

## NONLINEAR DATA CODING IN WIRELESS SENSOR NETWORKS

Vasyl Yatskiv, Su Jun, Nataliya Yatskiv, Anatoly Sachenko

Ternopil National Economic University  
3 Peremoga Square, Ternopil, Ukraine, 46020  
jazkiv@ukr.net, sjhosix@gmail.com, jatskiv@ukr.net, as@tneu.edu.ua

**Abstract:** A proposed nonlinear method of sensors data coding allows to increase the useful capacity of the data transmission protocols for Wireless Sensor Networks in 3 – 5 times by the integration of sensors data with different digits capacity and reducing the transmissions number.

**Keywords:** residue number system; wireless sensor networks; multilevel coding, protocol.

### 1. INTRODUCTION

Wireless sensor networks (WSN) with the domain of scientific (theoretical and experimental) research are increasingly moving into the region of the practical use in various areas. WSN can capture information about the condition of physical environment, can perform the simple data processing and transmit them to the remote server. There are such tasks among the practical applications of the Wireless Sensor Networks technologies: water resources monitoring, glacier motion, forest fires,

building and manufacture organization, “intelligent” measurements and increasing of the power consumption effectiveness, medicine and public health, sport medicine, patient care and patient monitoring [1].

Wireless Sensor Networks more often use the cluster tree topology in the tasks of monitoring (Fig.1) [1]. In this topology the neighboring nodes are the sources of information and the retransmitters of the data from another nodes simultaneously.

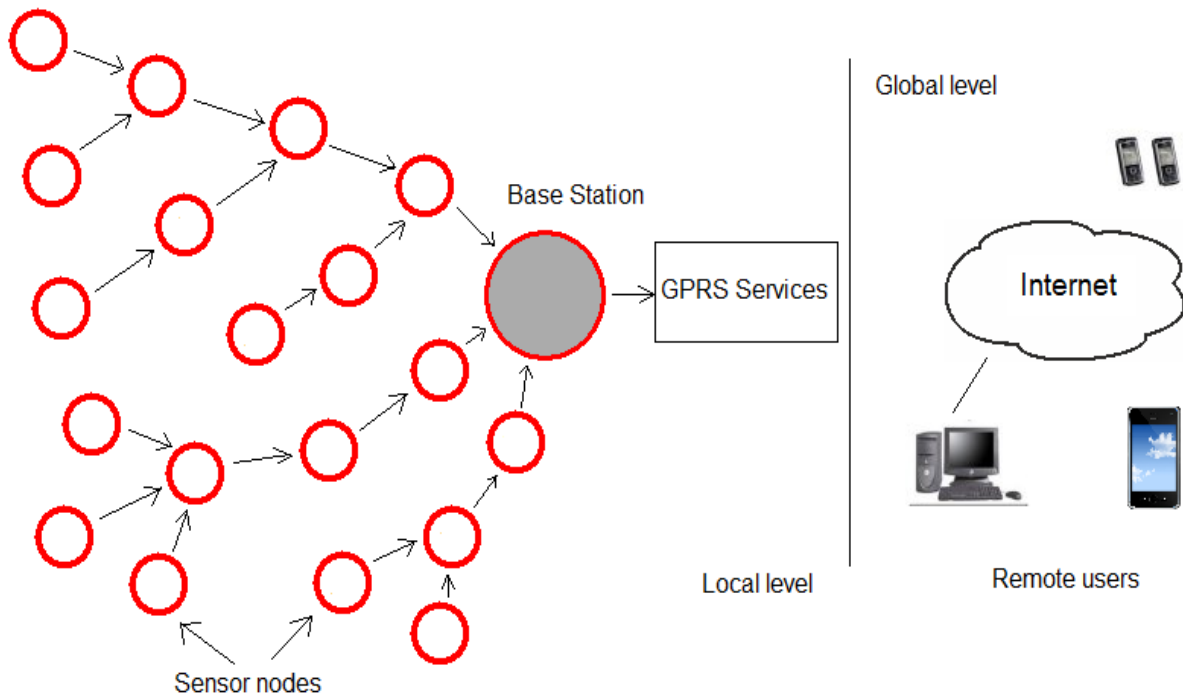


Fig. 1 – The structure of WSN

As a rule the base station transmits data through the wireless communication channel (GSM channel) to the networked remote server.

It is shown in [2] that energy expenses for data transmission in WSN occupy about 70 percent of all expenses. Take into consideration that one bit transmission on radio channel (even ultralow power) is equivalent to the performance of thousand operations by the processor of wireless node [3]. So, the use of the data processing methods in wireless controller is promising. The proposed additional data processing in the nodes is less energy-intensive than redundant data transmission. Taking into account this fact as well as low data transfer rate in WSN (250 Kbit/s) a development of coding methods and effective data transmission protocols oriented on increasing the useful bandwidth of communication channels is the actual issue.

