

A folyamatbányászat (Process mining)

Manapság a legtöbb szervezet használ információs rendszereket az üzleti folyamataik támogatásához. Néhány példa ilyen támogató rendszerre: Workow Management Systems (WMS), Customer Relationship Management (CRM), Enterprise Resource Planning (ERP). Ezen rendszerek tartalmazhatnak nyílt modelleket a folyamatokról (például a munka-folyamat rendszerek: Staffware, COSA), vagy támogathatják a folyamatokban szereplő feladatokat anélkül, hogy definiálnának egy nyílt folyamat modellt (pl.: az ERP rendszerek, mint az SAP R/3). Van, hogy egyszerűen csak nyomon követik (ellenőrzési célra főként) a már végrehajtott feladatokat anélkül, hogy bármilyen támogatást nyújtana az aktuálisan végrehajtásra kerülő feladatoknak (például egyedi információs rendszereknél kórházakban). Mindezek alapján elmondható, hogy ezen információs rendszerek jellemzően rendelkeznek naplózási képességgel, annak érdekében, hogy rögzítsék mely események/folyamatok hajtódtak végre a szervezetben. Ezek az elkészített naplók általában tartalmaznak adatokat a szervezetben végrehajtott eseményekről, a végrehajtott feladatok idejéről, emberekről vagy rendszerekről, amik szerepeltek a feladatoknál és egyéb más típusú adatokat. Ezek a naplófájlok (logs) szolgálnak kiindulási alapnak a folyamat bányászathoz és általában, eseménynaplóknak (event logs) nevezzük őket. A folyamatbányászat célja az eseménynaplókból az információ automatikus felderítése. A feltárt információk alkalmasak lehetnek új rendszerek telepítésére, amelyek támogatják az üzleti folyamatok végrehajtását, vagy visszajelző eszközként is használhatók, ami segíthet a vizsgálatoknál és az elemzéseknel, illetve javíthatja a már elfogadott üzleti folyamatokat. Az információ objektív összegyűjtése a fő haszna a folyamat bányászatnak. Az eseménynaplóban lévő adatok típusa határozza meg a folyamatbányászat során feltárt információk perspektíváit.

Ha a logfájl tartalmazza a folyamatban végrehajtott feladatokat és következtetni lehet az egyes eseményeknél ezen feladatok végrehajtására és kapcsolataira, akkor alkalmazható az úgynevezett ellenőrzött folyam (control-flow perspective) bányászat.

Ha a naplófájl rendelkezik információval a személyekről/rendszerekről melyek a végrehajtás során szerepelnek akkor (organizational perspective) szervezeti szempontból elemezhető. A szervezet szerinti bányászat során olyan információkhoz juthatunk, mint például a folyamatokban előforduló szociális hálózat, ami a munka továbbításán, vagy a szervezeti dolgokhoz (egységekhez, szerepekhez) kapcsolódó elosztási szabályokon alapul.

Ha a napló tartalmaz további információkat a feladatokról; adatmezők értékei, amik a végrehajtást módosítják, akkor az eset (case perspective) szerinti bányászat alkalmazható, vagyis az események és adatok közti kapcsolatok vizsgálatával nyerhetőek új információk.

A menedzserek leggyakrabban előforduló kérdései:

1. Melyik a leggyakrabban használt út a folyamat modellekénél?
2. Hogyan oszlanak el az események a folyamatban a különböző utakon?
3. Mennyire felelnek meg az esetek az elkészített folyamat modellel? Hol vannak a problémák? Milyen gyakori a (nem)megfelelés?
4. Mik az összeköttetési valószínűségek az egyes elosztott feladatoknál?
5. Mekkora az események átlagos (minimum/maximum) átviteli ideje?
6. Melyik utak tartanak túl sok ideig az átlaghoz képest? Hány eset követi ezeket az összeköttetéseket? Melyek a kritikus részek ezen utakon?
7. Mennyi az átlagos feldolgozási idő az egyes feladatokra?
8. Mennyi idő telik el két bármely feladat között a folyamat modellben?
9. Az aktuális események hogyan hajtották végre?
10. Mik az üzleti szabályok a folyamat modellben?
11. Valóban érvényesülnek a szabályok a modellben?
12. Hány ember érintett egy eseménynél?
13. Mik a kommunikációs struktúrák és függések az emberek között?
14. Mennyi átvitel történik egyik szereplőtől a másikig?
15. Kik a fontos emberek a kommunikációáramlásban (leggyakoribb áramlás)?
16. Kik dolgoznak kinek (ki kinek a beosztottja)?
17. Kik dolgoznak ugyanazon a feladaton?

