

Mineração de processos em plantas industriais - Caso de estudo

Adriana Guedes^[74545], Marco Barbosa^[75278]
e Ricardo Certo^[75315]

Universidade do Minho, Braga, Portugal

Resumo. Ao longo deste relatório serão apresentadas e descritas todas as etapas do trabalho realizado no âmbito da Unidade Curricular de Laboratórios em Engenharia Informática, segundo a orientação do professor Orlando Belo. O projeto consiste essencialmente na aplicação de *Process Mining* em plantas industriais, mais concretamente numa olaria. Antes de se partir para a realização concreta do projeto, houve um estudo prévio de forma a ser perceptível, por um lado, as bases e característica fundamentais no *Process Mining*, mas também acerca do funcionamento concreto de uma olaria, de forma a que as conclusões e resultados obtidos sejam o mais fiáveis, concisos, aceitáveis e aproximados da realidade possível. Ao longo deste relatório serão também aprofundados todos os passos realizados e também decisões tomadas ao longo do projeto. Num primeiro momento, será feita uma breve introdução juntamente com a motivação e objetivos do projeto, seguida de uma introdução ao *Process Mining*, analisando por isso as suas aplicações e quais os sistemas onde este poderá ser aplicado (com base em trabalhos relacionados.). Na apresentação do caso em estudo é explicada a metodologia de trabalho adotada, caracterizando e detalhando o caso em particular. Depois de introduzido e explicado o modo de funcionamento do caso de estudo, será apresentado o laboratório de trabalho representativo desse mesmo caso, laboratório esse utilizado para gerar e tratar os dados necessários à aplicação do *Process Mining*. Numa fase posterior, será abordado todo o processo de mineração realizado e também feita uma análise de resultados obtidos, a partir dos quais será decidida qual é a melhor solução a aplicar. Seguidamente, serão mencionadas todas as ferramentas utilizadas durante todo este projeto, e por último serão apresentadas as conclusões retiradas, apontando os seus pontos fortes e fracos, através de uma avaliação crítica ao trabalho realizado e possíveis evoluções a serem feitas no trabalho presente.

Keywords: Process Mining, Laboratório de trabalho, Mineração.

1. Introdução

1.1 Contextualização

Na indústria da atualidade, com o avanço e desenvolvimento tecnológico, associado ao grande crescimento das empresas, os processos de fabrico estão cada vez

mais complexos e de maiores dimensões (envolvendo um elevado número de trabalhadores e máquinas), daí que se torne cada vez mais difícil, do ponto de vista humano, o controlo e monitorização de tais processos.

Por outro lado, a indústria procura cada vez mais uma produção de qualidade, de baixo custo e que aproveite da melhor forma possível os recursos disponíveis, tornando-se assim cada vez mais importante não só avaliar o normal funcionamento do sistema, mas também o seu desempenho com o objetivo de o melhorar.

Perante tal necessidade e através deste projeto desenvolvido, foi vista uma janela de oportunidade para uma técnica designada de *Process Mining*, com o intuito de, essencialmente, apresentar propostas de solução para as necessidades que se sentem nos processos industriais atuais.

Recorrendo a um caso particular de uma planta produtiva, pretende-se por isso ilustrar a utilidade que a técnica de *Process Mining* trará para que se evidenciem possíveis problemas e aumente a eficiência de plantas produtivas.

1.2 Motivação e Objetivos

Na realização deste projeto, o ponto fulcral será a aplicação de processos de *mining* num laboratório representativo de uma planta industrial, no caso concreto da linha de produção de uma olaria.

A linha de produção idealizada implicou um estudo prévio sobre todo o seu funcionamento e elementos que a compõem, de forma a que esta seja representada da maneira mais concreta e próxima da realidade possível, tendo em conta que quase todas as plantas produtivas são sinónimo de perdas de tempo/eficiência devido a filas de espera.

O grande objetivo e meta a atingir com este projeto desenvolvido será a simplificação de todo o processo de produção idealizado numa olaria, tendo por base dados recolhidos aos quais foi aplicado o processo de mineração. As simplificações a serem feitas baseiam-se essencialmente na escolha da melhor máquina de uma etapa para um dado produto a ser produzido, ou então no peso custo/benefício que a adição de mais máquinas trará a uma dada etapa.

Após terem sido definidos os pontos centrais e objetivos do ambiente em causa, partiu-se para a montagem do laboratório de trabalho referido, mantendo sempre o foco em pontos de estrangulamento e problemas que poderão surgir.

Com este projeto desenvolvido, pretende-se assim caracterizar não só o funcionamento de uma planta produtiva em particular e todo o seu processo produtivo, mas acima de tudo retratar a utilidade da aplicação de *Process Mining* neste tipo de ambientes, com vista a sua otimização, melhoria e evolução.

