# MINERAÇÃO DE PROCESSOS APLICADA À LOGÍSTICA DE UMA EMPRESA DE ÓLEO E GÁS

**Nickolas Gueller Rocha**
Pontifícia Universidade Católica do Rio de Janeiro (PUC-RIO)
R. Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
nickolasgheller@gmail.com

**Leila Figueiredo Dantas**
**Silvio Hamacher**
**Luiza Fiorencio**
**Bruno Leite Mariani**
Instituto Tecgraf, PUC-RIO
R. Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22451-900
leilaffdantas@gmail.com, hamacher@puc-rio.br, luiza.fiorencio@tecgraf.puc-rio.br,
brunomke@tecgraf.puc-rio.br

**Pedro Henrique Sousa**
CENPES / Petrobras
Av. Horácio Macedo, 950 - UFRJ, Rio de Janeiro - RJ, 21941-915
pedro.marins@petrobras.com.br

## RESUMO

A mineração de processos é uma prática usada na elaboração de modelos, monitoramento de atividades e análise e aprimoramento de processos. O presente trabalho descreve a aplicação de técnicas de mineração de processos ao setor de logística de uma empresa brasileira de óleo e gás. O objetivo é encontrar o modelo atual de processo a partir da análise do histórico de eventos, verificar seu desempenho e conformidade e fornecer descobertas sobre os processos – apresentando também uma metodologia que pode ser aplicada a outros tipos de dados. Seguindo as etapas de Descoberta, Conformidade e Aprimoramento, foi possível identificar desvios e oportunidades de melhorias. Os resultados apresentaram uma taxa de 37% de ocorrência de desvios no processo, permitindo a elaboração de um modelo real com adequação de 100%. A análise de desempenho permitiu identificar a transição mais demorada do processo, a duração de todos os fluxos existentes e detalhar a atividade cancelamento.

**PALAVRAS CHAVE. Mineração de processos; Logística; Indústria de óleo e gás.**
**Tópicos: PO na Área de Energia, Petróleo e Gás**

## ABSTRACT

Process mining is a practice used in modeling, activity monitoring, and process analysis and improvement. This paper describes the application of process mining techniques to the logistics sector of a Brazilian oil and gas company. The goal is to find the current process model from event log analysis, verify its conformance with a normative model, and provide insights – besides presenting a methodology that can be applied to other types of data. Following the Discovery, Conformance and Enhancement steps, it was possible to identify deviations and opportunities for improvement. The results presented a 37% rate of occurrence of deviations in the process allowing the elaboration of a real model with fitness of 100%. The performance analysis of the activities allowed to identify the duration of all existing flows, the longest transition of the process and to detail the cancellation activity.

**KEYWORDS. Process Mining; Logistic; Oil and gas industry.**
**Paper topics: OR in Energy, Oil and Gas**

## 1. Introdução

A indústria de óleo e gás representa um importante pilar para a economia brasileira. Dados divulgados pelo Programa de Mobilização da Indústria Nacional de Petróleo e Gás Natural, em 2013, apontam que setores de petróleo e gás já chegaram a compor 12% do Produto Interno Bruto (PIB) brasileiro [Prominp 2013]. Segundo a Agência Nacional do Petróleo (ANP), estimativas apontam que a demanda mundial de petróleo atingirá seu pico em 2030 [ANP 2018], fato que acaba por destacar a necessidade de o setor brasileiro ser capaz de elevar sua produtividade neste prazo.

A exploração de óleo e gás depende diretamente da construção de poços marítimos, onde ocorrem as etapas de extração e produção. A construção destes poços envolve diversas questões logísticas, tais como a movimentação de recursos, elaboração de cronogramas, programação de sondas e prevenção de riscos geológicos e ambientais, além da integração com a cadeia de suprimentos. As atividades em plataforma de alto mar *(off-shore*), por conta do espaço e recursos limitados, exigem uma logística bem planejada, podendo acarretar em atrasos e prejuízos na ocorrência de erros.

Segundo um estudo realizado pela FGV Energia em 2016 [FGV Energia 2016], houve um aumento significativo de multas aplicadas ao setor. Dados divulgados pela ANP indicam uma redução de 89% na conclusão de poços exploratórios do ano de 2012 para 2018 [ANP 2018]. Os desafios deste setor envolvem ampliar sua produção, de forma que o território brasileiro seja melhor explorado, assim como atingir a capacidade de produção esperada. A tomada de decisão neste setor também é um problema que deve ser superado, visto que decisões erradas podem gerar prejuízos de bilhões de reais.

Sendo assim, a fim de superar os problemas citados, faz-se cada vez mais necessário entender o funcionamento dos processos logísticos envolvidos e identificar as oportunidades de melhoria. Para isso, o uso de novas ferramentas como a mineração de processos (*Process Mining*) pode ser útil, sobretudo por seu histórico de aplicações em logística, se mostrando como uma alternativa viável para melhorias serem alcançadas.

A mineração de processos começou a ser desenvolvida no início de 1999 por um grupo de pesquisadores da universidade de TUE, Eindhoven. A técnica, ainda em desenvolvimento na época, era chamada de "*workflow mining*", e se baseava na ideia de extrair informações de histórico de eventos (*event log*), adaptando técnicas de *data mining* e elaborando algoritmos para tal.

O presente trabalho propõe aplicar as técnicas de mineração de processos ao setor de logística de uma grande empresa brasileira de óleo e gás. Essa aplicação tem como intuito encontrar o modelo real do processo, verificar a conformidade, avaliar o desempenho e fornecer descobertas sobre o processo a partir da análise do histórico de eventos, com o principal objetivo de identificar desvios e oportunidades de melhorias. Espera-se que melhorias possam ser sugeridas e incorporadas à logística da empresa. Além disso, o trabalho também apresenta uma metodologia de mineração de processos que pode ser aplicada a outros tipos de bases de dados.

