

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

Julio Schenato Fonini

**ANÁLISE DE DESEMPENHO DO SISTEMA DE ATENDIMENTOS DE
SERVIÇOS EM CONCESSIONÁRIA DE ENERGIA**

Santa Maria, RS
2016

Julio Schenato Fonini

**ANÁLISE DE DESEMPENHO DO SISTEMA DE ATENDIMENTOS DE SERVIÇOS
EM CONCESSIONÁRIA DE ENERGIA**

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do título de **Mestre em Engenharia de Produção.**

Orientador: Prof. Dr. Vinícius Jacques Garcia

Santa Maria, RS
2016

Ficha catalográfica elaborada através do Programa de Geração Automática da Biblioteca Central da UFSM, com os dados fornecidos pelo(a) autor(a).

Schenato Fonini, Julio
ANÁLISE DE DESEMPENHO DO SISTEMA DE ATENDIMENTOS DE
SERVIÇOS EM CONCESSIONÁRIA DE ENERGIA / Julio Schenato
Fonini.- 2016.
95 p.; 30 cm

Orientadora: Vinícius Jacques Garcia
Dissertação (mestrado) - Universidade Federal de Santa
Maria, Centro de Tecnologia, Programa de Pós-Graduação em
Engenharia de Produção, RS, 2016

1. Simulação 2. Gerenciamento de serviços 3. Teoria das
filas I. Jacques Garcia, Vinícius II. Título.

Universidade Federal de Santa Maria – UFSM
Centro de Tecnologia – CT/UFSM
Programa de Pós-Graduação em Engenharia de Produção

A comissão examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**ANÁLISE DE DESEMPENHO DO SISTEMA DE ATENDIMENTOS
DE SERVIÇOS EM CONCESSIONÁRIA DE ENERGIA**

elaborada por
Julio Schenato Fonini

Como requisito parcial para obtenção do grau de
Mestre em Engenharia de Produção

COMISSÃO EXAMINADORA:


Vinicius Jacques Garcia, Dr.
(Presidente/Orientador)


José Federico Vizcaino González, Dr.


Daniel Pinheiro Bernardon, Dr.

Santa Maria, 05 de agosto de 2016.

DEDICATÓRIA

Dedico esse trabalho a minha família, pois sem ela nada seria possível. Aos meus pais, Paulo e Rosângela, por me darem a vida, me educar, amar e proporcionar uma base familiar sólida, onde pude usufruir de uma infância muito feliz.

Aos meus irmãos Jonas e Alice, que fizeram parte de toda essa infância e me ajudaram a ser quem sou hoje.

À minha amada esposa Patricia, que esteve ao meu lado ao longo da minha vida adulta, sendo fundamental na execução desse projeto, mas principalmente, por me presentear com as maiores dádivas que poderia receber: os meus filhos Théo e Malu, que hoje são a razão do meu presente e me dão esperanças para acreditar e tentar fazer um futuro melhor a todos e assim proporcionar a eles uma vida tão próspera e feliz como a vida que as pessoas que citei aqui me proporcionaram.

AGRADECIMENTOS

Ao professor Dr. Vinícius Garcia, pela orientação, conhecimentos transmitidos e principalmente pelo companheirismo e motivação para a conclusão desse projeto.

Ao Marcos Gundel, meu líder na AES Sul, que me proporcionou um ambiente de trabalho desafiador, livre de limites conservadores. Propiciando a minha evolução profissional e pessoal de forma plena.

A UFSM e ao governo, indiferente de partido político, que me disponibilizaram um curso de pós-graduação de boa qualidade, com toda a infraestrutura necessária, de forma gratuita.

RESUMO

ANÁLISE DE DESEMPENHO DO SISTEMA DE ATENDIMENTOS DE SERVIÇOS EM CONCESSIONÁRIA DE ENERGIA

AUTOR: Julio Schenato Fonini
ORIENTADOR: Vinícius Jacques Garcia

Este trabalho apresenta uma metodologia que permite o estudo do comportamento de um ambiente de distribuição de ordens de serviço em empresas de energia elétrica. Os componentes deste ambiente possuem uma diversidade de comportamentos, pois tanto a variação no ingresso de demandas como no comportamento dos recursos cria um ambiente com decisões conflitantes e, principalmente, de difícil análise do resultado da ação a médio e longo prazo, ocasionando uma diminuição do nível de serviço prestado pela empresa. Assim, a proposta do trabalho é modelar o problema estocástico em questão, possibilitando aumentar o conhecimento sobre o problema e a partir disso simular diferentes políticas de gerenciamento de serviços. Considerando este cenário, o problema foi modelado e um software de simulação foi desenvolvido, integrando o comportamento estocástico das variações dos seus componentes com a solução determinística de roteamento já implantado na empresa AES Sul. Após a validação do processo de simulação, quatro cenários foram criados e testados na aplicação de simulação. Estes cenários alteraram a quantidade de equipes disponíveis, níveis de priorização das ordens de serviço e a quantidade de segmentações da área de atendimento. Os resultados obtidos foram satisfatórios tanto com relação à obtenção de conhecimento sobre o comportamento do sistema quanto à formatação mais adequada do ambiente.

Palavras-chave: Simulação. Gerenciamento de serviços. Teoria das filas.

ABSTRACT

PERFORMANCE ANALYSIS OF A SERVICE MANAGEMENT SYSTEM IN POWER UTILITY

AUTHOR: Julio Schenato Fonini
ADVISOR: Vinícius Jacques Garcia

This research presents a methodology that allows the study of the behavior of a distribution environment of work orders in electric utilities. The components in this environment have a diversity of behaviors, which complicates decision-making during the management. Considering this scenario, the problem was modeled and simulation software has been developed by integrating the stochastic behavior of its components with the deterministic routing solution already deployed on AES Sul. After the validation of the simulation process, four scenarios were created and tested in simulation application. These scenarios have changed the amount of available crews, levels of prioritization of work orders and service area segmentation. The results obtained were satisfactory both in relation to obtaining knowledge about the behavior of the system as the most appropriate formatting. The scenario that obtained the best results was completed with a low segmentation of areas and a setting that increased the level of criticality of the orders closer to the due date and who had the most relevant services.

Keywords: Simulation. Service management. Queuing theory.

LISTA DE FIGURAS

Figura 1- Sistema básico de filas.....	16
Figura 2 - Função densidade probabilística da distribuição exponencial.....	23
Figura 3 - Diagrama de taxas para o processo de nascimento e morte.....	24
Figura 4 - Modelo para gerenciamento da capacidade e recurso	30
Figura 5 - Área de concessão.....	37
Figura 6 – Divisão da área de concessão.....	37
Figura 7 – Evolução diária de execução de serviços.....	40
Figura 8 – Evolução ordens de serviços pendentes	41
Figura 9 – Processo de despacho	42
Figura 10 – Planejamento de rotas de serviços para o dia subsequente	42
Figura 11 – Representação geográfica das ordens de serviço	46
Figura 12 – Disponibilidade efetiva versus execução de serviços	48
Figura 13 - Tempo de execução serviço TO562 pela equipe 213	49
Figura 14 - Distribuição de frequência dos tempos de execução da equipe 213.....	50
Figura 15 – Representação do processo	52
Figura 16 – Gerenciamento de serviços	53
Figura 17 – Sistema básico de filas para o problema	54
Figura 18 - Média de atendimentos e desvio padrão	55
Figura 19 - Cobertura da demanda emergência pelo recurso disponível	58
Figura 20 - Ciclo reforçador do nível de priorização	59
Figura 21 - Sequência da rota programada.....	60
Figura 22 – Exemplo das etapas de planejamento e execução de rotas	62
Figura 23 – Fonte de entrada conectada a área de execução	66
Figura 24 - Formação do posto técnico e definição da fila	67
Figura 25 - Relação posto técnico, área de execução, recurso e demanda.	68
Figura 26 - Relação entre objetos e variáveis.....	69
Figura 27 - Fontes geradoras (entrada).....	70
Figura 28 - Atribuição de atendimentos de emergência.....	72
Figura 29 - Fluxograma do programa de simulação.....	74
Figura 30 - Diagrama UML do sistema.....	78
Figura 31 - Histograma dos dados de ingresso de ordens comerciais.....	81
Figura 32 - Histograma distribuição do ingresso de ordens de ligação em quartas-feiras	81
Figura 33 - Histograma da variação de turno	82
Figura 34 - Histograma do tempo de deslocamento das equipes	83
Figura 35 - Histograma com os tempos de serviço da ordem de mudança de local.....	83
Figura 36 - Histograma com os tempos de atendimentos de emergência	84
Figura 37 - Rotas desequilibradas em função da prioridade	89

LISTA DE TABELAS

Tabela 1 - Iterações equação de equilíbrio	25
Tabela 2 – Atributos genéricos das ordens de serviço.....	44
Tabela 3 – Atributos que diferenciam as ordens emergenciais e programáveis.....	44
Tabela 4 – Ordens de serviços	45
Tabela 5 – Variação da capacidade	47
Tabela 6 – Comportamento da disponibilidade da equipe 213	47
Tabela 7 – Atributos das equipes.....	51
Tabela 8 - Grupos de ordens programadas	63
Tabela 9 – Hierarquia de prioridade entre os grupos.....	64
Tabela 10 - Prioridade das ordens de serviço	64
Tabela 11 - Roteamento programado para a equipe 195	65
Tabela 12 - Roteamento programado para equipe 121	65
Tabela 13 - Atributos dos componentes do problema	69
Tabela 14 - Estatísticas relevantes.....	80
Tabela 15 - Tipos de dados.....	85
Tabela 16 - Relação área de execução e posto técnico.....	86
Tabela 17 - Relação equipes e posto técnico	86
Tabela 18 - Matriz de prioridades.....	86
Tabela 19 - Resultado Cenário A	87
Tabela 20 - Relação equipes e posto técnico	88
Tabela 21 - Resultado Cenário B.....	88
Tabela 22 - Matriz de prioridade	90
Tabela 23 - Resultado Cenário C.....	90
Tabela 24 - Relação áreas de execução e posto técnico	91
Tabela 25 - Relação equipes e posto técnico	92
Tabela 26 - Resultado Cenário D	92

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVO	12
1.1.1	Objetivo geral	12
1.1.2	Objetivos específicos	13
1.2	PROPOSTA DE TRABALHO	13
1.3	JUSTIFICATIVA	13
2	REFERENCIAL TEÓRICO	15
2.1	TEORIA DAS FILAS	15
2.1.1	Fonte de entrada	16
2.1.2	Fila	16
2.1.3	Disciplina da fila	17
2.1.4	Mecanismo de atendimento	17
2.1.5	Modelagem de um sistema de filas	18
2.1.6	Relevância da distribuição de probabilidade dos dados	21
2.1.7	Processo de nascimento e morte	23
2.1.8	Modelo de filas de disciplina de prioridades	26
2.2	GERENCIAMENTO DE SERVIÇOS	29
2.2.1	Segmentação da demanda	30
2.2.2	Desenvolver serviços complementares	33
2.2.3	Compartilhamento da capacidade	34
2.2.4	Criação de capacidade ajustável	34
2.2.5	Funcionários multifuncionais	34
2.2.6	Programação de turnos de trabalho	34
3	DESENVOLVIMENTO	36
3.1	MAPEAMENTO DO PROCESSO	36
3.1.1	Área de concessão	36
3.1.2	Processos de despacho	39
3.1.3	Despacho de serviços programáveis	41
3.1.4	Atributos de serviço	43
3.1.5	Capacidade	46
3.1.6	Configuração do problema	51
3.2	CONFIGURAÇÃO DO PROBLEMA COMO SISTEMA DE FILAS	53
3.2.1	Fonte de entrada	55
3.2.2	Fila	56
3.2.3	Instalações de atendimento	56
3.2.4	Disciplina da fila	58
3.2.5	Priorização no despacho programado	60
3.2.6	Definição das prioridades de serviço	63
3.3	DESENVOLVIMENTO DO MODELO DE SIMULAÇÃO	66

3.3.1	Modelagem do problema	66
3.3.2	Desenvolvimento do software de simulação	74
3.3.2.1	<i>Define parâmetros</i>	75
3.3.2.2	<i>Simula ingresso de emergência do dia</i>	75
3.3.2.3	<i>Carrega demandas pendentes do posto</i>	75
3.3.2.4	<i>Carrega equipes com calendário do posto</i>	75
3.3.2.5	<i>Executa algoritmo de despacho</i>	76
3.3.2.6	<i>Simula a disponibilidade da equipe</i>	76
3.3.2.7	<i>Define disponibilidade para rota planejada</i>	76
3.3.2.8	<i>Simula tempos de deslocamento e serviço</i>	76
3.3.2.9	<i>Grava serviços executados</i>	76
3.3.2.10	<i>Simula ingresso de serviços comerciais</i>	77
3.3.2.11	<i>Calcula indicadores</i>	77
3.4	ESTUDO DE CASO	79
3.4.1	Coleta e tratamento dos dados	79
3.4.1.1	<i>Estatísticas de ingresso de ordens comerciais por tipo e dia da semana (λ_0, d)</i>	80
3.4.1.2	<i>Variação de turno das equipes</i>	82
3.4.1.3	<i>Tempo de deslocamento</i>	82
3.4.1.4	<i>Duração do serviço comercial</i>	83
3.4.1.5	<i>Duração do atendimento de emergência</i>	84
3.4.2	Cenários simulados	85
3.4.2.1	<i>Cenário A</i>	85
3.4.2.2	<i>Cenário B</i>	88
3.4.2.3	<i>Cenário C</i>	89
3.4.2.4	<i>Cenário D</i>	91
4	CONCLUSÃO	94
4.1	SUGESTÕES DE TRABALHOS FUTUROS	94
	REFERÊNCIAS	95

1 INTRODUÇÃO

As empresas do setor de distribuição de energia elétrica são afetadas por uma diversidade elevada de problemas, tanto de operação, como planejamento e obras, para suprir os clientes com energia dentro dos níveis de serviço estipulados pela Agência Nacional de Energia Elétrica (ANEEL), que é o órgão regulador. Sendo que, os indicadores mais relevantes se dão em relação às falhas nesse sistema de distribuição, denominadas de faltas e geralmente ocorrem em decorrência de eventos climáticos severos (tempestades), sem descartar sobrecargas em equipamentos e eventos causados por agentes externos (colisão de veículo, entre outros), conforme Zhu (2009).

Para manter o sistema de distribuição em funcionamento, as empresas geralmente possuem um centro de operações, que é incumbido de mitigar as falhas nesse sistema através de um atendimento ágil e apoio técnico às equipes de campo (CURCIC et al. 1996). De modo que as equipes são os recursos que efetuam os serviços em campo, tanto emergenciais como programados com o intuito de eliminar falhas futuras ou restabelecer o fornecimento de energia elétrica quando a falta já ocorreu (CHUN-FENG et al. 2011). Porém, esse sistema não é afetado apenas por falhas no fornecimento de energia elétrica, existindo uma gama de outros serviços comerciais - ligação de um novo cliente, cortes para reforma, troca de medidor - que também são de responsabilidade da empresa. Assim, as equipes de campo são determinadas como multitarefas e executam tanto serviços vinculados a manutenções corretivas e preditivas, como ordens de serviços comerciais, criando um cenário de despacho multiobjetivo, que dificulta a tomada de decisão por parte do centro de operações.

Conforme Perrier et al. (2013a), nas últimas décadas existe um avanço da literatura especializada em abordar temas relacionados à problemas de resposta emergencial em sistemas de distribuição e conseqüentemente problemas associados à tomada de decisão. Nesse contexto, as equipes de campo são contratadas para atender a essas demandas com o objetivo de manter a qualidade do atendimento dentro de um nível de serviço especificado. Porém, em muitos momentos o sistema de atendimento a demandas não está equilibrado, e conseqüentemente nessas ocasiões existe uma quantidade muito maior de demandas do que a capacidade de execução, agravando a qualidade do serviço prestado. Por outro lado, em períodos de baixa demanda pode ocorrer ociosidade de equipes. Para essa situação, faz-se necessário a utilização de técnicas de apoio à tomada de decisão, que possibilitem minimizar o impacto negativo gerado com a conseqüente extrapolação do nível de serviço ou contratação de recurso em quantidade superdimensionada, onerando a empresa financeiramente.

A utilização de técnicas de pesquisa operacional pode ser utilizada nesse sentido, já que a solução para problemas relacionados ao tema necessita de extrema sofisticação (PERRIER et al. 2013a e PERRIER et al. 2013b). Assim, em um ambiente de gerenciamento de capacidade e demanda uma série de abordagens poderiam ser utilizadas, como bem destacado por Fitzsimmons, J. e Fitzsimmons, M. (2010), que apresentam, basicamente, um modelo que divide as ações em duas estratégias básicas: controlar a demanda ou a capacidade, onde a partir da análise das particularidades de cada problema define-se as ações específicas de cada estratégia.

