

**KOMPONENTE ZA PODRŠKU SISTEMA ZA OPTIMIZACIJU RASPOREDA NASTAVE
FAKULTETA TEHNIČKIH NAUKA****COMPONENTS SUPPORTING THE SCHEDULE OPTIMIZATION SYSTEM OF THE
FACULTY OF TECHNICAL SCIENCES**

Ksenija Prčić, Nikola Luburić, *Fakultet tehničkih nauka, Novi Sad*

Oblast – ELEKTROTEHNIKA I RAČUNARSTVO

Kratka sadržaj – U okviru rada razvijene su komponente u sistemu za kreiranje rasporeda nastave na Fakultetu tehničkih nauka u Novom Sadu. Fokus komponenti je parsiranje značajnih dokumenata za kreiranje rasporeda, transformacija dobijenih podataka i izvoz finalnog rasporeda u različite formate. Rad detaljno istražuje domen problema, uključujući specifične izazove i ograničenja, kao i korisničke zahteve koji su ključni za efikasno funkcionisanje sistema. Opisane su faze prikupljanja i obrade podataka korišćenjem ETL (Extract, Transform and Load) metodologije, arhitektura sistema i implementacija razvijenog rešenja. Predloženi su i potencijalni putevi za unapređenje sistema. Krajnji proizvod ovog istraživanja je funkcionalan sistem za kreiranje i optimizaciju rasporeda nastave, koji se integriše u informacioni sistem fakulteta, omogućavajući korisnicima da upravljaju kompletnim procesom, od obrade ulaznih podataka do konačnog prikaza rasporeda u različitim formatima.

Ključne reči: rukovanje dokumentima, problem planiranja rasporeda, predstava domenskog znanja

Abstract – As part of this paper, components supporting the teaching schedule optimization system at the Faculty of Technical Sciences in Novi Sad were developed. Focus of the components is parsing of input documents significant for teaching schedule creation, transformation of extracted data and exports of the optimized schedule into various formats. The thesis thoroughly investigates the problem domain, including specific challenges and constraints, as well as user requirements essential for the system's effective functioning. The stages of data collection and processing using the ETL (Extract, Transform, and Load) methodology are detailed, along with the system's architecture and implementation. Potential ways for system enhancements are also discussed. The final product of this research is a system that facilitates the creation and optimization of teaching schedules and is also integrated into a broader informational system of the faculty, allowing users to manage the entire process, from processing input data to the final display of schedules in various formats.

Keywords: document management, scheduling problem, representation of domain knowledge

1. UVOD

Problem kreiranja rasporeda predstavlja izazov u različitim aspektima društva, od individualne organizacije vremena do organizacije rada zaposlenih u velikim organizacijama. Fokus istraživanja su izazovi kreiranja rasporeda nastave u obrazovnim ustanovama, gde efikasnost organizacije nastave značajno utiče na kvalitet obrazovanja i zadovoljstvo studenata i nastavnog osoblja. Istraživanje se bavi kreiranjem rasporeda nastave za Fakultet tehničkih nauka u Novom Sadu, gde je uočena potreba za sistemom za optimizaciju rasporeda. Trenutno ne postoji sistematičan pristup prikupljanju i ažuriranju podataka, što dovodi do gubitka informacija.

Rad obuhvata prikupljanje i obradu podataka koji se koriste za kreiranje rasporeda. Prvi korak je identifikacija i prikupljanje podataka koji utiču na kreiranje rasporeda. Na osnovu dostupnih podataka formira se raspored i optimizuje iskorišćenost svih resursa.

Za obradu podataka dizajniran je parser podataka koji adaptira Extract, Transform and Load (ETL) [1] metodologiju. Ovaj proces uključuje ekstrakciju podataka iz različitih izvora dostupnih u raznim formatima, njihovu transformaciju u format pogodan za generator rasporeda i integrisanje transformisanih podataka u sistem za kreiranje i dalju optimizaciju rasporeda. Cilj istraživanja je razvoj sistema za obradu ulaznih podataka različitih struktura i njihovu transformaciju u format pogodan za korišćenje u širem sistemu optimizacije rasporeda. Sistem takođe kreira različita dokumenta za prikaz rasporeda zainteresovanim stranama. Naredno poglavlje razmatra specifične izazove i ograničenja domena problema prilikom izrade rešenja. Poglavlje 3 opisuje arhitekturu sistema. U 4. poglavlju opisana je implementacija sistema. Poslednje poglavlje rada je zaključak koji sadrži sažetak istraživanja i moguća buduća unapređenja.