O trabalho se estrutura da seguinte forma: A Seção 2 apresenta uma revisão bibliográfica a respeito do assunto. A Seção 3 realiza uma descrição do problema em questão, enquanto a Seção 4 explica os métodos e ferramentas utilizadas. A Seção 5 apresenta os resultados obtidos na pesquisa. Por fim, a Seção 6 conclui o trabalho realizado.

## 2. Revisão Bibliográfica

Atualmente o mundo é capaz de gerar, em cerca de 10 minutos, mais informação do que foi gerada desde a pré-história até o ano de 2003. A preocupação agora, portanto, não se trata mais de como armazenar dados, mas sim como transformar toda essa informação em algo de valor [Aalst 2016]. As técnicas de mineração de processos surgiram como uma solução para o problema citado. O ramo cria uma ponte entre o *Data Science* e o *Business Process Management*, estabelecendo uma combinação das metodologias de análise de processos e análise de dados, possibilitando ações que tais técnicas individualmente não são capazes de efetuar, além de possuir diversas relações com o *Data Mining* [Aalst 2011].

A aplicação destas técnicas é possível em uma variedade de situações práticas, permitindo a descoberta, análise, diagnóstico e melhoria de processos por meio de análises do histórico de eventos [Weerdt et al. 2013]. Além disso, a técnica também permite a descoberta de modelos, capazes de mostrar à empresa os desvios de processos [Aalst 2016], possibilitando uma comparação entre o "teórico" e o "real".

As técnicas de mineração de processos podem ser aplicadas em praticamente todos os setores, desde a área de finanças à logística e cadeia de suprimentos. O único requisito necessário, assim como o ponto de partida da técnica, é a existência de um histórico de eventos, isto é, uma base de dados contendo os registros de todas as atividades realizadas dentro da organização em um espaço de tempo considerável, levando em conta diversos detalhes das atividades [Aalst 2016].

A abordagem da técnica se baseia em 3 etapas: descoberta, conformidade e aprimoramento. A primeira etapa consta da descoberta de um modelo de processos a partir do histórico de eventos. Esse modelo elaborado objetiva o entendimento do funcionamento real de uma empresa. A segunda etapa promove uma comparação entre o modelo de processo existente e o *log* de eventos, identificando as diferenças e coincidências encontradas. Por fim, a última etapa promove adaptações no modelo, considerando o diagnóstico da etapa anterior, procurando aproximá-lo do processo real [Aalst 2011]. A Figura 1 traz uma síntese a respeito das etapas e suas relações com outros fatores.

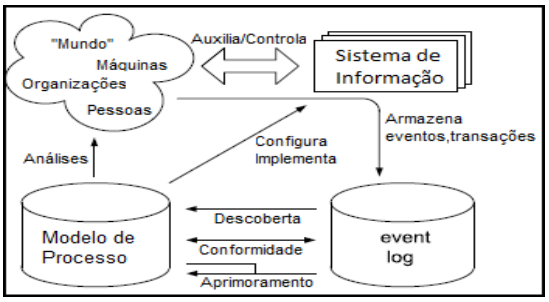



Figura 1 – Visão geral do ambiente da mineração de processos. Adaptado de [Aalst 2011]

[Becker et al. 2017] reconheceram a mineração de processos como uma ferramenta potente para a análise e aprimoramento de processos logísticos, útil para a elaboração de modelos e monitoramento de atividades. [Weerdt et al. 2013] mostraram o uso da técnica em uma companhia de seguros, por meio da qual descobriram erros em transferências bancárias. Estudos na China utilizaram mineração de processos na melhoria de resgates em acidentes de minas de carvão [He et al. 2019]. Também com aplicações na área da saúde, as técnicas de mineração de processos foram utilizadas em um hospital holandês, onde os resultados mostraram a capacidade de melhorias em fluxos assistenciais por meio da técnica [Mans et al. 2008].

[Beest e Maruster 2007] desenvolveram um trabalho na indústria de óleo e gás, onde foi feita uma abordagem por mineração de processos para otimizar os processos de produção e logística. O estudo permitiu a elaboração de um modelo realístico dos processos envolvidos, assim como a especificação dos fluxos de processo, duração de atividades e gargalos envolvidos, permitindo a previsão e controle de possíveis erros nos processos.

Após pesquisas na literatura brasileira a respeito da mineração de processos e suas aplicações em logística, chega-se à conclusão que há uma escassez de trabalhos que unam essa área e ferramenta. Sendo assim, além de obter melhorias para a empresa em estudo, o presente trabalho também propõe a junção destes assuntos, agregando conhecimento para pesquisadores e especialistas da área.

## 3. Descrição do Problema

A partir de 2016 iniciou-se o desenvolvimento de um sistema computacional que visava melhor controlar e otimizar os processos envolvidos na logística de equipamentos de construção de poços marítimos da empresa em estudo. Sendo assim, aliando a necessidade de melhor entender os fluxos de atividades logística e a oportunidade de extração do registro de eventos, a técnica de

mineração de processos começou a ser utilizada no monitoramento constante e análise de suas atividades.

Dentre diversas atividades recorrentes na cadeia, foram selecionadas as relacionadas à Solicitação de Movimentação (SM), que tratam da movimentação de recursos (itens), tais como suspensores, conjuntos de vedação, entre outros. As atividades a serem analisadas, portanto, tomam início na criação da SM, e são finalizadas com o Atendimento ou Cancelamento da SM.

O objetivo da aplicação da mineração de processos é permitir a análise do fluxo ao longo deste processo, verificando a existência de problemas em todas as etapas, desde a programação até a conclusão da movimentação dos itens. O trabalho foi iniciado conforme a necessidade de entender os processos envolvidos, identificar durações de atividades e verificar a conformidade entre os processos reais e o modelo normativo. Com a identificação de problemas existentes, torna-se possível a geração de soluções que auxiliam a gerência do sistema e otimizam esta parcela da cadeia de suprimentos da empresa, reduzindo o gasto de recursos e tempo com processos menos eficientes.