No problema em questão, dois fatores podem causar uma diminuição do nível de serviço prestado por uma empresa de distribuição de energia elétrica, são eles: elevação na quantidade de ingresso de serviços, tanto atendimentos de falhas (emergenciais) como comerciais (programáveis), ou a ocorrência de indisponibilidade de equipes, reduzindo a capacidade de atendimento. Esse aumento no ingresso de serviços ou a redução da capacidade ocasiona uma elevação da quantidade de serviços pendentes e conseqüentemente aumentando a relação da demanda sobre a capacidade, dificultando a gestão sobre o nível de serviço. Porém, mesmo existindo uma elevação da demanda ou a redução da capacidade, o principal fator que interfere na tomada de decisão é a oscilação entre elevação e redução de ambos, sendo esse um problema estocástico, de difícil análise sem ferramentas computacionais devidamente modeladas.

Para entender o modo de funcionamento desse ambiente em questão, esse trabalho propõe a modelagem do processo de despacho de serviços como um sistema de filas, baseado na teoria das filas, conforme definido por Hillier e Lieberman (2013) e a partir desse modelo definido, simular as estratégias de gerenciamento de serviços conforme proposto por Fitzsimmons, J. e Fitzsimmons, M. (2010).

Portanto, o trabalho aborda o gerenciamento de serviços em uma empresa de distribuição de energia elétrica, buscando criar uma metodologia que possibilite o estudo desse comportamento e assim melhorar o nível de serviço e a utilização dos recursos.

1.1 OBJETIVO

1.1.1 Objetivo geral

Desenvolver metodologia para simular o processo de gerenciamento de serviços em uma empresa de distribuição de energia elétrica, desenvolvendo uma plataforma que facilite o

estudo do tema e assim melhorar os níveis de serviços utilizando os recursos de maneira adequada.

1.1.2 Objetivos específicos

- a) Modelar o problema de despacho como um sistema de fila;
- b) Desenvolver um protótipo de software de simulação para testar o modelo e analisar os resultados;
- c) Desenvolver indicadores de monitoramento de processo;
- d) Discutir estratégias diferentes para interferir nos indicadores de processo.

1.2 PROPOSTA DE TRABALHO

O método de pesquisa seguido neste trabalho é do tipo estudo de caso, devido esse ser baseado em uma análise de um contexto específico do segmento de distribuição de energia elétrica. Onde objetiva-se a modelagem do processo de despacho de serviços.

A pesquisa tem um caráter exploratório e conceitos tanto de análise qualitativa, por meio de análise de percepções e opiniões dos profissionais envolvidos no processo em questão, e quantitativa, devido à necessidade de tratar de forma adequada as informações e comportamento do processo a partir de dados e fatos devidamente registrados no banco de dados da empresa.

A pesquisa se dará basicamente nas seguintes etapas:

- Definição explícita do problema;
- Mapeamento do processo em questão;
- Desenvolvimento do modelo;
- Desenvolvimento de protótipo de software
- Validação do modelo;
- Análise do comportamento do modelo.

1.3 JUSTIFICATIVA

As empresas de distribuição de energia elétrica possuem a necessidade de atender de forma adequada os seus clientes. Para que isso ocorra, uma série de decisões são tomadas para

definir e adequar uma capacidade de recursos suficientes para atender ao nível de serviço esperado. Porém, esse ambiente possui uma interferência de diversos fatores externos que ocasionam uma variabilidade elevada no processo, dificultando a definição e adequação dos recursos a fim de atender as demandas. Como exemplo, pode-se citar a interferência do clima no ingresso de atendimentos emergenciais, a variação do trânsito que interfere no deslocamento das equipes, a variação das tarefas necessárias em cada local de trabalho, o desempenho individual dos eletricitistas, entre outros.

Essas variações dificultam o entendimento e interferem nas percepções dos analistas de processo, dificultando a tomada de decisão e assim aumentando as chances de que as escolhas feitas não surtam o resultado esperado.

Devido a isso, faz-se necessário o estudo dessa problemática, com a proposta de desenvolver/aplicar métodos mais adequados para análise e simulação da variabilidade do processo, possibilitando aumentar a eficácia nas tomadas de decisão.

2 REFERENCIAL TEÓRICO

Partindo da proposta do trabalho, que visa desenvolver um software de simulação para estudar o comportamento de um processo de geração, despacho e execução de serviços em empresas de distribuição de energia elétrica. A busca por referencial teórico será direcionada para as áreas de teoria das filas e gerenciamento de serviços. Deste modo, a organização deste capítulo inicia com tópicos de teoria das filas e é concluído com gerenciamento de serviços.

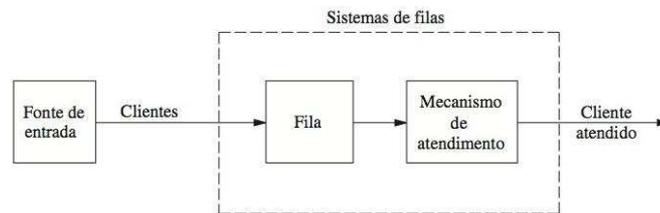
2.1 TEORIA DAS FILAS

A espera é um fenômeno que ocorre com elevada frequência no cotidiano das pessoas e processos, o que, na maioria das vezes, ocasiona um desagrado por quem está à espera da prestação do serviço. Entretanto, podem causar diversos outros impactos negativos ao sistema no qual está inserido: desde a desistência do cliente, como os custos vinculados ao ambiente destinado à espera do cliente ou os custos do próprio cliente parado sem gerar algum benefício esperado, como uma máquina parada à espera da manutenção corretiva gera congestionamento a jusante no processo produtivo, impossibilitando a produção efetiva do processo.

Hillier e Lieberman (2013) definem a teoria das filas como o estudo dessas esperas em suas mais diversas formas, buscando caracterizar cada sistema de fila por modelos no qual suas equações determinem o seu modo de funcionamento, propiciando a análise e simulação do comportamento desse sistema ao longo do tempo e conseqüentemente possibilitando ao analista definir meios mais eficientes para operar um sistema congestionado. Sendo que, esse modo mais eficiente de operar o sistema visa equilibrar o custo de serviço (operação) e o tempo de espera.

A composição mais básica de um modelo de fila possui estrutura conforme a Figura 1. Onde uma fonte de entradas gera clientes ao longo do tempo, introduzindo-os diretamente no sistema de filas, que, por sua vez, é dividido entre a fila e o mecanismo de atendimento, sendo que para o cliente receber atendimento, primeiro deverá aguardar ou não na fila, até que algum tipo de regra, disciplina da fila, autorize o início do atendimento. Após o atendimento, o cliente deixa o sistema de filas finalizando o processo (HILLIER e LIEBERMAN, 2013).

Figura 1- Sistema básico de filas



Fonte: (HILLIER e LIEBERMAN, 2013).

2.1.1 Fonte de entrada

Conforme Hillier e Lieberman (2013), a fonte de entrada, também chamada de população solicitante, possui duas características importantes:

- a) Tamanho: número de clientes distintos que podem requisitar um atendimento ao sistema. O tamanho pode ser definido como finito ou infinito, onde, pela facilidade dos cálculos geralmente utiliza-se o tamanho infinito, porém quando a população é relativamente pequena e a quantidade de clientes já concentrados dentro do sistema de filas começa a impactar na quantidade de população restante fora do sistema (possíveis clientes), o modelo de tamanho finito deve ser utilizado;
- b) Padrão estatístico para a geração de clientes: geralmente, o número de clientes gerados até um determinado momento possui uma distribuição de Poisson e o tempo entre cada cliente gerado segue uma distribuição exponencial, sendo que o tempo entre cada cliente gerado é conhecido como tempo entre chegadas. Sendo representado por λ_n (taxa média de chegada de novos clientes quando n clientes se encontram no sistema). Porém, se λ_n é constante para todo n então poderá ser representado somente por λ . Pois, deste modo, a taxa média de chegada não é influenciada por n , ou seja, a quantidade de clientes sendo atendidos e em atendimento, não influencia na decisão de um novo cliente a entrar na fila.

2.1.2 Fila

A fila é o local onde os clientes aguardam o atendimento, de modo que a capacidade máxima de clientes aguardando na fila possui um padrão finito ou infinito, assim existirá ou não um limite superior ao tamanho da fila. Porém, do mesmo modo que o processo ‘fonte de

entrada', o modelo infinito possui maior facilidade de computar e mesmo em situações em que a fila é finita, mas de um tamanho relativamente grande, o modelo infinito é utilizado. Porém, se no sistema em estudo o limite máximo de clientes é atingido com frequência, o modo finito deve ser empregado (HILLIER e LIEBERMAN, 2013).

2.1.3 Disciplina da fila

A disciplina da fila é a regra utilizada para determinar qual cliente será selecionado da fila para iniciar o atendimento. A regra mais conhecida e utilizada é a FIFO, onde o primeiro a entrar na fila será o primeiro a sair da fila, porém existem diversas outras regras possíveis, como: o último a chegar é o primeiro a sair (produtos empilhados), prioridades entre classes (fila de hospitais, onde o estado de saúde do cliente determina a sua classe e consequentemente uma prioridade), entre outras (ANDRADE, 2009).

2.1.4 Mecanismo de atendimento

Hillier e Lieberman (2013) dividem um mecanismo de atendimento em duas partes, formadas por:

- a) Instalações de atendimento: Corresponde a uma atividade do atendimento, onde esta pode, ou não, atender completamente a necessidade do cliente. Caso mais atividades sejam necessárias, outras instalações de atendimento em série deverão existir para atendê-lo. Cada Instalação de atendimento é formada por um ou mais canais de atendimento paralelo, também chamados atendentes.
- b) Canais de atendimento paralelo (Atendentes): Ao cliente ingressar na instalação de atendimento, ele é atendido por somente um atendente, que executa por completo a atividade relativa à instalação em que se encontra.

Consequentemente, a modelagem de um mecanismo de atendimento, passa pela análise da quantidade de instalações de atendimento, determinação da inter-relação entre as instalações de atendimento, ligações em série, paralelo e a sequência entre elas, quantidade de atendentes dentro de cada instalação de atendimento, tempo médio de atendimento ($1/\mu$) - definido como o tempo entre o início e fim de atendimento em cada instalação de atendimento -, distribuição de probabilidade com que μ ocorre. Hillier e Lieberman (2013) salientam que a distribuição que se supõe com maior frequência na prática é a distribuição exponencial.

2.1.5 Modelagem de um sistema de filas

A modelagem de um sistema de filas é baseada na análise de inter-relações dos componentes citados anteriormente e da coleta dos tempos de chegada e de atendimento dos atendentes. Sendo assim, uma análise estatística dos dados coletados é de extrema relevância para caracterizar uma distribuição de probabilidade, e assim garantir que o processo está em um ‘estado estacionário’ (ou estável). Ou seja, o comportamento dos dados se repete com um mesmo padrão ao longo do tempo (ANDRADE, 2009). Porém, Hillier e Lieberman (2013) explicam que os sistemas de fila, logo após começarem a operar, se encontram em condição transiente (não estacionário), devido a sofrerem um grande impacto da situação inicial do sistema e o tempo decorrido, entretanto conforme passa o tempo o sistema atinge o estado estacionário, facilitando assim a modelagem do problema.

Para caracterizar as variáveis relevantes ao processo de modelagem, pode-se adotar o seguinte padrão de terminologia e notação, conforme Hillier e Lieberman (2013):

- a) Estado do sistema: número de clientes no sistema;
- b) Comprimento da fila: número de clientes que aguardam o início do atendimento;
- c) N_t : número de clientes no sistema de fila no instante t ($t > 0$);
- d) $P_{n(t)}$: probabilidade de exatamente n clientes no sistema de filas no instante t , dado o número no instante 0 ;
- e) s : número de atendentes;
- f) λ_n : taxa média de chegada de clientes quando n clientes estão no sistema;
- g) μ_n : taxa média de atendimento para o sistema global quando n clientes estão no sistema, ou seja, quantidade média de clientes que encerra o atendimento por unidade de tempo.

Entretanto, quando:

A taxa média de chegada de clientes (λ_n) for constante para todo n , essa constante pode ser representada por λ ;

A taxa média de atendimento (μ_n) por atendente ocupado for constante para $n \geq 1$, então essa constante pode ser representada por μ . Neste caso, $\mu_n = s\mu$ quando $n \geq s$.

Assim, o fator de utilização ρ pode ser caracterizado por: $\rho = \frac{\lambda}{(s \cdot \mu)}$.

Para Andrade (2009), criar medidas de efetividade do sistema é extremamente relevante, citando as seguintes medidas como elementares nesse objetivo:

- a) O percentual de tempo em que o posto de atendimento permanece ocioso ou ocupado;
- b) O tempo médio que cada cliente aguarda na fila;
- c) Tempo médio que o cliente fica dentro do sistema de filas, ou seja, o tempo médio entre a chegada do cliente na fila e o final do atendimento;
- d) Número médio de clientes na fila por unidade de tempo (tamanho médio da fila);
- e) Número médio de clientes dentro do sistema de fila.

Cabendo ao analista identificar quais os melhores indicadores para monitorar o sistema em estudo, possibilitando assim agir de maneira mais eficaz no sistema. Mas, Hillier e Lieberman (2013) apresentam praticamente os mesmos indicadores, conforme a seguir:

P_n : probabilidade de n clientes estarem no sistema de filas;

L : número de clientes esperado no sistema de filas:

$$L = \sum_{n=0}^{\infty} n \cdot P_n \quad (1)$$

L_q : comprimento esperado da fila:

$$L_q = \sum_{n=s}^{\infty} (n - s) \cdot P_n \quad (2)$$

w : tempo de espera no sistema para cada cliente individual:

$$W = E(w) \quad (3)$$

w_q : tempo de espera na fila para cada cliente individual:

$$W_q = E(w_q) \quad (4)$$

E assim possibilitar a relação entre L , W , L_q , W_q :

$$L = \lambda \cdot W \quad (5)$$

$$L_q = \lambda \cdot W_q \quad (6)$$

$$W = W_q + \frac{1}{\mu} \quad (7)$$

Quando:

- λ_n constante para todo n , caso contrário: pode ser substituído por $w \cdot \bar{\lambda}$;

- $\frac{1}{\mu}$ constante para todo $n \geq 1$.

Onde:

W : tempo médio de espera no sistema;

W_q : tempo médio de espera na fila.

Sendo que, as notações indicadas acima se supõem que o sistema se encontre em uma condição de estado estável.

Hillier e Lieberman (2013) salientam que essas relações são muito importantes, pois permitem encontrar os valores das quatro variáveis fundamentais (L , W , L_q , W_q) a partir de apenas uma. Sendo que, em determinados casos, algumas variáveis são mais fáceis de serem calculadas do que outras.

Para exemplificar uma modelagem de sistema de fila básico, Andrade (2009) apresenta um modelo elementar com um atendente e fila única com população infinita.

As chegadas apresentam um comportamento de distribuição de Poisson com média λ chegadas/tempo, os tempos de atendimento seguem a distribuição exponencial negativa com média $1/\mu$, em consequência de o número de atendimentos possuírem um comportamento com distribuição de Poisson com média μ , a disciplina da fila é do modo FIFO, o tamanho da fonte de entrada é suficientemente grande para ser caracterizada como infinita, definindo o problema como M/M/1.

Equações básicas do modelo (ANDRADE, 2009 apud WAGNER, 1972):

a) Probabilidade de existir n clientes no sistema:

$$P_n = \left(\frac{\lambda}{\mu}\right)^n \cdot \left(\frac{\mu - \lambda}{\mu}\right) \quad (8)$$

b) Probabilidade de que o número de clientes no sistema (n) seja maior que uma quantidade r :

$$P_{(n>r)} = \sum_{r+1}^{\infty} \left(\frac{\lambda}{\mu}\right)^n \cdot \left(\frac{\mu - \lambda}{\mu}\right) = \left(\frac{\lambda}{\mu}\right)^{r+1} \quad (9)$$

c) Probabilidade de que o sistema esteja ocioso (taxa de ociosidade):

$$P_{(n=0)} = \left(\frac{\lambda}{\mu}\right)^0 \cdot \left(\frac{\mu - \lambda}{\mu}\right) = \left(\frac{\mu - \lambda}{\mu}\right) \quad (10)$$

d) Probabilidade de que o sistema esteja ocupado (taxa de ocupação):

$$P_{(n>0)} = \frac{\lambda}{\mu} = \rho \quad (11)$$

e) Número médio de clientes no sistema (L):

$$L = \frac{\lambda}{\mu - \lambda} \quad (12)$$

f) Número médio de clientes na fila (L_q):

Incluindo as filas de tamanho zero:

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \quad (13)$$

Para filas > 0 :

$$L_{q(fila>0)} = \frac{\lambda}{\mu - \lambda} \quad (14)$$

g) Tempo médio de espera na fila por cliente (W_q):

$$W_q = \frac{1}{\mu(\mu - \lambda)} \quad (15)$$

h) Tempo médio gasto no sistema por cliente (W):

$$W = \frac{1}{\mu - \lambda} \quad (16)$$

2.1.6 Relevância da distribuição de probabilidade dos dados

O comportamento do sistema de filas é determinado, em grande parte, por duas propriedades estatísticas: tempos entre chegadas e tempos de atendimento, de onde se derivam as taxas λ e μ . Deste modo, a distribuição de probabilidade dessas duas variáveis é de extrema relevância ao processo de modelagem, onde essas podem assumir diversas formas, desde que os tempos não assumam valores negativos. Porém, além da forma dever representar o comportamento da variável, ela deve ser relativamente simples possibilitando o seu tratamento matemático. Partindo disso, Hillier e Lieberman (2013) explicam que a distribuição exponencial é a forma mais adequada para representar um sistema de filas, pois

além de um tratamento simples, ela representa a maioria dos sistemas desse tipo, devendo ser utilizada sempre que possível. Essa adequação é adequada a partir das seguintes análises:

Sendo T uma variável aleatória que represente o tempo entre chegadas, pode-se atribuir a ela uma distribuição exponencial com parâmetro α se sua função de densidade probabilística for:

$$f_T(t) = \begin{cases} \alpha \cdot e^{-\alpha t}, & t \geq 0 \\ 0, & t < 0, \end{cases} \quad (17)$$

Neste caso, as probabilidades acumuladas:

$$P\{T \leq t\} = 1 - e^{-\alpha t}; \quad (t \geq 0) \quad (18)$$

$$P\{T > t\} = e^{-\alpha t}; \quad (t \geq 0) \quad (19)$$

O valor esperado:

$$E(T) = \frac{1}{\alpha} \quad (20)$$

A variância de T:

$$\text{var}(T) = \frac{1}{\alpha^2} \quad (21)$$

Porém, para determinar que T possui uma distribuição exponencial no modelo de filas, Hillier e Lieberman (2013) analisam o comportamento da variável T em relação a seis propriedades fundamentais da distribuição exponencial:

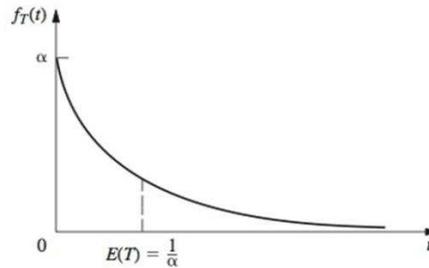
a) Propriedade 1: $f_T(t)$ é uma função estritamente decrescente de t ($t \geq 0$).

