

ANALÝZA VÝROBNÍHO PROCESU A PŘEDPŘÍPRAVA DAT PRO PROCESS MINING

Analysis of production process and data preprocessing for process mining

Miroslav Dišek¹, Dominik Musil², Jan Kolesár³

¹ Slezská univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné,
Univerzitní nám. 1934/3, 73340 Karviná
Email: O150611@opf.slu.cz

² Slezská univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné,
Univerzitní nám. 1934/3, 73340 Karviná
Email: O150860@opf.slu.cz

³ Slezská univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné,
Univerzitní nám. 1934/3, 73340 Karviná
Email: O160358@opf.slu.cz

Abstrakt: Cílem tohoto příspěvku je analýza výrobního procesu automobilové společnosti, sběru dat a následná příprava výstupních dat pro jejich další analýzu prostřednictvím process miningu. Začátek publikace obsahuje stručné představení společnosti a přiblížení výrobního procesu. Po identifikaci potřebných dat jsme vytvořili potřebnou strukturu dat pomocí počáteční konverze, pomocí které jsme data sjednotili a připravili jsme je pro následné kroky, mezi které například patří finální konverze či příprava potřebné struktury pro nástroj DISCO.

Abstract: Main objective of this publication is analysis of the automobile company's manufacturing process, gathering and preparation data for next analysis using process mining. Beginning of the publication consist of brief company introduction and description of manufacturing process. After the identification needed data we created required structure using initial conversion which merged the data. Then we prepared data for the following steps, for instance final conversion or preparation required structure for DISCO tool.

Klíčová slova: analýza dat, data, data mining, informační systém, process mining, SGS, struktura, výrobní proces

Keywords: analysis of data, business, data, data mining, information system, manufacturing process, process mining, SGS, structure

JEL classification: C88, L62

Úvod

V současnosti se k tradičním výrobním faktorům přidávají i informace. Důležitost informací neustále roste, hovoříme o informační společnosti. Globální komunikační infrastruktura umožňuje vytvoření zhodnocené sítě z hlediska dat, informací a znalostí. Růst dynamiky trhu, výrobních technologií, globalizace trhu a volný přístup k informacím přináší rostoucí složitost rozhodování a zvýšené nároky na management firem. Pro správné a rychlé rozhodování jsou

nutné včasné informace o vnitropodnikových procesech a aktivitách, ale také informace z okolí organizace, které vytváří podmínky, ale mnohdy i pravidla pro činnost organizací.

Pro podporu optimálního rozhodování je nezbytný kvalitní informační systém, který umožňuje rychlé a efektivní zpracování informací, členění a poskytování v ucelené struktuře tak, aby byly k dispozici ve chvíli, kdy jsou nutné při rozhodování managementu firmy.

Informační systém je komplex lidí, informací a systému řízení chodu IS, který zabezpečuje těsné a logické propojení na prostředí, systému organizace práce spojeného s provozem a využitím IS, technických prostředků a metod zabezpečujících sběr, přenos, aktualizaci, uchování a další zpracování dat pro tvorbu a prezentaci informací pro potřeby uživatelů a použité informační technologie. Cílem je získávání, spravování a poskytování potřebných informací na odpovídající místa ve vhodném čase, v potřebném rozsahu a vhodné formě. Kvalita informačního systému se odvíjí od všech fází jeho životního cyklu. Vybudování kvalitního informačního systému je nelehkým úkolem (Vaněk, Šperka, 2013), v některých případech je klíčovou součástí pro samotnou existenci či růst podniku.

Podnikový informační systém je komplexní nástroj, který je nápomocen při operativním řízení podniku ale i při dlouhodobém strategickém plánování. Po úspěšném a kvalitním zavedení IS do podniku se nám zvýší efektivita práce, konkurenceschopnost, ziskovost společnosti, a hlavně úspora nákladů a času (Basl, Blažiček, 2012).

