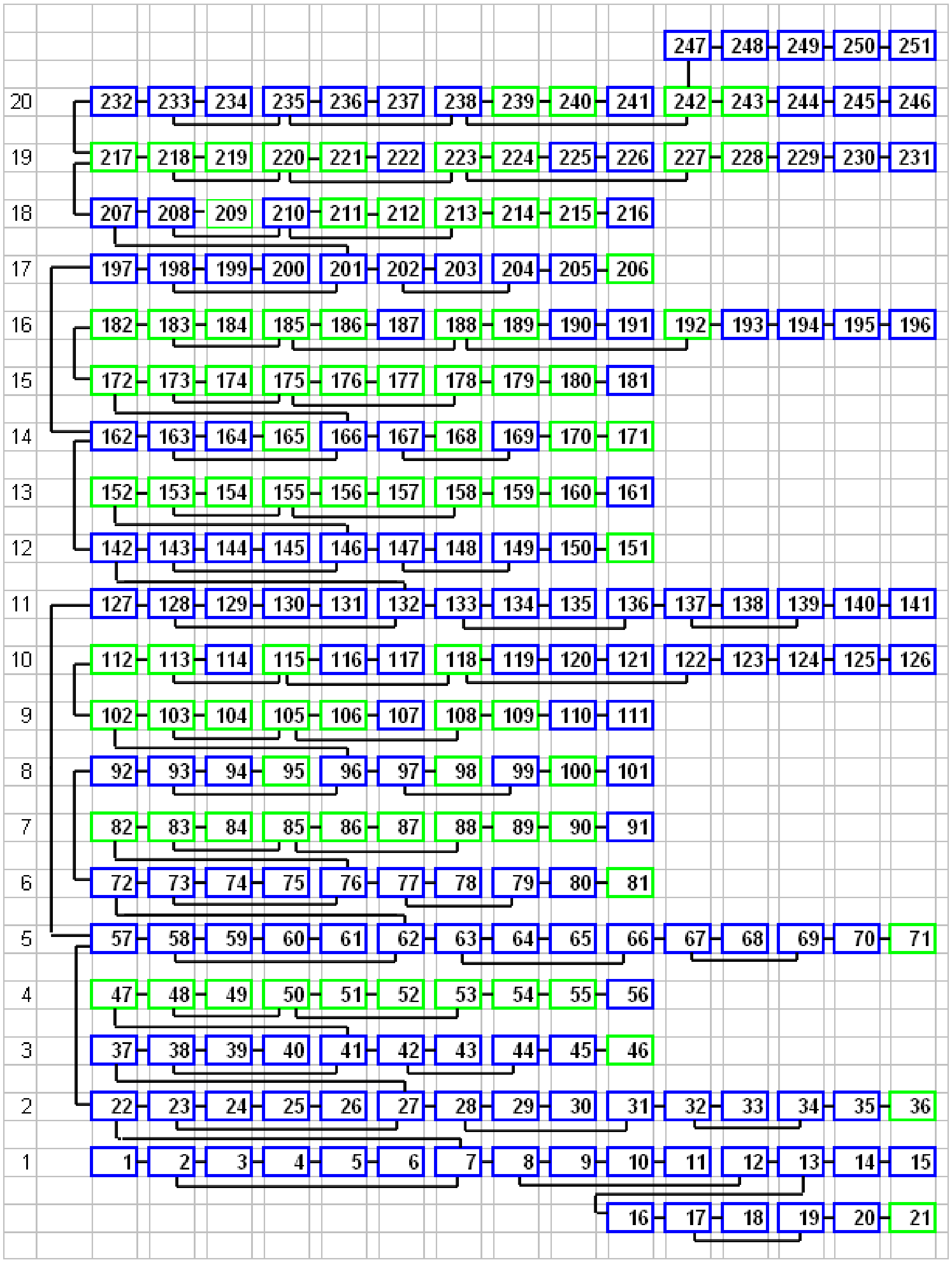
Y: 512__ Двоичные единицы = , время (сек) = из файла: E:\EMN\8к82км от 5.11.08\1 ступень\ОДМ82_1А.tmi

Figure B-3: Results of Algorithm A_4 Application

NOTE – Telemetry from the LV Proton-M first stage, 05.11.2008, $f_{044,\ 090}$, Reception & Record Stations MK-11A, 15A, 26A, 28A, 32A). Graphs shown in the figure (from top to bottom): MK-11A, 15A, 26A, 28A, 32A generated data streams.

