Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2018) 26: 582 – 592

doi:10.3906/elk-1704-263

TEMPEST font counteracting a noninvasive acquisition of text data

Ireneusz KUBIAK*
Military Communication Institute, Zegrze, Poland

Received: 23.04.2017 • **Accepted/Published Online:** 14.11.2017 • **Final Version:** 26.01.2018

Abstract: The protection of information against electromagnetic penetration has huge significance. Different solutions, technological or organizational, are used to protect the information-limiting levels of unintentional emissions. In particular, technological solutions are implemented in commercial devices to limit levels of electromagnetic emissions. Nevertheless, such solutions cannot always be used. This is connected with the construction of these devices. Very often the devices do not have enough space on the inside to install new elements such as filters, electromagnetic shielding, and others. In the paper a new solution is proposed. The solution does not change the construction of the devices (e.g., printers, screens). The new method is based on computer fonts called TEMPEST fonts (safe fonts). In contrast to traditional fonts (e.g., Arial or Times New Roman), the new fonts are devoid of distinctive features. Without these features the characters of the new fonts are similar each other. At the output of the information infiltration channel the characters of the TEMPEST fonts limit the possibilities of recognition of characters that appear on the reconstructed image. The image is reconstructed based on an electromagnetic emission, which is correlated with the process information. In this paper the dominance of the TEMPEST fonts over the traditional fonts and Sang Mun fonts in the protection of information against electromagnetic penetration is demonstrated. The TEMPEST fonts have wider use. They are also resistant to optical character recognition programs. The use of the TEMPEST fonts in special systems allows the safe processing of data while displaying them on the screen as well as printing them on laser printers, especially on laser printers with LED slats.

Key words: Electromagnetic emission, safe computer font, protection of information, electromagnetic eavesdropping, image and signal processing

1. Introduction: the state of the art

Data protection is an important element of each institution. Various solutions are used to prevent the risk of loss of such data. Limited access to office space is the most frequently used method (organizational methods [1,2]). Access to the Internet is limited in the case of computer workstations. However, electronically processed information could be intercepted in a noninvasive way. The process of noninvasive data acquisition does not leave any traces [3–6]. You do not find out about it until the information is disclosed.

Electromagnetic emission correlated with the processed data has to occur in the case of the mentioned phenomenon. The most important case in this regard consists of the possibility of intercepting data and presenting them in a form that is comprehensible and legible for a person. In particular, this concerns emission sources that process data in a graphic form. Such sources are VGA and DVI standard transmission lines [7–10]. The laser systems of typical laser printers could also be such sources of emission [11–14].

*Correspondence: i.kubiak@wil.waw.pl

Given the possibility of acquisition of presented data (graphic data) in a noninvasive way, the sources of the electromagnetic emission correlated with the data are subject to particular analysis and protection. The best and most versatile solutions are sought for all sources of revealing emissions. One of the most effective methods consists of shielding every element of IT equipment that could be a potential source of sensitive emission [2]. However, given the purchase price and appearance of such equipment, other more user-friendly solutions are sought. One of the suggested methods consists of filtering video signals (Figure 1 [15,16]).

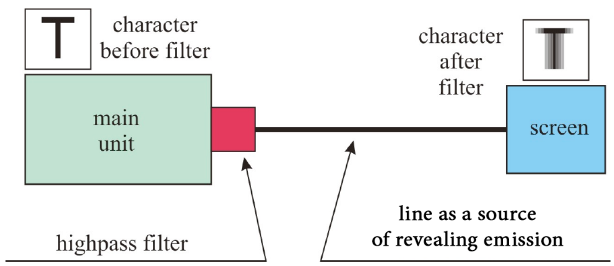

Figure 1. Filtering of the video signal – VGA standard.

