

Stanje in trendi na področju rudarjenja procesov

Gregor Polančič¹ in Mateja Kocbek Bule¹

¹ Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Koroška cesta 46, 2000 Maribor
gregor.polancic@um.si, mateja.kocbek@um.si

Izvleček

Upravljanje poslovnih procesov je uveljavljena metodologija, katere poglavitni namen je učinkovito izboljšanje organizacijskih procesov z uporabo različnih tehnik in tehnologij, ki v veliki meri temeljijo na modelih procesov. Predpogoj za uspešno upravljanje je zagotavljanje veljavnih modelov, kar pa je v praksi pogosto težko doseči, saj so le ti, pogosto zaradi človeških dejavnikov, nepopolni, nepravilni oziroma neažurni.

Alternativni oziroma komplementarni pristop, ki se v zadnjih letih uveljavlja na področju upravljanja poslovnih procesov, je zato na realnih podatkih temelječe upravljanje procesov, ki jih, predvsem v obliki dnevnikov dogodkov, generirajo poslovne informacijske rešitve. Rudarjenje procesov je sinonim za množico tehnik in tehnologij, ki omogočajo avtomatsko generiranje in vizualizacijo modelov procesov na osnovi dnevnikov dogodkov, z njihovo pomočjo pa je možno procese odkrivati, analizirati, preverjati skladnost izvajanih procesov in odkrivati morebitna odstopanja.

V prispevku bomo predstavili in umestili rudarjenje procesov v življenjski cikel upravljanja poslovnih procesov, na osnovi korakov algoritma Alfa pa bo pojasnjen postopek avtomatskega oblikovanja modela procesa iz podatkov systemskega dnevnika. Osrednji del prispevka je namenjen predstavitvi zmožnosti sodobnih rešitev za rudarjenje procesov, ki postajajo v sodobnih poslovnih okoljih nepogrešljiv člen poslovne analitike, z njihovo pomočjo pa je možno sprejemati natančne procesne oziroma poslovne odločitve, ki temeljijo na realnih operativnih podatkih.

Ključne besede: poslovni procesi, upravljanje poslovnih procesov, rudarjenje procesov

Abstract

Business process management is an established methodology the main purpose of which is to effectively improve organizational processes by applying a variety of techniques and technologies that are largely based on process models. A prerequisite for the successful management is the provisioning of valid models, which is often difficult to achieve in practice as commonly, due to human factors, process models tend to be incomplete, inaccurate, or outdated.

An alternative or complementary approach, which has been established in recent years in the field of business process management, considers empirical process-related data, which especially in the form of event logs, are generated by business applications. Process mining is a synonym for a multitude of techniques and technologies that enable automatic event log-based generation and visualization of process models that make possible the detection and analysis of processes, checking the compliance of implemented processes and detection of potential deviations.

In this paper, we will present and position process mining in the business process management lifecycle. Process mining will be demonstrated with the steps of a common process mining algorithm – Alpha. The central part of the paper is aimed at presenting the capabilities of modern tools for process mining, which are becoming an indispensable part of business analytics in modern business environments. With their help, it is possible to make precise process or business decisions based on real operational data.

Keywords: Business process, business process management, process mining

1 UVOD

Upravljanje poslovnih procesov (angl. business process management, BPM) je uveljavljena metodologija, katere poglobitveni namen je povečanje uspešnosti, učinkovitosti in prilagodljivosti organizacijskih procesov in posledično poslovanja podjetja. Predstavlja skupek praks, tehnik in tehnologij, ki so običajno urejene v obliki (teoretičnega) življenjskega cikla, sestavljenega iz naslednjih, medsebojno odvisnih, faz [1]: identificiranje procesa (angl. process identification), odkrivanje procesa (angl. process discovery), analiza procesa (angl. process analysis), prenova procesa (angl. process redesign), implementacija procesa (angl. process implementation) in spremljanje oz. nadzorovanje procesa (angl. process monitoring and control).

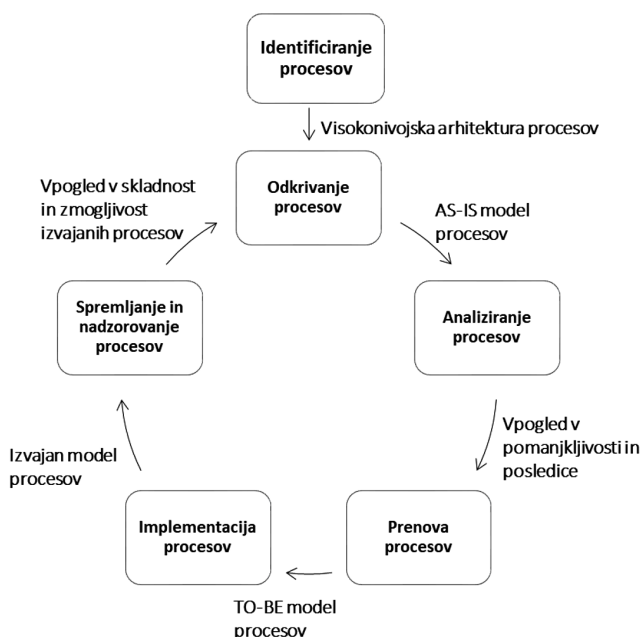
Ker so (poslovni) procesi neotipljiva sredstva (angl. intangible assets), jih običajno upravljamo preko modelov procesov (angl. process model), le ti pa so običajno predstavljeni v grafični obliki oziroma v obliki diagrama (angl. process diagram) [2]. Jedro tradicionalnega upravljanja poslovnih procesov tako predstavljajo modeli procesov, ki služijo različnim namenom kot so: analiziranje procesov, komuniciranje o procesih, implementacija procesov in spremembe oz. izboljšave procesov. Ker odločitve upravljanja poslovnih procesov v veliki meri temeljijo na modelih procesov, je poglobitveni izziv zagotavljanje veljav-

nih modelov procesov (angl. validity), kar pomeni, da modeli predstavljajo dejanske procese oziroma operativno izvajanje in da le te opisujejo v celoti (angl. completeness) [1].

V praksi se je izkazalo, da je ravno zagotavljanje veljavnih in pravilnih modelov procesov šibek člen njihovega upravljanja. Zaradi nenehnih sprememb v poslovnem okolju se poslovni procesi kontinuirano spreminjajo in prilagajajo. Modeli poslovnih procesov tako hitro postanejo neskladni z dejansko izvajanimi procesi kakor tudi s tehničnim okoljem v katerem se izvajajo. Modeliranje poslovnih procesov je prav tako podvrženo človeškim dejavnikom, saj je pretežno odvisno od spretnosti, znanj in razpoložljivosti analitika, kar vpliva na izdelane modele procesov. Tradicionalno odkrivanje in modeliranje procesov je drago in časovno potratno tudi zaradi vrzeli v poslovnem znanju deležnikov in pomanjkanja objektivnih validacij modelov [3]. Modeli so zato pogosto nepopolni (ne opisujejo celotnega procesa), neskladni (ne predstavljajo dejanskega procesa) ali nepravilni (ne upoštevajo pravil diagramskega jezika, na primer BPMN, in so zato nerazumljivi tako za ljudi kakor tudi za izvajalna okolja). Poslovne odločitve, ki se sprejemajo na neustreznih modelih procesov, so tako lahko napačne.

