OPEN ACCESS

AJSTSS
AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND SOCIAL SCIENCES

AFRICAN JOURNAL OF SCIENCE, TECHNOLOGY AND SOCIAL SCIENCES

Journal website: https://journals.must.ac.ke

MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

A Publication of Meru University of Science and Technology

Errors, error detection and correction efficiency in the container number code

Mutugi Praise Kiogora[1*], Josephine Mutembei[1] and Loyford Njagi[1]

[1]Meru University of Science and Technology, Meru, Kenya

ARTICLE INFO

KEYWORDS

Container Numbers

Error Detection

ABSTRACT

The container number is an 11-digit code that uses a single parity check of modulo 11 to check the authenticity of the container numbers. The world relies on shipment containers to transport goods. Various types of cargos have been erroneously delivered in different places. The modulo 11 container number is divided into four parts. The owner code that consists of three capital letters of the Latin alphabet to indicate the owner or principal operator of the container. Such code needs to be registered at the Bureau International des Conteneurs in Paris to ensure uniqueness worldwide (per ISO 6346). The equipment category identifier consists of one of the following three capital letters of the Latin alphabet U for all freight containers, J for detachable freight container-related equipment, Z for trailers and chassis. The serial number consists of six numeric digits, assigned by the owner or operator, uniquely identifying the container within that owner/operator's fleet. The check digit consists of one numeric digit providing a means of validating the recording and transmission accuracies of the owner code and serial number. The objective of this study is to determine the efficiency of the current modulo 11 container number in the error detection and correction. While the modulo 11 checksum calculation method is an effective way to detect errors in container numbers, it is not foolproof and errors can still occur. Modulo 11 container number code does not detect some transposition and substitution errors. The container number code and its check digit do not possess inherent error correction capabilities. This study therefore recommends a new modulo 13 container number code with higher error correction and detection capability.

Introduction

In today's interconnected world, where global trade relies heavily on shipping containers, the accuracy and reliability of container numbers are paramount. These numbers, serve as unique identifiers for each container, ensuring goods reach their intended destinations smoothly. Yet, in this digital era, where even minor errors can lead to significant disruptions, the efficiency of error detection and correction mechanisms in container codes becomes a critical concern. This paper delves into the mathematical intricacies of error

*Corresponding author: Mutugi Praise Kiogora Email: mutugi.kiogora1@gmail.com

https://doi.org/10.58506/ajstss.v2i2.166

detection and correction in container number codes, seeking to optimize their effectiveness. In doing so, it addresses the fundamental question: Can mathematical approaches enhance the security and efficiency of global trade by minimizing errors in container identification?

Error detection and correction techniques are used in various fields, including communication, storage, and computing, to ensure that data is transmitted or stored accurately without errors.

The most commonly used techniques for error detection include checksum, cyclic redundancy check (CRC), and parity bit. In the checksum technique, a mathematical calculation is performed on the data, and the result is sent along with the data. The receiver performs the same calculation and compares the result with the one received. If they match, the data is assumed to be correct. If they don't match, an error is detected. (W. W. Peterson, and D. T. Brown, 1961)

In the CRC technique, a checksum is generated based on the contents of the data frame. This checksum is sent along with the data frame and is checked by the receiver to detect any errors in the data. (W. W. Peterson, and D. T. Brown, 1961)

The parity bit technique adds an extra bit to the data, which is set to 1 or 0 depending on the number of 1s in the data. The receiver checks this parity bit to detect any errors in the data. (Raymond, H., 1986)

Shannon specifically, defined communication as the procedures by means of which one mechanism affects another mechanism.