## 4. Metodologia e ferramentas

O presente trabalho seguiu a aplicação da mineração de processos proposta por [Aalst 2016], a qual se define em 3 etapas: descoberta, conformidade e aprimoramento. Para a realização das etapas, foram utilizados os softwares DISCO 2.0, YASPER 2.1b4 e PROM 6.8. A Figura 2 traz uma síntese e contribuição a respeito da metodologia adotada neste trabalho, que pode ser reaplicada a outros tipos de dados.

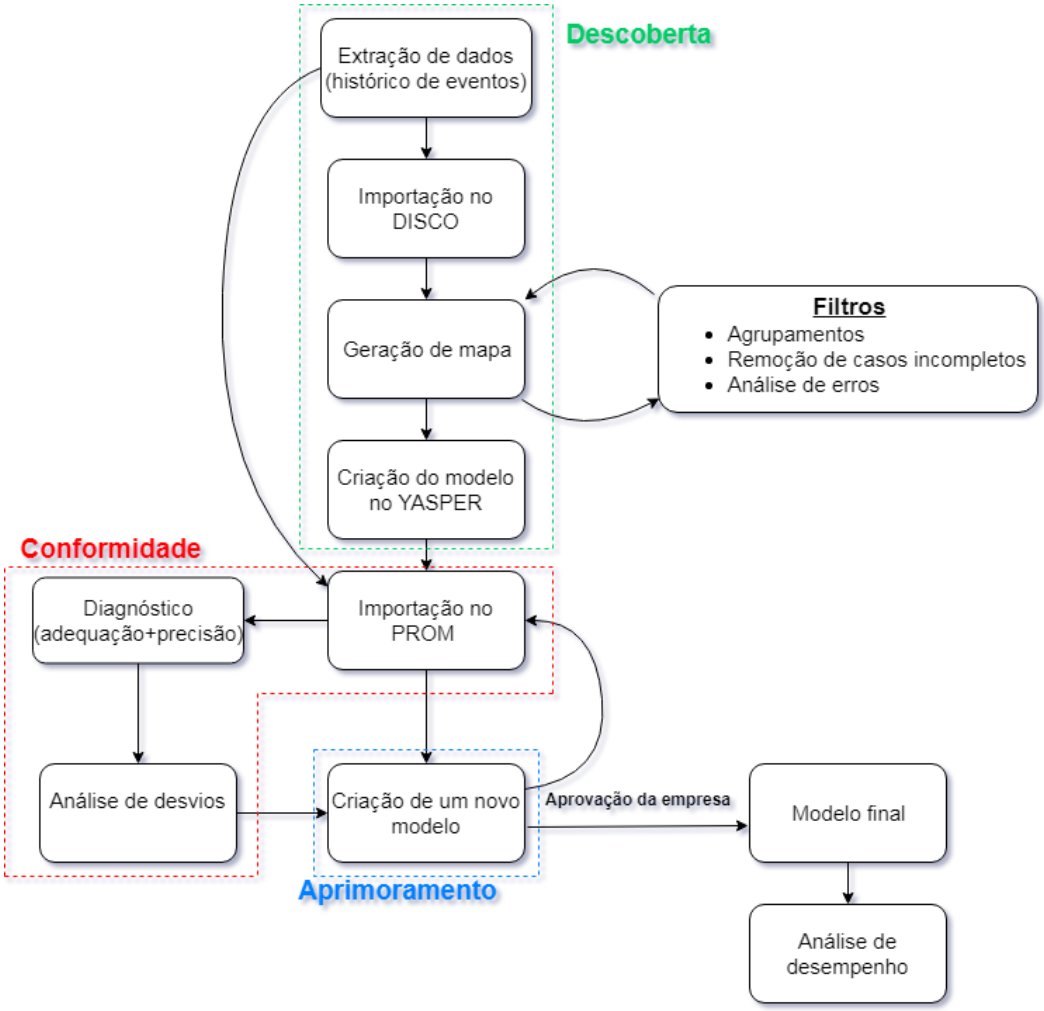



Figura 2 – Visão geral da metodologia utilizada

### 4.1. Descoberta

A primeira etapa consta da descoberta de um modelo de processos a partir do histórico de eventos. O modelo descoberto (Modelo *AS-IS*) apresenta o fluxo real das atividades da empresa, ou seja, o caminho percorrido por cada Solicitação de Movimentação.

Os dados do sistema computacional foram extraídos em XML. Este histórico de eventos pode ser interpretado como uma coleção de casos únicos, e cada caso pode ser visto como uma sequência de eventos. Isso significa que cada evento do histórico se refere a um fluxo de atividades e está relacionado a um caso particular (*case ID*).

Foram extraídos registros referentes ao período de 02/04/2018 a 27/03/2019, contendo no total 469 casos e 3861 eventos. Cada ocorrência presente no registro de eventos é composta das seguintes informações:

- Identificação da solicitação de movimentação (SM) – *Case ID*;
- Atividade – podendo ser classificada em Ação ou Estado;
- Data e hora em que a atividade foi executada;
- Responsável pela execução da atividade (recurso); e
- Identificação do projeto ao qual a SM está associada (tratado como atributo da SM).

As atividades selecionadas foram:

- Atividades de Ação: Criar SM; Submeter SM; Validar SM; Reprovar SM; Cancelar SM; SM Em andamento;
- Atividades de Estado: SM Criada; SM Aguardando Aprovação; SM Aprovada; SM Recusada; SM Cancelada; SM Atendida.

As atividades de ação representam ações que necessitam de um recurso, isto é, alguém para realizá-las, estabelecendo fluxos de trabalho. As de estado, por sua vez, apenas indicam estados das atividades no sistema, mostrando que estão no aguardo da próxima atividade de ação. Esta divisão foi realizada para auxiliar a análise de tempo das atividades, porém, algumas das atividades de estado não precisaram ser analisadas individualmente, fato que será melhor abordado adiante.

