On the Error Detection Capability of One Check Digit

Yanling Chen, Markku Niemenmaa, A. J. Han Vinck, and Danilo Gligoroski

***Abstract*— In this paper, we study a check digit system which is based on the use of elementary abelian p-groups of order p^k. This paper is inspired by a recently introduced check digit system for hexadecimal numbers. By interpreting its check equation in terminology of matrix algebra, we generalize the idea to build systems over a group of order p^k, while keeping the ability to detect all the: 1) single errors; 2) adjacent transpositions; 3) twin errors; 4) jump transpositions; and 5) jump twin errors. Besides, we consider two categories of jump errors: 1) t-jump transpositions and 2) t-jump twin errors, which include and further extend the double error types of 2)–5). In particular, we explore R_c, the maximum detection radius of the system on detecting these two kinds of generalized jump errors, and show that it is 2^k-2 for $p=2$ and $(p^k-1)/2-1$ for an odd prime p. Also, we show how to build such a system that detects all the single errors and these two kinds of double jump-errors within R_c.**

***Index Terms*— Check digit system, elementary abelian group, error detection, matrix algebra.**

I. Introduction

ACCORDING to the empirical investigations by D. F. Beckley [2] and J. Verhoeff [3], when transmitting a sequence of digits, the most common transmission errors made by human operators are the following:

1) single error: $\cdots a \cdots \to \cdots b \cdots$;
2) adjacent transposition: $\cdots ab \cdots \to \cdots ba \cdots$;
3) twin error: $\cdots aa \cdots \to \cdots bb \cdots$;
4) jump transposition: $\cdots abc \cdots \to \cdots cba \cdots$;
5) jump twin error: $\cdots aca \cdots \to \cdots bcb \cdots$.

Among them, single errors and adjacent transpositions, are the most prevalent ones. The recognition of these errors is usually done by appending a check digit a_{n+1} to a given sequence $a_1 \cdots a_n$ of information digits.

Manuscript received December 12, 2012; revised July 22, 2013; accepted July 30, 2013. Date of publication October 30, 2013; date of current version December 20, 2013. This paper was presented at the 2012 IEEE International Symposium on Information Theory.

Y. Chen was with the Department of Telematics, Norwegian University of Science and Technology, Trondheim N-7049, Norway. She is now with the Department of Electrical Engineering and Information Technology, Ruhr University Bochum, Bochum D-44780, Germany (e-mail: yanling.chen-q5g@rub.de).

M. Niemenmaa is with the Department of Mathematical Sciences, University of Oulu, Oulu 90014, Finland (e-mail: markku.niemenmaa@oulu.fi).

A. J. Han Vinck is with the Institute for Experimental Mathematics, University of Duisburg-Essen, Essen D-45326, Germany (e-mail: vinck@iem.uni-due.de).

D. Gligoroski is with the Department of Telematics, Norwegian University of Science and Technology, Trondheim N-7049, Norway (e-mail: danilog@item.ntnu.no).

Communicated by A. Ashikhmin, Associate Editor for Coding Techniques.
Digital Object Identifier 10.1109/TIT.2013.2287698

Some well-known examples of the check digit systems used in practice are the European Article Number (EAN) Code, the Universal Product Code (UPC), the International Standard Book Number (ISBN) Code and the system of the serial numbers of the former (i.e. pre-euro) German banknotes.

A. Previous Studies

Since single errors and adjacent transpositions are the most prevalent ones, research attention was first brought to design systems over groups with anti-symmetric mappings which ensure these two kinds of errors to be detected. One can refer to a long list of research articles such as [4]–[9] and a survey of anti-symmetric mappings in different groups in [10].

In addition, the possibility of constructing error detecting codes based on quasigroups was discussed in [6]. Necessary and sufficient conditions were established in order to detect adjacent transpositions and jump transpositions (but only in the information digits). In [11], the control digit involving in both errors was taken into account. Research on check digit system using quasigroups was continued in [12], and a comprehensive investigation was conducted in [12], [13], where necessary and sufficient conditions were established in order to detect each of the five error types. So far, the approaches taken, are in general analytical.

Recently, M. Niemenmaa in [1] proposed a check digit system for hexadecimal numbers, based on a suitable automorphism of the elementary abelian group of order 16. Its design is concise and elegant, with the capability of detecting all the five types of errors as listed above. (Note that it is not the first attempt to design a system over groups detecting all the five error types. For instance, one can refer to [14] for an earlier study which shows that the Sylow 2-subgroups of nearly all Chevalley groups in even characteristic allow the definition of a check digit system with such an error detection capability.) Two use cases of hexadecimal numbers in real life applications are worth mentioning. They are the International Standard Audiovisual Number (ISAN) which enables the identification of any kind of audiovisual works and the International Mobile Equipment Identifier (MEID) which is unique for each mobile station.

B. Our Approach

In this paper, we first briefly review the check digit system proposed in [1]. By interpreting its check equation in the terminology of matrix algebra [19], we generalize its idea for hexadecimal numbers to a group of order p^k, for any $k \geq 1$ and prime p. To serve the purpose of detecting all the five types

of errors, we explore the properties of the desirable matrices which are suitable for the system design, the number of them and the asymptotic probability of finding them as well.

Moreover, we look into the following two categories of jump errors, which include and further extend the error types 2)-5).

6) t-jump transposition:

$$\cdots ab_1 \cdots b_t c \cdots \to \cdots cb_1 \cdots b_t a \cdots$$

7) t-jump twin error:

$$\cdots ab_1 \cdots b_t a \cdots \to \cdots cb_1 \cdots b_t c \cdots$$

Let $t \geq 0$, and as $t = 0$ we have $b_1 \cdots b_t \in \emptyset$. It is easy to see that, the error types 2) and 4): adjacent transposition and jump transposition, can be regarded as t-jump transpositions for $t = 0$ and $t = 1$, respectively; 3) and 5): twin error and jump twin error, can be regarded as t-jump twin errors for $t = 0$ and $t = 1$, respectively.

These two kinds of errors were first considered in [15] and treated as transposition and twin errors on places $(i, i+t+1)$, where $1 \leq i \leq n$ and $i + t + 1 \leq n$. They are of our interest, not only because they simplify the list of the error types, but also because they may occur more frequently than expected, especially when people input data while using a new keyboard with an unexpected layout, or when they forget to switch the input language to the right one they intend to use.

For a given check digit system, we denote t^* to be the largest jump length such that for any $t \leq t^*$, all the t-jump transpositions and t-jump twin errors will be detected. Intuitively we have the *detection radius*: $R = t^* + 1$, reflecting the capability of the system to detect these two kinds of generalized jump errors. Note that a system capable of detecting error types 1)-5) has an detection radius $R \geq 2$ by definition. Furthermore, we denote t_c to be the maximum t^* that could be achieved, and accordingly $R_c = t_c + 1$ to be the *longest detection radius*. Then a check digit system capable of detecting all the single errors and double errors of types 6) and 7) within R_c, is of great interest due to its desirable performance on the error detection capability.

Consider our proposed systems over a group of order p^k. One of the main contributions of this paper is to show that

$$R_c = \begin{cases} 2^k - 2 & p = 2; \\ \frac{p^k - 1}{2} - 1 & p \text{ odd prime.} \end{cases} \tag{1}$$

Furthermore, we provide easy construction of the systems which detect all the single errors and achieves the maximum detection radius R_c. Our approach is simple, constructive and universal for any prime p and $k \geq 1$. Moreover, our design is not bound by the size of the information digits n. In particular, it outperforms several well known and widely used check digit system as we will demonstrate in Section V. The trick is to simply employ a matrix (the only parameter to be determined in the check equation) of a special kind. Any of those with primitive polynomials (of degree k in the prime field $\mathbf{F}_p$) as their characteristic polynomials will serve as a good candidate for the purpose of achieving the longest possible error detection radius.