Z višanjem stopnje avtomatizacije (procesov) so se pojavile priložnosti za reševanje omenjenih izzivov, ki omogočajo bolj neposredno povezovanje aktivnosti upravljanja procesov z dejansko izvajanimi procesi in sicer preko podatkov, ki jih poslovne informacijske rešitve generirajo v fazi operativne izvedbe procesov, predvsem v obliki dnevnikov dogodkov (angl. event log). Te podatke je možno z usmerjenimi tehnikami rudarjenja podatkov (angl. data mining) pretvoriti v obliko, ki je uporabna za sprejemanje odločitev upravljanja procesov in se imenuje rudarjenje procesov (angl. process mining). Rudarjenje procesov predstavlja uspešen primer prehoda akademskih zamisli in rešitev v poslovna okolja [3].

V prispevku bomo predstavili in umestili rudarjenje procesov v življenjski cikel upravljanja poslovnih procesov, na osnovi korakov algoritma Alfa pa bo pojasnjen postopek avtomatskega oblikovanja modela procesa iz podatkov systemskega dnevnika. Osrednji del prispevka je namenjen predstavitvi zmožnosti sodobnih rešitev za rudarjenje procesov, ki v postajajo v razvitih državah nepogrešljiv člen poslovne analitike, z njihovo pomočjo pa je možno



Slika 1: Življenjski cikel upravljanja poslovnih procesov [1]

sprejemati natančnejše procesne odločitve, ki temeljijo na operativnih podatkih.

2 OSNOVE RUDARJENJA PROCESOV

Kot že besedna zveza pove, je osnovni namen rudarjenja procesov »pridobivanje« procesov oziroma njihovih modelov iz osnovnih sestavin, to je (empiričnih) podatkov o procesih. Pri tem je pomembno izpostaviti, da lahko rudarjenje poteka popolnoma samodejno, s pomočjo algoritmov, ki na osnovi podatkov operativnega izvajanja procesov, generirajo modele procesov. Značilnost tako nastalih modelov procesov je visoka stopnja veljavnosti (a ne nujno popolnosti) oziroma dobro prileganje dejanski izvedbi tehničnega okolja in (skoraj) realno-časovna ažurnost modelov. Formalno je rudarjenje procesov definirano kot: »*tehniške, orodja in metode odkrivanja, spremljanja in izboljševanja realnih procesov, ki temeljijo na pridobivanju znanja iz dnevnikov dogodkov informacijskih sistemov*« [4].

3 DNEVNIK DOGODKOV

Empirične podatke oziroma vhodno točko rudarjenja procesov najpogosteje predstavlja (transakcijski) dnevnik dogodkov, ki ga je možno s pomočjo algoritmov preoblikovati v model procesa, če so upoštevane naslednje predpostavke: (1) dnevnik dogodkov beleži dogajanje enega procesa; (2) posamezne dogodke (vrstice) je možno povezati s primerkom oziroma instanco procesa (angl. case, instance); (3) zapisani dogodki so rezultat izvedene aktivnosti procesa in (4) dnevnik dogodkov je popoln. Primer poenostavljenega dnevnika dogodkov, ki izpolnjuje omenjene predpostavke, je v tabelarični obliki ponazorjen v Tabela 1, kjer: (1) atribut »CaseID« predstavlja primerek procesa; (2) »Activity« predstavlja naziv izvedene aktivnosti; in (3) »time:timestamp« predstavlja časovno značko dogodka zaključka aktivnosti.

V praksi se izkaže, da vhodni podatki v rudarjenje podatkov niso tako dobro strukturirani kot prika-

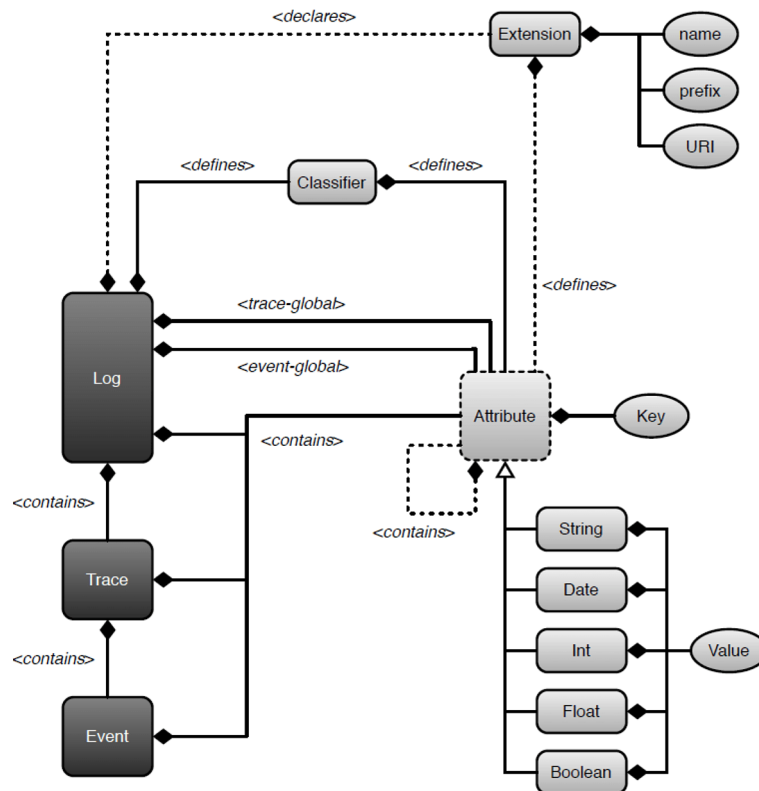
Tabela 1: Tabelarični prikaz(poenostavljenega) medicinskega dnevnika dogodkov

Case ID	Activity	time:timestamp	lifecycle: transition	CRP	DiagnosticUrinary Sediment
XJ	ER Registration	7.11.2019 08:18	complete		TRUE
XJ	ER Triage	7.11.2019 08:29	complete		
XJ	ER Sepsis Triage	7.11.2019 08:37	complete		
XJ	CRP	7.11.2019 08:51	complete	16.0	
XJ	IV Liquid	7.11.2019 09:05	complete		
XJ	IV Antibiotics	7.11.2019 10:05	complete		
XJ	Admission NC	7.11.2019 11:11	complete		
XJ	Leucocytes	8.11.2019 08:00	complete		
I	ER Registration	9.11.2019 09:21	complete		TRUE
I	ER Triage	9.11.2019 09:34	complete		
I	Leucocytes	9.11.2019 09:42	complete		
I	CRP	9.11.2019 09:42	complete	9.0	

zuje Tabela 1, (po kakovosti jih je možno umestiti v pet razredov [4]) zato je potrebno pred samo izvedbo rudarjenja podatkov le te ustrezno pripraviti za obdelavo, kar je poznano pod okrajšavo ETL (angl. Extract, Transform, Load) oziroma: (1) pridobivanje podatkov iz informacijskih rešitev; (2) sintaktična in semantična pretvorba podatkov v obliko, ki je primerna za rudarjenje procesov; (3) nalaganje podatkov v ciljni sistem, na primer podatkovno skladišče ali relacijsko podatkovno zbirko. Omenjeni postopek bistveno olajša zapisovanje dnevnikov dogodkov v

standardnih formatih, kot sta IEEE 1849-2016 XES (eXtensible Event Stream) in njegov predhodnik MXML (Mining eXtensible Markup Language). Slika 2 prikazuje Meta model standarda XES v obliki UML razrednega diagrama.