Selecionadas as atividades para compor o histórico de eventos, a aplicação da mineração de processos foi iniciada utilizando a ferramenta DISCO. A ferramenta serviu para a geração de um modelo referente aos processos envolvidos, o modelo real, permitindo uma fácil visualização dos fluxos presentes no histórico de eventos. Antes de seguir para a criação do modelo em si, algumas análises foram realizadas a partir do mapa gerado.

A primeira análise a ser feita foi a dos tempos de duração entre as atividades. Conforme explicado anteriormente, as atividades foram separadas entre “ação” e “estado”, o que possibilitou a verificação da existência de atrasos no sistema em questão. Durante a análise, observou-se que o tempo de duração entre algumas atividades de ação e estado eram correlacionadas, como por exemplo “criar SM” e “SM criada”, onde a duração entre elas era praticamente nula (menos de um segundo). Levando em conta essa observação, além do fato destas atividades serem realizadas sempre pelo mesmo usuário, foi realizado um agrupamento de atividades, fato que tornou o mapa exibido mais compreensível. Segue na Tabela 1 o agrupamento realizado, promovendo uma redução de quatro atividades. Após o agrupamento realizado, conforme esperado, a média de duração dos casos não se alterou.

Tabela 1 – Atividades Ação e Estado agrupadas

| **Atividade (pós agrupamento)** | **Atividades agrupadas** |
|---|---|
| Criar SM | {Criar SM; SM Criada} |
| Submeter SM | {Submeter SM; SM aguardando aprovação} |
| SM Cancelada | {Cancelar SM; SM Cancelada} |
| SM Recusada | {Reprovar SM; SM Recusada} |

Após as atividades serem agrupadas, o próximo passo foi remover os casos incompletos do *log* de eventos. Sistemas de informações normalmente realizam registro de forma contínua, o que indica que ao extrair dados é muito provável a ocorrência de casos ainda não finalizados. Sendo assim, os casos incompletos foram removidos, permitindo a análise apenas de SMs que entraram e

já saíram do sistema. Foram considerados como incompletos todos os casos que não possuíam "SM Atendida" ou "Cancelar SM" como atividade fim, levando em conta também que todos os casos se iniciam em "Criar SM". Por fim, foram removidos erros de processo identificados, que serão melhor explicados na Seção 5. A Figura 3 exibe de forma resumida os critérios de exclusão aplicados, junto do impacto quantitativo recorrente de cada um deles.

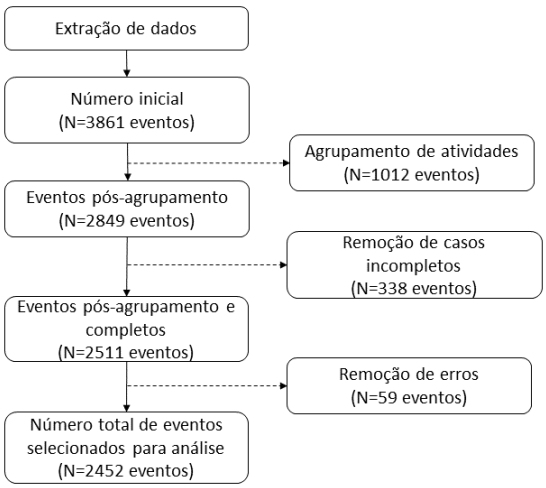



Figura 3 – Critérios de exclusão de eventos durante a descoberta do modelo

Após as alterações, pode-se observar o mapa do modelo *AS-IS* na Figura 4, contendo 385 casos completos e 2452 eventos. Os números dentro dos quadrados representam a frequência absoluta da atividade em questão, enquanto os números nas setas representam a quantidade de casos que transitaram de uma atividade para a outra. Tomando como exemplo o fluxo "Criar SM" → "Submeter SM", a imagem mostra que 385 SMs foram criadas e 374 destas seguiram para "Submeter SM", onde a atividade ocorreu 375 vezes, levando em conta repetições. Os círculos apresentam o início e fim do processo.

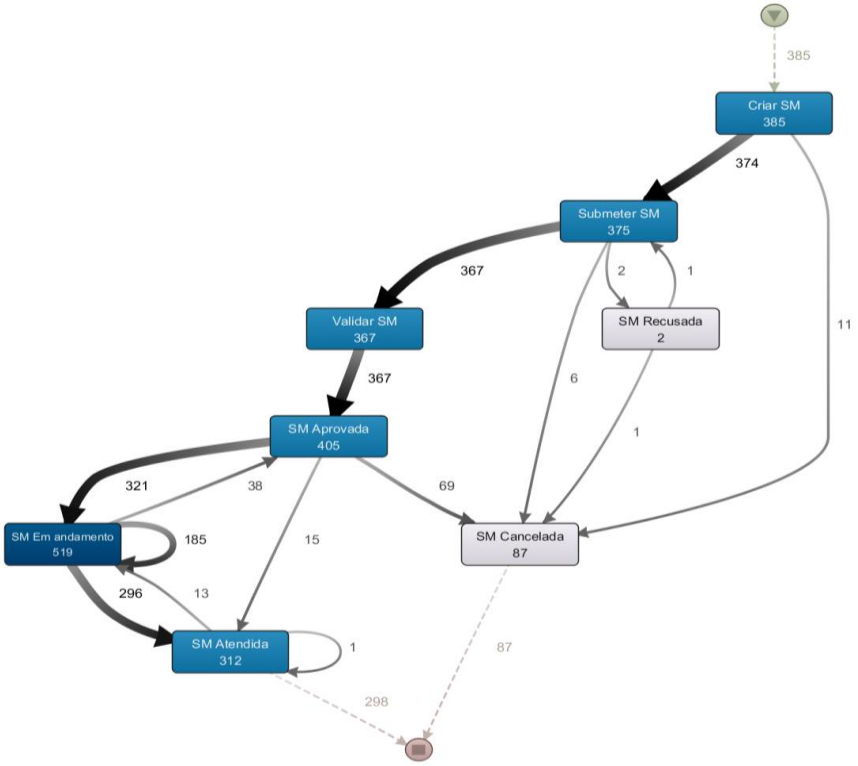



Figura 4 – Mapa do processo real da empresa exibido no DISCO após aplicação dos critérios de exclusão de eventos.