The rest of the paper is organized as follows: First in Section II, we briefly review the recently introduced check digit system for hexadecimal numbers. Its generalization to a group of order p^k is presented in Section III. In Section IV, we discuss on the maximum error detection radius of the proposed systems. In Section V, we give some concrete examples. Some necessary proofs are present in Section VI, VII and Section VIII. Finally we conclude in Section IX.

II. Check Digit System for Hexadecimal Numbers

In this section, we briefly review the check digit system for hexadecimal numbers proposed in [1].

A. Group

In [1], the hexadecimal numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E and F are represented as elements of the abelian group $\mathbf{G} = \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_2$ by denoting $0 = (0,0,0,0)$, $1 = (0,0,0,1), \ldots, 9 = (1,0,0,1)$, $\mathrm{A} = (1,0,1,0)$, $\mathrm{B} = (1,0,1,1)$, $\mathrm{C} = (1,1,0,0)$, $\mathrm{D} = (1,1,0,1)$, $\mathrm{E} = (1,1,1,0)$, and $\mathrm{F} = (1,1,1,1)$. It is easy to see that in $\mathbf{G}$, 0 is the identity element and the inverse of each element is the element itself.

B. Check Equation

Suppose that the information digits $a_1, \ldots, a_n$ and the check digit a_{n+1} are interpreted as elements of the group $(\mathbf{G}, +)$. Then a_{n+1} can be determined by the check equation

$$\mathrm{P}(a_1) + \mathrm{P}^2(a_2) + \cdots + \mathrm{P}^n(a_n) + \mathrm{P}^{n+1}(a_{n+1}) = 0, \tag{2}$$

where P is a permutation of the set $\mathbf{G}$.

C. Check Digit System

In [1], P is defined to be a mapping: $\mathbf{G} \to \mathbf{G}$ by

$$\mathrm{P}((a,b,c,d)) = (a+b, c, d, a). \tag{3}$$

By using this specific P, the system can detect all the five types of errors listed in the introduction. As noted, as a permutation, P is 15-cycle and fixes only the element 0. So is P^2 as well.

D. Remark

We slightly abuse the notion P by letting

$$\mathrm{P} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4}$$

Then equation (3) has the following matrix interpretation:

$$\mathrm{P}((a,b,c,d)) = (a,b,c,d) \cdot \mathrm{P}. \tag{5}$$

As a direct result, the check equation (2) can be rewritten as

$$a_1 \cdot \mathrm{P} + a_2 \cdot \mathrm{P}^2 + \cdots + a_n \cdot \mathrm{P}^n + a_{n+1} \cdot \mathrm{P}^{n+1} = 0. \tag{6}$$

III. Check Digit System Over a Group of Order p^k

In this section, we provide a general design of check digit systems which are based on the use of elementary abelian p-groups of order p^k. The proposed systems keep the ability to detect all the five error types 1)-5).

Consider a check digit system over a set of p^k numbers: $0, 1, 2, \ldots, p^k - 1$, where p is a prime and $k > 0$. Similarly to the approach described in Section II-A, we represent these p^k numbers as elements of the abelian group

$$\mathbf{G} = \underbrace{\mathbf{Z}_p \oplus \mathbf{Z}_p \oplus \cdots \oplus \mathbf{Z}_p}_{k} = (\mathbf{Z}_p)^k,$$

where $\mathbf{Z}_p = \mathbf{F}_p$. Therefore, each number corresponds to a k-tuple in $\mathbf{G}$. For simplicity, we can take the k-tuple as the base p representation of the number.

Denote the information digits to be $a_1, \ldots, a_n$ and the check digit to be a_{n+1}, all as elements of the group $\mathbf{G}$. We apply the check equation (6):

$$a_1 \cdot \mathrm{P} + a_2 \cdot \mathrm{P}^2 + \cdots + a_n \cdot \mathrm{P}^n + a_{n+1} \cdot \mathrm{P}^{n+1} = 0,$$

where P is a $k \times k$ matrix over $\mathbf{F}_p$. It is easy to see that

- if P is nonsingular, then all single errors will be detected since

$$1)' \quad a \cdot \mathrm{P} \neq b \cdot \mathrm{P} \quad \text{for all } a \neq b \in \mathbf{G}.$$

- in order to detect all the adjacent transpositions, P has to be such that

$$2)' \quad a + b \cdot \mathrm{P} \neq b + a \cdot \mathrm{P} \quad \text{for all } a \neq b \in \mathbf{G}.$$

  That is, $\mathrm{P} - \mathrm{I}$ must be nonsingular as well. Here I is the $k \times k$ identity matrix.
- in order to detect all the twin errors, P has to be such that

$$3)' \quad a + a \cdot \mathrm{P} \neq b + b \cdot \mathrm{P} \quad \text{for all } a \neq b \in \mathbf{G}.$$

  That is, $\mathrm{P} + \mathrm{I}$ must be nonsingular.
- for the jump transpositions, the following must hold

$$4)' \quad a + b \cdot \mathrm{P} + c \cdot \mathrm{P}^2 \neq c + b \cdot \mathrm{P} + a \cdot \mathrm{P}^2 \quad \text{for all } a \neq c \in \mathbf{G}.$$

  It requires $\mathrm{P}^2 - \mathrm{I}$ to be nonsingular.
- for the jump twin errors, we have the condition

$$5)' \quad a + c \cdot \mathrm{P} + a \cdot \mathrm{P}^2 \neq b + c \cdot \mathrm{P} + b \cdot \mathrm{P}^2 \quad \text{for all } a \neq b \in \mathbf{G}.$$

  It requires $\mathrm{P}^2 + \mathrm{I}$ to be nonsingular.

As a conclusion, we have the following theorem:

Theorem 3.1: A check digit system for p^k numbers, with the ability to detect all the errors of types 1)-5), can be designed by using the check equation (6) and choosing a matrix P over $\mathbf{F}_p$ such that P, $\mathrm{P} - \mathrm{I}$, $\mathrm{P} + \mathrm{I}$ and $\mathrm{P}^2 + \mathrm{I}$ are all nonsingular.

A. $p = 2$

In the case of $p = 2$, over $\mathbf{F}_2$, $\mathrm{P} + \mathrm{I}$ is the same as $\mathrm{P} - \mathrm{I}$. Furthermore, we note that as a binary matrix, $\mathrm{P}^2 + \mathrm{I} = (\mathrm{P} + \mathrm{I})^2$. Straightforwardly, we can simplify Theorem 3.1 to be the following theorem.

Theorem 3.2: A check digit system for 2^k numbers, with the ability to detect all the errors of types 1)-5), can be designed by using the check equation (6) and choosing a binary $k \times k$ matrix P such that both P and P+I are nonsingular.

According to Theorem 3.2, binary matrices with no eigenvalues of 0 or 1 are of our interest. As proved in [16], the number of such $k \times k$ matrices P is $2^{k(k-1)/2} N_k$, where

$$N_0 = 1, \quad N_k = (2^k - 1) N_{k-1} + (-1)^k. \tag{7}$$

The sequence begins $0, 2, 48, 5824, 2887680, \ldots$, for $k = 1, 2, 3, 4, \ldots$, respectively. As an example, for $k = 4$, P as defined in (4) is one of those 5824 suitable ones.

As k increases, the asymptotic probability of such a matrix is approaching 0.0833986 as shown in [16].

