METHODS FOR IMPLEMENTING DOCUMENT FLOW BASED ON THE USE OF A DECISION TREE

[1]Marysheva L.T., [2]Rakhmonov A.T., [3]Latipova N.Kh., [4]Abduvalieva Z.A.
[1,2,3]Associate Professor of TUIT
[4]TUIT doctoral student of TUIT
https://doi.org/10.5281/zenodo.11521400

Abstract. *Working with office documents is complex and routine; it is complicated by the fact that their approval requires approval from several departments. Enterprise employees are faced with such problems as contacting various departments to approve a document and it is not always the fact that management is on site. To solve this problem, we have fully applied the decision tree. It allows you to submit documents for approval when you log in to the system; it allows you to decide which employee is responsible for the implementation of a particular document and whether the document should be transferred to the right department or not.*

Keywords*: decision tree, graph, tree, document flow, incoming data subsystem, outgoing document subsystem.*

Introduction As a subsystem for automating the process of optimizing document flow, the document flow system implies the automatic transfer of documents, execution in accordance with certain criteria, and execution of assigned tasks. In an organization, when an employee has a file that requires the approval of managers, for confirmation, they pass from the lower department to the higher one and only after receiving the approval of the lower department, the document can be transferred to the higher department for approval. The decision tree is designed to make a decision about sending a document for approval in order to do this step by step. Having received approval from the desired department, the system decides to transfer the file to the next department.

The task of organizing electronic document management in a higher educational institution is to determine which category a document belongs to and, depending on the recognition result, to find ways to pass documents within the organization.

Determining which category, a document belongs to is carried out using keywords. The predictive analysis module recognizes a document by keywords and determines which category a particular document belongs to and redirects it to the required department for processing [1]. To determine the category of documents, a preliminary expert analysis of the definition of keywords was carried out. Based on keywords, the recognition module determines which category the document belongs to and sends it to the right department. Words that occur more often than others are taken as key words [2]. In documents, you may encounter the fact that the same words can appear in several documents at the same time. In this case, you should use phrases. For example, the word higher education appears in several documents at the same time. In this case, when searching for keywords, the phrase “Higher and secondary specialized education” was taken, and in another document “Law on Higher Education”. Also, keywords should be considered a word that defines the essence of the document and is unique for this document. This resolves the issue of keyword identity. Below is a table that clearly shows the identification of keywords and phrases. [2]. For convenience, I divided all input data into 4 categories. First category: laws and decrees of the President of the Republic of Uzbekistan. Second category Ministry of Higher and Secondary Education of the Republic of Uzbekistan, Ministry of Digital Technologies of the Republic of

Uzbekistan. The third category is a certificate of orders, instructions and their implementation by the rector and vice-rectors of the university. Each of these categories consists of one folder. The fourth category is department documents, curriculum, syllabus, department reports, information about graduates and students. The fourth category includes 18 folders.

One way is to model using a data structure - a decision tree. A tree is a data structure that is a collection of elements and relationships that form a hierarchical structure of these elements. Each element of the tree is called a vertex (node) of the tree.

Below is a system of decision trees for passing a document:

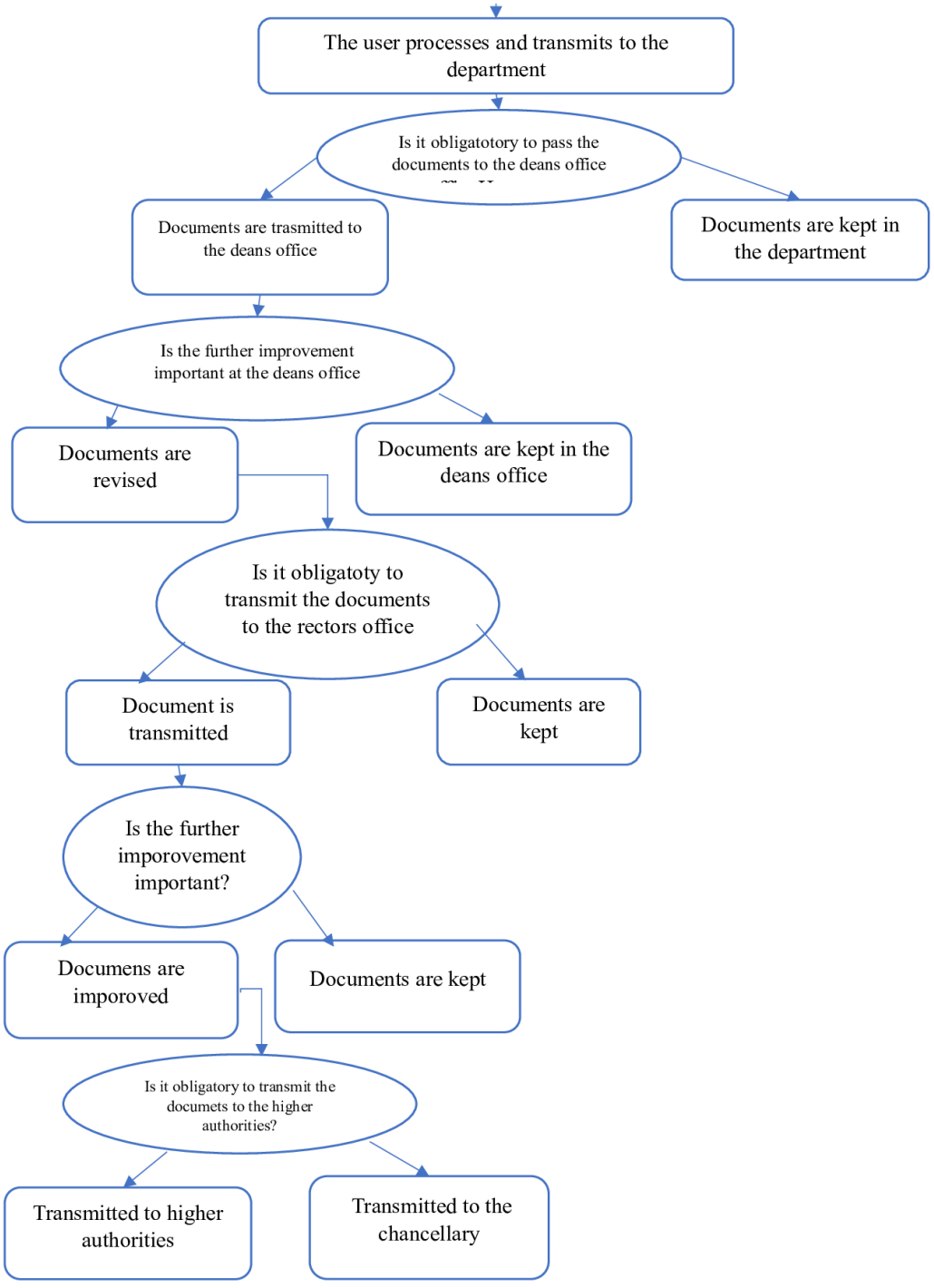

Diagram 1.1 Scheme of document passage from sender to executor

For each keyword, we determine its weighting coefficient as follows. In the documents we determine the keyword with the greatest participation [3]. Then the weight coefficient of this keyword is determined by the formula [3].

Mainly, 19 inputs are considered, which consist of 1292 documents. For each input data, several keywords are defined. In total, for 19 inputs, we have identified 258 keywords that can be used to recognize inputs into four categories. For the keywords of each category, its weighting coefficient is also determined [3]. Here are the keywords of the first category with weighting coefficients.

Table 2. Fragment of keywords with weighting coefficients.

Decrees and decisions of the President of the Republic of Uzbekistan				
	Key words	Signs	Weighting coefficient	Receiving and executing objects
	President of the Republic of Uzbekistan	X_1	$k_1 = 0.7$	1-object
	Social	X_2	$k_2 = 0.2$	
	Decision	X_3	$k_3 = 0.8$	
	Minister	X_4	$k_4 = 0.3$	
	accordingly	X_5	$k_5 = 0.4$	
	Information technology	X_6	$k_6 = 0.3$	
	Governor	X_7	$k_7 = 0.18$	
	Resolution	X_8	$k_8 = 0.25$	
	communication	X_9	$k_9 = 0.18$	
	President	X_{10}	$k_{10} = 0.62$	
	development	X_{11}	$k_{11} = 0.37$	
	State	X_{12}	$k_{12} = 0.43$	
	Economics	X_{13}	$k_{13} = 0.43$	
	academy	X_{14}	$k_{14} = 0.31$	
	Corruption	X_{15}	$k_{15} = 0.25$	
	Cabinet	X_{16}	$k_{16} = 0.5$	

Based on the above data, we will construct an algorithm that determines the input data stream. Let us introduce the vector $x = (x_0, x_1, x_2, ..., x_{258})$, where x_i is the frequency of occurrence of the keyword in the input data stream [3].

$$\bar{x}_1 = \frac{1}{16}\sum_{i=1}^{16} k_i\, x_i\,, \quad (1)$$

$$\bar{x}_2 = \frac{1}{9}\sum_{i=17}^{25} k_i\, x_i\,, \quad (2)$$

$$\bar{x}_3 = \frac{1}{7}\sum_{i=26}^{33} k_i\, x_i \quad , (3)$$

$$\bar{x}_4 = \frac{1}{224}\sum_{i=34}^{256} k_i\, x_i \ , (4)$$

When the input $\bar{x}_1 \neq 0$ data belongs to the first category, when the input $\bar{x}_2 \neq 0$ data belongs to the second category, when the input $\bar{x}_3 \neq 0$ data belongs to the third category, when the input $\bar{x}_4 \neq 0$ data belongs to the 4th category [3]..

