Quasigroup Redundancy Check Codes For Safety-Critical Systems

Nataša Ilievska

Faculty of Computer Science and Engeneering
Ss. Cyril and Methodius University
Skopje, Republic of Macedonia
e-mail: natasa.ilievska@finki.ukim.mk

Danilo Gligoroski

Department of Telematics
Norwegian University of Science and Technology
Trondheim, Norway
e-mail: danilog@item.ntnu.no

***Abstract*—We define error-detecting codes based on linear quasigroups. We prove that the probability of undetected errors of the defined codes, does not depend on the distribution of the characters in the input message. Next we calculate the probability of undetected errors and identify the best class of linear quasigroups of order 8 for these codes. Also, we explain how the probability of undetected errors can be controlled. At the end, we compare these codes with several CRC codes and conclude that our code has smaller probability of undetected errors than the CRC codes when code rate and block lengths are equal.**

Keywords–error-detecting codes; CRC; linear quasigroups; Safety-Critical Systems.

I. Introduction

A Cyclic Redundancy Check (CRC) is one of the most frequent mechanisms for error detection used in communication networks and storage devices. The idea presented first in 1961 in the work of Peterson and Brown [1] is for every block of data to produce a short check value attached to it. That check value is computed by an algorithm based on cyclic codes.

Very soon after their introduction, CRCs became very popular in communication and computer industry due to their mathematical simplicity and their properties to be implemented easily both in hardware and in software. Many variants of cyclic redundancy check codes have been proposed and standardized such as: CRC-8 [2], CRC-8-WCDMA (Wideband Code Division Multiple Access) [3], CRC-12 [4], CRC-ANSI (American National Standards Institute) [19], CRC-CCITT (Comité Consultatif Internationale Télégraphique et Téléphonique) [20], CRC-32 [6], CRC-64-ISO (International Organization for Standardization) [7], and many others.

Additionally, many other alternatives not based on cyclic codes have been proposed such as: Fletcher-16, Fletcher-32, Fletcher-64 [8][9] and Adler-32 [10].

Beside their typical use in digital networks, CRC codes (or their similar replacements) have been frequently used in so-called Safety-Critical Systems [11][21] that involves process control where toxic, flammable or explosive materials are used, in transportation systems such as railways, avionics and automotive systems and in nuclear power stations.

The motivation and justification of our work in this paper is closely connected with construction of redundancy check codes that will be more suitable in some use-case scenarios for those Safety-Critical Systems. This means that, while in some properties (such as the rate of the code) our codes are not that good as CRC codes, from the perspective of the probability to detect errors, our codes outperform CRCs by one or two orders of magnitude.

The paper is organized as follows. In Section II, we present the mathematical preliminaries to describe our codes. In particular it briefly defines the algebraic structures of quasigroups and linear quasigroups. In Secttion III, we describe our Linear Quasigroup Redundancy Check Codes and in Section IV, we thoroughly analyse the probability of undetected errors with our codes. In Section V, we identify a class of linear quasigroups of order 8 that give the best probabilities for error detections. In Section VI, we compare the error detection probability of our codes with three other CRC codes and we conclude the paper in Section VII.

II. Mathematical Preliminaries

Previous work with error-detecting codes based on quasigroups found that the best results are obtainned with linear quasigroups [12][15].

Definition 1: Quasigroup is algebraic structure $(Q, *)$ such that

$$(\forall u, v \in Q)(\exists! x, y \in Q) \quad (x * u = v \ \& \ u * y = v) \qquad (1)$$

Definition 2: The quasigroup $(Q, *)$ of order 2^q is linear if there are non-singular binary matrices A and B of order $q \times q$ and a binary matrix C of order $1 \times q$, such that

$$(\forall x, y \in Q) \quad x * y = z \Leftrightarrow \boldsymbol{z} = \boldsymbol{x}A + \boldsymbol{y}B + C \qquad (2)$$

where $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are binary representations of x, y and z as vectors of order $1 \times q$ and $+$ is binary addition.

When Q is a quasigroup of order 2^q, then we take that $Q = \{0, 1, ..., 2^q - 1\}$.

Example 1: One linear quasigroup of order 8 is defined with the following non-singular binary matrices

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and the matrix $C = [0\ 0\ 0]$. In order to calculate $5 * 6$, for example, first we turn 5 and 6 into binary form. Thus, $5 = [1\ 0\ 1], 6 = [1\ 1\ 0]$ and substitute them in (2). We calculate

$$[1\ 0\ 1]A + [1\ 1\ 0]B + [0\ 0\ 0] = [1\ 1\ 1]$$

Now, $[1\ \ 1\ \ 1]$ turned into decimal form is 7 so $5 * 6 = 7$. In the same manner we calculate all other products and obtain the following linear quasigroup:

$*$	0	1	2	3	4	5	6	7
0	0	3	2	1	7	4	5	6
1	7	4	5	6	0	3	2	1
2	1	2	3	0	6	5	4	7
3	6	5	4	7	1	2	3	0
4	5	6	7	4	2	1	0	3
5	2	1	0	3	5	6	7	4
6	4	7	6	5	3	0	1	2
7	3	0	1	2	4	7	6	5

Figure 1. Example of a linear quasigroup of order 8.

III. Linear Quasigroup Redundancy Check Codes

Let $(Q,*)$ be linear quasigroup of order 2^q and let $a_0a_1a_2...a_{n-1}$ be an input block of length n. The redundant characters are defined in the following way:

$$d_i = a_i * a_{i+1},\ i \in \{0,1,...,n-1\} \tag{3}$$

where all operations in indexes are per modulo n. This means that $d_0 = a_0 * a_1$, $d_1 = a_1 * a_2$, $\ldots$, $d_{n-2} = a_{n-2} * a_{n-1}$, $d_{n-1} = a_{n-1} * a_0$. Now, the extended message $a_0a_1a_2...a_{n-1}d_0d_1d_2...d_{n-1}$, previously turned into binary form, is transmitted through the binary symmetric channel. A block of length n is extended into a block with length $2n$, from where it follows that the rate of the code is $1/2$.