Iz zgornje slike je razvidno da XES datoteka oziroma dokument vsebuje 1 dnevnik dogodkov, ki je sestavljen iz množice sledi (angl. traces). Vsaka sled predstavlja zaporedje dogodkov (angl. event), ki pripadajo določenemu primerku izvedbe procesa. Tako dnevnik, kakor tudi sledi in dogodki, lahko imajo



Slika 2: Meta model XES [5]

množico dodatnih atributov, ki se lahko urejajo v hierarhično obliko. Določeni atributi se lahko opredelijo kot obvezni, na primer: vsaka sled mora imeti poimenovanja, vsak dogodek pa mora imeti časovno značko. Semantiko atributov določajo t.i. razširitve (angl. extensions).

Rudarjenje procesov iz dnevnikov dogodkov je povezani z nekaterimi izzivi [6]. Kot prvo, morajo biti dogodki povezani z določenim primerkom procesa, kar je lahko problematično v primeru dogodkov, ki so razpršeni v obsežni podatkovni bazi. Drugi izziv je povezan z granulacijo in formatom časovnih značk, ki se lahko razlikujejo v različnih tabelah podatkovne baze. Tretji izziv lahko predstavlja omejen doseg dnevnika dogodkov, ki morda beleži le del celotnega procesa. Četrty izziv lahko predstavljajo obsežne podatkovne zbirke poslovnih informacijskih sistemov, ki otežijo definiranje dnevnikov dogodkov za rudarjenje procesov. Peti izziv je povezan s stopnjo granulacije beleženih dogodkov, ki je lahko glede na dejanski proces večja ali manjša.

4 ALGORITMI RUDARJENJA PROCESOV

Avtomatizirano odkrivanje procesov predstavlja tehniko pridobivanja modela procesa na osnovi dnevnika dogodkov z uporabo ustreznega algoritma. Algoritem, ki se najpogosteje uporablja za ponazoritev delovanja avtomatiziranega odkrivanja procesov je algoritem α (alfa) (angl. alpha algorithm) in njegove izpeljanke ($\alpha+$, $\alpha++$, $\alpha\#$ in $\alpha\$$) [6].

Kot samo rudarjenje procesov, ima tudi algoritem α osnove v akademskem okolju, natančneje v teoriji zaporedij (multimnožice, kjer je pomembna urejenost oziroma vrstni red). Algoritem analizira relacije urejenosti med pari opravi v sledih (angl. traces) (zaporedje dogodkov, urejenih po časovnih značkah, ki pripadajo istemu primerku dnevnika dogodkov kot so: neposredni naslednik (angl. direct successor), vzročnost (angl. causality), sočasnost (angl. concurrency) in ekskluzivnost (angl. exclusiveness). Rezultat algoritma α je mreža delovnega toka »W«, ki ohranja omenjene relacije dnevnika dogodkov L, kar zapišemo kot:

$$\alpha L = W.$$

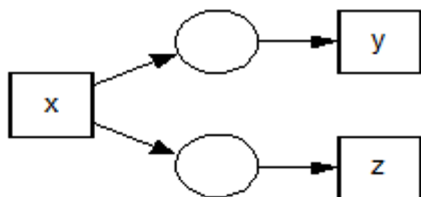
```
<log xes.version="1.0" xes.features="nested-attributes" openxes.version="1.0RC7" xmlns="http://www.xes-standard.org/">
<extension name="Lifecycle" prefix="lifecycle" uri="http://www.xes-standard.org/lifecycle.xesext"/>
<extension name="Organizational" prefix="org" uri="http://www.xes-standard.org/org.xesext"/>
<extension name="Time" prefix="time" uri="http://www.xes-standard.org/time.xesext"/>
<extension name="Concept" prefix="concept" uri="http://www.xes-standard.org/concept.xesext"/>
<extension name="Semantic" prefix="semantic" uri="http://www.xes-standard.org/semantic.xesext"/>
<global scope="trace">
<string key="concept:name" value="__INVALID__"/>
</global>
<global scope="event">
<string key="concept:name" value="__INVALID__"/>
<string key="lifecycle:transition" value="complete"/>
</global>
<classifier name="MXML Legacy Classifier" keys="concept:name lifecycle:transition"/>
<classifier name="Event Name" keys="concept:name"/>
<classifier name="Resource" keys="org:resource"/>
<string key="source" value="Rapid Synthesizer"/>
<string key="concept:name" value="exercise4.xml"/>
<string key="lifecycle:model" value="standard"/>
<trace>
<string key="concept:name" value="Case3.0"/>
<event>
<string key="org:resource" value="UNDEFINED"/>
<date key="time:timestamp" value="2009-06-04T17:21:03.471+02:00"/>
<string key="concept:name" value="a"/>
<string key="lifecycle:transition" value="complete"/>
</event>
<event>
<string key="org:resource" value="UNDEFINED"/>
<date key="time:timestamp" value="2009-06-04T17:22:03.471+02:00"/>
<string key="concept:name" value="b"/>
<string key="lifecycle:transition" value="complete"/>
</event>
<event>
<string key="org:resource" value="UNDEFINED"/>
<date key="time:timestamp" value="2009-06-04T17:23:03.471+02:00"/>
<string key="concept:name" value="d"/>
<string key="lifecycle:transition" value="complete"/>
</event>
<event>
<string key="org:resource" value="UNDEFINED"/>
<date key="time:timestamp" value="2009-06-04T17:24:03.471+02:00"/>
<string key="concept:name" value="c"/>
<string key="lifecycle:transition" value="complete"/>
</event>
<event>
<string key="org:resource" value="UNDEFINED"/>
<date key="time:timestamp" value="2009-06-04T17:25:03.471+02:00"/>
<string key="concept:name" value="e"/>
<string key="lifecycle:transition" value="complete"/>
</event>
</trace>
</log>
```

Slika 3: Izsek dnevnika dogodkov, zapisanega v formatu XES

Kot primer vzemimo vzporedni razcep oziroma razcep 'in' (angl. AND split) v mreži delovnega toka (Slika 4).

Kot je razvidno iz Slika 4, je vzporedni razcep (angl. AND split) sestavljen iz treh medsebojno povezanih opravil {x,y,z} v naslednjih zgoraj predstavljenih relacijah: vzročnost med opraviloma <x,y> in <x,z> ter relacijo sočasnosti med opraviloma <y,z>, kar zapišemo kot:

$$x \rightarrow y, x \rightarrow z \wedge y // z,$$



Slika 4: Vzporedni razcep predstavljen v Petrijevi mreži

Podobno kot za vzporedni razcep, algoritem α definira preslikave še za druge osnovne vzorce kontrolnega toka (angl. workflow patterns), kot so: zaporedje (angl. sequence), vzporedno združevanje (angl. AND join), ekskluzivni razcep (angl. XOR split) in ekskluzivno združevanje (angl. XOR join).

$$L_1 = [\langle a, b, c, d \rangle^3, \langle a, c, b, d \rangle^2, \langle a, e, d \rangle]$$

	a	b	c	d	e
a	# _{L₁}	→ _{L₁}	→ _{L₁}	# _{L₁}	→ _{L₁}
b	← _{L₁}	# _{L₁}	_{L₁}	→ _{L₁}	# _{L₁}
c	← _{L₁}	_{L₁}	# _{L₁}	→ _{L₁}	# _{L₁}
d	# _{L₁}	← _{L₁}	← _{L₁}	# _{L₁}	← _{L₁}
e	← _{L₁}	# _{L₁}	# _{L₁}	→ _{L₁}	# _{L₁}