The preliminary analysis has shown that WSN efficiency can be improved due to [4, 5]:

- (i) The reduction of service data in network packet;
- (ii) The redundancy elimination of measured sensors data.

The different methods are used for the sensors data redundancy decreasing. The mobile-agent-based computing model can decrease the data volume in sensor networks using data redundancy elimination by determined methods [6, 7]. But such approach has a limited application area because of restriction on the network architecture, a distance between sensors and cluster core and etc. To eliminate the data redundancy authors [8] proposed to optimize the traffic using data decorrelation. However such technique can be implemented in data flows for close located nodes only where the same physical quantities are measuring, so it has a limited application.

The service data proportion can be improved in the packet's structure increasing a size of data field. However control and monitoring systems based on WSN technology are characterized by short length of information messages. For example a volume of sensor data in measurements of physical quantities (temperature, pressure, humidity, radiation background and etc) has 8-16 bits. To solve this problem the data from different nodes are integrated in one packet. The simplest method of the packets size increasing is data concatenation. If  $\alpha = a_1 \dots a_n$  and  $\beta = b_1 \dots b_m$  words of alphabet A, then concatenation of word A and word B will be the word  $\gamma$  of the same alphabet A:  $\gamma = \alpha \beta = a_1 \dots a_n b_1 \dots b_m$  [9]. However existing methods of data concatenation work properly for sensor data with a same digit capacity only, its application for data with different digit capacity

leads to the essential redundancy increasing.

In this work we are exploring the directions of network protocols improvement by means of increasing network packets size for sensors with different digit capacity.

## 1. THE EVALUATION OF WSN PROTOCOLS STRUCTURE

Let's explore the protocols of the IEEE 802.15.4 standard and protocol SimpliciTI of the Texas Instruments to determine the volume of control information in the structure of the WSN protocols.

The standard IEEE 802.15.4 and ZigBee protocol stack are developed for the coordination of the equipment of different manufacturers and for the simplification of deploying WSN. ZigBee protocol stack is hierarchical model, which is built on the principle of OSI model that defines a networking framework for implementing protocols in seven layers (Open System Interconnection) [10].

ZigBee protocol stack includes layers of the IEEE 802.15.4 standard, the data-link layer, the network layer, application support sublayer and device profiles. The IEEE 802.15.4 standard defines the specifications of the physical layer (PHY) and sub layer the Media Access Control (MAC) for the low-rate wireless networks.

Physical layer (PHY) provides the data transmission service, performs channel selection and energy and signal management functions. The medium access control (MAC) layer manages access to the physical channel and network. It provides different access mechanisms to the physical channel. MAC layer supports "peer-to peer", "star" and "cluster tree" topologies, secure data exchange, and streaming and batch communication. It also defines automatic confirmation of packets reception, provides data transmission in appointed time intervals and uses 128-bit keys to implement its security mechanisms, 16- and 64-bit addressing.

The SimpliciTI protocol, developed by the Texas Instruments is the alternative for the ZigBee protocol stack [11]. The SimpliciTI protocol is destined to the networks with the limited number of nodes and autonomous power supply. This protocol uses physical layer of radio access, elaborated by the Texas Instruments instead of IEEE.802.15.4. The retransmitters and gateways, which are powered by electrical supply network are used for the increasing of distances for networks on the basis of SimpliciTI protocol. The implementation of SimpliciTI protocol requires minimum resources of microcontroller, so the system has low cost in comparison with systems, which used ZigBee protocol. The SimpliciTI protocol supports the transmission of packets with fixed and variable length. The maximum packet size

is 255 byte.

The useful network channel capacity means information rate [5]. The next factors have an influence on useful channel capacity: the interference, waiting of the transmission access, the interframe spaces, the service data units (packet header of MAC and physical layer), confirmation and retransmission of the frame. The size of service data units has most important influence on useful channel capacity in WSN.

The frame structures have been designed to keep the complexity to a minimum while at the same time making them sufficiently robust for transmission on a noisy channel. Each successive protocol layer adds to the structure with layer-specific headers and footers. This standard defines four frame structures [10]:

- a beacon frame, used by a coordinator to transmit beacons;
- a data frame, used for all transfers of data;
- an acknowledgment frame, used for confirming successful frame reception;
- a MAC command frame, used for handling all MAC peer entity control transfers.

