

Proces mining: istraživanje podataka o procesima u svrhu njegova poboljšanja

Raguž, Nikolina

Master's thesis / Diplomski rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Ekonomski fakultet u Osijeku**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:145:796634>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-20**



Repository / Repozitorij:

[EFOS REPOSITORY - Repository of the Faculty of Economics in Osijek](#)



Sveučilište Josipa Jurja Strossmayera u Osijeku

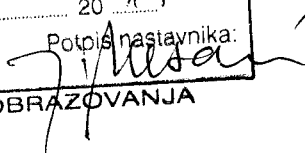
Ekonomski fakultet u Osijeku

Diplomski studij Poslovna Informatika

Nikolina Raguž

PROCES MINING: ISTRAŽIVANJE PODATAKA O PROCESIMA U SVRHU NJEGOVA POBOLJŠANJA

Diplomski rad

Diplomski rad iz predmeta	
UPRAVLJANJE POSLOVNIM PROCESIMA	
ocijenjen ocjenom	12,00 (5)
Osijek, 18.07.	20 19
Potpis nastavnika:	
	
II RAZINA OBRAZOVANJA	

Osijek, 2019

Sveučilište Josipa Jurja Strossmayera u Osijeku

Ekonomski fakultet u Osijeku

Diplomski studij Poslovna Informatika

Nikolina Raguž

**PROCES MINING: ISTRAŽIVANJE PODATAKA O
PROCESIMA U SVRHU NJEGOVA POBOLJŠANJA**

Diplomski rad

Kolegij: Upravljanje poslovnim procesima

JMBAG: 0010211929

e-mail: nikraguz@efos.hr

Mentor: prof. dr. sc. Josip Mesarić

Osijek, 2019

Josip Juraj Strossmayer University of Osijek
Faculty of Economics in Osijek
Graduate Study Business informatics

Nikolina Raguž

**PROCESS MINING: PROCESS DATA RESEARCH FOR THE
PURPOSE OF IMPROVEMENT**

Graduate paper

Osijek, 2019

IZJAVA

O AKADEMSKOJ ČESTITOSTI, PRAVU PRIJENOSA INTELEKTUALNOG VLASNIŠTVA, SUGLASNOSTI ZA OBJAVU U INSTITUCIJSKIM REPOZITORIJIMA I ISTOVJETNOSTI DIGITALNE I TISKANE VERZIJE RADA

1. Kojom izjavljujem i svojim potpisom potvrđujem da je diplomski
(navesti vrstu rada: završni / diplomski / specijalistički / doktorski) rad isključivo
rezultat osobnoga rada koji se temelji na mojim istraživanjima i oslanja se na objavljenu
literaturu. Potvrđujem poštivanje nepovredivosti autorstva te točno citiranje radova
drugih autora i referiranje na njih.
2. Kojom izjavljujem da je Ekonomski fakultet u Osijeku, bez naknade u vremenski i
teritorijalno neograničenom opsegu, nositelj svih prava intelektualnoga vlasništva u
odnosu na navedeni rad pod licencom *Creative Commons Imenovanje –
Nekomercijalno – Dijeli pod istim uvjetima 3.0 Hrvatska*. 
3. Kojom izjavljujem da sam suglasan/suglasna da se trajno pohrani i objavi moj rad u
institucijskom digitalnom repozitoriju Ekonomskoga fakulteta u Osijeku, repozitoriju
Sveučilišta Josipa Jurja Strossmayera u Osijeku te javno dostupnom repozitoriju
Nacionalne i sveučilišne knjižnice u Zagrebu (u skladu s odredbama Zakona o
znanstvenoj djelatnosti i visokom obrazovanju, NN br. 123/03, 198/03, 105/04, 174/04,
02/07, 46/07, 45/09, 63/11, 94/13, 139/13, 101/14, 60/15).
4. izjavljujem da sam autor/autorica predanog rada i da je sadržaj predane elektroničke
datoteke u potpunosti istovjetan sa dovršenom tiskanom verzijom rada predanom u
svrhu obrane istog.

Ime i prezime studenta/studentice: **Nikolina Raguž**

JMBAG: 0010211929

OIB: 48996630836

e-mail za kontakt: **nikraguz@efos.hr**

Naziv studija: **Diplomski studij Poslovna informatika**

Naslov rada: **Proces mining: istraživanje podataka o procesima u
svrhu njegova poboljšanja**

Mentor/mentorica rada: **prof.dr.sc. Josip Mesarić**

U Osijeku, 2019. godine

Potpis

Raguž Nikolina

Proces mining: istraživanje podataka o procesima u svrhu njegova poboljšanja

SAŽETAK

Mnoge tvrtke postaju svjesne važnosti upravljanja procesima i kako valjanim upravljanjem mogu brže i efikasnije postići željene ciljeve. Pošto su u današnje vrijeme sve popularnije tehnike rudarenja podataka, tvrtke pokušavaju pronaći stručnjake kako bi došli do potrebnih podataka, ali većina nije svjesna da postoji i proces mining, tj. rudarenje procesa. Dok druge tehnike rudarenja mogu otkriti podatke o tvrtki, proces mining ima mogućnost analizirati postojeći proces kojim se tvrtka bavi i poboljšati ga u različite svrhe (smanjenja troškova, bolje organizacije, bržeg izvođenja, automatizacije ili bolje kontrole izvođenja). Svrha ovoga rada je pobliže objasniti koncept proces mininga, a cilj je na procesu zaprimanja recenzija demonstrirati tehnike proces mininga. Korištena metodologija u ovom radu će od jednostavnijih pojmova u proces miningu oblikovati cijelu sliku i demonstrirati tehnike. U rezultatima istraživanja će se pobliže prikazati vrste proces mininga i koje tehnike je potrebno koristiti u kojoj fazi. Zatim će se dati detaljna analiza rezultata, predložiti će se određeni načini kako se proces može poboljšati i navest će se problemi s kojima se proces mining susreće. Na kraju je zaključeno da su prednosti istraživanja što se približio pojam proces mininga koji je još mlada disciplina i potrebno ga je više proučavati, a nedostaci su što se nisu koristili veći podaci na kojima bi se mogle demonstrirati naprednije tehnike. Ovako istraživanje bi moglo koristiti svima koji znaju nešto o upravljanju poslovnim procesima i htjeli bi svoje znanje unaprijediti, a može biti korisno i početnicima jer se rad bavi osnovnim tehnikama što nije pre komplicirano za svladati.

Ključne riječi: proces mining, zapis događaja, modeli, otkrivanje procesa, analiza procesa

Process mining: process data research for the purpose of improvement

ABSTRACT

Many companies are becoming aware of the importance of process management and how good management can achieve goals faster and more efficiently. As data mining techniques are increasingly popular these days, companies try to find experts to get the information they need, but most are unaware that process mining exists. While other mining techniques may reveal business information, process mining can analyse business practices and improve their processes. The purpose of this paper is to explain the concept of process mining, and the aim was to demonstrate process mining techniques on review process. The methodology used in this paper will use simple concepts to explain process mining and later shape the whole picture and demonstrate techniques. The results of the research will outline the types of process mining and what technique to use at which stage. Detailed analysis of the results will be provided, some steps for how the process can be improved will be suggested, and the problems that process mining will encounter will be outlined. In the end, it was concluded that the advantages of the research were bringing the concept of process mining, which is still a young discipline and needs to be studied more, as close as possible to the reader. The disadvantages are that no larger data were used, on which more advanced techniques could be demonstrated. This kind of research can be used by everyone who knows something about business process management and would like to have their knowledge expanded, it can also be useful for beginners because basic techniques have been used that are not too complicated to learn.

Keywords: process mining, event log, models, process discovery, process analysis

SADRŽAJ

1. Uvod	1
2. Teorijska podloga i prethodna istraživanja	2
2.1. Što je proces mining i kako utječe na poslovanje	2
2.1.1. Event Log	5
2.1.2. PetriNet	7
2.1.3. BPMN	8
2.1.3.1. Objekti toka	9
2.1.3.2. Objekti povezivanja	9
2.1.3.3. Plivaće staze	10
2.1.3.4. Artefakti	11
2.1.3.5. Prošireni skup modelskih elemenata	11
2.2. Vrste proces mininga	15
2.3. Dimenzije kvalitete	18
2.4. Područje primjene proces mininga i prethodna istraživanja	19
3. Metodologija rada	21
4. Opis istraživanja i rezultati istraživanja	22
4.1. Pogled na event log	22
4.2. Primjena tehnika proces mininga	25
4.2.1. Otkrivanje	25
4.2.2. Provjera usklađenosti	30
4.2.3. Poboljšanje	32
3.2.3.1. Ogranizacijska perspektiva	34
3.2.3.2. Vremenska perspektiva	36
3.2.3.3. Perspektiva slučaja	39
5. Rasprava	41
6. Zaključak	44
Literatura	45
Popis tablica	47
Popis slika	48

1. Uvod

Razvojem tehnologije počeli su se pojavljivati sve napredniji alati i tehnike kako bi se što brže i učinkovitije došlo do potrebnih informacija kojih iz dana u dan ima sve više. Povećanje broja informacija u svijetu uzrokovalo je pojavljivanje znanosti o podacima. Znanost o podacima zadužena je za prikupljanje i analizu podataka, i pošto podataka ima mnogo, razvile su se pod discipline u tom području u koje pripada i proces mining. Proces mining ima veze sa svim tehnikama znanosti o podacima i potrebno je za učinkovito upravljanje tvrtkom i efikasno ostvarivanje ciljeva.

Predmet istraživanja ovog rada je proces mining i kako se upotrebom istog može doći do relevantnih podataka vezanih za tvrtku i njene procese. Proces mining ima mogućnost analize procesa i ukazivanja na greške koje se događaju tijekom izvođenja ili na greške koje su napravljene tijekom dizajniranja procesa. Brojna istraživanja su pokazala kako učinkovitim vođenjem procesa tvrtka ima mogućnost ostvariti konkurentsku prednost, i upravljanjem procesima igra važnu ulogu. Iako se brojim istraživanja ukazala važnost proces mininga, mnoge tvrtke nisu upoznate s tim pojmom ili jednostavno odbijaju provesti nove tehnike nad svojim sustavom iz razloga što nisu spremni vidjeti na koji način mogu poboljšati svoj proces. Proces mining kao područje znanosti o podacima je istovremeno i problematično primijeniti jer su potrebni posebni stručnjaci za provođenje. Stručnjaci za proces mining moraju biti upoznati s metodama rudarenja podataka, strukturom cijelog sustava i menadžmentom poslovnih procesa. Oni zapravo moraju znati mnogo različitih disciplina kako bi uspješno napravili analizu poslovnog procesa.

Cilj ovog rada je na primjeru prikazati i objasniti tehnike proces mininga koje mogu pomoći pri analizi prednosti i nedostataka procesa. Tehnike proces mininga se mogu napraviti u specifičnim programskim alatima, a u ovom radu će biti korišten alat ProM te će se prikazati njegova funkcionalnost.

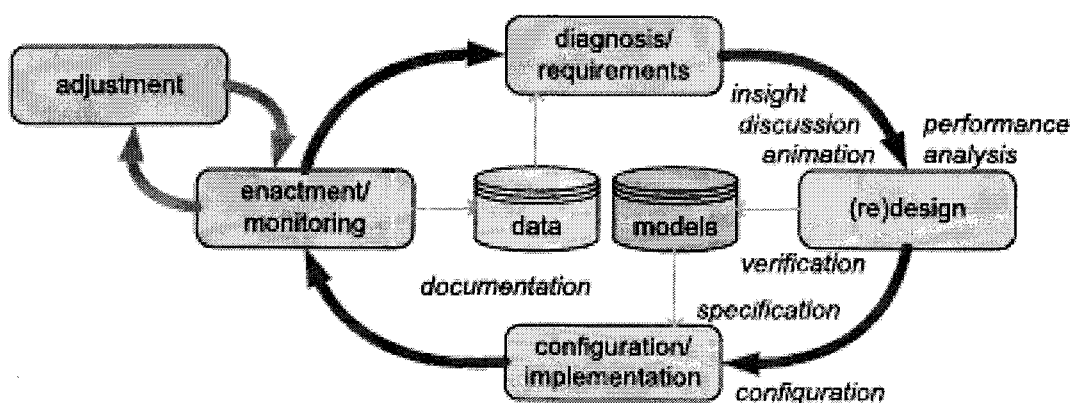
2. Teorijska podloga i prethodna istraživanja

Kako bi se u potpunosti objasnio proces i zaključile prednosti i nedostaci potrebno je definirati određene pojmove koji su bitni za proces i proces mining. Iz tog razloga u ovome dijelu će biti definiran proces mining i objašnjeni dijelovi potrebni za razumijevanje procesa.

2.1. Što je proces mining i kako utječe na poslovanje

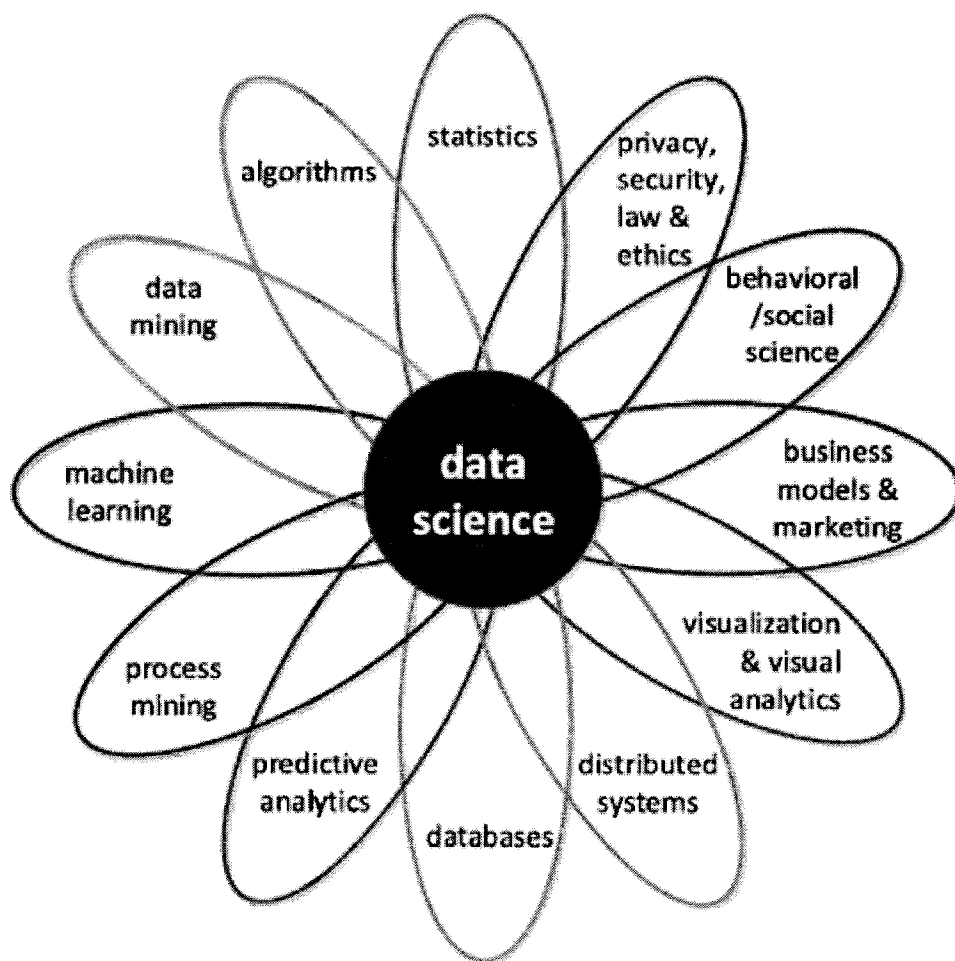
„Proces mining je skupina tehnika za analizu performansi i sukladnosti poslovnih procesa na temelju zapisa događaja nastalih tijekom njihovog izvođenja“ (Dumas i dr., 2018:419).