2. Trabalho Relacionado

2.1 Introdução ao *Process Mining* e as suas aplicações

O *process mining*, essencialmente, corresponde a uma técnica de análise de processos com base em *logs* de eventos. Uma log de eventos é o ponto de partida do *Process Mining*, onde cada evento nessa log se refere a uma atividade (um passo bem definido em algum processo) e é relacionado com um caso em particular (isto é, uma instância do processo). Esta acaba assim por consistir num multiconjunto de transações, onde cada transação descreve o “ciclo de vida” de um caso particular em termos de atividades executadas. Durante este processo, algoritmos específicos de mineração de dados são aplicados a uma log, de forma a serem identificadas tendências, padrões e detalhes contidos nesses registos.

Após algum estudo sobre esta técnica de mineração, podemos salientar os seguintes objetivos:

- encontrar pontos de estrangulamento, que não são mais que pontos críticos no sistema, pontos esses que poderão ter grande afluência e comprometer toda a eficiência do sistema em que se inserem.
- compreender os fatores que contribuem para o aparecimento dos pontos de estrangulamento mencionados.
- detetar desvios e anomalias no sistema em análise, permitindo medir a sua severidade e avaliar o nível geral de aceitação de tais anomalias.
- prever custos, atrasos e riscos futuros.
- recomendar e evidenciar medidas a serem tomadas face aos resultados obtidos após a aplicação de algoritmos específicos.

Ao fazer uso do *Process Mining*, normalmente percorrem-se cinco etapas interdependentes e fundamentais, que permitem descobrir, monitorizar e melhorar processos de vários domínios aplicacionais, através da extração de conhecimento a partir de logs de eventos.

Sendo assim, podemos ver a primeira etapa como a identificação e compreensão global do sistema onde a mineração será feita, determinando características e detalhes relevantes nesse mesmo sistema. Depois disso, a segunda etapa consiste na obtenção da referida log de eventos com base no sistema em causa e a terceira etapa no respetivo tratamento dessa log obtida, reduzindo a existência de

ruído nos dados (valores absurdos) e ajustando certos parâmetros de forma a tornar este registro o mais perceptível, completo e adequado.

Na quarta etapa, através da utilização de uma ferramenta de *Process Mining*, algoritmos de mineração são aplicados sobre a log de eventos e obtém-se um modelo do processo em causa, onde depois na etapa seguinte esse modelo é analisado, avaliando o quão bem retrata a realidade que se pretende representar e refinado através de informação contida na própria log.

Acrescenta-se também que após a fase de análise do modelo de processo obtido após a mineração, o sistema poderá ser modificado, uma nova log gerada com base nesse sistema reformulado, e um novo modelo criado que, à partida, será mais eficiente e menos crítico que o modelo inicial.

Estas reformulações não têm um número definido, pelo que normalmente acontecem até que se encontre um ponto de concordância entre as pessoas às quais a aplicação do *Process Mining* tem interesse.

Na seguinte imagem podemos observar todas etapas que caracterizam a aplicação do *Process Mining*:

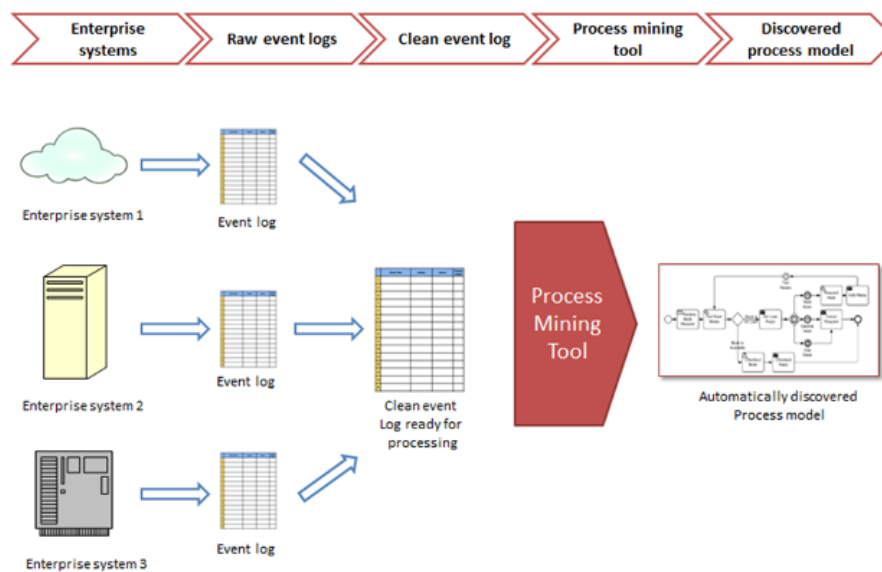


Figura 1- Etapas de aplicação do Process Mining

3. Caso Estudo

3.1 Metodologia de trabalho adotada

Para o projeto em questão, optou-se por seguir uma metodologia que permitisse a criação de um "kit" laboratorial para, não só permitir a representação do layout produtivo idealizado, mas também para gerar, limpar e preparar os dados para o processo de mineração, sendo que posteriormente seja possível efetuar uma análise da planta produtiva esquematizada, identificando as principais situações críticas e condicionantes para o sistema produtivo, e de que forma é que podem ser resolvidas.