Slika 5: Sledi dnevnika dogodkov L1 (zgoraj) in njegov odtis (spodaj)

Na osnovi odtisa dnevnika dogodkov (angl. footprint) (Slika 5), ki predstavlja relacije kavzalnih (\rightarrow), sočasnih (II) in ekskluzivnih (#) relacij vseh parov dogodkov vseh sledi dnevnika dogodkov L (Slika 5), nato algoritem α izdelava celovito mrežo delovnega toka v naslednjih osmih korakih: (1) določitev množice aktivnosti oziroma prehodov Petrijeve mreže, ki se pojavijo v dnevniku dogodkov; (2) določitev vhodnih prehodov, ki predstavljajo začetne prehode v sledih dnevnika dogodkov; (3) določitev izhodnih prehodov, ki predstavljajo končne prehode v sledih dnevnika dogodkov; (4) identifikacija parov med-vzročnih (angl. inter-causal) in notranje-sočasnih (angl. intra-concurrent) aktivnosti in (5) redukcija pridobljenih množic na način, ki ohranja le nad-množice; (6) definiranje mest P , ki jih dodamo med prehode; (7) sintaktično ustrezno povezovanje mest in prehodov z relacijo toka F (ang. flow relation) in (8) združevanje rezultatov v z izrisom mreže delovnega toka.

Jedro algoritma α predstavljata četrti in peti korak, kjer je poglobitveni izziv določitev mest in pripadajočih povezav v mreži delovnega toka, na način, da se določita množici vhodnih in izhodnih prehodov zgoraj omenjenih mest (angl. places). Omenjeno je možno doseči s preureditvijo vrstic in stolpcev odtisa dnevnika dogodkov (Slika 5), na način, da se oblikujeta dve množici dogodkov, katerih elementi so med množicama vzročno povezani (povezava v mreži delovnega toka), medtem, ko znotraj posameznih množic niso v vzročni zvezi.

Med omejitve algoritma α , spadajo t.i. kratke zanke v procesu, ki imajo dolžino ena (v sledi je več zaporednih enakih dogodkov) ali dva (v sledi je več zaporedno ponavljajočih parov dogodkov). V takšnih primerih algoritem α ne zazna povezav s pripadajočimi prehodi.

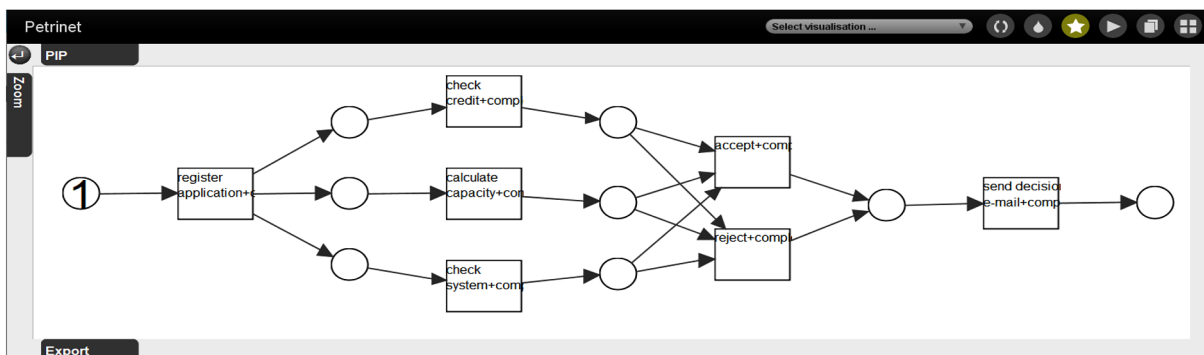
Poleg algoritma α obstaja še množica algoritmov za odkrivanje procesov, kot so na primer [6]: heuristično rudarjenje (angl. heuristic mining), genetsko rudarjenje procesov (angl. genetic process mining), področno rudarjenje (angl. region-based mining), induktivno rudarjenje (angl. inductive mining) in njihove številne variacije, vključno z lastniškimi algoritmi, ki so vključeni v komercialne rešitve (na primer, Celonis). Orodje Apromore uporablja algoritem razcepljenega rudarjenja (angl. split miner) [7].

Razlog za poplavo algoritmov je v pomanjkljivostih posameznih algoritmov, kot so [6]: (1) stopnja prilaganja (angl. fitness) ki opredeljuje, kako dobro se izdelan model procesa prilagaja podatkom v dnevniku oziroma kako dobro je možno le tega replicirati iz izdelanega modela procesa; (2) enostavnost (angl. simplicity), ki opredeljuje preprostost izdelanega modela procesa; (3) generalizacija (angl. generalization), ki opredeljuje kako dobro algoritem predvideva dejanski model, ki morda presega zapise v dnevniku dogodkov (le ti najverjetneje niso popolni) in (4) natančnost (angl. precision), ki je nasprotna lastnost generalizacije in določa, v kolikšni meri model opisuje natančno to kar je opredeljeno v sledih dnevnika dogodkov oziroma, da model ne dovoljuje obnašanja, ki ni evidentirano v dnevniku dogodkov.

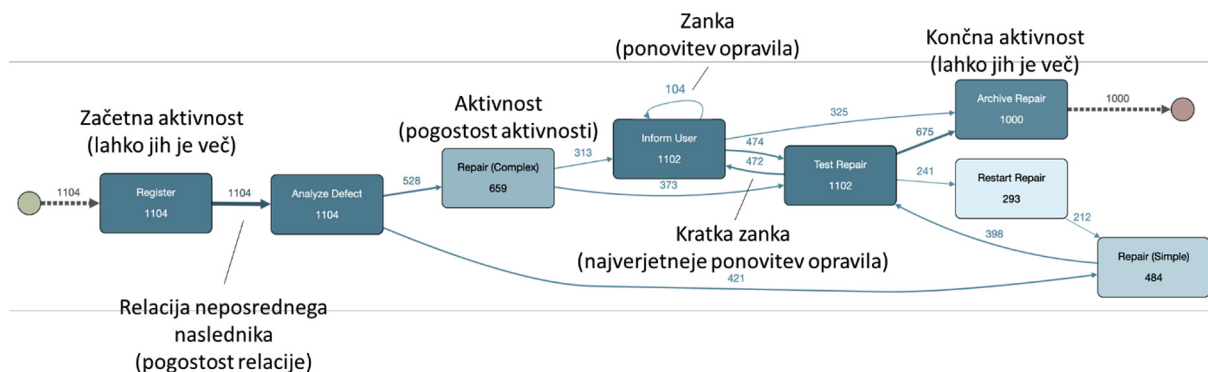
5 VIZUALIZACIJA PROCESOV

Modeli procesov, ki jih generirajo algoritmi za rudarjenje procesov, so najpogosteje vizualizirani v eni izmed naslednjih procesnih notacij: mreže delovnih tokov (angl. workflow maps), procesne mape (angl. process maps) in BPMN diagrami (Business Process Model and Notation).

