

## Mineração de Processos: Preparação de Dados em Máquinas CNC

Joceir Chaves (PUC-PR) [joceir.chaves@creapr.org.br](mailto:joceir.chaves@creapr.org.br)

Júlio Silva (PUC-PR) [jcsmg10@gmail.com](mailto:jcsmg10@gmail.com)

Eduardo A. P. Santos (PPGEPS, PUC-PR) [eduardo.portela@pucpr.br](mailto:eduardo.portela@pucpr.br)

Eduardo F. R. Loures (PPGEPS, PUC-PR) [eduardo.loures@pucpr.br](mailto:eduardo.loures@pucpr.br)

### Resumo:

O uso da mineração de processos ajuda as organizações principalmente na tomada de decisões, este artigo apresentará uma breve revisão dos conceitos de mineração de processos e analisará a possibilidade de aplicação da metodologia na gestão da manutenção de uma máquina CNC com base nas informações que podem ser extraídas de seu processo, e principalmente as dificuldades encontradas no processo de preparação dos dados e a importância do conhecimento do processo a ser tratado, de maneira que facilite o entendimento desta etapa inicial. A principal matéria prima para a mineração de processos são os logs de eventos, e sua correta limpeza e preparação é fundamental para proporcionar uma boa análise.

**Palavras chave:** Mineração de Processos, Manutenção, CNC.

## Process Mining: Data Preparation in CNC Machines

### Abstract

The use of process mining helps organizations mainly in decision making, this article will present a brief review of the concepts of process mining and will analyze the possibility of applying the methodology in the management of the maintenance of a CNC machine based on the information that can be extracted from its process, and especially the difficulties encountered in the data preparation process and the importance of knowing the process to be treated, in a way that facilitates the understanding of this initial stage. The main raw material for process mining are event logs, and their correct cleaning and preparation is critical to provide good analysis.

**Key-words:** Process Mining, Maintenance, CNC.

### 1. Introdução

A mineração de processos tem recebido um crescente interesse pela comunidade acadêmica com estudos em áreas que envolvam processos com fluxo de atividades, dos quais utilizava como fonte de dados para análise os logs de eventos gerados por sistemas de informações (VAN DER AALST e WEIJTERS, 2004; VAN DER AALST et al., 2007; KONDO et al., 2012; PARK e KANG, 2016). Entretanto, o sucesso da aplicação das técnicas de mineração de processo está diretamente ligado a qualidade dos dados utilizados, tanto em relação ao seu formato quanto seu conteúdo (SUIRADI et al., 2017). De fato, antes da análise resultante da mineração de processo, deve existir um trabalho preliminar na preparação e formatação dos

dados para um modelo adequado e aceitável nos softwares dedicados para mineração de processo (VAN DONGEN e VAN DER AALST, 2005), já que são gerados por diversos sistemas e em formatos variados (VAN DER AALST et al., 2007).

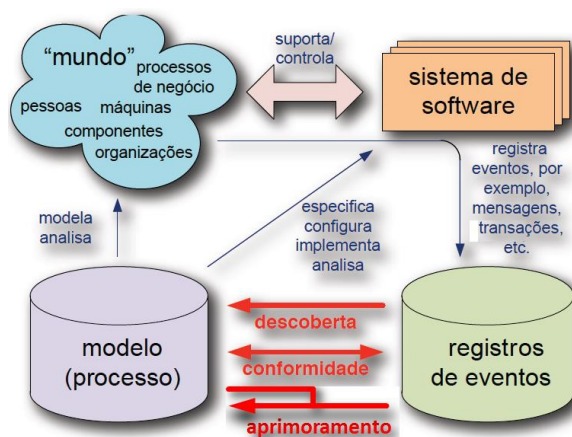
As fontes de dados utilizados são oriundos de uma variedade de sistemas como ERPs, WfMSs, CRMs (VAN DER AALST et al., 2007) sob a forma de registro do histórico de eventos, sendo assim, todo o sistema que possui a função de registro de eventos, tem a possibilidade de obter benefícios com a mineração de processos.

O objetivo desse artigo é avaliar a preparação de dados de uma máquina automática CNC (*computer numeric control*), desde sua extração entendimento dos eventos, limpeza e formatação dos dados para aplicação na ferramenta de mineração de processos ProM na versão 5.2, utilizando arquivos de entrada no formato MXLM (*Mining Extensible Markup Language*). Por se tratar de um processo automatizado com um mínimo de intervenção humana, gera uma oportunidade de análise em um ambiente pouco explorado na mineração de processos.

## 2. Mineração de Processos

A mineração de processos pode ser definida como o uso de técnicas e ferramentas para extrair conhecimento de sistemas de informações sob a forma de modelos (VAN DONGEN e VAN DER AALST, 2005), obtendo informações através dos logs de eventos gerados e armazenados. Cada evento normalmente é atribuído a uma atividade e também a uma instância de processo (*case*), e esses eventos são registrados de forma ordenada (VAN DONGEN e VAN DER AALST, 2005), onde dados diversos relativos ao processo podem estar armazenados.

Segundo van der Aalst et al. (2011), há três tipos principais de mineração de processos: a descoberta onde é obtido o modelo do processo, a conformidade no qual um modelo existente é utilizado para submeter um log de eventos e avaliar desvios, e por fim o aprimoramento, onde melhorias são feitas no processo após análise do log de eventos (VAN DER AALST, 2011). Na figura 1 é possível ver de forma resumida esses tipos de mineração de processos.



Fonte: Adaptado de van der Aalst et al. (2011)

Figura 1- Principais tipos de mineração de processo

De fato, para aplicar a mineração de processo é necessário ter os dados de entrada que é o log de eventos. Entretanto, segundo van Dogen e van der Aalst (2005) deve possuir requisitos mínimos para que possam ser aplicados em ferramentas de mineração de processos:

- Cada registro cronológico de uma atividade deve ser um evento atômico em um determinado ponto na linha do tempo;
- Cada registro cronológico de uma atividade deve se referir a uma atividade exclusivamente identificável;
- Cada registro cronológico deve se referir a um case específico (instância de processo);
- Cada instância de processo deve pertencer a exatamente um processo.