Com o objetivo de comparar o modelo teórico (inicial) da empresa com o fluxo real (Figura 4), foi gerado, manualmente, através do software YASPER e utilizando-se da representação de "Rede Petri", um modelo teórico para a Solicitação de Movimentação na Figura 5. Este modelo contém todas as atividades do processo, mas não todos os fluxos e desvios, fazendo com que sejam necessários ajustes a fim de se obter um modelo mais realístico.

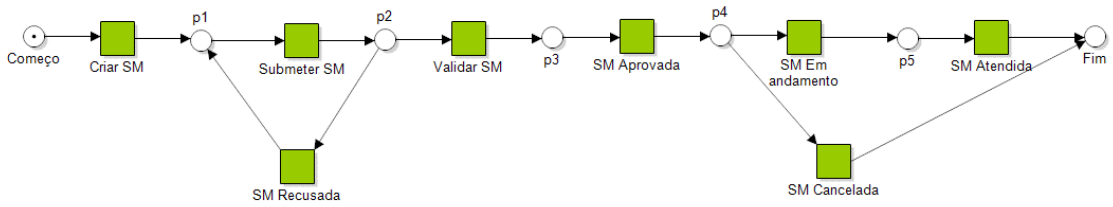



Figura 5 – Modelo inicial/teórico da empresa, elaborado no YASPER

Nos modelos Rede Petri, os quadrados verdes representam atividades a serem realizadas, os nós de transição, enquanto os círculos brancos são os nós de posição. Os modelos Rede Petri deste trabalho podem também conter atividades nomeadas "tau", que se tratam de "atividades invisíveis", com o único propósito de possibilitar a ocorrência de fluxos específicos.

### 4.2. Conformidade

A etapa de conformidade se baseia em comparar o modelo de processo inicial (Figura 5) com o histórico de eventos correspondente, comparando o comportamento representado por cada um. De forma concisa, é realizado um diagnóstico do modelo e ajustes são feitos no modelo de "Rede Petri" com o objetivo de obter modelos mais reais.

A partir das diferenças e semelhanças encontradas, é possível medir o nível de adequação (*fitness*) e de precisão (*precision*) do modelo e histórico de eventos, onde ambos variam entre 0 e 1. A adequação é um indicador que representa uma comparação entre os comportamentos presentes no *log* de eventos e os que se encontram no modelo. Quanto mais próximo de 1, mais adequado e realístico será o modelo. Esta medida, de uma maneira geral, responde quantos traços contidos no log de eventos se encontram de fato dentro do modelo em questão, portanto, o valor "1" mostra que o modelo contém todos os traços do log de eventos e o valor "0" mostra que nenhum traço do log se encontra no modelo.

A precisão, por sua vez, é o indicador do percentual de comportamentos presentes no modelo que não se encontram no *event log*, isto é, ela responde quantos traços de eventos o modelo possibilita que não se encontram dentro do log de eventos. Quanto mais próximo de 1, menos genérico e mais preciso será o modelo. O ideal é balancear essas duas medidas de desempenho de acordo com o objetivo da análise, levando também em consideração a simplicidade do modelo. A ferramenta utilizada para a análise de conformidade foi o PROM.

O PROM é um software que reúne um conjunto de *plug-ins* que possibilitam diversos tipos de análises de mineração de processos. Para o presente trabalho foram utilizados o *plug-in* "*Replay a Log on Petri Net for Performance/Conformance Analysis*" para calcular o ajuste dos modelos e análise de cada desvio e o *plug-in* "*Check Precision based on Align- ETConformance*" para calcular os níveis de precisão.

### 4.3. Aprimoramento

A partir do diagnóstico realizado na etapa anterior, a atual etapa propõe mudanças no modelo de processo inicial, de forma a analisar os desvios e aproximá-lo do comportamento real presente no histórico de eventos.

A partir do uso iterativo das ferramentas PROM e YASPER, foram realizadas mudanças contínuas no modelo. Esta etapa do trabalho funciona da seguinte forma: a partir dos resultados da etapa de conformidade, adiciona-se os desvios/mudanças no modelo atual e gera-se um novo modelo. Esse processo é repetido até a criação de um modelo final aprovado pela empresa com

alto valor de *fitness* e boa precisão. Esse *loop* e o passo a passo utilizado na metodologia podem ser visualizados na Figura 2.

Após as etapas de descoberta, conformidade e aprimoramento, além de interações com a empresa, o modelo final foi encontrado e análises de desempenho permitiram identificar a transição mais demorada do processo, a duração de todos os fluxos existentes e detalhar a atividade de cancelamento.

## 5. Resultados

A aplicação da mineração de processos proposta pelo trabalho foi capaz de fornecer informações relevantes sobre os processos contemplados pelo sistema, facilitando a tomada de decisões. A aplicação da metodologia proposta possibilitou visualização dos desvios e erros presentes, elaboração de um modelo de processo real, identificação da duração das atividades e outras análises de desempenho, como a respeito dos cancelamentos ocorridos.

### 5.1. Erros e desvios identificados

O primeiro mapa observado no DISCO, ainda na etapa de descoberta, já indicou a ocorrência de fluxos que não deveriam ocorrer no sistema. Foram observados 11 casos que faziam o caminho “Validar SM → SM Atendida” e três casos que passaram por “Validar SM → SM Em Andamento”. Porém, de acordo com a empresa, todos os casos de SM deveriam passar pelo estado “Aprovada” antes de entrar em andamento ou ser atendida.