A implementação deste "kit" laboratorial foi feita recorrendo ao uso do programa *Arena*, onde através desta ferramenta de simulação foi possível simular a planta produtiva de uma olaria e obter um registo de eventos de todos os processos aí efetuados.

Após o laboratório de trabalho montado, decidiu-se submeter a linha produtiva idealizada ao fabrico de uma panóplia de 16 produtos diferentes, através de uma ordem de fabrico de 500 unidades para cada um desses produtos. Esta decisão baseou-se no facto de, por um lado, um sistema produtivo atual lidar com o fabrico de muitos produtos e por vezes em grandes quantidades, e também pelo facto de que a simulação da planta produtiva ter que corresponder a um fabrico suficientemente grande de produtos, para assim potenciar a aplicação de *Process Mining* na deteção de pontos de fabrico em plantas produção discreta, evidenciando potenciais problemas e zonas críticas/de estrangulamento.

3.2 Apresentação e caracterização geral do estudo

O tema principal do projeto desenvolvido corresponde à mineração de processos em plantas industriais, mais concretamente no caso de uma olaria.

Ainda antes de se partir para a montagem do laboratório representativo desta planta produtiva, houve necessidade de alguma contextualização e perceção da área de domínio em causa, analisando os diversos passos produtivos e modos de funcionamento, principais máquinas que figuram na linha de produção e também os diversos produtos que podem ser feitos. Desta forma, podemos assim definir uma olaria como sendo um local destinado à produção de objetos que utilizam o barro ou argila como matéria-prima, sendo que existem, normalmente, 6 etapas diferentes por onde um produto pode passar.

Acrescenta-se também que um produto não passa obrigatoriamente em todas as etapas, essas etapas são sequenciais, podem ter um número variável de máquinas e a sequência delas encontra-se definida para cada um dos produtos.

A seleção dos produtos e das quantidades de produção de cada um são definidas por uma ordem fabríco, onde numa secção posterior deste relatório haverá um detalhe sobre a ordem de fabríco referida.

De forma a ser verdadeiramente possível aplicar o *Process Mining* ao caso de estudo mencionado, houve necessidade de simular um "laboratório de trabalho" representativo desse mesmo caso de estudo, com base nos factos retirados do estudo feito acerca da caracterização do *layout* produtivo de uma olaria.

3.3 Caracterização do laboratório para mineração de processos

Com base em pesquisas realizadas sobre plantas produtivas de olarias e suas características, decidiu-se que o laboratório de trabalho a implementar consistiria na simulação de uma olaria responsável por fabricar 16 produtos diferentes, composta por 6 etapas de fabríco diferentes, sendo cada uma delas composta por 3 máquinas, exceto uma que possui apenas 2 máquinas.

A criação deste laboratório teve em conta que, em cada etapa, o critério de escolha de máquina fosse a que tivesse menor fila de espera, semelhante ao que acontece nos casos reais, e que os tempos de trabalho dos produtos nas diversas máquinas fosse também o mais próximo possível da realidade, tornando por isso o laboratório o mais aproximado e concordante com uma planta produtiva real.

Este laboratório permite assim não só simular o funcionamento da linha produtiva referida, mas também registar também todos os eventos que foram acontecendo em todas as máquinas, com base numa ordem de fabríco que indicava a quantidade e os produtos que iam ser fabricados e suas respetivas sequências.

Desta forma, surge por isso o grande objetivo da criação de tal laboratório, que não é mais do que gerar os dados necessários para a posterior aplicação de algoritmos de mineração, possibilitando assim uma visão completa acerca do que realmente está a acontecer na linha de produção, identificando possíveis pontos críticos, defeitos e melhorias a implementar.

3.3.1 Apresentação das Etapas do processo de fabríco

As 6 etapas mencionadas anteriormente, que são parte constituinte da linha de produção da olaria simulada são as seguintes:

- Preparação (1): a matéria-prima é refinada, processada e limpa de impurezas, ficando assim pronta para ser trabalhada.