Apesar da mineração de processos ser relativamente recente, o seu crescente interesse proporcionou o surgimento de ferramentas sofisticadas para sua aplicação como algumas de uso acadêmico (EMiT, Little Thumb, InWoLvE, Process Miner, e MinSoN) e de uso comercial (ARIS PPM, HP BPI, e ILOG JViews) conforme van Dogen e van der Aalst (2005), além de outras mais recentes como o Disco (<https://fluxicon.com/disco/>) e o Celonis (<https://www.celonis.com/>). Entretanto, a falta de um padrão inicial dos dados dos logs de eventos a serem utilizados gera um trabalho complexo na preparação e análise, já que cada ferramenta trata de maneira diferente com o formato de entrada dos dados, além de que os sistemas geram os logs de eventos de maneiras diferentes.

Dessa forma, foi proposto por van Dogen e van der Aalst (2005) um padrão inicial para o formato dos dados de entrada de forma que as principais informações relevantes para a mineração de processo estejam presentes. Dessa forma, um correto mapeamento e interpretação dos log de eventos podem ser levados em consideração na preparação dos dados a serem analisados nas ferramentas de mineração de processo. O formato proposto é baseado no padrão XML (*eXtensible Markup Language*) e foi desenvolvido para ser utilizado em uma ferramenta também apresentada no trabalho de van Dogen e van der Aalst (2005), que é uma plataforma para mineração de processos denominada ProM.

A partir da modelagem dos requisitos necessários para que um log de eventos seja viável, e baseado na estrutura do padrão XML, é proposto um padrão inicial para a formatação de dados (VAN DOGEN e VAN DER AALST, 2005). O formato MXML (*Mining Extensible Markup Language*) permite que informações relevantes ao processo estejam presentes em um log de eventos, conforme fragmento apresentado na figura 2.



Fonte: Adaptado de van Dogen e van der Aalst (2005)

Figura 2 - Fragmento de um arquivo no padrão MXML

Adicionalmente no mesmo trabalho, van Dogen e van der Aalst (2005) apresentam um framework para mineração de processos (ProM), que utiliza o formato proposto no padrão MXML e permite a utilização de grande número de plug-ins para análise dos dados. Dessa forma, o objetivo do trabalho apresentado nesse artigo é parte inicial da mineração de processos, que consiste na extração de um log de eventos de uma máquina automática CNC e a preparação desses dados conforme modelo MXML proposto, finalizando com a carga dos dados na plataforma ProM na versão 5.2 e uma visão prévia da funcionalidade de alguns plug-ins como forma de validação do processo.

### 3. Metodologia

A metodologia aplicada nesse artigo é composta fundamentação teórica do tema principal que proporcionou um grande avanço para a mineração de processos, seguido pela escolha sistema originado de uma máquina automática CNC, a extração e tratamento do log de eventos, conversão do formato dos dados e avaliação dos dados em uma ferramenta de mineração de processos, seguido pelas conclusões e pontos de melhorias. Na representação esquemática da figura 3, é possível ver de forma resumida os passos aplicados para o desenvolvimento desse trabalho.

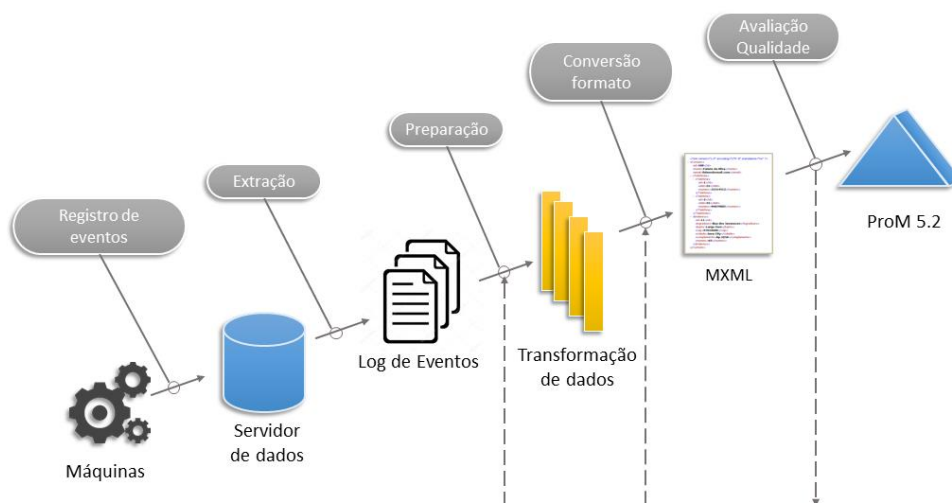


Figura 3 - Estrutura do trabalho

Nas sessões subsequentes desse artigo serão detalhadas as etapas relacionadas ao trabalho, com exceção da fundamentação teórica sobre o tema mineração de processos já tratada na sessão dois.

#### 3.1 Origem dos Dados Selecionados

Os dados utilizados nesse trabalho foram originados de uma máquina automática de uma linha de usinagem de blocos de motores composta por 33 máquinas, cujo nome da empresa será omitido nesse artigo. A máquina selecionada aleatoriamente possui seis eixos servo-acionados com posicionamento preciso comandado por um controlador CNC Siemens 840D PL. Os dados são coletados via rede local até um servidor onde um sistema de monitoramento da produção fornece recursos para identificar gargalos na linha de produção, perdas por disponibilidade, desempenho e qualidade. Além desses recursos de monitoração é possível obter o estado em tempo real da produção, com mensagens e alarmes, e, adicionalmente possibilita a exportação dos dados gravados sob a forma de log de eventos no formato Excel

(xlsx). Na figura 4 é possível observar de forma resumida, a tipologia de rede utilizada entre as máquinas e o servidor.