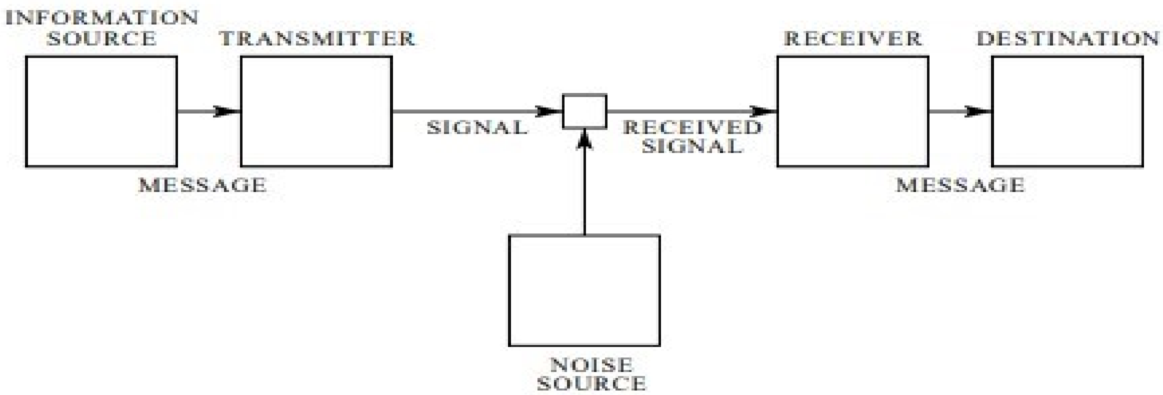

Figure 1: *A diagrammatic representation of a communication system*

The symmetric key cryptography is a good illustration of this communication system. Suppose two parties want to exchange a sensitive message. They agree on a specific key that is private to them that they will use to encrypt and decrypt their message, the sender uses the secret key to encrypt the data and the sends the message to the receiver where the receiver uses the agreed secret key to decrypt the message. If this key falls to the wrong party then they could distort the message, in this case this the Noise source. The symmetric cryptography is used in banks to authenticate ID and transactions and other institutions. (Matt Kerr, 2017)

Definitions

Correction efficiency is a measure of the effectiveness of a correction or error correction system in correcting errors or mistakes. It is usually expressed as a percentage, and it indicates the proportion of errors that have been corrected successfully.

Error detection is the process of identifying errors or mistakes in data or information. It involves examining data to check for any discrepancies, inconsistencies, or anomalies that may indicate errors or inaccuracies.

A *permutation* of a set A is a function from A to itself that is both one-to-one and onto. The following definition shows one way in which permutations are used in coding theory. (Sarah Spence, 2008)

If the encoder maps k-tuples of symbols from the message alphabet A in a one-to-one way to n-tuples of symbols from the input alphabet X of the channel (independent of the other input k-tuples), the resulting set of $|A|$ k output n-tuples is called *a block code*. For the elements of a block code one uses the name codeword. (Tilborg 1993.)

Example

{0000, 1111} is a code of length 4 with two code words whose digits come from the alphabet $F = \mathbb{Z}_2 = \{0, 1\}$ An error word e is detected by a code if a + e is not a code word for any code word a. If a + e is a code word, then e is undetected. (Michael Toymil, 2010

Example

Suppose the only messages we wish to send are "yes" and "No" and the encoder decodes a "yes" to "0" and a. "no" to "1"

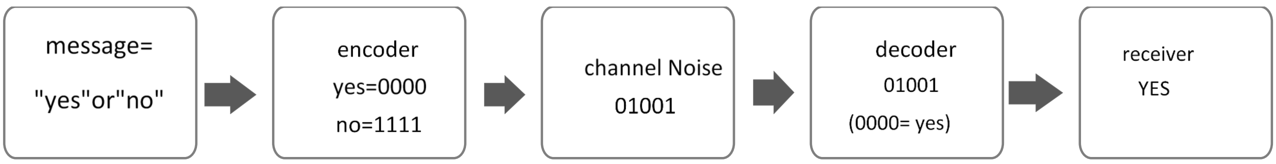

Figure 2: *Sample messages for decoding*

Here two errors occurred and the decoder has decoded the received vector "01001" as the nearest code word, which is "0000" or "yes"

The *hamming distance* between code words $u = (a_1, a_2 \dots a_n)$ and $v = (b_1, b_2, \dots b_n)$ is the number of places where they differ if the number of integers i for which $a_i \neq b_i$ and is denoted by d (u, v)
The distance of a code C denoted by d(c) or simply by d is the smallest of the distances between distinct code words. that is
$d = \min (d (u, v)) \mid u, v$ belongs to C, $u \neq v$.