Nicméně organizace mají problémy při analyzování a využívání těchto informací. Cílem process miningu je použít údaje o událostech a extrahovat informace o souvisejícím procesu. Jednou z výhod process miningu je možnost automaticky znázornit model procesu, sledující například události zaznamenané podnikovým informačním systémem (van der Aalst, 2011). Process mining je relativně nový a rychle se rozvíjející vědní obor, který postupně nachází uplatnění při optimalizaci procesů ve výrobě, logistice, službách atd. Z hlediska softwarových nástrojů patří k nejznámějším opensource programům ProM (Rozinat, van der Aalst, 2006) a z komerčních pak DISCO (Gunther, Rozinat, 2012) od firmy Fluxicon.

Process mining se zaměřuje na detekci, analýzu a optimalizaci podnikových (výrobních) procesů na základě dat z logovacích souborů, které jsou dostupné v informačních systémech. Process mining reprezentuje dosud chybějící spojení mezi klasickou analýzou podnikových procesů a dolováním dat (van der Aalst, 2011).

Tento příspěvek je dílčím výstupem SGS projektu “Pokročilé metody získávání dat a simulační techniky v oblasti podnikových procesů”, který navazuje na dílčí interdisciplinární výzkum v oblasti simulace podnikových procesů a process miningu, který probíhá na Katedře informatiky a matematiky a na Katedře podnikové ekonomiky a managementu s cílem zapojit do vědeckého výzkumu studenty doktorského studia a studenty navazujícího magisterského studia. V rámci spolupráce s podnikatelským prostředím budou analyzována reálná data v podobě výrobních procesů, která jsou uložena v logovacích souborech informačních systémů. V našem článku se používají konkrétní data společnosti, která se zabývá výrobou nákladních automobilů, a zaměříme se na určitý výrobní proces ve firmě. Tyto procesy generují data, která se zapisují do relačních databází postupně a našim úkolem bylo data konvertovat z těchto relačních databází a následně je zredukovat. První kapitola se bude zabývat seznámením s výrobním procesem automobilové společnosti a sběrem dat. V další kapitole jsme přiblížili strukturu dat uložených v databázi na SQL Serveru. Poslední část článku popisuje přípravu dat pro Process mining. Na základě této

přípravy budou data výrobního procesu zhodnoceny a připraveny pro následné zpracování nástrojem DISCO.

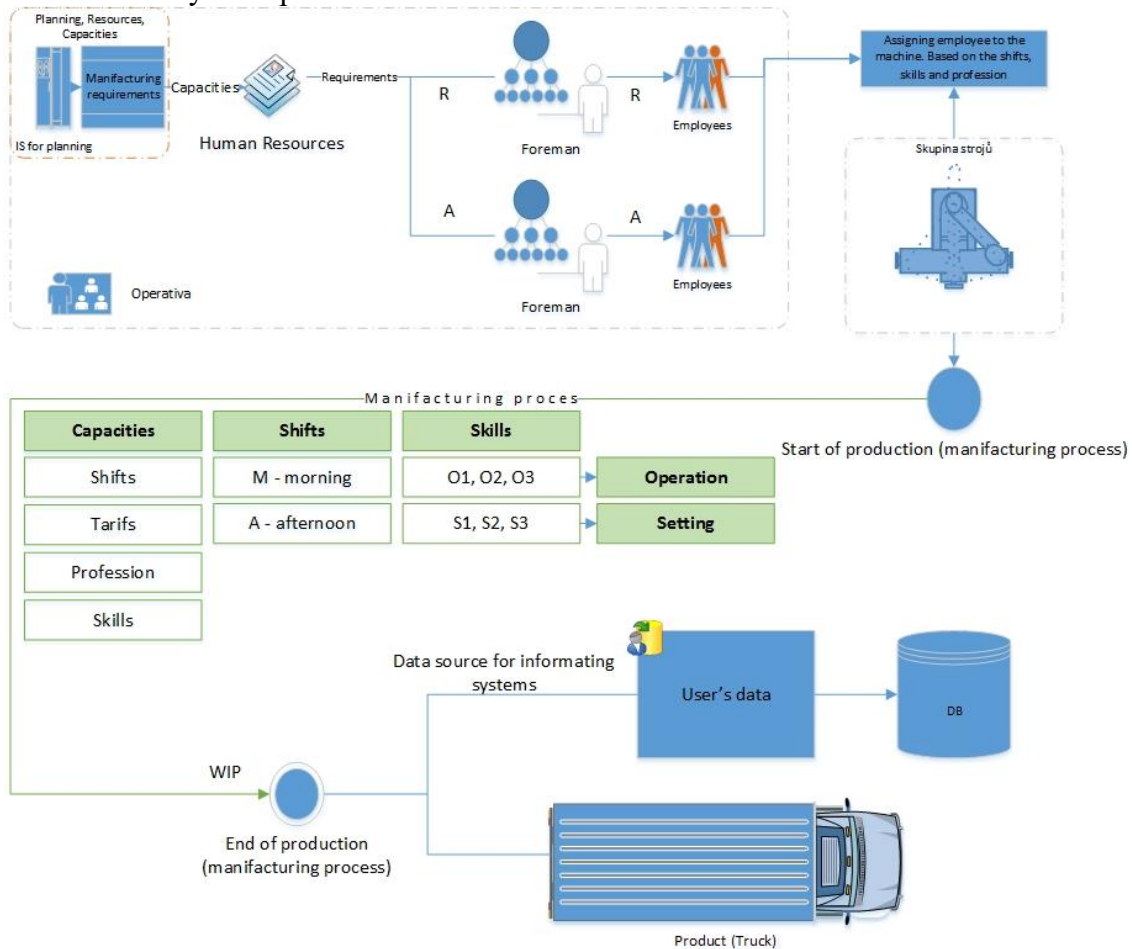
1 Seznámení s výrobním procesem automobilové společnosti a sběrem dat