Under the influence of the noises in the channel, some of the characters may not be correctly transmitted. After receiving the block, the receiver checks if all equations (3) are satisfied. If there is some $i \in \{0,1,...,n-1\}$ for which the equation is not satisfied, the receiver concludes that there is an error in transmission and it asks the sender to send the block once again. But, since the redundant characters are transmitted through the binary symmetric channel, it is possible that they are incorrectly transmitted too, in a way that all equations (3) are satisfied, although some of the information characters $a_0, a_1, ..., a_{n-1}$ are incorrectly transmitted. For this reason, it is possible to have undetected errors in transmission.

IV. The Probability of Undetected Errors

With $P\{i \to j\}$ we will denote the probability that i will be transferred into j through the binary symmetric channel. The following Lemma can be easily shown.

Lemma 1: For all binary vectors $\boldsymbol{a}, \boldsymbol{b}$ i $\boldsymbol{c}$, it is true that

$$P\{\boldsymbol{a}+\boldsymbol{b} \to \boldsymbol{c}+\boldsymbol{b}\} = P\{\boldsymbol{a} \to \boldsymbol{c}\}$$

where $+$ is a binary addition on vectors.

Theorem 1: The probability of undetected errors for the considered code is independent from the distribution of the characters in the input message and form the matrix C.

Proof: Let us consider two disjoint strings of consecutive characters, say $a_ia_{i+1}...a_{i+s}$, $s \geq 0$ and $a_ja_{j+1}...a_{j+r}$, $r \geq 0$, from the input message $a_0a_1 \ldots a_{n-1}$, such that there is at least one character between them, i.e., the two strings do not form a string of consecutive characters. Note that the two strings can have length one (if $s = 0$ or $r = 0$). Since, there is at least one character between the two strings, they act on different redundant characters: The string $a_ia_{i+1}...a_{i+s}$ acts on redundant characters $d_{i-1}, d_i, d_{i+1}, ..., d_{i+s}$ while the string $a_ja_{j+1}...a_{j+r}$ acts on $d_{j-1}, d_j, d_{j+1}, ..., d_{j+r}$ and this two sets of redundant characters are disjoint. For this reason, the random events:

S : the string $a_ia_{i+1}...a_{i+s}$ is incorrectly transmitted and the error is not detected;

R : the string $a_ja_{j+1}...a_{j+r}$ is incorrectly transmitted and the error is not detected;

are independent, from where it follows that $P(QR) = P(Q)P(R)$. For this reason, the probability of undetected errors will be function of the probabilities that r consecutive characters of the input message are incorrectly transmitted and the error is not detected. Therefore, in order to show that the probability of undetected errors is independent from the distribution of the characters in the input message and from the matrix C, it is enough to show that the probability that r consecutive characters of input message are incorrectly transmitted and the error is not detected is independent from the distribution of the characters in the input message and from the matrix C, for arbitrary r.

For this purpose, we introduce the following random events:

A_i : Exactly i consecutive characters from the input message $a_0a_1...a_{n-1}$ are incorrectly transmitted and the error is not detected, $i = 1, 2, ..., n$.

First, let calculate the probability $P(A_1)$, i.e., the probability that exactly one character (let say a_i) is incorrectly transmitted and the error is not detected.

Let H_j be the random event: the true value of a_i is j, $j = 0, 1, 2, ..., 2^q - 1$.

Then, using the formula for total probability, we obtain:

$$P(A_1) = \sum_{j=0}^{2^q-1} P(A_1|H_j)P(H_j) \tag{4}$$

$$\begin{aligned}
&P(A_1|H_j) = \\
&= \sum_{\substack{k=0\\k\neq j}}^{2^q-1} P\{a_i \to k\}P\{d_{i-1} \to a_{i-1} * k\}P\{d_i \to k * a_{i+1}\} \\
&= \sum_{\substack{k=0\\k\neq j}}^{2^q-1} P\{j \to k\}P\{a_{i-1} * j \to a_{i-1} * k\}\cdot \\
&\qquad \cdot P\{j * a_{i+1} \to k * a_{i+1}\} \\
&= \sum_{\substack{k=0\\k\neq j}}^{2^q-1} P\{\boldsymbol{j} \to \boldsymbol{k}\}P\{\boldsymbol{a_{i-1}}A + \boldsymbol{j}B + C \to \boldsymbol{a_{i-1}}A + \boldsymbol{k}B + C\}\cdot \\
&\qquad \cdot P\{\boldsymbol{j}A + \boldsymbol{a_{i+1}}B + C \to \boldsymbol{k}A + \boldsymbol{a_{i+1}}B + C\} \\
&= \sum_{\substack{k=0\\k\neq j}}^{2^q-1} P\{\boldsymbol{j} \to \boldsymbol{k}\}P\{\boldsymbol{j}B \to \boldsymbol{k}B\}P\{\boldsymbol{j}A \to \boldsymbol{k}A\} \\
&= \sum_{\substack{k=0\\k\neq j}}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{k}+\boldsymbol{j}\}P\{\boldsymbol{0} \to (\boldsymbol{k}+\boldsymbol{j})B\}P\{0 \to (\boldsymbol{k}+\boldsymbol{j})A\}
\end{aligned} \tag{5}$$

In the last two equations in (5), Lemma 1 is used. We introduce replacement $\boldsymbol{l} = \boldsymbol{k} + \boldsymbol{j}$ in the last expression of (5). Since j is fixed and k runs through all values from $\{0, 1, ..., 2^q - 1\} \setminus \{j\}$, l will run through all values from $\{0, 1, ..., 2^q - 1\} \setminus \{0\} = \{1, ..., 2^q - 1\}$:

$$P(A_1|H_j) = \sum_{l=1}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{l}\}P\{\boldsymbol{0} \to \boldsymbol{l}B\}P\{\boldsymbol{0} \to \boldsymbol{l}A\}, \qquad \forall j \in \{0, 1, 2, ..., 2^q - 1\} \tag{6}$$

Form (4) and (6) it follows that

$$P(A_1) = \sum_{l=1}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{l}\}P\{\boldsymbol{0} \to \boldsymbol{l}B\}P\{\boldsymbol{0} \to \boldsymbol{l}A\} \qquad (7)$$

Equation (7) means that $P(A_1)$ is independent from the true values of a_{i-1}, a_i and a_{i+1}, i.e., it is independent from the distribution of the characters in the input message. Also, $P(A_1)$ does not depend on the matrix C.