Assim:

$$P\{0 \leq T \leq \Delta t\} > P\{t \leq T \leq t + \Delta t\}; \quad (\Delta t, t > 0) \quad (22)$$

Deste modo, qualquer Δt posicionado à esquerda na reta t , tem maior probabilidade de ocorrer que o mesmo Δt posicionado em um ponto mais à direita da mesma reta, conseqüentemente o valor esperado tem uma menor probabilidade de ocorrer do que valores à sua esquerda na reta t . Hillier e Lieberman (2013) explicam, que do mesmo modo como em grande parte dos sistemas de fila, os ingressos de novos clientes ao sistema, apresentam um comportamento que intercalam momentos com agrupamento de chegadas seguido de um período sem chegadas, e que esta composição tem um caráter extremamente aleatório.

Figura 2 - Função densidade probabilística da distribuição exponencial



Fonte: (HILLIER e LIEBERMAN, 2013).

b) Propriedade 2: Ausência de memória

Onde a ocorrência de T não está relacionada com as ocorrências anteriores de T , e é representada matematicamente da seguinte maneira:

$$P\{T > t + \Delta t | T > \Delta t\} = P\{T > t\} \quad (23)$$

- c) Propriedade 3: o mínimo de diversas variáveis aleatórias exponenciais tem uma distribuição exponencial;
- d) Propriedade 4: a relação entre a distribuição exponencial e a distribuição de Poisson;
- e) Propriedade 5: para todo $t > 0$: $P\{T \leq t + \Delta t | T > t\} \approx \alpha \Delta t$, para Δt pequeno;
- f) Propriedade 6: não é afetada por agregação ou desagregação.

2.1.7 Processo de nascimento e morte

Conforme Hillier e Lieberman (2013) os modelos de filas mais básicos relacionam o processo de clientes ingressando ao sistema de filas e concluindo o atendimento com o processo de nascimento e morte. Onde, o termo nascimento caracteriza a chegada de um cliente ao sistema de filas e o termo morte caracteriza o término do atendimento ao cliente e conseqüentemente a sua saída do sistema. Deste modo, o processo de nascimento e morte é quem determina o estado do sistema, como $N_{(t)}$ é o número de clientes no sistema de filas no momento t (para $t > 0$), logo o processo de nascimento e morte descreve probabilisticamente como $N_{(t)}$ se comporta com o passar do tempo.

O processo de nascimento e morte é caracterizado pelas seguintes hipóteses:

- Hipótese 1: Dado $N_{(t)} = n$, a distribuição de probabilidade é exponencial com parâmetro λ_n ($n = 0, 1, 2, \dots$), para o tempo que falta entre o momento atual e a próxima ocorrência de nascimento;
- Hipótese 2: Dado $N_{(t)} = n$, a distribuição de probabilidade é exponencial com parâmetro μ_n ($n = 0, 1, 2, \dots$), para o tempo que falta entre o momento atual e a próxima ocorrência de morte;
- Hipótese 3: As variáveis aleatórias λ_n e μ_n são independentes entre si.

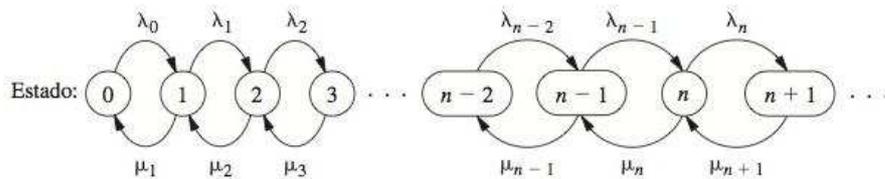
Isso significa dizer que o processo de nascimento e morte, depende somente do estado do sistema e que as transições entre estados se dá:

Nascimento: $n \rightarrow n + 1$;

Morte: $n \rightarrow n - 1$;

No diagrama da Figura 3 as setas indicam os movimentos válidos para a alteração do estado do sistema, identificando a relação encadeada entre os estados e que para chegar a um determinado estado deve obrigatoriamente seguir uma combinação de sequências de nascimentos e mortes que levem ao estado esperado:

Figura 3 - Diagrama de taxas para o processo de nascimento e morte



Fonte: (HILLIER e LIEBERMAN, 2013).

Assim, o estado futuro do sistema depende diretamente das taxas de chegada (nascimentos) e de saída (mortes) e do estado atual do sistema.

Partindo da Figura 3, Hillier e Lieberman (2013) fazem uma análise sobre as probabilidades do sistema se encontrar em algum determinado estado. Onde, considerando o estado 0, o processo somente pode ingressar nesse estado a partir do estado 1. Deste modo, a probabilidade do estado se encontrar em $n = 1$ (P_1) representa a proporção de tempo que o estado tem a possibilidade de ser igual a 0, e caso o estado se encontre em $n > 1$, então a probabilidade é zero de o estado migrar imediatamente para 0.

Como a partir do estado 1, somente com a ocorrência de uma morte o estado será 0, então a taxa média de entrada no estado 0 é igual a μ_1 . Assim:

$$\mu_1 P_1 + 0(1 - P_1) = \mu_1 P_1 \quad (24)$$

Consequentemente, a taxa média de saída do estado 0 deve ser $\lambda_0 P_0$, definindo assim a equação de equilíbrio para o estado 0:

$$\mu_1 P_1 = \lambda_0 P_0 \quad (25)$$

Assim a equação do equilíbrio é definida da seguinte forma:

$$\text{taxa que entra} = \text{taxa que sai}$$

Tabela 1 - Iterações equação de equilíbrio

Estado	taxa que entra = taxa que sai
0	$\mu_1 P_1 = \lambda_0 P_0$
1	$\lambda_0 P_0 + \mu_2 P_2 = (\lambda_1 + \mu_1) P_1$
2	$\lambda_1 P_1 + \mu_3 P_3 = (\lambda_2 + \mu_2) P_2$
⋮	⋮
$n - 1$	$\lambda_{n-2} P_{n-2} + \mu_n P_n = (\lambda_{n-1} + \mu_{n-1}) P_{n-1}$
n	$\lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} = (\lambda_n + \mu_n) P_n$
⋮	⋮

Fonte: (HILLIER e LIEBERMANN, 2013).

De modo que, a probabilidade de $n = 1$ dependente da probabilidade de $n = 2$ ocorrer (P_2). Assim, essa regra se alastra para localizar a probabilidade de ingressar em qualquer estado n , conforme Tabela 1.

Notação simplificada:

$$C_n = \frac{\lambda_{n-1} \lambda_{n-2} \cdots \lambda_0}{\mu_n \mu_{n-1} \cdots \mu_1}, \text{ para } n = 1, 2, \dots, \quad (26)$$

Então ao definir $C_n = 1$ para $n = 0$, define-se consequentemente as probabilidades para o estado estável:

$$P_n = C_n P_0 \text{ para } n = 0, 1, 2, \dots \quad (27)$$

Com a exigência de:

$$\sum_{n=0}^{\infty} P_n = 1 \quad (28)$$

Implica que:

$$\left(\sum_{n=0}^{\infty} C_n \right) P_0 = 1 \quad (29)$$

De modo que:

$$P_0 = \left(\sum_{n=0}^{\infty} C_n \right)^{-1} \quad (30)$$

A partir dessas definições as medidas de desempenho fundamentais do sistema de filas podem ser facilmente calculadas após definir o valor de P_n . Conforme já definido anteriormente:

$$L = \sum_{n=0}^{\infty} n P_n, \quad L_q = \sum_{n=s}^{\infty} (n - s) P_n. \quad (31)$$

$$W = \frac{L}{\bar{\lambda}}, \quad W_q = \frac{L_q}{\bar{\lambda}}, \quad (32)$$

Onde $\bar{\lambda}$ é a taxa média de chegadas em longo prazo, então:

$$\bar{\lambda} = \sum_{n=0}^{\infty} \lambda_n P_n \quad (33)$$

2.1.8 Modelo de filas de disciplina de prioridades

Esse tipo de modelo é utilizado quando os clientes que ingressam no sistema possuem criticidades diferentes de atendimento entre si, diferindo-se por algum tipo de característica. Nessa situação o sistema deve possuir uma disciplina de fila que organize a fila em vistas a atender os clientes mais prioritários com um tempo de espera mais reduzido do que clientes

com menos prioridade. Hillier e Lieberman (2013) citam como exemplo o caso de um hospital aonde os pacientes chegam ao sistema de filas e possuem diferentes sintomas, sendo que os pacientes em estado mais grave devem ser atendidos primeiro.

Hillier e Lieberman (2013) apresentam dois modelos de disciplina de prioridade, que partem das mesmas premissas descritas nos modelos elementares anteriores: Pressupõe-se da existência de um processo de entrada de Poisson, tempos de atendimentos exponenciais, estado estável do sistema e tempos de atendimentos esperados iguais entre as prioridades (classes). Porém, se distinguem entre possuir um caráter de prioridades não preemptivas, onde um cliente que está sendo atendido deverá ter seu atendimento concluído para então ser priorizado um caso de prioridade elevada, e preemptivas, onde ao ingressar um cliente de prioridade elevada paralisará um atendimento de prioridade inferior para iniciar o atendimento.

Para os dois modelos, as fórmulas das variáveis de desempenho do sistema permanecem as mesmas (L , L_q , W e W_q), o que muda é a distribuição de tempos de espera entre as prioridades aumentando em muito a variância entre as classes, pois as classes prioritárias tendem a permanecer muito menos tempo aguardando, devido às regras de priorização.

a) Modelo para prioridades não preemptivas:

$$W_k = \frac{1}{AB_{k-1}B_k} + \frac{1}{\mu}, \text{ para } k = 1, 2, \dots, N, \quad (34)$$

Onde:

$$A = s! \frac{s\mu - \lambda}{r^2} \sum_{j=0}^{s-1} \frac{r^j}{j!} + s\mu \text{ para } s > 1 \quad (35)$$

$$A = \frac{\mu^2}{\lambda} \text{ para } s = 1 \quad (36)$$

$$B_0 = 1 \quad (37)$$

$$B_k = 1 - \frac{\sum_{i=1}^k \lambda_i}{s\mu}, \quad (38)$$

$$r = \frac{\lambda}{\mu}, \quad (39)$$

s : número de atendentes;

μ : taxa média de atendimento por atendente ocupado;

λ_i : taxa média de chegada por classe de prioridades i .

$$\lambda = \sum_{i=1}^N \lambda_i \quad (40)$$

Para garantir a possibilidade da classe de prioridade k entrar em estado estável:

$$\sum_{i=1}^k \lambda_i < s\mu \quad (41)$$

O número esperado de clientes da prioridade k no sistema de filas:

$$L_k = \lambda_k W_k \quad \text{para } k = 1, 2, \dots, N \quad (42)$$

O tempo espera previsto na fila para clientes da prioridade k :

$$Wq_k = W_k - \frac{1}{\mu} \quad (43)$$

Tamanho esperado da fila de clientes da prioridade k :

$$Lq_k = \lambda_k \cdot Wq_k \quad (44)$$

b) Modelo para prioridades preemptivas:

Para esse modelo a única equação modificada é o cálculo de W_k , sendo:

Para $s = 1$:

$$W_k = \frac{1/\mu}{B_{k-1}B_k}, \quad \text{para: } k = 1, 2, \dots, N \quad (45)$$

Sendo que para $s > 1$ efetua-se calculo iterativo:

$$W_k = \sum_{i=k}^1 \left(\frac{\lambda_{i-1}}{\lambda_{i-1} + \lambda_i} W_i \right) \quad (46)$$

Os modelos de filas com prioridades preemptivas e não preemptivas são de extrema relevância nesse trabalho, pois o sistema de despacho em distribuição de energia elétrica apresenta uma gama de serviços diferentes com prioridades diferentes e possivelmente com a

existência de ambos os tipos de priorização, dependendo da relação entre os tipos de serviços e seus prazos de vencimento.

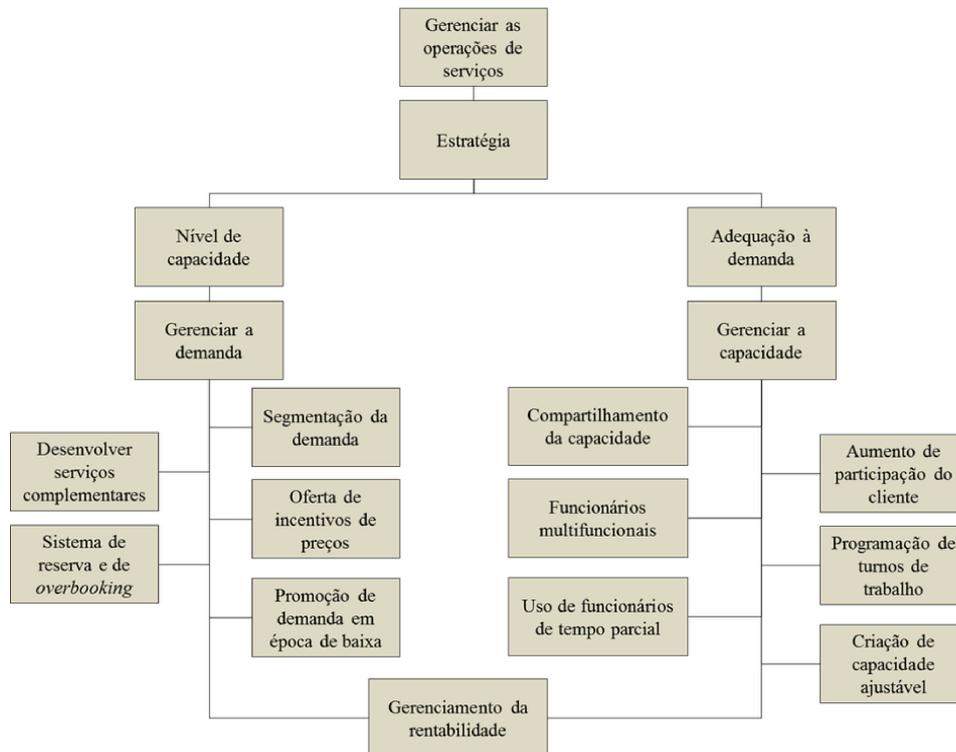
2.2 GERENCIAMENTO DE SERVIÇOS

O gerenciamento de capacidade e demanda é apresentado por Fitzsimmons, J. e Fitzsimmons, M. (2010) como proposta de ações possíveis e efetivas para minimizar o impacto do desequilíbrio entre esses fatores. Basicamente, eles apresentam duas estratégias de gerenciamento de capacidade: nível de capacidade e adequação à demanda. Onde a primeira é orientada principalmente ao *marketing* e a segunda voltada à operação. A proposta é que o analista possa escolher os pontos de maior alavancagem ou de maior facilidade para atingir resultados positivos, sem depender que todo o sistema compartilhe de melhorias. Onde, se a empresa tem menor flexibilidade de alterar o seu modo produtivo possa tomar ações vinculado ao lado externo, gerenciando a demanda. E o contrário também é verdadeiro, se a empresa possui um mercado restrito, os processos internos podem ser modificados para maximizar a lucratividade. Claro que a utilização de ações em ambos os lados devem ser analisadas e aplicadas conforme a possibilidade. A Figura 4 apresenta a estrutura de gerenciamento de serviços proposta por Fitzsimmons, J. e Fitzsimmons, M. (2010).