Essa análise de erros sugeriu melhorias no sistema, que como consequência geraram mudanças na forma de registrar as atividades. Ou seja, a análise realizada através da mineração de processos permitiu que alguns fluxos fossem bloqueados no sistema, evitando que voltassem a acontecer. Sendo assim, esses 14 casos (59 eventos) foram excluídos do histórico de eventos (Figura 3) por serem considerados erros. Os erros identificados, portanto, compõem 3,5% do número de casos total.

Após a análise de erros, foi feita uma análise dos desvios presentes no processo. Foram classificados como desvios todos os fluxos que não faziam parte do modelo inicial/teórico da empresa da Figura 5. Por exemplo, foram identificados 13 casos de SMs que, apesar de passarem pelo “SM Atendida”, voltaram para o sistema e continuaram no processo passando por outras atividades, onde 8 deles se encerraram em “SM Cancelada” e 5 em “SM Atendida”.

Dois *loops* foram observados na Figura 4: SM Em andamento→ SM Em andamento; SM Atendida →SM Atendida. O primeiro *loop* apresenta uma frequência de 74 casos e 185 eventos, sendo possível um mesmo caso executar esta atividade várias vezes, gerando um número excessivo de eventos. Com essa informação, a empresa decidiu bloquear a ocorrência do *loop*. O segundo *loop* citado apresentou apenas uma ocorrência, onde ela foi identificada e enviada para avaliação da empresa. Todos os desvios identificados no processo podem ser observados na Tabela 2. Eles não são considerados erros, apenas desvios do modelo teórico e, portanto, não são descartados do modelo real do processo.

Tabela 2 – Desvios identificados

| **Fluxo desviante** | **Frequência (eventos)** |
|---|---|
| Criar SM→SM Cancelada | 11 |
| Submeter SM→SM Cancelada | 6 |
| SM Recusada→SM Cancelada | 1 |
| SM Aprovada→SM Atendida | 15 |
| SM Atendida→SM Em andamento | 13 |
| SM Em Andamento→SM Aprovada | 38 |
| SM Em Andamento→SM Em Andamento | 185 |
| SM Atendida→SM Atendida | 1 |

Através do framework PROM, foi identificado um total de 143 casos que apresentavam desvios. Vale ressaltar que para serem considerados desvios, bastava a ocorrência de um fluxo que

não estivesse presente no modelo teórico/inicial. Os desvios identificados, portanto, compõem 37% do número de casos.

Os resultados adquiridos com a análise de erros e desvios foram repassados à empresa, que sugeriram melhorias ao sistema em questão, juntamente da análise particular de determinados desvios identificados, que de acordo com a empresa "não deveriam ocorrer".

### 5.2. Modelo final elaborado

A elaboração do modelo final se deu a partir da extensão e aprimoramento do modelo inicial de acordo com os desvios observados. Este é um resultado direto da aplicação da metodologia proposta na Figura 2 seguindo as etapas de conformidade e aprimoramento.

O modelo inicial gerado na Figura 5 apresentava índices de adequação e precisão de 0,92 e 0,99, respectivamente. Seguindo o proposto pela metodologia, foram realizadas mudanças no modelo inicial em que fluxos/desvios foram adicionados. Para cada alteração realizada no modelo usando o YASPER, os níveis de adequação e precisão eram recalculados através do PROM, permitindo avaliar a relevância da mudança realizada. Vale ressaltar que o mesmo *plug-in* utilizado para o cálculo do *fitness* (*Replay a Log on Petri Net for Performance/Conformance Analysis)* também exibe as diferenças entre o modelo e o *log* de eventos, auxiliando na análise dos fluxos. Esta funcionalidade facilita a realização das alterações necessárias para o modelo se aproximar da realidade, assim como permite que os desvios, explicados anteriormente, sejam quantificados.

A partir disto, foi obtido um modelo que representa o processo real com níveis de adequação e precisão de 0,98 e 0,83, respectivamente, exibido na Figura 6. Os níveis alcançados foram considerados satisfatórios, já que o principal objetivo era identificar os processos que realmente ocorrem no sistema em questão, priorizando um nível mais elevado de adequação, mesmo que isto signifique uma diminuição na precisão.

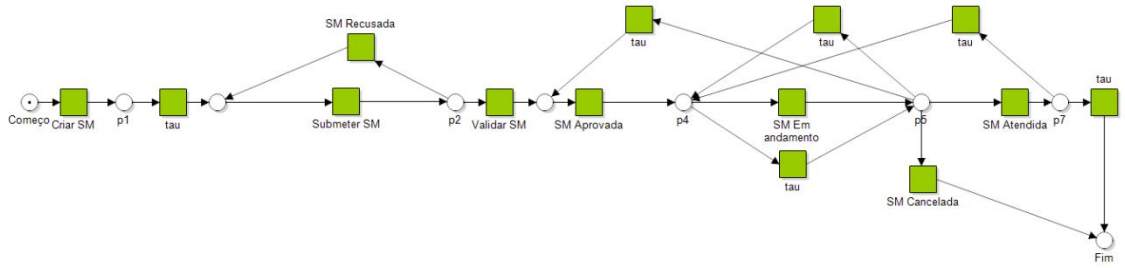



Figura 6 – Modelo final elaborado no YASPER.

### 5.3. Análise de desempenho

Além dos resultados anteriores, foi realizada uma análise de desempenho das atividades, que permitiu uma análise da duração de cada fluxo existente no processo. Como consequência, foi possível identificar que a transição mais demorada do processo se encontra em "SM em andamento → SM Atendida", com uma frequência de 296 casos, média de 19 dias e mediana de 10 dias. A Tabela 3 apresenta a duração e frequência de todos os eventos.