Mreže delovnih tokov so podzvrst Petrijevih mrež (natančneje, imajo natanko en začetek in konec ter bi



Slika 6: Mreža delovnega toka v okolju ProM

Slika 7: Primer procesne mape, ki jo je generiralo okolje Apromore³

v primeru vpeljave prehoda med začetkom in koncem izkazovale lastnost »močne povezanosti«. Zaradi močnega teoretičnega ozadja in možnosti tako formalnega, kot vizualnega prikaza so zelo primerne za algoritmično obdelavo, prikazujejo pa lahko tako statično kakor tudi dinamično sliko procesa. Slika 6 prikazuje mrežo delovnega toka procesa pridobivanja kredita, ki je bila pridobljena iz testnega dnevnika dogodkov v formatu XES¹ s pomočjo algoritma Alfa implementiranega v orodju ProM².

Slabost mreže delovnih tokov se kaže predvsem v njihovi praktični uporabnosti, saj zaradi konceptualne preprostosti (vsebujejo le tri teoretične koncepte: prehod, stanje in povezavo) ne nudijo neposredne podpore za višje-nivojske koncepte poslovnih procesov in so tako slabše razumljive za poslovne uporabnike.

Procesne mape (angl. process map, directly follows graph) so neformalna in preprosta notacija za prikaz poslovnih procesov, ki najpogosteje vključujejo le dva gradnika: (1) aktivnost, ki je predstavljena z vozliščem (najpogosteje pravokotnik) in (2) povezavo med vozlišči, ki predstavlja zaporedje izvedbe aktivnosti oz. pripadajočih dogodkov. V primeru rudarjenja procesov so procesne mape priljubljene zaradi preprostosti in dejstva, da rudarjenje procesov najpogosteje rezultira prav v modelih, ki predstavljajo množice aktivnosti in relacij med njimi.

Kot je razvidno iz Slika 7, so elementi procesnih map pogosto dopolnjeni še s »statističnim slojem«, na primer: trajanje ali pojavnost elementa. Le te je moč pridobiti s tehnikami rudarjenja procesov, predstavlja pa pomembne informacije za procesne odločitve.

Zaradi kompleksnosti realnih procesov, predvsem iz vidika števila variacij izvedb enega procesa (varianta predstavlja vse sledi procesa z istim zaporedjem aktivnosti), so običajno tudi procesne mape kompleksne, zato orodja običajno nudijo možnosti abstrakcije pridobljenih modelov in sicer se najpogosteje omejuje prikaz glede na pogostost izvedbe aktivnosti ali povezav (na primer prikaz najpogostejše ali najredkeje izvedenih variant procesa). Slika 8 prikazuje procesno mapo, ki jo generira okolje Celonis. Na desni strani so v grafični obliki ponazorjene frekvence variant izvedbe procesa in območje ter delež variant, ki so ponazorjene v prikazani procesni mapi. Iz Slika 8 je tako razvidno da procesna mapa prikazuje potek treh najpogostejše izvajanih sledi, ki skupno pokrivajo 77% vseh primerkov procesa.

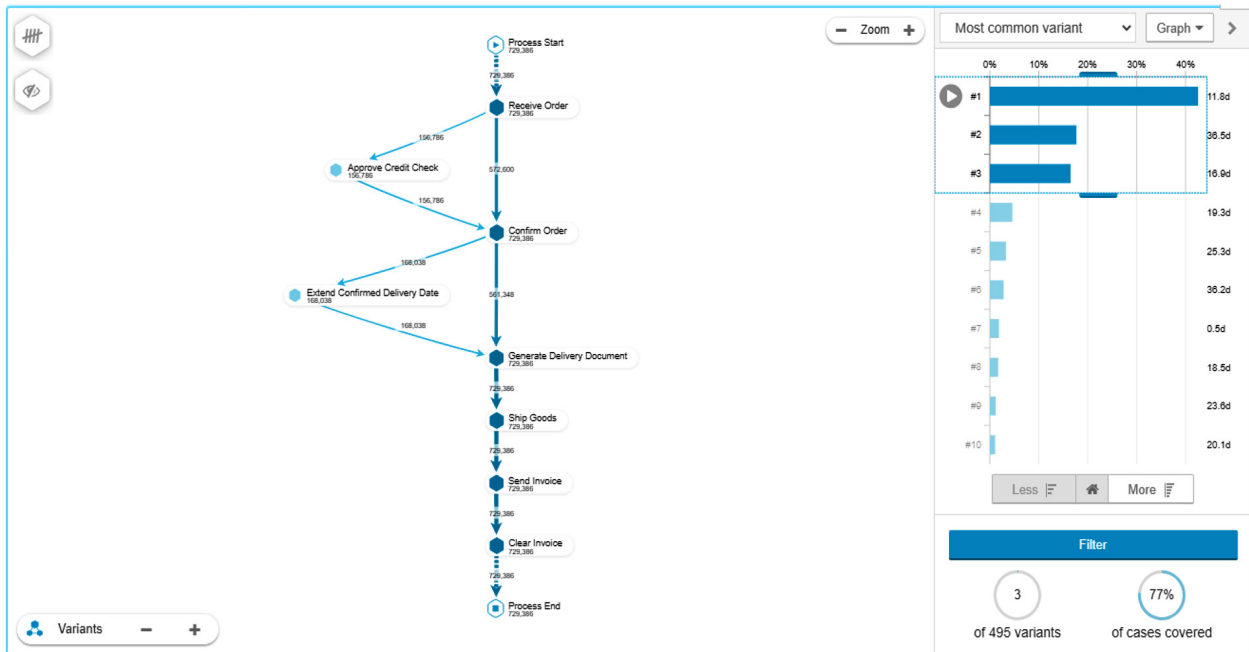
Drug prikaz istega procesa oziroma dnevnika dogodkov prikazuje Slika 9, ki vizualizira 495 variant procesa, ki pa skupaj predstavljajo le 8% vseh izvedb oziroma primerkov procesa.

Med slabosti procesnih map spada pomanjkanje procesnih konceptov, na primer težje je prepoznati pogojne in vzporedne tokove. Alternativa, ki odpravlja omenjeno slabost je uporaba BPMN 2.0, ki je de-facto in ISO standard (ISO/IEC 19510:2013) za modele procesov. Zaradi standardiziranosti tako notacije, kot tudi meta-modela, nudi le ta številne prednosti, kot so: bogat nabor konceptov poslovnih procesov, ki jih notacija podpira, prenosljivost modelov med orodji, možnosti simulacije in avtomatizacije procesov. Slika 10 prikazuje segment BPMN modela, ki ga je generiralo okolje Apromore na dnevniku iz Tabela

¹ <http://www.promtools.org/prom6/downloads/FutureLearn%20-%20Process%20mining%20with%20ProM%20-%20Event%20logs.zip>.

² <https://www.promtools.org/>.

³ <https://apromore.org/>.



Slika 8: Abstrakcije modelov procesov v okolju Celonis⁴ s prikazom najbolj pogostih variant

1 z uporabo »split miner« algoritma, ki je sposoben zaznati koncepte BPMN brez dodatnih transformacij med notacijami [3].

Slabost uporabe notacije izhaja predvsem iz omejitve orodij za rudarjenje procesov, ki generirajo že na videz drugačne modele, kot jih izdelajo analitiki in tudi v »omejenosti« algoritmov, ki so zmožni pre-

poznati le osnovne koncepte in vzorce kontrolnega toka. Zato je običajno izkoriščen le majhen del gradnikov notacije BPMN 2.0.