Similarly, we derive a formula for $P(A_2)$ - the probability that exactly two consecutive characters (let say a_i and a_{i+1}) form the input message are incorrectly transmitted and the error is not detected. We introduce the random events H_{jk} : the true value of a_i is j and the true value of a_{i+1} is k, $j, k = 0, 1, 2, ..., 2^q - 1$. Then,

$$P(A_2) = \sum_{j=0}^{2^q-1} \sum_{k=0}^{2^q-1} P(A_2|H_{jk})P(H_{jk}) \qquad (8)$$

$$\begin{aligned}
&P(A_2|H_{jk}) = \\
&= \sum_{\substack{l=0\\ l\neq j}}^{2^q-1} \sum_{\substack{s=0\\ s\neq k}}^{2^q-1} P\{a_i \to l, a_{i+1} \to s\}P\{d_{i-1} \to a_{i-1} * l\}\cdot \\
&\qquad \cdot P\{d_i \to l * s\}P\{d_{i+1} \to s * a_{i+2}\} \\
&= \sum_{\substack{l=0\\ l\neq j}}^{2^q-1} \sum_{\substack{s=0\\ s\neq k}}^{2^q-1} P\{j \to l, k \to s\}P\{a_{i-1} * j \to a_{i-1} * l\}\cdot \\
&\qquad \cdot P\{j * k \to l * s\}P\{k * a_{i+2} \to s * a_{i+2}\} \\
&= \sum_{\substack{l=0\\ l\neq j}}^{2^q-1} \sum_{\substack{s=0\\ s\neq k}}^{2^q-1} P\{\boldsymbol{j} \to \boldsymbol{l}\}P\{\boldsymbol{k} \to \boldsymbol{s}\}\cdot \\
&\qquad \cdot P\{\boldsymbol{a_{i-1}}A + \boldsymbol{j}B + C \to \boldsymbol{a_{i-1}}A + \boldsymbol{l}B + C\}\cdot \\
&\qquad \cdot P\{\boldsymbol{j}A + \boldsymbol{k}B + C \to \boldsymbol{l}A + \boldsymbol{s}B + C\}\cdot \\
&\qquad \cdot P\{\boldsymbol{k}A + \boldsymbol{a_{i+2}}B + C \to \boldsymbol{s}A + \boldsymbol{a_{i+2}}B + C\} \\
&= \sum_{\substack{l=0\\ l\neq j}}^{2^q-1} \sum_{\substack{s=0\\ s\neq k}}^{2^q-1} P\{\boldsymbol{j} \to \boldsymbol{l}\}P\{\boldsymbol{k} \to \boldsymbol{s}\}P\{\boldsymbol{j}B \to \boldsymbol{l}B\}\cdot \\
&\qquad \cdot P\{\boldsymbol{j}A + \boldsymbol{k}B \to \boldsymbol{l}A + \boldsymbol{s}B\}P\{\boldsymbol{k}A \to \boldsymbol{s}A\} = \\
&= \sum_{\substack{l=0\\ l\neq j}}^{2^q-1} \sum_{\substack{s=0\\ s\neq k}}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{l} + \boldsymbol{j}\}P\{\boldsymbol{0} \to \boldsymbol{s} + \boldsymbol{k}\}\cdot \\
&\qquad \cdot P\{\boldsymbol{0} \to (\boldsymbol{l} + \boldsymbol{j})B\}\cdot \\
&\qquad \cdot P\{\boldsymbol{0} \to (\boldsymbol{l} + \boldsymbol{j})A + (\boldsymbol{s} + \boldsymbol{k})B\}\cdot \\
&\qquad \cdot P\{\boldsymbol{0} \to (\boldsymbol{s} + \boldsymbol{k})A\}
\end{aligned} \qquad (9)$$

We introduce replacement: $\boldsymbol{t} = \boldsymbol{l} + \boldsymbol{j}$ and $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{k}$ in the last expression of (9). When l gets all values from $Q \setminus \{j\}$, t will get all values from $Q \setminus \{0\}$. Similarly, when s gets all values from $Q \setminus \{k\}$, r will get all values from $Q \setminus \{0\}$. We obtain:

$$\begin{aligned}
&P(A_2|H_{jk}) = \\
&\sum_{t=1}^{2^q-1} \sum_{r=1}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{t}\}P\{\boldsymbol{0} \to \boldsymbol{r}\}P\{\boldsymbol{0} \to \boldsymbol{t}B\}P\{\boldsymbol{0} \to \boldsymbol{t}A + \boldsymbol{r}B\}\cdot \\
&\qquad \cdot P\{\boldsymbol{0} \to \boldsymbol{r}A\}, \quad \forall j, k \in \{0, 1, ..., 2^q - 1\}
\end{aligned} \qquad (10)$$

Using (8) and (10) we derive that:

$$\begin{aligned}
P(A_2) = \sum_{t=1}^{2^q-1} \sum_{r=1}^{2^q-1} &P\{\boldsymbol{0} \to \boldsymbol{t}\}P\{\boldsymbol{0} \to \boldsymbol{r}\}P\{\boldsymbol{0} \to \boldsymbol{t}B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t}A + \boldsymbol{r}B\}P\{\boldsymbol{0} \to \boldsymbol{r}A\}
\end{aligned} \qquad (11)$$

From (11) we see that $P(A_2)$ is independent from the true values of a_{i-1}, a_i, a_{i+1} and a_{i+2} , i.e., it is independent form the distribution of the characters in the input message. Obviously, $P(A_2)$ does not depend on the matrix C, too.