Each packet transmitted in WSN contains the control information which is required for the safe data delivery and data processing. According to the IEEE 802.15.4, SimpliciTI service data unit consists of preamble sequence, start of frame delimiter, frame length, frame control, sequence number, addressing fields, and checksum (Fig. 2).

| Preamble Sequence | Start of Frame Delimiter | Frame Length | Addressing Fields | Data Payload | FCS |
|-------------------|--------------------------|--------------|-------------------|--------------|-----|
| 32/64             | 16/32                    | 8            | 8                 | 255          | 16  |

a)

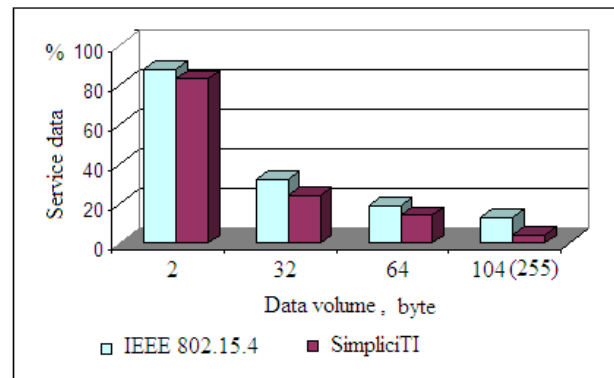
| Preamble Sequence | Start of Frame Delimiter | Frame Length | Frame Control | Sequence Number | Addressing Fields | Data Payload | FCS |
|-------------------|--------------------------|--------------|---------------|-----------------|-------------------|--------------|-----|
| 32                | 8                        | 8            | 16            | 8               | 32                | 104          | 16  |

b)

**Fig. 2 – Schematic view of the data frame and the PHY packet: a) SimpliciTI; b) IEEE 802.15.4**

The analysis of the frame structure of WSN protocol (IEEE 802.15.4, SimpliciTI) – conducted by authors of presented paper – confirmed the decreasing of useful bandwidth when user data size is reduces. The percentage of control information in this protocol makes up from 4% to 80 % – and it depends on data field size (Fig. 3).

Let’s evaluate useful channel capacity in relation to data area size without regard for the interframe spaces and waiting time of medium access.



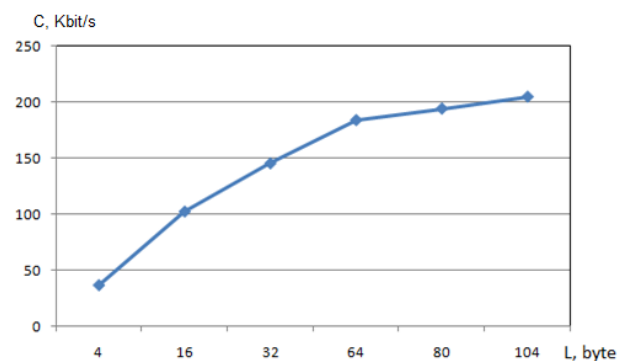
**Fig. 3 – Percentage of service data in the packet for IEEE 802.15.4 and SimpliciTI protocols**

For this we calculate maximum number of minimum length frames. According to IEEE 802.15.4 standard minimum frame length with preamble sequence is  $L_1 = 27$  bytes or 216 bit, so if the transmission rate is equal to 250 Kbit/s it is necessary to 864 microseconds for it’s transmission, one bit transmission time is  $t = 4$  microseconds. The maximum possible network bandwidth is equal to  $m_1 = 1157$  frame/s [5].

According to IEEE 802.15.4 standard maximum frame length with preamble sequence is  $L_2 = 127$  bytes or 1016 bit. The time of frame transmission is 4064 microseconds. So, the maximum possible network bandwidth is equal to  $m_2 = 246$  frame/s. At the same time it is expedient to calculate useful capacity with the use of maximum length frames and minimum length frames.

For the minimum length frames with user’s data size  $d_1 = 4$  bytes useful channel capacity is (Fig.4):

$$C_1 = m_1 \cdot d_1 \cdot t = 37 \text{ Kbit/s, and for the maximum length frames with user’s data size } d_2 = 104 \text{ bytes useful channel capacity is: } C_2 = m_2 \cdot d_2 \cdot t = 204,7 \text{ Kbit/s.}$$



**Fig. 4 – Dependence of useful channel capacity on data field size in frame of IEEE 802.15.4 standard protocol**

As follows from fig. 4 the useful channel capacity reduced to 5 times if we have minimum data area size in protocol structure.

The transmission of the packets with the maximum length of user’s data is more optimal,

because the percent of service data relative to user's data considerably less than in packets with the minimum length. Consequently, useful channel capacity approximates to nominal value – 250 Kbit/s (see Fig.3).