2. DOMEN PROBLEMA

Rešenje problema optimizacije rasporeda u obrazovnim ustanovama je problem na svetskom nivou, što potvrđuje i postojanje međunarodnog takmičenja posvećenog ovoj temi [2]. U radu [3] opisan je postupak organizacije takmičenja kao i broj timova (preko 490) i država učesnica (preko 66), gde ovi brojevi govore o značajnosti problema.

Rad opisuje postupak rešavanja problema kreiranja rasporeda nastave na Fakultetu tehničkih nauka (FTN) u Novom Sadu sa fokusom na načine prikupljanja i obrade podataka potrebnih za formiranje rasporeda.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio docent dr Nikola Luburić.

Problem optimizacije rasporeda FTN-a u Novom Sadu opisan je u radu [4]. Glavni fokus ovog rada je postupak ekstrakcije podataka iz trenutnog informacionog sistema i korišćenje istih za formiranje novog rasporeda. Oba rada su bazirana na sličnom domenskom modelu koji uključuje: organizacione jedinice fakulteta (departmane i katedre), studentske programe, studentske grupe, predmete, predavače, prostorije, vremenske intervale i termine u okviru rasporeda nastave.

Proces ekstrakcije, transformacije i učitavanja podataka poznat je u svetu softverskog inženjerstva [5], posebno u sferi skladištenja podataka [6]. U radu [7] detaljno je opisan ETL proces primenjen u akademskom okruženju za usklađivanje podataka u distribuiranim bazama. Kako je trenutna struktura informacionog sistema FTN-a distribuirana i sastoji se iz većeg broja različitih komponenti, često dolazi do nekonzistentnosti podataka. U cilju centralizovanog rukovanja podacima značajnim za kreiranje i optimizaciju rasporeda, dizajniran je ETL proces koji prikuplja potrebne podatke, transformiše ih u oblik pogodan za obradu i čuva ih u dokument orijentisanoj bazi radi daljeg održavanja [8].

3. SPECIFIKACIJA SISTEMA

U ovom poglavlju su detaljnije prikazani korisnički zahtevi i opisan je način integracije komponenti opisanih u ovom radu u postojeću informacionu infrastrukturu fakulteta. Rad je deo šireg sistema koji ima za cilj razvoj procesa koji obezbeđuje konzistentno rukovanje podacima neophodnim za kreiranje rasporeda. U nastavku je opisan ceo proces, od prikupljanja podataka, preko kreiranja rasporeda, do dostavljanja informacija o rasporedu svim zainteresovanim stranama.

3.1. Zainteresovane strane i izvori ulaznih podataka

Pre razvoja sistema potrebno je analizirati trenutnu strukturu dokumenata FTN-a, način na koji se oni kreiraju, održavaju i koriste u izradi rasporeda. Takođe, važno je utvrditi ko su zainteresovane strane za konačni raspored, kako se on formira, održava i dostavlja zainteresovanim stranama.

Prvi koraci u razvoju sistema su uključivali analizu strukture fakulteta i identifikaciju sektora koji učestvuju u kreiranju rasporeda ili za koje je raspored od značaja za rad. Na osnovu razgovora sa domenskim ekspertom, profesorima u nastavi, kao i sa profesorima koji se pored nastave nalaze i na organizacionim pozicijama u okviru FTN-a izdvojeni su sektori koji utiču na kreiranje rasporeda: studentska služba, kadrovska služba, katedre, departmani, dekanat i informaciono tehnološka (IT) služba.

Na osnovu inicijalnog istraživanja, prikupljeni su dokumenti koji utiču na kreiranje rasporeda. Izvori ovih dokumenata su: IT služba, domenski ekspert i zvanična stranica FTN-a. Pored sektora koji imaju uticaj na kreiranje rasporeda, utvrđene su i zainteresovane strane za finalni raspored: studenti, nastavno osoblje, nenastavno osoblje i domenski ekspert.