Proces mining dio je BPM-a (Business Proces Management), a BPM je „umjetnost i znanost nadzora nad radom organizacije u cilju postizanja konzistentnih rezultata i iskorištavanja mogućnosti poboljšanja“ (Dumas i dr., 2018:1). Kako bi se proces mining uspješno pozicionirao unutar određene tehnike, potrebno je prvo razjasniti životni ciklus BPM-a koji je vidljiv na slici 1. Životni ciklus BPM-a započinje dizajnom procesa u dizajn fazi. Nakon dizajniranja modela potrebno ga je konfigurirati i implementirati u određeni sustav. Nakon implementiranja dizajna u sustav potrebno je provjeravati odgovara li on i ako je potrebno napraviti određene izmjene. Ako je ustanovljeno da je sve u redu nastavlja se na fazu dijagnoze gdje se evaluira proces i provjeravaju dodatni zahtjevi. Ako je performansa procesa loša, životni ciklus BPM-a napravi puni krug i slijedi redizajn procesa. Proces mining može iznimno pomoći u BPM životnom ciklusu jer može povezati modele koji su potrebni za dizajn i implementaciju, i podatke koji su potrebni za nadgledanje i dijagnozu.



Slika 1. BPM životni ciklus (prema: Aalst, 2018)

Proces mining također je i dio znanosti o podacima i bavi se analizom podataka u svrhu poboljšanja procesa u tvrtkama. Kada se spominje znanost o podacima u većini slučajeva se misli na rudarenje podataka, statistiku ili strojno učenje, ali znanost o podacima u sebi sadrži puno više polja za proučavanje koji su bitni za funkcioniranje tvrtke. Znanost o podacima je spoj različitih, preklapajućih disciplina, što je vidljivo na slici 2 koja želi prikazati odnos između disciplina, ali one ne moraju tako izgledati. Neka područja mogu biti veća nego druga, ili se čak mogu više ili manje preklapati.



Slika 2. Discipline u znanosti o podacima (prema: Aalst, 2018)

Iz slike je također vidljivo da je proces mining ključni dio znanosti o podacima. Proces mining spaja tradicionalne analize bazirane na modelu i tehnike za analizu bazirane na podacima, i ta tehnika pokušava spojiti podatke o događajima (tj. promatrano ponašanje) i modele procesa (ručno radene ili automatski otkrivene). „Proces mining se može smatrati kao poveznica između strojnog učenja i rudarenja podataka, i modeliranje procesa i analize“ (Aalst, 2018:31). Razlog zašto se proces mining treba koristiti uz ostale tehnike je taj što rudarenje podacima, statistika i strojno učenje ne uzimaju u obzir cjelovite modele procesa, a znanost o procesima je centrirana na proces ali se često usredotoči na modeliranje nego na učenje iz podataka. Pozicija proces mininga unutar svih tehnika čini ga moćnim alatom za iskorištavanje stalno rastuće količine podataka u svrhu unapređenja cjelovitih procesa.

Cilj proces mininga je pretvoriti podatke u uvide i akcije. Proces mining je relativno mlada disciplina koja se nalazi između strojnog učenja i rudarenja podataka. Ideja proces mininga je otkriti, promatrati i poboljšati postojeće poslovne procese koristeći event logove. Proces mining se može koristiti kako bi se provjerila usklađenost u procesu, otkrila odstupanja, precizirala „uska grla“ itd. Autori W. van der Aalst i dr. (2012) smatraju da postoje određene zablude vezane za proces mining. Naglašavaju da određeni proizvođači softvera, analitičari i istraživači limitiraju upotrebu proces mininga na posebne tehnike za rudarenje podataka koje se mogu koristiti samo za slučajeve koji su gotovi i koji su offline. Iz tog razloga autori naglašavaju slijedeće tri karakteristike proces mininga:

Proces mining nije limitiran samo na otkrivanje toka procesa. Otkrivanje toka procesa je najčešći način korištenja proces mininga i možda najzabavniji, ali nije limitirano samo na tu perspektivu. Proces mining može i analizirati podatke vezane za proces. Osim otkrivanja toka procesa, veliku ulogu u proces miningu igraju i organizacijska, vremenska i perspektiva slučaja. Proces mining nije samo specifična vrsta rudarenja podataka. Proces mining može koristiti tehnike rudarenja podataka i davati slične analize, ali to je samo dio proces mininga. Tehnike za rudarenje podataka nisu potaknute procesom, dok proces mining jeste. Iz tog razloga je u proces miningu potrebno koristiti potpuno nove tehnike uz one koje se koriste u rudarenju podataka.

Proces mining nije limitiran samo na offline analizu. Iako se u proces miningu koriste podaci iz procesa koji je završen, rezultati analize tih podataka mogu se koristiti na procesima koji su trenutno u tijeku.

2.1.1. Event Log

Kada se proces odvija na BPMS-u (Business Process Management Software) ili nekom drugom odgovarajućem programu, sustav bi trebao imati mogućnost koordinacije zaposlenika i individualnih slučajeva u tvrtki. Ako se tvrtka koristi takvim sustavom, zaposlenici će imati zadane zadatke koje mogu pregledati u tom sustavu i koje moraju obaviti. Zaposlenici bi zatim trebali zabilježiti svoju aktivnost u sustavu, upisati što su radili ili makar označiti u sustavu da su obavili aktivnost koju im je sustav zadao. Na taj način tvrtka može pratiti kako se odvija proces. Nakon korištenja takvih sustava iz baze podataka se mogu izvući podatci o aktivnostima u procesu, zapisi događaja tj. event logovi. „Event log je kolekcija zapisa vremenski označenih događaja. Svaki zapis događaja govori nešto o izvršenju radnog predmeta (a time i zadatka) procesa (npr. da je zadatak započeo ili dovršen) ili nam govori da se događaj, eskalacija ili drugi relevantni događaj dogodio u kontekstu danog slučaja u procesu“ (Dumas i dr., 2018:422). Proces mining koristi zapise događaja koji su u stvarnosti ispis događanja u sustavu. Oni se mogu pronaći svugdje i mogu se koristiti u različite svrhe. Jednostavni zapisi događaja najčešće su prikazani u tabličnom formatu i spremljeni su u CSV (Comma-Separated-Values) formatu. Za velike i kompleksne zapise događaja CVS format nije dovoljan i za takve zapise češće se koristi XES (eXtensible Event Stream) format. U zapisu događaja dobiva se uvid u slijedeće: koja aktivnost je obavljena, tko je obavljao aktivnost, u koje vrijeme je aktivnost započela, u koje vrijeme je završila itd. Svi ti podaci pomažu pri otkrivanju zastoja u sustavu i prikazuju gdje sustav treba poboljšanja. Mnogi jednostavni zapisi događaja mogu se pretvoriti iz CSV u XES format, ali u većini slučajeva podaci koji su potrebni za rudarenje procesa nisu direktno dostupni u tim zapisima, nego se trebaju izvući iz drugih izvora. Dumas i dr. (2018) smatraju kako se u pribavljanju podataka mogu identificirati četiri izazova: izazov korelacije, izazov vremenske oznake, izazov dugovječnosti i izazov zrnatosti. Izazov korelacije odnosi se na problem identificiranja kojem događaju pripada neki slučaj. U procesu je potrebno istražiti koji atribut može služiti kao identifikator slučaja, pošto sustavi u većini slučajeva nemaju točno definiran postupak. Izazov vremenske oznake odnosi se na problem što sustavima nije primaran zadatak zapisivati vrijeme, i može doći do situacije gdje su dva zadatka zabilježena u isto vrijeme, a u stvarnosti je sustav kasnije zabilježio zadatak jer je čekao vrijeme kada nema puno događanja. Izazov dugovječnosti se odnosi na miješanje gotovih i ne gotovih slučajeva i dobivanje krivih rezultata. Izazov zrnatosti se odnosi na raščlambu jednog zadatka. Postoji mogućnost da se u sustavu zabilježi jedan zadatak, a taj zadatak se u stvarnosti sastoji od mnogo pod-zadataka koje je potrebno izvršiti. Teško je definirati precizno mapiranje takvih zadataka.

Postoje određeni izazovi koje treba riješiti kako bi se dobio savršeni zapis događaja. Zapis događaja može biti nepotpun tj. zapis događaja može sadržavati premalo događaja i nije moguće dobiti točan model iz takvih podataka. Zapis događaja može u sebi sadržavati i nepotrebne podatke kao što su rijetke aktivnosti koje nisu reprezentativne za ukupni proces.

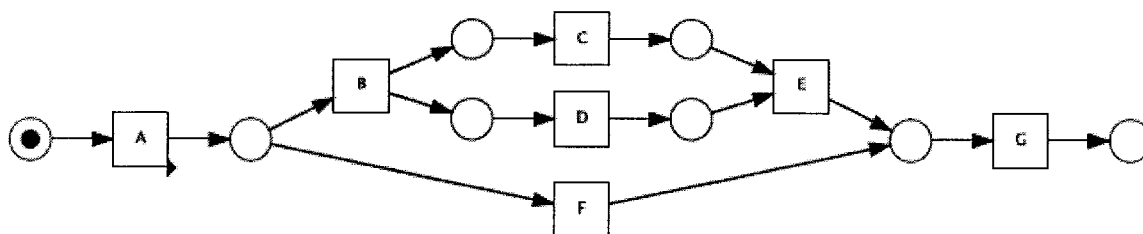
Zapis događaja se učitava u program (npr. ProM ili Disco) i pomoću algoritama se obrađuju podaci. Kada program obradi podatke, podaci mogu biti vizualizirani putem tablica, grafikona i dijagrama. Najčešća je uporaba dijagrama jer se na njima vide rezultati na pregledan i jednostavan način. Najčešće korišteni dijagrami su PetriNet i BPMN dijagrami.

Case id	Event id	Properties				
		Timestamp	Activity	Resource	Cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	register request	Pete	50	...
	35654522	30-12-2010:15.06	examine casually	Mike	400	...
	35654524	30-12-2010:16.34	check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	decide	Sara	200	...
	35654526	06-01-2011:12.18	reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	check ticket	Pete	100	...
	35654531	09-01-2011:09.55	decide	Sara	200	...
	35654533	15-01-2011:10.45	pay compensation	Ellen	200	...

Slika 3. Event log (prema: Aalst, 2018)

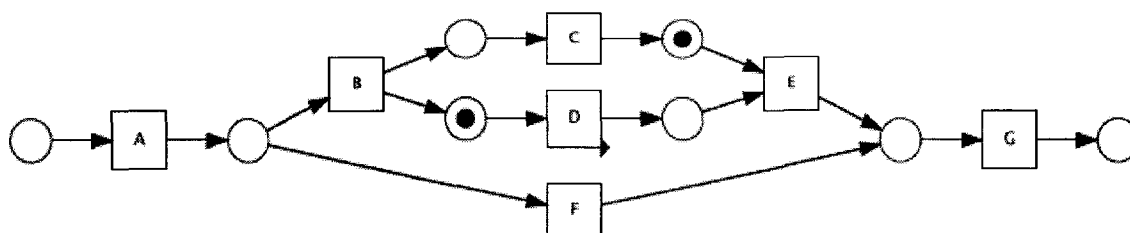
2.1.2. PetriNet

PetriNet (Petrijeve mreže) je naziv za grafički jezik (grafički alat) koji se koristi u svrhu reprezentacije procesa. Ako je poznato kako se određeni proces izvršava i poznate su sve aktivnosti, on se može prikazati PetriNet dijagramom radi lakše analize i boljeg razumijevanja toka procesa. PetriNet se sastoji od dva elementa: tranzicija i mjesta. Tranzicije predstavljaju aktivnosti koje mogu biti izvršene, a mjesta predstavljaju stanje (prijelazno ili konačno) koje proces može postići. Na slici 4 je vidljiv primjer nekog procesa prikazan pomoću PetriNet dijagrama, i unutar tog modela krugovi predstavljaju mjesta, a kvadrati tranzicije. PetriNet model se također sastoji od mjesta gdje se aktivnosti razdvajaju. Takva mjesta mogu imati AND (paralelno) ili XOR (ekskluzivno) grananje. AND grananje predstavlja grananje dviju ili više aktivnosti koje se zatim paralelno izvode. Na primjeru koji je vidljiv na slici 4, AND grananje se događa nakon aktivnosti B. Nakon što se izvrši aktivnost B može se izvršiti aktivnost C, zatim D ili obrnuto, i nakon izvršenja obje aktivnosti može se izvršiti aktivnost E. XOR grananje predstavlja međusobno isključivanje dvaju ili više aktivnosti. Na slici 4 XOR grananje se odvija nakon aktivnosti A. XOR grananje označava da postoji mogućnost izvođenja ili aktivnosti B ili aktivnosti F. Nakon odabira aktivnosti za izvršenje, druga aktivnost se ignorira i proces nastavlja odabranom putanjom. Ako se npr. odabere izvršenje aktivnosti F, aktivnost B se ignorira. Model izvodi aktivnost F, pa zatim G.



Slika 4. PetriNet (prema: Burattin, 2015)

PetriNet funkcionira na principu tokena i pravilu pokretanja. Tokeni se postavljaju u tranzicije kako bi se mogla vidjeti pravilnost modela i kojom kojim tranzicijama model šalje tokene. „Tranzicija unutar PetriNeta može biti pokrenuta ako sva ulazna mjesta (mjesta s krajevima koji ulaze u tranziciju) sadrže barem jedan token“ (Burattin, 2015:14). Na slici 4 je vidljiv model koji je identičan kao i na slici 4, samo što je model na slici 5 „pokrenut“ i u sebi sadrži tokene.



Slika 5. Tokeni unutar PetriNeta (prema: Burattin, 2015)

ProM alat i njegovi plug-inovi u većini vremena će modele prikazivati kao PetriNet zbog matematičke pozadine i zbog lakšeg prikaza. Određeni plug-inovi pružaju mogućnost i prikaza u BPMN modelu, ali u većini slučajeva ako se želi prikazati BPMN potrebno je iskoristiti plug-in koji će zadani PetriNet model pretvoriti u BPMN.

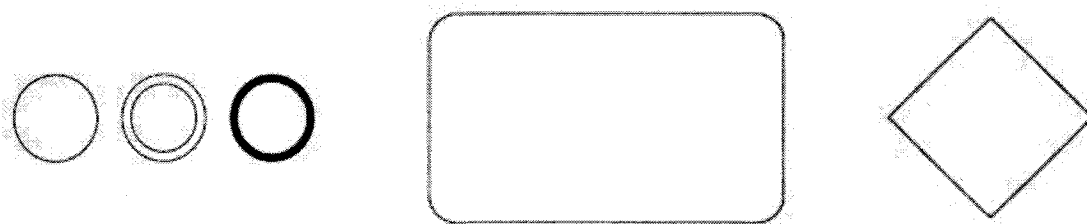
2.1.3. BPMN

BPMN (Business Process Modeling Notation) je druga vrsta modela koja se koristi u proces miningu, iako je češće korišten PetriNet. „Primarni cilj BPMN-ovog razvoja bio je pružiti zapis koji je razumljiv svim poslovnim korisnicima, od poslovnih analitičara koji stvaraju početne nacрте procesa, do tehničkih programera odgovornih za implementaciju tehnologije koja će ih izvoditi i, konačno, poslovnim ljudima koji će upravljati i nadzirati te procese“ (White, 2004:1). BPMN model je nešto teže za naučiti i razumjeti nego PetriNet jer ima više komponenata, ali s njim je lakše analizirati model i predstaviti poslovnim korisnicima kako bi ga bolje razumjeli.

BPMN ima mogućnost nositi se sa vrlo kompleksnim poslovnim procesima i istovremeno kreirati jednostavne modele za takve procese. BPMN se, isto kao i PetriNet, sastoji od određenih elemenata. Četiri glave kategorije elemenata su: objekti toka, objekti povezivanja, plivaće staze i artefakti. Postoji još i prošireni skup modelskih elemenata.

2.1.3.1. Objekti toka

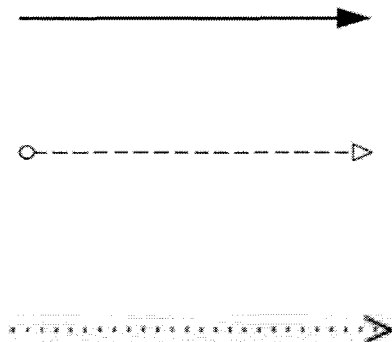
Objekti toka vidljivi na slici 6 se sastoje od tri elementa: događaj, aktivnost i skretnica. Događaj je označen krugom i prikazuje „nešto“ što se događa u procesu. Aktivnosti su prikazane kvadratom sa zaobljenim rubovima i generički je izraz za posao koji neka tvrtka obavlja. Skretnica se prikazuje oblikom dijamanta i koristi se za kontrolu razdvajanja i spajanja dijelova procesa. Isto kao i PetriNet sastoji se od AND (paralelnih) i XOR (ekskluzivnih) skretnica, ali BPMN dodaje još jednu OR (inkluzivnu) skretnicu, gdje točan alternativni put ne isključuje ostale tj. svi alternativni putevi se mogu uzeti u obzir.