A fenti és azokhoz hasonló kérdések megválaszolására, illetve a különböző bányászati módszerek alkalmazására szükség van egy megfelelő eszközrendszer használatára.

A ProM keretrendszer

A ProM keretrendszer egy nyílt forráskódú eszköz, amely támogatást nyújt a különböző folyamatbányászati pluginek fejlesztéséhez, bővítéséhez.

Azokat a plugineket, amelyek alapjául kizárólag az események adatai szolgálnak *discovery* plugineknek nevezzük, mert nem használnak semmilyen már létező információt az alkalmazott modellekről.

Azon plugineket, amelyek azt vizsgálják, hogy az eseménynaplókban szereplő adatok az előírásoknak megfelelően mennyire egyeznek meg a telepített modellel, *conformance* plugineknek nevezzük.

Ha az információ kinyeréséhez szükség van a modellekre és a hozzájuk tartozó logokra a modell továbbfejlesztéséhez, növeléséhez, akkor *extension* pluginekről beszélünk.

A ProM keretrendszerben a logok szűrésére beépített Log Filtereket használhatunk. A folyamatbányászaton a filtereket annak érdekében használjuk, hogy kiválasszuk, melyik folyamatot vegyük figyelembe a bányászati algoritmus futtatása során. Ilyen filterek például a *Processes*, *Event types*, *Start events*, *End events* és *Events*.

Az *Event types* filter alkalmazása lehetővé teszi az események típusának megválasztását, amelyeket figyelembe akarunk venni a bányászaton során.

A *Start events* úgy szűri a logokat, hogy csak azok a tracek/események maradnak meg amelyek a megjelölt feladattal kezdődnek.

Az *End events* az előzőhöz hasonló módon működik, csak a szűrt események a kívánt feladattal végződnek.

Az *Events* filterrel egyszerűen megadhatjuk, hogy mely események szerepeljenek a szűrt logban a folyamatbányászaton során.

Folyamatbányászat a *Control-Flow Perspective* módszerrel

A módszer megállapítja a folyamat feladatai között lévő összefüggéseket. Segítségével választ kaphatunk arra, hogy mely feladatok előznek meg más feladatokat, vannak-e párhuzamos feladatok vagy hurkok a folyamatban. Egyszóval arra alkalmas, hogy megadja azt a folyamat modellt, ami a naplóban lévő események alapján összefoglalja a lejátszódó folyamatokat. Az információ fontossága abban rejlik, hogy visszajelzést ad arról, hogyan hajtódnak végre az aktuális események a szervezetben. A Control-flow bányászatot a ProM rendszer több különböző pluginnal is támogatja. Az egyik itt használható bányászati eszköz az Alfa algoritmus (*Alpha algorithm*), ami Petri-hálót használ a folyamat modell ábrázolására. Az ezen a módszeren alapuló bányászathoz használható számos algoritmus közül különösen hasznosnak bizonyulók még például a *Multi-phase Macro plugin*, *Heuristics miner* és a *Genetic algorithm plugin*.

Folyamatbányászat a *Case Perspective* módszerrel

Az eseményekben lévő kapcsolatok vizsgálatára alkalmas módszer a *Performance Sequence Diagram Analysis*. Diagramon ábrázolhatóak a munkafolyamat során az egyes lépések közötti kapcsolatok, a leggyakoribb útvonalak, valamint az is meghatározható, hogy az egyes feladatokhoz mely szereplők tartoztak, kik vettek részt bennük. Egy esemény során végrehajtott folyamat időbeli lefolyása is megjeleníthető. Információt kaphatunk arról, hogy a folyamat során mely munkaterületnél fordul elő idővesztés, hol van a hiba, illetve ki/kik lehetnek felelősek érte.

Folyamatbányászat az *Organizational Perspective* módszerrel

Ezen módszerrel a szervezetre, illetve annak működésére irányuló kérdésekre kaphatunk választ. Az itt alkalmazott bányászati pluginok a *Social Network Miner* és az *Organizational Miner*. Az elemzéshez pedig az *Analyze Social Network* plugin használható.