B3 EXPLANATIONS TO SECTION 5

B3.1 RESULTS OF ALGORITHM A$_4$ EVALUATION



‘blue’ top, E = 0; ‘green’, E = 1

Figure B-4: Results of Algorithm A$_4$ Evaluation via Simulated Data

B3.2 RATIONALE FOR USING THE RRS (DIVERSE CHANNELS) RATIONAL CONTENT DEFINITION METHOD

As shown by the example of spacecraft TMI (reference [C1]), the correlated telemetry reliability depends on a large number of uncertain factors.

A ‘contribution’ of a separate RRS to the correlated telemetry reliability is manifested concurrently with ‘contributions’ from other mutually redundant RRS. Therefore this ‘contribution’ is not evident. So, in one case (figure B-5a), a time interval with TIL (T_{1_TIL}) is replaced by an interval ($T_{2\_PIL\_2}$) with PIL, and time intervals with PIL ($T_{2\_PIL\_1}$ and $T_{2\_PIL\_3}$) are replaced by intervals with valid TMI; in the other case (figure B-5b), a difference in time of beginning the TIL ($T_{1\_TIL\_1}$ and T_{2_TIL}) intervals gives the ability to supplement RRS_1 TMI in the interval $T_{1\_TIL\_1}$, with $T_{1\_TIL\_1} + T_{1\_TIL\_2} << T_{2\_TIL}$; in the third case (figure B-5c), by using TMI from the obviously worse (in terms of reliability) RRS_2 and RRS_3, as compared to RRS_1, the same reliability can be ensured ($T_{1\_TIL} = T_{2\_TIL} \cap T_{3\_TIL}$), etc. (Here, ‘TIL’ is the total data loss and ‘PIL’ is the partial data loss.)

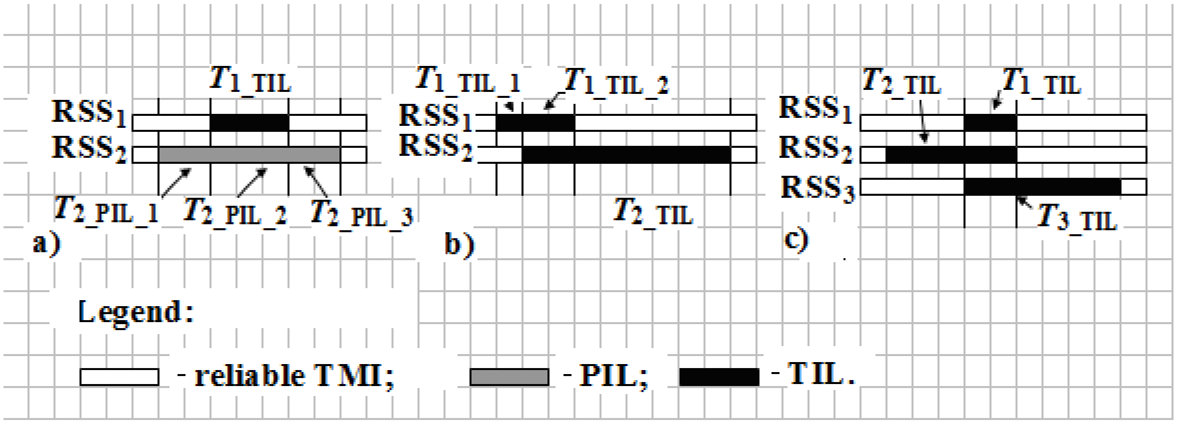


Figure B-5: Examples of RRS Complementarity

By changing a method for correlated data generation (selection of other algorithms), evaluations of options involving RSS may be changed. Reliability of correlated data relies on reliability of data acquired via diverse channels and correcting capabilities of algorithms applied for correlated data generation (see, for example, comparative characteristics of correlated data reliability ensured by algorithm A_4, auto-selection, and majorization, reference [C1], figure B-4). It follows that reliability of data acquired from different diverse channels (options with mutually redundant RRS) also depends on the applied algorithms for correlated data generation. Assuming the technology for correlated data generation is changed (other algorithms for correlated data generation are used), the options involving RSS could be reconsidered.

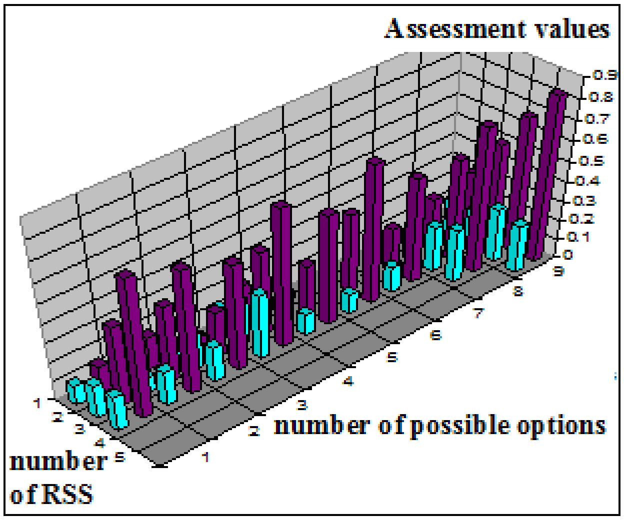
The necessary condition for excluding any RSS from the launch site telemetry facilities shall be a stable (i.e., recurring from launch to launch) *minor* degradation (calculated from rated estimates [equation 3, equation 6]) of correlated telemetry data reliability, with TMI from these RSS being ignored while generating the correlated telemetry data.