Automobilová společnost je světově známou společností v oblasti automobilového průmyslu. Její zaměření je především na nákladní automobily, ale také se zaměřuje na zakázkovou výrobu či speciální vojenské vozy. Motto společnosti zní - „Vše souvisí se vším“. Celý výrobní proces je založen převážně na plánované produkci, materiálové bilanci a jeho toku. Pracovníci operativního oddělení kontrolují výrobní proces za pomoci různých informačních systémů, které hlídají například materiálovou bilanci, výrobní objednávky a další. Popis částí výrobního procesu je pro lepší pochopení znázorněn na obrázku č. 1.

Výrobní proces se dělí:

- Plánování výroby a plánovaná produkce
- Operativa – produkční dispečink
- Vlastní zdroje – zaměstnanci, stroje, přípravky
- Materiálová bilance a jeho tok
- Výrobní proces – výrobní cyklus
- Vytvoření výstupních dat a komunikace s produkčními a správními informačními systémy

Obrázek 1: Výrobní proces a sběr dat



Zdroj: Vlastní zpracování

2 Struktura dat

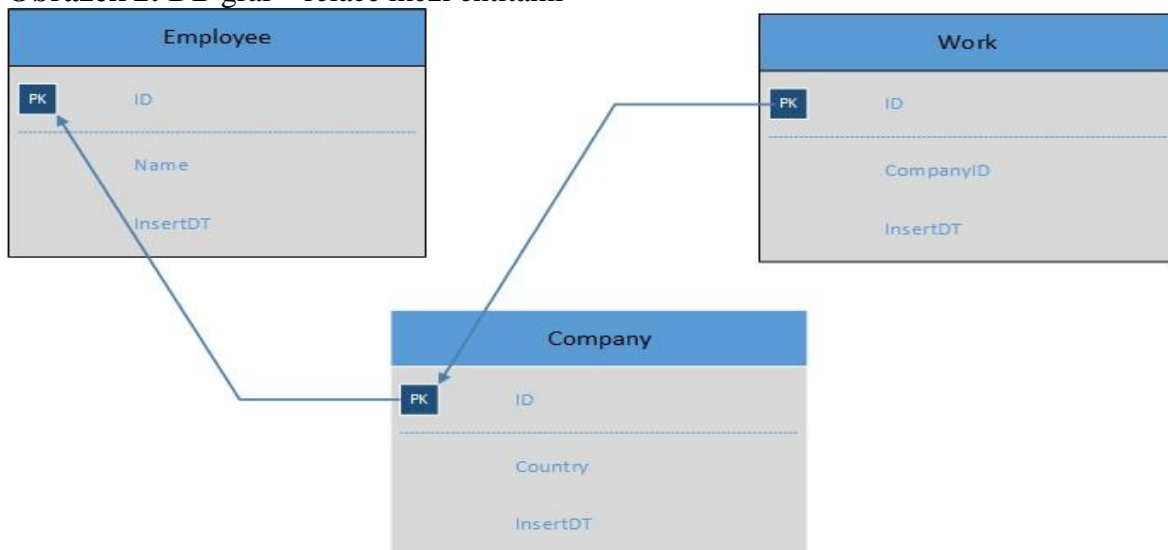
Veškerá data jsou uložena v databázi na SQL Serveru. Tento server slouží jako úložiště veškerých výstupů. Tyto výstupy tvoří zejména:

1. Manuální evidence dat (Zaměstnanci, Stroje, Část materiálů a přípravků)
2. Události, které jsou potřebné pro správu databáze (Eventy, Logy v případě chyb)
3. Výstupy z informačních systémů (TECHMNG, VOMNG, WIP – Odvádění práce zaměstnanců s časovou stopou)

Před přiblížením těchto tří výstupů je potřebná definice a charakteristika celkové databázové a datové struktury.

Například: Pokud budeme chtít smazat zaměstnance s ID (Identifikátor) č. 1, tak ho nemůžeme smazat, pokud má stále aktivní vazby například v tabulce se stroji. Pokud smažeme záznam o zaměstnanci č. 1 v tabulce se stroji, poté smažeme také vytvořenou relaci mezi dvěma záznamy a poté můžeme zaměstnance smazat. Tuto ochranu relací nám poskytují již samotné nástroje, jakým je například MS SQL Server Management viz. obrázek č. 2.

Obrázek 2: DB graf – relace mezi entitami



Např.: Zobrazit jméno zaměstnance, kde EmployeeID = CompanyID a CompanyID = WorkID

Zdroj: Vlastní zpracování

V reálných datech je relace mezi všemi tabulkami. Jednotlivé informační systémy tyto tabulky naplňují ve spolupráci s vnitřním scriptem tabulky. Ukázka je znázorněná v tabulkách č. 1, 2, 3, 4 a 5.

Tabulka 1: IS TECHMNG – Data pro tabulku s technologickými postupy - TechOp

V-TCH-OP-Id	Code	Description	U-Duration	Op-Type	P-temID
5483	340174/20	Drill	3,3	Production	1196
5518	340179/53	Point quote	24	Point	1201
5523	340179/70	Drill	9,24	Production	1201
5544	340181/169	Adjusting operations	0,75	Setting	1203
5546	340181/179	Adjusting operations	0,75	Setting	1203
5566	340182/10	Drill	43,3333	Production	1204

Zdroj: Vlastní zpracování z SQL DB

Tabulka 2: IS VOMNG – Data pro tabulku s výrobními objednávkami – VOOp

P-temID	P-ID	Pplan-Id	R-Count	O-Code	RequestedDeliveryDT
877	1209	2	100	340006	2015-10-19 00:00:00.000
878	1231	2	18	340007	2015-10-21 00:00:00.000
883	3044	2	50	340012	2015-10-14 00:00:00.000
889	5856	2	50	340018	2015-10-13 00:00:00.000
909	6353	2	200	340038	2015-10-19 00:00:00.000
911	6613	2	100	340040	2015-10-20 00:00:00.000

Zdroj: Vlastní zpracování z SQK DB

Tvorba dat není pořizována pouze za pomoci informačních systémů, ale také manuálně. Část vyhrazených pracovníků spravuje vybrané DB a eviduje změny ručně. Tabulky, které patří mezi manuálně udržované, jsou tyto:

- Zaměstnanci
- Stroje
- Materiál (část tabulky)
- Produkty (část tabulky)

Tabulka 3: Tabulka Materiály

M-Id	Code	Description	Avai-Amount
1	111113001000	Technical petrol	419
2	111213024000	MOGUL DIESEL	1291
3	111313003000	Plastic oil	0
4	111442004000	GRAFIT	4
5	1201045001	Interposer	0

Zdroj: Vlastní zpracování z SQL DB

Tabulka 4: Tabulka Stroje

M-Id	Location	Code	M-Category-Id
6	ABC	4	6
7	ABC	11	6
8	ABC	65	6
9	ABC	00C	7
10	ABC	00D	8