Markus Kuhn suggested this approach to the problem [15,16]. He brings forward the concept of using a filtering system assembled on the output video port (VGA standard) of, e.g., a computer central unit. Of course, this produces the desired effect. Kuhn limits over 50% of signal spectrum. Given the rejection of higher components of the spectrum of the measured sensitive emission signal, it is practically not measurable. In this case the observed letters on an "eavesdropped" monitor are significantly blurred. It is tiresome to read the signs. However, a very important aspect should be noted here. Such a solution is not versatile. It could be used for analogue systems such as the VGA standard. Moreover, a change of resolution of the displayed image results in changed parameters of the vision signal. Using the same filter for all image resolutions is not going to be effective. In this case the use of an adaptive filter is necessary, for every resolution of the displayed image.

A digital standard such as DVI has other characteristics that make it impossible to introduce significant changes to the spectrum of the signal. In this case false pixels could occur and a monitor could switch off in extreme cases. Therefore, the desired effect is not reached. In the case of the DVI standard a combination of colors of letters and backgrounds could be used for a higher level of electromagnetic protection of the processed data [7,15]. The USB standard offers similar properties. It is used for work with, e.g., a laser printer. In this case a solution in the form of signal filtering cannot be used [2] to prevent electromagnetic infiltration of the printing device.

Sang Mun suggests another solution [17,18]. His method is based on a computer font, in which individual letters and digits are enriched with additional elements (e.g., noise; Figure 2). The Arial font is the basis for the set of fonts. It is easily identified in the electromagnetic infiltration process. The primary purpose of Sang Mun fonts is to prevent the effectiveness of optical character recognition (OCR) programs. The possibility of using fonts in electromagnetic protection of processed text data was verified. In this case, the method of correlation between signs at the output of a radiated type of information infiltration channel (IIC) was used. It turns out that for very weak signals and poor legibility of presented signs the correlation method is likely to indicate correct signs.

The Null Pointer font is another solution that could be deemed effective in counteracting electromagnetic infiltration (Figure 3 [19]). It is a generally available commercial font. It has some characteristics of a solution described in the article. Individual characters have vertical and horizontal lines with various widths. However,

Figure 2. The characters of Sang Mun Noise font.

the structure of characters shows no consistency. This primarily relates to the diversity of line widths that allow identification of a given sign easily. The correlation method used in the analysis allows to clearly identify the sought characters in a presented image.

Figure 3. The characters of the Null Pointer font.

The electromagnetic protection of processed text information based on safe fonts [17,20] was also proposed. This is a highly versatile solution that protects text data against electromagnetic infiltration regardless of the source of the sensitive emission: DVI digital standard, VGA analogue standard, single or double diode system of laser printers, or printers with a LED slat [14]. Moreover, safe fonts prevent use of OCR programs. Direct transfer from paper version to electronic version is practically impossible.

2. The essence of safe computer fonts

The design of safe fonts takes into account the characteristics of a radiated-type IIC. This is a high-pass filter described by the following relation:

$$x'\left(t\right) = \lim_{\Delta t \to 0} \frac{x\left(t + \Delta t\right) - x\left(t\right)}{\Delta t} + s(t), \tag{1}$$

where $x'(t)$ is signal $x(t)$ on the output of the IIC, $x(t)$ is the signal of the revealing emission on the input of the IIC, and $s(t)$ is noise and disturbances coming from natural sources and from sources built by man.

Therefore, presented images based on revealing emission signals include only vertical and diagonal edges of graphic elements. They correspond to beginnings and ends of electrical impulses that decide about displaying appropriate shapes of signs on a monitor. Taking this characteristic into account, a safe font called Asymmetrical Safe was proposed, with shape shown in Figure 4.

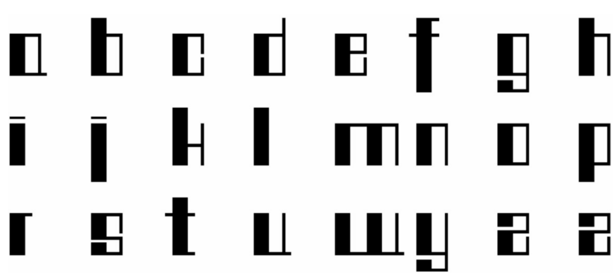

Figure 4. The characters of Asymmetrical Safe font.

