ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 8, Issue 3 , March 2021

Increasing the credibility of information in the systems of electronic circulation of documents of enterprises

Djumanov Olimjon Israilovich, Kholmonov Sunatillo Makhmudovich, Shukurov Lazizbek Elbek ug'li

Candidate of Technical Sciences, Associate Professor, Department of Information Technologies, Samarkand State University, Samarkand, Uzbekistan
PhD in Technical Sciences, Department of Information Technologies, SamarkandStateUniversity, Samarkand, Uzbekistan
Graduate student, Department of Information Technologies, SamarkandStateUniversity, Samarkand, Uzbekistan

ABSTRACT: The problem of increasing the reliability of the information in electronic documents (ED) is formulated on the basis of information redundancy of different nature. Methods for increasing the reliability of information in electronic document management systems based on search, recognition and classification of requisites, attributes, formats of electronic documents for the purpose to detect and correct distortions have been investigated. Principles and methods of reference checks based on fractal processing of image characteristics and lexicological synthesis of the structure of electronic documents are proposed. A software-algorithmic complex has been implemented, on the example of which a raising in the information reliability indicator has been proved in comparison with existing technologies.

KEYWORDS: electronic document, credibility, reference check, fractal processing, lexicological synthesis, document structure.

I. INTRODUCTION

The analysis of modern scientific literature has shown that in the existing service technologies, insufficiently resolved issues of ensuring the reliability of the information in electronic documents (ED) are not sufficiently resolved, which are one of the key factors in the effective functioning of electronic document management systems (EDMS) of enterprises and organizations [1]. The lack of effective tools that perform the functions of optimizing the credibility of the transmission and processing of electronic documents creates additional difficulties and also leads to a decrease in the quality of functioning of control systems [2].

The decrease in the credibility of the transmission and processing of information in the EDMS is due to the probability of operator errors allowed when entering information, due to errors in scanning and recognition devices, the impact of interference in communication channels, failures, and failures of electronic devices, as well as spelling errors of users [3].

This work substantiates the study and development of a software-algorithmic complex focused on increasing the reliability of the information in the EDMS based on application of search mechanisms, recognition and classification, control of the reliability of details, attributes, ED formats, as well as the detection and distortion correction [4].

The following tasks were solved: development of mechanisms for detecting and correcting distorted information, as well as methods for ensuring the credibility, integrity and safety of ED; development of algorithms for operational analysis and processing of ED, making the necessary corrections in the documents; development of algorithms for analysis, synthesis, generation of texts, translation from one type to another [5].

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 8, Issue 3 , March 2021

A software-algorithmic complex for increasing the reliability of information (SACIRI) has been developed, which includes the following typical modules: integrator - means of settings, associations, import, export of data; semantic synthesizer, text analyzer and ED generation tools; an adapter built on the basis of mechanisms for regulating variable data processing algorithms; data loading interface, libraries of keywords and dictionaries; converter - data converter [6].

II. CONSTRUCTIVE APPROACHES, PRINCIPLES AND MECHANISMS FOR INCREASING THE RELIABILITY OF INFORMATION IN ELECTRONIC DOCUMENTS

When the EDMS is launched, the input are scattered versions of ED, a library of ED classes and programs, tasks for processing ED, the finite presentation format of the template document. Input ED are presented in the form of records, the constituents of which break into a specified number of fields [7].

To perform the function of searching, analyzing, ensuring the reliability, integrity and safety of information, the paintings are reflected of links between the components of ED are reflected in the following levels: the first reflects the paintings of links between ED; the second reflects the links between the attributes of the ED; the third reflects the links between the signs of ED; on the fourth, the rules for controlling the credibility, integrity and safety of information are linked; the fifth reflects the connections of inputs and outputs of a fuzzy semantic hyper network (FSHN); the sixth is the output layer of the network for making decisions about the reliability of the ED information [8].