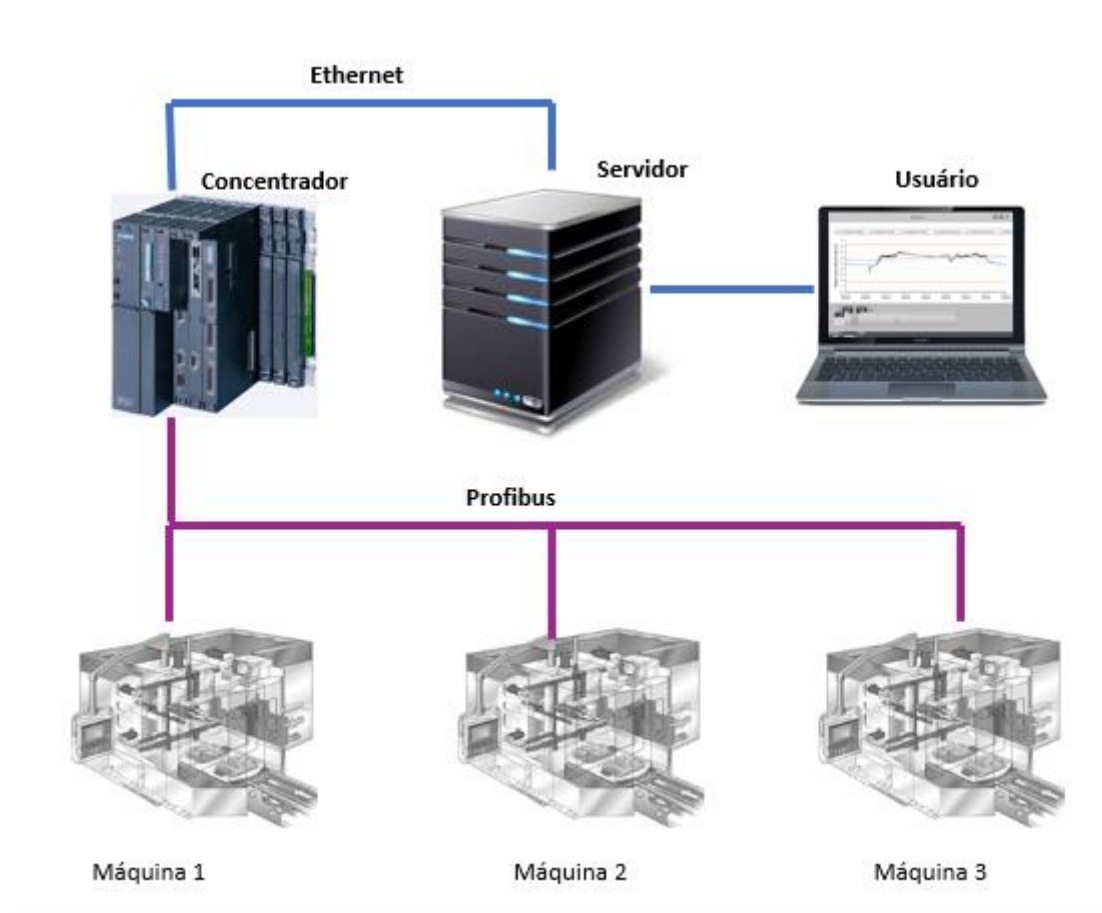


Figura 4 - Arquitetura simplificada do sistema

Avaliando a estrutura no qual o sistema opera, recebendo registros sendo a maior parte automáticos de equipamentos, e registrando em um servidor, é possível assumir que o sistema associado a esse processo possui requisitos que podem obter vantagens da mineração de processos.

### 3.2 Limpeza e Preparação dos Dados

No estudo de caso relacionado a esse artigo, o sistema responsável pela geração de dados incluindo o log de eventos por equipamento está em operação desde o ano de 2000. Logo, não seria um fato novo a necessidade de preparação dos dados obtidos no formato de planilha antes da conversão para MXLM para ser importado na ferramenta ProM 5.2 para posterior análise.

A preparação e limpeza de um log de eventos já foi uma fonte de pesquisa encontrado em outro artigo, Suiradi et al. (2017) tratam de padrões de imperfeições encontradas em log de eventos e como contornar reduzindo a perda de informações relevantes. Apesar da preocupação na preparação do log de eventos extraídos de sistemas de informações, alguns aspectos relativos à qualidade devem ser levados em consideração conforme Manifesto de Mineração de Processos (VAN DER AALST et al., 2011). O manifesto destaca como critérios que: os eventos devem ser fidedignos assegurando que realmente foram executados, devem ser completos sem eventos faltantes, devem ter uma semântica bem definida e devem ser

seguros com relação às operações de registros.

Com base nessa premissa foi feito a análise preliminar do log de eventos extraído, que compreendia em um período de trinta dias do mês de julho de 2017. O arquivo exportado diretamente no formato de planilha (xlsx) possui 7485 registros relacionados a eventos, parte do conteúdo bruto pode ser observada conforme tabela 1 onde foram suprimidas as colunas dos atributos relacionados a cada evento. As colunas principais do log de eventos possuem a seguinte semântica:

- Área de produção: identifica o tipo de produto que é produzido no equipamento;
- Área de trabalho: identifica a operação que realizada no processo de usinagem;
- Horário do servidor: momento em que foi registrado cada evento;
- Identificação do estado: código único para cada tipo de evento registrado;
- “Name”: a função ou modo de operação relacionado ao código da identificação do estado;
- Descrição: o estado atual do equipamento, também relacionado à identificação de estado.