The design of a safe font is based on the concept of maximizing similarity between individual characters. Lack of certain elements of structural characters makes it likely to be recognized as a different alphabetic character. Distinctive features of characters at IIC input might disappear at its output. Therefore, a given character might be recognized as a different one. Such phenomenon exists for each source of sensitive emission identified as a source of emission processing data in a graphic form (screen – VGA (DVI) standard, laser printer (one and two-diode laser system and LED slat)). This was the purpose of developing a font that would allow for secure processing of text data from the electromagnetic infiltration effectiveness point of view.

Each character of letters and digits of the font is built of various widths of vertical and horizontal lines. Some basic characteristics of two-element writing have been retained:

- a narrower vertical line exists only on the right side of the characters (vertical line) and in their upper and lower parts (horizontal lines);
- a wide vertical line exists only on the left side of a character;
- none of the characters bear any decorative elements (Figure 5).

This allows standardization of the font characters while maintaining sufficient legibility.

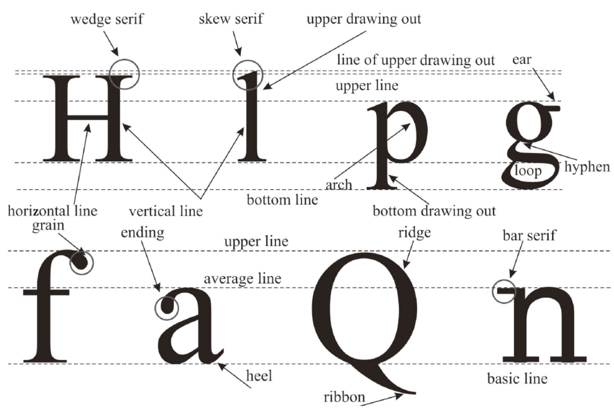

Figure 5. Letter elements.

Very high-class receivers are used in tests of TEMPEST devices (screens and printers). One of them is the DSI-1550A measurement system (Dynamic Sciences Corporation, range of frequencies: 10 kHz to 22 GHz, BW = up to 200 MHz; Figure 6). The video output of the receiver allows recording of the revealing emission signals. This is, however, an absolute value of the signal. That is why presented images show positive impulses marking the mentioned vertical and diagonal edges of characters, pursuant to the following relation [2]:

$$x''(t) = \left| \lim_{\Delta t \to 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} + s(t) \right|. \tag{2}$$

During the tests the distance between the receiving antenna and the source (monitor or laser printer) of revealing emissions was equal to 1 m in accordance with the MIL-STD-461F "Requirements for the control of electromagnetic interference. Characteristics of subsystems and equipment", in an anechoic chamber. In this way, the source of emission was separated from additional, unwanted sources of electromagnetic disturbances. The used computer was of the TEMPEST class. The sensitive emissions were measured at frequencies of 525 MHz and 685 MHz and for the measuring band BW = 20 MHz. Signal sampling frequency was equal to 15 Ms/s. The displayed and printed document contained signs written in the Arial font. In order to establish

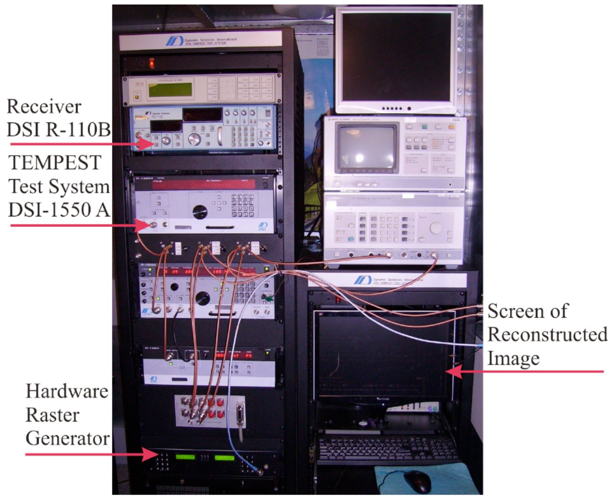

Figure 6. TEMPEST test system R-1550A.

the suitability of safe fonts in the protection of processed electromagnetic information, the research was also conducted for signs written in the Asymmetrical Safe font.