B. p as an Odd Prime

According to Theorem 3.1, in order to design a check digit system with the ability to detect all the errors of types 1)-5) by the check equation (6), we need to use a matrix P over $\mathbf{F}_p$ such that P, $\mathrm{P} - \mathrm{I}$, $\mathrm{P} + \mathrm{I}$ and $\mathrm{P}^2 + \mathrm{I}$ are all nonsingular.

a) a prime p: $p \equiv 1 \mod 4$: It is known that for a prime p, if $p \equiv 1 \mod 4$, then -1 is a quadratic residue modulo p. In this case, there is an $\alpha \in \mathbf{F}_p$ such that $\alpha^2 \equiv -1 \mod p$. Straightforwardly, $\mathrm{P}^2 + \mathrm{I}$ can be further factored into $(\mathrm{P} + \alpha\mathrm{I}) \cdot (\mathrm{P} - \alpha\mathrm{I})$. We note that if $\mathrm{P}^2 + \mathrm{I}$ is nonsingular, it implies that both $\mathrm{P} + \alpha\mathrm{I}$ and $\mathrm{P} - \alpha\mathrm{I}$ are nonsingular as well.

According to Theorem 3.1, the matrices over $\mathbf{F}_p$ with no eigenvalues of $0, 1, -1, \alpha, -\alpha$ are of our interest. Let γ_k be the number of invertible $k \times k$ matrices over $\mathbf{F}_p$, and let ρ_k be the number of $k \times k$ matrices over $\mathbf{F}_p$ with no eigenvalues of $0, 1, -1, \alpha, -\alpha$. We have the following lemma.

Lemma 3.3:

$$1 + \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k = \frac{1}{1-u} \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^4, \qquad |u| < 1. \tag{8}$$

Proof: The proof is given in Section VI. ■

As k increases, the asymptotic probability of such a matrix is $\prod_{r \geq 1} (1 - 1/p^r)^5$ by the following theorem.

Theorem 3.4:

$$\lim_{k \to \infty} \frac{\rho_k}{p^{k^2}} = \prod_{r \geq 1} (1 - \frac{1}{p^r})^5. \tag{9}$$

Proof: The proof is given in Section VII. ■

b) a prime p: $p \equiv 3 \mod 4$: If a prime p have $p \equiv 3 \mod 4$, then -1 is not a quadratic residue modulo p. In this case, $z^2 + 1$ is an irreducible polynomial over $\mathbf{F}_p$. As a direct result, $\mathrm{P}^2 + \mathrm{I}$ can not be further factored.

Consider the $k \times k$ matrices P over $\mathbf{F}_p$ which are suitable for the digit check system defined in Theorem 3.1, i.e., those P such that P, $\mathrm{P} - \mathrm{I}$, $\mathrm{P} + \mathrm{I}$, $\mathrm{P}^2 + \mathrm{I}$ are all nonsingular. We have the following Lemma:

Lemma 3.5: The $k\times k$ matrices P over $\mathbf{F}_p$ which are suitable for the digit check system defined in Theorem 3.1, are those without factors of $z, z-1, z+1$ or z^2+1 in their characteristic polynomial.

Proof: Let $f(z)$ be the characteristic polynomial of P, i.e., $f(z)=\det(z\mathrm{I}-\mathrm{P})$; $m(z)$ be the minimal polynomial of matrix P. By definition, $m(z)$ is a monic polynomial over $\mathbf{F}_p$ of least degree such that $m(\mathrm{P})=0$. Let $r(z)$ be any of the irreducible polynomials $z, z-1, z+1$ or z^2+1. Assume that $r(z)\mid f(z)$.

According to [20, Theorem 9.16], the characteristic polynomial and the minimal polynomial of a matrix have the same irreducible factors. Since $r(z)$ is irreducible and $r(z)\mid f(z)$, we have $r(z)\mid m(z)$ as well. Thus $m(z)$ can be factored into $m(z)=r(z)q(z)$, where $\deg(q(z))<\deg(m(z))$. Replacing z by P, we have $m(\mathrm{P})=r(\mathrm{P})q(\mathrm{P})=0$. Recall that $r(\mathrm{P})\in\{\mathrm{P},\mathrm{P}-\mathrm{I},\mathrm{P}+\mathrm{I},\mathrm{P}^2+\mathrm{I}\}$. By Theorem 3.1, $r(\mathrm{P})$ is nonsingular over $\mathbf{F}_p$ for $p\equiv 3 \mod 4$. Thus, we have $q(\mathrm{P})=0$ which is a contradiction to the fact that $m(z)$ is the minimal polynomial of P, since $\deg(q(z))<\deg(m(z))$. As a result, the assumption that $r(z)\mid f(z)$, is not true. ∎

Let v_k be the number of $k\times k$ matrices P over $\mathbf{F}_p$ without factors of $z, z-1, z+1$ or z^2+1 in their characteristic polynomial $f(z)$. We have

Lemma 3.6:

$$1+\sum_{k=1}^{\infty}\frac{v_k}{\gamma_k}u^k=\frac{1}{1-u}\prod_{r=1}^{\infty}(1-\frac{u}{p^r})^2(1-\frac{u^2}{p^{2r}}),\qquad |u|<1. \tag{10}$$

Proof: The proof is given in Section VI. ∎

As k increases, the asymptotic probability of such a matrix is

$$\prod_{r=1}^{\infty}(1-1/p^r)^3(1-1/p^{2r})$$

by the following theorem.

Theorem 3.7:

$$\lim_{k\to\infty}\frac{v_k}{p^{k^2}}=\prod_{r=1}^{\infty}(1-\frac{1}{p^r})^3(1-\frac{1}{p^{2r}}). \tag{11}$$

Proof: Similar to the proof of Theorem 3.4. ∎

IV. Check Digit System With Detection Radius R_c

The check digit systems proposed in last section are able to detect all the errors of types 1)-5) as listed in the introduction. In this section, we explore their detecting capability beyond that. In more detail, we aim to detect errors of type 1), 6) and 7). Note that error types 6) and 7) not only include but also generalize the error types 2)-5) in the term of jump length.

We recall the longest jump length t_c, such that for any $t\le t_c$, all the t-jump transpositions and t-jump twin errors will be detected. Note that t_c, or accordingly R_c (the maximum detection radius defined by $R_c=t_c+1$), reflects the capability range of a check digit system with check equation (6) on detecting double errors of types 6) and 7). A system capable of detecting all the single errors, and double errors of types 6) and 7) within R_c, is desirable since it further improves the error detection capability of the system proposed in last section.

Consider the check digit system for p^k numbers over the group $\mathbf{G}=(\mathbf{Z}_p)^k$. We apply the check equation (6).

- To detect all the t-jump transpositions, P has to satisfy the condition

  $6)'\quad a+c\cdot\mathrm{P}^{t+1}\neq c+a\cdot\mathrm{P}^{t+1}\quad$ for all $a\neq c\in\mathbf{G}$.

  It requires that $\mathrm{P}^{t+1}-\mathrm{I}$ must be nonsingular.
- For the t-jump twin errors, we have the condition

  $7)'\quad a+a\cdot\mathrm{P}^{t+1}\neq c+c\cdot\mathrm{P}^{t+1}\quad$ for all $a\neq c\in\mathbf{G}$.

  It requires that $\mathrm{P}^{t+1}+\mathrm{I}$ must be nonsingular. We note that when P is over $\mathbf{F}_2$, $\mathrm{P}^{t+1}-\mathrm{I}$ and $\mathrm{P}^{t+1}+\mathrm{I}$ are the same.