Zdroj: Vlastní zpracování z SQL DB

Tabulka 5: Tabulka Stroje

E-Id	Code	Name	Profession	EmployedFrom
1	1131	Employee1	Comptroller mechanical engineering	1997-06-03
2	11815	Employee2	Brusic metal	2009-12-01
3	1228	Employee3	Operation and adjustments	2015-05-04
4	13293	Employee4	Metal turner	1983-01-01
5	13374	Employee5	Operation and adjustments	1983-12-24

Zdroj: Vlastní zpracování z SQL DB

3 Příprava dat pro process mining

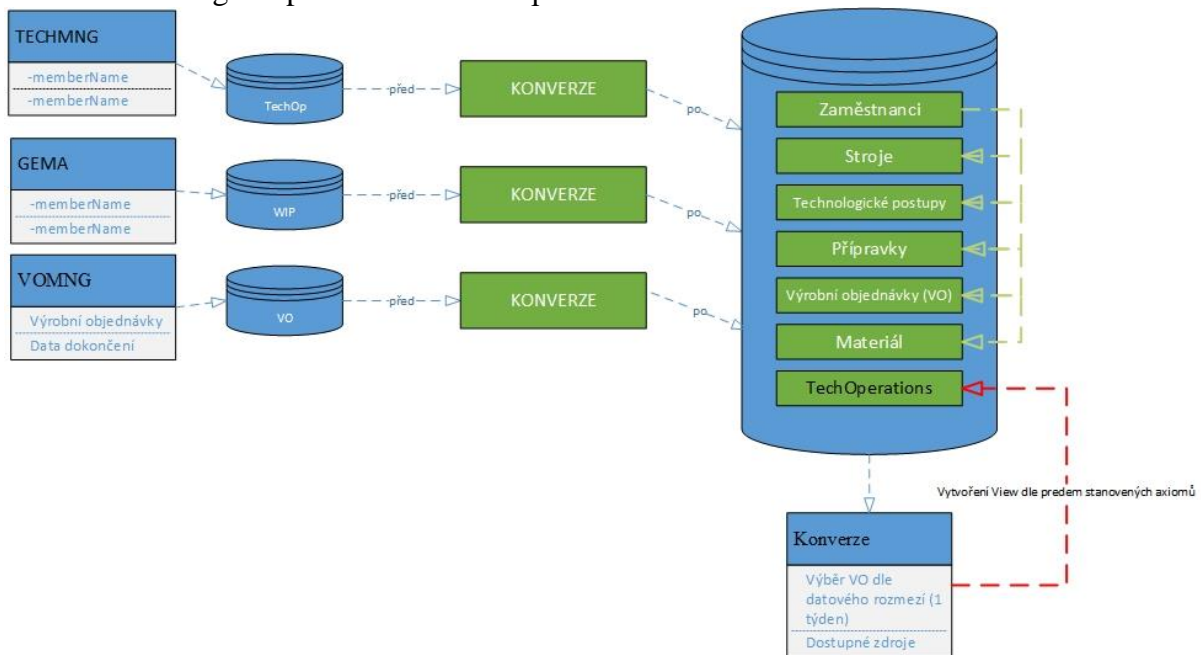
Příprava dat pro následný process mining bude následující:

1. Konverze dat
2. Příprava náhledu tzv. View tabulky, která bude obsahovat veškeré axiomy pohromadě
3. Druhá konverze – seřazení dat pro náhled
4. Redukce dat – výběr podstatných dat a jejich množství
5. Korekce zredukovaných dat
6. Finální příprava dat do potřebného formátu pro nástroj DISCO

3.1 Konverze dat

Na obrázku č. 3 lze vidět průběh konverze potřebných dat. Tato konverze je důležitá, jelikož je potřeba dostat data do jedné databáze, ve které se následovně s daty pracuje. Konverze je plně automatická. Po dokončení konverze jsou vytvořeny tabulky s požadovanou strukturou. Tyto tabulky obsahují také pole, která se dají považovat za pole povinná.