| Área de produção   | Área de trabalho | Horário do servidor | Identificação do estado | Name                          | Descrição                                 |
|--------------------|------------------|---------------------|-------------------------|-------------------------------|-------------------------------------------|
| MD13 Cylinderblock | 030              | 2017-07-03 07:21:31 | 10000                   | Manual mode                   | No alarm                                  |
| MD13 Cylinderblock | 030              | 2017-07-03 07:21:52 | 10000                   | Manual mode                   | No alarm                                  |
| MD13 Cylinderblock | 030              | 2017-07-03 07:21:55 | 20000                   | Auto/Cycle not active         | No alarm                                  |
| MD13 Cylinderblock | 030              | 2017-07-03 07:21:58 | 30100                   | Auto/Cycle active/Cont. cycle | No waiting/No alarm/100%                  |
| MD13 Cylinderblock | 030              | 2017-07-03 07:22:03 | 31102                   | Auto/Cycle active/Cont. cycle | Wait Blocked/No part change/No alarm/100% |
| MD13 Cylinderblock | 030              | 2017-07-03 08:25:50 | 32100                   | Auto/Cycle active/Cont. cycle | Part change/No alarm/100%                 |
| MD13 Cylinderblock | 030              | 2017-07-03 08:25:57 | 1122                    | Part out                      | Part out                                  |
| MD13 Cylinderblock | 030              | 2017-07-03 08:25:57 | 1123                    | No part in                    | Part in signal off                        |
| MD13 Cylinderblock | 030              | 2017-07-03 08:26:06 | 31101                   | Auto/Cycle active/Cont. cycle | Wait Starved/No part change/No alarm/100% |
| MD13 Cylinderblock | 030              | 2017-07-03 08:26:49 | 32100                   | Auto/Cycle active/Cont. cycle | Part change/No alarm/100%                 |
| MD13 Cylinderblock | 030              | 2017-07-03 08:26:58 | 1121                    | Part in                       | Part in                                   |
| MD13 Cylinderblock | 030              | 2017-07-03 08:26:58 | 1124                    | No part out                   | Part out signal off                       |
| MD13 Cylinderblock | 030              | 2017-07-03 08:27:03 | 30100                   | Auto/Cycle active/Cont. cycle | No waiting/No alarm/100%                  |
| MD13 Cylinderblock | 030              | 2017-07-03 08:38:56 | 31102                   | Auto/Cycle active/Cont. cycle | Wait Blocked/No part change/No alarm/100% |
| MD13 Cylinderblock | 030              | 2017-07-03 08:40:47 | 32100                   | Auto/Cycle active/Cont. cycle | Part change/No alarm/100%                 |
| MD13 Cylinderblock | 030              | 2017-07-03 08:40:54 | 1122                    | Part out                      | Part out                                  |
| MD13 Cylinderblock | 030              | 2017-07-03 08:40:54 | 1123                    | No part in                    | Part in signal off                        |
| MD13 Cylinderblock | 030              | 2017-07-03 08:41:03 | 31101                   | Auto/Cycle active/Cont. cycle | Wait Starved/No part change/No alarm/100% |
| MD13 Cylinderblock | 030              | 2017-07-03 08:41:46 | 32100                   | Auto/Cycle active/Cont. cycle | Part change/No alarm/100%                 |
| MD13 Cylinderblock | 030              | 2017-07-03 08:41:55 | 1121                    | Part in                       | Part in                                   |

Tabela 1 - Amostra do log de eventos extraído

Dentre os 26 atributos relacionados a cada atividade registrada no log de eventos, 17 foram descartados por não terem nenhum conteúdo registrado ou não ter nenhuma mudança de estado significativa que pudesse ser utilizado na fase de descobrimento do processo. A descrição desses atributos não está incluída nesse trabalho devido à limitação de espaço.

Um agravante encontrado no log de eventos extraído foi a ausência de um case ID que identificasse o início e fim de cada instância de processo. Entretanto, para o equipamento em questão cada instância é feita sequencialmente não ocorrendo eventos fora da linha de tempo de uma instância completa. Foi assumido nesse caso que cada instância representa o processo de uma operação de usinagem em uma peça distinta. Outra informação relevante é a atividade de início e fim de cada instância que são bem definidas. Um processo inicia-se com a atividade de carga automática de peças descrita na coluna “Identificação do estado” com o código 1121, que significa o carregamento de peças. O fim do processo ocorre com o descarregamento automático de peças com o equipamento pronto para iniciar um novo ciclo, descrito na “Identificação do estado” com o código 32100.

De posse das informações do início e fim de cada instância foi possível separar cada case Id em seu grupo de atividades sequenciais. Com relação à descrição da tarefa de cada registro, seria possível utilizar os códigos presentes na coluna “Identificação do estado”, entretanto, para a leitura em uma ferramenta de mineração de processos ficaria pouco informativo, sendo necessário recorrer a uma tabela com a descrição de cada código. Porém, só o conteúdo da coluna “Name” também não seria possível utilizar, porque existem conteúdos iguais com descrição diferente, portanto, a solução encontrada foi unir os conteúdos das colunas “Name” e “Descrição” em uma nova coluna identificada como “task description”, levando em consideração somente a união de conteúdos diferentes.

Com base nessa análise preliminar foi feita a primeira limpeza do log de eventos descartando todas as colunas dos atributos não relevantes, corrigido e modificado os rótulos das colunas remanescentes conforme a descrição a seguir:

- Área de produção: será utilizado para a escrita do case Id relacionado a cada instância;
- Área de trabalho: seu conteúdo corresponde ao originador da atividade, que no caso desse log de eventos é o número da operação que é executado na máquina;
- Horário do servidor: será utilizado como *timestamp* de cada tarefa;
- Identificação do estado: renomeado para “TaskId” e mantido na planilha, porém não será utilizado na conversão porque traz pouca informação com relação à tarefa;
- “Name” e Descrição: serão mantidas e renomeadas para “Task” e “Description”, e posteriormente concatenadas na conversão representando a “task description”;
- Demais colunas dos atributos: os atributos não removidos serão mantidos com seu próprio rótulo e sem alteração de conteúdo.