Az eseménynapló tulajdonságának ellenőrzése

Gyakori, hogy a szervezetekben a folyamatoknak bizonyos szabályoknak vagy elveknek kell megfelelnie. Az egyik ilyen előforduló példa a „négy szem elve”, ami meghatározza, hogy egy személy nem tudjon végrehajtani egyedül bizonyos feladatokat (például költségvetés készítése és annak jóváhagyása). Az ilyen jellegű szabályokat és elveket alkalmazzák például a leszállított termékek minőségbiztosítására és/vagy a csalások elkerülésére. Az egyik módja a szabályok betartásának ellenőrzésére a logok megvizsgálása (auditálása), amik megmutatják mi történt a szervezetben. A ProM rendszerben az eseménynaplók vizsgálatára/auditálására használhatjuk a *Semantic LTL Checker* analizáló plugint.

Vannak olyan kérdések is, melyek vizsgálatához szükség van az eseménynapló és a folyamat modell egyidejű elemzésére, kezelésére. A ProM rendszerben az ilyen feladatokat segítő pluginek felfedik az eltéréseket a modellben leírt folyamat és az eseménynaplóban leírt tényleges folyamat között. Manapság sok cég használ folyamat felügyelő programokat/rendszereket (például: PAIS), azonban ezek nem tökéletesek, nem teljesekek. Alkalmasak a folyamatok felügyeletére, azonban nem alkalmazkodnak megfelelően a változásokhoz. Minden esetre egy ilyen felügyelő eszköz mindenképpen hasznos visszajelzést tud adni a rendszerről.

Azt az elemzést, ami megvizsgálja, hogy a logban lévő folyamatok mennyire egyeznek meg a modellel, valamint kiemeli a különbségeket, a ProMban a Megfelelőség Ellenőrző eszköz (*Conformance Checker*) végzi.

Egy másik, a *Conformance Checker*-hez hasonlóan inputként egy naplót és egy Petri-hálót használ a Petri hálón alapuló teljesítményelemző (*Performance Analysis with Petri net*) eszköz.

Ahhoz, hogy fényt derítsünk azokra az üzleti szabályokra, feltételekre, amelyek befolyásolják a döntési pontokat a modellben, egy újabb eszköz lehet segítségünkre; a *Decision Point Analysis* plug-in.

A ProM keretrendszer által biztosított funkciók

A ProM képes kommunikálni, kapcsolatot teremteni a különböző már meglévő rendszerekkel, mint például munkafolyamat irányító rendszerekkel, mint a Staffware, Oracle BPEL, Eastman Workow, Websphere, InConcert, FLOWer, Caramba, és YAWL. Különböző szimulációs eszközökkel is, mint az ARIS, EPC Tools, Jasper, és a CPN Tools, vagy ERP rendszerekhez, mint a PeopleSoft és SAP. Olyan elemző eszközökkel való kapcsolatra is használható, mint az AGNA, NetMiner, Viscosity, AlphaMiner, vagy éppen az ARIS PPM. Több mint 20 különböző rendszerrel képes a ProM a folyamat modellek és eseménynaplók cseréjére. Lehetőség nyílik a különböző modellek közvetlen betöltésére/exportálására/importálására.

A ProM nyílt forráskódú, és ez lehetővé teszi a forráskód kiegészítését, megváltoztatását, azonban lehetőséget biztosít saját (általunk létrehozott) pluginek hozzáadására a rendszerhez. Ezzel a módszerrel könnyedén új funkciókkal bővíthető a ProM, anélkül hogy különösebben ismerni kellene a teljes keretrendszert, illetve módosítani kellene a forráskódot. Jelenleg több mint 70 beépített plugin áll rendelkezésünkre és a rendszer 5 különböző típusú plugint támogat:

- **Bányászati:** jellemzően egy logból állítanak össze egy modellt. (*Alpha algorithm, Multi-phase Macro plugin, Heuristics miner, Genetic algorithm plugin, Social Network Miner, Fuzzy Miner, Organizational Miner*)
- **Import:** általában betöltnek fájlból egy modellt, és esetenként naplófájl használják a modellben lévő fontos objektumok azonosítására.
- **Export:** jellemzően a modellt fájlba exportálják.
- **Konvertáló:** egyik modellből egy másikba konvertálnak. (*Petri hálóból EPC-be, Fuzzy modellbe, Heurisztikus vagy WF hálóba, vagy YAWL modellbe*)
- **Elemző pluginek:** feladatuk a modellek elemzése esetenként az eseménynaplókat is felhasználva. (*Performance Sequence Diagram Analysis, Analyze Social Network, (Semantic) LTL Checker, Conformance Checker, Performance Analysis with Petri net, Decision Point Analysis, Originator by Task Matrix, Log Summary*)