Obrázek 3: Diagram potřebné konverze pro finální stav



Zdroj: Vlastní zpracování

Při tvorbě tzv. View se vytvoří *script*, který na základě jeho definice spojí požadované tabulky, jejich sloupce a následně jejich data. Jako View (Pohled) se označuje databázový objekt, který poskytne uživateli potřebné data ve struktuře, jakou má tabulka, ale s tím rozdílem, že nezobrazuje přímo uložená data, ale zobrazí je dle předem definovaného předpisu. Díky tomuto lze s daty dále pracovat s větší pohodlností a vlastní definicí.

Mezi nejdůležitější axiomy patří:

- Seřazení předchází produkci
- Rozmístění pracovníků dle dovedností po celém objektu
- Externí operace papírově nesmí prodlužovat dobu výroby – fiktivní úzká místa
- Interní operace se bere jako samostatný job
- Koridory – nárazové vstupy VO pro zakázkovou výrobu – mají přednost před ostatními VO
- Operace vždy na sebe logicky navazují
- Pojistná doba

- Pojistný bod – čas, kdy může výrobek cestovat z bodu A do bodu B – nebere se jako prostoj
- Písemná a elektronická evidence – tvorba dat
- Plánuje se vždy jen na 5 pracovních dní
- Přeplánování výroby a plánované výroby probíhá vždy dvakrát týdně

Závěr

Tento příspěvek je součástí prvního roku projektu a uvedli jsme v něm základní informace o process miningu, který používáme v rámci našeho SGS projektu. První rok projektu je zaměřen kromě teoretických východisek v oblasti process miningu především na získání reálných dat z logovacích souborů a jejich následné převedení do prostředí nástroje DISCO pro detailní analýzu závislostí. V našem článku používáme data společnosti, která se zabývá výrobou nákladních automobilů, a zaměřujeme se na určitý výrobní proces ve firmě. Tyto procesy generují data, která se zapisují do relačních databází postupně a našim úkolem bylo tato data konvertovat. V rámci práce pro první rok nás čekají ještě body jako příprava náhledu, druhá konverze za účelem seřazení dat pro náhled, redukce dat pro výběr podstatného množství dat, korekce zredukovaných dat a v neposlední řadě finální příprava dat do potřebného formátu pro nástroj DISCO

Ve druhém roce projektu bude probíhat analýza dat z výrobních procesů s cílem nalezení závislostí, definice matematických modelů, vytvoření formální definice modelu a také návrh optimalizace výrobního procesu a re-designu tohoto procesu, což v praxi znamená, že výstup z programu DISCO znázorní datový i grafický průběh výrobního procesu. Tento výrobní proces bude možno díky výstupům z programu DISCO a našich předešlých zjištění detailněji zkoumat. Tímto plynule přejdeme z čištění dat a přípravy do process miningu. Následující postup bude jednoznačný, a to vyhledat slabá místa ve výrobním procesu (úzká hrdla), pokusit se blíže definovat matematický model a výrobní proces co nejvíce zefektivnit.

Poděkování

Tento publikační výstup byl podpořen projektem SGS/19/2016 s názvem “Pokročilé metody získávání dat a simulační techniky v oblasti podnikových procesů”.

Literatura

- [1] AALST, W. V. D., 2011. *Process mining: discovery, conformance and enhancement of business processes. 1st ed.* New York: Springer., ISBN 978-3-642-19344-6.
- [2] BASL, J. a R. BLAŽÍČEK., 2012. *Podnikové informační systémy: podnik v informační společnosti. 3., aktualiz. a dopl. vyd.* Praha: Grada., ISBN 978-80-247-4307-3.
- [3] GUNTHER, C.W. and A. ROZINAT., 2012. *Disco: Discover your processes. In: Proceeding of BPM Demos.* CEUR Workshop Proceedings, vol. 940, pp. 40-44
- [4] ROZINAT, A. and W.V.D. AALST., 2006. *W.M.O.: Decision Mining in ProM.* In: Dustdar, S., Fiadeiro, J.L., Sheth, A.P. (eds.) BPM .
- [5] VANĚK, J. a R. ŠPERKA., 2013. *Informační systémy: pro prezenční a kombinovanou formu studia.* Karviná: Slezská univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné., ISBN 978-80-7248-855-1.