The design of individual characters was based on the concept of making them as similar to each other as possible while also making them distinguishable, which allows for smooth reading. Moreover, widths of the characters should not be different from widths of characters in traditional fonts: Arial and Times New Roman. This guarantees compliance with the basic statistical data concerning the number of characters with spaces per sheet of paper [17].

The tests of revealing emission during text data processing were carried out conducted in ideal measurement conditions (inside the anechoic chamber; Figure 7). The real conditions improve the level of electromagnetic protection of processed data.

Figure 7. Measuring system inside of the anechoic chamber.

3. Character correlation coefficient of computer font characters

The usefulness of special computer fonts in the processing of information and counteraction of electromagnetic infiltration was assessed based on visual analysis. This results from the fact that the graphic line of the device is a source of sensitive emission signals. The visual analysis can be supported by a correlation analysis and character error rate (CER) values estimated based on a character correlation coefficient [21]. The character correlation coefficient measured for individual characters is calculated for appropriate fragments of the analyzed image including the searched characters and the pattern image including the searched character, pursuant to the following relation:

$$R^{z}_{j,i} = \frac{\sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} (x_{n+j\cdot N_w, m+i\cdot N_w} - \bar{x}_{j,i}) \times (y_{n,m} - \bar{y})}{\sqrt{\sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} (x_{n+j\cdot N_w, m+i\cdot M_w} - \bar{x}_{j,i})^2 \times \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} (y_{n,m} - \bar{y})^2}}, \tag{3}$$

where:

$$\bar{x}_{j,i} = \frac{1}{N_w \cdot M_w} \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} x_{n+j\cdot N_w, m+i\cdot M_w}, \tag{4}$$

$$\bar{y} = \frac{1}{N_w \cdot M_w} \sum_{n=0}^{N_w-1} \sum_{m=0}^{M_w-1} y_{n,m}, \tag{5}$$

and

$0 \le j \le b-1$; $0 \le i \le d-1$; $b = N_a/\ N_w$ – number of rows of correlation image; $d = M_a/\ M_w$– number of columns of correlation image; $x_{n,m}$– pixel amplitude of analyzed image; $y_{n,m}$– pixel amplitude of pattern image (single character of computer font); M_a– number of columns of analyzed image; N_a– number of rows of analyzed image; M_w– number of columns of pattern image; N_w– number of rows of pattern image; i– column number of correlation image; m– column number of pattern image; j– row number of correlation image; n– row number of pattern image.

Eq. (3) is correct for the character correlation calculated at both the input and output of the IIC. High R^Z values (taking into account sufficient recognizability of characters and their legibility at the input of the IIC) make distinctive features of given characters partly disappear at the output of the IIC. Consequently, visually recognizing characters at the output of the IIC becomes impossible and R^Z values do not allow to correctly mark a character in a location where it exists. In this case, there are high CER values.

The character correlation coefficient values calculated according to Eq. (3) for the analyzed fonts are shown in Tables 1–3 and Figure 8.

Table 1. The number of relationships between the font characters on the output of the IIC for which the values of correlation coefficients R^Z are included in the suitable ranges of similarity – laser system of laser printer as a source of sensitive emissions.

Value of character correlation coefficient R^Z	Number of relationships				
	Asymmetrical Safe	Arial	Times New Roman	Sang Mun Noise	Null Pointer
$0.0 \le R^Z < 0.2$	0	44	46	119	52
$0.2 \le R^Z < 0.4$	18	94	107	122	190
$0.4 \le R^Z < 0.7$	157	110	97	12	11
$0.7 \le R^Z < 0.9$	78	5	3	0	0
$0.9 \le R^Z \le 1.0$	0	0	0	0	0

Table 2. The number of relationships between the font characters on the output of the IIC for which the values of character correlation coefficients R^Z are included in the appropriate ranges of similarity – VGA line as a source of sensitive emissions.