### 3. THE PROPOSED APPROACH

The difference of the proposed method consists in representation of nodes data in residue number system (RNS) (bl. 1) and their transformation in positional number system (bl. 2). In the receiving site (server) inverse transformation of data (bl. 4) in RNS (bl. 5) is occurred (Fig. 5).

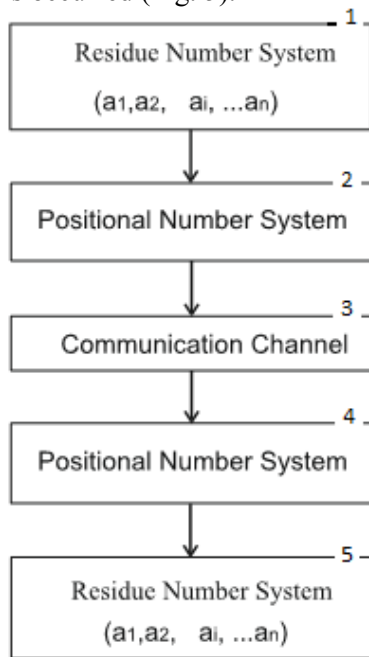


Fig. 5 – Block diagram of proposed coding method

To explain details let's assume that any positive integral number in RNS is represented by the set of smallest positive residues after this number above was divided on the fixed positive integers  $p_1, p_2, \dots, p_i, \dots, p_n$ , which are called as modules. Let's designate the smallest positive residue from a division  $A$  on  $p_i$  as

$$a_i = A - \left[ \frac{A}{p_i} \right] \cdot p_i, \quad (1)$$

where  $[ \bullet ]$  is rounding to smaller integer. Then a number  $A$  in RNS will be written down as:

$$A = \{ \alpha_1, \alpha_2, \dots, \alpha_i, \dots, \alpha_n \}.$$

The coding concept is based on the following steps. The coprime modules are selected per each level according to a condition

$$p_i > a_{imax}, \quad (2)$$

where  $a_{imax}$  is max value of data.

The sensors data are united into packets by the formula [12]:

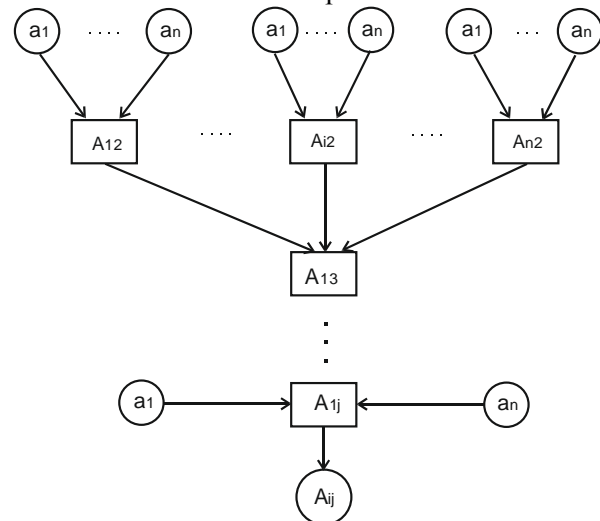
$$A_{ij} = \sum_{i=1}^n a_{ij} \cdot B_{ij} \pmod{P_{ij}}, \quad (3)$$

where  $a_{ij}$  – data of sensor,  $B_{ij}$  – orthogonal or bases,  $P_{ij} = \prod_{i=1}^n p_i$ ,  $p_i$  – coprime numbers,  $n$  – a number of modules,  $i = \overline{1, n}$ ,  $j$  – level number.

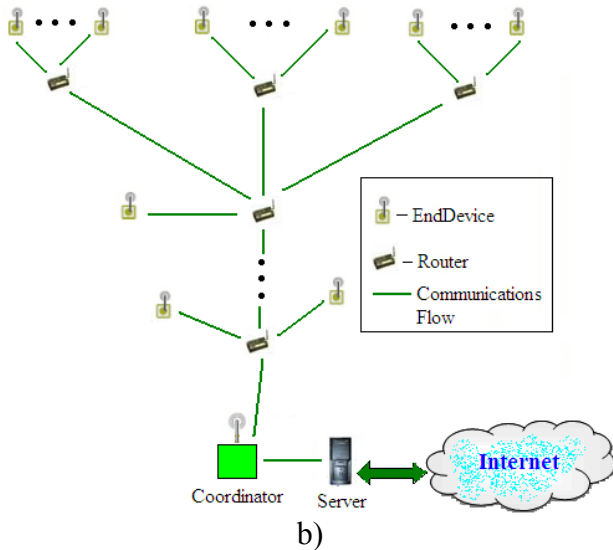
$$B_{ij} = \frac{P_{ij}}{p_i} \cdot \delta_i \equiv 1 \pmod{p_i}, \quad (4)$$

where  $0 < \delta_i < p_i$  – a weight of the orthogonal element.