In general, to derive formula for $P(A_r)$ - the probability that exactly r consecutive characters $a_i, a_{i+1}, ...a_{i+r-1}$ from the input message are incorrectly transmitted and the error is not detected, we introduce random evens
$H_{j_0 j_1 \cdots j_{r-1}}$: the true value of a_i is j_0, the true value of a_{i+1} is j_1, the true value of a_{i+2} is j_2,..., the true value of a_{i+r-1} is j_{r-1}, where $j_0, j_1, ..., j_{r-1} \in \{0, 1, ..., 2^q - 1\}$.

Now,

$$P(A_r) = \sum_{j_0=0}^{2^q-1} \sum_{j_1=0}^{2^q-1} \cdots \sum_{j_{r-1}=0}^{2^q-1} P(A_r|H_{j_0 j_1 \cdots j_{r-1}})P(H_{j_0 j_1 \cdots j_{r-1}}) \qquad (12)$$

$$P(A_r|H_{j_0 j_1 \cdots j_{r-1}}) = \sum_{\substack{s_0=0\\ s_0\neq j_0}}^{2^q-1} \sum_{\substack{s_1=0\\ s_1\neq j_1}}^{2^q-1} \cdots \sum_{\substack{s_{r-1}=0\\ s_{r-1}\neq j_{r-1}}}^{2^q-1} B_{s_0}^{s_{r-1}} \qquad (13)$$

where in a same way as (9) we obtain:

$$\begin{aligned}
B_{s_0}^{s_{r-1}} = &P\{\boldsymbol{0} \to \boldsymbol{s_0} + \boldsymbol{j_0}\}P\{\boldsymbol{0} \to \boldsymbol{s_1} + \boldsymbol{j_1}\}P\{\boldsymbol{0} \to \boldsymbol{s_2} + \boldsymbol{j_2}\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{s_{r-1}} + \boldsymbol{j_{r-1}}\}P\{\boldsymbol{0} \to (\boldsymbol{s_0} + \boldsymbol{j_0})B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to (\boldsymbol{s_0} + \boldsymbol{j_0})A + (\boldsymbol{s_1} + \boldsymbol{j_1})B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to (\boldsymbol{s_1} + \boldsymbol{j_1})A + (\boldsymbol{s_2} + \boldsymbol{j_2})B\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to (\boldsymbol{s_{r-2}} + \boldsymbol{j_{r-2}})A + (\boldsymbol{s_{r-1}} + \boldsymbol{j_{r-1}})B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to (\boldsymbol{s_{r-1}} + \boldsymbol{j_{r-1}})A\}
\end{aligned} \qquad (14)$$

By introducing replacement $\boldsymbol{t_u} = \boldsymbol{s_u} + \boldsymbol{j_u}$, $u = 0, 1, 2, ..., r-1$ in the expression (14) and replacing it in (13), we get:

$$\begin{aligned}
P(A_r|H_{j_0 j_1 \cdots j_{r-1}}) = &\sum_{t_0=1}^{2^q-1} \sum_{t_1=1}^{2^q-1} \cdots \sum_{t_{r-1}=1}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{t_0}\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_1}\}P\{\boldsymbol{0} \to \boldsymbol{t_2}\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-2}}\}P\{\boldsymbol{0} \to \boldsymbol{t_{r-1}}\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_0}B\}P\{\boldsymbol{0} \to \boldsymbol{t_0}A + \boldsymbol{t_1}B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_1}A + \boldsymbol{t_2}B\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-2}}A + \boldsymbol{t_{r-1}}B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-1}}A\}
\end{aligned} \qquad (15)$$

From (12) and (15), we derive

$$\begin{aligned}
P(A_r) = &\sum_{t_0=1}^{2^q-1} \sum_{t_1=1}^{2^q-1} \cdots \sum_{t_{r-1}=1}^{2^q-1} P\{\boldsymbol{0} \to \boldsymbol{t_0}\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_1}\}P\{\boldsymbol{0} \to \boldsymbol{t_2}\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-2}}\}P\{\boldsymbol{0} \to \boldsymbol{t_{r-1}}\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_0}B\}P\{\boldsymbol{0} \to \boldsymbol{t_0}A + \boldsymbol{t_1}B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_1}A + \boldsymbol{t_2}B\} \cdot ...\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-2}}A + \boldsymbol{t_{r-1}}B\}\cdot \\
&\cdot P\{\boldsymbol{0} \to \boldsymbol{t_{r-1}}A\}
\end{aligned} \qquad (16)$$

This means that $P(A_r)$ is independent from the distribution of the characters in the input message and from the matrix C. Thus, the theorem is proven. ■

Using the fact that the probability of undetected errors is independent from the distribution of the characters in the input message the following theorem holds (proved in [16]):

Theorem 2: Let $f(n, p)$ be the probability that at most 4 characters of the input message with length n are incorrectly transmitted through a binary symmetric channel with probability of bit-error p and the error is not detected. If linear quasigroup of order 2^q is used for the code defined with (3) then

$$\begin{array}{rcl} f(2,p) & = & 2v_0v_1 + r_2 \\ f(3,p) & = & 3v_0^3v_1 + 3v_0v_2 + r_3 \\ f(4,p) & = & 4v_0^5v_1 + 4v_0^3v_2 + 2v_0^2v_1^2 + 4v_0v_3 + r_4 \\ f(n,p) & = & nv_1v_0^{2n-3} + nv_2v_0^{2n-5} + \dfrac{n(n-3)}{2}v_1^2v_0^{2n-6} \\ & & +nv_3v_0^{2n-7} + n(n-4)v_2v_1v_0^{2n-8} \\ & & +\dfrac{n(n-4)(n-5)}{6}v_1^3v_0^{2n-9} + nv_4v_0^{2n-9} \\ & & +n(n-5)v_3v_1v_0^{2n-10} + \dfrac{n(n-5)}{2}v_2^2v_0^{2n-10} \\ & & +\dfrac{n(n-5)(n-6)}{2}v_2v_1^2v_0^{2n-11} \\ & & +\dfrac{n(n-5)(n-6)(n-7)}{24}v_1^4v_0^{2n-12}, \quad n \geq 5 \end{array} \tag{17}$$