The frame variant of the ED scenarios presentation using the basic frames template is proposed and implemented [9]:

$$M_i^{U_i} \; ; \; \Phi_{S_1}^{U_1 \Leftrightarrow U_{1+n}} \text{ and their elements - } M_{i,j}^{Pic}.$$

The frame template reflects the formed logical and structural connections between the elements of the ED, their properties, the number of levels, as well as the involved modules in the form

$$M_i^{U_i} \in [U_{1 \to n} \in (S_{1 \to m})],$$

where $M_i^{U_i}$ - an array of extracted and specified properties of ED:

U_i - the level of connections between the components of ED;

S_i - components of ED belonging to the FSHN level;

i - counter of the search for ED elements;

n - number of levels of FSHN;

m - number of ED components.

The frame and connections between the ED components are specified in the form

$$\hat{O}_{S_1}^{U_1 \Leftrightarrow U_{1+n}} \in \left[U_i (S_i^{U_i} \to \frac{U_{\frac{i}{i+k}} (S_i^{\frac{U_i}{U_{i+k}}})}{k_{S(B_i)}} \right],$$

where $\hat{O}_{S_1}^{U_1 \Leftrightarrow U_{1+n}}$ - frame belonging to U_i levels, containing a "frame of relations" links between components of S_i ED;

$\hat{O}_{S_1}^{U_1 \Leftrightarrow U_{1+n}}$ – a formed graph that which the stacking frame;

k -number of analysis frames;

$k_{S(B_i)}$ - coefficient of parallelization of data processing.

SACIRI to ensure the integrity, safety, and reliability of the information in the EDMS includes the following program modules:

- of the program menu, process graphs of the analysis and processing of ED;
- generation, analysis, synthesis, broadcasts of ED texts;

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 8, Issue 3 , March 2021

- performing functions of parameter setting and integration with other modules;
- computational schemes of the structural components of the FSHN, based on modules, using the properties of fuzzy identification algorithms;
- ED information control, error detection and correction; withdrawal of the results of the system as a whole [10].

III. THE MECHANISM OF RATING ASSESSMENT OF THE RELIABILITY OF ED INFORMATION

A method has been developed and implemented, in which the lots of promotion jobs increasing the reliability of ED information is ordered, which are used to evaluate the results according to some r_{run} "rating" [11]. For what, lots of n assignments is formed and the carried out J_{run} task with the greatest rating is allocated. The n value is determined by the condition

$$n = \begin{cases} N, & if \ \ a \ \ \mathrm{n}_{\mathrm{run}} = 0; \\ \min\{ n_{run}, N, M \}, & if \ \ a \ \ \mathrm{n}_{\mathrm{run}} > 0, \end{cases}$$

where N - total number of assignments;

M - the number of data processing modules in the complex;

n_{run} – integer module parameter.

It is required to determine the execution time of a batch of tasks with the length of the $p.l$ on the multitude $c \in j.C$ by the ratio of

$$\vartheta = (j.f).\mu \cdot (p.l) / c.r ,$$

where $(j.f).\mu$ - average execution time of a complex on a standard processor;

$c.r$ - complex performance.

It is believed that the time ϑ from the multitudes of $j.C$ is the same for all modules. The packet size limit is superimposed proceeding on the condition

$$1 \le p.l \le c.v .$$

Variations of possible ϑ values are checked by the condition

$$(j.\mu) \cdot \max_{c \in C} (1/(c.r)) \le \vartheta \le (j.\mu) \cdot \min_{c \in C} ((c.v)/(c.r)) .$$

Denoting the left and right sides of the inequality by m and M and we obtain

$$\vartheta = m + k_{check}^{1} M ,$$

where $k_{check}^{1} \in [0,1]$ - module parameter.

The packet size is determined by fixing the ϑ value from the specified range

$$p.l = \vartheta \cdot (c.r) / ((j.f).\mu) .$$

The ϑ parameter is considered the main one for setting up the complex in the EDMS environment

$j.t_{check} = k_{check}^{2} \vartheta$, where $k_{check}^{2} \in [0,1]$ - module parameter.