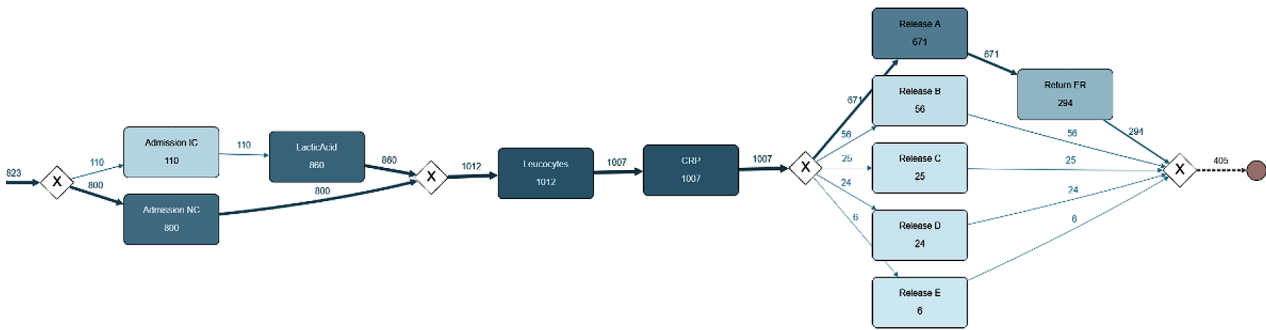
6 ZMOŽNOSTI RUDARJENJA PROCESOV

Ključna lastnost rudarjenja procesov je močna povezanost med modeli procesov in realnostjo, ki je



Slika 9: Abstrakcije modelov procesov v okolju Celonis s prikazom najredkeje izvajanih variant

⁴ <https://www.celonis.com/>.

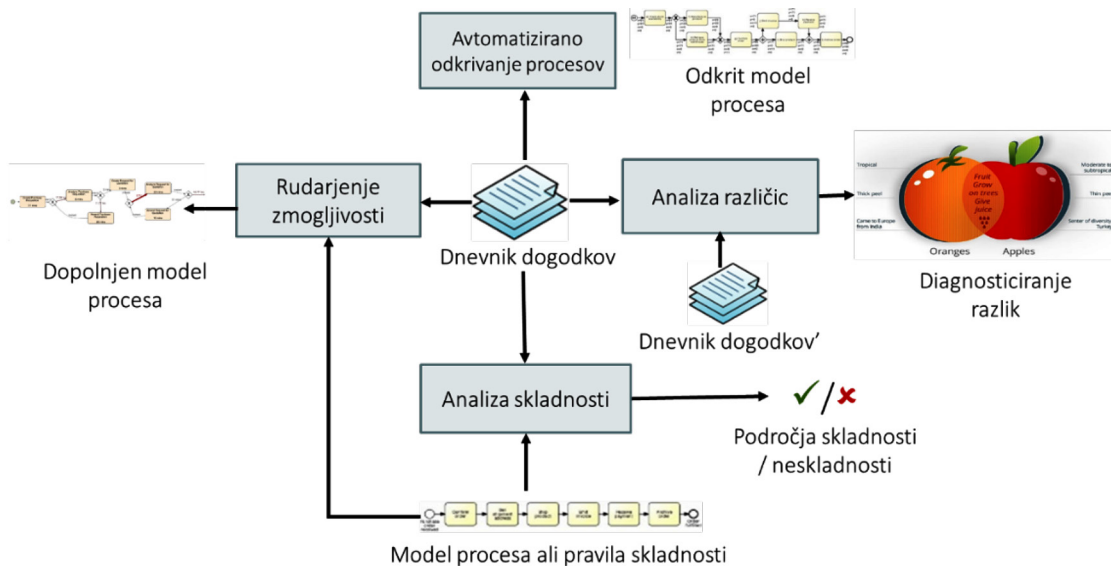


Slika 10: Segment modela BPMN, z eksplicitnimi elementi kontrolnega toka (BPMN), ki ga je generiralo okolje Apromore

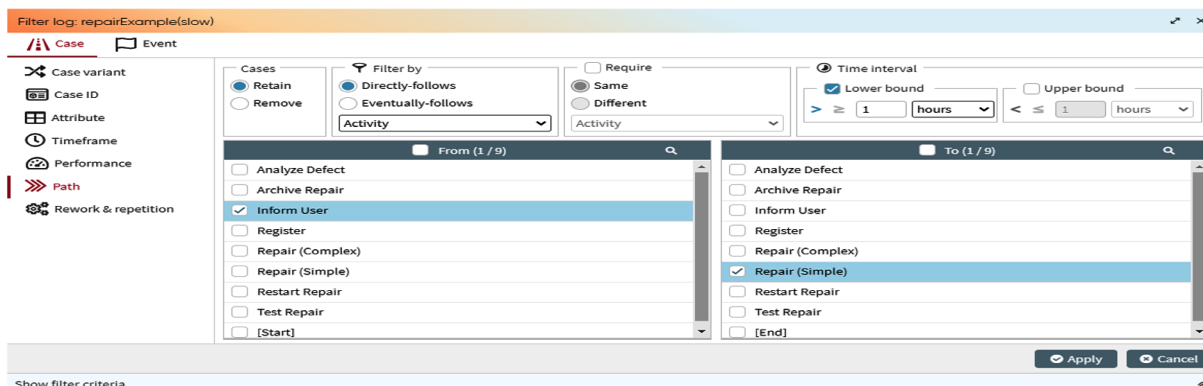
lahko definirana na tre načine: »play-in«, »play-out« in »replay«. Primer »play-in« je generiranje modela procesa na osnovi dnevnika dogodkov. »Play-out« predstavlja nasprotno aktivnost in sicer generiranje dnevnika dogodkov na osnovi definirane modela procesa. Zadnji način »replay« predstavlja kombinacijo predhodnih načinov, kjer se na osnovi dnevnika dogodkov generira model procesa, le tega pa nato uporabimo za generiranje novega dnevnika dogodkov. Način »replay« je primeren za ugotavljanje skladnosti med dejanskim in generiranim dnevnikom dogodkov ali za ponazoritev dodatnih informacij v modelu procesa (na primer: frekvenca in trajanje). Zato lahko s tehnikami in orodji za rudarjenje procesov poleg avtomatiziranega odkrivanja procesov, ki je bilo podrobneje predstavljeno v predhodnem poglavju, izvajamo še druge procesne aktivnosti (Slika 11), ki so predstavljene v nadaljevanju.