Esta define as seguintes ações para cada modo de gestão:

- a) Gerenciar a demanda:
 - Segmentação da demanda;
 - Desenvolver serviços complementares;
 - Oferta de incentivos de preços;
 - Sistema de reserva e de overbooking;
 - Promoção de demanda em época de baixa.
- b) Gerenciar a capacidade:
 - Compartilhamento da capacidade;
 - Aumento da participação do cliente;
 - Funcionários multifuncionais;
 - Programação de turnos de trabalho;
 - Uso de funcionários de tempo parcial;
 - Criação de capacidade ajustável.

Figura 4 - Modelo para gerenciamento da capacidade e recurso



Fonte: (FITZSIMMONS, J. e FITZSIMMONS, M., 2010).

Partindo da estrutura de gerenciamento de serviços, iremos descrever os tópicos mais relacionados ao problema em questão.

2.2.1 Segmentação da demanda

Basicamente, essa estratégia é apresentada como um método para identificar padrões de comportamento de grupos de clientes e assim identificar oportunidades para encaixar os diferentes comportamentos gerando demandas de forma equilibrada, conforme o recurso está disponível. Assim, evita-se o desenvolvimento de expectativa do cliente, onde o prazo é respeitado, pois a demanda é gerada no momento adequado ao recurso disponível, e a empresa define a hierarquia sobre a priorização dos serviços a partir do retorno que cada um deles fornece (FITZSIMMONS, J. e FITZSIMMONS, M., 2010).

O problema de segmentação da demanda é um dos pontos relevantes do problema em questão, pois o desenvolvimento de hierarquia na escolha da ordem para ser executada a partir de atributos dessa demanda, ou cliente, torna esse sistema mais aderente à importância e relevância de cada cliente ou necessidade desses. Porém, por outro lado, afeta a produtividade

dos recursos devido à existência de deslocamentos entre os serviços, que foram sequenciados sem levar em consideração a sua posição geográfica.

Yan et al. (2010) também apresenta um estudo para lojistas on-line, onde ele aplica simulação computacional para comparar o comportamento de seis políticas de gerenciamento de filas, com vistas a atender dois indicadores de variação de backlog: BV (workload) que é calculado pela divisão do coeficiente de variação da quantidade de ordens pendentes no período pelo coeficiente de variação do ingresso de ordens no período, e BV (time) que é calculado pela divisão do coeficiente de variação do tempo médio de duração dos pedidos e o coeficiente de variação do tempo médio entre ingressos de pedidos. Conforme equação apresentada por Yan et al. (2010):

$$BV(workload)_t = \frac{sd(x_t, x_{t-1} \dots, x_1) / mean(x_t, x_{t-1} \dots, x_1)}{sd(y_t, y_{t-1} \dots, y_1) / mean(y_t, y_{t-1} \dots, y_1)} \quad (47)$$

Onde:

x_t : Quantidade de ordens pendentes no dia t ;

y_t : Quantidade de ordens ingressadas no dia t ;

sd : desvio padrão;

$mean$: média aritmética.

$$BV(time)_t = \frac{c.o.v \left(\left(\sum_{i=1}^{x_t} \frac{T_i}{x_t} \right), \dots, \left(\sum_{i=1}^{x_1} \frac{T_i}{x_1} \right) \right)}{c.o.v \left(\left(\sum_{i=1}^{y_t} \frac{T_i}{y_t} \right), \dots, \left(\sum_{i=1}^{y_1} \frac{T_i}{y_1} \right) \right)} \quad (48)$$

Onde:

T_i : Prazo da ordem i .

Sendo que a avaliação desses indicadores se dá em relação às seguintes políticas de gerenciamento de fila:

- a) Política FIFO: primeiro pedido a chegar será o primeiro a ser atendido;
- b) Política de Lead Time: postergação de ordens, seguida de uma priorização das ordens com menor tempo até o vencimento;
- c) Política de liberação na data de vencimento: como o próprio nome diz, o pedido somente é enviado no último dia de prazo, aumentando o backlog;

- d) Política de limite de inventário: a liberação de pedidos ocorre conforme a FIFO, porém até um limite mínimo de inventário, ao inventário retornar ao valor padrão, inicia-se a liberação de pedidos novamente com o comportamento FIFO;
- e) Política de limite de Inventário-Lead time: os pedidos são recebidos e postergados, seguindo a liberação da sistemática da política de Lead time até um limite mínimo de inventário, ao inventário retornar ao valor padrão inicia-se a liberação conforme política de Lead time;
- f) Política de liberação na data de vencimento com limite de inventário: mistura a política de Liberação na data de vencimento e limite de inventário, onde as ordens são postergadas até o último dia do prazo e então são liberadas priorizando as ordens com menor prazo até o limite mínimo de inventário.

Yan et al. (2010) utiliza um sistema computacional para simular o comportamento das políticas acima em relação aos indicadores de variação de backlog. Sendo que os melhores resultados foram encontrados para a política de liberação na data de vencimento, onde os autores listam os benefícios da postergação como sendo: melhor controle da variação na cadeia de fornecimento, melhor utilização dos recursos de produção e possibilita aumentar receitas aos clientes desejarem pagar por entregas mais rápidas. Porém, aumenta-se o risco de não cumprir os prazos de entrega, sendo esse, um ponto de atenção ao lojista.

Huang et al. (2011) também trabalha em um problema de segregação de clientes em prioridades e propõe um modelo heurístico para solucionar um problema de gerenciamento de demanda e capacidade de lojas on-line, onde os clientes podem aceitar algum tempo de espera devido a facilidade de efetuar a compra e diversidade dos produtos, porém, podem desistir do pedido após certo período em espera posterior ao prazo estabelecido. Para atender esses pedidos, o lojista pode ter produto estocado, aumentando os custos devido ao estoque, ou não, tudo depende do tempo que os clientes aceitam esperar e do comportamento da rede de fornecedores que atendem ao lojista. Outra derivação deste problema é que o lojista, às vezes, pode optar por possuir um fornecedor principal, que atende em um tempo mais longo a um custo mais barato, a outro fornecedor que atende em um menor tempo a um custo mais alto. Logo, a decisão do problema é sobre a escolha do melhor momento para efetuar o pedido de fornecimento e quanto de produto será solicitado, para isso Huang et al. (2011) utiliza duas medidas de custos para onerar o lojista pela minimização do estoque, que são, basicamente, um custo de replanejamento (postergação) dos pedidos e um custo de pedido cancelado, sendo que o primeiro varia conforme o tempo de postergação e o segundo é um preço fixo vinculado ao valor de lucratividade da venda do produto.

Por sua vez, Taylor (1996) apresenta um problema de gestão da manutenção, quando existem diversas ordens de serviços com criticidades diferentes, de modo que o impacto de uma ordem de serviço ultrapassar o seu prazo estipulado é maior que se outras vencerem. Podendo citar como exemplo: máquinas gargalo de uma linha de produção ou máquinas que apresentam um estado crítico e que a quebra está iminente. O autor salienta que indicadores de gerenciamento de backlog devem levar em consideração a diferença entre as ordens, caso contrário ordens irão vencer ou horas extras deverão ser feitas para atender a ordem em curto prazo. O modelo apresentado por Taylor, basicamente, divide as prioridades em três grupos, e garante a execução em sequência de prioridades, garantindo que todas as prioridades 1 serão efetuadas antes das prioridades 2 e todas as prioridades 2 serão efetuadas antes das prioridades 3. Então, aloca os recursos visando minimizar custos com mão de obra e atender as ordens prioritárias.

Wang et al. (2013) também procuram resolver um problema de gerenciamento de backlog com níveis de prioridades diferentes entre pedidos. Onde, apresentam um modelo heurístico para atender dois níveis de prioridades, gerenciando os momentos adequados para solicitar o ingresso de mais pedidos para produção, o momento de reprogramar as ordens e o momento de priorizar (nível de racionamento crítico). Outra problemática citada é o modo como os indicadores de atendimento entre as prioridades se relacionam, no caso em questão o objetivo é possibilitar comparar ordens de uma prioridade em relação ao seu nível de serviço e outras a um custo de penalização.

2.2.2 Desenvolver serviços complementares

Utilizado para gerar receita adicional a partir de recursos ociosos, desenvolvendo serviços complementares ao produto principal, buscando atender outras necessidades dos clientes que já frequentam o estabelecimento. Também utilizado para mitigar impacto nas oscilações, gerando receitas em época de baixa demanda. Essa abordagem é relevante para tratar os períodos no qual a oscilação do ingresso de serviços está no seu nível inferior, onde a demanda é menor do que a capacidade, evitando com que o recurso fique ocioso, utilizando-se dele para execução de serviços complementares que agreguem algum tipo de valor para a empresa (FITZSIMMONS, J. e FITZSIMMONS, M., 2010).

2.2.3 Compartilhamento da capacidade

Consiste em utilizar o recurso de maneira mais sofisticada, não o deixando vinculado a um posto e sim definindo outras tarefas e postos de trabalho ao recurso em questão.

2.2.4 Criação de capacidade ajustável

Consiste em flexibilizar a utilização do recurso, tentando acioná-lo de maneira mais aderente ao comportamento da demanda. É muito próxima a solução aplicada por Huang et al. (2011), onde ele habilita uma série de fornecedores para atender a demanda, mas inicializa cada um deles em cenários diferentes, alterando os custos conforme a restrição de prazo, onde possui um fornecedor a um custo mais baixo com prazo elevado e que o atende com frequência alta (fornecedor de rotina) e outro fornecedor com custo elevado mas com prazo extremamente curto, possibilitando o acionamento deste quando ocorre uma elevação dos clientes em fila (pedidos), de modo que esse aumento da fila pode afetar a eficácia no cumprimento do prazo limite de entrega.

2.2.5 Funcionários multifuncionais

Está relacionado diretamente com os itens compartilhamento da capacidade e criação de capacidade ajustável, onde a capacidade do funcionário em exercer diversas tarefas, torna o recurso mais dinâmico, possibilitando se encaixar nas tarefas conforme ocorre a oscilação de demanda sobre cada uma.

2.2.6 Programação de turnos de trabalho

Basicamente, essa ação se baseia na análise do comportamento de ingresso como forma de dimensionar os horários de turnos de trabalho dos funcionários, com vistas a distribuí-los de maneira adequada à geração de demandas.

As propostas de Fitzsimmons, J. e Fitzsimmons, M. (2010) estão diretamente conectadas às duas variáveis relevantes no processo: demanda e capacidade, onde a busca por estratégias que manipulem o comportamento dessas variáveis torna a proposta relevante pela amplitude de modos diferentes de alcançar o resultado, que é o equilíbrio entre as duas variáveis.

Deste modo, o trabalho permeia a definição do modelo como um sistema de filas, possibilitando a simulação de cenários diferentes, e o gerenciamento de serviços que devem conduzir os cenários propostos. Identificando meios melhores de gerenciar o nível de serviço em empresas de distribuição de energia elétrica.

3 DESENVOLVIMENTO

Para uma empresa de distribuição de energia elétrica, existe uma gama de serviços a serem prestados ao cliente. Estes serviços são desde a ligação e manutenção do contrato de um cliente, como manutenções preditivas ou corretivas. Para a execução desses, faz-se necessário que a empresa utilize recursos (equipes de campo) com uma capacidade adequada para dar vazão a essa demanda, minimizando a execução desses fora do prazo estipulado ao cliente e assim, evitar as penalizações do órgão regulador e satisfazer a expectativa desse cliente.

Porém, esse processo de atender as demandas de serviço é extremamente complexo, devido à existência de variabilidade tanto no ingresso como na execução dos serviços. Esse problema é aumentado devido ao elevado custo do recurso necessário para atender a essa demanda (equipes), tornando a utilização de superdimensionamento de capacidade inviável para a empresa. Portanto, parte-se de um princípio que ordens ficam estocadas, evitando a ociosidade do recurso. Mas, por outro lado, esse estoque de ordens de serviço, ocasiona uma diminuição no nível de serviço prestado pela empresa, pois quanto mais se postergar a execução do serviço, mais próximo ao fim do prazo previsto ele se encontrará e assim aumentará as chances de este não ser atendido no prazo. Deste modo, torna-se necessário o estudo dos componentes relevantes ao processo e suas variações, e assim propor uma metodologia para equilibrar a relação entre capacidade de execução e demandas de serviço, atendendo ao nível de serviço esperado.

Para abordar o tema, o desenvolvimento do trabalho é estruturado da seguinte forma: na sessão 3.1 apresenta-se o mapeamento detalhado do processo; na sessão 3.2 é configurado o problema em função de um sistema de fila; na sessão 3.3 é desenvolvido o modelo e sistema de simulação e a sessão 3.4 finaliza o desenvolvimento com o estudo de caso.

3.1 MAPEAMENTO DO PROCESSO

3.1.1 Área de concessão

As concessionárias de distribuição de energia elétrica geralmente atendem a uma grande área geográfica e as demandas de serviço podem ser geradas em qualquer cliente e equipamento da região de concessão. Para a empresa que foi estudada nesse trabalho, a AES Sul Distribuidora Gaúcha de Energia, essa área corresponde a mais de 99.000 km², formada

por 118 municípios, e atende a mais de 1.300.000 clientes. Na Figura 5 é possível visualizar a representatividade territorial da empresa na distribuição de energia elétrica no estado do Rio Grande do Sul:

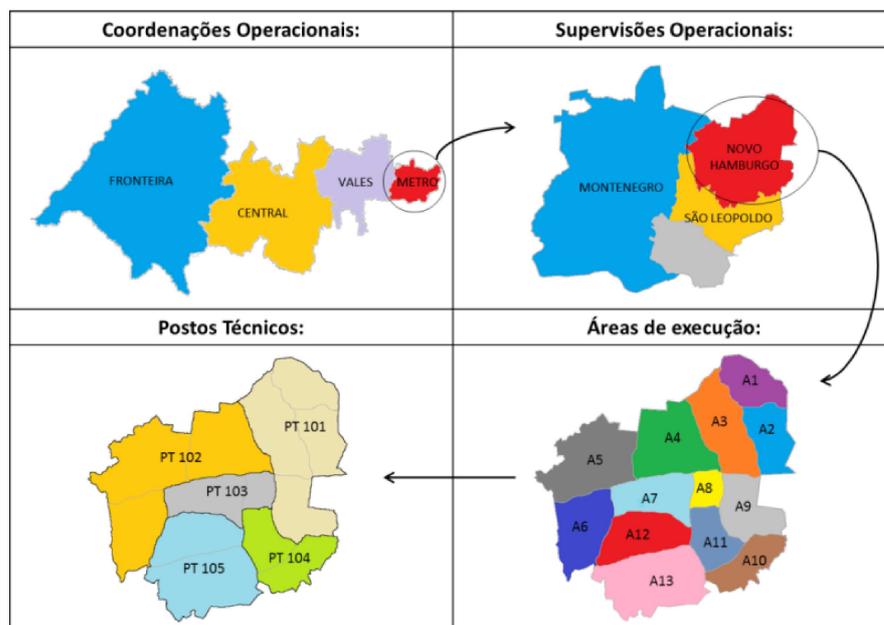
Figura 5 - Área de concessão



Fonte: AES Sul.

Devido ao tamanho da área de concessão e quantidades de demandas geradas, em média são gerados em torno de 500.000 serviços por ano - entre serviços comerciais, técnicos e emergenciais (os tipos de serviço serão abordados na sessão 3.1.4) -, a empresa segregou sua área de concessão em regiões menores dividindo assim o problema em instâncias menores e conseqüentemente minimizando a complexidade de gestão.

Figura 6 – Divisão da área de concessão



Fonte: Desenvolvido pelo autor.

Como pode ser visto na Figura 6, a área de concessão da empresa é dividida nos seguintes tipos de conjuntos:

- a) Coordenações Operacionais: a AES Sul é dividida em 4 macrorregiões denominadas coordenações: Metropolitana, Vales, Central e Fronteira. Essa divisão tem um caráter relacionado à gestão de recursos humanos. Onde para cada coordenação existe um coordenador responsável por gerir a estrutura de recursos humanos das supervisões vinculadas.
- b) Supervisões Operacionais: são vinculadas a uma coordenação operacional, e possuem a função de gerenciar diretamente as equipes de campo garantindo que o recurso esteja disponível de maneira apropriada, nos horários pré-determinados e com um rendimento adequado.
- c) Áreas de execução: menor divisão geográfica da área total de concessão da empresa. São definidas por polígonos que limitam cada área em um software de cadastro gráfico (GIS), segregando cada supervisão operacional em diversas subáreas. De modo que, no sistema de operação (software) os serviços e equipes ficam vinculados a ela;
- d) Posto técnico: é uma estrutura de dados que possibilita criar conjuntos de áreas de execução. Esse atributo espacial é o que segrega efetivamente a área de concessão da empresa quando se estuda o despacho de serviços. Sendo esse o atributo de ligação entre a localização de demandas e equipes, e assim determinando um âmbito de atendimento. Analisando a Figura 6 identifica-se que os postos técnicos, que compõem a supervisão de Novo Hamburgo, possuem a seguinte conformação:

- PT 101 = {A1, A2, A3, A9};
- PT 102 = {A4, A5, A6};
- PT 103 = {A7, A8};
- PT 104 = {A11, A10};
- PT 105 = {A12, A13}.

