

## Enhancing spoken connected-digit recognition accuracy by error correction codes – A novel scheme

SUNIL K KOPPARAPU and P V S RAO

Cognitive Systems Research Laboratory, Tata Infotech Limited, Plot No 14,  
Sector 24, Vashi-Turbe, Navi Mumbai 400 705, India  
e-mail: {SunilKumar.Kopparapu, pvs.rao}@TataInfotech.com

MS received 25 July 2002; revised 5 July 2004

**Abstract.** Recognizing spoken connected-digit numbers accurately is an important problem and has very many applications. Though state-of-the-art word recognition systems have gained acceptable accuracy levels, the accuracy of recognition of current connected spoken digits (and other short words) is very poor. In this paper, we develop a novel scheme to enhance the accuracy of recognizing a connected number. The basic idea proposed in this paper is to increase the number of digits in a number and use these appended digits to increase the overall accuracy of recognizing the number, as is done in the error-correcting code literature. We further show that the developed scheme is able to uniquely and exactly correct single-digit errors.

**Keywords.** Connected-digit recognition; spoken number recognition; error-correcting codes.

### 1. Introduction

Error-correcting codes are a well-researched area and are often used to take care of errors introduced when data streams are transmitted from one place to another through some communication channel. Useful and efficient error-correcting codes have several objectives.

- (1) Detection and correction of errors introduced in the channel. This is the most important objective of an error-correcting code; ideally one would like a code that is capable of correcting all errors due to noise;
- (2) efficient transmission of data, so as to not waste time sending extra data;
- (3) easy encoding and decoding schemes; this is also desirable for computational efficiency.

In general, the more errors a code needs to correct per message digit, the less efficient the transmission and also probably the more complicated the encoding and decoding schemes. A good code balances these objectives (Rosenquist 2003).

Much of the study in the area of error-correcting codes has been in the area of binary block codes, meaning that the transmitted message is encoded as a string of 1's and 0's and the code

under consideration is itself of constant length. Binary error-correcting codes are defined on a field  $\{0, 1\}$  with addition, multiplication and division operations. Non-binary linear codes have also been addressed in literature, but are restricted to a field  $\{0, 1, \dots, p\}$ , where  $p$  is a prime. Without going into the specifics of different error-correcting codes, we might state that, broadly speaking, these have the same components: the generator matrices, parity check matrices and syndrome decoding.

Non-binary, non-prime information is often used to code information such as postal code, bar code and ISBN<sup>1</sup> library code to name a few. For example, in character-recognition systems decimal numbers are popularly used (as in postal codes) and are processed directly, not transformed to binary (0's and 1's) data. Wagner (2002) enumerates different error-detecting schemes that are used in practice. Wangner & Putter (1989) specifically look at codes capable of detecting errors in decimal system while Sethi *et al* (1978) talk of error-correcting codes for alphanumeric data. Gumm (1985) presents a system which is capable of detecting all single-digit errors and adjacent-digit transpositions using a single check digit for arbitrary number systems. Brown (1973) talks of construction of error detection and correction codes for any arbitrary number system.

In this paper, we present a novel scheme to detect and correct single-digit (decimal number, hence non-binary and non-prime) errors. The idea is to enhance the accuracy of recognition of connected spoken digit recognition, which is otherwise very poor as reported in literature and discussed elsewhere in this paper.

Speech as a medium is being increasingly used in human computer interface applications. There has been fair success in terms of achieving accuracy in word-based speech recognition systems and this has led to speech recognition and speaker verification systems being deployed in practical applications such as auto-attendant and personal authentication biometric systems. Connected-digit recognition is important and has use in several application areas. Several applications (for example banks, courier or postal agencies to track parcels, railway or travel agencies to confirm reservations, B2C portals to give credit card numbers etc.) have need for a reliable connected-digit recognizer. In addition, as speech gains popularity as the medium of interaction between human and the computer, and the confidence of the users increases, most of the applications which earlier used the telephone key pad to *obtain* numbers are being speech-enabled. While word based (and even phoneme based) recognition systems have reached acceptable accuracy levels, this can not be said of recognition of connected-digit numbers. For this reason, one does as yet not find in the market systems which are able to reliably recognize spoken connected-digit numbers.

Current connected-digit recognition systems (for example, Buhrke *et al* 1994, Kopec & Bush 1985) use sophisticated algorithms to recognize spoken digit strings using whole word models<sup>2</sup>. These algorithms are generally based on maximum likelihood recognisers such as hidden Markov modeling and neural networks. A good and a reliable connected-digit recognizer is therefore the need of the hour.

In this paper, we describe a simple, yet novel scheme to increase the recognition accuracy of a connected-digit recognizer without increasing the digit recognizing accuracy *per se*. The main idea is to append extra digits and use the resulting longer number to increase the accuracy of recognizing a connected-digit number.

---

<sup>1</sup>“International Standard Book Number” is a unique machine-readable book identification number

<sup>2</sup>Isolated word recognition

## 2. Connected-digit recognition

The accuracy of recognition of spoken connected-digits is low, even though the accuracy of recognizing longer words has reached acceptable levels. This is due to several reasons, for example as below.

- (1) Spoken digits are of short acoustic duration, typically a few seconds of speech,
- (2) some digits are acoustically very similar to each other (for example, *one* and *nine*), and
- (3) detecting a connected number involves identifying several individual digits that form the number; hence, overall accuracy falls to a much lower level.