An error is a deviation from accuracy or correctness which is commonly caused by noisy communication channels such as thermal noise, imperfections in equipment, human errors among others. If a code word u is transmitted through a noisy communication there is a possibility that a different code word v will be received instead of the original code word u hence presence of an error in the code word. Error detection is the identification of errors in a code word of which it may be discarded and request for retransmission made. Error correction is the detection of errors in a code word and reconstruction of the original error free data. There are several types of errors that occur during data entry (Gallian, J., 1991)

Identification codes are codes that assign numbers or symbols to objects, items or even human beings for easy identification. Error-detection identification codes are codes that add extra digits to the identification numbers by formulas or algorithms that allow detection of various types of errors such as single errors, transposition errors among others (Sutherland, D., 1999)

Methodology

How to calculate the check digit of the container number code modulo 11.

A container check digit-calculator is used to determine if the 11 digits in the code are correct using the check digit. The container code format is divided into 4 parts. The owner code consists of three capital letters of the Latin alphabet to indicate the owner or principal operator of the container. Such code needs to be registered at the Bureau International des Conteneurs in Paris to ensure uniqueness worldwide. The equipment category identifier consists of one of the following three capital letters of the Latin alphabet: U for all freight containers, J for detachable freight container-related equipment and Z for trailers and chassis. The serial number consists of 6 numeric digits, assigned by the owner or operator, uniquely identifying the container within that owner/operator's fleet. The check digit consists of one numeric digit providing a means of validating the recording and transmission accuracies of the owner code and serial number. An example of the container number is MRKU9530406. (ISO 6346:1995).

Step one

An equivalent numerical value is assigned to each letter of the alphabet, beginning with 10 for the letter A (11 and multiples thereof are omitted): (ISO 6346:1995)

The individual digits of the serial number keep their numeric value.

A	B	C	D	E	F	G	H	I	J	K	L	M
10	12	13	14	15	16	17	18	19	20	21	23	24
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
25	26	27	28	29	30	31	32	34	35	36	37	38

Table 1:

Step two

Each of the numbers calculated in step 1 is multiplied by 2position, where position is the exponent to basis 2. Position starts at 0, from left to right. (ISO 6346:1995)

1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
0	1	2	3	4	5	6	7	8	9
2	2	2	2	2	2	2	2	2	2
1	2	4	8	16	32	64	128	256	512

Table 2

Step three

Sum up all results of step 2

Divide them by 11

Erase all decimal digits of the division (i. e. make the result an integer value)

Multiply the integer value by 11

Subtract result of d) from result of a): This is the check digit! If the final difference is 10, then the check digit becomes 0.

Example:

M	R	K	U	9	5	3	0	4	0	6
24	29	21	32	9	5	3	0	4	0	
1	2	4	8	16	32	64	128	256	512	
24	58	84	256	144	160	192	0	1024	0	
						SUM			1942	
						MOD (1942, 11)			6	

Table 3
(ISO 6346:1995)

Results and findings

Strengths of the container number code modulo 11

The implementation of a check digit algorithm in container numbers offers several advantages that contribute to the reliability and efficiency of container tracking and management systems. Firstly, the algorithm is designed to yield a high probability of error detection, effectively identifying errors such as transpositions, substitutions, or missing digits within the container number. This mathematical approach ensures a robust mechanism for error identification. (Crowley 2018).