Assim, ao analisar o Posto Técnico 101, identifica-se que ele é formado pelo conjunto das áreas de execução A1, A2, A3 e A9 e, conseqüentemente, todos os serviços e equipes vinculados a essas áreas de execução formam, respectivamente, as demandas e recursos deste Posto Técnico.

Portanto, o Posto Técnico é o componente central desse sistema onde, a partir da sua delimitação, é possível associar as demandas e recursos e efetivamente executar um processo de despacho de serviços.

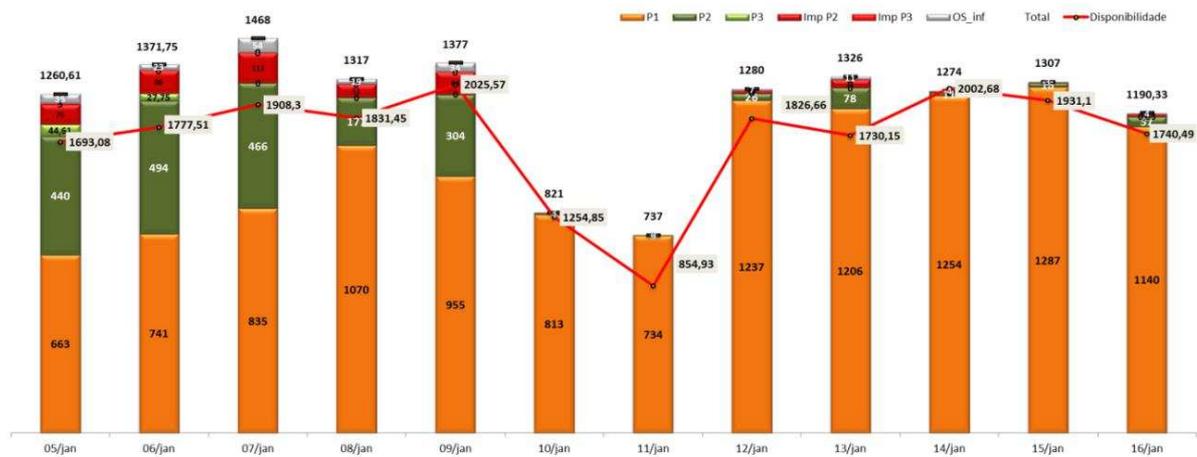
3.1.2 Processos de despacho

Atualmente a empresa possui dois processos distintos de despacho:

- a) Programado: Esse processo é efetuado somente para as ordens comerciais e técnicas, onde essas, possuem um prazo previsto de execução maior que 1 dia útil, possibilitando que sejam roteirizadas na madrugada anterior ao dia de serviço das equipes. Desse modo, existe um processo automático que é executado todos os dias às 2 horas da manhã que, para cada Posto Técnico, selecionam-se todas as demandas comerciais e técnicas pendentes mais todas as equipes vinculadas a ele e gera uma instância do problema a ser resolvido por um algoritmo de roteirização. Este, por sua vez, define a rota para todas as equipes do Posto Técnico para o dia que está iniciando.
- b) Tempo Real: Esse processo é efetuado para as ordens de emergência que ingressam ao longo do dia. As ordens emergenciais são serviços derivados de falhas no fornecimento de energia elétrica ao cliente e, conseqüentemente, precisam ser atendidos no menor tempo possível, minimizando o impacto ao cliente. Então, um despacho em tempo real se faz necessário e esse, por sua vez, irá postergar a seqüência de atendimentos determinados pelo processo programado mencionado anteriormente. E, salvo em eventos climáticos severos (temporais), esse tipo de serviço tende a ficar pouco tempo esperando para ser atendido, diferentemente das ordens programadas.

Assim, a interferência do despacho em tempo real sobre o despacho programado, ocasiona uma variação na execução prevista dos serviços programados, conforme se visualiza na Figura 7:

Figura 7 – Evolução diária de execução de serviços



Fonte: AES Sul.

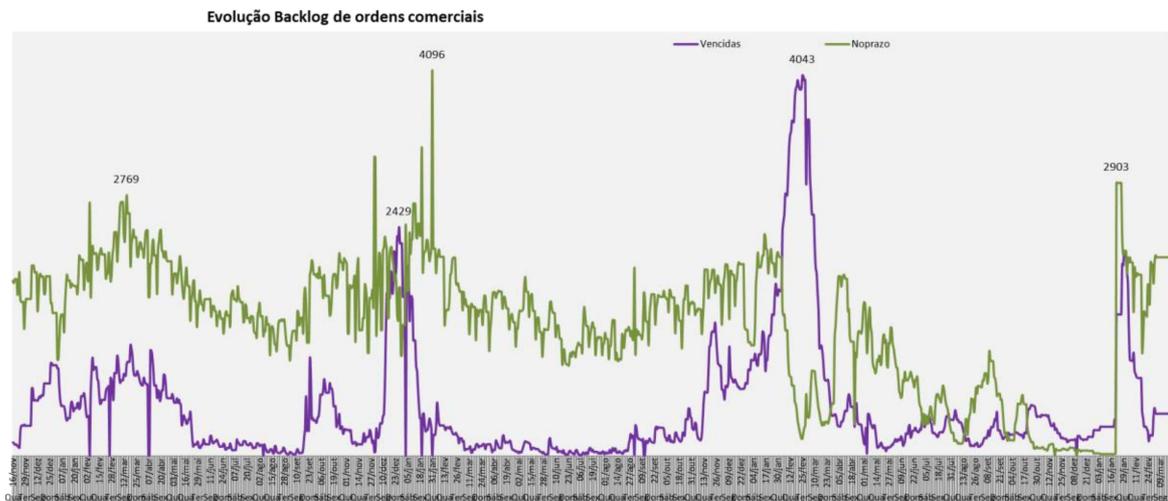
A Figura 7 é um gráfico de controle da empresa que identifica a quantidade de serviços executados por dia e estratificado por grupo de serviço:

- P1: Ordem emergencial;
- P2: Ordem comercial;
- P3: Ordem técnica;
- imp. P2: improdutivo de P2;
- imp. P3: improdutivo de P3;
- OS_inf: outras ordens de serviço.

Nesse, fica evidenciado a diminuição dos atendimentos de ordens programáveis (comerciais e técnicas) devido ao aumento de ordens emergenciais: Na segunda-feira (05/01/2015) foram executadas 440 ordens comerciais contra 663 ordens emergenciais, entretanto na segunda-feira subsequente (12/01/2015) foram executadas somente 26 ordens comerciais contra 1.237 ordens emergenciais. Esse comportamento evidencia a priorização do atendimento emergencial (falhas) em detrimento ao despacho de serviços comerciais e técnicos.

Outro ponto relevante, agora relacionado ao despacho programado, é a variação na quantidade de ordens pendentes, e dessa, a relação de ordens dentro e fora do prazo de resolução estipulado. Na Figura 8, é identificada a oscilação na quantidade de ordens em relação ao prazo estipulado:

Figura 8 – Evolução ordens de serviços pendentes



Fonte: AES Sul.

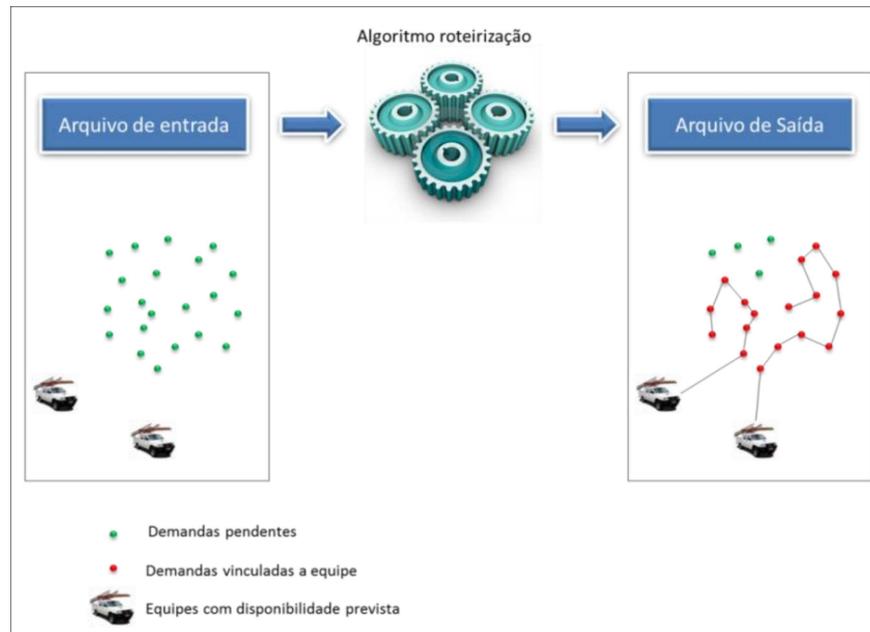
Esse ponto é de extrema relevância, pois em alguns momentos o nível de serviço é fortemente prejudicado, chegando ao pico de 4.043 ordens fora de prazo. Essa postergação, da execução de ordens pode ter várias causas, entre elas: elevação da taxa de ingresso de serviços, redução da taxa de atendimento, ou até mesmo a interferência do despacho de tempo real, conforme evidenciado na Figura 7.

Esse problema, afeta o equilíbrio do ambiente, interferindo diretamente no despacho tanto programado como de tempo real, pois em um cenário crítico as ordens comerciais poderão ter importâncias diferentes entre si, mas principalmente, poderá haver uma alteração das definições iniciais, onde o atendimento de emergência sempre teria prioridade em relação ao atendimento programado, mas devido ao estado das ordens comerciais chegarem a um nível tão crítico que seja insustentável a sua postergação e assim interferirem na decisão de despacho de tempo real, sendo que o limiar de decisão não é de fácil definição.

3.1.3 Despacho de serviços programáveis

Conforme discutido no item 3.1.2, para cada posto técnico são identificadas as equipes com disponibilidade prevista para o dia seguinte e todas as demandas programáveis pendentes de execução. Essas informações são agrupadas gerando diversas instâncias para um processo de roteirização automático, onde ao final desse processo é definida a rota programada de todas as equipes da empresa para o dia seguinte.

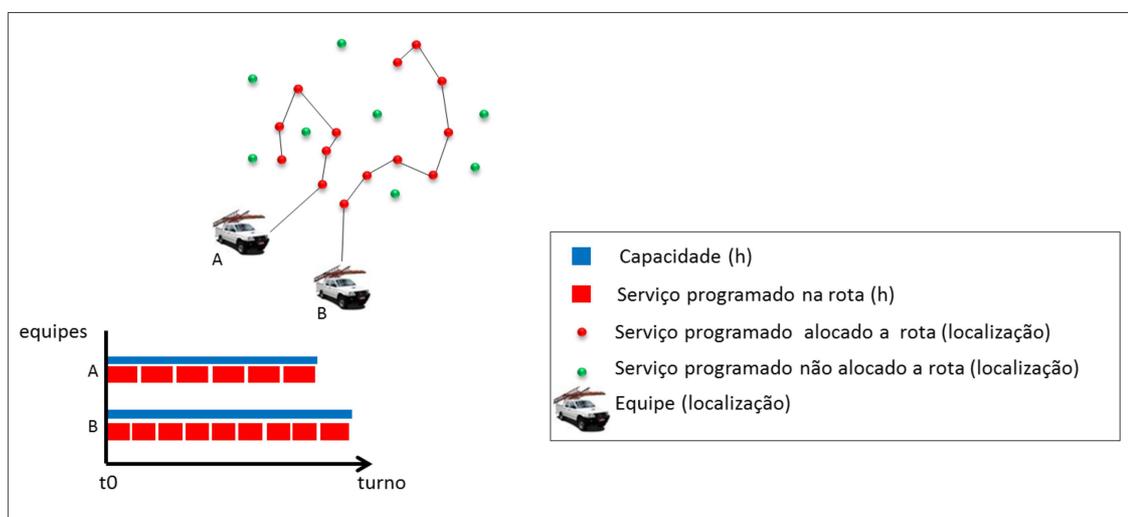
Figura 9 – Processo de despacho



Fonte: Desenvolvido pelo autor.

A partir da solução do algoritmo, as rotas são efetivadas no sistema (gravadas no banco de dados) e enviadas para o terminal portátil das equipes. Deste modo, conforme as equipes iniciam o turno o terminal portátil é sincronizado com o sistema e o despacho é executado.

Figura 10 – Planejamento de rotas de serviços para o dia subsequente



Fonte: Desenvolvido pelo autor.

Na Figura 9 e na Figura 10 identifica-se que existem demandas de serviço pendentes não vinculados a rotas de equipes (pontos verdes), isso porque a quantidade de recurso disponível é menor do que a demanda necessita, conforme a visualização da rota planejada das equipes na Figura 10, onde a capacidade das equipes em tempo (linha azul) é igual ao somatório dos tempos de deslocamento e serviço de cada ordem de serviço programada na rota (retângulos vermelhos), não restando capacidade ociosa. Portanto, existe uma decisão inicial: quais ordens de serviço terão alguma possibilidade de ser executada no dia. Pois as ordens que não forem alocadas pelo processo automático, somente concorrerão a ingressar em uma rota na próxima execução, neste caso, permanecendo mais um dia sem chances de serem executadas.

Porém, a alocação de uma ordem de serviço a alguma rota não garante a sua execução, pois variações ocorrem ao longo do dia, geralmente, postergando a execução desses serviços na agenda da equipe. Inicialmente podemos citar os seguintes problemas como mais impactantes na garantia de execução da rota programada da equipe:

- a) Despacho de tempo real: Ingresso de ordens de serviços emergenciais, que devem ser atendidas no menor prazo possível, ocasionando uma interferência na roteirização programada das equipes e consequentemente aumentando os riscos de as ordens programadas não serem executadas, conforme ilustrado na Figura 7.
- b) Diminuição da capacidade: Fatores que fazem a capacidade do recurso ser menor que a contratada. Como exemplo pode-se citar reuniões, treinamentos, faltas de funcionários, quebras de veículos, ociosidade da equipe, falta de eficiência na execução da tarefa, entre outros.

Assim, para descrever um modelo que represente a realidade deste sistema, faz-se necessário analisar mais a fundo as demandas de serviço e seu comportamento de ingresso, e o comportamento dos recursos envolvidos na execução desses serviços, para isso a seção 3.1.4 busca descrever o comportamento individual das demandas, a seção 3.1.5 descreve o comportamento dos recursos e a seção 3.1.6 descreve o comportamento da inter-relação entre capacidade e demanda, definindo a configuração do problema.

3.1.4 Atributos de serviço

As demandas de serviço são divididas em dois grupos principais, ordens programáveis e emergenciais, que se distinguem basicamente pela urgência e consequentemente pelo prazo previsto de execução com que devem ser atendidas, porém essas ordens compartilham de

vários atributos comuns, conforme pode ser visualizado na Tabela 2 – Atributos genéricos das ordens de serviço. Esses atributos têm por objetivo qualificar o atendimento em função da sua localização, prazos, custos, abrangência, entre outros.

Tabela 2 – Atributos genéricos das ordens de serviço

Atributo	Descrição
Data de ingresso	Data/hora que ingressou o atendimento.
Tipo	Programável ou emergencial.
Subtipo	Subdivisão dos tipos de serviço.
Prazo	Tempo máximo previsto para a execução da ordem.
Valor	Alguns tipos de ordens possuem um valor de débito ou crédito aos clientes. Em ambos os casos a postergação da ordem onera a empresa.
Coordenadas geográficas	Latitude e longitude da ordem de serviço. Atributo herdado do cadastro do cliente ou do equipamento que o serviço é gerado.
Tempo de serviço previsto	Tempo de execução do serviço em campo, é herdado do subtipo da ordem.
Criticidade	Atributo que identifica uma alteração na prioridade desta devido a um problema extraordinário, geralmente vinculado ao cliente necessitar urgência na execução (risco de acidente, hospitais, entre outros).
Área de execução	Menor divisão geográfica da área de concessão. É um atributo herdado do cadastro do cliente.

Fonte: Desenvolvido pelo autor.

Por outro lado, Tabela 3 identifica os atributos que diferenciam as ordens programadas e emergenciais:

Tabela 3 – Atributos que diferenciam as ordens emergenciais e programáveis

Atributo	Emergencial	Programável
Subtipo	Instalação (afeta um equipamento e vários clientes) ou fornecimento (afeta um cliente)	Ligação de energia, troca de medidor, mudança de medição, corte para reforma, religue de reforma, ordens de manutenção, ordens de corte por inadimplência.
Prazo	Unidade de medidas em horas	Unidade de medidas em dias.
Valor	Custos vinculados à transgressão das metas de fornecimento de energia elétrica.	Custo vinculado a multas relacionadas ao não cumprimento de prazos de ordens de serviço.

Fonte: Desenvolvido pelo autor.

A distinção entre as ordens de emergência e programável se dá pelo caráter urgente das ordens emergenciais, que estão relacionadas a falhas no fornecimento de energia elétrica

ao cliente, onerando o mesmo com o passar do tempo, pois o impossibilita de utilizar o serviço já contratado da empresa, fornecimento de energia elétrica.