For example, if the probability of recognizing a single-digit ( $d1$ ) correctly is  $p_{d1}^c$ , then the probability of recognizing a connected number made up of  $n$  digits (represented by  $dn$ ) is

$$P_{dn}^c = (p_{d1}^c)^n.$$

Clearly, connected-word recognition accuracy is much *poorer* compared to the individual digit recognition accuracy. Even a respectable single-digit accuracy of 0.95 gives a 6-digit number recognition accuracy of  $(0.95)^6$  or only 0.735! This paper discusses methods for increasing this accuracy of connected numbers, without actually increasing digit accuracy. Figure 1 shows the probability of correctly recognizing connected  $n$ -digit numbers (for  $n = 1, 2, 6, 8$ ) with increasing single-digit recognition accuracy.

## 3. Increasing connected-digit recognition accuracy

Consider a spoken number, made up of  $\alpha$  digits, that has to be recognized. Let  $p_{d1}^c$  denote the probability of correct recognition of a single-digit; then  $p_{d1}^e = 1 - p_{d1}^c$  denotes the probability of error in recognition of a single-digit. Note that  $0 \leq p_{d1}^c, p_{d1}^e \leq 1$ . The accuracy of recognition of an  $\alpha$ -digit number (probability of correct recognition of all the  $\alpha$  digits) is

$$P_{d\alpha}^c = (p_{d1}^c)^\alpha. \tag{1}$$

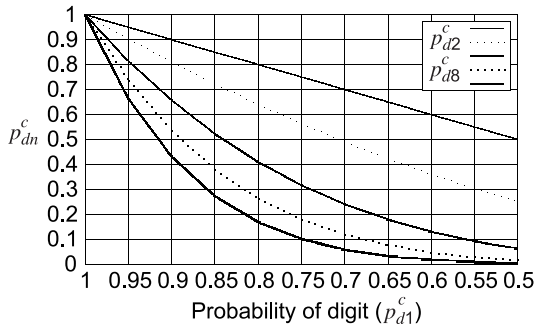
Suppose we append  $\beta$  extra digits<sup>3</sup> to the  $\alpha$ -digit number; then the accuracy of recognition (probability of correctly recognizing all the  $\gamma \stackrel{def}{=} (\alpha + \beta)$  digits) of the number is

$$P_{d\gamma}^c = (p_{d1}^c)^\gamma.$$

Clearly,  $p_{d\gamma}^c \leq p_{d\alpha}^c$ . Adding extra digits reduces the accuracy of number recognition (compare the curves corresponding to  $p_{d\alpha=6}^c$  and  $p_{d\gamma=8}^c$  in figure 1. Note that the advantage of adding extra digits is  $p_{d\gamma}^c - p_{d\alpha}^c$ , which is  $< 0$ . Now, suppose the  $\gamma$ -digit number is constructed such that it is possible to identify and correct *exactly* and *only*  $k$  digits in error in the  $\gamma$ -digit number, then, the advantage of appending the extra  $\beta$  digits to the  $\alpha$  digit number, using the theory of binomial expansion, is

$$P_{ad}^{\alpha \rightarrow \gamma}(k) = \{p_{d\gamma}^c - p_{d\alpha}^c\} +^\gamma C_k (p_{d1}^c)^{\gamma-k} (p_{d1}^e)^k. \tag{2}$$

<sup>3</sup>Use of the extra  $\beta$  digits is explained in the later part of the paper



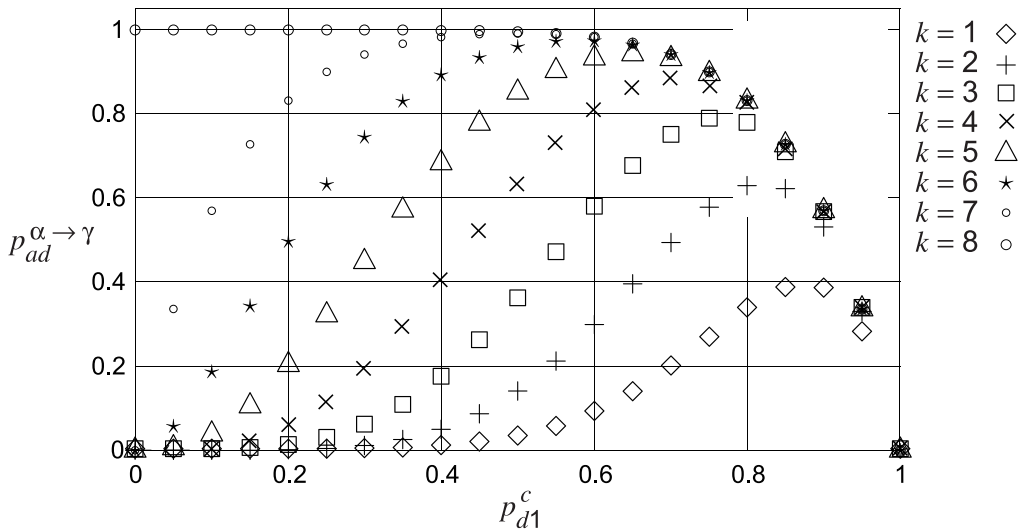
**Figure 1.** Number recognition accuracy versus digit recognition accuracy.