To achieve the longest detection radius R_c, it is required that $\mathrm{P}-\mathrm{I},\mathrm{P}^2-\mathrm{I},\ldots\mathrm{P}^{R_c}-\mathrm{I}$ and $\mathrm{P}+\mathrm{I},\mathrm{P}^2+\mathrm{I},\ldots,\mathrm{P}^{R_c}+\mathrm{I}$ are all nonsingular matrices over $\mathbf{F}_p$. In general, we have the following upper bound for R_c.

Lemma 4.1:

$$R_c\le p^k-2.$$

Proof: Consider $f(z)$, the characteristic polynomial of a nonsingular matrix P with $\deg f(z)=k$. By [19, Lemma 3.1], there exists a positive integer $e\le p^k-1$ such that $f(z)\mid z^e-1$. Replacing z by P in $f(z)$ we have $f(\mathrm{P})=0$. In addition that $f(z)\mid z^e-1$, we obtain $\mathrm{P}^e-\mathrm{I}=0$, for $e\le p^k-1$. In order to fullfil 6)', the detection radius $R=t^*+1$ must be less than e and so is R_c. Therefore, we have $R_c\le p^k-2$. ∎

Let $f(z)$ be the characteristic polynomial of a nonsingular matrix P. We have the following lemma.

Lemma 4.2: If $f(z)$ is irreducible, then for any non-zero polynomial $g(z)$, $g(\mathrm{P})$ is nonsingular if $f(z)\nmid g(z)$.

Proof: If $f(z)$, as the characteristic polynomial of P, is irreducible, then it is also the minimal polynomial of P such that $f(\mathrm{P})=0$, where 0 is the $k\times k$ all-zero matrix. If $f(z)\nmid g(z)$, by Euclid's algorithm, there are $p(z)$ and $q(x)$ such that

$$1=p(z)f(z)+q(z)g(z).$$

Replacing z by P and 1 by I, we obtain the following:

$$\mathrm{I}=p(\mathrm{P})f(\mathrm{P})+q(\mathrm{P})g(\mathrm{P})=q(\mathrm{P})g(\mathrm{P}),$$

which implies $g(P)$ is nonsingular. ∎

A. $p=2$

If $f(z)$, as the characteristic polynomial of P, is a primitive polynomial over $\mathbf{F}_2$, then by definition, the smallest integer e such that $f(z)$ divides z^e-1 is $e=2^k-1$. In other words, we have $f(z)\nmid z^t-1$ for any $t<2^k-1$. Then by Lemma 4.2, we have that $\mathrm{P}^t-\mathrm{I}$ is nonsingular for all $t<2^k-1$. Therefore, a digit check system built on such P is capable of detecting t-jump errors of types 6) and 7) within jump length $<2^k-2$. This gives us $R_c\ge 2^k-2$. In addition that $R_c\le p^k-2$ by Lemma 4.1, we conclude with the following theorem.

Theorem 4.3: For $p=2$, $R_c=2^k-2$.

Theorem 4.4: Let P be a matrix whose characteristic polynomial is a primitive polynomial over $\mathbf{F}_2$. Then a check digit

system built on P is able to detect all the single errors and double errors of type 6) and 7) within detection radius R_c.

It is easy to check that the binary matrix P as defined in (4), and used for the check digit system for hexadecimal numbers proposed in [1], actually has the primitive polynomial z^4+z^3+1 as its character polynomial. According to Theorem 4.4, the system could detect not only the error types 1)-5), but all the t-jump transpositions and t-jump twin errors for $t \leq 13$.

B. p as an Odd Prime

If $f(z)$, as the characteristic polynomial of P, is a primitive polynomial over $\mathbf{F}_p$, then by definition, the smallest integer e such that $f(z) \nmid z^e - 1$ is $e = p^k - 1$. Let $m = e/2$. We can factor $z^e - 1$ into $(z^m - 1)(z^m + 1)$. Since $f(z) \nmid z^m - 1$, we have $f(z) \mid z^m + 1$. So far, the detection radius R of the check digit system built on such P, satisfies that $R \leq m - 1$.

For any $m' < m$, clearly $f(z) \nmid z^{m'} - 1$, since $m' < e$. Now we suppose that there is an $m' < m$ such that $f(z) \mid z^{m'} + 1$. Then we have $f(z) \mid z^{2m'} - 1$ with $2m' < 2m = e$, which is contrary to the fact that e is the smallest integer such that $f(z) \mid z^e - 1$ holds. As a conclusion of the above discussion, we have that $f(z) \nmid z^{m'} - 1$ and $f(z) \nmid z^{m'} + 1$ hold for any $m' < m$. By Lemma 4.2, it follows that for all $1 \leq t < m$, $\mathrm{P}^t - \mathrm{I}$ and $\mathrm{P}^t + \mathrm{I}$ are nonsingular.

Therefore, a digit check system built on such P is capable of detecting all t-jump transpositions and t-jump twin errors within jump length $< m - 1$. This gives us $R = m - 1$ for systems based on P of this special kind, and establishes the achievability of Theorem 4.5.

Theorem 4.5: For an odd prime p, $R_c = (p^k - 1)/2 - 1$.

Proof: We have shown in above discussion that by employing P of a special kind, one can build a check digit system with $R = (p^k - 1)/2 - 1$. By definition of R_c, we easily obtain $R_c \geq (p^k - 1)/2 - 1$.

To establish the equality, we need to prove the converse part, i.e., $R_c \leq (p^k - 1)/2 - 1$. To do this, we show that for any possible choice of P, the check digit system built on P has a detection radius R where $R \leq (p^k - 1)/2 - 1$ always holds. Thus we derive $R_c \leq (p^k - 1)/2 - 1$. We refer to Section VIII for a detailed proof of this part. ∎

Theorem 4.6: Let P be a $k \times k$ matrix whose characteristic polynomial is a primitive polynomial (or its scalar multiplication) over $\mathbf{F}_p$. Then a check digit system over p^k numbers built on P is able to detect all the single errors and double errors of type 6) and 7) within detection radius R_c.

An example, which is able to detect all the single errors and double errors of type 6) and 7) with detection radius $R_c = (p - 1)/2 - 1$, is given in [15, Example 23] for a check digit system over $\mathbf{Z}_p$ for $p \geq 7$. One can see that our results contribute the case in at least twofold. First, our result is more universal since it is not bound by the size of the information bits n. Secondly, our proposal works for a check digit system over p^k numbers, where p is an odd prime (not necessarily ≥ 7), and $k \geq 1$ in a constructive manner.

Note that in case of $k = 1$, by Theorem 4.5, we easily obtain $R_c = 0$ for $p = 3$, and $R_c = 1$ for $p = 5$, respectively. This implies that a check digit system with check equation (6) is unable to detect all the errors of types 1)-5). This result is consistent with the note as stated in [15, Remark 5].

C. Choices of P

We recall that the companion matrix of a monic polynomial $g(z) = c_0 + c_1 z + \cdots + c_{k-1} z^{k-1} + z^k$ of a positive degree k over $\mathbf{F}_p$ is defined to be the following $k \times k$ matrix

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & 0 & \dots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -c_{k-1} \end{pmatrix}.$$

The characteristic polynomial of such a matrix is exactly $g(z)$ [19, p. 93].

It is known that there are $\phi(p^k - 1)/k$ primitive polynomials of degree k over $\mathbf{F}_p$, where $\phi(\cdot)$ is Euler's Totient function (one can refer to [19, Theorem 3.5 & Theorem 3.16] for this result). So given any of the primitive polynomials, an easy construction of matrix P suitable for Theorem 4.4 and Theorem 4.6, is to take the companion matrices of them.