Slika 6. Objekti toka: događaj, aktivnost i skretnica (prema: White, 2004)

2.1.3.2. Objekti povezivanja

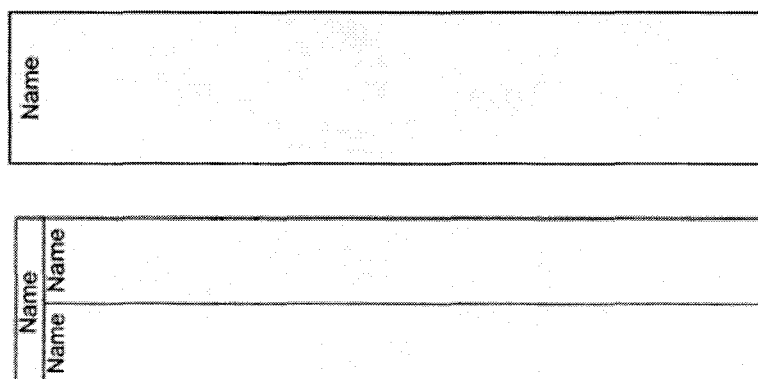
Objekti povezivanja u BPMN dijagramu koriste se za povezivanje događaja i aktivnosti kako bi model imaju osnovnu strukturu. Objekti povezivanja su vidljivi na slici 7. BPMN se sastoji od tri objekata povezivanja: sekvencijski tok, tok poruke i asocijacija. Sekvencijski tok unutar BPMN-a prikazuje se punom linijom i ispunjenom strelicom na vrhu. Koristi se za prikaz redoslijeda izvršavanja aktivnosti. Tok poruke se prikazuje iscertanom linijom i praznom strelicom na vrhu. Tok poruke koristi se za prikaz razmjene poruka između dva sudionika u procesu. Asocijacija se prikazuje točkastom linijom i ona se koristi za asocijaciju podataka, teksta i drugih artefakata sa objektima toka.



Slika 7. Objekti povezivanja: sekvencijski tok, tok poruke i asocijacija (prema: White, 2004)

2.1.3.3. Plivaće staze

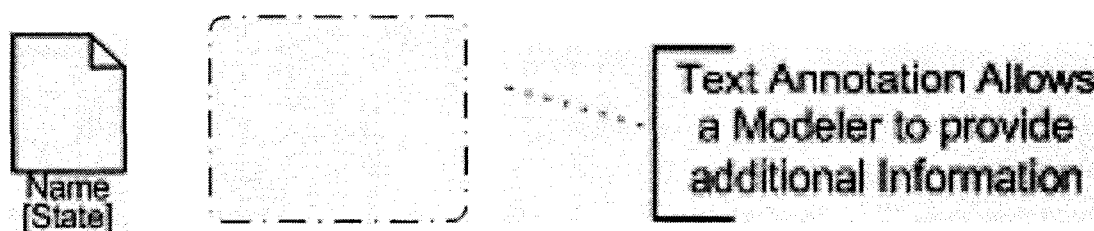
Mnoge metodologije za modeliranje koriste plivaće staze kao mehanizam za organiziranje procesa u različite kategorije radi bolje analize. BPMN također podržava plivaće staze sa dva glavna elementa: polje ili bazen i staza, vidljivi na slici 8. Polje ili bazen predstavlja sudionika u procesu. Djeluje kao spremnik za sve aktivnosti koje jedan sudionik odrađuje. Bazeni se koriste kada u procesu sudjeluju dva ili više sudionika i oni su uz pomoć bazena razdvojeni. Staza je particija bazena i dugačka je koliko i bazen. Ona se koristi za organizaciju i kategorizaciju aktivnosti unutar procesa.



Slika 8. Plivaće staze: polje (bazen) i staze (prema: White, 2004)

2.1.3.4. Artefakti

Artefakti su dizajnirani kako bi omogućili fleksibilnost u modeliranju procesa i kako bi se moglo dodati potrebno objašnjenje modelu. Elementi, vidljivi na slici 9, koji pripadaju artefaktima su: podatkovni objekt, grupa i anotacija. Podatkovni objekti su mehanizmi koji prikazuju koji podaci potrebni aktivnostima ili koje podatke aktivnost „konzumira“ tj. koristi. Grupa je označena iscrtanim kvadratom i može se koristiti za grupiranje analize ili dokumentacije. Anotacija služi za pružanje dodatnih informacija o dijagramu



Slika 9. Artefakti: podatkovni objekt, grupa i anotacija (prema: White, 2004)

2.1.3.5. Prošireni skup modelskih elemenata

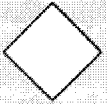
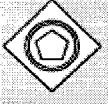
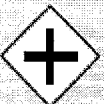
Osim glavnim skupinama elemenata, BPMN dijagram se može još opisivati i sa proširenim skupom modelskih elemenata. Prošireni skup predstavlja elemente koji pripadaju glavnim kategorijama, ali imaju posebnu svrhu korištenja i proširuju mogućnost upotrebe glavnih elemenata. U prošireni skup modelskih elemenata pripadaju prošireni događaji, proširene skretnice i proširene aktivnosti. Prošireni skup događaja se, kao i osnovni elementi, označavaju krugom, ali u sredini imaju određeni znak koji predstavlja posebnu upotrebu takvog događaja. U tablici 1 vidljiv je prošireni skup događaja i pripadajući opis za svaki element. Navedeni su početni, intermedijarni i finalni događaji. Početni događaji prikazuju gdje proces ili određeni događaji započinju. Intermedijarni događaji se odvijaju između početnog i završnog događaja. Oni utječu na tijek procesa ali ne mogu započeti ili direktno završiti proces. Finalni događaji prikazuju gdje proces ili određeni događaji završavaju. U tablici je prikazan i izgled događaja pri prihvatanju i odašiljanju. Početni događaji i određeni intermedijarni događaji mogu reagirati na okidač i tako započeti proces, dok finalni događaji i drugi dio intermedijarnih događaja stvaraju rezultat i tim putem označavaju završetak procesa ili događaja.

Tablica 1. Prošireni skup događaja (prema: Mesarić, Šebalj: 2017)

	Početni	Intermedijarni	Finalni		
	Prihvaćanje		Odašiljanje		
Opći					Opći događaji - pokazuju gdje proces započinje i završava
Poruka					Primanje i slanje poruka
Timer					Ciklični vremenski događaj, započinje u neko vrijeme, traje ili završava
Pogreška					Prihvaćanje i izdavanje događaja u kojem se javlja pogreška
Poništavanje					Reakcija kojom se prekida ili inicira prekid transakcije
Kompenzacija					Rukovanje kompenziranjem ili započinjanje kompenziranja
Uvjeti					Reakcija na promijenjene uvjete ili integraciju poslovnih pravila
Signal					Signaliziranje kroz procese. Odašlani signal može se prihvatiti više puta
Višestruki					Prihvaćanje ili odašiljanje jednog od više mogućih događaja
Veza					Prekidni poveziivači . Upotrjebljeni zajedno jednaki su sekvencijskom toku
Prekid					Okida neposredni prekid procesa

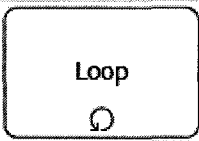
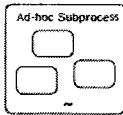
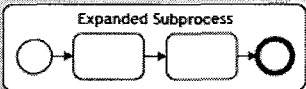
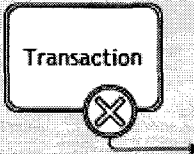
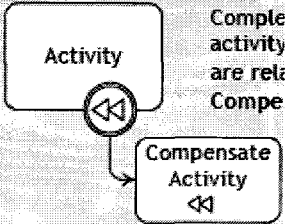
Prošireni skup skretnica se također odnosi na dodatne vrste skretnica koje se mogu koristiti u BPMN dijagramu. Dodatne vrste skretnica prikazane su u tablici 2. U tablici je prikazan izgled osnovnih skretnica (ekskluzivna, inkluzivna i paralelna) i dodatnih skretnica koje se mogu koristiti u dijagramu pri modeliranju kompleksnijih procesa.

Tablica 2. Prošireni skup skretnica (prema: Mesarić, Šebalj: 2017)

		Podatkovno bazirana ekskluzivna odluka. Kada razdjeljuje, usmjerava sekvencije toka na samo jednu od mogućih. Pri spajanju čeka se jedna od mogućih grana prije okidanja daljnjeg toka
		Na događaju bazirana ekskluzivna odluka. Slijedi nakon prihvata događaja ili primitka zadatka. Sekvencijski tok se usmjerava prema slijednom događaju/zadatku koji se događa prvi.
		Paralelna skretnica (odluka) Kada se koristi za razdjeljivanje sekvencijskog toka, sve izlazne grane se aktiviraju istovremeno. Kada se spajaju paralelne grane, čeka se na sve dolazeće grane da završe prije nego li se okine izlazni tok
		Inkluzivna skretnica (odluka) Kada se koristi za razdjeljivanje, jedna ili više grana se aktiviraju na temelju uvjeta grananja. Pri spajanju čekaju se sve aktivne za kompletiranje
		Kompleksna skretnica (odluka) Aktivira jednu ili više grana na temelju kompleksnih uvjeta ili verbalnog opisa. Treba se koristiti s oprezom zbog nejasne semantike

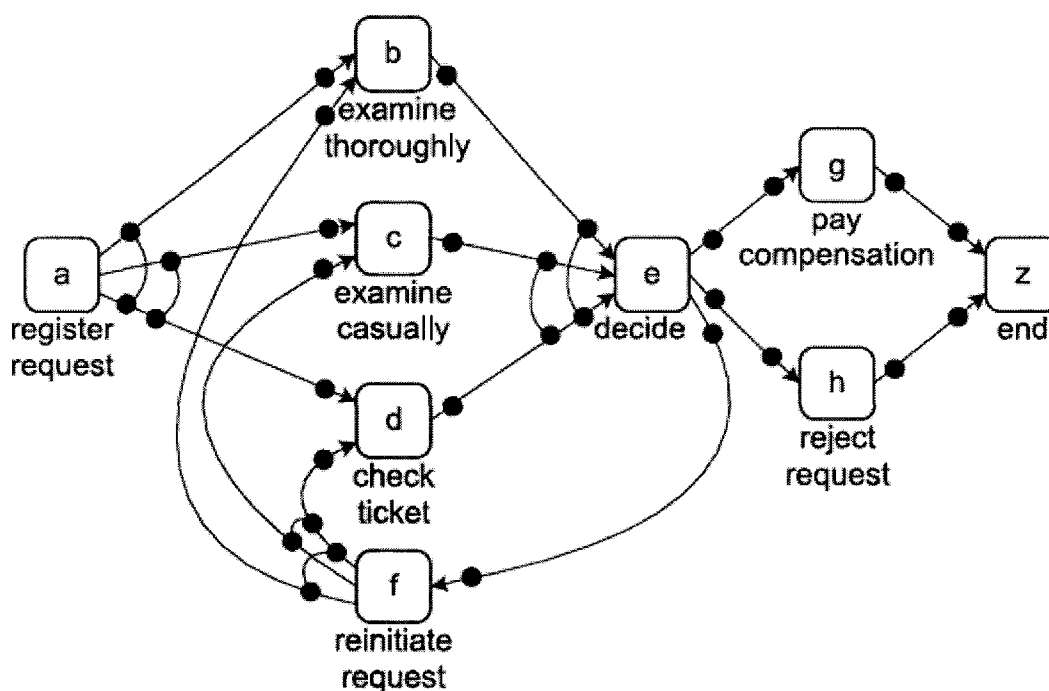
Prošireni skup aktivnosti i transakcija vidljiv je u tablici 3. Tablica prikazuje dodatne vrste aktivnosti koje se mogu koristiti u modeliranju procesa kako bi se grupirale određene aktivnosti i prikazuje transakcije koje pripadaju aktivnostima ali se koriste za transakcijske događaje.

Tablica 3. Prošireni skup aktivnosti i transakcija (prema: Mesarić, Šebalj: 2017)

	Višestruka instanca iste aktivnosti koja započinje ili paralelno ili sekvencijski za svaki pojedini slučaj
	Aktivnost u petlji koja se višestruko ponavlja dok je uvjet za ponavljanje zadovoljen. Uvjet se može testirati prije ili nakon izvršenja aktivnosti
	Ad-hoc podproces sadrži samo zadatke. Svaki se zadatak može izvršiti arbitrarno, često nakon što se ispune određeni uvjeti
	Zadatak (task) je jedinica rada, atomarna aktivnosti, posao koji se mora izvršiti
	Sažeti podproces – sadrži više aktivnosti koje se mogu vidjeti kad se raširi (ekspandira)
	Ekspandirani podproces – sadrži valjane oznake BPMN-a
	Transakcija je skup aktivnosti koje logički pripadaju zajedno. Mogu slijediti nakon određenog transakcijskog protokola
	Pridruženi događaj za poništenje ukazuje na reakciju otkazivanja (poništenja). Aktivnosti unutar transakcije se kompenziraju nakon poništenja
	Kompletirana aktivnost može se kompenzirati. Aktivnost i kompenzirana aktivnost se odnose uz dodavanje intermedijarnog kompenzacijskog događaja

2.2. Vrste proces mininga

Proces mining se može, obzirom na svrhu provođenja podijeliti na tri vrste: otkrivanje, provjera usklađenosti i poboljšanje. Prva vrsta proces mininga je otkrivanje. Tehnika otkrivanja procesa koristi zapis događaja i pokušava na tom temelju napraviti model bez prethodnog unošenja dodatnih informacija. Postoji više vrsta tehnika s kojima se može provesti faza otkrivanja kao što su: α -algoritam, heuristic miner, genetički proces mining, inductive mining i mining baziran na regiji. Najčešća i najlakša tehnika za provesti je α -algoritam, a u ovom radu pažnja će biti posvećena α -algoritmu, heuristic mineru i inductive mineru. α -algoritam „uzima event log i proizvodi PetriNet koji objašnjava ponašanje koje se nalazi u event logu“ (Aalst, 2016:33). α -algoritam ima mogućnosti modelirati PetriNet bez dodatnih informacija. Ako se u event logu nalaze i informacije poput resursa (npr. osoba koja obavlja neku aktivnost), α -algoritam može otkriti modele koji su vezani za resurse, npr. koji resurs obavlja najviše aktivnosti u jednom danu. α -algoritam ne uzima u obzir frekvencije događanja aktivnosti, stoga je jednostavan za korištenje. Inductive miner je napredna tehnika otkrivanja modela i može se nositi sa velikim modelima i ne toliko čestim događanjima u event logu. Inductive miner je trenutno najpopularnija tehnika za otkrivanje modela jer je fleksibilan, skalabilan i pruža određene kriterije ispravnosti koje druge tehnike ne mogu pružiti. Heuristic miner koristi prikaz podataka sličan uzročnim mrežama koja je vidljiva na slici 10 („Uzročna mreža je grafikon na kojem čvorovi predstavljaju aktivnosti, a lukovi predstavljaju uzročne ovisnosti. Svaka aktivnost ima skup mogućih veza ulaza i skup mogućih izlaznih veza“ (Aalst, 2016:72).) Ovakvi algoritmi u obzir uzimaju i frekvencije događanja i aktivnosti za modeliranje. Ideja heuristic minera je da se putovi u modelu koji nisu često korišteni (npr. ako zaposlenici imaju izbor za obaviti aktivnost A ili B, češće će odabrati aktivnost A, onda se aktivnost B smatra da nije toliko često korištena.) ne bi trebali prikazati u modelu. Upotreba frekvencija čini pristup mnogo robusnijim od većine drugih pristupa.