Assuming that the ability to identify and correct  $k$  digits in error implies the ability to identify and correct  $k$  or less digits in error, we have

$$p_{ad}^{\alpha \rightarrow \gamma} = \{p_{d\gamma}^c - p_{d\alpha}^c\} + \sum_{j=1}^k \gamma C_j (p_{d1}^c)^{\gamma-j} (p_{d1}^e)^j. \tag{3}$$

Figure 2 shows the recognition advantage derived due to  $k(= 1, 2, 3, 4, 5, 6, 7, 8)$  digits in error correction capability for  $\alpha = 6$  and  $\gamma = 8$ . Notice in figure 2 the following.

- (1) For a given  $k$ ,  $p_{ad}^{\alpha \rightarrow \gamma}$  peaks at a certain value of  $p_{d1}^c$ .  
 This is because, for a given  $k$ , the advantage due to error correction capability is less for smaller values of  $p_{d1}^c$  as also for larger values of  $p_{d1}^c$ , though for different reasons. For larger  $p_{d1}^c$  there is more probability of the  $\gamma$  digit number being recognized correctly even without the error correction capability coming into picture, meaning less advantage due to the error-correcting capability. For smaller  $p_{d1}^c$  there is a higher probability of larger than



**Figure 2.** Recognition advantage of using  $\gamma = 8$  digits instead of  $\alpha = 6$  for  $k = 1, 2, 3, 4, 5, 6, 7, 8$  digits in error-correction capability versus probability of single-digit recognition.

$k$  (number of digits that can be corrected by the error correction capability) digits being in error which renders the error-correcting capability useless. Somewhere in between, given by (5), the advantage of using the error-correcting capability peaks.

- (2) The advantage derived by using the error correction capability increases with increase in  $k$  (maximum value of  $p_{ad}^{\alpha \rightarrow \gamma}$  for  $k = 7$  is greater than maximum value of  $p_{ad}^{\alpha \rightarrow \gamma}$  for  $k = 1$ ).
- (3) The peak value of  $p_{ad}^{\alpha \rightarrow \gamma}$  for a given  $k$  shifts to a lower  $p_{d1}^c$  with increasing  $k$ .  
This behaviour is to be expected because with increasing  $p_{d1}^c$  there is less chance of making an error and hence no advantage of using the error detection and-correction capability; but as  $p_{d1}^c$  decreases there is more likelihood of more digits in the number being in error and hence the error correction capability is advantageous.

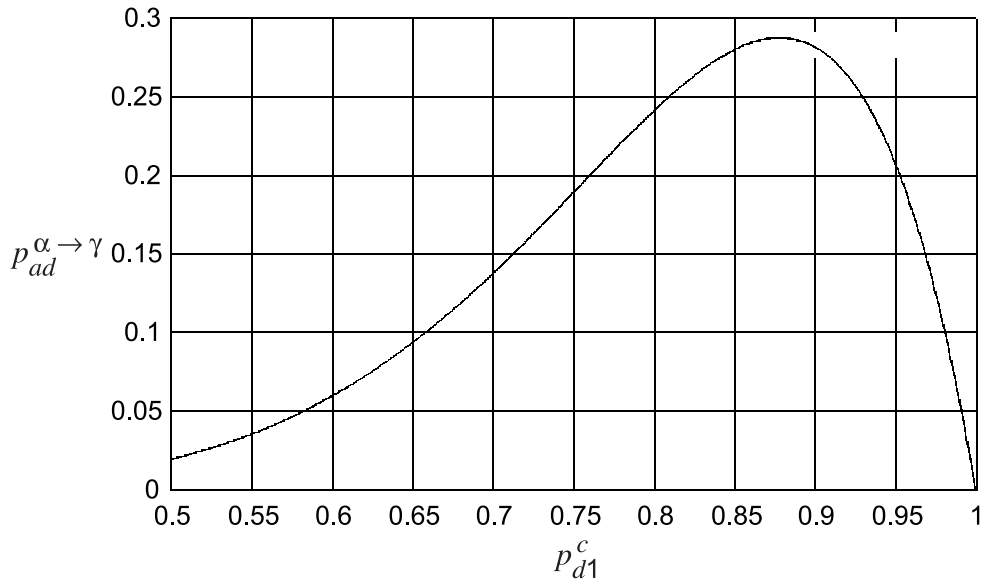
Specifically for single-error digit correction ( $k = 1$ ), the advantage of using  $\gamma$  digits instead of  $\alpha$  digits is (figure 3, for  $\alpha = 6, \gamma = 8$  and  $k = 1$ ),

$$p_{ad}^{\alpha \rightarrow \gamma} = \{p_{d\gamma}^c - p_{d\alpha}^c\} +^\gamma C_1(p_{d1}^c)^{\gamma-1}(p_{d1}^e)^1. \tag{4}$$

Clearly, there is a distinct advantage in using  $\gamma$  digits instead of  $\alpha$  digits, provided the extra  $\beta$  digits can be used to identify and correct errors of  $k$  digits or less.

*Note 1.* Figure B1 (appendix B) shows the plot of advantage of using  $\gamma = 8$  digits (with the capability of identifying and correcting  $k = 1$  error digit) instead of  $\alpha = 6$ , namely,  $p_{ad}^{\alpha \rightarrow \gamma}$  versus  $p_{d1}^c$ , the correct recognition of a single-digit. Notice from the plot that there is always an advantage ( $p_{ad}^{\alpha \rightarrow \gamma}$  is positive) of using  $\gamma$  digits instead of  $\alpha$  digits, for  $p_{d1}^c > 1/(\gamma - 1)$  (see appendix A, provided we are able to identify and correct  $k = 1$  digits in error).