Regarding the number of matrices which have primitive polynomials as their characteristic polynomials, one can refer to [21]. For instance, over $\mathbf{F}_2$ it is in total [21]

$$\prod_{i=1}^{k-1} (2^k - 2^i) \frac{\phi(2^k - 1)}{k}.$$

V. Examples

In this section, we compare our system with some well known and wildly used system such as ISBN, MEID and ISAN.

A. Check Digit System Over $\mathbf{Z}_{11}$

Example 5.1: The International Standard Book Number code (ISBN) is over $\mathbf{Z}_{11}$, and uses the check equation:

$$\sum_{i=1}^{10} i \cdot x_i \equiv 0 \mod 11. \tag{12}$$

This system detects all the errors of types 1)-5) with the exception of the twin error at places $(5, 6)$ as $5 + 6 = 11$.

In [15, Example 18], a modification of the ISBN code is proposed over $\mathbf{Z}_{11}$, which uses the following check equation:

$$\sum_{i=1}^{5} i \cdot x_i + \sum_{j=6}^{10} (5 - j) \cdot x_j \equiv 0 \mod 11. \tag{13}$$

This system improves the ISBN code in the manner that it detects all the errors of types 1)-5) without exceptions.

In our approach, we apply Theorem 4.6 for $p = 11$ and $k = 1$. Note that over $\mathbf{Z}_{11}$, the primitive elements are $2, 6, 7, 8$. We use $\alpha \in \{2, 6, 7, 8\}$ in the following check equation:

$$\sum_{i=1}^{10} \alpha^i \cdot x_i \equiv 0 \mod 11. \tag{14}$$

TABLE I
THE FORMAT OF MEID

MEID														Check Digit
Manufacturer Code								Serial Number						C
R	R	X	X	X	X	X	X	Z	Z	Z	Z	Z	Z	
x_{14}	x_{13}	x_{12}	x_{11}	x_{10}	x_9	x_8	x_7	x_6	x_5	x_4	x_3	x_2	x_1	

TABLE II
THE FORMAT OF ISAN

ISAN																
Root												Episode				Check Digit
R	R	R	R	R	R	R	R	R	R	R	R	E	E	E	E	C
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}

For instance, if we take $\alpha = 6$, Equation (14) has coefficients $\{\alpha^i \mod 11, 1 \leq i \leq 10\} = \{6, 3, 7, 9, 10, 5, 8, 4, 2, 1\}$. In particular, the MOD 11-2 system specified in ISO/IEC 7064 [22] can be considered as a special case of our system by taking $\alpha = 2$.

By Theorem 4.6, our system could detect not only the errors of types 1)-5), but also more generalized errors of types 6)-7) within detection radius $R_c = 4$. It is easy to check that this holds also for the modification system defined by (13). In fact, both systems are able to detect all the single errors, all the t-jump transposition errors, and almost all the t-jump twin errors with the only exception of the 4-jump twin error. The reason is that in our system, $\alpha^5 \equiv 10 \mod 11$ and thus $\alpha^i + \alpha^{i+5} = 11 \cdot \alpha^i$, for $1 \leq i \leq 5$; whilst in the modification system proposed in [15, Example 18], the coefficients at places $(i, i+5)$ are i and $-i$, respectively, which sum to 0. As a direct result, the 4-jump twin errors remain undetected in both systems.

Moreover, our system outperforms the original ISBN code and the modification system on detecting the phonetic errors $(\cdots a0 \cdots \to \cdots 1a \cdots$, where $a \geq 2)$. It only fails to detect the phonetic errors as $x_i = 10$ for $1 \leq i \leq 9$. However, in the ISBN code, number 10 is not used for $x_i, 1 \leq i \leq 9$ at all. So we could say that our system is able to detect all the possible phonetic errors occurred in ISBN; whilst both the original ISBN code and the modification proposed in [15, Example 18] fail to do so (one can refer [15, Example 17 & Example 18] for a detailed list of undetectable phonetic errors for both systems).

B. Check Digit System Over Hexadecimal Numbers

Example 5.2: A Mobile Equipment IDentifier (MEID) [23] is a globally unique 14-digit hexadecimal identification number for a physical piece of mobile station equipment. It is composed mainly of two basic components, the manufacturer code and the serial number, as shown in Table I.

For an MEID which contains at least one hexadecimal digit in the RR digits, the check digit is calculated using a slight modification of the Luhn formula, in the manner that all arithmetic is performed in base 16. One can refer to [24, Annex B] for the calculation. Note that the check digit is not part of the MEID and is not transmitted when the MEID is transmitted.

It is easy to check that MEID as a hexadecimal check digit system, could not detect the following errors:

- t-jump transposition error: $(\mathrm{F}, 0) \leftrightarrow (0, \mathrm{F})$ at places $(i, i+t)$ for all odd t satisfying that $1 \leq i \leq 14$, $i + t \leq 14$.
- t-jump twin error: $(x, x) \leftrightarrow (x+5, x+5)$, where $3 \leq x \leq 7$, at places $(i, i+t)$ for all odd t satisfying that $1 \leq i \leq 14, i + t \leq 14$.

Example 5.3: The International Standard Audiovisual Number (ISAN) [25] is a numbering system that enables the unique and persistent identification of any audiovisual works. An ISAN consists of 16 hexadecimal digits, which can be divided into two segments: root segment and episode segment, as shown in Table II. An appended check digit is calculated over the 16 ISAN digits according to a MOD 37, 36 system specified in accordance with ISO/IEC 7064 [22].

However, according to [22], the MOD 37, 36 system is unable to detect all the errors of types 1)-5). In fact, it fails to detect about 0.16% of error type 2), 2.8% of error types 3) and 5), and 1.7% of error type 4).

So both MEID and ISAN, two widely used hexadecimal check digit systems, fail to detect all the errors of types 1)-5).

Example 5.4: Following our approach as described in Theorem 4.4, one can construct alternative systems for MEID and ISAN as follows:

- Represent the hexadecimal numbers as elements of the abelian group $\mathbf{G} = \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_2 \oplus \mathbf{Z}_2$ as described in Section II-A.
- Find a 4×4 matrix P whose characteristic polynomial, i.e. $\det(z\mathrm{I} - \mathrm{P})$, is either $z^4 + z + 1$ or $z^4 + z^3 + 1$ (both are primitive polynomial of degree 4). For instance, we can follow Section IV-C and choose the companion matrix of $z^4 + z + 1$,

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

  as an easy choice of such P.
- Apply the following check equation to calculate the check digit a_{n+1} from the n information digits $a_1, \ldots, a_n$:

$$\sum_{i=1}^{n+1} a_i \cdot \mathrm{P}^i = 0.$$

It is easy to see that the above system serves as an alternative for MEID by letting $n = 14$, and an alternative for ISAN by

letting $n = 16$. Comparing to both MEID and ISAN, our system is capable of detecting all the errors of types 1)-5). Beyond that, it is able to detect both the t-jump transposition and t-jump twin errors within its detection radius $R_c = 14$.

In more detail, our system, as an alternative for MEID, is capable of detecting all the possibles errors of types 1) and 6)-7); whilst as an alternative for ISAN, is capable of detecting all the single errors, and almost all the possible jump errors of types 6)-7) with the only exception of 14-jump transposition and twin errors at places $(1, 16)$ and $(2, 17)$.

VI. Proof of Lemma 3.3 and Lemma 3.6

In order to count the number of matrices over $\mathbf{F}_p$ which are suitable for the check digit system proposed in Section III, we use the method of cycle index for matrices over finite field which is introduced by Kung [17] and extended by Stong [18]. First we recall the following lemma.