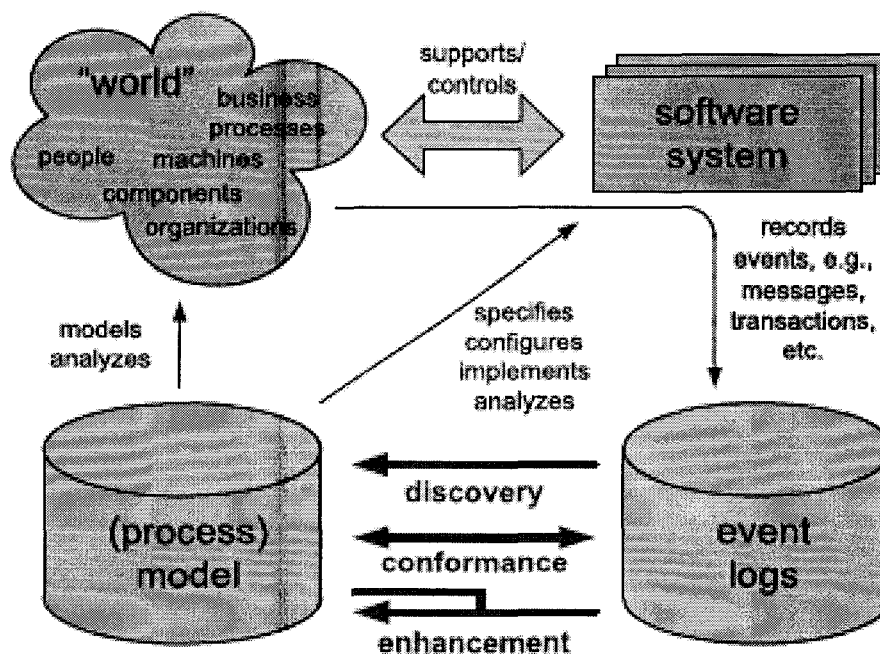


Slika 10. Primjer uzročne mreže (prema: Aalst, 2018)

Druga vrsta proces mininga je provjera usklađenosti. Ovdje se postojeći model procesa uspoređuje sa event logom od istog procesa. Provjera usklađenosti se koristi kako bi se stvarnost (koja je zabilježena u event logu) usporedila sa postojećim modelom i obrnuto. Koristi se također za pronalaženje, lociranje i objašnjavanje devijacija, te mjerenje ozbiljnosti tj. jačine tih devijacija. Provjera usklađenosti može se izvršiti uz pomoć tzv. „uzročnih otisaka“ i reprodukcije tokena. Te dvije tehnike se mogu ručno izvršiti na manjim event logovima pošto je potrebno bilježiti redoslijed obavljanja aktivnosti i izračunati fitness. Pošto je event log koji je korišten u radu pre velik za ručno izračunavanje, koristit će se plug-in u ProM-u koji može automatski provjeriti usklađenosti.

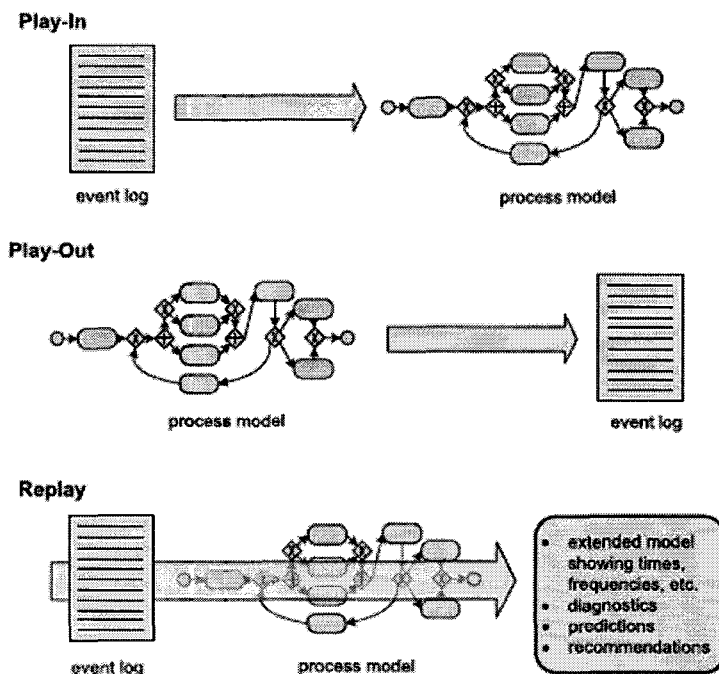
Treća vrsta proces mininga je poboljšanje. Ideja poboljšanja je proširiti ili usavršiti postojeći model koristeći informacije o stvarnom procesu koje su pohranjene u event logu. Poboljšanje je slično kao provjera usklađenosti, ali provjera usklađenosti se odnosi na usklađivanje modela i stvarnosti, dok je cilj poboljšanja promijeniti ili proširiti postojeći model. Jedna vrsta poboljšanja je popravak. Popravak se odnosi na izmjenjivanje modela kako bi bolje reflektirao stvarnost tj. ono što se nalazi u event logu. Druga vrsta poboljšanja je proširenje. Proširenje se odnosi na dodavanje nove perspektive u modelu kako bi više odgovarao stvarnosti.

Nakon definiranja svih vrsta proces mininga, moguće ih je pozicionirati. Na slici 11 vidljiva je pozicija svih vrsta proces mininga, i one se nalaze između zapisa događaja i procesa tj. modela. Korištenjem svih vrsta proces mininga zapisi događaja i modeli se mogu u potpunosti povezati kako bi proces uspješno funkcionirao.



Slika 11. Pozicioniranje vrsta proces mininga (prema: Aalst, 2018)

Osim vrsta proces mininga, postoje i tri načina korištenja proces mininga: play-out, play-in i replay. Play-out se odnosi na klasično korištenje proces mininga. Cilj play-outa je uz pomoć modela (PetriNet ili BPMN dijagram) generirati ponašanje. Play-in je suprotni proces od play-out, iz event loga se stvara model koji prikazuje ponašanja. Replay koristi event log i model kao input. Event log se „ponovno odigrava“ preko modela za različite svrhe kao što su provjera usklađenosti event loga i modela, proširenje modela frekvencijama i vremenskim informacijama, izgradnju prediktivnih modela i operativna podrška. Slika 12 prikazuje tri načina korištenja.



Slika 12. Način korištenja proces mininga (prema: Aalst, 2018)

2.3. Dimenzije kvalitete

Postoje određena pravila kojih se treba pridržavati kada se žele upotrijebiti tehnike proces mininga, a to su ostvarivanje dimenzija kvalitete. Postoje četiri dimenzije kvalitete, a to su sposobnost, preciznost, generalizacija i jednostavnost i prikazane su na slici 13.

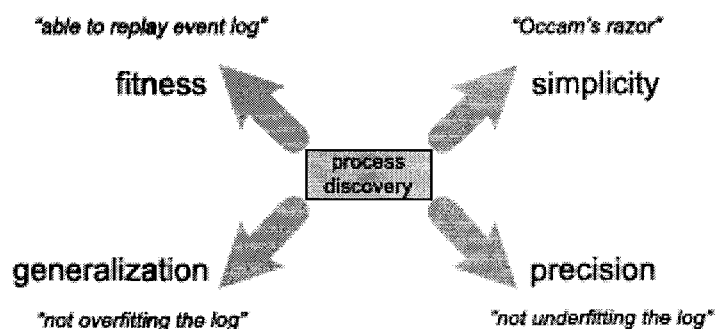
Sposobnost modela se odnosi na prikaz ponašanja koja su zabilježena u event logu. Model sa dobrom sposobnošću prikazat će ponašanja iz event loga. Model ima savršenu sposobnost ako se sva ponašanja iz event loga mogu ponoviti u modelu od početka do kraja.

Jednostavnost modela se odnosi na Occamovu oštricu. To je „princip koji se pripisuje logističaru iz 14. stoljeća Williamu od Ockhama. Princip govori da ne bi trebalo povećavati, iznad onoga što je dovoljno, broj entiteta koji su potrebni za objašnjavanje ičega, npr. potrebno je tražiti najjednostavniji model koji može objasniti ono što je promatrano u event logu“ (Aalst, 2016:120). Dakle, jednostavan model je onaj koji može objasniti ponašanje koje je viđeno u event logu.

Model je precizan ako ne dozvoljava „previše“ ponašanja. Model koji nije precizan previše generalizira ponašanje u event logu kao npr. dozvoljava ponašanje koje je bitno drukčije nego što su ponašanja u event logu.

Model treba generalizirati i ne ograničavati ponašanja prikazana u event logu. Dakle, generalizacija je suprotna od preciznosti, i potrebno je između dvije dimenzije pronaći „zlatnu sredinu“.

Općenito govoreći, nijedan model ne može u potpunosti zadovoljiti sve četiri dimenzije. Npr. prejednostavan model u većini slučajeva neće imati sposobnost prikaza realnosti ili neće biti dovoljno precizan.



Slika 13. Dimenzije kvalitete (prema: Aalst, 2018)

2.4. Područje primjene proces mininga i prethodna istraživanja

Kako se proces mining bavi analizom procesa, može se zaključiti kako se ta tehnika može primijeniti u bilo kojem području koje sadrži neki proces. Autor Aalst (2018) ističe kako su alat ProM primijenili u više od 150 organizacija, od općina i bolnica do financijskih institucija i proizvođača visokotehnoloških rješenja. Kao primjer u kojim se sve područjima zapravo može iskoristiti proces mining mogu se uzeti sva istraživanja i studije slučaja provedenih u različitim područjima. Autori J. D. Weerdt, A. Schupp, A. Vanderloock i B. Baesens (2013) proveli su studiju slučaja primjene proces mininga u organizaciji financijskih usluga. Primijenili su tehnike proces mininga, otkrivanje, usklađivanje i proširenje društvenim mrežama i otkrili određene devijacije u procesu koje su istaknute menadžmentu organizaciji. Istraživanje je pomoglo organizaciji da promijeni problematične dijelove u procesu i omogućilo im je bolje poslovanje. Autori M. Jans, J. M. van der Werf, N. Lybaert i K. Vanhoof (2011) proveli su istraživanje kako se proces mining može iskoristiti u svrhu ranog otkrivanja transakcijskih prijevара. Iskoristili su tehnike proces mininga kako bi otkrili pravi proces i analizirali nedostatke i kako bi otkrili slučajeve koji ne idu prema propisima. Zaključili su kako je upotrebom proces mininga moguće obaviti reviziju nad slučajevima i u ranim fazama otkriti

radi li se o transakcijskoj prijevari ili ne. Osim u financijskim institucijama i u svrhu otkrivanja prijevara, proces mining se može koristiti i u bolnicama. Autori R.S. Mans, M.H. Schonenberg, M. Song, W.M.P. van der Aalst i P.J.M. Bakker (2008) proveli su proces mining tehnike u Nizozemskoj bolnici. U istraživanju su pokazali kako je moguće otkriti kompleksne bolničke procese i analizirati ih. Uspješno su osigurali podatke koji se mogu usporediti s podacima o dijagnostičkoj putanji zdravstvenog procesa ginekološke onkologije. Tim putem su bolnici uspješno ukazali na nedostatke i načine kako popraviti proces imajući pacijenta na umu. Posljednje istraživanje koje pokazuje gdje se sve može iskoristiti proces mining proveli su autori E. Schwaickardt i M. J. Pereira Dantas (2018). Istraživanje je provedeno u svrhu analize dostupne literature koja se bavi proces miningom u procesu upravljanja opskrbom kako bi se vidjelo na koji način je moguće iskoristiti tehnike i čemu treba posvetiti pažnju u budućnosti. Koristeći sva nabrojana istraživanja može se zaključiti kako je proces mining moguće koristiti u različite svrhe, ali pošto je proces mining još uvijek mlada tehnika za analizu, potrebno je dodatno usavršavati njene programe i načine na koje može analizirati procese.

3. Metodologija rada

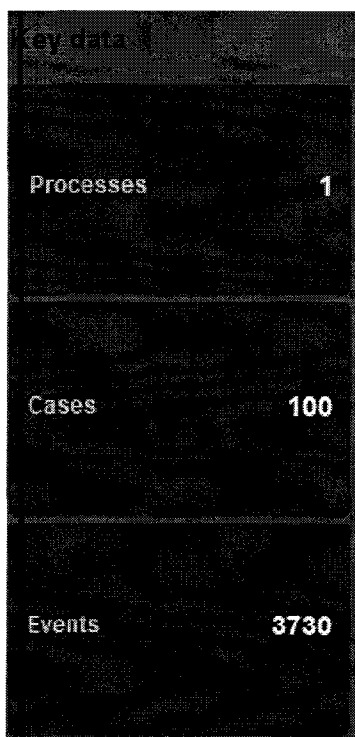
Predmet istraživanja ovog rada su tehnike proces mininga na primjeru načina postupanja s recenzijama za određeni časopis na kojem će se upotrijebiti tehnike proces mininga. Izvori podataka korišteni u radu sastoje se od vlastitih i sekundarnih istraživanja. Vlastito istraživanje je provedeno nad podacima vezanim za recenziju kako bi se demonstrirale tehnike proces mininga. Alat korišten u vlastitim istraživanjima je ProM. ProM je alat koji podržava različite tehnike proces mininga u obliku plug-in datoteka i razvijen je na fakultetu tehnologije Eindhoven. ProM je alat koji je kompliciran za korištenje, ali pošto podržava toliko različitih vrsta analiza iznimno je koristan. Sekundarni podaci korišteni u radu su podaci o recenzijama koji su korišteni u knjizi Process Mining, knjige, stručni članci i studije slučaja. Kao primarni izvor literature koja je oblikovala u najvećem dijelu ovaj rad koristila se knjiga Process Mining: Data Science in Action čiji je autor Wil van der Aalst. Metode korištene u ovom završnom radu su induktivna metoda, metoda sinteze i metoda deskripcije. „Induktivna metoda je sistematska i dosljedna primjena induktivnog načina zaključivanja u kojem se na temelju pojedinačnih ili posebnih činjenica dolazi do zaključka o općem sudu, od zapažanja konkretnih pojedinačnih slučajeva i fakata dolazi se do općih zaključaka, od poznatih pojedinačnih slučajeva polazi se nepoznatom općem, od izučenog neizučenom, od većeg broja pojedinačnih pojava vrše se uopćavanja“ (Zelenika, 1998:323). Induktivna metoda u ovom radu omogućuje da se na temelju pojedinačnih dijelova procesa u sustavu visokih učilišta dođe do zaključka o potrebi za izmjenom procesa. Metoda sinteze predstavlja „postupak znanstvenog istraživanja i objašnjavanja stvarnosti putem spajanja, sastavljanja jednostavnih misaonih tvorevina u složene i složenijih u još složenije, povezujući izdvojene elemente, pojave, procese i odnose u jedinstvenu cjelinu u kojoj su njezini dijelovi uzajamno povezani“ (Zelenika, 1998:329). Metodom sinteze spojiti će se pojedinačni dijelovi ili elementi procesa u sustavu kako bi se moglo doći do jedinstvene cjeline. Metoda deskripcije je „postupak jednostavnog opisivanja ili ocrtaavanja činjenica, procesa i predmeta u prirodi i društvu te njihovih empirijskih potvrđivanja odnosa i veza, ali bez znanstvenog tumačenja i objašnjavanja“ (Zelenika, 1998:338). Metoda deskripcije se u radi koristi u svrhu opisivanja procesa radi lakšeg shvaćanja i tumačenja. Pretpostavka je da će analiza promatranog načina postupanja s recenzijama imati određene negativne aspekte i na osnovu toga će biti dani određeni prijedlozi za poboljšanje procesa.

4. Opis istraživanja i rezultati istraživanja

Zapis događaja korišten u radu opisuje način postupanja s recenzijama za određeni časopis. Zapis sadrži 100 slučajeva (recenzija) i 3730 aktivnosti vezanih uz recenzije. Svaka recenzija nakon zaprimanja se šalje trima različitim recenzentima i oni moraju napisati izvještaj. U određenim slučajevima recenzenti ne odgovaraju i nije moguće napraviti odluku o recenziji. Kao rezultat toga potrebno je pozvati dodatne recenzente. Taj proces se ponavlja dok se ne dođe do konačne odluke (prihvatiti ili odbiti recenziju). Podaci korišteni u radu su također korišteni u knjizi Process Mining: Data Science in Action kao primjer.