3.2. Izvorna dokumenta

Jedan od najvećih izazova koji je primećen već u početnim fazama istraživanja jeste nekonzistentnost

podataka koji potiču iz različitih izvora. Različiti sektori koji učestvuju u organizovanju nastave vode sopstvene evidencije podataka između kojih ne postoji sinhronizacija.

Najznačajniji dokument za kreiranje rasporeda bio je trenutni raspored nastave, prikazan na Slici 3.1. Zbog njegove složene i nepregledne strukture, odlučeno je da će trenutni Excel dokumenti koji sadrži raspored u novom sistemu biti zamenjen bazom podataka.

Učionica	06:30	06:45	07:00	07:15	07:30	07:45	08:00	08:15	08:30	08:45	09:00	09:15	09:30	09:45	10:00	10:15
A1				Svi	Svi	Svi	Svi									
A2								Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi
A3								Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi
A4																
101								Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi
102								Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi	Svi
103																
104																
105																
107																
108																
108A																
109																
109A																
201																
202																
203																
204																
205																
206																

Slika 3.1. Isečak Excel dokumenta rasporeda nastave

3.3. Slučajevi korišćenja

Rad opisuje komponente koje čine deo većeg sistema za kreiranje i optimizaciju rasporeda. Sistem podržava jedan tip korisnika, administratora, koji je odgovoran za rukovanje ulaznim podacima potrebnim za izradu rasporeda kao i za sam proces kreiranja rasporeda. Komponente imaju tri glavna cilja koja se integrišu u sistem na različitim nivoima:

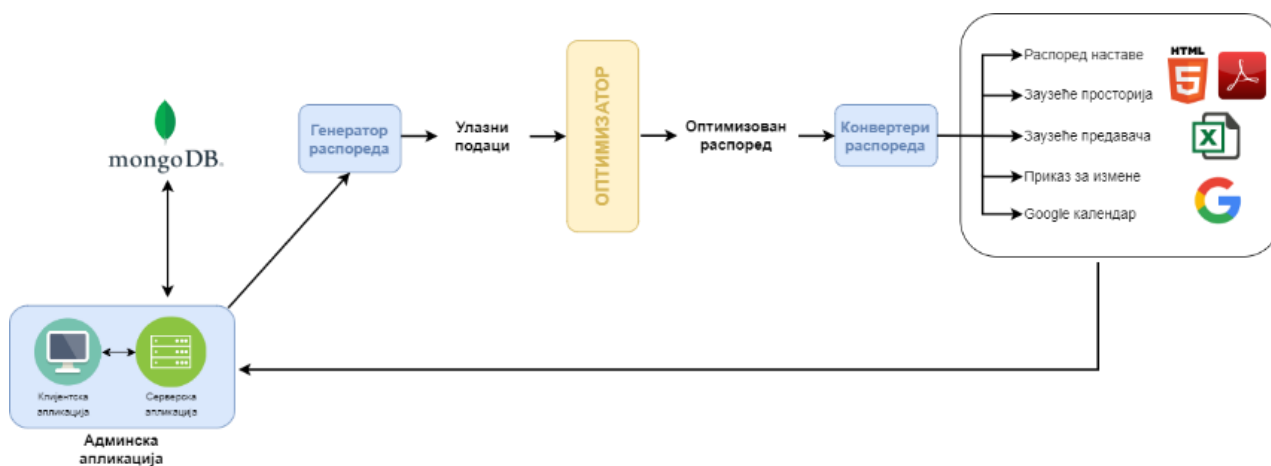
1. Priprema ulaznih podataka:
 - a. Ekstrakcija podataka iz dokumenata bitnih za kreiranje rasporeda,
 - b. Transformacija podataka,
 - c. Učitavanje podataka u JSON formatu.
2. Generisanje termina za raspored na osnovu transformisanih podataka u format kompatibilan sa optimizator komponentom,
3. Izvoz rezultata optimizacije u različite formate prilagođene potrebama svake od zainteresovanih strana:
 - a. Izvoz rasporeda za studente,
 - b. Izvoz rasporeda za predavače,
 - c. Izvoz rasporeda za prostorije,
 - d. Izvoz rasporeda za dalje izmene,
 - e. Izvoz zauzeća predavača u odgovarajućem formatu za uvoz u Google kalendar.