7 PREVERJANJE SKLADNOSTI

Preverjanje skladnosti (angl. conformance checking) omogoča primerjavo izvajane modela (oziroma dnevnika dogodkov) z definiranimi poslovnimi pravili ali definiranim modelom procesa (angl. prescribed process model) in je s tem relevantno za usklajevanje poslovanja z zahtevami in za presojanje poslovanja (angl. auditing). Primeri poslovnih pravil, ki jih lahko preverjamo, so: (1) omejitve kontrolnega toka, kot je analiza izvajanja obveznih aktivnosti (na primer: odobritve zahtevkov ali obvezna kontrola kakovosti); (2) omejitve nivoja storitev oziroma SLA (angl. service level agreement), kot je najdaljši dovoljen čas izvajanja aktivnosti, regije ali procesa; (3) omejitve virov kot je »ločevanje dolžnosti« (na primer: ista oseba ne sme izvesti dveh zaporednih aktivnosti) in (4) identifikacija redkih primerkov izvedbe, ki so potencialno neskladni s



Slika 11: Zmožnosti rudarjenja procesov [8]



Slika 12: Časovna analiza skladnosti v okolju Apromore

poslovnimi pravili. Rezultat analize preverjanja skladnosti je seznam odstopanj od pravil ali definiranega modela procesa. Slika 12 prikazuje analizo časovne skladnosti izvedbe storitve popravila, kjer imamo na primer opredeljeno časovno pravilo »čakanje med informiranjem uporabnika in izvedbo preprostega opravila ne sme trajati več kot eno uro«. Orodje Apromore lahko odkrije vse primerke procesa, ki trajajo dlje od navedenega.

8 RUDARJENJE ZMOGLJIVOSTI

Z rudarjenjem zmogljivosti pridobimo dodatne informacije o modelih procesov, ki lahko vodijo v njihove izboljšave. Rezultat rudarjenja zmogljivosti so grafi zmogljivosti in modeli procesov, ki so dopolnjeni z informacijami kot so trajanja aktivnosti ali pogledi na procese iz vidika določenega vira. Na tak način lahko odkrivamo ozka grla (angl. bottleneck analysis) v izvajanju, kot so: (1) aktivnost je počasna zato predstavlja ozko grlo; (2) vse vhodne povezave v aktivnost so počasne, zato je najverjetneje ozko grlo vir, ki mu je dodeljena izvedba aktivnosti; in (3) počasna je predaja dela (angl. handoff) med dvema viroma (odebeljena povezava »21 mins« na Slika 13).

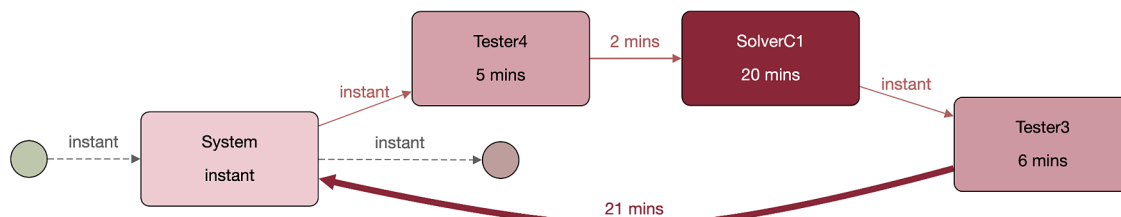
Pogled virov lahko prav tako omogoči identifikacijo potencialno preobremenjenih udeležencev pro-

cesa (velika frekvenca ali visoko trajanje izvedbe) ali premalo obremenjenih udeležencev procesa (nizka frekvenca ali kratko trajanje izvedbe).

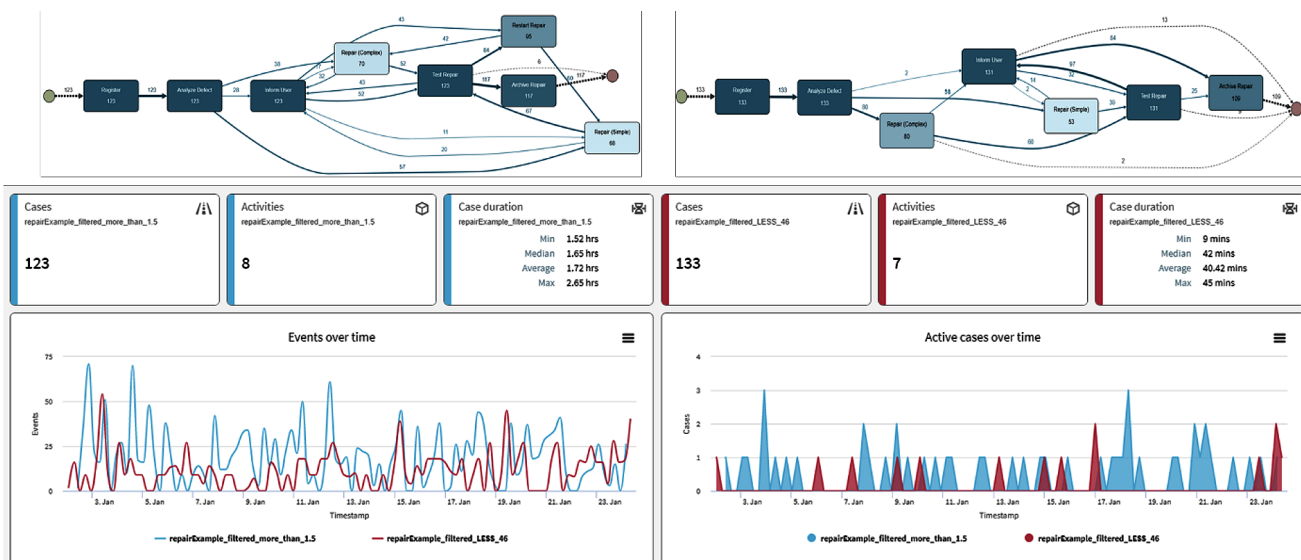
Na zmogljivost delovanja lahko prav tako bistveno vplivajo ponovitve opravil, ki so lahko posledica slabo ali nepopolno opravljenega dela, ki se kaže kot: (1) zankanje ene aktivnosti (angl. self loop); (2) kratke zanke, kjer se izmenično večkrat izvedeta dve aktivnosti (angl. short loop, ping-pong behavior); ali (3) posredne ponovitve, kjer se po istem vzorcu ponavlja ista skupina aktivnosti (angl. indirect repetition).

9 ANALIZA RAZLIČIC

Analiza različic (angl. variant analysis) temelji na primerjavi dveh ali več različic dnevnikov dogodkov istega procesa, ki tako predstavljajo različne variante procesa (na primer, primerjava vseh primerkov procesa, ki so se uspešno zaključili glede na neuspešne). Primerjava variant procesov poda vpogled na vprašanja tipa »zakaj?«, na primer: (1) zakaj se določeni primerki procesa (ki smo jih združili v eno izmed variant procesa) izvajajo hitreje kot drugi? (2) Zakaj se določeni primerki procesa uspešno zaključijo, medtem ko se drugi neuspešno? (3) Zakaj je vir, ki je vključen v izbrano varianto procesa manj učinkovit kot drugi?



Slika 13: Analiza ozkih grl iz pogleda virov v okolju Apromore



Slika 14: Primerjava modelov primerkov istega procesa, ki trajajo manj kot 45 minut (levo) s primerki ki trajajo več kot 90 minut (desno) in statistična primerjava v nadzorni plošči (spodaj)

Odgovore na zgornja vprašanja lahko pridobimo z enim izmed naslednjih pristopov analize različic procesov: (1) analizo metrik uspešnosti (angl. performance measures) različic procesov; (2) analizo različic procesov glede na čas izvedbe (na primer: analiza in primerjava različic procesov, ki so se izvajali pred in med pandemijo); (3) analiza različic procesov glede na lastnosti primerkov (na primer: vrsta izdelka, segment kupcev, geografsko področje, ipd.)