- Moldagem (2): a matéria-prima processada é moldada e definida segundo a forma do produto a ser feito.
- Corte (3): por vezes, certos produtos implicam cortes mais precisos e concretos que não podem ser feitos através da moldagem, e esta etapa serve para esse fim.
- Cozedura (4): esta etapa permite consolidar a forma final do produto a ser feito, endurecendo-o e retirando-lhe a capacidade de voltar a ser facilmente moldado ou destruído.
- Pintura (5): por vezes, certos produtos apresentam-se coloridos e com desenhos, e não na cor da matéria-prima que os compõem, sendo que esta etapa tem esse mesmo propósito.
- Envernizamento (6): de forma a dar um acabamento melhor e mais duradouro, por vezes certos produtos sofrem um envernizamento que lhes dá esse mesmo efeito.

A etapa da cozedura é composta por 2 máquinas, sendo todas as outras constituídas por 3, perfazendo assim um total de 17 máquinas inicialmente. Todas essas máquinas seguem uma distribuição normal de valores que caracterizam o tempo (valor mínimo, média, valor máximo), sendo estes valores atribuídos com base no tempo previsto que cada etapa iria demorar.

Mais uma vez, acrescenta-se também que todas as etapas são sequenciais, mas não obrigatórias, ou seja, um produto tem uma sequência de etapas pelas quais tem de passar (podendo não ser todas as etapas da linha produtiva), etapas essas que são abordadas sequencialmente.

Na imagem seguinte pode-se visualizar uma das etapas criada na ferramenta de simulação utilizada:

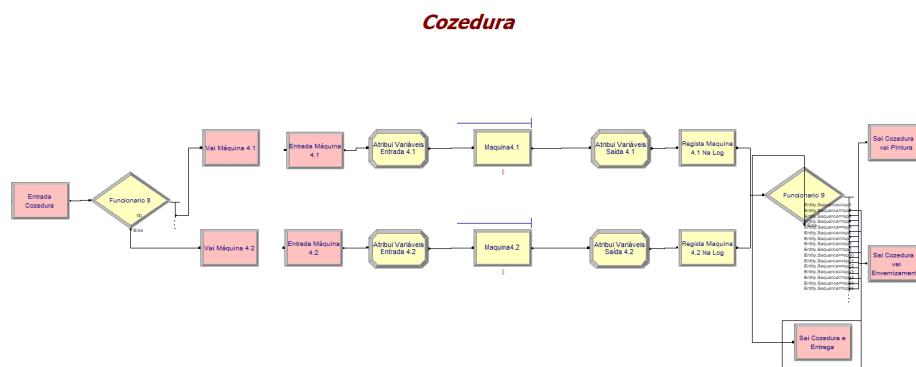


Figura 2 - Exemplo de uma etapa do laboratório desenvolvido

3.3.2 Apresentação dos produtos e das suas sequências

De seguida vamos apresentar numa tabela a lista dos 16 produtos possíveis de serem feitos na linha de produção simulada (podendo este número ser facilmente aumentado) e suas respectivas sequências, onde, a título de exemplo, podemos ver que um Vaso Simples terá que passar pelas etapas de Preparação (1), Moldagem (2) e Cozedura (4).

Nome Produto	Sequência
<i>Vaso Simples</i>	<i>1,2,4</i>
<i>Azulejo Decorativo</i>	<i>1,3,4,5,6</i>
<i>Jarra</i>	<i>1,2,4,6</i>
<i>Floreira Pintada</i>	<i>1,2,4,5</i>
<i>Tijolo</i>	<i>2,4</i>
<i>Taça ondulada</i>	<i>1,2,3,4,6</i>
<i>Prato Simples</i>	<i>1,2,3,4,6</i>
<i>Prato Pintado</i>	<i>1,2,3,4,5,6</i>
<i>Tijolo Cortado</i>	<i>2,3,4</i>
<i>Cinzeiro</i>	<i>1,2,3,4,6</i>
<i>Copo Festivo</i>	<i>1,2,3,4,5</i>
<i>Jarra Pintada</i>	<i>1,2,4,5,6</i>
<i>Cão de Barro</i>	<i>1,2,3,4,5,6</i>
<i>Garrafeira de Barro</i>	<i>1,2,3,4</i>
<i>Pipo de Barro Pintado</i>	<i>1,2,3,4,5,6</i>
<i>Assador de Chouriço em Barro</i>	<i>1,2,3,4,6</i>

Podemos também constatar que, por exemplo, a Taça Ondulada e o Prato *Simples* possuem a mesma sequência, mas possuem tempos diferentes de produção devido à Taça Ondulada ser mais complexa ao nível da sua conceção.