Furthermore, the simplicity and efficiency of the check digit algorithm enhance its practicality. Calculations can be swiftly performed without requiring complex computations or additional data, making it a straightforward and efficient process to integrate into various systems (Crowley 2018). This simplicity contributes to the quick verification of container numbers.

In terms of storage requirements, the check digit is a single digit appended to the end of the container number, resulting in minimal impact on storage or transmission size. This aspect facilitates efficient data management and communication without significantly increasing the overhead (Crowley 2018).

The adoption of standardized container number formats, as outlined by ISO 6346, adds another layer of reliability. This standardization ensures consistency across different shipping lines and countries, promoting interoperability and simplifying error detection and correction procedures. Universally applying the same rules facilitates a streamlined and standardized approach to container number management (ISO 6346:1995).

The widespread adoption of container number codes with check digits in the shipping industry is a testament to its effectiveness. This universal acceptance ensures that error detection and correction capabilities remain consistent across various organizations and systems involved in container tracking and management (Crowley 2018).

Lastly, the quick verification process is a notable feature of the check digit system. The simple algorithm allows for rapid calculation and verification, requiring minimal computational resources. This characteristic makes it well-suited for real-time validation or batch processing scenarios, further enhancing its practical utility in diverse operational contexts.

Weaknesses Of The Container Number Code Modulo 11

The check digit calculation method used in container numbers is generally effective in detecting errors or discrepancies in the number. It serves as a simple and quick way to verify the accuracy of the container number by performing a mathematical calculation.

The effectiveness of the check digit method lies in its ability to detect certain types of errors, such as single-digit substitutions, transpositions, or missed digits. If any of these errors occur in the container number, the calculated check digit will not match the actual check digit in the number. By comparing the calculated check digit with the provided check digit, it is possible to identify whether an error has been introduced during data entry, transmission, or storage. If the check digit does not match, it suggests that there is a potential error in the container number.

The container number code and its check digit do not possess inherent error correction capabilities. The check digit is primarily designed for error detection, not correction. While the checksum calculation method is an effective way to detect errors in container numbers, it is not foolproof and errors can still occur. Below is a list of errors this paper highlights that the current modulo 11 container number code doesn't detect. (ISO 6346:1995)

Transposition error

Definition: A transposition error is a specific type of error that occurs when two adjacent characters or digits are swapped or reversed in a sequence.

Example: Suppose we have a jump twin error where SUDU3070079 interchange to SUDU0770309 we prove that the check digit calculator does not detect jump twin errors.

S	U	D	U	0	7	7	0	3	0	9
30	32	14	32	0	7	7	0	3	0	
1	2	4	8	16	32	64	128	256	512	
30	64	56	256	0	224	448	0	768	0	
						SUM			1846	
						MOD (1846, 11)			9	

Table 4

The sum of this numbers: = 1846
1846 mod 11 = 9
This completes our proof.

The modulo 11 code is designed to detect certain types of errors, such as single-digit substitutions, some transpositions, or missed digits, but it may not be sensitive to all types of transposition errors. the specific pattern of transposition results in a sum (1846) and modulo 11 operation (9) that align with a valid check digit, of the original sum (4486) and MOD (4486, 11) is 9, thus failing to flag the transposition error.

Single Substitution Errors.

Definition: This is when a single digit or the whole code is written wrongly. (Lin Shu and Daniel J. Costello, 2011)

Example: Suppose we have an error where MRKU9530406 interchange to MRUU9530406

M	R	U	U	9	5	3	0	4	0	6
24	29	32	32	9	5	3	0	4	0	
1	2	4	8	16	32	64	128	256	512	
24	58	128	256	144	160	192	0	1024	0	
						SUM			1986	
						MOD (1986, 11)			6	

Table 5

The sum of this numbers: = 1986
mod 11 = 6
The check sum digit didn't detect the error on this,

The error was not detected because the sum of the original code and the erroneous code was increased by a multiple of 11. In modular arithmetic, if two numbers differ by a multiple of the modulus (in this case, 11), their remainders after division by that modulus will be the same. Since 1986 and 1986 + 11 (or any multiple of 11) have the same remainder when divided by 11, the modulo 11 operation fails to distinguish between the original and erroneous codes, resulting in the error going undetected. (Burton, D. M. 2011).