10 ZAKLJUČEK

Digitalna preobrazba in optimizacija poslovanja je prepletena s tehnološkimi inovacijami, ki morajo zagotavljati hiter in veljaven vpogled v delovanje organizacij, temelječ na realnih podatkih, na način, ki je razumljiv vsem vpletenim. Med ključne tehnike za doseganje navedenega spada rudarjenje procesov, ki izkorišča močno povezanost med fizično in informacijsko realnostjo (digitalni dvojček) poslovanja, v kateri se poslovni dogodki beležijo v realnem času, le ti pa so uporabljeni za usmerjanje, prilagajanje in nadzor poslovnih procesov. Rudarjenje procesov postaja del rutine večjih podjetij razvitih držav [3], k čemu je pripomogel tudi širok nabor »enterprise-ready« orodij za rudarjenje procesov. Le ta so dostopna različnim vrstam in potrebam organizacij in omogočajo relativno nizek vstopni prag v aktivnosti rudarjenja procesov. Slabost vpeljave orodij je pogosto, da so le ta v podjetjih implementirana v omejenem obsegu in zato ne pokrivajo celotnega poslovanja.

Poglavitna izziva širše vpeljave rudarjenja procesov ostajata kakovost podatkov in človeški dejavniki. Izkušnje kažejo, da je okoli 80% časa potrebnega za lociranje, izbiranje, pridobivanje in transformacijo podatkov, pogosto pa omenjene aktivnosti odkrijejo tudi težave s kakovostjo podatkov, ki jih je potrebno odpraviti neodvisno od rudarjenja procesov. Človeški dejavniki so pogosto povezani z nepoznavanjem področja rudarjenja procesov in »strahom« pred odkritjem dejanskih procesov, ki bi lahko izpostavili pomanjkljivo vodenje, neučinkovitosti ali neskladnosti med operativnim delovanjem in predpisi.

V preteklem desetletju se je rudarjenje procesov uveljavilo predvsem kot tehnologija za odkrivanje in vizualizacijo realnih procesov, v prihodnosti pa se obeta razvoj področja v smer optimizacije procesov, podprte z umetno inteligenco [9]. Prvi trend je robotsko rudarjenje procesov (angl. robotic process mining), ki omogoča odkrivanje rutinskih digitalnih opravil iz dnevnikov uporabniških vmesnikov (angl. UI log), s pomočjo katerih se lahko nato generirajo avtomatske skripte. Drugi trend je vzročno rudarjenje procesov (angl. causal process mining), ki omogoča odkrivanje vzročno-posledičnih povezav med specifikami primerkov procesov in njihovimi rezultati (na primer, primerek procesa se izvede drugače v kolikor je stranka iz določene regije). Odkrivanje omenjenih povezav lahko vodi v izboljšave procesov in višje zadovoljstvo deležnikov. Tretji trend je »kaj – če« rudarjenje procesov (angl. what-if process mi-

ning), ki namesto analize obstoječega stanja omogoča simulacije delovanja procesov v primeru spremembe določenih vhodnih podatkov. Na primer, »kako se bo odzval proces, če se število zahtevkov podvoji?«. Četrty trend je predpisano ali normativno spremljanje procesov (angl. prescriptive process monitoring), ki s pomočjo strojnega učenja omogoča predvidevanja (negativnih) rezultatov primerkov (na primer, ali se bo določen primerek zaključil pravočasno). Peti trend rudarjenja procesov v prihodnosti je avtomatizirano izboljševanje procesov (angl. automated process improvement), ki za razliko od trenutnih pristopov omogoča avtomatizirano vpeljavo sprememb procesov, ki izboljšujejo izbran indikator (na primer: stroški, čas in stopnja napak).

Izjemne razmere, kot je pandemija, so pokazale kako pomembna je digitalizacija, procesi, zanesljivi podatki in zmožnost hitrih prilagoditev novim razmeram. Nagle spremembe v poslovanju se odražajo v naglih spremembah procesov, ki jih je možno zagotoviti z ustrežno procesno avtomatizacijo in spremljati s tehnikami in tehnologijami rudarjenja procesov.

LITERATURA

- [1] M. Dumas, M. L. Rosa, J. Mendling, in H. A. Reijers, *Fundamentals of Business Process Management*. Springer Berlin Heidelberg, 2018. [Na spletu]. Dostopno na: <https://books.google.si/books?id=KgVTDwAAQBAJ>
- [2] G. Jošt, J. Huber, M. Heričko, in G. Polančič, »Improving cognitive effectiveness of business process diagrams with opacity-driven graphical highlights«, *Decision Support Systems*, let. 103, str. 58–69, nov. 2017, doi: 10.1016/j.dss.2017.09.003.
- [3] M. Kerremans, S. Searle, T. Srivastava, in K. Iijima, »Market Guide for Process Mining«, Gartner, sep. 2020. Pridobljeno: sep. 08, 2021. [Na spletu]. Dostopno na: <https://www.gartner.com/doc/reprints?id=1-24ARMY34&ct=201002&st=sb>
- [4] W. van der Aalst *idr.*, »Process Mining Manifesto«, v *Business Process Management Workshops*, let. 99, F. Daniel, K. Barkaoui, in S. Dustdar, Ur. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, str. 169–194. doi: 10.1007/978-3-642-28108-2_19.
- [5] G. Acampora, A. Vitiello, B. N. D. Stefano, W. M. P. van der Aalst, C. W. Günther, in E. Verbeek, »IEEE 1849: The XES Standard: The Second IEEE Standard Sponsored by IEEE Computational Intelligence Society [Society Briefs]«, *IEEE Comput. Intell. Mag.*, let. 12, št. 2, str. 4–8, 2017, doi: 10.1109/MCI.2017.2670420.
- [6] W. M. P. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. 2016 edition. New York, NY: Springer, 2016.
- [7] A. Augusto, R. Conforti, M. Dumas, M. La Rosa, in A. Polyvyanyy, »Split miner: automated discovery of accurate and simple business process models from event logs«, *Knowl Inf Syst*, let. 59, št. 2, str. 251–284, maj 2019, doi: 10.1007/s10115-018-1214-x.
- [8] »Process Mining 101«. Pridobljeno: sep. 09, 2021. [Na spletu]. Dostopno na: <https://apomore.org/process-mining-101/>
- [9] M. Dumas, »Process Mining in 2021 and Beyond«. mar. 09, 2021. Pridobljeno: sep. 09, 2021. [Na spletu]. Dostopno na: <https://apomore.org/whitepaper-process-mining-in-2021-and-beyond/>

Gregor Polančič je znanstveni svetnik in izredni profesor na področju informatike. Spada med vodilne raziskovalce na področjih modeliranja poslovnih procesov in konceptualnega modeliranja v informatiki, tehnik in tehnologij upravljanja poslovnih procesov, tehnologij komuniciranja in sodelovanja. Bil je gostujoči profesor in raziskovalec na številnih tujih akademskih institucijah, recenzent številnih vrhunskih znanstvenih revij in član več odborov znanstvenih srečanj. Njegova bibliografija obsega preko 300 zapisov, od tega preko 30 izvornih znanstvenih člankov s faktorjem vpliva.

Mateja Kocbek Bule je asistentka in doktorska študentka na Fakulteti za elektrotehniko, računalništvo in informatiko, Univerze v Mariboru. Med njene interesne dejavnosti spada predvsem področje upravljanja poslovnih procesov.