Tabela 3 – Análise de duração e frequência de eventos.

| Atividades | Média | Mediana | Máximo | Mínimo | Frequência absoluta | Frequência relativa |
|---|---|---|---|---|---|---|
| Criar SM→ Submeter SM | 3,8 horas | 0 | 11 dias | 0 | 374 | 19,88% |
| Criar SM→ SM Cancelada | 49 dias | 3,1 dias | 23,4 semanas | 2 min | 11 | 0,58% |
| Submeter SM→ Validar SM | 10,2 horas | 1 min | 57 dias | 0 | 367 | 19,51% |
| Submeter SM→ SM Recusada | 12,5 min | 12,5min | 24 min | 1 min | 2 | 0,11% |
| Submeter SM→SM Cancelada | 12,5 dias | 12,2 horas | 37 dias | 1 min | 6 | 0,32% |
| SM Recusada→ Submeter SM | 3,8 horas | 3,8 horas | 3,8 horas | 3,8 horas | 1 | 0,05% |
| SM Recusada→SM Cancelada | 22,9 horas | 22,9 horas | 22,9 horas | 22,9 horas | 1 | 0,05% |
| Validar SM→SM Aprovada | 0 | 0 | 0 | 0 | 367 | 19,51% |
| SM Aprovada→SM Em andamento | 3,5 dias | 2,5 horas | 19,2 semanas | 1 min | 321 | 17,07% |
| SM Aprovada→SM Cancelada | 7,3 dias | 31 min | 14,1 semanas | 0 | 69 | 3,67% |

| Atividades | Média | Mediana | Máximo | Mínimo | Frequência absoluta | Frequência relativa |
|---|---|---|---|---|---|---|
| SM Aprovada→SM Atendida | 51,3 horas | 0 | 32 dias | 0 | 15 | 0,80% |
| SM Em andamento→SM Atendida | 19 dias | 10 dias | 24,1 semanas | 0 | 296 | 15,74% |
| SM Em andamento→SM Aprovada | 12,9 dias | 60,9 horas | 15,7 semanas | 1 min | 38 | 2,02% |
| SM Atendida→SM Em andamento | 6,5 dias | 20min | 56,2 dias | 1 min | 13 | 0,69% |

Todos esses valores identificados foram repassados para empresa e permitiram que a mesma tomasse iniciativas para melhorar a eficiência do processo. Com relação à transição de maior duração, é importante que ela seja observada constantemente pela empresa, visto que se trata da atividade que mais demanda trabalho manual dentre todas presentes.

Além da identificação acima, também foi possível a identificação do fluxo mais presente no processo (Figura 7), que, apesar de ser o mais frequente, engloba apenas 52,73% dos casos. Os casos em questão apresentaram uma duração média de 18,2 dias e mediana de 10 dias, o que foi capaz de mostrar à empresa o tempo esperado de duração dos processos em geral, ficando a seu critério se o tempo pode ser considerado razoável ou não.

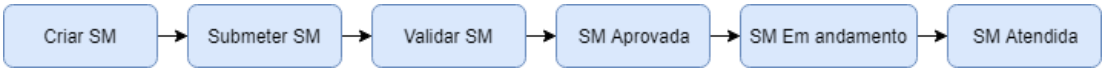



Figura 7 – Fluxo com maior frequência

### 5.3.1. Análise de cancelamentos

Durante a aplicação da técnica foi observado um elevado número de cancelamentos de SMs (87 casos), representando 22,6% do total de casos completos. Realizando-se uma expansão do modelo final descoberto (Figura 6), isto é, adicionando fluxos que permitam o cancelamento da SM em qualquer momento do processo, foi realizada a análise dos tempos de cancelamento envolvidos, a fim de permitir uma visualização dos padrões de cancelamento presentes em cada etapa do processo. O novo modelo encontrado na Figura 8 apresentou índices de adequação e precisão de 1,0 e 0,81, respectivamente. O índice de adequação máximo indica que o modelo mostra todos os desvios presentes no processo, assim como a precisão de 0,81 indica que o modelo não permite a ocorrência de muitos fluxos que não se encontram no processo real.

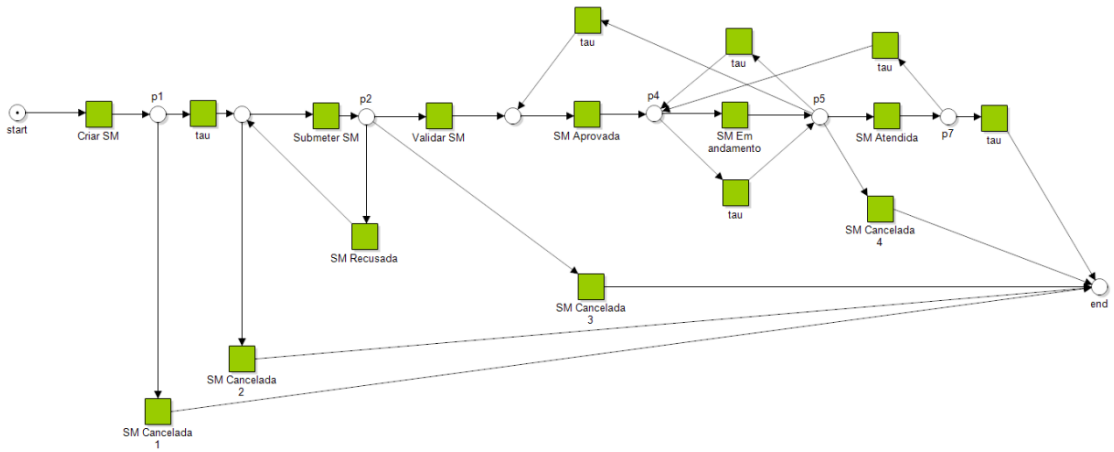



Figura 8 – Modelo elaborado no YASPER para medir tempo de cancelamentos em cada etapa do processo

Para a análise dos tempos e frequência de cancelamentos, foi utilizado novamente o *plug-in* "*Replay a Log on Petri Net for Performance/Conformance Analysis"*, que possui como opção, além do cálculo de adequação e análise de desvios, uma ferramenta que apresenta os tempos envolvidos nos fluxos do processo. Os resultados da análise podem ser observados na Tabela 4,

onde as frequências exibidas se referem ao número de casos e todos os casos nela presentes se referem a SMs que foram canceladas em diferentes pontos do fluxo de processo.