To define the possibilities for reducing a number of involved RSS, the technique enabling the rational choice of RSS for mutual redundancy has been developed (reference [C1]) (see 5.2.2.2).

B3.3 APPLICATION SPECIFICS OF THE TECHNIQUE FOR RSS (DIVERSE CHANNELS) RATIONAL CONTENT DEFINITION

The correlated data reliability as per 5.2.2.2 b) may be compared in different ways, specifically by using criteria described in 5.1.1. Also compared were time intervals with TIL, PIL, and valid TMI defined from results produced by analyzing the plots of reference parameters (reference [C1]) (see figure B-6).

With the limited capabilities for selecting the basic options (particularly, due to a frequently changing priority of RSS involvement), the bootstrap technique (reference [C4]) is employed (reference [C1]) for the imitative replication of statistic samples (its application results are shown in figure B-6). For example, if at the i-n launch a set of S_{thr_i} RSS is employed for mutual redundancy ($S_{thr\_i} = \{11, 12, 13, 22, 32, 34\}$), at the ($i$+1)-n launch, S_{thr_i+1} ($S_{thr\_i+1} = \{12, 13, 14, 21, 32, 34\}$), and at the ($i$+2)-n launch, S_{thr_i+2} ($S_{thr\_i+2} = \{12, 13, 32, 34, 51\}$), the basic option of involving RSS can be selected on the following condition: $S_b = S_{thr\_i} \cap S_{thr\_i+1} \cap S_{thr\_i+2}$, A $S_b = \{12, 13, 32, 34\}$.



NOTES

1 Dark and light columns of diagram relate to Q_{an_MV} and Q_{an_MSD}, respectively.

2 Composition of basic options:

1) LV Proton, f02, RRS 13Б, 14Б, 26, 31Б;
2) LV Proton f01, RSS 13A, 14A, 25, 31A;
3) LV «Proton», f074, RRS 11, 14, 31, 51;
4) LV «Proton», f154, RSS 12, 13, 32, 34;
5) LV «Soyuz», f094, RRS 12, 14, 22, 32;
6) LV «Soyuz», f154, RRS 11, 12, 21, 34;
7) LV «Soyuz», f154, RRS 12, 13, 32, 34;
8) LV «Soyuz», f154, RRS 11, 12, 13, 32, 34;
9) LV «Soyuz», f154, RRS 11, 12, 13, 21, 32, 34.

Figure B-6: Improvement of Correlated Telemetry Data Reliability

NOTE – Chart shows the IMPROVEMENT of correlated telemetry data reliability, depending on the number of RSS involved for mutual redundancy (Q_{an_MV} are the mean values of options with RSS involved for mutual redundancy; Q_{an_MSD} are the mean-square deviations of options with RSS).

The chart (figure B-6) shows that, as the number of RSS increases, so do the mean values of options, and it becomes evident that reliability is apt to be *sufficiently* improved. Herewith, the relatively large values of the calculated mean-square deviations suggest that the considered options with the involved RSS are unequal.

As follows from the evaluation results (reference [C1]), exclusion from the launch site telemetry facilities of any RSS currently involved for mutual redundancy will make probable the *sufficient* deterioration of correlated telemetry data reliability. The number of RSS needs to be at least 4–6.

ANNEX C

INFORMATIVE REFERENCES

(INFORMATIVE)

[C1] V. L. Vorontsov. *Methods for Diverse Reception of Telemetry Information and Conditions for Their Application in the Process of the Launch Site Telemetry System Evolution*. 2nd ed. Naberezhnye Chelny: Kama State Academy of Engineering and Economics, 2009.

[C2] R. E. Blahut. *Theory and Practice of Error Control Codes*. Reading, Massachusetts: Addison Wesley, 1983.

[C3] L. E. Varakin. *Signal System Theory*. Moscow: Sov. Radio, 1978.

[C4] A. I. Orlov. “Techniques for Data Selection Replication (Bootstrap Methods).” Chapter 11.4 in *Econometrics*. Moscow: Exam, 2002.

[C5] V. Y. Turin. *Transmission of Data via Channels with Memory*. Moscow: Communication, 1977.

NOTE – Normative references are presented in 1.9.