Value of character correlation coefficient R^Z	Number of relationships				
	Asymmetrical Safe	Arial	Times New Roman	Sang Mun Noise	Null Pointer
$0.0 \leq R^Z < 0.2$ [1)]	0	125	126	191	0
$0.2 \leq R^Z < 0.4$ [2)]	6	77	84	57	67
$0.4 \leq R^Z < 0.7$ [3)]	125	48	43	5	155
$0.7 \leq R^Z < 0.9$ [4)]	110	3	0	0	29
$0.9 \leq R^Z \leq 1.0$ [5)]	12	0	0	0	2

[1)]Lack of relationship, [2)]weak relationship, [3)]moderate relationship, [4)] rather strong relationship, [5)]very strong relationship.

Table 3. The number of relationships between the font characters on the output of the IIC for which the values of correlation coefficients R^Z are included in the suitable ranges of similarity – DVI line as a source of sensitive emissions.

Value of character correlation coefficient R^Z	Number of relationships				
	Asymmetrical Safe	Arial	Times New Roman	Sang Mun Noise	Null Pointer
$0.0 \leq R^Z < 0.2$	0	25	43	107	45
$0.2 \leq R^Z < 0.4$	46	137	152	144	190
$0.4 \leq R^Z < 0.7$	171	90	58	2	18
$0.7 \leq R^Z < 0.9$	36	1	0	0	0
$0.9 \leq R^Z \leq 1.0$	0	0	0	0	0

The level of similarity between signs is different for each font. Therefore, the values of the character correlation coefficient were divided for five ranges: lack of relationship ($0.0 \leq R^Z < 0.2$), weak relationship ($0.2 \leq R^Z < 0.4$), moderate relationship ($0.4 \leq R^Z < 0.7$), rather strong relationship ($0.7 \leq R^Z < 0.9$), and very strong relationship ($0.9 \leq R^Z \leq 1.0$). This allows carrying out the quantitative analysis of values of the character correlation coefficient belonging to each range mentioned above. A large number of values belonging to the moderate relationship, rather strong relationship, and very strong relationship ranges is very positive from the point of view of electromagnetic protection of information. In this case the recognition of signs could not be correct, in contrast to the Arial font for which the values of the character correlation coefficient first of all belong to the lack of relationship and weak relationship ranges. It allows recognizing almost every sign because it has strong distinctive features.

The obtained values of the character correlation coefficient for the safe font are substantially different than the values calculated for traditional font characters and for Sang Mun font and Null Pointer font. The safe font characters are more similar, which has a direct influence on errors in attempts to find a correct character in noisy images. Characters of traditional and Sang Mun fonts having characteristic decorative elements that could be much more easily identified in spite of distortion of the regularity of their shapes. Font Null Pointer cannot be classified as a safe font in spite of having some characteristics of one. The structure of individual characters lacks consistency in appropriate proportions of line thickness and their locations.