Tabela 4 – Tempos de cancelamento conforme o momento de cancelamento da SM

| Cancelamentos | Fluxo correspondente | Frequência (%) | Média | Mediana | Tempo mínimo | Tempo Máximo |
|---|---|---|---|---|---|---|
| SM Cancelada 1 | Criar SM → SM Cancelada | 11 (*12,64%*) | 49 dias | 3,1 dias | 2 min | 163 dias |
| SM Cancelada 2 | Criar SM →Submeter SM → SM Recusada → SM Cancelada | 1 (*1,15%*) | 22,9 horas | 22,9 horas | 22,9 horas | 22,9 horas |
| SM Cancelada 3 | Criar SM →Submeter SM → SM Cancelada | 6 (*6,90%*) | 12,5 dias | 12,2 horas | 1 min | 37 dias |
| SM Cancelada 4 | Criar SM →…→SM Aprovada→…→SM Cancelada | 69 (*79,31%*) | 7,3 dias | 31 min | 0 | 125 dias |

Os resultados mostram que cerca de 80% dos cancelamentos ocorreram quando a SM se encontrava em um estágio avançado, após passar por diversas etapas e consumido diversos recursos. Ademais, observou-se que as SMs canceladas logo após serem criadas possuem uma duração média muito superior às demais. Investigando a causa deste fato, dentre as 11 ocorrências em "SM Cancelada 1", foi observado um caso com duração de 163 dias, assim como outros 2 que duraram 118 dias, casos estes que deixaram a média tão alta.

Os resultados deste estudo foram reportados à empresa, dando transparência ao andamento e desempenho dos processos avaliados e permitindo aprofundar o estudo de casos específicos que possam gerar melhorias.

## 6. Conclusão

A aplicação das técnicas de mineração de processos foi capaz de trazer benefícios à empresa em questão. As análises possibilitaram a descoberta de desvios de processo que representavam 37% dos casos, permitindo a elaboração de um modelo realístico dos processos envolvidos na Solicitação de Movimentação, com adequação de 100%, e um melhor conhecimento do sistema logístico da empresa. Além disso, a especificação dos fluxos e duração de atividades permite a previsão, controle e análise de todas as atividades do processo, como é o caso da atividade de cancelamento.

O trabalho realizado tratou de uma análise sob a perspectiva de SM. Cada SM analisada contém diversos objetos chamados de "itens", que, assim como a SM, também contam com uma série de possíveis operações e atividades. Os trabalhos futuros estão sendo realizados sob a perspectiva dos itens, permitindo análises mais detalhadas a respeito do que foi abordado. Além das análises por item, também se propõe análises quanto à sensibilidade dos resultados em relação a períodos maiores ou menores de avaliação.

## Referências

Aalst, W.M.P. (2011). Process Mining – Discovery, Conformance and Enhancement of Business Processes, Springer-Verlag, Berlin.

Aalst, W.M.P. (2016). Process Mining: Data Science in Action. Springer-Verlag, Berlin.

ANP. A Indústria de Petróleo e Gás no Brasil. 2018. Web Page. `http://www.anp.gov.br/images/Palestras/Decio_Oddone_Goldman%20Sachs_Outubro_2018.pdf`. Acessado: 2019-03-01.

Becker, T. e Intoyoad, W. (2017). Context Aware Process Mining in Logistics. *Procedia CIRP*, 63:557-562

Beest, N. R. T. P. e Maruster, L. (2007) A Process Mining Approach to Redesign Business Processes - A Case Study in Gas Industry. In *Ninth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC 2007)*, p. 541-548, Timisoara.

DISCO. 2.0. Fluxicon, 2019. Web Page. https://fluxicon.com/disco/. Acessado: 2019-04-15.

FGV Energia (Rio de Janeiro). FGV Energia avalia cenário atual do setor de petróleo no Brasil. 2016. Web Page. https://portal.fgv.br/noticias/fgv-energia-avalia-cenario-atual-setor-petroleo-brasil. Acessado: 2019-03-02.

He, Z., Wu, Q., Wen, L. e Fu, G. (2019). A Process Mining approach to improve emergency rescue processes of fatal gas explosion accidents in Chinese coal mines. *Elsevier, Safety Science*, 111:154-166.

Mans, R.S., S., Schonenberg, M. H., Song, M., Aalst, W.M.P. e Bakker, P.J.M. (2008) Application of *Process Mining* in Healthcare – A Case Study in a Dutch Hospital. In: Fred A., Filipe J., Gamboa H. (eds) Biomedical Engineering Systems and Technologies. BIOSTEC 2008. Communications in Computer and Information Science, vol 25. Springer, Berlin, Heidelberg

PROM. 6.8. TU/e, 2019. Web Page. http://www.promtools.org/doku.php/. Acessado: 2019-04-15.

PROMINP. Participação do seor de petróleo e gás chega a 13% do PIB brasileiro. 2013. Web page. http://www.promimp.com.br/prominp/pt_br/noticias/participacao-do-setor-de-petroleo-e-gas-chega-a-13-do-pib-brasileiro-3.htm. Acessado: 2019-03-01.

Weerdt J., Schupp, A., Vanderloock, A. e Baesens, B. *Process Mining* for the multi-faceted analysis of business processes - A case study in a financial services organization. Elsevier, Computers in Industry, 64v, 57-67, 12 out. 2013.

YASPER. 2.1b4. TU/e, 2019. Web page. http://www.yasper.org/. Acessado: 2019-04-15.