In the blocks  $A_{12}, A_{i2}, A_{n2}$ , data  $(a_1, \dots, a_n)$  of upper level nodes associates according to the formula (3) (Fig. 6 a). In the block  $A_{13}$  data from the previous level (blocks  $A_{12}, A_{i2}, A_{n2}$ ) are associated and etc. As a result of multilevel coding we get message in weighted (number) system in which measured nodes data are represented.



a)

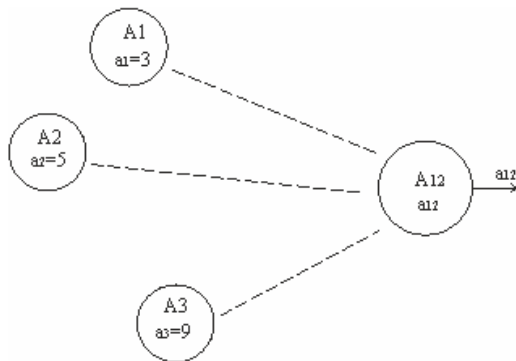


**Fig.6 – Multilevel coding of sensors data: a) block diagram; b) functional diagram.**

Decoding is running via a sequential operation residue of relevant modules:

$$\begin{aligned}
 A_{ij-1} &= A_{ij} \pmod{P_{ij-1}}; \\
 A_{ij-2} &= A_{ij-1} \pmod{P_{ij-2}}; \quad (5) \\
 &\dots \\
 A_{i1} &= A_{i2} \pmod{P_{i1}}.
 \end{aligned}$$

Let's consider the example of the data integration of three nodes  $A_1, A_2, A_3$  (Fig. 7). Let assume that nodes generate the information in the range:  $A_1 \in \{0, 6\}$ ,  $A_2 \in \{0, 10\}$ ,  $A_3 \in \{0, 12\}$ , according to the condition (2) we choose coprime modulus:  $p_1 = 7$ ,  $p_2 = 11$ ,  $p_3 = 13$  and the nodes value  $a_1 = 3$ ,  $a_2 = 5$ ,  $a_3 = 9$ . According to the formula (4) we calculate base numbers:  $B_1 = 715$ ,  $B_2 = 364$ ,  $B_3 = 924$ . The result of calculation for the  $A_{12}$  by formula (3) is equal to:  $a_{12} = 296$ . The value  $a_{12}$  is transmitted to server by communication channel.



**Fig.7 – Data integration on the on basis of RNS**

Such approach allows integrate data with different digit capacity and there is no necessity to append node address, because for each node is

assigned the unique module of Residue Number System, which performs an address functions.

Note growing the frame size increases the network capacity only in a case of network reliable operation (data not distorted neither lost) (see Fig.4). Otherwise increasing the size of the packet may cause bandwidth decreasing since the network retransmits the lost packets.

#### 4. THE ERROR DETECTION

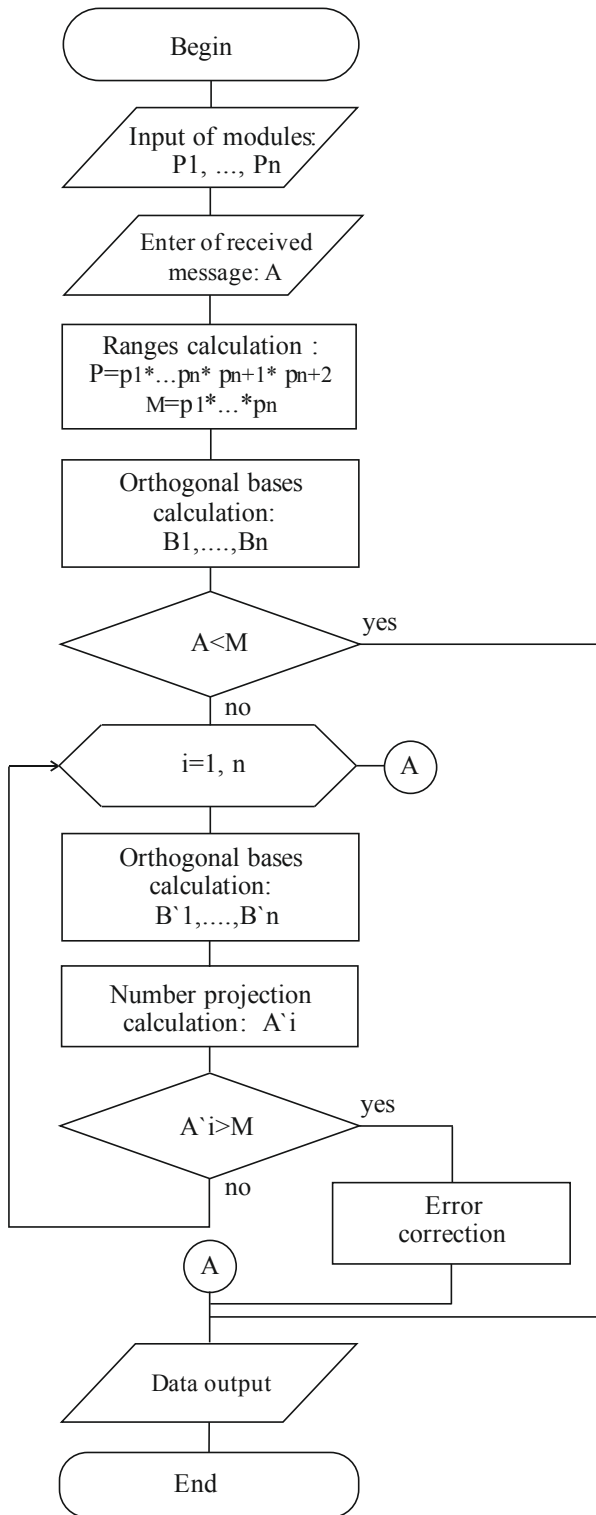
The wireless networks are characterized by high bit error rate (BER) because of electromagnetic obstacles influence in comparison with cable network. For WSN the typical values of BER are equal to  $10^{-4} \div 10^{-6}$  [11].