Double Substitution Errors.
Definition: This is when two digits are written wrongly. (Raymond, H. 1986)

Example: Suppose we have an error where MRKU9530406 interchange to MHUU9530406

M	H	U	U	9	5	3	0	4	0	6
24	18	32	32	9	5	3	0	4	0	
1	2	4	8	16	32	64	128	256	512	
24	36	128	256	144	160	192	0	1024	0	
						SUM			1964	
						MOD (1964, 11)			6	

Table 6

The sum of this numbers: = 1964 mod 11 = 6
The check sum digit didn't detect the error on this.

The modular arithmetic property holds true in this case as well. If two numbers differ by a multiple of the modulus (11 in this case), their remainders after division by that modulus will be the same. Consequently, the modulo 11 operation fails to distinguish between the original and erroneous codes, leading to the error going undetected.

This underscores the importance of choosing a suitable checksum method and modulus that effectively detect the types of errors relevant to the application. (Burton, D. M. 2011).

Both substitution and transposition errors.

Definition: This is when a single digit or the whole code is written wrongly and another two adjacent characters or digits are swapped or reversed in a sequence. (Raymond, H. 1986)

Example: Suppose we have an error where MRKU9530406 interchange to WHAU9530406

W	H	A	U	9	5	3	0	4	0	6
35	18	10	32	9	5	3	0	4	0	
1	2	4	8	16	32	64	128	256	512	
35	36	40	256	144	160	192	0	1024	0	
						SUM			1887	
						MOD (1887, 11)			6	

Table 7

The sum of this numbers: = 1887 mod 11 = 6

The check sum digit didn't detect the error on this.

In this case, there are multiple errors, involving both substitution and transposition. The differences between the original and typed code are as follows: The letter "M" has been substituted with "W."

The letter "R" has been transposed with "H."

The letter "K" has been substituted with "A."

Therefore, the errors in typing "MRKU9530406" as "WHAU9530406" involve two substitutions (M to W, K to A) and one transposition (R and H).

Similar to the previous examples, the reason for the error not being detected lies in the modular arithmetic property. When a single digit or the entire code is substituted (substitution error), and two adjacent characters or digits are swapped or reversed (transposition error), the resulting sum can still align with the original modulo 11 result if the difference between the original and erroneous codes is a multiple of 11. In this specific case, the sum of the original code (1942) and the erroneous code (1887) is congruent to 6 modulo 11. Since both the original and erroneous codes yield the same remainder when divided by 11, the check sum digit fails to identify the presence of the error. This illustrates the challenges in designing a checksum method that is sensitive to a combination of substitution and transposition errors. (Burton, D. M. 2011).

Container Number Code Modulo 11 Dictionary Size

The container number is an 11 digits code. The container code format is divided into 4 parts. The owner code that consists of 4 alphabets the equipment category consists of one digit, the serial number consists of 6 numbers and the check digit. We use the permutation formula with repetition because the container code the order of the digits of the code have an impact on the final digit code.

The modulo 11 dictionary size

Since there is a 4-digit alphabet, with the fourth alphabet being picked from just 3 alphabets that is U, J, and Z, The Size of the code is calculated in 3 parts.

First, the Owner Code, where n is 26 and r is 3 since we choosing for the first 3 alphabets, MRK. Then n is 26 that is 26 alphabets and r is 3 the number of alphabets. With repetition.

$n^r = 26^3$

= 17,576

Second, the equipment category, this is one digit choose from 3 alphabets that is U, J and Z.