Conclusion: As a result of the research, the software package “Optimization of electronic document flow in institutions” was developed.

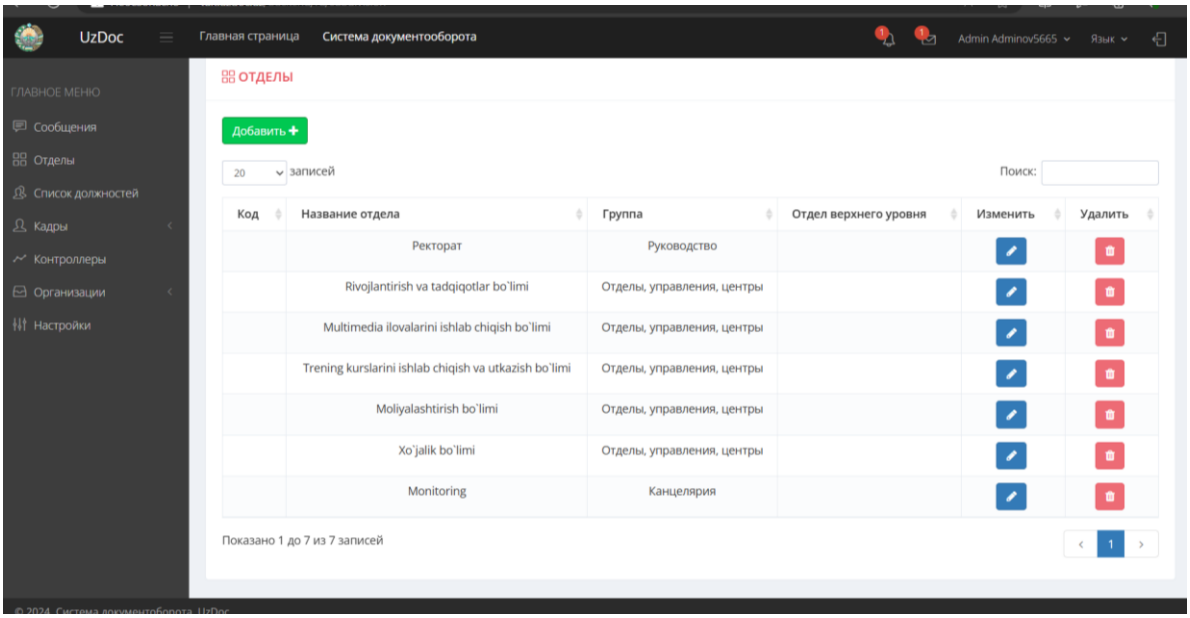

Figure 1.2. Dialog window of the software package for optimizing document flow in institutions

As a result of studying the documents of the department, an intelligent module was developed for recognizing and distributing documents by object [4]. Based on the obtained key phrases, an automatic document recognition module was built [7]. The resulting algorithm allows you to build an intelligent module that allows you to recognize documents by keywords [8]. An algorithm for constructing software is proposed.

REFERENCES

1. Abduvalieva Z.A , Marisheva L.T , Latipova N.X , Sheyna N.E. Structure and functional features of document management systems on the example of the department. Journal of Northeastern University, Volume 25 Issue 04, 2022.
2. Ricardo Campos a,e,∗ , Vítor Mangaravitee,g , Arian Pasquali e , Alípio Jorge b,e , Célia Nunes c,f , Adam Jatowt d. YAKE! Keyword extraction from single documents using multiple local features. 2019 Elsevier Inc. All rights reserved.
3. Rakhmanov A.T., Abduvalieva Z.A. DOCUMENT MANAGEMENT BASED ON MATHEMATICAL MODELING. Muhammad Al-Khorazmiy Avlodlari (06/2/2024)
4. Isaeva, M., Yoon, H., Y.: Paperless university — How we can make it work?. In: 15th International Conference on Information Technology Based Higher Education and Training (ITHET). pp. 1–8 (2016).
5. Luo, H., Fan, Y., Wu, C.: Overview of Workflow Technology. J. Softw. 11, 78-82

6. Fan, Yusun: Base on Workflow Management Technology. Beijin:Tsinghua University Press, 32, (2001)
7. Chen, Hong-na, Zu, Xu, Zhou, Feng: On the Developing Situation, Research Content and Trend of Workflow Technology. Journal of Chongqing Instiute of Technology. 20(2), 65-69 (2006)
8. Li, Zhao, Qing, Li, Farong, Zhong: A Visual Modeling Framework of Workflow Systems Based on CCS. Semantics, Knowledge and Grid. Fifth International Conference. pp. 200-207 (2009)
9. Dinesh, P, Mital, Goh, Week, Leng School of Electrical & Electronic Engineering Nanyang Technology University Nanyang Avenue, Singapur 2263. Text segmentation for automatic document processing. Rosemont, IL, USA. pp. 132-133 (1995).