In the formulas, we use the following notations:
v_t - the probability of undetected errors when exactly t consecutive characters of the initial message $a_0a_1 \ldots a_{n-1}$ are incorrectly transmitted (the characters $a_i, a_{i+1}, \ldots, a_{i+t-1}$ are incorrectly transmitted, but a_{i-1} and a_{i+t} are correctly transmitted), $t = 1, 2, 3, 4$;
v_0 - the probability of correct transmission of a character;
r_t - the probability of undetected errors in a block with length t when all t characters are incorrectly transmitted, $t = 2, 3, 4$.

Although the Theorem 2 in [16] is formulated for fractal quasigroups of order 4, from the proof it can be seen that it holds if for coding is used quasigroup of arbitrary order for which the probability of undetected errors is independent from the distribution of the characters in the input message.

The formula (17) gives us an approximate formula for the probability of undetected errors. Namely, the probability that 5 or more characters of the input message are incorrectly transmitted and the error is not detected is inconsiderably small for small values of a bit-error p. For this reason and the fact that in the real channels the probability of bit-error p is very small, the formula $f(n,p)$ given with (17) gives a good enough approximation of the probability of undetected errors.

The parameters v_t in Theorem 2 are practically $P(A_t)$ from the proof of Theorem 1. The parameters r_2, r_3 and r_4 occur for the following reason. Let say that the two consecutive characters a_0 and a_1 are incorrectly transmitted. The information character a_0 affects the redundant characters d_{n-1} and d_0, while a_1 affects d_0 and d_1. If the block length is greater then or equal to 3, then a_0 and a_1 have one common redundant characters and both of them affect d_0. But if the block length is equal to 2, then the characters a_0 and a_1 have two common redundant characters that are affected: d_0 and d_1. For this reason, the probability that two consecutive characters are incorrectly transmitted and the error is not detected for the blocks with length two is different than the probability for the blocks with length greater than two. Therefore, this case should be considered separately from the general one, and requesting the parameter r_2 to be introduced. A similar situation is for r_3 and r_4. The formulas for these parameters are obtained analogously to the formulas for v_t:

$$\begin{array}{l} r_2 = \displaystyle\sum_{t=1}^{2^q-1}\sum_{r=1}^{2^q-1} P\{\mathbf{0} \to \boldsymbol{t}\}P\{\mathbf{0} \to \boldsymbol{r}\}P\{\mathbf{0} \to \boldsymbol{t}A + \boldsymbol{r}B\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{r}A + \boldsymbol{t}B\} \end{array} \tag{18}$$

$$\begin{array}{l} r_3 = \displaystyle\sum_{t=1}^{2^q-1}\sum_{r=1}^{2^q-1}\sum_{s=1}^{2^q-1} P\{\mathbf{0} \to \boldsymbol{t}\}P\{\mathbf{0} \to \boldsymbol{r}\}P\{\mathbf{0} \to \boldsymbol{s}\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{t}A + \boldsymbol{r}B\}P\{\mathbf{0} \to \boldsymbol{r}A + \boldsymbol{s}B\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{s}A + \boldsymbol{t}B\} \end{array} \tag{19}$$

$$\begin{array}{l} r_4 = \displaystyle\sum_{t=1}^{2^q-1}\sum_{r=1}^{2^q-1}\sum_{s=1}^{2^q-1}\sum_{h=1}^{2^q-1} P\{\mathbf{0} \to \boldsymbol{t}\}P\{\mathbf{0} \to \boldsymbol{r}\}P\{\mathbf{0} \to \boldsymbol{s}\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{h}\}P\{\mathbf{0} \to \boldsymbol{t}A + \boldsymbol{r}B\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{r}A + \boldsymbol{s}B\}P\{\mathbf{0} \to \boldsymbol{s}A + \boldsymbol{h}B\}\cdot \\ \qquad\qquad \cdot P\{\mathbf{0} \to \boldsymbol{h}A + \boldsymbol{t}B\} \end{array} \tag{20}$$

In order to calculate the probability of undetected errors for a given linear quasigroup, one should first calculate the values of the parameters, using (16), (18), (19) and (20), and then to substitute these values into (17).

V. Results With Linear Quasigroups of Order 8

A. The Smallest Probability of Undetected Errors

The values of v_t depend on matrices A and B (see (16)), from where it follows that they depend on the chosen quasigoup for coding. Since the probability of undetected errors depends on v_t, it follows that this probability depends on the chosen quasigroup for coding. It is best the probability of undetected errors to be as small as possible. For this reason, we applied the formula (17) on each pair (A, B) of non-singular binary matrices of order 3 and found that the smallest probability of undetected errors is the following:

$$\begin{array}{rcl} f(2,p) & = & (1-p)^2p^3(4 - 20p + 56p^2 - 96p^3 + 96p^4 \\ & & -55p^5 + 22p^6 - 3p^7) \\ f(3,p) & = & (1-p)^3p^3(3 - 30p + 162p^2 - 580p^3 + 1470p^4 \\ & & -2658p^5 + 3394p^6 - 3024p^7 + 1866p^8 - 787p^9 \\ & & +213p^{10} - 27p^{11} + p^{12}) \\ f(4,p) & = & (1-p)^4p^4(8 - 88p + 404p^2 - 784p^3 - 808p^4 \\ & & +9440p^5 - 29720p^6 + 57432p^7 - 77044p^8 \\ & & +74352p^9 - 51892p^{10} + 25960p^{11} - 9179p^{12} \\ & & +2268p^{13} - 378p^{14} + 44p^{15} - 3p^{16}) \\ f(n,p) & = & \frac{1}{24}np^4(1-p)^{6(n-4)} \times \Big[24 - 384p + 2832p^2 \\ & & -12384p^3 + 12(n+2807)p^4 - 48(2n+1009)p^5 \\ & & +48(6n-499)p^6 + 326976p^7 + 4(n^2 - 603n \\ & & -239092)p^8 + 48(144n + 37283)p^9 - 24(n^2 \\ & & +271n + 102896)p^{10} + 24(5n^2 - 455n \\ & & +110576)p^{11} + (n^3 + 78n^2 + 39863n \\ & & -2273238)p^{12} + 8(n^3 - 96n^2 - 6943n \\ & & +194358)p^{13} + 4(5n^3 - 16n^2 + 13801n \\ & & -213770)p^{14} + 8(n^3 + 99n^2 - 4924n + 46854)p^{15} \\ & & -2(13n^3 + 420n^2 - 10207n + 64386)p^{16} \\ & & -8(n^3 - 86n^2 + 1007n - 4234)p^{17} + 4(5n^3 \\ & & -90n^2 + 583n - 1626)p^{18} - 8(n^3 - 12n^2 + 53n \\ & & -102)p^{19} + (n^3 - 10n^2 + 35n - 50)p^{20}\Big], n \geq 5 \end{array} \tag{21}$$

The graphic of this function, for different values of the block length n is given in Figure 2.

B. Controlling the Error

As we can see from Figure 2, when the block length increases the probability of undetected errors decreases and the sequence of maximums converges to zero. This means that there is some natural number n_0, such that the maximum of $f(n,p)$ will be smaller than ε for all natural numbers n that are greater than or equal to n_0 and the maximum of $f(n,p)$ will be greater than ε for all natural numbers n that are smaller than n_0. So, if we want the probability of undetected errors to be smaller than some previous given value ε, we will choose the block length n to be the smallest natural number such that the maximum of the function $f(n,p)$ is smaller than ε (i.e., $n = n_0$). Since the maximum of $f(n,p)$ will be smaller than ε, follows that $f(n,p)$ will be smaller than ε for all values of $p \in (0, \frac{1}{2})$. We choose n to be the smallest natural number for which the maximum of the function $f(n,p)$ is smaller than ε since we want if there are errors in transmission the smallest possible blocks to be retransmitted.

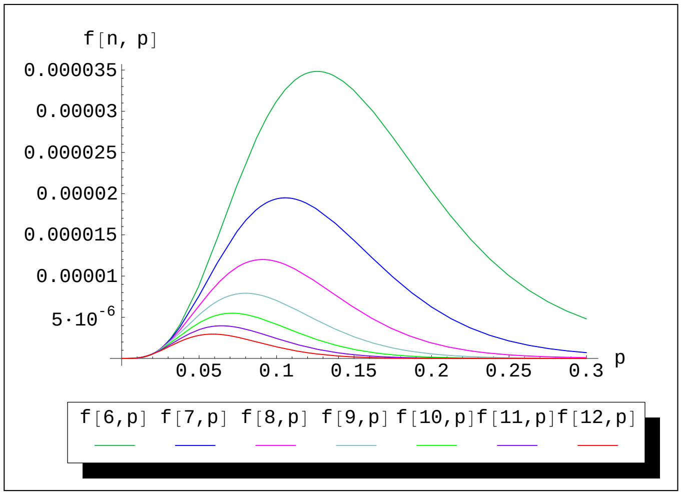

Figure 2. The smallest probability of undetected errors for different values of the block length n if for coding are used quasigroups of order 8.

C. The Best Class of Linear Quasigroups of Order 8

We define the best class of linear quasigroups of order 8 to be the class that contains exactly those linear quasigroups of order 8 that achieve the smallest probability of undetected errors, i.e., the probability given with (21). The quasigroups from this class are best for coding.

We obtained that the best class contains the quasigroups for which the matrix A contains exactly two zeros and the matrix B is determined by the matrix A in the following way. For each non-singular binary matrix A of order 3 with exactly 2 zeros, there are two possible choices for the matrix B: one in which the rows of A are cyclically shifted one position up, and the other in which the rows of A are cyclically shifted two positions up. Namely, if $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is non-singular binary matrix with exactly two zeros, than this matrix A determine two quasigroups of order 8 which have the smallest probability of undetected errors. One of them is obtained when $B = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$ and the other is obtained when $B = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$.

There are 18 binary non-singular matrices with two zeros that can be chosen for the matrix A and for each one of them there are two possible choice for the matrix B. Therefore, there are 36 pairs of non-singular binary matrices (A,B) which give the smallest probability of undetected errors. These quasigroups are given below. Since for every of these 36 pairs of matrices (A,B), any of the eight binary matrices of order 1×3 may be chosen as matrix C, there are 288 linear quasigroups of order 8 in the best class of linear quasigroups of order 8.

The 36 pairs of non-singular binary matrices (A,B) in the best class are given in Figure 3.