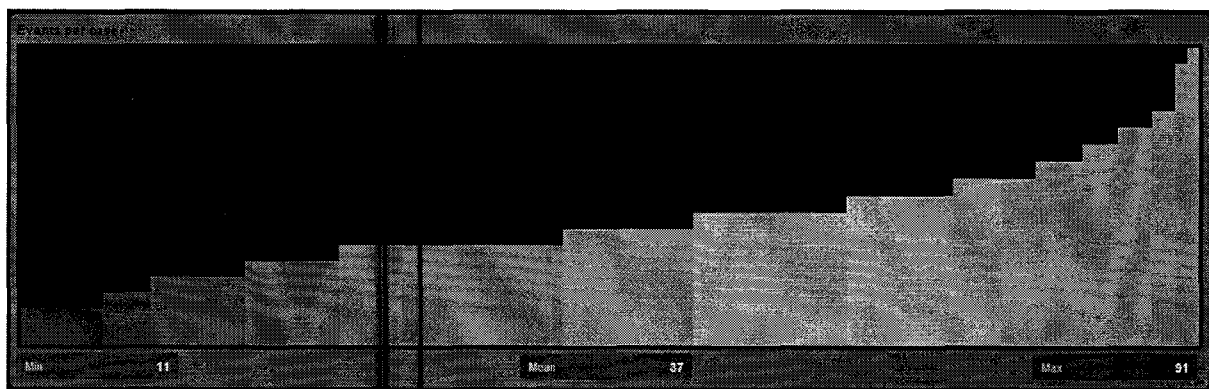
4.1. Pogled na event log

Nakon pribavljanja potrebnih podataka u primjerenom formatu, event log je potrebno učitati u ProM alat. ProM alat u sebi sadrži mogućnost pogleda event loga tzv. Log Visualizer u kojem ProM prikaže određene statistike vezane uz učitani zapis događaja. Na slici 14 prikazan je Log Visualizer na kojem je vidljivo da zapis u sebi sadrži jedan proces na kojem se radi, 100 slučajeva zaprimljenih u proces i 3730 aktivnosti vezanih za proces.



Slika 14. Ključni podaci vezani uz event log (izrada autora)

Nakon promatranja ključnih podataka vezanih uz event log, u Log Visualiseru se mogu pregledati određeni grafikoni koji sadrže dodatne podatke o event logu. Kao primjer, na slici 15 prikazan je grafikon koji predstavlja broj aktivnosti po slučaju. Iz grafikona je vidljivo da je najmanji broj aktivnosti po slučaju 11, najveći 91, a u prosjeku event log u sebi sadrži 37 aktivnosti po jednom slučaju zaprimljenom.



Slika 15. Broj događaja po slučaju (izrada autora)

Unutar Log Visualisera postoji i mogućnost pregleda individualnog slučaja kako bi vidjeli koje aktivnosti je taj slučaj sadržavao. Primjer jednog takvog slučaja i pripadajućih aktivnosti vidljiv je na slici 16 gdje je na vrhu prikazan broj slučaja i koliko događaja u sebi sadrži. Ispod broja slučaja prikazani su aktivnosti uz koje je vidljivo koji resurs je obavio događaj i u koje vrijeme je događaj započet.



Slika 16. Slučaj sa pripadajućim događajima (izrada autora)

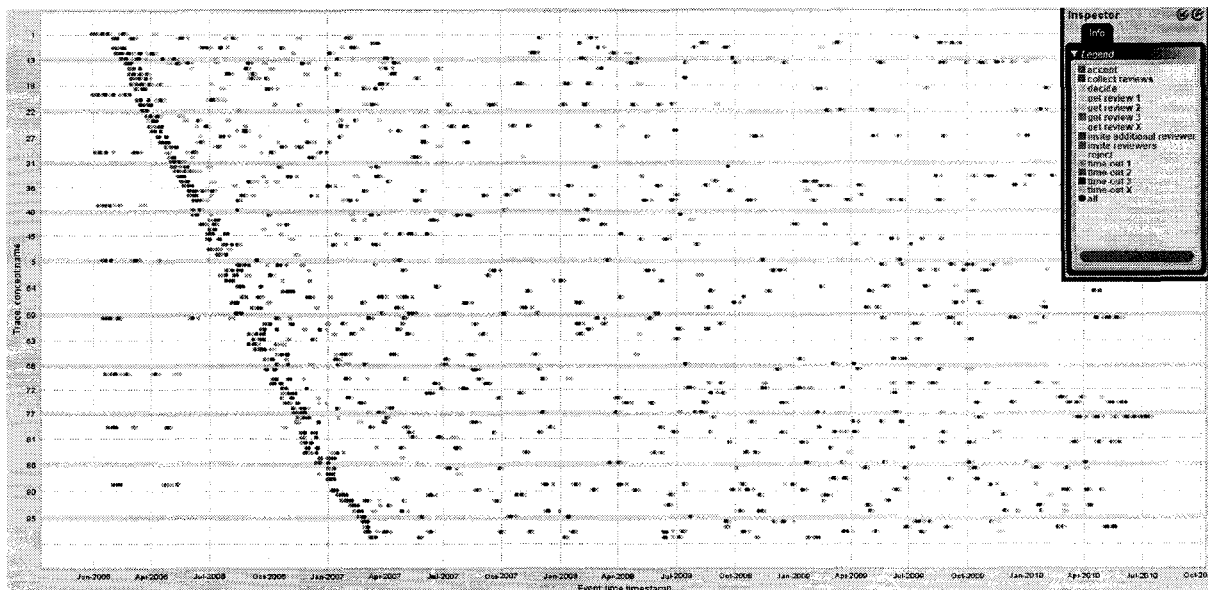
Posljednja stvar koju je korisno pregledati unutar Log Visualisera je učestalost pojavljivanja aktivnosti, tj. koja aktivnost se najčešće obavlja. Na slici 17 vidljivo je da se u ovom event logu najčešće odvijala aktivnost odluke, pojavljujući se ukupno 626 puta, tj. 16.78%.

Class	Occurrences (absolute)	Occurrences (relative)
decide+complete	626	16.783%
decide+start	626	16.783%
invite additional reviewer+start	526	14.102%
invite additional reviewer+complete	526	14.102%
get review X+complete	263	7.051%
time-out X+complete	263	7.051%
collect reviews+start	100	2.681%
invite reviewers+start	100	2.681%
collect reviews+complete	100	2.681%

Slika 17. Učestalost pojavljivanja događaja (izrada autora)

Zapis događaja u programu ProM se može također vizualizirati kao dotted chart tj. točkasti dijagram. Na slici 18 vidljiv je zapis događaja koji je vizualiziran kao točkasti dijagram i na njemu su vidljivi svi događaji koji su se dogodili u promatranom procesu. Na X osi dijagrama nalazi se vrijeme u kojem se proces odvijao, a na Y osi nalaze se slučajevi. Svaka točka na točkastom dijagramu predstavlja jedan događaj koji se dogodio unutar procesa. Detaljnom

analizom točkastog dijagrama može se dobiti uvid u vrijeme potrebno određenom slučaju da se izvrši. Pogledom na dijagram vidljivo je da se iznimno puno događaja sporo odvijalo iz određenog razloga jer između točaka koje predstavljaju događaj ima veliki razmak koji predstavlja vrijeme.



Slika 18. Točkasti dijagram (izrada autora)

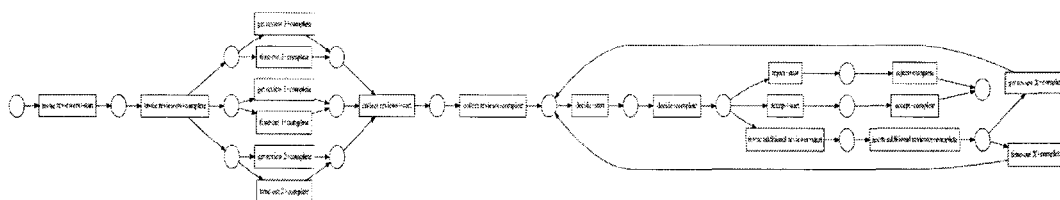
4.2. Primjena tehnika proces mininga

U nastavku rada primijenit će se tri tehnike proces mininga na podacima o recenzijama kako bi se one demonstrirale i objasnile.

4.2.1. Otkrivanje

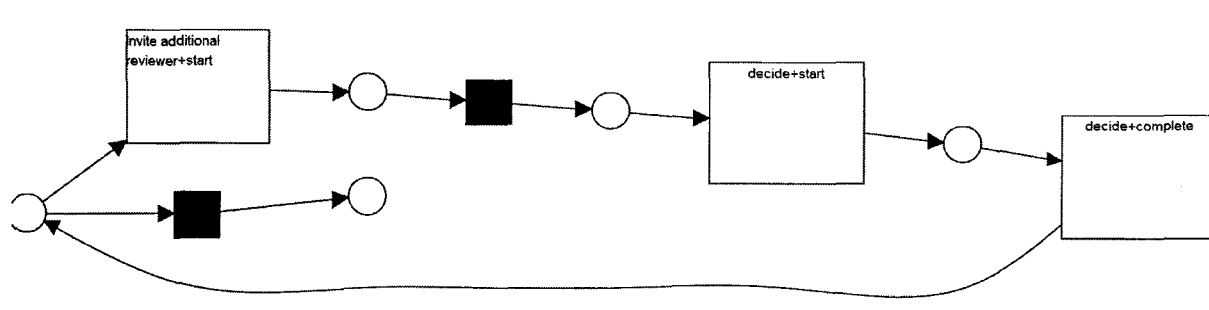
Primjena proces mininga započinje fazom otkrivanja, tj. iz event loga će se automatski generirati model. U ovom slučaju koristit će se α -algoritam koji će prikazati PetriNet model. Iz zapisa događaja se žele dobiti modeli koji mogu imati petlje, paralelne dijelove i koji mogu imati mogućnost izbora, a α -algoritam je jedan od prvih algoritama koji je bio u mogućnosti to napraviti, jamčeći određena svojstva modela. Primjenom α -algoritma na zapis događaja ProM alat prikazuje PetriNet model, vidljiv na slici 19. Na modelu su vidljive sve aktivnosti koje su provedene i kojim tokom one idu. Također je vidljivo da model ima dvije XOR skretnice i jednu AND skretnicu, prva nakon „odluči-završeno“, a druga nakon „pozovi dodatne recenzente-završeno“. AND skretnica započinje nakon „pozovi recenzente-završeno“. Zapis događaja

korišten u radu je jednostavan kako bi se lakše mogle demonstrirati tehnike proces mininga, ali ako se model pokaže previše kompleksnim, on se može modificirati kako bi se dobio jednostavniji pogled radi bolje analize procesa.



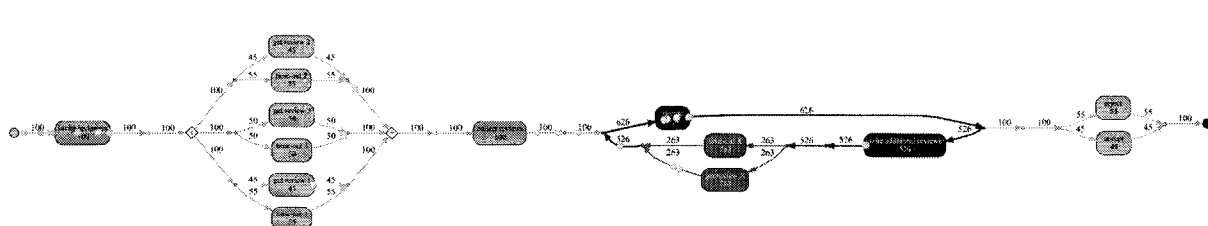
Slika 19. PetriNet model (izrada autora)

Nakon uređivanja modela radi bolje kvalitete, dobije se model prikazan na slici 20. Takav model je lakše za analizirati i moguće je utvrditi koje aktivnosti se najčešće obavljaju, jer modificiranjem se iz modela izbacuju aktivnosti i putanje koje nisu prečesto korištene. Ako se ustanovi da je model i dalje prevelik i teško ga je analizirati, potrebno je iskoristiti druge metode otkrivanja modela, kao što su inductive miner i heuristic miner.



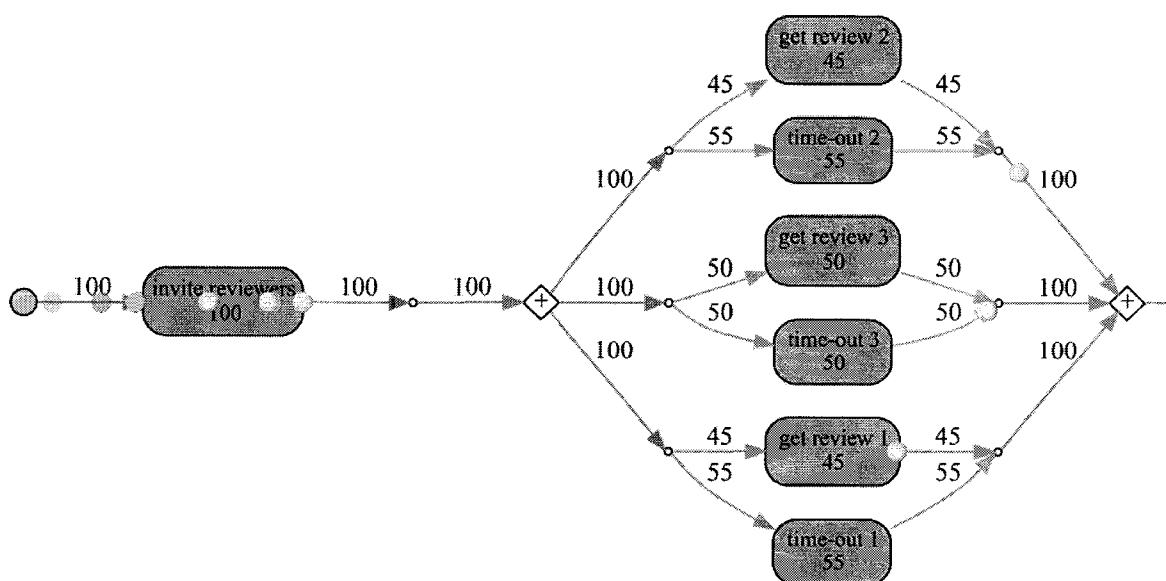
Slika 20. Modificirani model (izrada autora)

Kako bi se dobio dobar model u ovom radu koristit će se inductive miner. Inductive miner je tehnika koja prikazuje jednostavniji BPMN dijagram i njega je lakše analizirati i ima bolje dimenzije kvalitete nego što α -algoritam može pružiti. Na slici 21 vidljiv je model dobiven inductive minerom i vidljivo je da se model može jednostavnije analizirati.



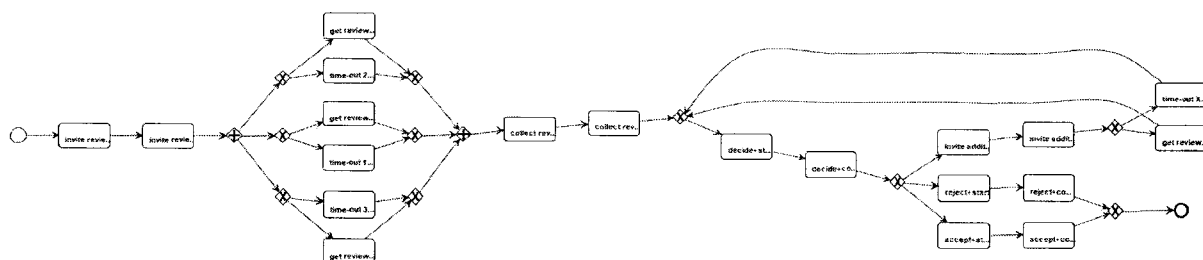
Slika 21. BPMN dijagram generiran inductive minerom (izrada autora)

Inductive miner ima mogućnost prikaza tijeka slučaja i pokazivanja frekvencije događaja što je prikazano na slici 22. Na ovakvom prikazu dijagrama vidljive su frekvencije koje se nalaze iznad linija dijagrama. Npr. prvi vidljiv broj, 100, pokazuje da se prvi događaj dogodio 100 puta, onoliko puta koliko je ukupno bilo slučajeva, jer su svi slučajevi započeli s istom aktivnosti. Na slici su također vidljive i žute točke, koje prikazuju tijek odvijanja pojedinog slučaja i koje događaje taj slučaj sadrži.