Estes casos de produtos com sequências iguais e tempos de execução diferentes permite assim efetuar a distinção entre os produtos, da mesma maneira que acontece numa linha produtiva real.

3.4 Geração e tratamento de dados na log de processos

A *log* de eventos obtida através do laboratório de trabalho criado foi obtida com base numa linha de produção de uma olaria, usando para isso a ferramenta de simulação Arena.

Para essa *log*, foi definida qual seria a sua estrutura, sendo que esta possui assim informações sobre o tipo de produto a ser fabricado, máquina pela qual o produto passou, sequência definida para o produto, data/hora de entrada e também de saída na máquina, salientando que todos estes atributos, no momento em que são registados no laboratório, vêm em formato numérico.

Note-se que, face ao objetivo de aumentar a eficiência da planta produtiva, houve necessidade de registar na *log* os tempos de permanência dos produtos nas máquinas (através dos atributos data/hora), de maneira a que o processo de mineração consiga salientar eventuais pontos de estrangulamento e críticos, ou seja, máquinas na linha produtiva que estão a comprometer o desempenho de toda ela.

Para uma melhor utilização da ferramenta de *Process Mining* escolhida, o *Disco*, uma melhor interpretação dos dados *input* por parte dela e para uma análise mais fácil/compreensível dos resultados obtidos, alguns valores numéricos dos atributos da *log* de eventos foram convertidos em valores nominais, através de um script próprio. Este script criado segundo na linguagem de programação *python*, recebe os dados não formatados que foram recolhidos no laboratório de trabalho e, por exemplo, se o tipo de produto for 1, este valor é transformado em Vaso Simples, sendo também concatenadas as datas e horas de início e de fim a modo de ficarem apenas dois *timestamps* (um registo de entrada do produto na máquina e um outro registo de saída).

O script é também responsável por alterar alguns tempos de produtos com a mesma sequência por serem mais demorados (mais complexos) a produzir, uma vez que através da ferramenta de simulação usada não era possível distinguir como diferentes produtos com a mesma sequência produtiva.

3.5 Processo de Mineração

Após a log de eventos gerada mediante a simulação de 500 unidades de 16 produtos diferentes, surge o momento de realmente iniciar o processo de mineração através da ferramenta *Disco*.

Num primeiro momento, a log gerada e modificada pela script foi importada para a ferramenta referida, e aí foi definido o que cada coluna da log iria significar.

A primeira coluna da log, correspondente ao número de início de fabrico do produto (por exemplo, o valor de 1 significa que aquele produto foi o primeiro a ser produzido e assim sucessivamente) foi definida como o *CaseID* (identificador do caso), o atributo que define a máquina por onde o produto passou no momento de registo foi definido como *Activity* (o que faz com que, no modelo do processo, as máquinas sejam os nodos), os atributos das datas de entrada e saída do produto forma marcados com sendo *Timestamp* e os restantes atributos foram definidos como *Other attributes*.

Após este passo, obtém-se assim o modelo de processo para os dados contidos na log, e através da sua observação será feita uma análise e retiradas as respetivas conclusões, tendo em conta o objetivo definido inicialmente que não é mais do que aumentar a eficiência da planta produtiva da olaria que foi simulada.

3.6 Análise de Resultados

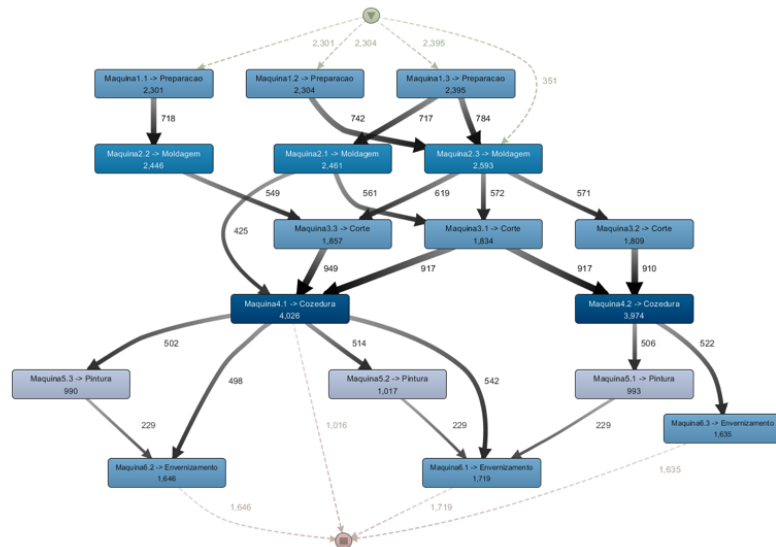


Figura 3 - Primeiro resultado obtido através do Disco

Pela análise do resultado obtido para o primeiro modelo da linha produtiva, pode-se ver que os pontos de estrangulamento da linha encontram-se na máquina 4.1

(4026 produtos) e na máquina 4.2 (3974), máquinas essas correspondentes à etapa de Cozedura e que se encontram no modelo marcadas com um tom de cor mais escuro. Este resultado seria um pouco expectável, na medida em que a existência de apenas 2 máquinas na etapa mais demorada de toda a linha produtiva (a cozedura), iria inevitavelmente conduzir a um grande estrangulamento, afluência e atraso de produtos nesta etapa.