	A			B			A			B	
1	0	1	0	1	1	1	0	1	1	1	1
0	1	1	1	1	1	0	1	1	1	0	1
1	1	1	1	0	1	1	1	1	0	1	1
1	0	1	1	1	1	1	0	1	0	1	1
1	1	1	0	1	1	1	1	1	1	0	1
0	1	1	1	0	1	0	1	1	1	1	1
1	1	0	0	1	1	1	1	0	1	1	1
0	1	1	1	1	1	0	1	1	1	1	0
1	1	1	1	1	0	1	1	1	0	1	1
1	1	0	1	1	1	1	1	0	0	1	1
1	1	1	0	1	1	1	1	1	1	1	0
0	1	1	1	1	0	0	1	1	1	1	1
1	1	1	0	1	1	1	1	1	1	0	1
0	1	1	1	0	1	0	1	1	1	1	1
1	0	1	1	1	1	1	0	1	0	1	1
1	1	1	0	1	1	1	1	1	1	1	0
0	1	1	1	1	0	0	1	1	1	1	1
1	1	0	1	1	1	1	1	0	0	1	1
1	1	1	1	0	1	1	1	1	0	1	1
1	0	1	0	1	1	1	0	1	1	1	1
0	1	1	1	1	1	0	1	1	1	0	1
1	1	1	1	1	0	1	1	1	0	1	1
1	1	0	0	1	1	1	1	0	1	1	1
0	1	1	1	1	1	0	1	1	1	1	0
0	1	1	1	0	1	0	1	1	1	1	1
1	0	1	1	1	1	1	0	1	0	1	1
1	1	1	0	1	1	1	1	1	1	0	1
0	1	1	1	1	1	0	1	1	1	0	1
1	1	1	1	0	1	1	1	1	0	1	1
1	0	1	0	1	1	1	0	1	1	1	1
1	1	0	1	0	1	1	1	0	1	1	1
1	0	1	1	1	1	1	0	1	1	1	0
1	1	1	1	1	0	1	1	1	1	0	1
1	1	0	1	1	1	1	1	0	1	0	1
1	1	1	1	0	1	1	1	1	1	1	0
1	0	1	1	1	0	1	0	1	1	1	1
1	1	1	1	0	1	1	1	1	1	1	0
1	0	1	1	1	0	1	0	1	1	1	1
1	1	0	1	1	1	1	1	0	1	0	1
1	1	1	1	1	0	1	1	1	1	0	1
1	1	0	1	0	1	1	1	0	1	1	1
1	0	1	1	1	1	1	0	1	1	1	0
0	1	1	1	1	1	0	1	1	1	1	0
1	1	1	1	1	0	1	1	1	0	1	1
1	1	0	0	1	1	1	1	0	1	1	1
1	0	1	1	1	1	1	0	1	1	1	0
1	1	1	1	1	0	1	1	1	1	0	1
1	1	0	1	0	1	1	1	0	1	1	1
0	1	1	1	1	0	0	1	1	1	1	1
1	1	0	1	1	1	1	1	0	0	1	1
1	1	1	0	1	1	1	1	1	1	1	0
1	0	1	1	1	0	1	0	1	1	1	1
1	1	0	1	1	1	1	1	0	1	0	1
1	1	1	1	0	1	1	1	1	1	1	0

Figure 3. The best class of linear quasigroups of order 8.

VI. COMPARISON WITH THE CRC CODES

We give a comparison of our codes with some CRC codes. The comparison is made from the aspect of the probability of undetected errors. Ability of the CRC code to detect errors depends on the chosen polynomial for coding. We will consider several cases of polynomials, accepted as a standard for coding. Namely, we will consider CRC-12 defined with the polynomial $g(x) = x^{12}+x^{11}+x^3+x^2+x+1$ [4], CRC-ANSI defined with $g(x) = x^{16}+x^{15}+x^2+x+1$ [19] and CRC-CCITT defined with $g(x) = x^{16}+x^{12}+x^5+1$ [20]. Unlike our code, CRC code adds constant number of redundant characters, regardless of the length of the information block. CRC-12 adds 12 bits, while CRC-ANSI and CRC-CCITT add 16 bits. In Table I the maximums of the probabilities of undetected errors for our code and CRC-12 code are given, while in Table II the maximums of the probabilities of undetected errors for our code, CRC-ANSI and CRC-CCITT codes for different values of the block length n are given. The values for CRC codes are taken from [18]. The block length is expressed in bits. We can see that our code has smaller probability of undetected errors than CRC-12 for all values of block length n which are greater than or equal to 12. For n greater than or equal to 15, our code has smaller probability of undetected errors than all considered

TABLE I. THE MAXIMUMS OF THE PROBABILITY OF UNDETECTED ERRORS FOR OUR CODE AND CRC-12. THE BLOCK LENGTH n IS EXPRESSED IN BITS

n	Our code	CRC-12
6	1.53809×10^{-2}	4.98239×10^{-4}
9	1.94931×10^{-3}	4.44429×10^{-4}
12	2.72046×10^{-4}	4.92904×10^{-4}
15	7.22452×10^{-5}	5.32493×10^{-4}
18	3.48346×10^{-5}	5.24102×10^{-4}
21	1.94926×10^{-5}	5.01021×10^{-4}
24	1.20057×10^{-5}	4.95408×10^{-4}
27	7.91486×10^{-6}	4.68967×10^{-4}
30	5.49112×10^{-6}	4.36575×10^{-4}
33	3.96425×10^{-6}	4.13106×10^{-4}
36	2.95482×10^{-6}	3.98553×10^{-4}
39	2.26092×10^{-6}	3.84785×10^{-4}
42	1.76831×10^{-6}	3.68207×10^{-4}
45	1.40897×10^{-6}	3.53279×10^{-4}
48	1.14074×10^{-6}	3.40974×10^{-4}