Slika 22. Prikaz frekvencija s inductive minerom (izrada autora)

Pošto proces mining može koristiti različite vrste jezika za modeliranje (PetriNet, BPMN itd.), program ProM u sebi sadrži mogućnost pretvaranja PetriNet modela u BPMN ako se netko odluči da je tako lakše analizirati model, iako većina dodataka u ProM-u zahtjeva da se koristi PetriNet model jer je algoritmima njega lakše analizirati. Na slici 23 vidljiv je prikaz PetriNet modela pretvorenog u BPMN model.



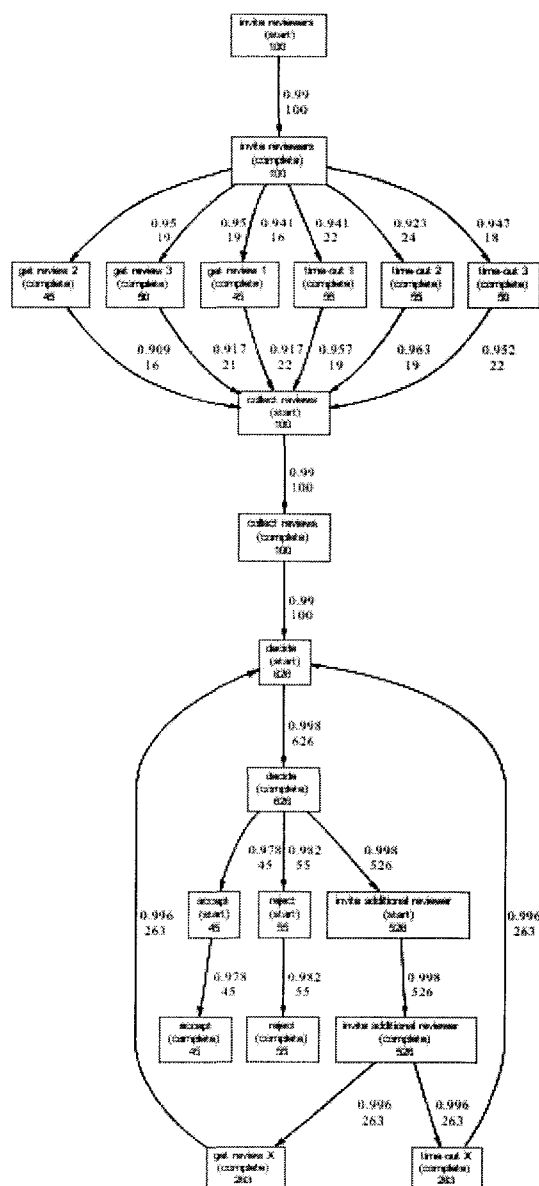
Slika 23. BPMN model (izrada autora)

Posljednja tehnika kojom će se otkriti model u ovom radu je heuristic miner. Heuristic miner je nasljednik α -algoritma i koristi naprednije tehnike za otkrivanje modela nego α -algoritam. Za razliku od α -algoritma, heuristic miner u obzir uzima i frekvencije događanja svakog događaja i uz pomoć tzv. „mjera ovisnosti“ otkriva model s prikladnim frekvencijama. Mjere ovisnosti se mogu i ručno analizirati, ali uz pomoć heuristic minera to nije potrebno. Za ručno analiziranje potrebno je pogledati zapis događaja i zapisati slijed svake aktivnosti i napraviti matricu. Nakon što se napravi matrica u kojoj je vidljivo koja aktivnost slijedi nakon određene aktivnosti, mogu se izračunati frekvencije „izravno slijedi nakon“ odnosa. Primjer jednog takvog izračuna frekvencija nalazi se na slici 24.

$ \Rightarrow_L $	a	b	c	d	e
a	$\frac{0}{0+1} = 0$	$\frac{11-0}{11+0+1} = 0.92$	$\frac{11-0}{11+0+1} = 0.92$	$\frac{13-0}{13+0+1} = 0.93$	$\frac{5-0}{5+0+1} = 0.83$
b	$\frac{0-11}{0+11+1} = -0.92$	$\frac{0}{0+1} = 0$	$\frac{10-10}{10+10+1} = 0$	$\frac{0-0}{0+0+1} = 0$	$\frac{11-0}{11+0+1} = 0.92$
c	$\frac{0-11}{0+11+1} = -0.92$	$\frac{10-10}{10+10+1} = 0$	$\frac{0}{0+1} = 0$	$\frac{0-0}{0+0+1} = 0$	$\frac{11-0}{11+0+1} = 0.92$
d	$\frac{0-13}{0+13+1} = -0.93$	$\frac{0-0}{0+0+1} = 0$	$\frac{0-0}{0+0+1} = 0$	$\frac{4}{4+1} = 0.80$	$\frac{13-0}{13+0+1} = 0.93$
e	$\frac{0-5}{0+5+1} = -0.83$	$\frac{0-11}{0+11+1} = -0.92$	$\frac{0-11}{0+11+1} = -0.92$	$\frac{0-13}{0+13+1} = -0.93$	$\frac{0}{0+1} = 0$

Slika 24. Primjer izračuna frekvencija (prema: Aalst, 2018)

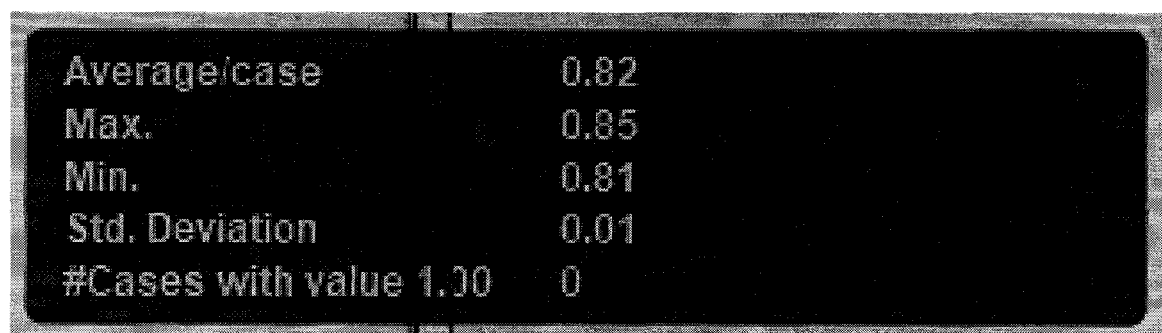
Nakon što se izračunaju frekvencije za dani zapis događaja, može se uz pomoć heuristic minera dobiti graf ovisnosti. Iz grafa ovisnosti (prikazan na slici 25) može se analizirati učestalost obavljanja aktivnosti i mogu se identificirati aktivnosti koje nisu često obavljane. Ovisno o pragu ovisnosti koji je izabran određene putanje na dijagramu mogu nestati ako nisu korištene dovoljno. U slučaju procesa recenzija, svi događaji se dovoljno često odvijaju prema pragu ovisnosti (0.9) i nijedna putanja iz modela nije izbrisana.



Slika 25. Heuristic miner (izrada autora)

4.2.2. Provjera usklađenosti

Kako bi se provjerila usklađenost modela potrebno je imati zapis događaja i normativni model. Normativni model je model koji prikazuje kakva je idealna situacija, tj. kako bi proces zapravo trebao izgledati. Normativni model se može određenim tehnikama izvući iz event loga ili se može ručno napraviti. Korištenjem tehnika za provjeravanje usklađenosti može se utvrditi gdje u procesu nastaju zastoji i koje aktivnosti najduže treba čekati za izvršenje. Primjenom tehnike za provjeru usklađenosti u ProM alatu, dobije se određena statistika. Iz statistike modela (prikazana na slici 26) može se zaključiti kako zapis događaja ne odgovara savršeno normativnom modelu jer se usklađenost gleda prema dimenziji sposobnosti opisane u poglavlju dimenzija kvalitete. Kada se gleda usklađenost zapisa i modela analizira se mogućnost odigravanja zapisa preko modela. Na modelu se pokušava odigrati svaki slučaj onako kako je zapisano u zapisu događaja, i ako model dozvoljava svaki slučaj koji je u zapisu, zapis i model su usklađeni i imaju 100% usklađenost tj. 1 fitness. Prema statistikama vidljivo je da je srednja vrijednost usklađenosti modela i zapisa 0.82, što nije savršeno ali prolazi kao dobar model jer ima vrlo visoku usklađenost. Najmanja usklađenost zapisa i modela iznosila je 0.81, a najviša 0.85.



Average/case	0.82
Max.	0.85
Min.	0.81
Std. Deviation	0.01
#Cases with value 1.00	0

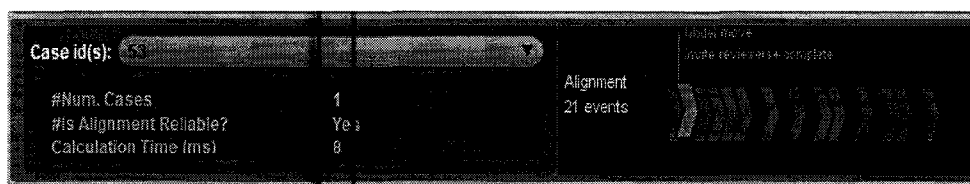
Slika 26. Statistika modela (izrada autora)

Na slici 27 prikazani su nasumični slučajevi kako bi se moglo analizirati koje aktivnosti nisu zapisane u zapisu događaja, a u modelu su stavljene da moraju biti odrađene.



Slika 27. Prikaz tehnike usklađivanja (izrada autora)

Na slici 28 vidljiv je slučaj broj 53, na kojem je program sa desne strane dao pregled aktivnosti koje su se odvijale označene uzastopnim strelicama. Strelice označene zelenom bojom označuju stanje u kojem je aktivnost zabilježena i u zapisu događaja i u modelu i ta aktivnost se odvija sinkronizirano. Ljubičastom bojom označene su aktivnosti koje su preskočene, tj. aktivnosti koje su zabilježene u modelu, ali nisu zabilježene u zapisu događaja, pa je model preskočio te aktivnosti. U programu postoji mogućnost i da je strelica žuto obojana, što bi značilo da je neka aktivnost umetnuta u model, tj. aktivnost je zabilježena u zapisu događaja ali nije u modelu, te program sam dodaje aktivnost. U korištenom zapisu događaja nema žuto obojanih strelica što znači da nema umetnutih aktivnosti u model. Prva aktivnost koja se u modelu i zapisu odvija i označena je zelenom strelicom je „pozvati recenzente-započeto“ i ta aktivnost se odvija sinkronizirano, tj. i model i zapis događaja imaju tu aktivnost. Druga aktivnost koja se odvija u normativnom modelu, ali ne i u zapisu događaja i označena je ljubičastom strelicom je „pozvati recenzente-završeno“. Ta aktivnost je zabilježena u normativnom modelu što znači da bi tako trebao izgledati savršen proces, ali ta aktivnost nije zabilježena u zapisniku, što bi značilo da zaposlenici u programu obilježavaju u koje vrijeme su započeli sa aktivnošću, ali ne obilježavaju kada su s njom završili. Takve ne usklađenosti modela i zapisa događaja se analiziraju u poboljšanju procesa o kojem će biti rečeno nešto više u nastavku rada.



Slika 28. Slučaj 53 (izrada autora)

4.2.3. Poboljšanje

Nakon otkrivanja modela iz zapisa događaja i usklađivanja modela sa zapisom, potrebno je otkrivene probleme u procesu ispraviti. Tu se koristi proces mining u smislu poboljšanja procesa. Poboljšanje procesa može biti popravak procesa ili proširenje procesa.

Popravak procesa odrađuju stručnjaci za modeliranje procesa koji moraju biti upoznati s tehnikama modeliranja i sa cijelim procesom tvrtke kako bi znali što u procesu treba promijeniti i na koji način proces promijeniti da on bolje funkcionira. Popravak procesa zahtjeva modeliranje postojećeg procesa. Popravak se vrši kada normativni model i zapis događaja nisu sukladni i potrebno je napraviti izmjene. „Problem popravka modela je transformirati model N koji nije u skladu s zapisom L u model N0 koji odgovara zapisu L, ali je, što je više moguće, bliži modelu N“ (Fahland,Aalst,2013:10). Cilj je popraviti normativni model da više odgovara zapisu događaja kako se ukupni proces ne mijenja previše jer to može narušiti radu određene tvrtke, ali da popravljeni model bude što bliži normativnom jer je to savršeni način na koji proces može funkcionirati. Pošto se u primjeru korištenom u radu radi samo o aktivnostima za koje nije zabilježen završetak, program ProM može izvršiti jednostavan popravak modela kako bi zapis događaja i model bili usklađeni. Na slici 29 vidljiv je izvršeni popravak modela u programu i vidljivo je da su sve strelice koje su prethodno bile označene ljubičasto sada označene zeleno i da su zapis događaja i model usklađeni.

Case id(s): 63		Alignment 21 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 81		Alignment 21 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 23		Alignment 16 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 62		Alignment 16 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 64		Alignment 16 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 76		Alignment 16 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	
Case id(s): 67		Alignment 11 events
#Num. Cases	1	
#Is Alignment Reliable?	yes	
Approximate memory used (kb)	1	

Slika 29. Usklađeni zapis događaja i model (izrada autora)

Pošto je tehnika popravka modela napredna tehnika u proces miningu rad će se fokusirati na tehnike proširenja procesa kako bi se dobio uvid u druge dimenzije poslovnog procesa.

Zapisi događaja osim aktivnosti mogu sadržavati puno više podataka koje je moguće iskoristiti u proces miningu. Koristeći te dodatne podatke može se dobiti uvid u druge perspektive procesa kao što su organizacijska perspektiva, perspektiva slučaja i vremenska perspektiva. Organizacijska perspektiva se može koristiti kako bi se dobio uvid u organizacijsku strukturu, društvene mreže i tipičan obrazac rada. Perspektiva slučaja se može koristiti kako bi se bolje razumio način donošenja odluka i razlike između slučajeva, a vremenska perspektiva pomaže pri otkrivanju zastoja u procesu i druge probleme vezane uz performansu.

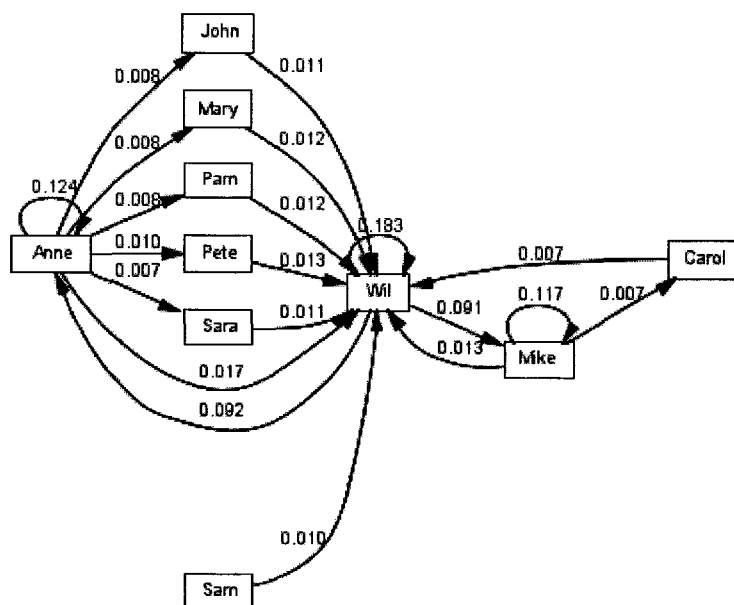
3.2.3.1. Organizacijska perspektiva

Jedna vrsta organizacijske perspektive su društvene mreže bazirane na primopredaji posla. Društvene mreže daju grafički prikaz podataka vezanih uz pojedince u procesu, i ova specifična društvena mreža prikazuje količinu predanog posla između zaposlenih. Slika 30 prikazuje sve primopredaje posla koje su se odvijale u procesu. Svaki iznos unutar matrice koji nije nula, prikazuje vezu predaje posla između osoba. Brojevi unutar matrice prikazuju srednju vrijednost predaje posla.