In the IEEE 802.15.4 standard Cyclic Redundancy Check Codes are used for error detection. The polynome  $G_{16}(x) = x^{16} + x^{12} + x^5 + 1$  provides the detection of one-multiple error with probability 100%, another errors number – with probability  $P = (1 - 2^{-n})$ , where  $n$  – is the number of digits of checksum [10].

In WSN after error detection data packet retransmission is executed. That is results to decrease of useful channel capacity and to increase of energy consumption.

We can increase useful channel capacity of wireless network by implementation of error-correcting codes. The use of error-correcting codes reduces the number of packet retransmissions, which were garbled. In the work the algorithms for the error detection and error correction on the basic of RNS correcting code is elaborated. In this algorithm correcting code of RNS with two check basis (see Fig.7).

To investigate the RNS's codes redundancy let's consider a system with bases  $p_1, p_2, \dots, p_i, \dots, p_n$  and a range  $M = p_1 \cdot p_2 \cdot \dots \cdot p_n$ , and call the range  $M$  as a working range. Moreover let's introduce a basis  $p_{n+1}, p_{n+2}$  which are coprime with any of the accepted basis and let's present numbers in the system with basis  $n+1, n+2$ . It means that we will transfer numbers and make operations at numbers (which are located in a range  $[0, M)$ ) in wider range  $[0, P)$  where  $P = M \cdot p_{n+1} \cdot p_{n+2}$ . Error detection and correction occurs following algorithm (Fig.8). After bases introduction (block 1) and message receiving (block 2) total range  $P = \prod_{i=1}^{n+2} p_i$  and operating range  $M = \prod_{i=1}^n p_i$  are calculated (block 3).



**Fig. 8 – Block diagram of the error detection and correction algorithm in RNS**

The orthonormal bases are calculated by the formula (4). In the block 5 (see Fig.8) a received message  $A$  is compared with the operating range  $M$ : if  $A < M$  so there is no errors in the received message, and calculations finish by the results output. In another case the serial calculation of the orthonormal bases is performed (block 7) and calculation of the number projection on every basis (block 8). Herewith number projection  $A'_i$  is calculated by

strikeout of the  $A a_i$  digit. If the number projection  $A'_i$  is greater than operating range  $M$  (block 9), the next basis is verified (blocks 6 – 9), otherwise there is error on this basis, correction of which is performed in block 10 by the formula [12]

$$a_i = \tilde{a}_i + \left[ \frac{p_i(1+p_{n+1})}{p_{n+1}} - \frac{\tilde{A}}{B_i} \right],$$

where  $\tilde{a}_i$  – is the value of erroneous digit;  $\tilde{A}$  – is received number projection.

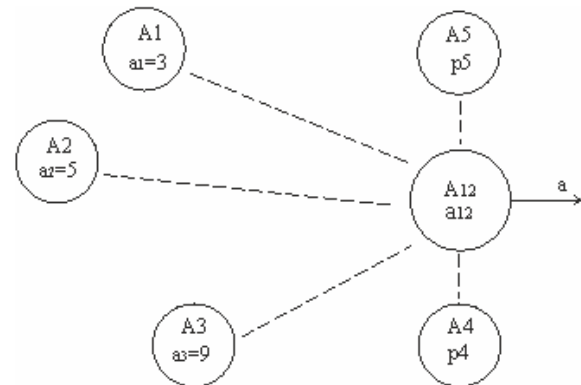
After error correction (block 10) the results of processing come to the block 11.