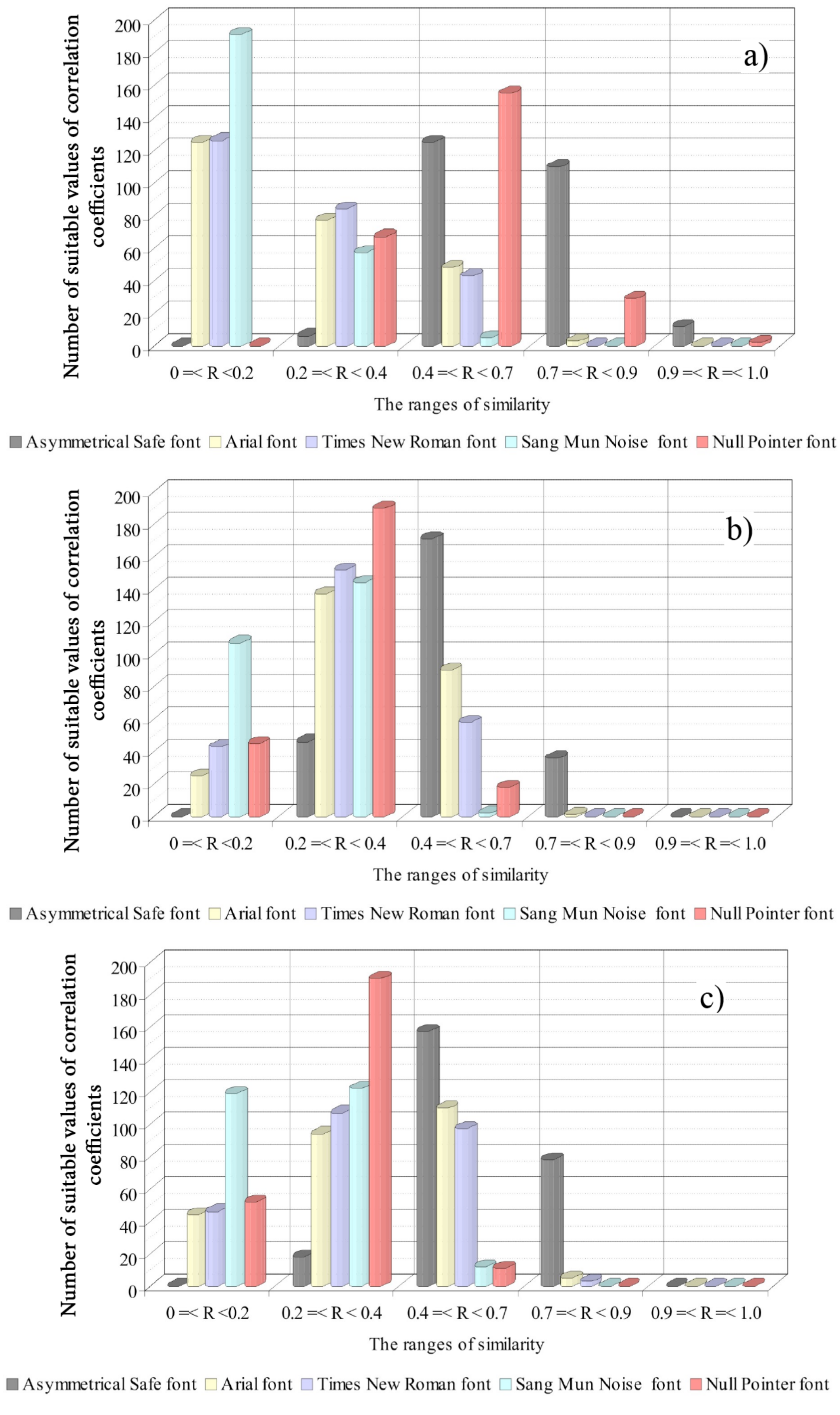

Figure 8. The values of character correlation coefficient R^Z for appropriate ranges of similarity: a) VGA standard, b) DVI standard, c) laser printer with dual diode of laser system.

4. Character error rate

Visual analysis is a basic method classifying emission sources in the form of graphic lines in tests of effectiveness of the electromagnetic infiltration process [22–24]. However, the method is support by computer analysis. The correlation method is the most frequently used one. CER is its measure. It defines the relation between the number of incorrectly recognized or unrecognized characters and the number of all characters found in an image. Determining CER values was limited to traditional fonts and the safe font.

Let's assume that r is number of characters existing in the analyzed image, c is number of characters incorrectly recognized, e is number of characters correctly recognized, k is number of characters unrecognized but wanted $(k = r - e)$, and f is number of all characters existing in the analyzed image. Then:

$$CER = \frac{c+k}{f} = \frac{c+(r-e)}{f}. \tag{6}$$

For the purpose of analyses it was assumed that $r = 31$. Then $f = 651$. Thirty-one characters were randomly located in a matrix of size 21 × 31. Examples of fragments of analyzed images are shown in Figure 9. Of course the quality of the images is very poor. This results from the fact that they are obtained using the rasterization method of very weak revealing emission signals. However, we can notice that Arial font characters could be distinguished in contrast to the safe font characters.