	Anne	Carol	John	Mary	Mike	Pam	Pete	Sam	Sara	Wil
Anne	0.1239067...	0.0043731...	0.0075801...	0.0084548...	0.0	0.0084548...	0.0096209...	0.0058309...	0.0069970...	0.0166180...
Carol	0.0011661...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0069970...
John	0.0017492...	0.0	0.0	0.0	5.8309037...	0.0	0.0	0.0	0.0	0.0113702...
Mary	2.9154518...	0.0	0.0	0.0	5.8309037...	0.0	0.0	0.0	0.0	0.0116618...
Mike	0.0	0.0072886...	0.0055393...	0.0052478...	0.1169096...	0.0061224...	0.0061224...	0.0055393...	0.0052478...	0.0125364...
Pam	2.9154518...	0.0	0.0	0.0	2.9154518...	0.0	0.0	0.0	0.0	0.0122448...
Pete	8.7463556...	0.0	0.0	0.0	8.7463556...	0.0	0.0	0.0	0.0	0.0134110...
Sam	8.7463556...	0.0	0.0	0.0	2.9154518...	0.0	0.0	0.0	0.0	0.0099125...
Sara	0.0011661...	0.0	0.0	0.0	2.9154518...	0.0	0.0	0.0	0.0	0.0110787...
Wil	0.0915451...	0.0	0.0	0.0	0.0909620...	0.0	0.0	0.0	0.0	0.1825072...

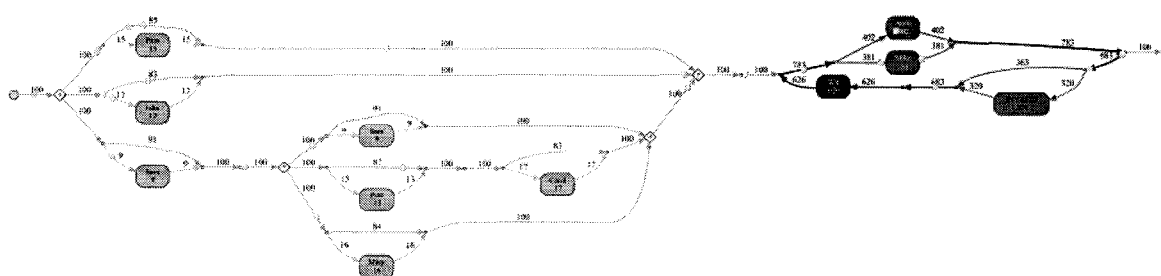
Slika 30. Matrica primopredaje posla (izrada autora)

Osim tablicom, društvena mreža se može prikazati i modelom. Za prikaz modela potrebno je odrediti prag vrijednosti kako se u modelu ne bi prikazale sve veze i kako model ne bi bio kompliciran. Za prikaz društvene mreže za proces recenzije koristio se prag vrijednosti od 0.006 kako bi model prikazao sve zaposlenike u procesu. Ako se želi vidjeti cijeli model sa svim interakcijama zaposlenika potrebno je prag vrijednosti postaviti na nulu i model će prikazati sve interakcije koje su zabilježene i u tablici. Ako se iz tablice i modela ustanovi da određena osoba zaprima puno posla od drugih, postoji mogućnost potrebe rasterećenja te osobe.



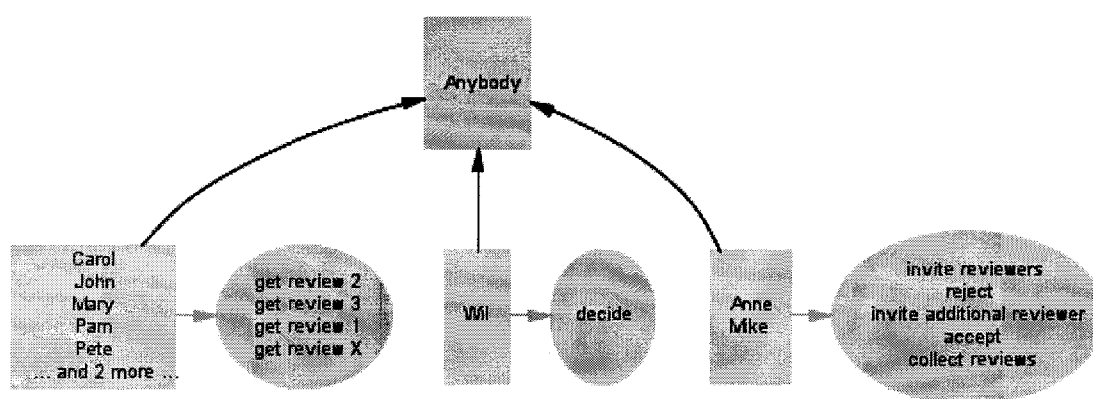
Slika 31. Model društvene mreže (izrada autora)

Primjenom inductive visual minera može se dobiti model koji sličan modelu korišten za otkrivanje modela, samo što u ovom slučaju u kvadratima nisu prikazane aktivnosti koje se odvijaju u procesu, nego su prikazana imena osoba koje su zadužene za aktivnosti. Iz dijagrama prikazanog na slici 32 vidljive su sve osobe koje sudjeluju u procesu i usporedbom s modelima koji prikazuju aktivnosti, može se vidjeti koju aktivnost određena osoba obavlja. Iz dijagrama je također vidljiv tok događaja u procesu i koliko slučajeva određena osoba zaprimi.



Slika 32. Inductive visual miner (izrada autora)

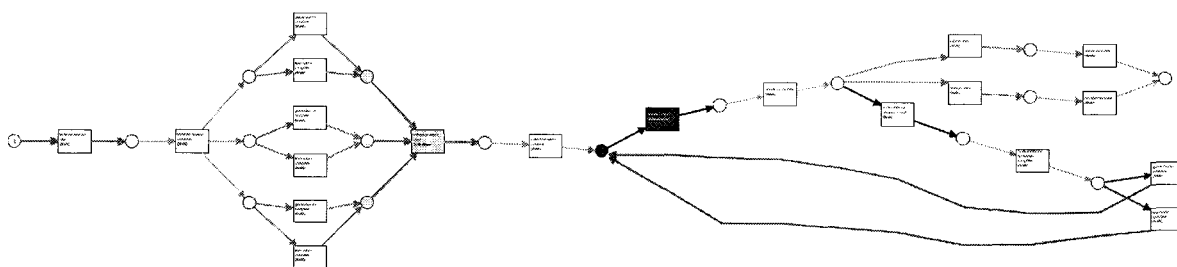
Postoji mnogo vrsta društvenih mreža, među koje pripada i mreža hijerarhije u kojem program analizira zaposlenike i uloge koje obavljaju, zatim ih pokušava rangirati prema važnosti obavljanja posla. Na slici 33 vidljiv je model hijerarhije i vidljivo je da u ovom procesu nema izričito definirane hijerarhije. Na vrhu hijerarhije prema modelu zabilježen je „bilo tko“ zato što zaposlenici u ovom procesu nemaju izričitog menadžera ili šefa koji kontrolira aktivnosti. Program je uspješno podijelio zaposlenike prema poslu kojem obavljaju, i na modelu je vidljivo da Carol, John i 5 drugih obavljaju poslove zaprimanja recenzija. Wil je zadužen za odlučivanje, tj. hoće li recenzija biti prihvaćena ili ne, a Anne i Mike su zaduženi za pozivanje dodatnih recenzenata i za prikupljanje recenzija od ostalih zaposlenika.



Slika 33. Model hijerarhije (izrada autora)

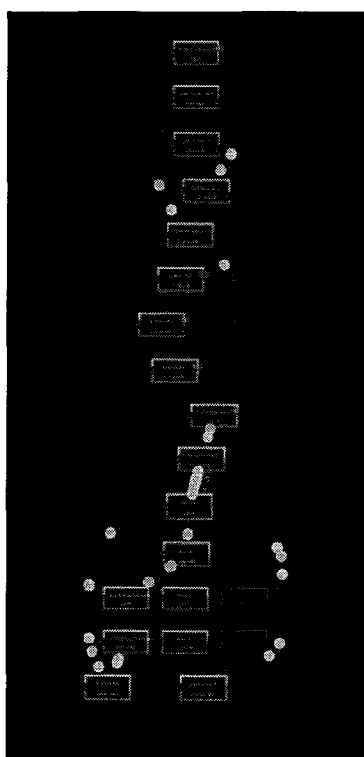
3.2.3.2. Vremenska perspektiva

Vremenska perspektiva zadužena je za vrijeme i frekvenciju događanja ili aktivnosti. Većina zapisa događaja u sebi sadrži i vremenske oznake kada je neka aktivnost započela i završila. Korištenjem tih informacija mogu se pobliže analizirati zastoji u sustavu. Na slici 34 vidljiv je model koji je napravljen tehnikom odigravanja zapisa događaja nad normativnim modelom i jedna aktivnost je označena crvenom bojom i ta aktivnost označava da se tu nalazi zastoj.



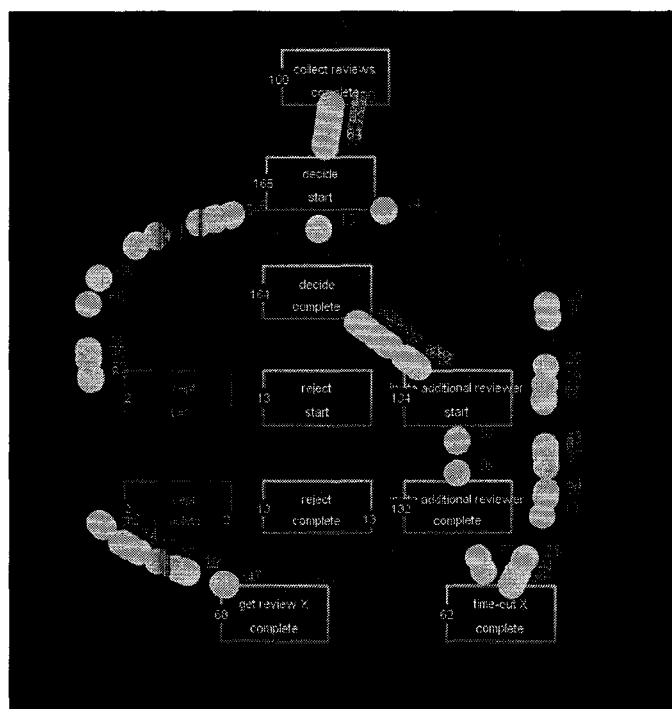
Slika 34. Odigravanje zapisa nad modelom (izrada autora)

Pošto je iz modela koji je napravljen za provjeru usklađenosti vidljivo da proces ima zastoje koje treba popraviti, potrebno je bolje pregledati zastoje jer prethodno korišteni model ne prikazuje što se točno događa sa zastojima. U tu svrhu će se koristiti tzv. fuzzy miner. Fuzzy miner ima mogućnost prikaza sličnih modela kao i prethodni modeli i može pojednostaviti model. Korištenjem fuzzy minera može se dobiti animacija zapisa događaja na modelu i može se vidjeti što se događa sa slučajima koji prolaze kroz proces. Na slici 35 prikazana je fuzzy animacija koja prikazuje realno vrijeme obavljanja zadataka i prikaz putanje kojom se slučajevi kreću. Aktivnosti su prikazane kvadratom, a slučajevi su prikazani krugom i vidljiva je njihova putanja kojom se kreću.



Slika 35. Fuzzy animacija (izrada autora)

Na slici 36 vidljiv je dio modela na kojem se stvaraju najviše zastoja. Prvi zastoji u procesu kreću između aktivnosti „prikupi recenzije“ i „odluči“ iz razloga što se sve prikupljene recenzije od svih recenzenata u isto vrijeme šalju na odluku, a jedna osoba obavlja posao prikupljanja recenzija. Drugi zastoj kreće kada je potrebno zvati dodatne recenzente i zastoj se nastavlja uzduž svih aktivnosti koje dodatni recenzenti i osobe vezane uz te aktivnosti obavljaju. Može se zaključiti kako proces najvećih poteškoća ima u prikupljanju recenzija i sa dodatnim recenzentima i taj dio procesa potrebno je dodatno usavršiti.

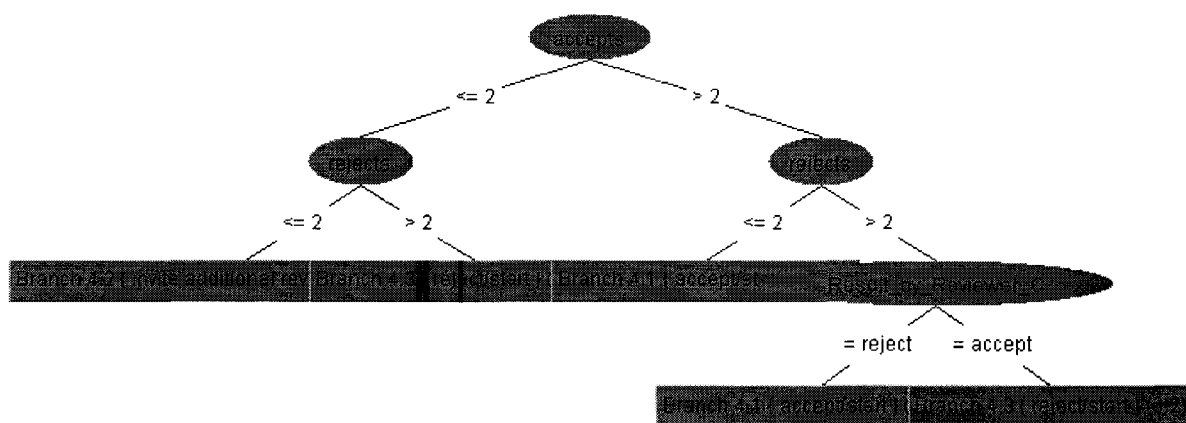


Slika 36. Dio procesa s najviše poteškoća (izrada autora)

Dodavanjem vremenske perspektive procesu može se dobiti uvid u različite informacije vezane uz performansu procesa kao što su: vizualizacija čekanja i vrijeme posluživanja, detekcija zastoja i analiza, SLA analiza (Service Level Agreement) i analiza frekvencija.

3.2.3.3. Perspektiva slučaja

Perspektiva slučaja se odnosi na svojstva pojedinog slučaja. Svaki slučaj je karakteriziran svojim atributima, atributima događaja, putanjom kretanja i informacijama izvođenja. Kako bi se otkrila perspektiva slučaja koristi se rudarenje odluka. Rudarenje odluka se koristi kako bi se otkrila pravila koja utječu na donošenje odluka u procesu. Kao što je objašnjeno u poglavlju otkrivanja modela, model u sebi sadrži 2 XOR skretnice na kojima se mora napraviti odluka koju aktivnost izabrati. Nad jednom skretnicom će se iskoristiti rudarenje odluka kako bi se vidjelo koja pravila su korištena u svrhu donošenja odluke koju aktivnost izvršiti. Rudarenje odluka kao prikaz koristi stabla odlučivanja što je tehnika klasifikacije. Analizirat će se XOR skretnica koja se nalazi nakon aktivnosti „odluči-završeno“ gdje zaposlenik treba odabrati hoće li prihvatiti recenziju, odbiti ili pozvati dodatnog recenzenta. Druga XOR skretnica neće biti analizirana iz razloga što se skretnica odnosi na situaciju u kojoj zaposlenik nije u mogućnosti dobiti recenziju i u toj situaciji nema drugog izbora nego ponovno tražiti recenzenta. Primjenom tehnika rudarenja odluka dobije se stablo prikazano na slici 37. Uz pomoć dobivenog stabla mogu se izvući određena pravila koja su korištena za donošenje odluke. Ako je recenzija prihvaćena 2 puta ili manje, i ako je odbijena 2 puta ili manje, zaposlenik će pozvati dodatnog recenzenta. Ako je recenzija prihvaćena 2 puta ili manje i odbijena više od dva puta, recenzija će biti odbijena. Ako je recenzija prihvaćena više od dva puta, a odbijena 2 puta ili manje, recenzija će biti prihvaćena. Ako je recenzija prihvaćena više od 2 puta, a također je i odbijena više od dva puta, gledaju se rezultati od recenzenta C. Ako su rezultati od recenzenta C odbijeni, prihvatiti će se prethodna recenzija, a ako su rezultati od recenzenta C prihvaćeni, odbit će se prethodna recenzija. Ovo je prvi primjer u modelu donošenja odluka prema kojem je vidljivo da se odluka donosi ovisno koliko puta je recenzija prihvaćena ili odbijena.