In Section 4 we propose a simple yet novel scheme which is able to identify and correct such errors by appending  $\beta$  extra digits to the original  $\alpha$  digits. It can be shown that, the



**Figure 3.** Advantage of using  $\gamma = 8$  digits instead of  $\alpha = 6$  for single-error digit correction ( $k = 1$ ) number recognition versus probability of single-digit recognition.

maximum increase in accuracy of recognition (because of being able to detect and identify *exactly*  $k$  error digits in an  $\gamma$  digit number) occurs at

$$p_{d1}^c = (\gamma - k)/\gamma, \tag{5}$$

and the maximum increase in probability of correct recognition is

$${}^\gamma C_k ((\gamma - k)/\gamma)^{\gamma-k} (k/\gamma)^k.$$

We show this in appendix B.1, namely, the resulting advantage in recognition due to *exactly*  $k$  digits in error correction.

*Note 2.* It must be noted that the ability to identify and correct  $k$  errors implies the ability to identify and correct  $k$  or less digits in error (3). As shown by figure B1 in Appendix B.2, there exists an optimal  $p_c$  where the recognition advantage is maximum due to use of our scheme (§ 4). We show in appendix B.2 that optimal  $p_c$  is a function of  $\gamma$  and  $k$ . To the best of our knowledge an analytical solution (solving a general polynomial in  $p_c$  of order  $k > 1$ ) for identifying the optimal  $p_c$  is not possible.

#### 4. Proposed scheme

Consider the case where  $\alpha = 6$ , namely a decimal number  $N_\alpha = d_1d_2d_3d_4d_5d_6$ , where  $0 \leq d_i \leq 9$ . We append  $\beta = 2$  extra digits to generate an  $\gamma = 8$ -digit number  $N_\gamma = d_1d_2d_3d_4d_5d_6d_7d_8$ . The construction of  $d_{\alpha+1}$  and  $d_{\alpha+2}$  is as follows,

$$d_{\alpha+1} = 10 - \text{modulo}(d'_{\alpha+1}, 10), \tag{6}$$

where

$$d'_{\alpha+1} = \sum_{i=1}^{\alpha} d_i.$$

Also,

$$d_{\alpha+2} = \text{modulo}(d'_{\alpha+2}, 11), \tag{7}$$

where

$$d'_{\alpha+2} = \sum_{i=1}^{\alpha+1} w_i d_i,$$

and  $\{w_i\}_{i=1}^{\alpha+1}$  are chosen such that

$$w_i \neq w_j, \quad \text{if } i \neq j. \tag{8}$$

As will be shown in this paper,  $d_{\alpha+1}$  is used to identify a single-error and  $d_{\alpha+2}$  to locate the position of the error. The  $\beta = 2$  digits can be used to not only detect a single-error but also to identify the error; hence all single-digit errors can be corrected. We know from (6) that

$$\text{modulo} \left( \sum_{i=1}^{\alpha} d_i, 10 \right) + d_{\alpha+1} = 10. \tag{9}$$

Suppose, there is a single-digit error in recognizing  $N_\gamma$  and the transmitted number is

$$\hat{N}_\gamma = d_1 \cdots d_{\eta-1} D_\eta d_{\eta+1} \cdots d_{\gamma-1} d_\gamma,$$

where,  $1 \leq \eta \leq \gamma$ . Now, find

$$\begin{aligned} E &= \text{modulo} \left( \left\{ \sum_{i=1}^{\eta-1} d_i + D_\eta + \sum_{i=\eta+1}^{\alpha+1} d_i \right\}, 10 \right) \\ &= \text{modulo} \left( \left\{ \sum_{i=1}^{\alpha+1} d_i + (D_\eta - d_\eta) \right\}, 10 \right) \\ &= \text{modulo} (\{10 + (D_\eta - d_\eta)\}, 10) \\ &= \text{modulo}((D_\eta - d_\eta), 10) \\ &= \text{modulo}(e, 10), \end{aligned} \tag{10}$$

where

$$e = (D_\eta - d_\eta). \tag{11}$$

If we know  $\eta$ , the position of the error, we can reconstruct  $N_\alpha$  as

$$N_\alpha = d_1 \cdots d_{\eta-1} d'_\eta d_{\eta+1} \cdots d_\alpha,$$

where

$$\begin{aligned} d'_\eta &= D_\eta - E, \quad \text{if } (D_\eta - E) \geq 0, \\ &= 10 + (D_\eta - E), \quad \text{if } (D_\eta - E) < 0. \end{aligned} \tag{12}$$

Since we do not know the exact location of the error, namely  $\eta$ , we construct  $(\alpha + 1)$  different numbers  $C_\zeta$ ,

$$C_\zeta = d_1 \cdots d_{\zeta-1} (d_\zeta - E) d_{\zeta+1} \cdots d_{\eta-1} D_\eta d_{\eta+1} \cdots d_{\alpha+1},$$

where  $1 \leq \zeta \leq \alpha + 1$  and

$$\begin{aligned} d_\zeta - E &= d_\zeta - E, \quad \text{if } (d_\zeta - E) \geq 0, \\ &= 10 + (d_\zeta - E), \quad \text{if } (d_\zeta - E) < 0. \end{aligned} \tag{13}$$