TABLE II. THE MAXIMUMS OF THE PROBABILITY OF UNDETECTED ERRORS FOR OUR CODE, CRC-ANSI AND CRC-CCITT. THE BLOCK LENGTH n IS EXPRESSED IN BITS

n	Our code	CRC-ANSI	CRC-CCITT
6	1.53809×10^{-2}	2.09564×10^{-4}	1.82571×10^{-4}
9	1.94931×10^{-3}	1.83062×10^{-4}	1.59587×10^{-4}
12	2.72046×10^{-4}	1.49497×10^{-4}	1.31108×10^{-4}
15	7.22452×10^{-5}	1.49435×10^{-4}	1.07281×10^{-4}
18	3.48346×10^{-5}	1.87672×10^{-4}	9.68045×10^{-5}
21	1.94926×10^{-5}	1.96955×10^{-4}	8.80828×10^{-5}
24	1.20057×10^{-5}	1.88110×10^{-4}	7.82445×10^{-5}
27	7.91486×10^{-6}	1.72350×10^{-4}	6.89410×10^{-5}
30	5.49112×10^{-6}	1.66609×10^{-4}	6.05393×10^{-5}
33	3.96425×10^{-6}	1.67740×10^{-4}	5.32930×10^{-5}
36	2.95482×10^{-6}	1.61975×10^{-4}	4.71277×10^{-5}
39	2.26092×10^{-6}	1.52149×10^{-4}	4.18904×10^{-5}
42	1.76831×10^{-6}	1.40941×10^{-4}	3.74422×10^{-5}
45	1.40897×10^{-6}	1.34158×10^{-4}	3.41088×10^{-5}
48	1.14074×10^{-6}	1.31914×10^{-4}	3.17809×10^{-5}

CRC codes (even in the case when the CRC checksum is for short lengths such that the CRC code has also a rate of 1/2).

Additionally, which is important for Safety-Crytical Systems, we can make the probability of undetected errors arbitrary small, which is not case with CRC codes. Namely, the probability of undetected errors for CRC code with c redundant bits tends to 2^{-c} when the block length n tends to infinity.

VII. CONCLUSION

We defined error-detecting codes based on linear quasigroups. We proved that the probability of undetected errors is independent from the distribution of the characters in the input message. Using this property, we found the best class of linear quasigroups of order 8 for such coding and we computed the corresponding probability of undetected errors. Finally, we explained how the probability of undetected errors can be made arbitrary small. We compare our codes with CRC-12, CRC-ANSI and CRC-CCITT and show that our code has smaller probability of undetected errors than the CRC codes when code rate and block lengths are equal.

ACKNOWLEDGMENT

This work was partially financed by the Faculty of Computer Science and Engineering at the "Ss.Cyril and Methodius" University.

REFERENCES

[1] W. W. Peterson and D. T. Brown, "Cyclic Codes for Error Detection," in Proceedings of the IRE, vol. 49, no. 1, 1961, pp. 228-235.

[2] P. Koopman and T. Chakravarty, "Cyclic Redundancy Code (CRC) Polynomial Selection For Embedded Networks," International Conference on Dependable Systems and Networks, 2004, pp. 145-154.

[3] A. Richardson, WCDMA Handbook, Cambridge, UK, Cambridge University Press, 2005, pp. 223.

[4] A. Perez, "Byte-Wise CRC Calculations," IEEE Micro vol. 3, no. 3, 1983, pp. 40-50.

[5] S. Blanc, A. Bonastre, and P. J. Gil, "Dependability assessment of by-wire control systems using fault injection," Journal of Systems Architecture, vol. 55, no. 2, 2009, pp. 102-113.

[6] P. Koopman, "32-Bit Cyclic Redundancy Codes for Internet Applications," The International Conference on Dependable Systems and Networks, 2002, pp. 459-468.

[7] D. T. Jones, "An improved 64-bit cyclic redundancy check for protein sequences," University College London, 2009.

[8] J. G. Fletcher, "An Arithmetic Checksum for Serial Transmissions," IEEE Transactions on Communications, vol 30, no. 1, 1982, pp. 247-252.

[9] J. Zweig and C. Partridge, "TCP Alternate Checksum Options," IETF RFC 1146, Mar. 1990.

[10] P. Deutsch and J.-L. Gailly, "ZLIB Compressed Data Format Specification Version 3.3," IETF RFC 1950, May 1996.

[11] G. Latif-Shabgahi, J. M. Bass, and S. Bennett, "A taxonomy for software voting algorithms used in safety-critical systems," IEEE Transactions on Reliability, vol. 53, no. 3, 2004, pp. 319-328.

[12] V. Bakeva and N. Ilievska, "A probabilistic model of error-detecting codes based on quasigroups," Quasigroups and Related Systems, vol. 17, no. 2, 2009, pp. 135-148.

[13] N. Ilievska and D. Gligoroski, "Error-Detecting Code using Linear Quasigroups," Advances in Intelligent Systems and Computing vol. 311, ICT Innovations 2014, Springer, 2014, pp. 309-318.

[14] N. Ilievska and D. Gligoroski, "An Error-Detecting Code based on Linear Quasigroups," in Proceedings of 11th International Conference for Informatics and Information Technology (CIIT 2014), Bitola, Republic of Macedonia, 2014, in press.

[15] N. Ilievska, "Proving the probability of undetected errors for an error-detecting code based on quasigroups," Quasigroups and Related Systems vol. 22, no. 2, 2014, pp. 223-246.

[16] N. Ilievska and V. Bakeva, "A Model of error-detecting codes based on quasigroups of order 4," in Proceedings of 6th International Conference for Informatics and Information Technology, Bitola, Republic of Macedonia, 2008, pp. 7-11.

[17] Y. Chen, M. Niemenmma, A.J. Han Vinck, and D. Gligoroski, "On the Error Detection Capability of One Check Digit," IEEE Transactions on Information theory, 2014, pp. 261-270.

[18] K. A. Witzke, "Exmanination of the undetected error probability of linear block codes," Thesis: M.A. Sc, University of British Columbia Department of Electrical Engineering, 1984.

[19] T.V. Ramabadran and S.S. Gaitonde, "A tutorial on CRC computations," Micro, IEEE, vol.8, no.4, Aug. 1988, pp. 62-75.

[20] P. Koopman and T. Chakravarty, "Cyclic Redundancy Code (CRC) Polynomial Selection For Embedded Networks," in Proceedings of the International Conference on Dependable Systems and Networks, 2004, pp. 145-154.

[21] J.C. Knight, "Safety Critical Systems: Challenges and Directions," in Proceedings of the 24th International Conference on Software Engineering, 2002, pp. 547-550.