Os atributos das ordens de serviços são definidos no momento da geração e somente alguns deles sofrem alterações ao longo do tempo de espera para a sua execução. Portanto, ao definir o local de atendimento, obrigatoriamente identifica-se o cliente ou o equipamento afetado, e assim o serviço gerado herda os atributos do local do atendimento: coordenada, Área de execução do cliente ou equipamento que está sendo gerada a ordem de serviço. Ao definir o tipo do atendimento, obrigatoriamente identifica-se o prazo da ordem e o tempo de serviço previsto, pois ambos são relativos ao tipo de atendimento. E conseqüentemente, ao definir o tipo do atendimento e o local do atendimento, os custos diretos vinculados às ordens também estarão definidos. Na Tabela 4 é apresentado um exemplo de duas ordens de serviços geradas, onde já possuem os atributos inicializados.

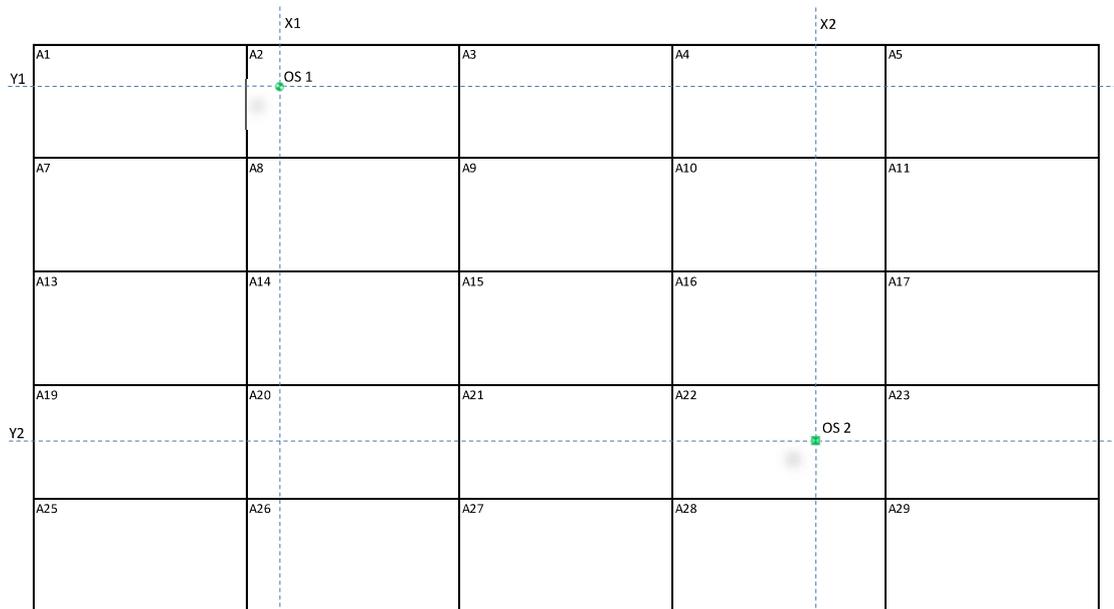
Na relação entre a Tabela 4 e a Figura 11, ao gerar as ordens de serviço OS1 e OS2, estas recebem os atributos do cliente: coordenadas e área de execução, esses dois atributos qualificam de forma clara a localização do atendimento. Onde, a OS1 é vinculada à área A2 e a OS2 é vinculada à área A22. Após isso, essas ordens estarão disponíveis para serem executadas, aguardando na fila de serviços pendentes o momento do atendimento. Porém, não necessariamente esses dois serviços estarão concorrendo em um mesmo processo de roteirização, pois as áreas A2 e A22 podem fazer parte de postos técnicos diferentes e conseqüentemente não estarão aguardando na mesma fila de atendimento.

Tabela 4 – Ordens de serviços

Atributos	OS1	OS2
Área de execução	A2	A22
Coordenadas	-29,893843975;-53,32435544	-29,82473975;-53,2655544
Tipo	Ligação Nova	Corte – Encerrado
Valor	0	R\$ 40,00
Prazo	3 dias	20 dias
Data geração	03/12/2015 11:11	03/12/2015 12:00

Fonte: Desenvolvido pelo autor.

Figura 11 – Representação geográfica das ordens de serviço



Fonte: Desenvolvido pelo autor.

3.1.5 Capacidade

A capacidade da empresa em executar os serviços é determinada pela quantidade de equipes e o tempo que cada uma delas permanece disponível em cada posto técnico. Por sua vez, a equipe é formada por um conjunto de recursos contratados pela empresa, são eles: camioneta, ferramentas, materiais e pessoas (funcionários).

Como o vínculo empregatício entre empresa e funcionário atende a legislação vigente, que determina uma quantidade máxima de horas de serviço mensal, dentre outras regras, a maior parte das equipes possui uma disponibilidade média mensal de 176 horas, devendo possuir um calendário previsto de turno diário, determinando a sua previsão de Turno diária.

Porém, existem diversos fatores que ocasionam uma diminuição da disponibilidade efetiva das equipes, conforme relacionado na Tabela 5:

Tabela 5 – Variação da capacidade

Quantidade de equipes contratadas
(*) turno previsto conforme legislação
(=) Disponibilidade contratada
(-) Disponibilidade não prevista em calendário
(=) Disponibilidade prevista (calendário com horários das equipes)
(-) Calendário não realizado (faltas, atrasos, férias, etc.)
(=) Disponibilidade realizada
(-) Tempo em atividade (reuniões, treinamentos, etc.)
(-) Tempo ocioso (períodos sem execução de tarefas)
(=) Disponibilidade efetiva

Fonte: Desenvolvido pelo autor.

Deste modo, a disponibilidade contratada não necessariamente gerará uma mesma disponibilidade efetiva, diminuindo a capacidade de execução de serviços. Na Tabela 6 é mostrado o histórico de turnos de trabalho (em horas) da equipe 213 entre os dias 01/02/2016 e 10/02/2016, ao avaliar o resultado da última coluna identifica-se que a equipe não cumpre de forma completa o turno previsto, diminuindo a sua capacidade de execução.

Tabela 6 – Comportamento da disponibilidade da equipe 213

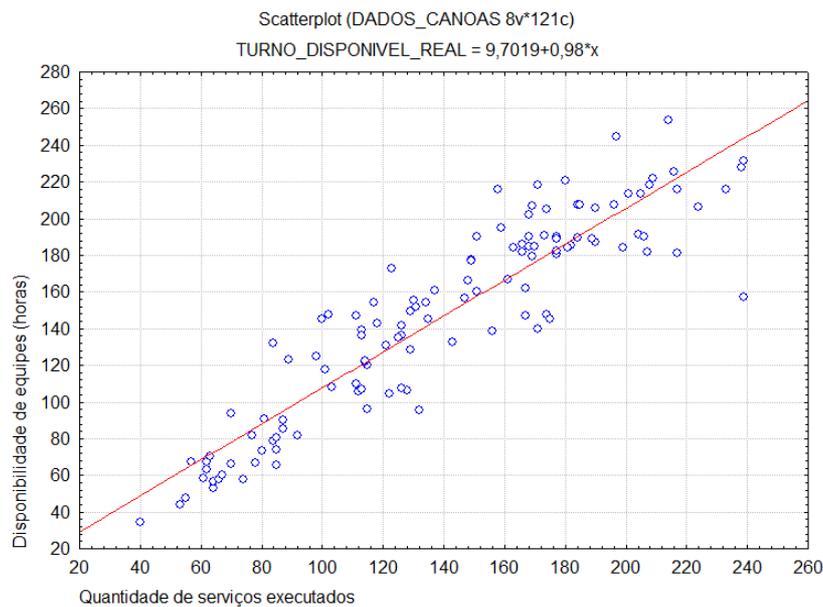
Dia	Turno previsto	Turno realizado	Atraso	Atividade	Tempo ocioso	Turno efetivo	Percentual efetivo
01/02/2016	7,8	8,3	0	2,8	0,53	4,97	64%
02/02/2016	8,8	8,78	0,02	0,87	0,8	7,11	81%
03/02/2016	8,8	7,65	1,15	0,68	1,58	5,39	61%
05/02/2016	8,8	8,1	0,7	0,03	0,43	7,64	87%
08/02/2016	8,8	8,98	0	2,42	1,07	5,49	62%
09/02/2016	9	6,8	2,2	0,33	0,6	5,87	65%
10/02/2016	8,8	8,4	0,4	0,57	1,97	5,86	67%

Fonte: Desenvolvido pelo autor.

Sendo esse um ponto evidente de interferência no processo de execução de serviços, já que a execução de serviços depende diretamente da duração do turno efetivo da equipe, conforme pode ser visualizado na Figura 12. Nesse gráfico de dispersão é relacionado o valor do somatório dos turnos efetivos das equipes com a quantidade de ordens executadas por essas mesmas equipes. Efetuando uma análise de correlação desses dados, identifica-se que a disponibilidade efetiva das equipes explica em 84% a execução de serviços no dia, possibilitando determinar que cada hora a mais de disponibilidade efetiva acarreta a execução

de 0,98 serviços. Deste modo, a definição da disponibilidade de forma adequada e a análise do comportamento da variação na execução dessa disponibilidade se torna um fator determinante para a execução das rotas planejadas e conseqüentemente da execução dos serviços.

Figura 12 – Disponibilidade efetiva versus execução de serviços



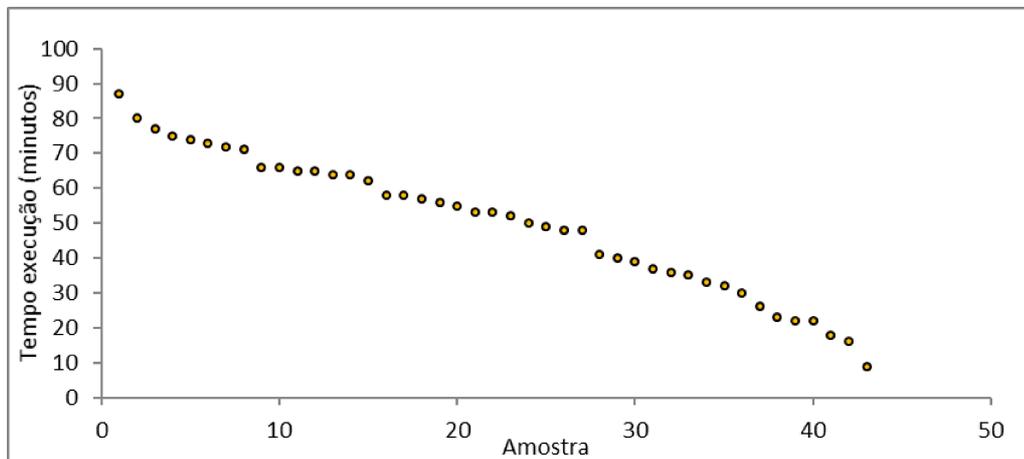
Fonte: Desenvolvido pelo autor.

Porém, a variação dos pontos ao redor da reta poderia ser mais bem explicada, com a inclusão de outras duas variáveis: tempo de deslocamento e tempo de serviço. Pois, passado o problema inicial de disponibilidade, enquanto a equipe está disponível efetuando os serviços, ocorre a variação na execução dos serviços, em relação basicamente a tempos de execução de deslocamento e execução da tarefa:

- a) **Deslocamento:** Como em um problema de roteirização de veículos, parte-se da premissa que as tarefas são executadas em locais distintos, em que o deslocamento é inerente ao ato de execução dos serviços. A velocidade do deslocamento, o caminho tomado, a disciplina da equipe e a quantidade de tráfego são fatores importantes para determinar o tempo de deslocamento entre dois pontos. Esse tempo de deslocamento pode variar de equipe para equipe, horário para horário e região para região.

- b) Tempo de serviço: Do mesmo modo que o deslocamento, toda ordem de serviço consome tempo do recurso para execução da tarefa em campo. Este tempo varia pelo tipo de atividade, cenário encontrado em campo no momento da tarefa, pelo desempenho da equipe, entre outros. Na Figura 13 apresenta-se uma amostra dos tempos de execução da ordem de serviço tipo TO562, efetuados pela equipe 213 entre os dias 01/02/2016 e 10/04/2016, nesse período a equipe efetuou 43 ordens desse tipo. Variando os tempos de execução entre 9 e 86 minutos de maneira relativamente uniforme.

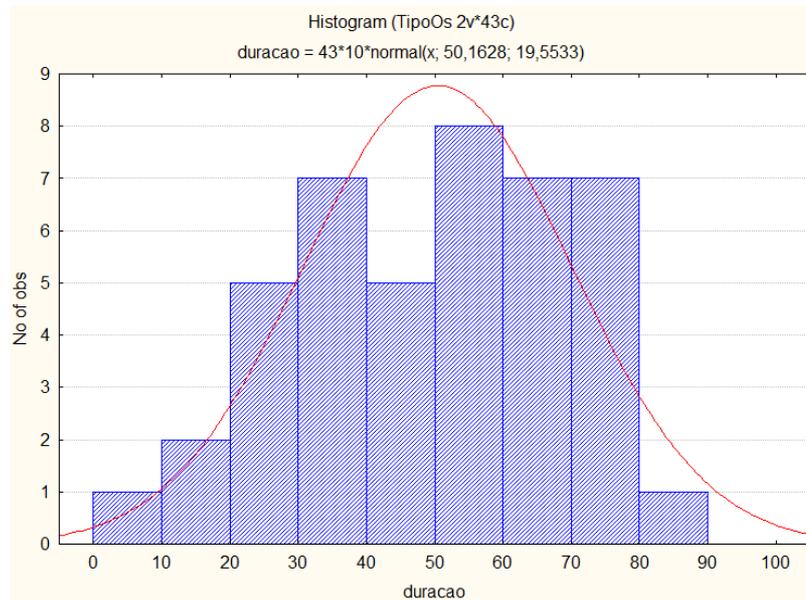
Figura 13 - Tempo de execução serviço TO562 pela equipe 213



Fonte: Desenvolvida pelo autor.

Porém ao analisarmos o gráfico de frequências, histograma, a distribuição não possui essa uniformidade, tão pouco fica aderente a alguma curva de distribuição tradicional, conforme pode ser visualizado na Figura 14.

Figura 14 - Distribuição de frequência dos tempos de execução da equipe 213



Fonte: Desenvolvido pelo autor.

Deste modo, é possível propor que existam outras variáveis interferindo no processo de execução. Sendo elas: tipo de atividade específica (pequenas variações dentro de um mesmo tipo de serviço), diversidade de cenários encontrados em campo no momento da tarefa, variabilidade no desempenho da equipe, clima, interferência do cliente, entre outros. Sendo que, esse tema será mais bem abordado no item 3.4.1.

Partindo do que já foi dito até aqui, podemos definir uma série de atributos para configurar as equipes e assim poder descrevê-las a partir do comportamento desses. Para isso, na Tabela 7, estão listados um conjunto de atributos mais relevantes e que devem ser monitorados no processo de modelagem do problema, nessa tabela também consta uma breve descrição de cada atributo.

Assim, cada equipe pode possuir um comportamento singular, o que ocasionará uma modelagem que contemple desempenhos diferentes no resultado de execução de serviços. Portanto, a individualidade dos recursos deve ser levada em consideração para um adequado dimensionamento da força de trabalho.

Tabela 7 – Atributos das equipes

Atributo	Descrição
Tipo	Especifica quais as habilidades que a equipe possui, influenciando os serviços que pode executar.
Coordenada inicial	Latitude e longitude na qual a equipe irá iniciar o turno.
Coordenada parcial	Latitude e longitude atualizada ao longo do turno de trabalho, coletada através de GPS.
Custo	Custo/ hora que custa a equipe.
Disponibilidade contratada	Carga horária contratada.
Disponibilidade prevista	Carga horária programada para a equipe.
Tempo de atividade	Tempo de atividade programada para equipe.
Tempo de ociosidade:	Tempo médio que a equipe fica sem efetuar tarefas durante a disponibilidade prevista.
Disponibilidade efetiva	Carga horária efetiva, após descontar todas as parcelas de indisponibilidade.
Eficiência deslocamento	Quanto a equipe desvia do tempo de deslocamento previsto.
Eficiência tempo de serviço:	Quanto a equipe desvia do tempo de execução de tarefas.
Área de execução:	Área de execução que a equipe está vinculada.

Fonte: Desenvolvido pelo autor.