Slika 37. Prikaz stabla odlučivanja za donošenje odluke (izrada autora)

Gledajući statistiku analize koju program izračuna, može se vidjeti koliko je program bio uspješan. Na slici 38 vidljivo je da je program ispravno klasificirao 107 slučajeva, što čini 62.94%, a neispravno je klasificirao 63 slučaja što čini 37.06%. Rezultati programa nisu savršeni ali su pomogli kako bi se otkrio cilj perspektive slučaja, a to je otkriti pravila koja su se koristila u svrhu donošenja odluka.

Correctly Classified Instances	107	62.9412 %
Incorrectly Classified Instances	63	37.0588 %

Slika 38. Statistika stabla odlučivanja (izrada autora)

5. Rasprava

Primjenom tehnika proces mininga na određeni proces može se doći do različitih rezultata koje na prvi pogled nisu vidljivi. Iako mnogi smatraju da je proces mining zapravo rudarenje podataka, rezultatima istraživanja se dokazalo kako proces mining koristi određene tehnike rudarenja podataka (kao što su stabla odlučivanja), ali za cjelovitu analizu procesa potrebno je koristiti drugačije i puno kompliciranije tehnike. Iako proces mining nije zapravo rudarenje podataka, ono se može smatrati kao poveznica između rudarenja podataka i tradicionalnom BPMA. Proces mining kao tehnika je još uvijek mlada i ne razvijena, i iz tog razloga stručnjaci još uvijek smišljaju i analiziraju nove tehnike za koje postoji mogućnost upotrebe. Isto tako, još uvijek nije dostupno mnogo programa koji bi sa sigurnom točnošću analizirali dane procese, te se u njima mogu pojaviti određene pogreške. Kao što se godinama uspostavljala potreba za korištenjem npr. rudarenja podataka i strojnog učenja, tako i proces mining treba određeno vrijeme i veliki broj istraživanja kako bi tvrtke, organizacije i sl. uvidjeli važnost primjene analize svojih procesa. Kako vrijeme prolazi tako raste broj podataka o procesima i što se tvrtke brže odluče priključiti korištenju proces mining tehnika, to će im lakše biti za analizu svojih procesa i sustava. Ako neka tvrtka na neki određeni način već sprema zapise događaja, to im je lakše primijeniti proces mining, jer postavljanje sustava za prikupljanje podataka je dugotrajan proces i potrebno ga je pravilno odraditi.

Teorijskom podlogom u radu se pokušalo pokriti dostatan broj termina potrebnih za shvaćanje što je proces mining i kako ga koristiti. Navođenjem prethodnih istraživanja zaključilo se kako je primjena proces mininga raznolika i ne treba se ustručavati primijeniti ga na bilo što, jer podaci o događajima postoje u svakom sustavu samo je potrebno pronaći bitne podatke.

Rezultatima istraživanja su se pokazale jednostavne tehnike proces mininga s kojima se mogu dobiti vrlo bitni podaci o procesu i takvi podaci se mogu primijeniti za poboljšanje procesa. Jedini nedostatak primjene čak i jednostavnih tehnika proces mininga je taj što se proces mining mora razumjeti u cijelosti i potrebna je određena edukacija o programu u kojem se analiza odluči provesti. Autori koji su osmislili program ProM na svojim internetskim stranicama omogućili su uvod u aplikaciju tako da se i početnici mogu početi koristiti programom. Korištenjem alata ProM za analizu procesa uvidjelo se da takav alat ne koristi mnogo ljudi, jer primjenom određenih dodataka u programu došlo je to tehničkih pogrešaka koje nije moguće ispraviti bez podrške ostalih korisnika koji su se sretali sa problemom. Do rješenja tehničkih problema se nije došlo i morali su se pronaći drugi načini za analizu jer skupina osoba koja

koristi program je premala i nisu se susretali sa sličnim problemima. Osim jednostavnih tehnika za analizu procesa, postoje i mnogo kompliciranije tehnike za koje su potrebne godine proučavanja i nisu ušle u opseg ovoga rada. Napredne tehnike u proces miningu bave se tzv. „lazanja“ i „špageti“ procesima koji zahtijevaju puno detaljniju analizu jer su oni po prirodi kompliciraniji od procesa koji su već završeni.

Analizom procesa zaprimanja recenzija može se zaključiti kako proces ima određenih problema koje je potrebno ispraviti. Proces miningom se otkrio dostatan model procesa koji se uspješno primjenio u daljnjoj analizi. U toj analizi je otkriveno da model zapravo ne odgovara savršeno normativnom modelu, ali program je uspješno ispravio model i nisu bili potrebni dodatni popravci. Dodatni popravci modela su mnogo kompliciraniji nego što je napravljeno u radu, iz razloga što je u određenim slučajevima potrebno redizajnirati cijeli proces, a za to su zaduženi stručnjaci. Analizom procesa je ustanovljeno da je model moguće proširiti određenim perspektivama, i kroz te perspektive su opaženi problemi zastoja u procesu. Problemi zastoja se rješavaju dodavanjem radnog osoblja u proces kako bi se rasteretila jedina osoba koja obavlja aktivnost koja uzrokuje zastoj. Zastoj se može riješiti i uvođenjem dodatnih aktivnosti i putanja kojom se aktivnosti kreću. Proces recenzija je zatim proširen i organizacijskom perspektivom, gdje se uvidjelo kako zaposlenici prosljeđuju posao, i može se zaključiti da ako određene osobe zaprimaju puno posla od drugih, to može uzrokovati zastoje u procesu jer osoba mora analizirati sav primljen posao. Takva situacija se također može riješiti uvođenjem dodatne osobe za bržu analizu procesa.

Koristeći proces mining za analizu procesa potrebno je obratiti pažnju na korake koje je potrebno poduzeti prije, tijekom, i nakon analize. Svakako je potrebno primijeniti tri vrste proces mininga u pravilnom redoslijedu. Na kraju je potrebno vratiti integrirani model koji ima dodane sve potrebne perspektive. Takav model se može koristiti i u daljnjoj analizi i u drugim alatima. Na takvom modelu je moguće primijeniti i naprednije tehnike proces mininga i ostalih tehnika. Proces mininga kao tehnika se sve više pokušava koristiti na stvarnim primjerima i potrebno je obratiti pažnju na određene izazove s kojima se proces mining susreće. Otkrivanje modela se smatra skoro najvećim izazovom u proces miningu zbog svih varijabli koje utječu na tu fazu. Ponekad zapisi događaja nisu dostatni da se otkrije dobar model, a ponekad program jednostavno nije u mogućnosti dati prikaz procesa koji se može s lakoćom analizirati. Najveća pogreška s kojom se može krenuti u otkrivanje procesa je pretpostavka da je zapis događaja savršen. Zapis događaja u sebi može sadržavati nepotrebne podatke kao što je već rečeno u poglavlju u kojem se pisalo o zapisu događaja. Iako jednostavni algoritmi, poput α -algoritma

nisu u mogućnosti otkriti dobar model, razvijeni su napredni modeli kao npr. inductive miner koji pružaju dobru potporu za algoritme koji će biti razvijeni u budućnosti. Još jedan problem s kojim se proces mining susreće je problem prezentacije. Dok se osobe koje su upoznate sa svim terminima BPMA i proces mininga lagano mogu snaći u svim modelima, osobe kojima je potrebno iznijeti analizu se neće u svakoj situaciji snaći. Modeli mogu biti vrlo komplicirani, a potrebno ih je na što lakši način objasniti osobama kojima je to potrebno. Iz tog razloga je potrebno pronaći način kako proizvesti modele koji imaju kvalitetu i koje je lagano za objasniti.

Potrebno je reći kako je proces mining za tvrtke lagano početi koristiti, zato što one koriste sustave u kojima mogu pronaći zapise događaja, samo je potrebno znati kako. Uz svu dostupnu literaturu moguće je jednostavno započeti s proces miningom i ostvariti konkurentsku prednost na tržištu.

6. Zaključak

U ovom radu istraživano je područje procesa mininga i kako primijeniti njegove tehnike. Proces mining je rastuća disciplina i poput rudarenja podataka ili tehnika znanosti o podacima, uskoro će se proces mining raširiti u upotrebi i biti će ključan za napredovanje tvrtke ili organizacije. Rješavan je problem primjene tehnika na procesu, kako bi se prikazale tehnike i vrste procesa mininga i kako bi se one približile čitatelju. Metode korištene u radu za istraživanje sustavno su krenule od jednostavnijih termina prema općim terminima i primjeni kako bi se što je više moguće analizirali dijelovi procesa mininga u svrhu boljeg razumijevanja. U rezultatima istraživanja vidljivo je da je analizirani proces imao određene probleme i za te probleme su se dali prijedlozi za poboljšanje u prethodnom poglavlju. Proces korišten za istraživanje ilustrira jednostavnost primjene tehnika procesa mininga, i kako problemi u procesu ne moraju biti preveliki. Uz proces mining probleme je lakše popraviti i moguće je dodatno usavršiti proces. Postoje određena ograničenja vezana uz ovo istraživanje, a to je nedostatak javno dostupnih podataka o procesima. Nedovoljno tvrtki koristi proces mining, a one koje koriste nisu volje javno podijeliti svoje procese vjerojatno iz razloga jer smatraju da će to uništiti reputaciju. Proces korišten u svrhu istraživanja je stvarni proces koji je korišten kao primjer za primjenu tehnika, i ograničenje vezano uz njega je što se smatra da je to mali zapis događaja. Takav zapis događaja se koristio jer su se htjele demonstrirati jednostavnije tehnike procesa mininga, a veći zapis događaja bi imao velike modele koje početnici u procesu miningu ne bi mogli analizirati. Ovakvo istraživanje se može poboljšati korištenjem zapisa događaja s većim brojem podataka i proširiti rad s naprednijim tehnikama procesa mininga. Proces mining kao disciplina je relativno mlada, i u budućnosti su potrebna dodatna proširenja radi bolje analize procesa.

Literatura

1. Burattin, A. (2015). *Process Mining Techniques in Business Environments*, Berlin: Springer
2. Dumas, M., Mendling, M.L.R.J. and Reijers, H.A. (2018) *Fundamentals of Business Process Management*, Berlin: Springer
3. Fahland, D. and Van der Aalst, W.M.P. (2013). *Model Repair - Aligning Process Models to Reality*. Eindhoven University of Technology. Raspoloživo na: <http://wwwis.win.tue.nl/~wvdaalst/old/publications/z9.pdf> [pristupljeno: 21. Kolovoza 2019].
4. Jans, M., van der Werf, J. M., Lybaert, N., Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Application*. 38(2011). pp. 13351-13359.
5. Mans, R.S., Schonenberg, M.H, Song, M., van der Aalst, W.M.P. and Bakker, P.J.M. (2008). Application of Process Mining in Healthcare – A CaseStudy in a Dutch Hospital. *Biomedical Engineering Systems and Technologies 2008*. Fred, A., Filipe, J. and Gamboa, H. (ur.) Madeira: Portugal. 28.-31.01.2008. Berlin: Springer. pp 425-438.
6. Mesarić, J. i Šebalj, D. (2017). *Upravljanje poslovnim procesima*. Ekonomski fakultet u Osijeku. Raspoloživo na: http://www.efos.unios.hr/upravljanje-poslovnim-procesima/wp-content/uploads/sites/196/2018/09/bpm_predavanja-2017-18.pptx [pristupljeno 28. kolovoza 2019].
7. Rozinat, A. and Van der Aalst, W.M.P. (2006). Decision Mining in ProM. *Business Process Management 2006*. Dustdar, S., Fiadeiro, J.L., Sheth,A. (ur.) Vienna: Austria. 05.-07.09.2006. Berlin: Springer. pp 420-425.
8. Schwaickardt, E. and Dantas, M.J.P. (2018). Process Mining Applied in Supply Management Processes. *European Journal of Scientific Research*. 151(2). pp. 160-171.
9. Van der Aalst, W.M.P. and van Dongen, B.F. (2013) Discovering Petri Nets From Event Logs. *5th International Summer School "Advanced Course on Petri Nets"*. Jensen K., van der Aalst, W.M. P., Balbo, G., Koutny M., Wolf, K. (ur.) Rostock, Germany. September 2010. Berlin: Springer. pp 372-422.
10. Van der Aalst, W.M.P. (2018). *Process Mining: Data Science in Action*, Berlin: Springer

11. Van der Aalst, W.M.P. (2012). Process Mining Manifesto. *BPM 2011 Workshops, Part I*. F.Daniel (ur.) Clermont-Ferrand, France. 29.08.2011. Berlin: Springer. pp. 169-194.
12. Weerdt, J. D., Schupp A., Vanderloock A. and Baesens, B. (2013). Process Mining for the multi-faceted analysis of business processes—A case study in a financial services organization. *Computers in Industry*. 64(2013). pp. 57-67.
13. White, S. (2004). *Introduction to BPMN*. [Online] New York: IBM. Raspoloživo na: <https://www.ibm.com> [pristupljeno: 20. kolovoza 2019].
14. Zelenika, R. (2000). *Metodologija i tehnologija izrade znanstvenog i stručnog djela*, Rijeka: Sveučilište u Rijeci.

Popis tablica

Tablica 1. Prošireni skup događaja (prema: Mesarić, Šebalj: 2017)	12
Tablica 2. Prošireni skup skretnica (prema: Mesarić, Šebalj: 2017)	13
Tablica 3. Prošireni skup aktivnosti i transakcija (prema: Mesarić, Šebalj: 2017).....	14

Popis slika

Slika 1. BPM životni ciklus (prema: Aalst, 2018)	2
Slika 2. Discipline u znanosti o podacima (prema: Aalst, 2018)	3
Slika 3. Event log (prema: Aalst, 2018)	6
Slika 4. PetriNet (prema: Burattin, 2015)	7
Slika 5. Tokeni unutar PetriNeta (prema: Burattin, 2015)	8
Slika 6. Objekti toka: događaj, aktivnost i skretnica (prema: White, 2004)	9
Slika 7. Objekti povezivanja: sekvencijski tok, tok poruke i asocijacija (prema: White, 2004)	10
Slika 8. Plivaće staze: polje (bazen) i staze (prema: White, 2004)	10
Slika 9. Artefakti: podatkovni objekt, grupa i anotacija (prema: White, 2004)	11
Slika 10. Primjer uzročne mreže (prema: Aalst, 2018)	16
Slika 11. Pozicioniranje vrsta procesa mininga (prema: Aalst, 2018)	17
Slika 12. Način korištenja procesa mininga (prema: Aalst, 2018)	18
Slika 13. Dimenzije kvalitete (prema: Aalst, 2018)	19
Slika 14. Ključni podaci vezani uz event log (izrada autora)	22
Slika 15. Broj događaja po slučaju (izrada autora)	23
Slika 16. Slučaj sa pripadajućim događajima (izrada autora)	24
Slika 17. Učestalost pojavljivanja događaja (izrada autora)	24
Slika 18. Točkasti dijagram (izrada autora)	25
Slika 19. PetriNet model (izrada autora)	26
Slika 20. Modificirani model (izrada autora)	26
Slika 21. BPMN dijagram generiran inductive minerom (izrada autora)	27
Slika 22. Prikaz frekvencija s inductive minerom (izrada autora)	27
Slika 23. BPMN model (izrada autora)	28
Slika 24. Primjer izračuna frekvencija (prema: Aalst, 2018)	28
Slika 25. Heuristic miner (izrada autora)	29
Slika 26. Statistika modela (izrada autora)	30
Slika 27. Prikaz tehnike usklađivanja (izrada autora)	31
Slika 28. Slučaj 53 (izrada autora)	31
Slika 29. Usklađeni zapis događaja i model (izrada autora)	33
Slika 30. Matrica primopredaje posla (izrada autora)	34
Slika 31. Model društvene mreže (izrada autora)	35

Slika 32. Inductive visual miner (izrada autora)	35
Slika 33. Model hijerarhije (izrada autora)	36
Slika 34. Odigravanje zapisa nad modelom (izrada autora)	37
Slika 35. Fuzzy animacija (izrada autora).....	37
Slika 36. Dio procesa s najviše poteškoća (izrada autora).....	38
Slika 37. Prikaz stabla odlučivanja za donošenje odluke (izrada autora)	40
Slika 38. Statistika stabla odlučivanja (izrada autora)	40