Lemma 6.1: [16] Let $\mathcal{A}$ be any set of monic irreducible polynomials with coefficients in $\mathbf{F}_p$. Let μ_k be the number of $k \times k$ matrices over $\mathbf{F}_p$ whose characteristic polynomial factors into powers of elements of $\mathcal{A}$. Then

$$1 + \sum_{k=1}^{\infty} \frac{\mu_k}{\gamma_k} u^k = \prod_{\phi \in \mathcal{A}} \prod_{r=1}^{\infty} (1 - \frac{u^{\deg \phi}}{p^{r \deg \phi}})^{-1}. \tag{15}$$

In Lemma 6.1, what is of interest is the coefficient of u^k for $k \geq 1$. In the following proof, we assume that $|u| < 1$.

Let Φ be the full set of monic irreducible polynomials with coefficient in $\mathbf{F}_p$. We apply Lemma 6.1.

Taking $\mathcal{A} = \Phi \backslash \{z\}$ gives us all the invertible matrices over $\mathbf{F}_p$ with $\mu_k = \gamma_k$ in Lemma 6.1. So

$$1 + \sum_{k=1}^{\infty} u^k = \prod_{\phi \in \Phi \backslash \{z\}} \prod_{r=1}^{\infty} (1 - \frac{u^{\deg \phi}}{p^{r \deg \phi}})^{-1}. \tag{16}$$

- If $p \equiv 1 \mod 4$, taking $\mathcal{A} = \Phi \backslash \{z, z-1, z+1, z-\alpha, z+\alpha\}$ gives us the set of matrices over $\mathbf{F}_p$ without any of $0, 1, -1, \alpha, -\alpha$ as eigenvalues. Thus we have $\mu_k = \rho_k$ in Lemma 6.1 and

$$1 + \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k = \prod_{\phi \in \Phi \backslash \{z, z-1, z+1, z-\alpha, z+\alpha\}} \prod_{r=1}^{\infty} (1 - \frac{u^{\deg \phi}}{p^{r \deg \phi}})^{-1}. \tag{17}$$

  In order to obtain the right side of the above equation, we multiply the right side of (16) by

$$\prod_{r=1}^{\infty} (1 - u/p^r)^4$$

  to take out the factors corresponding to the polynomials $z-1, z+1, z-\alpha, z+\alpha$. So we have the statement of Lemma 3.3:

$$\begin{aligned} 1 + \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k &= (1 + \sum_{k=1}^{\infty} u^k) \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^4 \\ &= \frac{1}{1-u} \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^4, \end{aligned}$$

  where the last equation is due to the fact that $(1-u)^{-1} = 1 + u + u^2 + u^3 + \cdots$ for $|u| < 1$.

- If $p \equiv 3 \mod 4$, $z^2 + 1$ is irreducible over over $\mathbf{F}_p$. Taking $\mathcal{A} = \Phi \backslash \{z, z-1, z+1, z^2+1\}$ gives us the set of matrices over $\mathbf{F}_p$ without $z, z-1, z+1, z^2+1$ in their characteristic polynomial. Thus we have $\mu_k = \upsilon_k$ in Lemma 6.1 and

$$1 + \sum_{k=1}^{\infty} \frac{\upsilon_k}{\gamma_k} u^k = \prod_{\phi \in \Phi \backslash \{z, z-1, z+1, z^2+1\}} \prod_{r=1}^{\infty} (1 - \frac{u^{\deg \phi}}{p^{r \deg \phi}})^{-1}.$$

  In order to obtain the right side of the above equation, we multiply the right side of (16) by

$$\prod_{r=1}^{\infty} (1 - u/p^r)^2 (1 - u^2/p^{2r})$$

  to take out the factors corresponding to the polynomials $z-1, z+1, z^2+1$. So we have the statement of Lemma 3.6:

$$\begin{aligned} 1 + \sum_{k=1}^{\infty} \frac{\upsilon_k}{\gamma_k} u^k &= (1 + \sum_{k=1}^{\infty} u^k) \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^2 (1 - \frac{u^2}{p^{2r}}) \\ &= \frac{1}{1-u} \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^2 (1 - \frac{u^2}{p^{2r}}), \end{aligned}$$

  where the last equation is due to the fact that $(1-u)^{-1} = 1 + u + u^2 + u^3 + \cdots$ for $|u| < 1$.

VII. Proof of Theorem 3.4

In this section, we give the proof of Theorem 3.4 in detail as follows.

First, we multiply $(1-u)$ to Eq. (8) and obtain

$$(1-u) + \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} (1-u) u^k = \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^4, \qquad |u| < 1.$$

Note that $0 \leq \rho_k / \gamma_k \leq 1$ by definition. Consider the case $u \neq 0$. For the series $\sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} |u|^k$, it is easy to see that for any $\epsilon > 0$, there exists $N_0 = \lfloor \ln_{|u|} (1 - |u|) \epsilon \rfloor$ such that

$$\sum_{k=N+1}^{N+m} \frac{\rho_k}{\gamma_k} |u|^k \leq \sum_{k=N+1}^{N+m} |u|^k < \frac{|u|^{N+1}}{1-|u|} < \epsilon$$

holds for any $N > N_0$ and $m \geq 1$. According to Cauchy's criterion for convergence, the series $\sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} |u|^k$ is convergent. Similarly, we can prove the same holds for $\sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} |u|^{k+1}$. In other words, both $\sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k$ and $\sum_{k=1}^{\infty} -\frac{\rho_k}{\gamma_k} u^{k+1}$ are absolutely convergent. Note that absolute convergence implies convergence. We have

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} (1-u) u^k &\overset{(a)}{=} \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k + \sum_{k=1}^{\infty} -\frac{\rho_k}{\gamma_k} u^{k+1} \\ &\overset{(b)}{=} \sum_{k=1}^{\infty} \frac{\rho_k}{\gamma_k} u^k + \sum_{k=2}^{\infty} -\frac{\rho_{k-1}}{\gamma_{k-1}} u^k \\ &\overset{(a)}{=} \frac{\rho_1}{\gamma_1} u + \sum_{k=2}^{\infty} (\frac{\rho_k}{\gamma_k} - \frac{\rho_{k-1}}{\gamma_{k-1}}) u^k, \end{aligned}$$

where (a) is due to the fact that if two series $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ are convergent, then the series $\sum_{k=1}^{\infty}(a_k + b_k)$ is convergent and $\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$; (b) is due to the fact that $\sum_{k=1}^{\infty} -\frac{\rho_k}{\gamma_k} u^{k+1}$ is absolutely convergent; and any rearrangements of the terms in the absolute convergent series will result in a new series that is converge to the same limit. Based on above discussions, we have

$$1 - u + \frac{\rho_1}{\gamma_1} u + \sum_{k=2}^{\infty} (\frac{\rho_k}{\gamma_k} - \frac{\rho_{k-1}}{\gamma_{k-1}}) u^k = \prod_{r=1}^{\infty} (1 - \frac{u}{p^r})^4.$$

Letting u go to 1, we obtain

$$\lim_{k \to \infty} \frac{\rho_k}{\gamma_k} = \prod_{r=1}^{\infty} (1 - \frac{1}{p^r})^4.$$

We conclude the proof by the fact

$$\lim_{k \to \infty} \gamma_k / p^{k^2} = \prod_{r=1}^{\infty} (1 - 1/p^r)$$

as in [26].

VIII. Proof of Theorem 4.5

The achievability of Theorem 4.5 has been established in the discussion in Section IV-B. In this section, we give the proof of the converse part. To do that, we show for any P, the detection radius R of the the system defined by (6) satisfies $R \leq (p^k - 1)/2 - 1$ and therefore $R_c \leq (p^k - 1)/2 - 1$ holds.