Now, one of these is the correct number and we are left with the task of selecting this; namely choose one of  $\{C_\zeta\}_{\zeta=1}^{\alpha+1}$ . We make use of the construction of  $d_{\alpha+2}$  for this purpose. We first find,

$$\begin{aligned} S_{C_\zeta} &= w_1 d_1 + \cdots + w_{\zeta-1} d_{\zeta-1} + w_\zeta (d_\zeta - E) + w_{\zeta+1} d_{\zeta+1} \\ &\quad + \cdots + w_{\eta-1} d_{\eta-1} + w_\eta D_\eta + w_{\eta+1} d_{\eta+1} + \cdots + w_{\alpha+1} d_{\alpha+1}, \end{aligned}$$

for  $1 \leq \zeta \leq \alpha + 1$  and find

$$P_{C_\zeta} = \text{modulo} (S_{C_\zeta}, 11). \tag{14}$$

The correct value of  $\eta$  is  $\delta$ , when

$$P_{C_\delta} = d_{\alpha+2}. \quad (15)$$

The scheme for generating and correcting single-errors is shown in algorithm 1.

1. **Given number:**  $N_\alpha$ , a  $\alpha$  digit number say  $d_1 \cdots d_\alpha$ .
2. **Generation:** Generate  $d_{\alpha+1}$  using (6) and  $d_{\alpha+2}$  using (7).  
 {the generated number to be spoken is  $N_\gamma$  (say,  $d_1 \cdots d_\alpha \cdots d_\gamma$ )}
3. **Spoken number:** The spoken number  $N_\gamma$  is recognized by the speech engine as  $d_1 \cdots d_{\eta-1} D_\eta d_{\eta+1} \cdots d_{\gamma-1} d_\gamma$   
 {A single-error at position  $\eta$  (for example)}.
4. **Single-error correction:** The error  $(d_\eta - D_\eta)$  is determined by (12), and the position of the error is determined by (15).

**Algorithm 1.** Scheme for generation and correction of single-errors.

*Note 3.* Note that the digit recognition error could occur in any of the  $\gamma$  digits. The proposed scheme does not impose (that the recognition error occur)  $1 \leq \eta \leq \alpha$  restriction. If the error occurs in the  $\gamma$  position then  $E$  in (10) returns 0 meaning no error. If there is an error in the  $(\alpha + 1)$ th position then  $E = \text{modulo}((D_{\alpha+1} - d_{\alpha+1}), 10)$  and (12) would be able to identify that the error was in the location  $\alpha + 1$ . Which means that the recognition error could occur in the range  $1 \leq \eta \leq \gamma$ .

#### 4.1 Uniqueness of $C_\delta$

We need to show that  $C_\delta$  that satisfies (15) is unique. This can be shown in two steps, namely, showing that the error ( $e$ ) and the location of the error are both uniquely determinable.

4.1a *Uniqueness of error identification:* We know that the error (11)  $e = (D_\eta - d_\eta) \in \{-9, 9\}$ . It follows that  $E = \text{modulo}(e, 10) \in \{0, 9\}$ . Now, one needs to show that, given the erroneously recognized digit sequence,  $d_1 \cdots D_\eta \cdots d_\gamma$  and  $E$ , one can uniquely determine  $d_\eta$  using (12).

We know that  $d \in \{0, 9\}$  (actual digit) and  $D \in \{0, 9\}$  (misrecognized digit). Consider the condition that the digit  $d = i \in \{0, 9\}$  is misrecognized as  $D = j \in \{0, 9\}$  by the speech recognition engine. We can *compute* the actual transmitted digit  $i'$  from (12) as

$$i' = j - E, \quad \text{if } j \geq E, \quad (16)$$

$$= 10 + (j - E), \quad \text{if } j < E, \quad (17)$$

where,  $E = \text{modulo}(j - i, 10)$ . We need to show that  $i'$  is indeed the actual  $i$ . There are two possible cases, namely,  $j - i \geq 0$  and  $j - i < 0$ . We look at them separately,

*Case  $j - i \geq 0$ :* We have  $E = j - i$  and hence we can estimate  $i' = j - (j - i) = i$ , using (16).



Case  $j-i < 0$ : We have  $E = 10+j-i$  and hence we can estimate  $i' = 10+j-(10+j-i) = i$ , using (16).

Clearly, we can obtain  $i$  the correct digit ( $d$ ) by estimating  $i'$  as above. This establishes the uniqueness of identification of error.