To analyze and evaluate the effectiveness of the complex, it is required to determine the following $x_1(t)$ parameters - the number of modules involved by the time of the t, when information distortions occurred in the ED;

$x_2(t)$ - the number of modules involved by the time of the t, when in the ED there was no distortion of information;

$x_3(t)$ - the total number of involved modules by the time t;

$x_4(t)$ - the number of modules involved by the time t, in which no errors were found in the information [12].

Evaluation of the efficiency functional of the complex as part of the S program modules will be written by the

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 8, Issue 3 , March 2021

$$\omega(S) = \sum_{i=1}^{4} c_i \lim_{T\to\infty} T^{-1} \sum_{t=0}^{T} E_S x_i(t),$$

where E_S - operator of averaging the efficiency of all involved modules S;

S - the given structure of the involved modules of the complex in the EDMS;

c_i - weighting coefficients corresponding to the criteria for assessing the minimum probability of undetected errors, “cost” and “labor intensity” of various data processing algorithms used to control the reliability of ED information.

IV. CONCLUSION

Thus, the effectiveness of the application of the technique to ensure the credibility of the transmission and processing of ED has been investigated and proved. Similarly, optimizing methods for processing data from a wide variety of production systems can be built.

REFERENCES

[1]. Keldysh N. V. Methodological bases of automated solution of problems of departmental electronic document management. // Scientific method. collection No. 56. - M., 2009. - p. 110-117.

[2]. Efimov A. N., Litvinyuk S. B. Analysis of the nature of errors allowed by optical character recognition programs / / Moscow: Mashinostroenie, Avtomatizatsiya i sovremennye tekhnologii. - 1997, No. 10. - p. 29-35.

[3]. Konovalov M. G. Methods of adaptive information processing and their applications. - M.: IPI RAS, - 2007. - p. 21-28.

[4]. Bessonov, S. V. Optimization of electronic document management in corporate systems: dis. kand. ekon. nauk. M., 2000,187 p.

[5]. Gudov, A. M. On a model for optimizing document flows, implemented when creating an electronic document management system, collection "Computing Technologies", 2006. - p. 53-65.

[6]. Zhumanov I. I., Akhatov A. R. Fuzzy semantic hypernet of information reliability control in electronic document management systems / / 4-th International Conference on Application of Information and Communication Technologies, Tashkent, 12-14 October 2010, Section 2, IEEE. – p.21-25.

[7]. Akhatov A.R., Jumanov I.I. Improvement of text information processing quality in documents processing systems // 2nd IEEE/IFIP International Conference In Central Asia On Internet ICI-2006, September 19-21, International Hotel Tashkent, Uzbekistan

[8]. Kemal Oflazer, Sergei Nirenburg, Marjorie McShane. Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning. // To appear in Computational Linguistics, 2001.

[9]. Akhatov A. R., Zaripova G. I., Zaletova T. Control of the reliability of text information in electronic document management systems based on a neural-fuzzy semantic hypernet. Control and management", TSTU. - Tashkent, 2012. - №6 (48). - Pp. 44-50.

[10]. Zhumanov I. I., Karshiev Kh. B., Tishlikov S. A. Improving the reliability based on the use of natural redundancy of information in electronic documents / / Informatics and its applications, 2021. Vol. 10. Issue 1. C.

[11]. Kholmonov S. M., Absalomova G. B. Methods and algorithms for improving the reliability of textual information of electronic documents / / International Scientific Journal "Science and the World" No. 4 (80), pp. 56-60.

[12]. Zhumanov I. I., Karshiev H. B. Methods of improving the efficiency of electronic document management systems based on multi-criteria optimization of information reliability / / Journal of Problems of Computational and Applied Mathematics, TUIT, Tashkent, 2020-No. 3. - pp. 51-70.