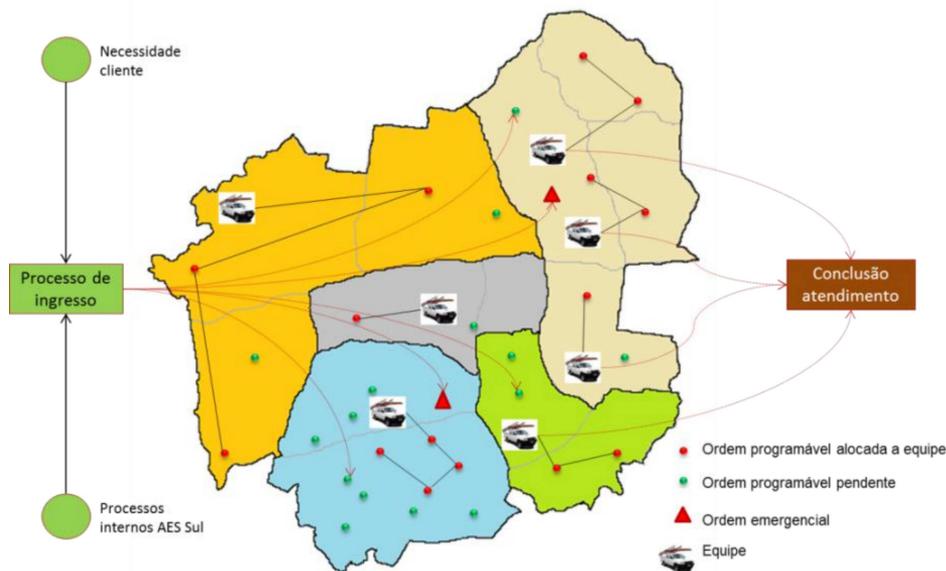
3.1.6 Configuração do problema

Ao analisar a descrição do problema até o momento identifica-se que a empresa possui uma quantidade de equipes contratadas e que pouco oscila, devido a ter constantes do tipo quantidade de carros, eletricitas e equipamentos, caracterizando um limite máximo para a capacidade, definido como capacidade fixa por Fitzsimmons, J. e Fitzsimmons, M. (2010), sendo que, nesse caso o que varia é a eficiência da equipe em executar a capacidade contratada. Por outro lado, o ingresso de demandas pode oscilar tanto na quantidade de serviços ingressados como no tipo desses, afetando diretamente o nível de serviço prestado pela empresa.

Basicamente, o problema em si se comporta de modo muito parecido com o sistema de filas apresentado por Hillier e Liebermann (2013), onde existe um comportamento de ingresso de serviços a partir de uma necessidade do cliente ou de algum processo interno da empresa, uma espera por parte do serviço para o atendimento e o atendimento propriamente dito efetuado pela equipe. Esse sistema pode ser representado de forma simplificada pela Figura 15. Possibilitando visualizar os atendimentos em execução, os serviços, esperando o atendimento e uma fonte geradora ingressando novos serviços para o sistema de filas, também

fica evidente o problema de deslocamento, pois as ordens de serviços estão posicionadas em locais distintos, desse modo cada nível de priorização existente interfere no processo de roteirização das equipes, pois para priorizar a execução de um serviço a rota deverá ser alterada, sendo que esse serviço priorizado pode aumentar o tempo de deslocamento na rota.

Figura 15 – Representação do processo



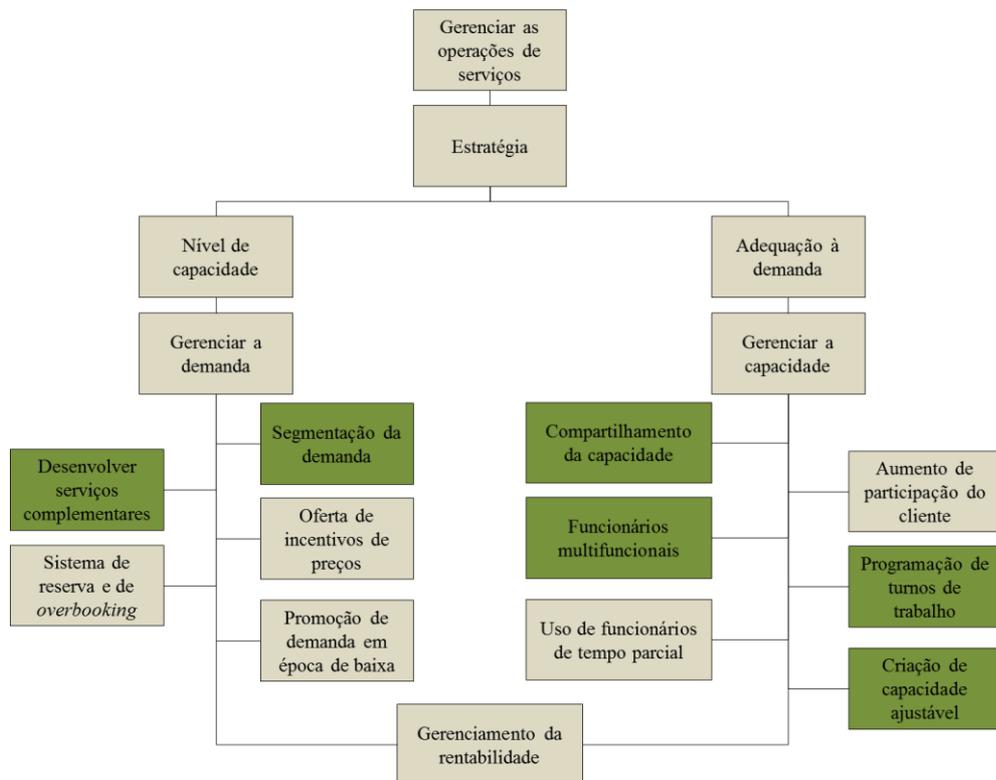
Fonte: Desenvolvido pelo autor.

Além do sistema de filas, o processo também está diretamente relacionado com o processo de gerenciamento de backlog, e como citado por Fitzsimmons, J. e Fitzsimmons, M. (2010) um modo de tratar esse tipo de problema está em duas estratégias genéricas: controlar o nível de capacidade, através do gerenciamento da demanda, e a adequação à demanda, através do gerenciamento da capacidade. Nessas estratégias existem algumas propostas de ação, sendo que a utilização destas depende do comportamento de cada problema.

Na Figura 16, a partir dessas propostas de ações, identifica-se em verde quais são mais aderentes ao problema em questão, são elas: segmentação da demanda – relacionado com a segregação de postos técnicos e conseqüentemente dos serviços –, desenvolver serviços complementares – geração de serviços de manutenção preditivas em momentos ociosos das equipes e assim possibilitar minimizar o ingresso de atendimentos emergenciais –, compartilhamento da capacidade – compartilhar equipes entre postos técnicos, utilizando-se de forma mais eficaz os recursos e conseguindo equilibrar o nível de serviço entre os postos técnicos –, funcionários multifuncionais – necessidade básica para possibilitar a dinamicidade

da utilização do recurso –, programação de turnos de trabalho – com vistas a dimensionar a força de trabalho adequada por horário e dia da semana de acordo com a geração e nível de serviço esperado –, criação de capacidade ajustável – abordagem relevante para aumentar a capacidade através de horas extras em momentos de aumento elevado do ingresso de serviços, com exemplo pode-se citar eventos climáticos severos (tempestades).

Figura 16 – Gerenciamento de serviços



Fonte: Ajustado de Fitzsimmons, J. e Fitzsimmons, M. (2010).

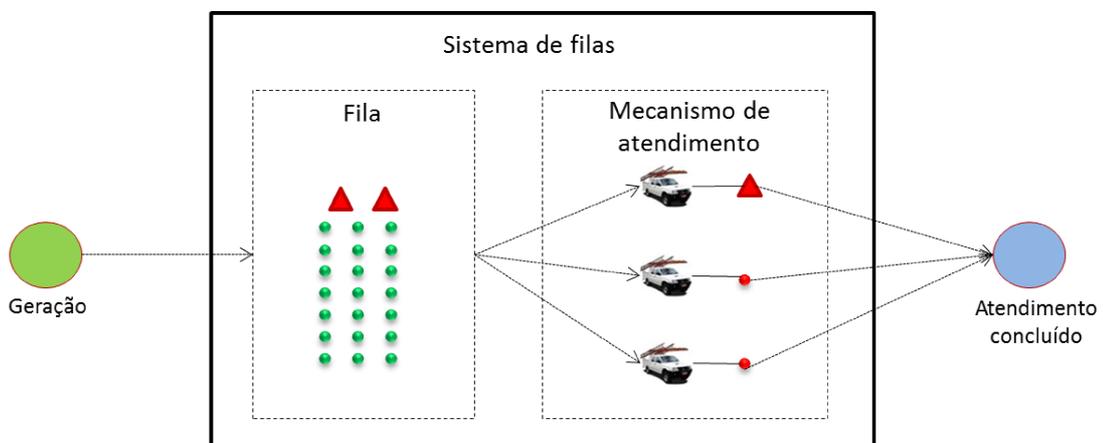
As abordagens marcadas na Figura 16, são de extrema relevância nesse problema, pois tratam de possíveis pontos de controle do problema em questão. Sendo que, essas estratégias serão discutidas ao longo da modelagem, possibilitando ao não, a sua inserção no modelo de fila sugerida.

3.2 CONFIGURAÇÃO DO PROBLEMA COMO SISTEMA DE FILAS

Nos itens anteriores, foram definidos os principais componentes do sistema: demanda e capacidade. Onde fica evidente a singularidade de cada indivíduo, ordem de serviço e

equipe, quando a localização, custo, tipo e prazo definem a individualidade do serviço e a localização, tipo, disponibilidade e o desempenho na execução definem a equipe. Logo, ao determinar a conformação dos postos técnicos e a atribuição de equipes a ele, determinaria o comportamento do backlog (ordens pendentes). Porém, é o estado do conjunto de ordens de serviço pendentes que solicita uma maior ou menor quantidade de recurso para atendê-lo, em vistas a reduzir o tempo de espera das ordens de serviços e assim aumentar o nível de serviço.

Figura 17 – Sistema básico de filas para o problema



Fonte: Desenvolvido pelo autor.

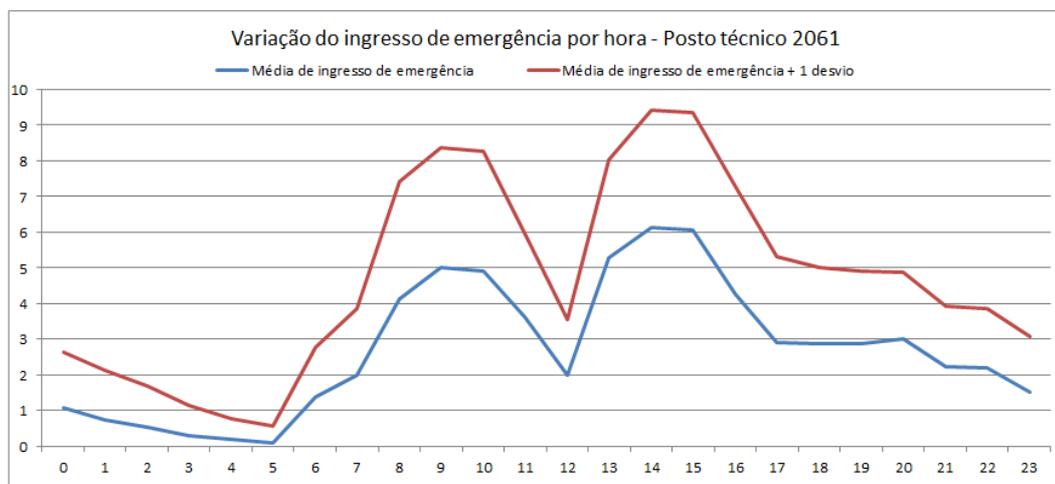
Conforme Hillier e Liebermann (2013), a teoria das filas compreende, basicamente, o estudo dos tempos de espera em um sistema, com o objetivo de entender o sistema de fila, e assim poder controlar o atendimento do nível de serviço desejado com a melhor utilização do recurso disponível. Assim, pode-se iniciar um paralelo entre o problema descrito e a teoria das filas. Sendo que, basicamente, o sistema se divide em uma fonte de entrada, a fila aonde os clientes esperam e o mecanismo de atendimento, local aonde os clientes são atendidos pelos servidores. Ao traçarmos um paralelo com o estudo de caso em questão, relacionando a Figura 1 com a Figura 17, se pode supor a existência de um subsistema de filas para cada posto técnico, onde existe uma geração de serviços e equipes de atendimentos específicos para cada um, conforme identificado na Figura 17.

3.2.1 Fonte de entrada

Entidade que gera um novo cliente para o sistema de filas, nesse caso em específico o cliente é representado pela ordem de serviço. Agrupando o problema em uma única instância, uma ordem poderia ser gerada em qualquer um dos quase 1.300.000 clientes da empresa, assim a população teria condições de ser considerada infinita. Porém, como cada instância corresponde a um posto técnico e o tamanho da fila (quantidade de ordens pendentes) não tem restrição dentro do sistema, sendo limitada somente pela quantidade total de clientes do posto, deverá ser analisado se para todos os postos técnicos o tamanho que a fila geralmente chega não interfere na quantidade de ordens a serem ingressadas (quantidade de clientes do posto que ainda não ingressaram serviços, os possíveis candidatos), se isso ocorrer a população deve ser considerada infinita, caso contrário será considerada finita, conforme orientação de Hillier e Liebermann (2013).

A taxa de ingresso (λ) representa a quantidade média de ingressos de ordens serviço por unidade de tempo. Neste ponto, existe uma discussão relevante sobre a unidade de tempo, como as ordens emergenciais tem por objetivo atender o mais cedo possível, a variação com que ocorrem ao longo de um dia é relevante, então a unidade de tempo poderá ser em torno de 1h ou até 0,5h, pois é plausível imaginar que em horários de maior consumo de energia, possa existir uma geração de demanda mais elevada em relação a um horário com menor demanda (exemplo: a madrugada), interferindo no dimensionamento dos recursos.

Figura 18 - Média de atendimentos e desvio padrão



Fonte: Desenvolvido pelo autor.

Conforme pode ser visualizado na Figura 18, ao longo do dia ocorre uma variação elevada na taxa de ingressos de emergência por faixa de horário.

Entretanto, como as ordens programáveis possuem mais do que 1 dia de prazo, e não são despachadas no mesmo dia no qual são ingressadas, não se faz necessário essa estratificação por horário, podendo ser agrupada em dias. De modo que, quanto maior o prazo de uma ordem de serviço, maior poderia ser a granularidade deste agrupamento. Assim, a taxa de ingresso poderá ser segregada conforme o tipo de serviço: λ_t .

3.2.2 Fila

É o local onde os clientes esperam, nesse caso a fila não existe de maneira física, pois os serviços estão esperando para serem efetuados nos locais em que se encontram, na casa do cliente. Assim o cliente não tem noção do tamanho da fila e somente saberá que terá problemas de não atendimento do prazo quando esse fato já tiver acontecido. Deste modo, λ_t não tem influência da quantidade de clientes que estão no sistema de fila (n), pois não existe a indução de desistência de aguardar na fila em função do tamanho da fila.

O tipo da fila pode variar, mas comumente poderá ser considerada como canais múltiplos (geralmente existe mais de uma equipe por posto técnico) e fase única (geralmente uma equipe resolve o atendimento em uma única visita).

3.2.3 Instalações de atendimento

A fonte de entrada poderá ingressar uma nova ordem de serviço em qualquer lugar da área de concessão. Porém, nem todas as equipes estarão aptas para atender. Assim, os serviços ingressados somente serão atendidos por equipes cadastradas no mesmo posto técnico da ordem de serviço, então o posto técnico é a entidade que “recebe” os serviços da região e possui as equipes como recurso para executar estas demandas. Portanto, ele (posto técnico) pode ser considerado a própria instalação de atendimento e a equipe, por sua vez, é o atendente (também conhecido como servidor).

Em relação à taxa média de atendimento por unidade de tempo (μ), existem algumas particularidades interessantes:

- a) Como as ordens de serviço estão espalhadas pelo posto técnico, existirá, além do tempo de serviço para executar a tarefa, o tempo de deslocamento para a equipe se deslocar entre um local de atendimento e outro, assim, o tempo de atendimento de

uma ordem de serviço é o somatório entre o tempo de deslocamento e o tempo de serviço utilizado;

- b) Conforme abordado anteriormente, o tempo de deslocamento e o tempo de serviço variam em relação à equipe que efetuará o atendimento, dependendo do seu desempenho individual de deslocamento e serviço;
- c) O tempo de serviço também varia em relação ao tipo de serviço, uma ordem programável de mudança de local de atendimento possui um tempo médio de execução em torno de 1,5 horas, por sua vez a ordem de ligação e instalação da medição possui um tempo médio de execução em torno de 0,75 horas.

Assim, a taxa média de atendimento por unidade de tempo possui a seguinte variação:

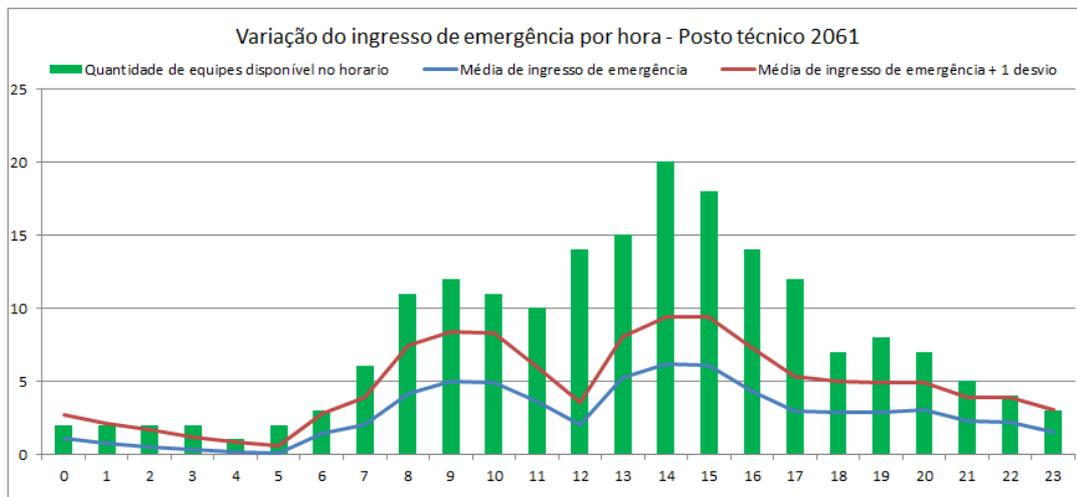
- $\mu_{e,t}$, taxa média de atendimento pela equipe “e” da ordem do tipo “t” por unidade de tempo.
- O número de atendentes é igual a quantidade de equipes vinculadas ao posto técnico.
- A capacidade do sistema está diretamente relacionada com $\mu_{e,t}$, neste caso não só a quantidade de equipes é relevante para variar a capacidade de execução, mas principalmente quais as equipes estão vinculadas ao posto e quais ordens estão pendentes de execução.