De maneira a eliminar/reduzir o estrangulamento nestes 2 pontos da linha, o modelo inicial da planta produtiva sofreu uma reformulação, tendo sido por isso duplicadas as máquinas presentes na cozedura. Sendo assim, o laboratório de trabalho passou a ter 4 máquinas na etapa de Cozedura, perfazendo assim um total de 19 máquinas em toda a linha de produção.

Posto isto, foi feita uma nova simulação para a mesma ordem de fabrico (16 produtos diferentes e 500 unidades de cada produto), que levou ao registo de uma nova log de eventos.

Essa log foi importada da mesma forma para a ferramenta de mineração, ferramenta essa gerou um novo modelo do processo como se pode ver na seguinte imagem:

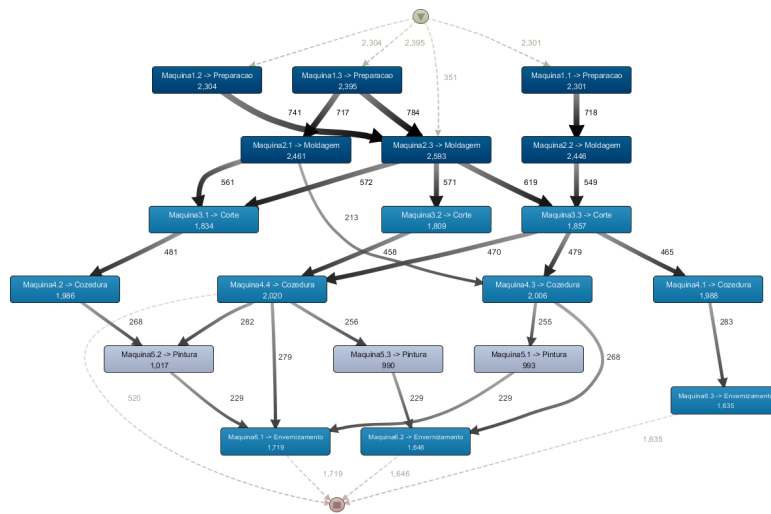


Figura 4 - Segundo resultado obtido no Disco

Como se pode ver, nos 2 pontos críticos identificados anteriormente, o número de produtos a passarem em cada máquina diminuiu consideravelmente (para cerca de metade), ficando a afluência de produtos nelas semelhante a outras máquinas de toda a planta.

Pode-se ver também que, segundo este novo modelo do processo, as máquinas da etapa de Preparação são passíveis de serem os novos pontos críticos. Perante tal evidência, percebe-se o porquê de isso acontecer, uma vez que ao aumentarmos o número de máquinas numa fase posterior da linha, estamos a aumentar o *throughput* (carga de trabalho por unidade de tempo) da linha, e por isso as máquinas do início são aquelas sofrem maior carga. Posto isto, tomou-se a decisão de não partir para uma nova reformulação da planta produtiva, não só porque as máquinas que supostamente são críticas têm uma afluência de produtos semelhante à maioria das outras máquinas (cerca de 2300 produtos), mas também porque uma nova reformulação, desta vez no início da linha produtiva, iria aumentar a carga nas máquinas a jusante e levar ao mesmo problema do primeiro modelo criado. Sendo assim, uma nova reformulação iria levar a um *loop* infinito de adição sucessiva de máquinas, uma vez que o aumento do número de máquinas a jusante na linha iria criar constrangimentos nas máquinas a jusante, e vice-versa.

Por outro lado, olhando segundo uma perspetiva real do problema, este método seria demasiado extensivo/custoso para a empresa da olaria, uma vez que esta não possuiria capital infinito para estar sempre a acrescentarem máquinas e porque também se atinge um certo ponto em que as reformulações da planta pouco contribuem para o aumento da sua eficiência.

Posto isto, foi definido o cenário ideal como sendo apenas a duplicação das máquinas da cozedura, uma vez que tal esquema produtivo gera um resultado aceitável e consistente entre todas as máquinas, apresentando também um equilíbrio entre custo de aquisições de novas máquinas e performance da planta produtiva.