Let $f(z)$ be the characteristic polynomial of the matrix P. Since P is nonsingular, we have $f(0) = \det(\mathrm{P}) \neq 0$. Denote the order of f by $\mathrm{ord}(f) = \mathrm{ord}(f(z))$, to be the least positive integer e for which $f(z)$ divides $z^e - 1$. Without loss of generality, we assume that $f(z)$ is a monic polynomial. By the definition of $\mathrm{ord}(f)$, we have $f(z) \mid z^{\mathrm{ord}(f)} - 1$. Clearly the detection radius R satisfies that $R \leq \mathrm{ord}(f) - 1$.

As a polynomial of degree k over $\mathbf{F}_p$, $f(z)$ must belong to one of the following three categories:

1) $f(z)$ *is irreducible, and in particular a primitive polynomial*. Then we have $\mathrm{ord}(f) = p^k - 1$ by definition. Following the discussion in Section IV-B, we have the detection radius $R = (p^k - 1)/2 - 1$ for the system built on such P.
2) $f(z)$ *is irreducible, but not a primitive polynomial*. Then according to [19, Corollary 3.4], we have $\mathrm{ord}(f) \mid p^k - 1$, but $\mathrm{ord}(f) \neq p^k - 1$. In this case we have $\mathrm{ord}(f) \leq (p^k - 1)/2$ for p as an odd prime. Thus we have the detection radius $R \leq (p^k - 1)/2 - 1$.
3) $f(z)$ *is reducible such that* $f = f_1^{b_1} f_2^{b_2} \cdots f_s^{b_s}$, where $s \geq 2$, and $\{f_i, 1 \leq i \leq s\}$ are distinct monic irreducible polynomials with $d_i = \deg(f_i)$ and $\sum_{i=1}^{s} b_i d_i = k$.
   - First consider the special case: $b_i = 1$ for $1 \leq i \leq s$. According to [19, Theorem 3.9], we have $\mathrm{ord}(f) = \mathrm{lcm}(\mathrm{ord}(f_1), \ldots, \mathrm{ord}(f_s))$ with $d_i = \deg(f_i)$ and $\sum_{i=1}^{s} d_i = k$.

$$\begin{aligned} \mathrm{ord}(f) &= \mathrm{lcm}(\mathrm{ord}(f_1), \ldots, \mathrm{ord}(f_s)) \\ &\leq \mathrm{lcm}(p^{d_1} - 1, \ldots, p^{d_s} - 1) \\ &\leq \frac{\prod_{i=1}^{s}(p^{d_i} - 1)}{\gcd(p^{d_1} - 1, \ldots, p^{d_s} - 1)} \\ &\leq \frac{\prod_{i=1}^{s}(p^{d_i} - 1)}{p - 1} \\ &< \frac{(p^{\sum_{i=1}^{s} d_i} - 1)}{p - 1} \\ &\leq \frac{(p^k - 1)}{2}. \end{aligned}$$

   As a direct result, we have in this case $R < (p^k - 1)/2 - 1$.
   - Let us look at the general case: there exists at least some $b_i > 1$, where $1 \leq i \leq s$. According to [19, Theorem 3.11], we have in this case $\mathrm{ord}(f) = p^l \cdot \mathrm{lcm}(\mathrm{ord}(f_1), \ldots, \mathrm{ord}(f_s))$, where l is the smallest integer with $p^l \geq \max(b_1, \ldots, b_s)$. (Note that the above special case is actually within this general case as $l = 0$.) Since $\{f_i, 1 \leq i \leq s\}$ are irreducible polynomials, we have $\mathrm{ord}(f_i) \mid p^{d_i} - 1$. Therefore,

$$\begin{aligned} \mathrm{ord}(f) &= p^l \cdot \mathrm{lcm}(\mathrm{ord}(f_1), \ldots, \mathrm{ord}(f_s)) \\ &\leq p^l \cdot \mathrm{lcm}(p^{d_1} - 1, \ldots, p^{d_s} - 1) \\ &\leq p^l \cdot \frac{\prod_{i=1}^{s}(p^{d_i} - 1)}{\gcd(p^{d_1} - 1, \ldots, p^{d_s} - 1)} \\ &\leq p^l \cdot \frac{\prod_{i=1}^{s}(p^{d_i} - 1)}{p - 1} \\ &< p^l \cdot \frac{(p^{\sum_{i=1}^{s} d_i} - 1)}{p - 1} \\ &\leq \frac{(p^{l + \sum_{i=1}^{s} d_i} - 1)}{2}. \end{aligned}$$

   Recall that l is the smallest integer with $p^l \geq \max(b_1, \ldots, b_s)$. Let $b_w = \max(b_1, \ldots, b_s)$. Then $l = \lceil \log_p b_w \rceil$. Considering that b_w and d_w are positive integers, we easily have

$$l = \lceil \log_p b_w \rceil \leq b_w - 1 \leq (b_w - 1) d_w.$$

   Therefore,

$$l + \sum_{i=1}^{s} d_i \leq (b_w - 1) d_w + \sum_{i=1}^{s} d_i \leq \sum_{i=1}^{s} b_i d_i = k.$$

   This gives us $\mathrm{ord}(f) < (p^k - 1)/2$ and thus $R < (p^k - 1)/2 - 1$.

Note that for any $k \times k$ matrix P over $\mathbf{F}_p$, its characteristic polynomial must belong to one of the three categories listed above. We show in above discussions that $R \leq (p^k - 1)/2 - 1$ always holds for p as an odd prime. As a conclusion, we have $R_c \leq (p^k - 1)/2 - 1$ and thus complete the proof.

IX. Conclusion

In this paper, we study a check digit system which is based on the use of elementary abelian p-groups of order p^k, for $k \geq 1$ and a prime p. The work is inspired by the check digit system for hexadecimal numbers proposed in [1] which is able to detect all the 1) single errors, 2) adjacent transpositions, 3) twin errors, 4) jump transpositions and 5) jump twin errors.

In our treatment, we interpret the check equation in terms of matrix equation. In order to detect all the five error types which are of most interest in the previous literature, we explore the suitable matrices with desirable properties. We also determine the number of such matrices for a given order k, and the asymptotic probability of finding them as k tends to ∞.

Furthermore, we consider two categories of jump errors: 6) t-jump transpositions and 7) t-jump twin errors, which include and further extend the double error types of 2)-5). In particular, a check digit system capable of detecting all the single errors and these two kinds of double errors with maximum detection radius R_c, is desirable due to its improved performance on error detection capability. We show that for $p = 2$, $R_c = 2^k - 2$; and for an odd prime p, $R_c = (p^k - 1)/2 - 2$. Last but not least, using matrices of a special kind, those whose characteristic polynomials are primitive polynomials, we show how to construct check digit systems that are desirable for being able to detect all the errors of type 1) single errors, and of types 6) and 7) within the longest detection radius.

Our approach is simple, constructive and easy to be adopted in practice. Its realizations over $\mathbf{Z}_{11}$ and hexadecimal numbers outperform the existing systems on detecting the most frequent single and double errors occurring in the transmission by human operation, and therefore serve as attractive alternatives for ISBN code, MEID and ISAN. In general, our study focuses on the check digit system with one check symbol over a special class of group of order p^k. Unfortunately, it does not cover the decimal number system which is widely used and favored in human operations. So far, the best known result for a check digit system over a group of order 10 is built over the Dihedral group D_5 [3]. Nevertheless, the system is able to detect all the errors of types 1)-2), but not all the errors of types 3)-5) or further. So it still remains an open problem in this specific case regarding the existence or design of a system which is capable of detecting all the errors of types 1)-5). Another interesting research direction to be considered is the maximum error detection capability that can be achieved by one single check symbol.