4.1b *Uniqueness of error location:* Once the error is uniquely determined, it remains to show that the location of the error is also uniquely determinable to establish that the scheme suggested is a valid scheme to correct single-digit errors. Given the error  $E$  (10), it is clear from (14) that

$$\begin{aligned}
 P_{C_\eta} &= \text{modulo}(S_{C_\eta}, 11) \\
 &= \text{modulo} \left( \sum_{i=1}^{\alpha+1} w_i d_i + w_\eta (D_\eta - E) - w_\eta d_\eta, 11 \right) \\
 &= \text{modulo} \left( \sum_{i=1}^{\alpha+1} w_i d_i + w_\eta (D_\eta - (D_\eta - d_\eta)) - w_\eta d_\eta, 11 \right) \\
 &= \text{modulo} \left( \sum_{i=1}^{\alpha+1} w_i d_i, 11 \right) \\
 &= d_{\alpha+2}.
 \end{aligned}$$

Hence, one is assured of being able to correct the occurrence of a single-error. Now, it remains to be shown that none of the other  $P_{C_j}$  for  $j \neq \eta$  is  $d_{\alpha+2}$ . To demonstrate this, consider

$$\begin{aligned}
 P_{C_\sigma} (\sigma \neq \eta) &= \text{modulo}(w_1 d_1 + \dots w_\eta D_\eta + \dots w_\sigma (d_\sigma - E) + \dots w_{\alpha+1} d_{\alpha+1}, 11) \\
 &= \text{modulo} \left( \sum_{i=1}^{\alpha+1} w_i d_i - w_\sigma (D_\eta - d_\eta) + w_\eta D_\eta - w_\eta d_\eta, 11 \right) \\
 &= \text{modulo} \left( \sum_{i=1}^{\alpha+1} w_i d_i + (w_\eta - w_\sigma)(D_\eta - d_\eta), 11 \right) \\
 &= d_{\alpha+2} \quad \text{if} \quad \begin{cases} \text{(i)} & D_\eta = d_\eta & \text{or} \\ \text{(ii)} & w_\eta = w_\sigma & \text{or} \\ \text{(iii)} & \text{modulo}(((w_\eta - w_\sigma)(D_\eta - d_\eta)), 11) = 0. \end{cases}
 \end{aligned}$$

None of the above three conditions is true

- We know that,  $D_\eta \neq d_\eta$ , because there exists an error<sup>4</sup> at location  $\eta$  (condition (i))
- $w_\eta \neq w_\sigma$  because of the choice of weights (8) (condition (ii))

<sup>4</sup> $D_\eta = d_\eta$  implies there is no error in the recognition of digits by the speech engine.

- We know that  $(D_\eta - d_\eta)$  is an integer in the range  $\{-9, 9\}$  and if we choose weights  $\{w_i\}_{i=1}^{\alpha+1}$  such that  $w_i - w_j \neq$  (any multiple of 11)  $\forall i, j \in \{1, \alpha + 1\}$  then condition (iii) is not true.

This shows that a single-error can be uniquely identified and corrected using the suggested scheme (algorithm 1).

#### 4.2 Advantage of using the proposed scheme

We can show that the proposed scheme enhances connected-digit recognition accuracy considerably. Consider an  $\alpha = 6$  digit number and assume that the digit recognition accuracy is  $p_{d1}^c = 0.9$ . The baseline connected-digit recognition (CDR) system would have a recognition accuracy of 53.14% ( $= (0.9)^6 = 0.53144$ , using (1)). Suppose we append  $\beta = 2$  digits to give the proposed scheme the ability to detect and correct  $k = 1$  digits in error. Then the  $\gamma = 8$  digit number would be recognised by the scheme proposed in the paper with an accuracy of 81.3% ( $= (0.9)^8 + {}^8C_1(0.9)^7(0.1)^1$  using (4)).

### 5. Conclusions

In this paper, we have proposed a simple yet novel scheme to increase the recognition accuracy of a connected-digit without actually increasing the recognition accuracy of the digit. We presented the scheme, which can be essentially classified as a non-prime, non-binary error-correcting code. The basic idea is to append extra digits and use them to detect and correct a single-digit error. We have shown that the error correction can *always* be done and that it is *unique*. This scheme can be used to enhance, spoken connected-digit recognition accuracy.

The authors thank the anonymous reviewer for the suggestions and comments, which have immensely helped in enhancing the content of the paper.

#### Appendix A. $p_{ad}^{\alpha \rightarrow \gamma}$ is positive for $p_{d1}^c > 1/(\gamma - 1)$ for $k = 1$ digit error correction

Observe that there is no advantage of using the error-correcting scheme when the RHS of (4) is  $\leq 0$ . Equating (4) to 0 and replacing  $p_{d1}^c$  by  $p$ , we have,

$$\begin{aligned} p^\gamma + \gamma p^{\gamma-1}(1-p) &= p^{\gamma-2}, \\ p^{\gamma-2} - p^\gamma - \gamma p^{\gamma-1} + \gamma p^\gamma &= 0, \\ (\gamma-1)p^\gamma - \gamma p^{\gamma-1} + p^{\gamma-2} &= 0. \end{aligned}$$

Dividing by  $p^{\gamma-2}$  (assuming that  $p > 0$ ) we have

$$(\gamma-1)p^2 - \gamma p + 1 = 0.$$

Solving the quadratic equation we get

$$p = 1, 1/(\gamma-1).$$

It can be easily verified that there is a positive advantage of using the error-correcting scheme for  $1/(\gamma-1) < p < 1$ .