3.7 Implementações Práticas

Mediante a análise dos resultados obtidos fruto da aplicação de *Process Mining* no caso de estudo em particular, chegou-se à conclusão que a duplicação das máquinas de cozedura (quando inicialmente eram duas) seria a melhor opção em termos práticos.

Essa opção leva a um melhor desempenho na produção, pois assim consegue-se a redução do fluxo de produtos nas máquinas para metade, ficando assim em conformidade com o fluxo de todas as outras máquinas das restantes etapas.

Na prática, esta solução encontrada abarca um custo extra, que corresponde ao custo de aquisição de 2 máquinas de cozedura, investimento este que a longo prazo seria revertido, uma vez que a eficiência da planta produtiva aumenta e leva a uma maior capacidade de produção diária.

O facto de não se ter continuado a minerar o processo da planta produtiva idealizada, apesar de isso garantir em teoria uma “solução ideal”, no contexto empresarial em que o problema se insere essa solução torna-se impraticável pelos elevados custos de aquisição e manutenção das inúmeras novas máquinas adquiridas.

Pode-se por isso, afirmar que não há grande interesse na obtenção de um desempenho máximo quando isso se torna insustentável para a empresa, uma vez que o orçamento disponível, o número médio de encomendas e outros fatores são variáveis muito importantes na decisão final sobre a melhor opção a tomar face aos resultados obtidos.

3.8 Ferramentas Utilizadas

Para a elaboração do projeto e todos os passos necessários até se atingir uma solução aceitável face aos resultados obtidos, foram utilizadas 3 ferramentas diferentes.

A primeira ferramenta foi o *Arena*, que possibilitou a implementação e posterior simulação do laboratório de dados. De seguida, recorreu-se a uma outra ferramenta designada de *PyCharm*, que é um IDE de *python*, e possibilitou a implementação do script responsável pelo tratamento dos dados da log. Por último, utilizou-se o *Disco* para realizar o processo de mineração e retirar as respetivas conclusões com os resultados obtidos.

4. Mineração Incremental

A mineração incremental surgiu da necessidade de assegurar uma maior eficiência aos processos de mineração de dados, isto é, evitar que as técnicas de mineração de dados sejam novamente executadas sempre que ocorram modificações nos dados (quando novos dados -incrementais- são adicionados ou dados antigos são removidos).

Por outro lado, no panorama da indústria atual, os eventos recolhidos por vezes apresentam uma grande dimensão e surge por isso o problema de que, aquando a aplicação de *Process Mining* sobre esses dados, os algoritmos de mineração utilizados por vezes não conseguem lidar com o seu tamanho e como tal não possibilitam a verdadeira mineração do processo industrial.

Face a estes problemas, o processo de mineração incremental pode assim surgir como uma solução para eles.

Esta solução consiste na adição de registos de eventos aos anteriores, de modo a obterem-se novos modelos, sendo que todos os dados já processados uma vez não necessitam de reproprocessamento, mas continuam a ser parte integrante dos inputs para a geração de um novo modelo (que essencialmente junta o modelo anterior ao novo registo de eventos). Estes novos registos de eventos seriam obtidos através de, no caso deste projeto, uma nova simulação, ou então a continuação de uma simulação (quando estas são muito extensas), mas nunca sendo reproprocessado o mesmo registo de eventos. Desta forma, a recolha de dados é feita de uma forma contínua e evita assim trabalho redundante, aumentando por isso eficiência do processo de mineração.

Caso esta abordagem fosse adotada, poderiam ser tiradas conclusões mais acertadas pois assim assegurava-se que o resultado final não iria ser influenciado por uma anomalia/ruído presente nos registos de eventos recolhidos.

Por outro lado, consegue-se também obter um espectro maior de dados e, consequentemente, uma linha temporal de análise maior permitindo que se tirem conclusões mais acertadas a longo prazo. Dado que esse espectro de dados vai sendo analisado de forma parcial e incremental, o tempo de processamento é reduzido e a consequente eficiência do processo de mineração aumenta.

Sendo assim, aplicar esta abordagem ao projeto realizado até então seria o próximo passo a fazer, podendo posteriormente ser feita a comparação dos resultados obtidos entre a aplicação de *Process Mining* numa planta produtiva industrial e a aplicação de *Process Mining* segundo uma abordagem incremental nessa mesma planta.