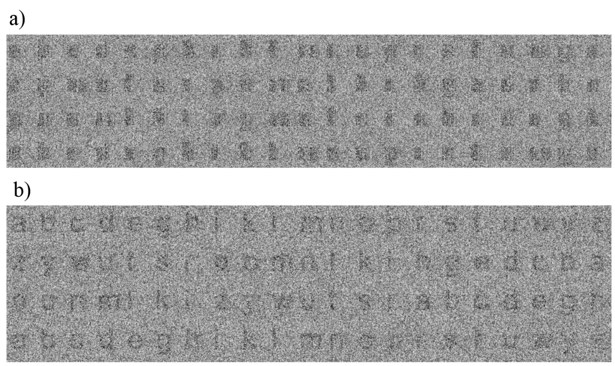

Figure 9. Fragments of analyzed images containing characters of letters for a) Asymmetrical Safe font, b) Arial font ($SNR = SNR_{DVI}3$, $SNR_{DVI}3 < SNR_{DVI}2 < SNR_{DVI}1$).

CER values for selected letters of the analyzed fonts at the output of the IIC for emission sources in the form of DVI, VGA video standard, and a laser printer with two-diode laser system are given in Table 4.

Some characters, for example "c", "e", "n", "o", "s", and "u", have high CER values unlike the others. It applies to the safe font. For this font the values between the mentioned signs and others are very high. Then for the accepted criterion the correct sign cannot be recognized. The determined value of the character correlation coefficient is this criterion. Because the correlation between signs of the safe font is very high, we could take three types of decisions that have an influence on values of CER:

- the correct recognition of a sign;
- the incorrect recognition of a sign;
- failure to recognize a sign.

5. Conclusions

This article presents a so-called safe computer font as a solution for counteracting electromagnetic eavesdropping. The suggested solution was compared to the existing solutions in the form of video signal filtration (Kuhn's

Table 4. Values of CER ($\times 10^{-3}$) for characters on the output of the IIC for different sources of revealing emissions and for SNR_{DVI} 3.

Character	VGA standard			DVI standard			Laser printer		
	A 1)	B 2)	C 3)	A 1)	B 2)	C 3)	A 1)	B 2)	C 3)
a	3	3	96	2	3	25	5	7	589
b	1	159	83	5	5	58	0	14	115
c	3	48	824	60	17	661	73	24	654
e	1	12	724	60	14	344	7	14	715
l	3	3	25	18	8	20	143	8	278
n	112	109	527	40	17	404	49	87	443
o	70	2	589	80	65	487	32	51	693
p	68	3	89	46	14	12	3	5	178
r	4	202	809	22	6	114	34	27	615
s	3	5	559	23	3	141	2	2	640
u	19	8	430	11	8	570	5	14	438
y	3	3	21	6	5	204	0	2	76
z	1	2	239	9	54	100	5	5	418

1) Arial font, 2) Times New Roman font, 3) Asymmetrical Safe font.

solution), Sang Mun Noise font, Null Pointer font, and the changes of colors of background and alphabet characters.

The comparisons consisted of an attempt to identify the characters included in the images reconstructed from the revealing emission signals. In each case the font characters, except the characters of the safe font, were identifiable using the visual method or the correlation method. The characters of the safe font are very poorly identifiable due to the elimination of distinctive features of characters. Thus, the reconstruction of the text data was impossible. CER values confirmed the effectiveness of the safe font (Asymmetrical Safe font) in electromagnetic protection of processed text data.

The filtration of the video line resulting in narrowing of a signal spectrum could be effective in particular for the VGA analogue standard. This standard is not very sensitive to changes of signal shapes. In the case of the DVI digital standard the filter application on the output of a graphic card results in the display of false pixels or a monitor switching off. Analogical effects could be observed during the operation of laser printers.

References

[1] Kubiak I, Przybysz A. Electromagnetic protection of information and communication technology systems and networks. Telecommunication Review and Telecommunication News 2006; 12: 371-374 (in Polish with an abstract in English).

[2] Grzesiak K, Kubiak I, MusiałS, Przybysz A. Electromagnetic Information Security. Warsaw, Poland: Publisher House of Military University of Technology, 2009 (in Polish with an abstract in English);

[3] Kubiak I, Musial S. Hardware raster generator as a tool that supports electromagnetic infiltration. Telecommunication Review and Telecommunication News 2011; 11: 1601-1607 (in Polish with an abstract in English).