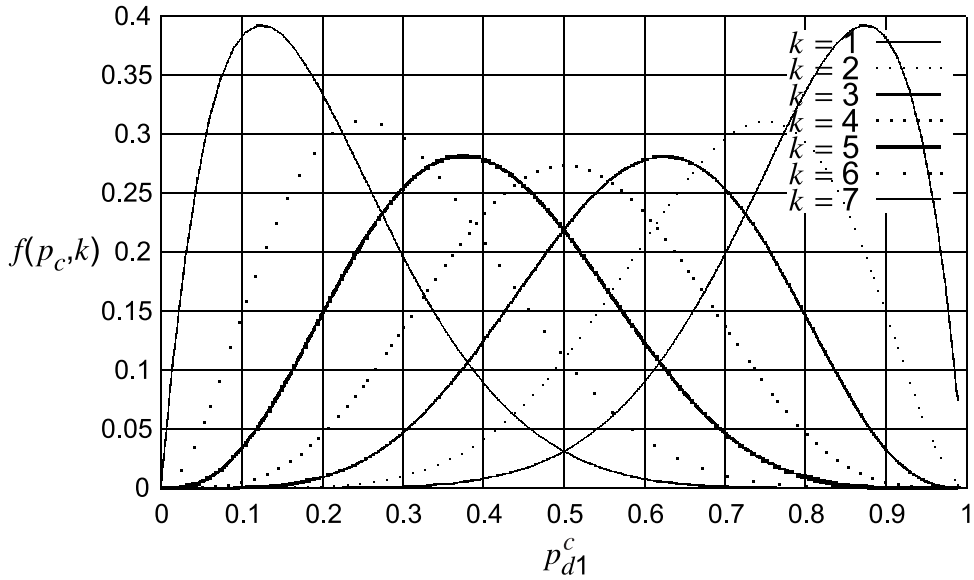


Figure B1. Advantage due to correction of exactly  $k$  digits in error.

**Appendix B. Recognition accuracy advantage**

B.1 Exactly  $k$  digits in error are corrected

Let  $p_c$  be the probability of correct recognition of a single-digit. Then from (2),

$$f(p_c, k) = {}^\gamma C_k p_c^{\gamma-k} (1 - p_c)^k, \tag{B1}$$

is the advantage of a scheme being able to detect and correct exactly  $k$  digits in error (see figure B1). We can find the extrema of  $f(p_c, k)$  by differentiating (B2) with respect to  $p_c$

$$\partial f(p_c, k) / \partial p_c = {}^\gamma C_k ((\gamma - k) p_c^{\gamma-k-1} (1 - p_c)^k - k p_c^{\gamma-k} (1 - p_c)^{k-1}),$$

and equating it to 0,

$$\begin{aligned} {}^\gamma C_k ((\gamma - k) p_c^{\gamma-k-1} (1 - p_c)^k - k p_c^{\gamma-k} (1 - p_c)^{k-1}) &= 0, \\ {}^\gamma C_k (p_c^{\gamma-k-1} (1 - p_c)^{k-1}) (\gamma(1 - p_c) - k) &= 0, \\ p_c &= ((\gamma - k) / \gamma). \end{aligned}$$

We can show that the extrema is a maximum by taking the second differential of (B1) and showing that it is negative for the extrema  $p_c = ((\gamma - k) / \gamma)$ .

$$\begin{aligned} \frac{\partial^2 f(p_c, k)}{\partial p_c^2} &= \frac{\partial}{\partial p_c} {}^\gamma C_k (p_c^{\gamma-k-1} (1 - p_c)^{k-1}) (\gamma(1 - p_c) - k) \\ &= {}^\gamma C_k \left( (\gamma(1 - p_c) - k) \frac{\partial}{\partial p_c} \{ p_c^{\gamma-k-1} (1 - p_c)^{k-1} \} \right. \\ &\quad \left. + p_c^{\gamma-k-1} (1 - p_c)^{k-1} (-\gamma) \right) \end{aligned}$$

$$\begin{aligned}
 &= {}^\gamma C_k (\gamma(1 - p_c) - k) \{(\gamma - k - 1)(1 - p_c)^{k-1} p_c^{\gamma-k-2} \\
 &\quad - p_c^{\gamma-k-1} (1 - p_c)^{k-2} (k - 1)\} - \gamma p_c^{\gamma-k-1} (1 - p_c)^{k-1} \\
 &= {}^\gamma C_k \{(1 - p_c)^{k-2} p_c^{\gamma-k-2} \{(\gamma(1 - p_c) - k)\{(\gamma - k - 1)(1 - p_c) \\
 &\quad + (k - 1)p_c\} - \gamma p_c(1 - p_c)\}\}.
 \end{aligned}$$

Substituting  $p_c = ((\gamma - k)/\gamma)$ , and observing that  $(1 - p_c) = (k/\gamma)$ , we have

$$\begin{aligned}
 \frac{\partial^2 f(p_c, k)}{\partial p_c^2} &= {}^\gamma C_k \left\{ \left(\frac{k}{\gamma}\right)^{k-2} \left(\frac{\gamma - k}{\gamma}\right)^{\gamma-k-2} \left\{ \overbrace{0}^{0} \right. \right. \\
 &\quad \left. \left. \left\{ (\gamma - k - 1) \left(\frac{j}{\gamma}\right) + (k - 1) \left(\frac{\gamma - k}{\gamma}\right) \right\} - \gamma \left(\frac{\gamma - k}{\gamma}\right) \left(\frac{k}{\gamma}\right) \right\} \right\} \\
 &= -{}^\gamma C_k \left\{ \left(\frac{k}{\gamma}\right)^{k-2} \left(\frac{\gamma - k}{\gamma}\right)^{\gamma-k-2} \gamma \left(\frac{\gamma - k}{\gamma}\right) \left(\frac{k}{\gamma}\right) \right\} \\
 &= -{}^\gamma C_k \left\{ \left(\frac{k}{\gamma}\right)^{k-1} \left(\frac{\gamma - k}{\gamma}\right)^{\gamma-k-1} \gamma \right\} \\
 &< 0.
 \end{aligned}$$