3.4. Arhitektura sistema

Arhitektura sistema prikazana je na Slici 3.2. Rad opisuje komponente sistema koje su međusobno nezavisne i obuhvataju obradu različitih tipova dokumenata i podataka.

Odabrana je odgovarajuća arhitektura kako bi sistem bio efikasan i funkcionalan. Zbog prirode problema, arhitektura sistema koristi više skripti napisanih u Python [9] programskom jeziku koji je izabran zbog dinamičkog tipiziranja i podrške za rad sa Excel i CSV dokumentima kroz biblioteku Pandas [10].

Tokom razvoja korišćeno je Jupyter Notebook [11] okruženje, dok su za produkciju razvijene Python skripte.



Slika 3.2. Arhitektura integracionog sistema

Za upravljanje zavisnostima koristi se alat pip [12], a za generisanje PDF dokumenata alat wkhtmltopdf [13].

3.5. Arhitektura integracionog sistema

Dizajniran je integracioni sistem koji omogućava konzistentno kreiranje rasporeda nastave, od obrade ulaznih podataka, preko njihove optimizacije, do generisanja rasporeda nastave.

Komponente koje cine integracioni sistem su:

- skladište podataka,
- administratorska aplikacija,
- generator rasporeda,
- optimizator i
- konverteri rasporeda.

4. SPECIFIKACIJA SISTEMA

Sistem se sastoji iz tri nezavisne komponente koje su implementirane upotrebom Python programskog jezika. Celokupan izvorni kod komponenti je javno dostupan¹. U daljem tekstu poglavlja opisana je svaka od komponenti, kao i tok podataka kroz njih.

4.1. Struktura parser komponente

Parser komponenta je najkompleksnija od tri opisane u radu i zaduzena je za obradu podataka. Kompleksnost je posledica velikog broja ulaznih dokumenata koje je potrebno parsirati i uskladiti. Proces parsiranja je logički podeljen u 6 koraka, gde se u svakom koraku obrađuje određeni set ulaznih dokumenata i njihov izlaz se ulančava u krajnju strukturu podataka koja je optimizovana za skladištenje u dokument orijentisanoj bazi podataka. Održavanje skladištenih podataka se vrši kroz administratorsku aplikaciju.

4.2. Struktura generator komponente

Generator termina je središnja komponenta u toku podataka kroz sistema optimizacije. Nakon parsiranja podataka, na osnovu realizacije nastave, potrebno je kreirati termine za odabrani semestar. Prilikom generisanja ter-

mina poštovana su pravila propisana Pravilnikom fakulteta [14].

Zbog velikog broja studenata i ograničenog kapaciteta prostorija, potrebno je rešiti problem optimalnog kombinovanja grupa u cilju bolje iskorišćenosti prostora. Ovaj problem je poznat kao problem pakovanja u prostoru [15] i spada u grupu NP-teških problema [16][17]. Generisanje termina rasporeda zahteva kreiranje algoritma za kombinovanje studentskih grupa i rešavanje opisanog NP-teškog problema.

Generisani termini koji su izlaz ove komponente se dalje prosleđuju optimizatoru koji je zaduzen za dodeljivanje prostorija i vremenskog intervala svakom terminu.

4.3. Struktura konverter komponente

Konverter komponenta je najjednostavnija među tri navedene komponente. Njena uloga je kreiranje prikaza rasporeda u odgovarajućim formatima za sve zainteresovane strane. Ulaz predstavlja raspored kreiran od strane optimizatora, sa informacijama o studijskom programu, grupama, predmetima, predavačima, učionicama i vremenskim intervalima. Konverter komponenta pretvara ovu ulaznu strukturu, predstavljenu JSON objektom, u dokumente pogodne za pregled od strane studenata, predavača i drugih zainteresovanih strana.

5. ZAKLJUČAK

U radu je predstavljen razvoj i implementacija komponenti za podršku sistema za kreiranje rasporeda nastave na Fakultetu tehničkih nauka u Novom Sadu. Motivacija za razvoj sistema je bolja organizacija nastave i veće zadovoljstvo studenata i nastavnog osoblja. Sistem podržava različite scenarije korišćenja i osigurava konzistentnost u kreiranju rasporeda, od obrade ulaznih podataka, preko transformacije podataka i generisanja termina rasporeda, do kreiranja izvoza pogodnih za sve zainteresovane strane. Glavna funkcionalnost je obrada podataka iz različitih izvora korišćenjem ETL procesa, omogućavajući lakšu integraciju sa postojećim informacionim sistemom fakulteta.