Let's consider an example of the use of extended modules system. For the data received from the blocks  $A_1, A_2, A_3$  we extend the range of data representation with the use of modules  $p_4, p_5$  in the block  $A_{12}$  (Fig.9) to provide opportunity of the error detection and error correction. In this case we obtain

$$a_4 = a_{12} \pmod{p_4};$$

$$a_5 = a_{12} \pmod{p_5}.$$

Thereafter, the values  $a_1, \dots, a_5$  integrate according to the formula (3).



**Fig.9 – The integration of data with the correcting moduli**

In the receiving side (server) received data are checked for errors and their correction according to algorithm (see Fig.8). For the decoding, decoder must know moduli, which were used in coding process. So, data area of protocol contains data and the numbers of moduli  $N_{M1} \dots N_{Mn}$  (Fig.10).

|          |          |     |          |      |
|----------|----------|-----|----------|------|
| $N_{M1}$ | $N_{M2}$ | ... | $N_{Mn}$ | Data |
|----------|----------|-----|----------|------|

**Fig. 10 – The structure of data area of communications protocol**

The number of modulo is used for the identification of sensor, which transmits data.

As a result of received data repeated division on

coprime moduli RNS we obtain initial data of nodes.

It is known, that numerical system in residual classes gives opportunity of the use of singular noise combating code for correction of errors, which appear in communication channels in time of data transmission and in information system during data processing [11].

RNS codes universality is explained not only by high correcting potential and group error resistance, but adaptive changes of correcting characteristics without change of coding method.

### 5. THE EXPERIMENTAL RESEARCH

Let's consider the useful channel capacity when we use the RNS error-correcting codes with two check modules. For the maximum length frames (104 – 2=102 bytes, check modules occupy 2 bytes) the useful channel capacity is equal to  $C_p = m_2 \cdot (d_2 - 2) = 200,7$  Kbit/s.

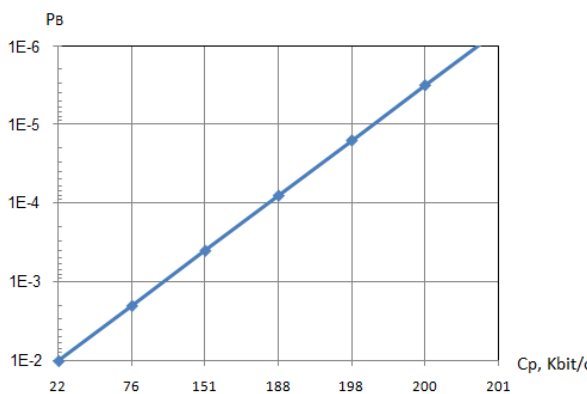
So, the use of the RNS error-correcting codes with two check modules provides error correction by any module and the useful channel capacity is 200 Kbit/s.

Let's calculate useful channel capacity for different error probability means:

$$C_p = \frac{m \cdot L}{1 + P_b \cdot L}$$

where  $m$  – is the number of the frames per second;  $L$  – is the length of data area;  $P_b$  – is the error probability.

In the Fig. 11 the results of calculation of the useful channel capacity with different bit error probabilities are showed.



**Fig. 11 –Dependence of useful channel capacity on errors probability for the maximum data block size (104 bytes)**

Consequently, if the error probability in channel more than  $10^{-5}$  the use of error-correcting codes, particularly RNS error-correcting codes, is expediency (see Fig. 10). If the error probability is

$10^{-3}$  the useful channel capacity reduces in 2 times as shown in Fig. 10.

Thus the proposed approach allows reducing the overhead number as well as increasing the effective capacity of data transmission protocols significantly. Moreover we provide the errors detection and correction implementing data coding by RNS transformation.

Taking into account the results above we plan to explore in future the optimal size of data packet for which the network capacity will reach the maximum value at different levels of a noise.

### 6. CONCLUSIONS

A proposed multilevel nonlinear method of sensors data coding allows to increase the useful capacity of the data transmission protocols for Wireless Sensor Networks in 3 – 5 times by the integration of sensors data with different digits capacity and reducing the transmissions number.

The experimental researches of the useful channel capacity depending on bit error probability with the use of RNS error-correcting codes with two check modules confirmed that the error-correcting codes application is reasonable in a case if the bit error probability exceeds  $10^{-5}$ .