Hence, the extrema  $p_c = ((\gamma - k)/\gamma)$  is a maxima. The maximum of (B1) is

$$f_{\max}(p_c, k) = {}^\gamma C_k ((\gamma - k)/\gamma)^{\gamma-k} (k/\gamma)^k. \tag{B2}$$

**B.2  $k$  or less digits in error are corrected**

Let  $p_c$  be the probability of correct recognition of a single-digit. Then from (3)

$$f(p_c) = \sum_{j=1}^k {}^\gamma C_j p_c^{\gamma-j} (1 - p_c)^j, \tag{B3}$$

is the advantage of a scheme being able to detect and correct  $k$  digits (or less) in error (see figure B2. Figure B2 shows that  $f(p_c)$  has an extrema and is a function of  $k$  and  $\gamma$ . We can find the extrema of  $f(p_c)$  by differentiating (B3) with respect to  $p_c$ ,

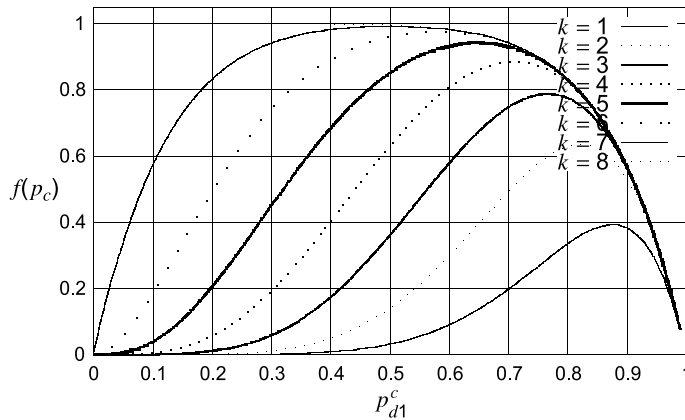
$$\frac{\partial f(p_c)}{\partial p_c} = \sum_{j=1}^k {}^\gamma C_j ((\gamma - j)p_c^{\gamma-j-1} (1 - p_c)^j - j p_c^{\gamma-j} p_c^{j-1}),$$

and equating it to 0,

$$\sum_{j=1}^k {}^\gamma C_j ((\gamma - j)p_c^{\gamma-j-1} (1 - p_c)^j - j p_c^{\gamma-j} (1 - p_c)^{j-1}) = 0,$$

$$\sum_{j=1}^k {}^\gamma C_j (p_c^{\gamma-j-1} (1 - p_c)^{j-1}) (\gamma(1 - p_c) - j) = 0,$$

$$g(\gamma, k) = 0,$$



**Figure B2.** Advantage due to correction of  $k$  digits or less in error (for  $\gamma = 8$ ).

where

$$\begin{aligned}
 g(\gamma, k) = & p_c^{\gamma-k-1} \{ \gamma (-^\gamma C_1 + ^\gamma C_2 - ^\gamma C_3 + \dots + ^\gamma C_k) p_c^k \\
 & + (\gamma - 1) (^\gamma C_1 - 2^\gamma C_2 + 3^\gamma C_3 + \dots + k^\gamma C_k) p_c^{k-1} \\
 & + (\gamma - 2) (^\gamma C_2 - 3^\gamma C_3 + 4^\gamma C_4 + \dots + k^\gamma C_k) p_c^{k-2} \\
 & + (\gamma - 3) (^\gamma C_3 - 4^\gamma C_4 + 5^\gamma C_5 + \dots + k^\gamma C_k) p_c^{k-3} \\
 & + \dots + (\gamma - k)^\gamma C_k \}.
 \end{aligned}$$

Observe that  $g(\gamma, k)$  is a polynomial of degree  $k$  in  $p_c$ . Unfortunately, there is no analytical solution possible, for a general  $k$ , although the plots in figure B2 show that there exists an optimal  $p_c$  which maximises  $f(p_c)$  for a given value of  $k$  and  $\gamma$ .

### References

Brown D 1973 Construction of error detection and error correction codes to any base. *Electron. Lett.* 9: 290

Buhrke E R, Cardin R, Normandin Y, Rahim M G, Wilpon J G 1994 Application of vector quantized hidden markov modeling to tele-phone network based connected digit recognition. In *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Process.* pp 105–108

Gumm H 1985 A new class of checkdigit methods for arbitrary number systems. *IEEE Trans. Inf. Theor.* 31: 102–105

Kopec G, Bush M 1985 Network-based connected-digit recognition. *IEEE Trans. Acoust., Speech, Signal Process.* 33: 1401–1413

Rosenquist T 2003 An Introduction to Error Correcting Codes. In <http://web.syr.edu/~rosenqu/ecc/main.htm>

Sethi A, Rajaraman V, Kenjale P 1978 An error-correcting coding system for alphanumeric data. *Inf. Process. Lett.* 7: 72–77

Wagner N, Putter P 1989 Error detecting decimal digits. *Commun. ACM* 32: 106–110

Wagner N R 2002 The laws of cryptography: Coping with decimal numbers. In <http://www.cs.utsa.edu/~wagner/laws/decimal.html>