Za buduća unapređenja, predlaže se integracija sistema sa adapterima za dinamičko preuzimanje podataka, povećavajući njegovu primenu i potencijal za korišćenje u različitim obrazovnim ustanovama. Takođe, proširenje

¹ <https://github.com/evaj10/raspored-parser>

sistema novom funkcionalnošću za izmenu kreiranog rasporeda, integracijom sa kalendarskim rešenjima, olakšalo bi rukovanje kreiranim terminima u okviru rasporeda nastave.

6. LITERATURA

- [1] Extract, Transform, Load процес, <https://www.ibm.com/topics/etl> [датум приступа мај 2024]
- [2] Интернационално такмичење у оптимизацији распореда наставе за 2019. годину, <https://www.itc2019.org/home> [датум приступа мај 2024]
- [3] Müller, T., Rudová, H., & Müllerová, Z. (2024). Real-world university course timetabling at the International Timetabling Competition 2019. *Journal of Scheduling*, 1-21.
- [4] Janković, E., & Luburić, N. (2023). SISTEM ZA OPTIMIZACIJU RASPOREDA NASTAVE FAKULTETA TEHNIČKIH NAUKA. *Zbornik radova Fakulteta tehničkih nauka u Novom Sadu*, 38(12), 1679-1682.
- [5] Vassiliadis, P. (2009). A survey of extract–transform–load technology. *International Journal of Data Warehousing and Mining (IJDWM)*, 5(3), 1-27.
- [6] Simitsis, A., Vassiliadis, P., & Sellis, T. (2005, April). Optimizing ETL processes in data warehouses. In *21st International Conference on Data Engineering (ICDE'05)* (pp. 564-575). Ieee.
- [7] Yulianto, A. A. (2019). Extract transform load (ETL) process in distributed database academic data warehouse. *APTİKOM Journal on Computer Science and Information Technologies*, 4(2), 61-68.
- [8] Vera, H., Boaventura, W., Holanda, M., Guimaraes, V., & Hondo, F. (2015, September). Data modeling for NoSQL document-oriented databases. In *CEUR Workshop Proceedings (Vol. 1478, pp. 129-135)*.
- [9] Званична документација Python програмског језика, <https://www.python.org/doc/> [датум приступа мај 2024]
- [10] Званична документација Pandas Python библиотеке, <https://pandas.pydata.org/docs/> [датум приступа мај 2024]
- [11] Званична документација Jupyter Notebook софтвера, <https://docs.jupyter.org/en/latest/> [датум приступа мај 2024]
- [12] Званична документација pip алата за руковање зависностима за Python програмски језик, <https://pip.pypa.io/en/stable/> [датум приступа мај 2024]
- [13] Званична документација wkhtmltopdf алата за генерисање PDF докумената на основу HTML докумената, <https://wkhtmltopdf.org/docs.html> [датум приступа мај 2024]
- [14] Правилник о изради распореда наставе, ФТН, <http://www.ftn.uns.ac.rs/520569743/pravilnik-o-izradi-rasporeda-casova-nastave-i-rasporeda-ispitana-fakultetu-tehnickih-nauka-od-31-01-2019-> [датум приступа мај 2024]
- [15] Lodi, A., Martello, S., & Vigo, D. (2002). Recent advances on two-dimensional bin packing problems. *Discrete Applied Mathematics*, 123(1-3), 379-396.
- [16] Knuth, D. E. (1974). Postscript about NP-hard problems. *ACM SIGACT News*, 6(2), 15-16.
- [17] Lodi, A., Martello, S., & Vigo, D. (2002). Recent advances on two-dimensional bin packing problems. *Discrete Applied Mathematics*, 123(1-3), 379-396.

Kratka biografija:



Ksenija Prčić рођена је 16. марта 1998. године у Новом Саду. Као носилац дипломе „Vuk Karadžić“ 2017. године завршава гимназију „Јован Јовановић Змај“, након чега уписује Факултет техничких наука, смер Softversko inženjerstvo i informacione tehnologije. По завршетку студија у року, уписује мастер академске студије и исте завршава 2024. године.