### 7. REFERENCES

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, *Wireless sensor networks: a survey*, *Computer Networks*, (38) (2002), pp. 393-422.
- [2] A. Perrig, R. Szewczyk, V. Wen, D. Culler and J. D. Tygar, SPINS: security protocols for sensor networks, in *International Conference on Mobile Computing and Networking (MobiCom 2001)*, Rome, Italy, 2001.
- [3] A. Zubinsky, Diffused intelligence, *Computer Review*, (8) (2003), pp. 58-61. (in Russian)
- [4] V. G. Olifer, N. A. Olifer, *Computer Networks. Principles, Technologies, Protocols*, SPb.: Piter, 2007. – 960 p. (in Russian).
- [5] A. S. Tanenbaum, *Computer Networks*, 4<sup>th</sup> Edition, 2003, Prentice Hall, Upper Saddle River, NJ 07458.
- [6] H. Qi, Y. Xu, and X. Wang, Mobile-agent-based collaborative signal and information processing in sensor networks, *Proceedings of the IEEE*, (91) 8 (2003), pp. 1172-1183.
- [7] Min Chen, Taekyoung Kwon, Yong Yuan, Yanghee Choi, and Victor C. M. Leung, Mobile agent-based directed diffusion in wireless sensor networks, *EURASIP Journal on Advances in Signal Processing*, (Vol. 2007), Article ID 36871, 13 pp.

doi:10.1155/2007/36871.

- [8] Ye. V. Tolstikova, Minimization of redundancy for transmission for data size transmission in radio sensors network, *Problems of Informatization and Monitoring*, (29) 1 (2010), pp.168-171 (in Russian).
- [9] Yu. V. Merekin, Lower bound of complexity for words concatenation circuit, *Discrete Analysis and Operations Research*, (3) 1 (1996), pp. 52-56. (in Russian).
- [10] ZigBee Alliance, IEEE 802.15.4, ZigBee Standard. <http://www.zigbee.org>.
- [11] A. Samarin, An attractive simplicity SimpliCI. *Electronic Components*, (2) (2008) (in Russian)
- [12] *Modular Parallel Computing Structures of Neuro Processing System*, N. I. Chervyakov, P. A. Sakhnyuk, A. V. Shaposhnikov, S. A. Ryadnov. Edited by N. I. Chervyakov, Moscow, Fizmatlit, 2003. – 288 p. (in Russian).
- [13] William Stallings, *Wireless Communications and Networking*, Upper Saddle River, NJ: Prentice Hall, 2002.

*Wireless Sensor Networks, Network Management, Information Security and Software Engineering.*



**Nataliya Yatskiv** is the associate professor of the Department of Information Computing Systems and Control, **Ternopil National Economic University**. She earned her engineer-physicist degree at Ivano-Frankivsk State Oil and Gas University in 1997 and her PhD Degree in Computer

Systems and Networks at L'viv Polytechnic University in 2003.

Her main Areas of Research Interest are Human-Computer Interaction, Data Processing in Wireless Sensor Networks.



**Anatoly Sachenko** is Professor and Head of the Department of Information Computing Systems and Control and Research advisor of the Research Institute for Intelligent Computer Systems, Ternopil National Economic University. He earned his B.Eng.

Degree in Electrical Engineering at L'viv Polytechnic Institute in 1968 and his PhD Degree in Electrical Engineering at L'viv Physics and Mechanics Institute in 1978 and his Doctor of Technical Sciences Degree in Electrical and Computer Engineering at Leningrad Electrotechnic Institute in 1988. Since 1991 he has been Honored Inventor of Ukraine, since 1993 he has been IEEE Senior Member.

His main Areas of Research Interest are Implementation of Artificial Neural Network, Distributed System and Network, Parallel Computing, Intelligent Controllers for Automated and Robotics Systems. He has published over 430 papers in areas above.



**Vasyl Yatskiv**. Eng. Automation of technological processes. Ivano-Frankivsk Technical University of Oil and Gas, Ukraine (1996). Ph.D. of Computers, system and network. Lviv Polytechnic University, Ukraine (2001).

Associate Professor of Department of Specialized

Computing Systems of Institute of Computer Information Technologies of **Ternopil National Economic University**.

His main Areas of Research Interest are Wireless Sensor Networks, Residue Number System, Free Space Optics



**Su Jun** received the B.S. degree from National Technical University of Ukraine, Ukraine, in 2002, and the M.S. degree from National Technical University of Ukraine, Ukraine, in 2004, all in Computer System and Network.

He is currently pursuing the Ph.D. degree at Ternopil National Economic University,

Ukraine. His main Areas of Research Interests