= 3C1

= 3

For the serial number n =10 and r=6

nPr(with repetition) = $\boldsymbol{n}^r$

$=10^6$

=1,000,000

Therefore, the combination of the entire serial number is the product of the combination of the alphabets and that of the serial number.

=17,576*1,000,000*3

= 52,728,000,000

Conclusion

The primary aim was to assess the efficiency of the container number code in its pivotal role of error detection and correction. This evaluation centered on the existing code's deployment of the modulo 11 code, which serves as the bedrock for identifying potential discrepancies within the 11-digit container code. This code structure is segmented into four essential components: the owner code, equipment category, serial number, and the check digit. The dynamic behind the error detection and correction process lies within the complex mechanism of the modulo 11 code, which computes the check digit to validate the accuracy of the rest of the digits in the code. The strengths of this approach are manifold: High Probability of Error Detection, Simplicity and Efficiency, Minimal Impact on Storage, Standardization and Adoption and Widespread Use. However, the existing code is not without limitations. While adept at detecting specific errors such as substitutions, transpositions, or missing digits, it falls short in identifying all potential errors that might occur. The comprehensive analysis of the modulo 11 container number code for error detection and correction validated its effectiveness in most scenarios, while also highlighting its inherent limitations in certain error types.

The check digit calculation method used in container numbers uses the single parity check digit and a single parity check digit only detects and corrects no errors therefore some transposition and substitution errors go undetected. As observed, there is a need to design and develop a new algorithm that improves the error detection capability of the container number code. Such a code is likely to use a double verification system in order to overcome the challenges in error detection of the container number.

References

Ali, S., & Raza, S. (2014). Standardizing the international identification system for containers.

Augot, D., Betti, E., & Orsini, E. (2009). An Introduction to Linear and Cyclic Codes.

Bowman, J. (2010). Coding theory and cryptography. Edmonton: University of Alberta. Burton, D. M. (2011). Elementary Number Theory (7th ed.). McGraw-Hill.

Causley, B. (2012). The secret behind the Luhn-ie. XRDS, 19(1), 1–2.

Celko, J. (2010). Joe Celko's data measurements and standards in SQL. Morgan Kaufmann Publishers.

Child's, L. (1979). A concrete introduction to higher algebra. New York: Springer Science and Business Media.

Container-xchange.com. (2022). Container numbers. Retrieved from https://www.containerxchange.com/blog/containernumber/#:~:text=A%20container%20number%20is%20a,International%20d es%20Containers%20(BIC)

Crowley. (2018). Top Five Advantages of Ocean Freight Shipping. Retrieved from
https://blog.crowley.com/advantages-of-ocean-shipping

Custom News - Specialized Page of Customs E-Magazine (Vu Thi Anh Hong, 2022).

El-Sharkawy, A. M., Elhoseny, M., & Abdelnaby, M. A. (2020). A secure container tracking system based on blockchain and IoT technologies. Journal of Ambient Intelligence and Humanized Computing, 11(4), 1689-1704.

Gallian, J. (1991). The mathematics of identification numbers. The College of Mathematics Journal, 22(3),
194–202.

Gallian, J. A. (1991). The Mathematics of Identification Numbers. College Mathematics Journal, 22(3), 194.

Ghandhi, K. (2015). Check digits. Retrieved from http://cosmos.ucdavis.edu/archives/2010/cluster6/ghandhikira.HTML. 6/8/2015.

Gupta, V., & Verman, C. (2012). Error detection and correction: An introduction. International Journal of
Advanced Research in Computer Science and Software Engineering, 2(11), 212–218.

GVCT. (2011). How is the check digit of a container calculated? Retrieved from http://www.gvct.co.uk/2011/09/how-is-the-check-digit-of-a-container-calculated/

Hamming, R. W. (1950). Error detecting and error correcting codes. Bell System Technical Journal, 2, 147–160. Hodge, J., Schlicker, S., & Sundstrom, T. (2013). Abstract algebra: An inquiry-based approach. Boca Raton: Chapman and Hall/ CRC Press.