Com as devidas alterações efetuadas no log de eventos, foi desenvolvido um código em VBA para a criação da identificação das instâncias (case id) com base nas definições descritas anteriormente nesse artigo, onde foi analisado o fluxo de atividades do equipamento em questão. Dessa forma os dados do log de eventos resultaram em um novo arquivo, conforme amostra da tabela 2, com os dados adequados para conversão no formato adequado.

| CaseId | Originator | TimeStamp        | TaskId | Task                          | Description                               |
|--------|------------|------------------|--------|-------------------------------|-------------------------------------------|
| 1      | 30         | 03/07/2017 08:26 | 1121   | Part in                       | Part in                                   |
| 1      | 30         | 03/07/2017 08:26 | 1124   | No part out                   | Part out signal off                       |
| 1      | 30         | 03/07/2017 08:27 | 30100  | Auto/Cycle active/Cont. cycle | No waiting/No alarm/100%                  |
| 1      | 30         | 03/07/2017 08:38 | 31102  | Auto/Cycle active/Cont. cycle | Wait Blocked/No part change/No alarm/100% |
| 1      | 30         | 03/07/2017 08:40 | 32100  | Auto/Cycle active/Cont. cycle | Part change init/No alarm/100%            |
| 1      | 30         | 03/07/2017 08:40 | 1122   | Part out                      | Part out                                  |
| 1      | 30         | 03/07/2017 08:40 | 1123   | No part in                    | Part in signal off                        |
| 1      | 30         | 03/07/2017 08:41 | 31101  | Auto/Cycle active/Cont. cycle | Wait Starved/No part change/No alarm/100% |
| 1      | 30         | 03/07/2017 08:41 | 32101  | Auto/Cycle active/Cont. cycle | Part change done/No alarm/100%            |
| 2      | 30         | 03/07/2017 08:41 | 1121   | Part in                       | Part in                                   |
| 2      | 30         | 03/07/2017 08:41 | 1124   | No part out                   | Part out signal off                       |
| 2      | 30         | 03/07/2017 08:42 | 30100  | Auto/Cycle active/Cont. cycle | No waiting/No alarm/100%                  |
| 2      | 30         | 03/07/2017 08:53 | 31102  | Auto/Cycle active/Cont. cycle | Wait Blocked/No part change/No alarm/100% |
| 2      | 30         | 03/07/2017 09:18 | 32100  | Auto/Cycle active/Cont. cycle | Part change init/No alarm/100%            |
| 2      | 30         | 03/07/2017 09:18 | 1122   | Part out                      | Part out                                  |
| 2      | 30         | 03/07/2017 09:18 | 1123   | No part in                    | Part in signal off                        |
| 2      | 30         | 03/07/2017 09:18 | 31101  | Auto/Cycle active/Cont. cycle | Wait Starved/No part change/No alarm/100% |
| 2      | 30         | 03/07/2017 09:19 | 32101  | Auto/Cycle active/Cont. cycle | Part change done/No alarm/100%            |
| 3      | 30         | 03/07/2017 09:19 | 1121   | Part in                       | Part in                                   |

Tabela 2 - Amostra de instâncias corrigidas

Com as devidas correções encontradas nessa fase de limpeza do log de eventos, correção dos rótulos dos dados e geração do case id de cada instância, o arquivo foi preparado para fase de conversão para o formato utilizável de ferramentas de mineração de processos. Neste artigo foi adotado o uso da ferramenta ProM 5.2 e, portanto, o formato a ser utilizado é com a extensão MXML.

### 3.3 Conversão do Log de Eventos

Um log de eventos que atenda aos requisitos de qualidade para mineração de processos, não caracteriza que esteja pronto para ser importado na ferramenta de análise. O arquivo deve estar no formato correto para importação e identificação de todos os dados relativos ao log de eventos.

Nesse artigo a ferramenta de mineração elegida é o ProM 5.2 (<http://www.promtools.org>) que aceita carga de arquivos no formato MXML, e o arquivo que foi extraído e preparado para análise é no formato de planilha (extensão xlsx), sendo necessário a conversão para o formato adequado. Poderia ser utilizada uma ferramenta como o Disco (<https://fluxicon.com/disco/>) que importa dados em csv e possibilita a exportação para MXML, o arquivo em questão seria facilmente exportado para csv. Entretanto, a ferramenta não resolveria todo o processo de limpeza e preparação preliminar efetuado conforme sessão anterior, e também algumas considerações a serem feitas durante a conversão. Outro ponto importante que o software Disco é um produto comercial, e na sua versão gratuita possibilita a leitura de apenas 100 registros, sendo que o arquivo a ser analisado possui 7485 registros de entrada podendo facilmente ser maior caso fosse utilizado um log de eventos com duração maior que 30 dias.

Para viabilizar a conversão do arquivo para análise, optou-se pelo desenvolvimento de uma macro em VBA formatando o log de eventos conforme modelo proposto por van Dogen e van der Aalst (2005) apresentando na sessão dois nesse artigo. Foram necessárias algumas considerações acerca do log de eventos durante para o processo de desenvolvimento do código:

- a) O primeiro case id (0) foi desconsiderado para conversão devido ao corte sofrido durante a extração do log, dessa forma não haverá atividades de início incorretas para análise no ProM;
- b) As atividades não estão bem definidas com relação à sua duração com o *timestamp* indicando somente o início (*event type start*). Como o término de uma atividade nesse processo em análise é precedido pelo início da próxima, foi considerado esse detalhe para gerar o *timestamp* do fim da atividade (*event type complete*);
- c) Os atributos de cada tarefa foram considerados somente no evento de início (*event type start*), por não trazerem ganho de informação relevante ao ser inserido no evento do fim de atividade (*event type complete*) já que seriam iguais;
- d) Foi considerado o formato padrão de data utilizado pelo ProM na conversão do arquivo ("yyyy-mm-ddThh:nn:ss.000Z") onde além da data e hora, o fuso horário também é incluído.