[4] Kubiak I. Power line as a source of sensitive emissions from laser printers. Prz Elektrotechniczn 2017; 6: 106-111 (in Polish with an abstract in English).

[5] Kubiak I. The unwanted emission signals in the context of the reconstruct possibility of data graphics. International Journal of Image, Graphics and Signal Processing 2014; 11: 1-9.

[6] Grzesiak K, Przybysz A. Software raster generator. Telecommunication Review and Telecommunication News 2011; 11: 1596-1600,(in Polish with an abstract in English).

[7] Kubiak I. Video signal level (colour intensity) and effectiveness of electromagnetic infiltration. B Pol Acad Sci Tech 2016; 64: 207-2018.

[8] Przybysz A. Emission security of DVI and HDMI interfaces. Telecommunication Review and Telecommunication News 2014; 7: 669-673 (in Polish with an abstract in English).

[9] Kubiak I. Digital (DVI) and analog (VGA) graphic standard in electromagnetic protection of text information. Telecommunication Review and Telecommunication News 2014; 6: 413-416 (in Polish with an abstract in English).

[10] Loughry J, Umphress DA. Information leakage from optical emanations. ACM T Inform Syst Se 2002; 5: 262-289.

[11] Kubiak I, Przybysz A. The impact of commercial equipment design to electromagnetic protection of data process. Prz Elektrotechniczn 2015; 11: 41-44 (in Polish with an abstract in English).

[12] Grzesiak K, Przybysz A. Emission security of laser printers. In: Military Communications and Information Systems Conference; 27–28 September 2010; Wroclaw, Poland. Warsaw, Poland: Publisher House of Military University of Technology. pp. 353-363.

[13] Kubiak I, Przybysz A. Printing technology and electromagnetic protection of information. Prz Elektrotechniczn 2016; 1: 177-181 (in Polish with an abstract in English).

[14] Kubiak I, Przybysz A. Laser printers and effectiveness of attack type TEMPEST. Warsaw, Poland: Publisher House of Military University of Technology, 2016 (in Polish with an abstract in English).

[15] Kuhn MG. Compromising Emanations: Eavesdropping Risks of computer displays, Technical Report. Cambridge, UK: University of Cambridge Computer Laboratory, 2003.

[16] Kuhn MG. Optical time-domain eavesdropping risks of CRT displays. In: Proceedings of the 2002 IEEE Symposium on Security and Privacy; 12–15 May 2002; Berkeley, CA, USA. New York, NY, USA: IEEE. pp. 3-18.

[17] Kubiak I. Computer Font Resistant to Electromagnetic Infiltration. Warsaw, Poland: Publisher House of Military University of Technology, 2014 (in Polish with an abstract in English).

[18] Kubiak I. Computer Font Resistant to Electromagnetic Infiltration. Results of Tests and Analises. Warsaw, Poland: Publisher House of Military University of Technology, 2014 (in Polish with an abstract in English).

[19] Kubiak I. Null Pointer computer font – electromagnetic safe or not? Telecommunication Review and Telecommunication News 2015; 1: 11-18 (in Polish with an abstract in English).

[20] Kubiak I. Computer font resistant to electromagnetic infiltration process. Prz Elektrotechniczn 2014; 6: 207-215 (in Polish with an abstract in English).

[21] Kubiak I. Methods of Analysis and Digital Processing of Images in Electromagnetic Infiltration Process. Warsaw, Poland: Publisher House of Military University of Technology, 2013 (in Polish with an abstract in English).

[22] Bahare J, Abdolah C. Persian sign language recognition using radial distance and Fourier transform. International Journal of Image, Graphics and Signal Processing 2014; 1: 40-46.

[23] Michael JM. The Pentagon worries that spies can see its computer screens, someone could watch what's on your VDT. Wall Street Journal 07.08.2000.

[24] Ketenci SS, Gangal A. Automatic reduction of periodic noise in images using adaptive Gaussian star filter. Turk J Electr Eng Co 2017; 25: 2336-2348.