International Organization for Standardization. (1995). Freight containers - Coding, identification and marking (ISO 6346: 1995).

Kamaku, W. (2012). Error detection and correction on the international standard book number. Unpublished PhD thesis: Jomo Kenyatta University of Agriculture and Technology.

Kamaku, W., Mwathi, C., & Kivunge, B. (n.d.). Limitations in the Convectional ISBN-10 Code.

Kumano, S., Miyamoto, K., Tamagawa, M., Ikeda, H., & Kan, K. (2004). Development of a container identification mark recognition system.

Kumar, K., & Kaur, P. (2015). Vulnerability detection of international mobile equipment identity number of smartphone and automated reporting of changed IMEI number.

Lang, S. (1971). Basic Mathematics. Addison-Wesley Publishing Company.

Lin, S., & Costello, D. J. (1983). Error Control Coding: Fundamentals and Applications.

Matt Kerr. (2019). Number Theory and Cryptography lecture notes.

Michael Toymil. (2010). Algebraic Coding Theory University of Puget Sound Math 434, Spring 2010. Moon, T. K., & Stirling, W. C. (2000). Mathematical Methods and Algorithms for Signal Processing.

Nyagah Machariaa, Waweru Kamaku, Joy Mutegic, Sitawa Wattanga. (2019) Modulo 10 Error Detection Efficiency Analysis for Bank Routing Number. Multimedia University of Kenya.

Nyquist, H. (1924). Certain Factors Affecting Telegraph Speed. Bell System Technical Journal, April 1924.

Peterson, W. W., & Brown, D. T. (1961). Cyclic codes for error detection.

Qiao, Y., & Li, Y. (2016). A new algorithm for auto-

matic container number recognition.
Raj, R., Singla, M., & Singh, D. (2021). An efficient security framework for containerized cargo based on blockchain technology. International Journal of Logistics Systems and Management, 38(4), 491-509.
Raymond, H. (1986). A first Course in Coding Theory. CLARENDON Press, U.S.A.
Rosen, K. H. (2012). Discrete Mathematics and Its Applications (7th ed.).
Discrete-Mathematics-and-Its-Applications-7Th-EditionRosen-1
Shannon, C. E. (1948). A Mathematical Theory of Communication. The Bell System Technical Journal.
Sun, H., Zhang, Y., & Zheng, X. (2021). A lightweight and efficient container tracking scheme based on dynamic clustering. IEEE Transactions on Industrial Informatics, 17(10), 7226-7236.
Sutherland, D. (1990). Error-detecting identification code for algebra students. School Science and Mathematics,
90(4), 283–290.
The Vancouver Sun. (2023). CBSA seizes record methamphetamine load headed down under. Retrieved from https://www.timesofisrael.com/idf-intercepts-major-iranian-weapons-shipment-to-gaza/ (2014) Tilborg, H. C. A. van. (1993). Coding Theory: A First Course.
Tsang, S. S., & Zhang, X. (2022). Research on the influence factors of container number recognition accuracy based on convolutional neural network. Journal of Marine Science and Technology, 30(2), 245-256.
Ugrinovi , M., & Babi , D. (2023). The role of smart contracts in the container shipping industry. Maritime Policy & & Management, 50(1), 114-128.
UNDP. (2008). UNDP-Shipping-Guide - Shipping and Incoterms.
United Nations Office on Drugs and Crime (UNODC). (2023). World Drug Report 2023. Retrieved from https://www.unodc.org/unodc/en/data-and-analysis/world-drug-report-2023.html
UNODC. (2021). UNODC Toolkit for mainstreaming Human Rights and Gender Equality.
Wachira, W., Waweru, K., & Nyaga, L. (n.d.). Transposition Error Detection in Luhn's Algorithm. Walton, J. (2007). Error Detection Based on Check Digit Schemes.
Wicker, S. B. (1995). Error Control Systems for Digital Communication and Storage.