Sendo assim, o arquivo foi convertido conforme o meta-modelo sugerido por van Dogen e van der Aalst (2005), com a extensão MXML e codificação em UTF-8 (8-bit *Unicode Transformation Format*) para correta leitura no ProM 5.2. Dessa forma, foi possível fazer a carga do arquivo em uma ferramenta de mineração de processos para avaliação e análise do log de eventos.

#### 4 Resultados

A correta conversão do arquivo executada na sessão anterior, pode ser observada pela sua importação na ferramenta ProM 5.2 na tela *Filter* onde parte dos eventos e suas características podem ser observadas conforme figura 5.

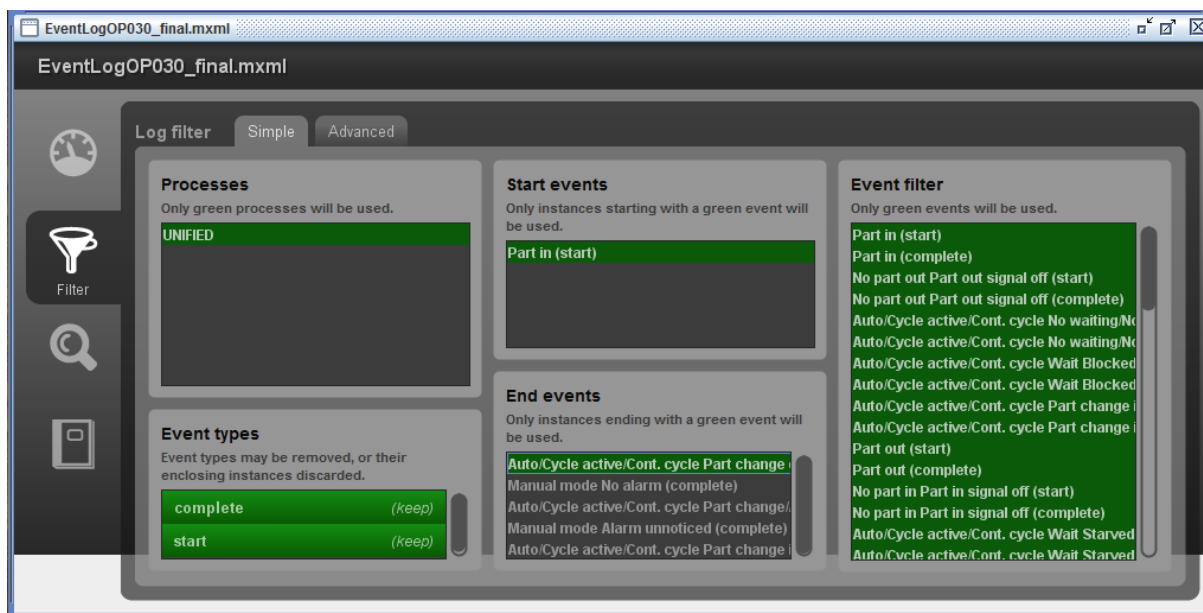


Figura 5 - Tela do ProM com o log de eventos carregado

Para uma visualização um pouco mais clara com relação aos registros do log de eventos utilizados nesse artigo, foi utilizado o plug-in *Fuzzy miner*, com parâmetro *Node Filter Significance cutoff* em 0,508 e os demais em ajuste padrão (figura 6). Nesse modelo é possível observar as atividades de início e fim desse processo conforme análise na fase de preparação do log de eventos. Finalizado a exportação do arquivo para o uso em ferramentas de mineração de processos, é possível fazer uma análise adequada apontando possíveis desvios e propondo pontos de melhorias para o processo em questão.

Outros plug-ins de mineração de processos podem ser aplicados ao log de eventos para análise e outros diagnósticos, dado a validade da conversão dos dados no formato suportado pelo ProM. Apesar do foco principal desse artigo ser a limpeza, preparação e formatação dos dados de um log de eventos, o arquivo será submetido a dois plug-ins para validação de resultados através da mineração de processos.

A partir do filtro considerando somente as instâncias completas, os eventos foram submetidos ao plug-in de análise *Perform Sequence Diagram Analysis* para identificação dos padrões existentes. Foram identificados 46 padrões relativos à sequência dos eventos, sendo que das 753 instâncias completas, 668 pertencem ao primeiro padrão encontrando que representa a sequência de funcionamento normal do equipamento. Entretanto, com base na estatística do tempo de atravessamento dessas atividades possível observar uma variação considerável entre os valores mínimos e máximos (12,93 min. – 89,27min.) conforme pode ser observado na figura 7.

A análise anterior pode ser complementada com o plug-in *Performance Analysis with Petry net*, no qual é possível ver a estatística do tempo total do processo relativo às instâncias analisadas no plug-in anterior. Entretanto, possibilita o foco dos tempos individuais entre as

transições das atividades, revelando possíveis pontos de gargalo durante a operação automática do equipamento (figura 8).