References

[1] M. Niemenmaa, "A check digit system for hexadecimal numbers," *Appl. Algebra Eng., Commun. Comput.*, vol. 22, no. 2, pp. 109–112, 2011.

[2] D. F. Beckley, "An optimum system with modulo 11," *Comput. Bull.*, vol. 11, pp. 213–215, Jan. 1967.

[3] J. Verhoeff, "Error detecting decimal codes," *Mathematisch Centre Tracts*, vol. 29. The Netherlands: Amsterdam, Mathematisch Centrum, 1969.

[4] H. M. Damm, "Check digit systems over groups and anti-symmetric mappings," *Archiv der Mathematik*, vol. 75, no. 6, pp. 413–421, 2000.

[5] H. P. Gumm, "A new class of check-digit methods for arbitrary number systems," *IEEE Trans. Inf. Theory*, vol. 31, no. 1, pp. 102–105, Jan. 1985.

[6] A. Ecker and G. Poch, "Check character systems," *Computing*, vol. 37, no. 4, pp. 277–301, 1986.

[7] J. A. Gallian and M. D. Mullin, "Groups with anti-symmetric mappings," *Archiv Math.*, vol. 65, no. 4, pp. 273–280, 1995.

[8] S. Heiss, "Anti-symmetric mappings for finite solvable groups," *Archiv Math.*, vol. 69, no. 6, pp. 445–454, 1997.

[9] R. H. Schulz, "Check character systems over groups and orthogonal Latin squares," *Appl. Algebra Eng., Commun. Comput.*, vol. 7, pp. 125–132, Jan. 1996.

[10] R. H. Schulz, "On check digit systems using anti-symmetric mappings," in *Numbers, Information and Complexity*. Boston, MA, USA: Kluwer Academic, 2000, pp. 295–310.

[11] R. H. Schulz, "A note on check character systems using Latin squares," *Discrete Math.*, vol. 97, nos. 1–6, pp. 371–375, 1991.

[12] G. B. Belyavskaya, V. I. Izbash, and G. L. Mullen, "Check character systems using quasigroups: I," *Designs, Codes Cryptograph.*, vol. 37, no. 2, pp. 215–227, 2005.

[13] G. B. Belyavskaya, V. I. Izbash, and G. L. Mullen, "Check character systems using quasigroups: II," *Designs, Codes Cryptograph.*, vol. 37, no. 3, pp. 405–419, 2005.

[14] C. Broecker, R. H. Schulz, and G. Stroth, "Check character systems using Chevalley groups," *Designs, Codes Cryptograph.*, vol. 10, no. 2, pp. 137–143, 1997.

[15] G. L. Mullen and V. Shcherbacov, "n-T-quasigroup codes with one check symbol and their error detection capabilities," *Comment Math. Univ. Carolinae*, vol. 45, no. 2, pp. 321–340, 2004.

[16] K. E. Morrison, (2013). *Matrices Over F_q With no Eigenvalues of 0 or 1* [Online]. Available: http://www.calpoly.edu/~kmorriso/Research/mnev01.pdf

[17] J. P. S. Kung, "The cycle structure of a linear transformation over a finite field," *Linear Algebra Appl.*, vol. 36, pp. 141–155, Mar. 1981.

[18] R. Stong, "Some asymptotic results on finite vector spaces," *Adv. Appl. Math.*, vol. 9, no. 2, pp. 167–199, 1988.

[19] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and their Applications*. Cambridge U.K.: Cambridge Univ. Press, 1986.

[20] S. Lipschutz and M. Lipson, *Schaum's Outline of Linear Algebra*. New York, NY, USA: McGraw-Hill, 2009.

[21] P. Choudhury, *Generating Matrices of Highest Order Over a Finite Field*. Ithaca, NY, USA: Cornell Univ. Press, Nov. 2005.

[22] *Information Technology—Security Techniques—Check Character Systems*, ISO/IEC Standard 7064, 2003.

[23] *3G Mobile Equipment Identifier (MEID)*, 3GPP2 S.R0048-A, Jun. 2005.

[24] *MAP Support for the Mobile Equipment Identity (MEID)*, 3GPP2 X.S0008-0, Jan. 2009.

[25] *Information and Documentation—International Standard Audiovisual Number (ISAN)*, ISO Standard 15706, 2002.

[26] L. S. Charlap, H. D. Rees, and D. P. Robbins, "The asymptotic probability that a random biased matrix is invertible," *Discrete Math.*, vol. 82, no. 2, pp. 153–163, 1990.

Yanling Chen received her B.S. degree in 2001 and her M.S. degree in 2004, both in applied mathematics from Nankai University, Tianjin, China. In 2007, she received her Ph.D. degree in engineering from the University of Duisburg-Essen, Germany. Currently she is a researcher at the Ruhr University Bochum, Germany. Her research interests include information theory, coding theory and cryptography.

Markku Niemenmaa is a professor of Algebra and its Applications at the university of Oulu, Finland. He received his PhD in mathematics from the university of Tampere, Finland in 1978. His research interests are check digit systems, group theory and non-associative binary systems.

A. J. Han Vinck is a full professor in Digital Communications at the University of Essen, Essen, Germany, since 1990. He studied electrical engineering at the University of Eindhoven, the Netherlands, where he obtained his Ph.D. in 1980.

He accepted the position of visiting professor (2010-2012) at the University of Johannesburg, South Africa. He was invited (2011) to be consultant professor at the Harbin Institute of Technology, Harbin, China. In 2003 he was an adjoint professor at the Sun Yat-Sen University in Kaohsiung, Taiwan. In 1986 he was a visiting scientist at the German space Agency in Oberpfaffenhofen, Germany. His interest is in Information and Communication theory, Coding and Network aspects in digital communications. He is the author of the book Coding Concepts and Reed-Solomon Codes.

Professor Vinck served on the Board of Governors of the IEEE Information Theory Society from1997 until 2006. In 2003 he was elected president of the IEEE Information theory Society. In 1997 he acted as Co-chairman for the 1997 IEEE Information Theory symposium in Ulm, Germany (704 participants). He was founding Chairman (1995-1998) of the IEEE German Information Theory chapter. In 1990 he organized the IEEE Information Theory workshop in Veldhoven, the Netherlands(125 participants). IEEE elected him in 2006 as a fellow for his Contributions to Coding Techniques. Professor Vinck is the initiator and organizer of the Japan-Benelux workshops on Information theory (now Asia-Europe workshop on Concepts in Information Theory) and the International Winterschool on Coding, Cryptography and Information theory in Europe. He started (Essen, 1997) and still supports the organization of the series of conferences on Power Line Communications and its Applications, ISPLC. In 2006 he received the IEEE ISPLC2006 Achievement award in Orlando (FL, USA) for his contributions to Power Line Communications and for facilitating the transition of ISPLC to a fully financially and technically sponsored IEEE Communications Society conference. The SAIEE annual award was presented to him for the best paper published in the SAIEE Africa Research Journal in the year 2008.

He is co-founder and president of the Leibniz foundation. This foundation stimulates research in the field of Information theory, Neuroscience and Biology.

Danilo Gligoroski is a Professor of Information Security and Cryptography at Norwegian University of Science and Technology (NTNU), Trondheim, Norway. He received the PhD degree in Computer Science from Institute of Informatics, Faculty of Natural Sciences and Mathematics, at University of Skopje Macedonia in 1997. His research interests are Cryptography (especially Hash functions, Stream ciphers, Block ciphers, Ultra Fast Public Key Algorithms), Information Security, Discrete algorithms and Information Theory and Coding.