A ideia essencial nesta abordagem e referida anteriormente pode ser traduzida na seguinte imagem:

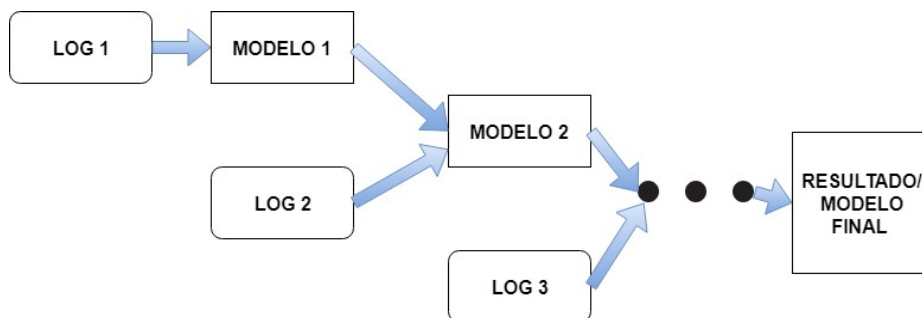


Figura 5 - Mineração incremental em process mining

5. Conclusões

Tendo em vista o grande objetivo definido para o projeto que é a melhoria, identificação de problemas e aumento da eficiência de uma planta produtiva recorrendo a técnicas de mineração de processos, em jeito de conclusão podemos afirmar que este foi cumprido. Através de um caso particular de uma linha produtiva (uma olaria), pretendeu-se demonstrar a vantagem do uso de *Process Mining*, uma vez que partindo de registos provenientes dessa linha, algoritmos de *mining* foram aplicados sobre eles e permitiram uma análise e reformulação do modelo em proveito da melhoria da linha produtiva.

A ferramenta de mineração utilizada identificou pontos de estrangulamento que existiam no sistema produtivo idealizado, possibilitou que fosse feita uma reformulação a tal sistema e comprovou que tal reformulação contribuiu para uma melhoria desse mesmo sistema.

Durante a fase de desenvolvimento, pode-se dizer que surgiram também alguns percalços e dificuldades.

Inicialmente, houve necessidade de um estudo prévio, não só sobre o que realmente era o *Process Mining* e quais as suas aplicações, mas também sobre plantas produtivas reais semelhantes àquela que se pretendeu simular, de forma a aproximar o mais possível a simulação da realidade e que as conclusões tiradas fossem o mais plausível para um ambiente real.

Após isto, no momento da montagem do laboratório de dados surgiu também uma dificuldade relacionada com traduzir para a ferramenta de simulação aquilo que realmente se pretendia realizar e implementar. Um dos problemas também encontrado na ferramenta de simulação prendeu-se com o facto de alguns tempos que eram registados serem um pouco absurdos e descontextualizados, daí que foi criado um script que facilmente os corrigiu e permitiu também fazer a distinção de diferentes produtos apesar das suas sequências produtivas serem as mesmas.

Aquando o uso da ferramenta de mineração para realmente aplicar o *Process Mining* não surgiram grandes problemas, tendo existido apenas a dificuldade relacionada com o facto do grupo nunca ter trabalho com a ferramenta e que facilmente foi superada após algumas utilizações da mesma.

Posto isto, pode-se concluir também que a mineração incremental, apesar de ser algo que não teve grande ênfase no projeto desenvolvido, constitui uma boa abordagem a seguir quando se recorre à mineração de processos devido às vantagens e potenciais problemas que poderão existir durante a mineração. Este tipo de abordagem, iria assim garantir todas as propriedades do *Process Mining* mesmo quando o laboratório de dados produzisse uma grande quantidade de registos/eventos, levando também a um aumento da eficiência e rapidez do processo de mineração uma vez que todos os eventos recolhidos seriam processados incrementalmente, apenas

novos eventos iriam contribuir para um ajuste do modelo do processo e, portanto, a análise redundante dos mesmos eventos já recolhidos seria também eliminada. Desta forma, considera-se que os objetivos previstos foram alcançados, as grandes dificuldades superadas e as análises realizadas o mais corretas e próximas da realidade possível mediante a abordagem utilizada, demarcando por isso a potencialidade que o *Process Mining* tem para a mineração de processos em plantas industriais.

Referências

1. Wil van der Aalst.: How to get started with process mining?
2. Wil van der Aalst.: Extracting Event Data from Databases to Unleash Process Mining.
3. Wil van der Aalst.: Service Mining: Using Process Mining to Discover, Check, and Improve Service Behavior.
4. Wikipedia Process Mining, https://en.wikipedia.org/wiki/Process_mining .
5. A.C. Kalsing, C. Iochpe, L.H. Thom: An Incremental Process Mining Algorithm.
6. E.R. Mahendrawathi, Hanim Maria Astuti, Ayu Nastiti: Analysis of Customer Fulfilment with Process Mining: A Case Study in a Telecommunication Company.
7. E.R. Taghiabadi, P.N.M. Kromhout, M. Nagelkerke: Process Mining: Let data describe your process.