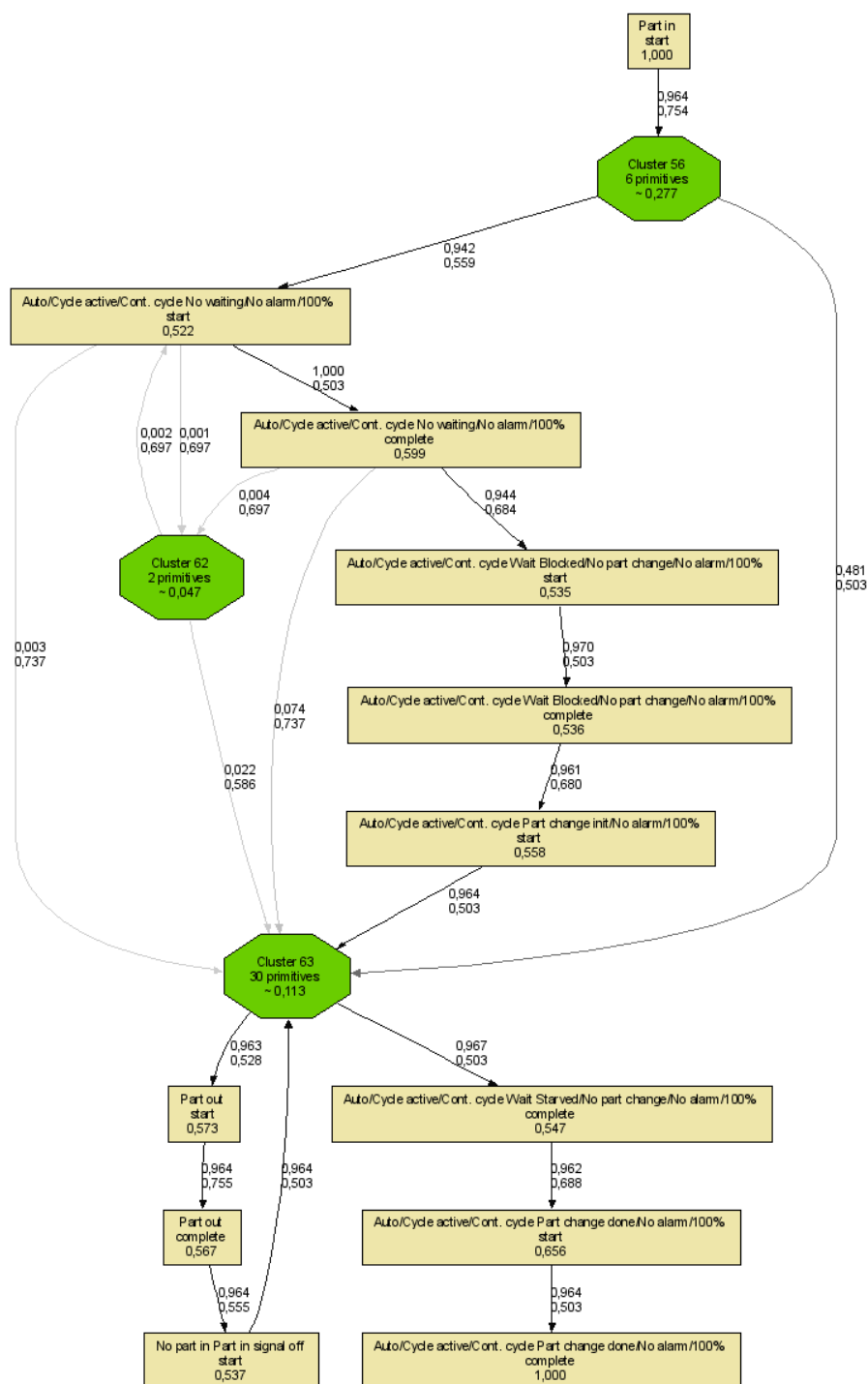


Figura 6 - Plug-in Fyzzzy miner aplicado ao log de eventos

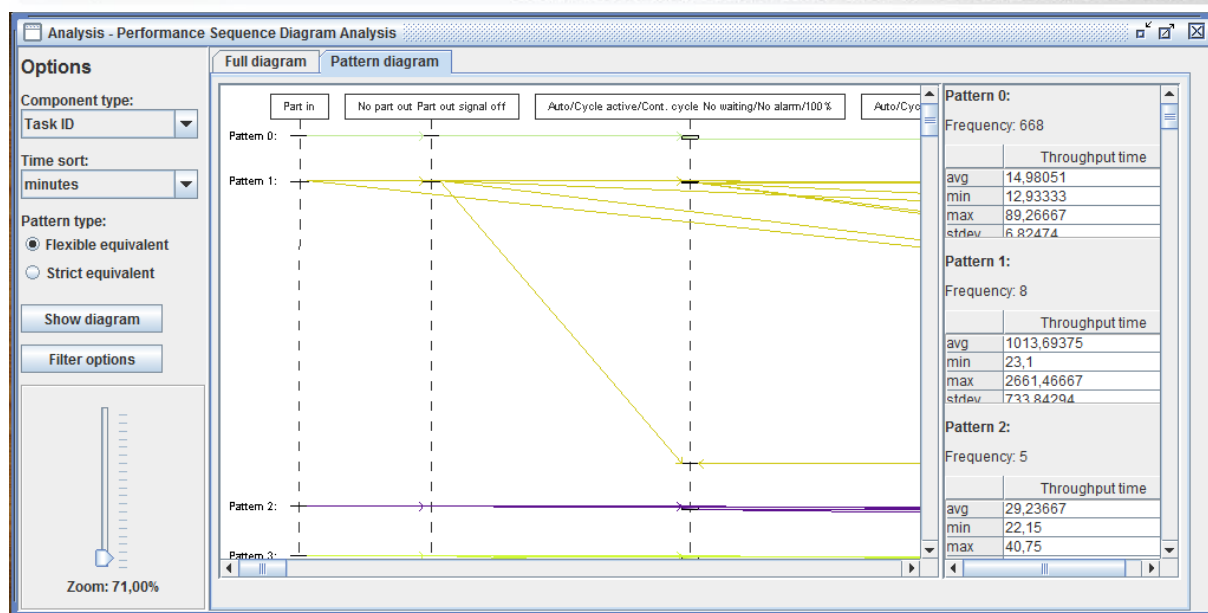


Figura 7 - Análise dos padrões do log de eventos

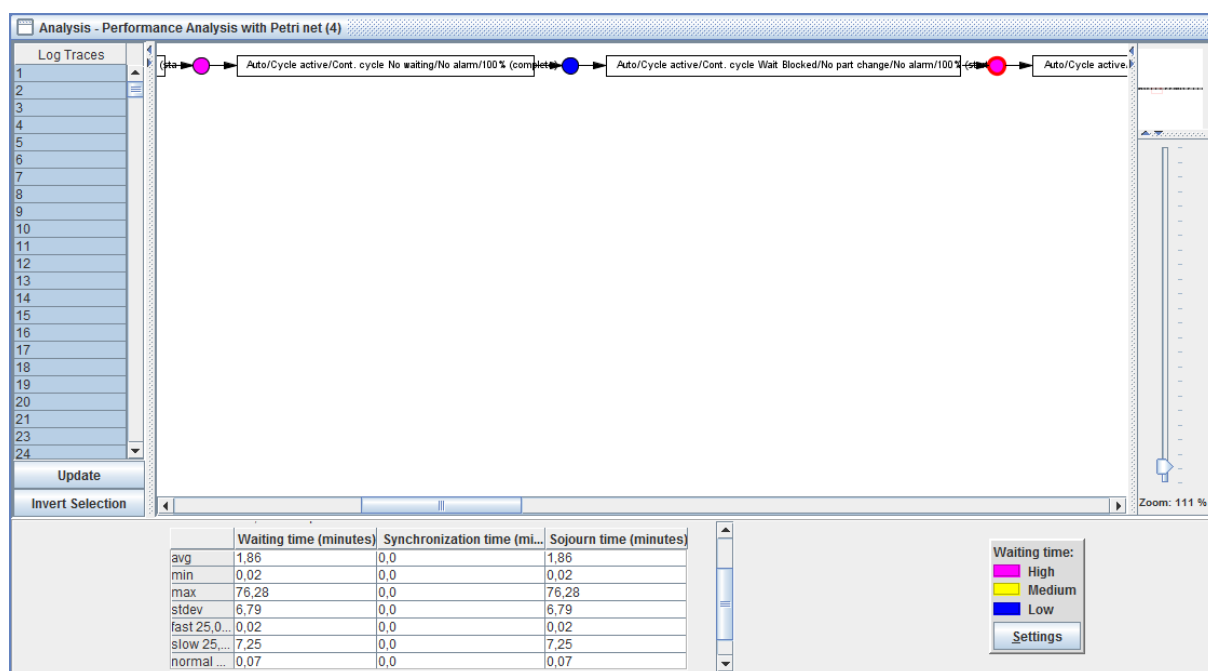


Figura 8 - Análise de gargalos do ciclo em automático

## 5 Conclusões Finais

Levando em consideração a importância dos dados para a mineração de processos, esse artigo manteve o foco na origem da principal matéria prima para a atividade. As ferramentas tem evoluído e suportam formatos diversos, entretanto a fonte onde esses dados são gerados não caminham na mesma velocidade, sendo necessário a sua correta limpeza e preparação antes de submeter a um processo de análise em uma ferramenta.

Para que um processo de limpeza e preparação de dados tenha sucesso, é imprescindível um bom conhecimento do processo a ser tratado, para que não fiquem resíduos desnecessários que não agregam em termos de informação, e dados importantes não sejam eliminados

comprometendo a fidelidade do log de eventos. A correta formatação conforme o meta-modelo de dados proposto, MXML no caso desse artigo, é fundamental para a importação do log de eventos sem erros adicionais.

Apesar de ser efetuada uma análise superficial do log de eventos para fins de validação dos dados, é possível observar resultados que não são facilmente visualizados em relação ao desempenho do equipamento. Como a análise foi feita sobre o padrão de funcionamento normal do equipamento, era esperado uma estabilidade em relação ao tempo das atividades automáticas, que não ocorreu, sendo necessário uma análise mais profunda para descobrir os motivos desses desvios.

Como proposta de trabalhos futuros, é sugerido uma análise detalhada de um log de eventos de uma máquina automática sob a lente da mineração de processos, que pode trazer benefícios com relação à performance do equipamento em relação à eventos não conectados a paradas registradas com alarmes. Outra possibilidade após análise, seria aumentar a granularidade dos eventos registrados mapeando todas as etapas de manufatura do equipamento, a ponto de ser possível correlacionar os principais componentes e seus respectivos modos de falha e o efeito do impacto na performance da máquina.

### **Referências**

**KONDO, RICARDO E.; SANTOS, EDUARDO A. P.; LOURES, EDUARDO F. R.** *Mineração de Processos Para a Identificação de Padrões De Falhas Para A Racionalização de Alarmes*. Anais do XIX Congresso Brasileiro de Automática, Campina Grande, v.unico, pp. 2958-2965, Set. 2012.

**PARK, SUNGBUM; KANG, YOUNG SIK.** *A Study of Process Mining-based Business Process Innovation*. Procedia Computer Science, vol. 91 pp. 734 – 743, 2016.

**VAN DER AALST, W.M.P.** *Process Mining: discovery, conformance and enhancement of business processes*. 1.ed: Springer-Verlag Berlin Heidelberg, 2011.

**VAN DER AALST, W.M.P. et al.** *Process Mining Manifesto*. Business Process Management Workshops 2011, Lecture Notes in Business Information Processing, Vol. 99, Springer-Verlag, 2011.

**VAN DER AALST, W.M.P.; REIJERS, H.A.; VAN DONGEN, B.F.; MEDEIROS, A.K. ALVES DE; SONG, M.; VERBEEK, H.M.W.** *Business process mining: An industrial application*. Information Systems. Vol. 32, pp. 713-732, 2007.

**VAN DONGEN, B.F. ; VAN DER AALST, W.M.P.** *A Meta Model for Process Mining Data*. In Proceedings of the CAiSE WORKSHOPS, vol. 2, pp. 309-320, 2005.

**VAN DER AALST, W.M.P.; WEIJTERS, A.J.M.M.** *Process mining: a research agenda*. Computers in Industry, vol. 53, pp. 231-244, 2004.