Outra situação relevante em relação a taxa média de atendimento por unidade de tempo, é vinculada a Figura 19, onde essa se relaciona com a abordagem de Programação de turnos de trabalho proposta por Fitzsimmons, James A. e Fitzsimmons, Mona J. (2010), onde a demanda de serviço emergencial oscila de maneira elevada ao longo das 24 horas.

Assim, a determinação de turnos de trabalho que visem distribuir os horários das equipes em busca da maximização de disponibilidade em períodos de maior frequência de falhas é um modo de atender de maneira mais rápida os atendimentos, aumentando o nível de serviço, conforme o exemplo da Figura 19, quando nos horários mais críticos existirão mais equipes disponíveis, reduzindo o tempo de espera de um cliente na fila.

Sendo que o equilíbrio entre demanda e capacidade se dá na relação entre λ_t e $\mu_{e,t}$, a necessidade informada para λ_t ser considerado em unidade de tempo diferente para cada tipo de ordem está vinculada com a expectativa de tempo ao equilíbrio do sistema. Para ordens emergenciais, espera-se que o sistema entre em equilíbrio dentro do mesmo dia, já para ordens programáveis, a expectativa é de dias para o sistema entrar em equilíbrio. Isso faz com que os problemas tenham que ser analisados com uma granularidade diferente.

Figura 19 - Cobertura da demanda emergência pelo recurso disponível



Fonte: Desenvolvido pelo autor.

3.2.4 Disciplina da fila

Devido a existência de diferentes tipos de ordens de serviços, e essas possuem impactos e prazos diferentes, a disciplina da fila é do tipo com prioridade, pois conforme o tempo passa, as ordens vão se deslocando mais para próximo do seu prazo de vencimento até o momento que ultrapassam esse prazo, mas a relevância dessa ordem não está vinculada diretamente ao seu prazo e sim em relação aos prazos e tipos das demais ordens pendentes dentro do mesmo posto técnico e a capacidade do sistema em executar esses serviços.

Conforme já discutido anteriormente, uma parcela da demanda pendente não será executada no dia, essa situação em conjunto com a existência de ordens com tipos, custos e localização diferentes, faz com que algum tipo de priorização seja atribuído à ordem de serviço, isso porque em um processo de roteirização normal o indicador mais relevante é a localização das ordens, e assim esse cria rotas que minimizam somente o deslocamento.

Deste modo, a relação entre capacidade e demanda afeta diretamente o nível de priorização entre as ordens e consequentemente o processo de sequenciamento das demandas no roteiro das equipes. Quanto maior o GAP entre a capacidade e demanda, maior deve ser a criticidade na priorização de serviços em detrimento da minimização de deslocamento.

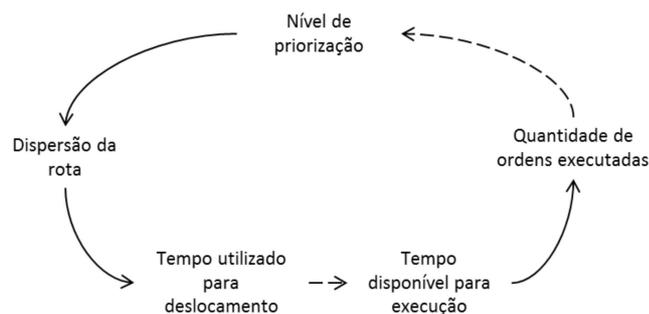
Como exemplo, poderia supor que em um posto técnico a quantidade de demandas pendentes somasse um tempo de serviço igual a 12,5 horas, porém existisse somente uma equipe disponível com 7,8 horas de turno previsto. Consequentemente, nem todas as ordens

seriam alocadas na rota programada da equipe, zerando a probabilidade de que todas as ordens serem executadas.

Porém, se somente nesse exemplo, descartarmos as ordens não atribuídas às rotas e considerarmos somente as alocadas, e que essas possuíssem o mesmo tipo, com o mesmo prazo, o processo de roteirização poderia utilizar como indicador de roteirização a maior utilização do recurso, sendo este a minimização do deslocamento. Porém, como as ordens possuem criticidades e prazos distintos em conjunto com o risco das ordens alocadas mais ao final da rota sofrerem impacto de postergação, pelo impacto dos fatores mencionados anteriormente, e não serem executadas, torna-se relevante que no processo de roteirização programada, as ordens mais críticas sejam inseridas no início das rotas das equipes, pois quanto mais ao final do turno maior é a probabilidade de esta ser postergada, “transbordando” para fora do turno da equipe, e assim não ser executada.

Contudo, essa decisão provoca a elevação do indicador de deslocamento para priorizar ordens mais críticas, o aumento do tempo de deslocamento diminui a capacidade de execução de serviços, decorrendo em um ciclo reforçador de piora no desempenho de execução de serviços, quando a quantidade de serviços pendentes está acima do limite de capacidade de execução.

Figura 20 - Ciclo reforçador do nível de priorização



Fonte: Desenvolvido pelo autor.

A Figura 20 representa o fluxo reforçador do processo de priorização, onde: Quanto maior é o nível de priorização maior será a dispersão da rota, quanto maior a dispersão da rota maior será o tempo utilizado para deslocamento, quanto maior o tempo utilizado para deslocamento menor é o tempo disponível para execução de tarefas, quanto menor o tempo disponível para execução das tarefas menor é a quantidade de ordens executadas e quanto

menor a quantidade de ordens executadas maior deverá ser o nível de priorização, considerando uma geração constante.

Devido a isso, o estudo da modelagem dessa dinâmica é extremamente relevante para a minimização dos impactos negativos da existência de uma elevada quantidade de ordens pendentes, com uma quantidade restrita de equipes e a existência de ordens com níveis de serviço diferentes.

3.2.5 Priorização no despacho programado

Para minimizar o impacto dessa variabilidade e aumentar a execução de ordens críticas, os serviços são divididos em quatro prioridades diferentes: 0, 1, 2 e 3. Sendo que, quando menor o valor numérico maior é a prioridade ao negócio. Deste modo, o algoritmo baseado em heurísticas de roteamento, efetua a roteirização garantindo que as ordens de maior prioridade sempre estejam a frente de ordens de menor prioridade. Porém, para minimizar o impacto da variação de execução de ordens programadas entre as equipes, o processo procura equilibrar a distribuição de ordens de mesma prioridade entre as equipes, conforme Figura 21:

Figura 21 - Sequência da rota programada

Equipe	1	2	3	4	5	6	7
A	0	0	0	1	1	2	2
B	0	0	0	1	2	2	2
C	0	0	1	1	1	1	2
D	0	0	1	1	1	1	2
E	0	0	0	1	1	2	2

Fonte: Desenvolvido pelo autor.

Assim, conforme as variações ocorrem ao longo do dia, impactando na execução da rota programada (conforme exemplo da Figura 22), as ordens mais críticas continuam com chances de serem executadas.

Garantido o equilíbrio de ordens e a sequência de prioridades, o algoritmo utiliza o tempo de espera como indicador para melhorar o roteamento.

Assim, quando o algoritmo que efetua o sequenciamento das ordens, recebe a instância com as informações de ordens pendentes e equipes disponíveis do posto técnico ele executa a seguinte regra para equilibrar as ordens de mesma prioridade nas equipes:

AgruparOrdens (Instância, N, T)

1. $it = 1$;
2. $sol = \text{ConstroiRotas}(\text{Instância}, N, T)$;
3. $sol = \text{MelhoraRotas}(\text{Instância}, N, T, sol)$;
4. Retorna sol ;

ConstroiRotas (Instância, N, T, sol)

1. Se $\text{TemOrdensP0}(\text{Instância})$ Então
2. $sol = \text{CriaRotasComP0}(\text{Instância}, N, T, sol)$;
3. $sol = \text{CriaRotasComP123}(\text{Instância}, N, T, sol)$;
4. Retorna (sol);

CriaRotasComP0 (Instância, N, T, sol)

1. $it = 1$;
2. Enquanto $it \leq nroOrdens(\text{Instância})$ e $\text{TemNosNaoAlocadosP0}(\text{Instância}, sol)$ Faça
3. Se $\text{NúmeroMedianas}(sol) < \text{NúmeroEquipes}(\text{Instância})$ Então
4. $sol = \text{CriaMediana}(\text{Instância}, sol)$;
5. $sol = \text{DesignaNós}(\text{Instância}, sol)$;
6. $sol = \text{RevisaPosiçãoMedianas}(\text{Instância}, sol)$;
7. $sol = \text{RevisaDesignaçãoNós}(\text{Instância}, sol)$;
8. $it = it + 1$;
9. Retorna (sol);

CriaRotasComP123 (Instância, N, T, sol)

1. $it = 1$; $p = 1$;
2. Enquanto $it \leq \text{NumeroOrdens}(\text{Instância})$ Faça
3. $nos = \text{SelecionaNosNaoAlocados}(\text{Instância}, sol, p)$;
4. Se $\text{Número}(nos, p) > 0$ Então
5. Se $\text{NúmeroMedianas}(sol) < \text{NúmeroEquipes}(\text{Instância})$ Então
6. $sol = \text{CriaMediana}(\text{Instância}, sol)$;
7. $sol = \text{DesignaNós}(\text{Instância}, sol, p)$;
8. $sol = \text{RevisaPosiçãoMedianas}(\text{Instância}, sol)$;
9. $sol = \text{RevisaDesignaçãoNós}(\text{Instância}, sol)$;
10. Senão
11. $sol = \text{RemoveDesignaçãoNós}(\text{Instância}, sol, p)$;
12. $sol = \text{DesignaNós}(\text{Instância}, sol, p)$;
13. $sol = \text{RevisaPosiçãoMedianas}(\text{Instância}, sol)$;
14. $sol = \text{RevisaDesignaçãoNós}(\text{Instância}, sol)$;
15. Se $p < 4$ Então
16. $p = p + 1$;
17. $it = it + 1$;
18. Retorna (sol);

MelhoraRotas (Instância, N, T, sol)

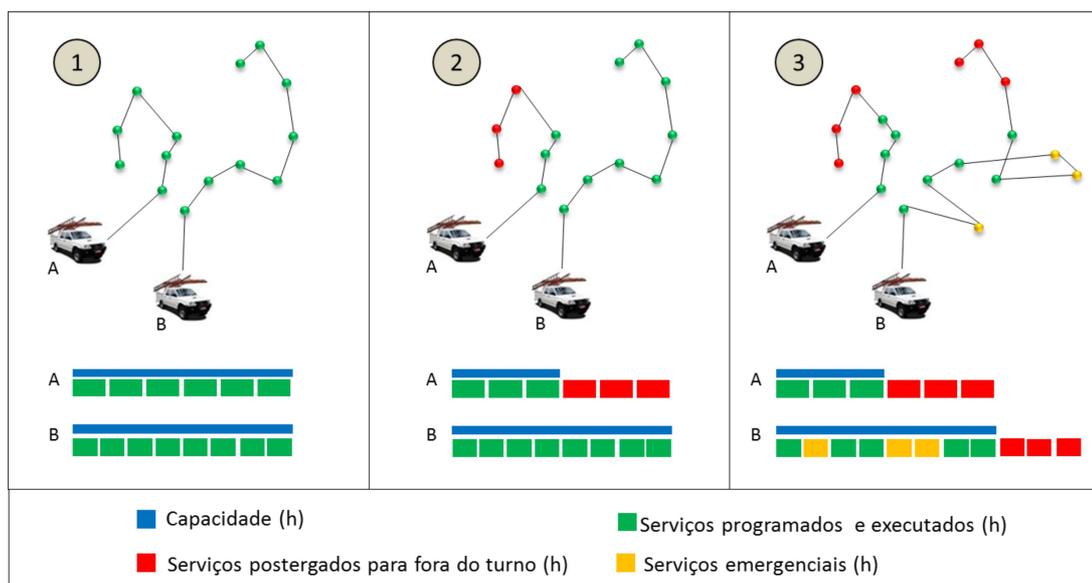
1. $it = 1$;
2. Para cada rota r de sol Faça
3. $sol = \text{BuscaLocal2opt}(\text{Instância}, sol, r)$;
4. $sol = \text{Avalia}(\text{Instância}, sol)$;
5. Retorna (sol);

Outro ponto relevante é o equilíbrio de ordens prioritárias entre as equipes. No processo de roteirização o algoritmo respeita um carregamento máximo para ordens com prioridade 0 e 1, sendo respectivamente 55% e 65% do turno disponível de cada equipe, isso

faz com que, mesmo que o indicador de roteirização tenda a incluir uma quantidade elevada de ordens para uma mesma equipe devido à pequena distância desse grupo de ordens entre si, ele está restrito a respeitar esses limites, obrigando a divisão dessas ordens entre duas ou mais equipes, como pode ser visualizado na Figura 21. Porém, caso a quantidade de ordens seja elevada a ponto de não conseguir respeitar esses limites, o algoritmo desconsidera essa restrição, possibilitando aumentar a roteirização com base no deslocamento, pois entende que o cenário é crítico e que devido a isso, deve-se maximizar a utilização das equipes.

Deste modo, as regras definidas nessa heurística de roteamento visam garantir que ordens mais críticas estejam em uma posição inicial na rota, aumentando a sua probabilidade de execução. Possibilitando que variabilidades individuais afetem o mínimo possível na execução de ordens críticas, conforme pode ser verificado na Figura 22:

Figura 22 – Exemplo das etapas de planejamento e execução de rotas



Fonte: Desenvolvido pelo autor.

Na Figura 22, o cenário 1 representa o calendário previsto de duas equipes conforme a programação efetuada pelo algoritmo de despacho, nessa visão todas as ordens seriam efetuadas ao longo do dia de trabalho. No cenário 2 a equipe A tem a sua disponibilidade reduzida pela metade, com essa diminuição de capacidade 3 ordens foram postergadas para após o fim do turno da equipe, impossibilitando a execução. No cenário 3 a equipe B recebe ao longo do turno de trabalho 3 atendimentos de emergência, esses atendimentos prioritários postergam mais três ordens para fora do turno e, conseqüentemente, também não serão

executadas. Assim fica evidente que quanto mais no início da rota uma ordem se encontra maiores são as chances desta ser executada.

3.2.6 Definição das prioridades de serviço

As ordens de serviço, como já dito anteriormente, possuem diferentes tipos e conseqüentemente níveis de serviço esperado distintos. A primeira diferença é entre ordens programáveis e emergenciais. Sendo que, as ordens emergenciais sempre são mais prioritárias e assim interferem na rota das ordens programáveis, para minimizar o impacto dessa interferência do despacho de tempo real no despacho programado, a definição de prioridade se torna relevante, e assim possibilitar o funcionamento adequado do processo, conforme caracterizado no item anterior.

Essa priorização se dá sobre as ordens programáveis e se comporta com as seguintes regras:

- a) As ordens são divididas em grupos, conforme demonstrado na Tabela 8;

Tabela 8 - Grupos de ordens programadas

Nome Ordem de serviço	Grupo
Ligação e instalação da medição	Regulada
Religie de reforma	Regulada
Mudança do ponto de medição	Cliente
Troca de medidor baixa tensão	Cliente
Aferição de medidor	Cliente
Retirada de equipamento de medição	Cliente
Corte CP ou Ramal	AES
Encerrado	AES
Manutenção de rede equipe tipo - A	Técnica
Manutenção de rede equipe tipo - C	Técnica

Fonte: Desenvolvido pelo autor.

- b) Os grupos podem ser hierarquizados em relação à sua importância, conforme demonstrado na Tabela 9.

As ordens se encontram na Tabela 9 em uma ordem de criticidade do maior para a menor. Porém, ocorre uma sobreposição entre essas criticidades conforme os prazos de cada ordem, valores envolvidos e riscos técnicos da não execução da tarefa. Deste modo, conforme a ordem da tabela acima e alguns atributos dos serviços as prioridades são definidas.

Para analisar o comportamento do processo de roteirização, com a influência da prioridade, para aumentar o nível de serviço de ordens mais críticas, se faz necessário analisar uma instância já resolvida pelo processo.

Tabela 9 – Hierarquia de prioridade entre os grupos

Grupo	Impacto ao vencer	Prioridade
Regulada	Multas, insatisfação cliente (cliente sem energia), ordem controlada pela ANEEL	0
Cliente	Multas e insatisfação cliente (cliente com energia)	1
AES	Inadimplência e perdas	2
Técnica	Aumento da probabilidade de ocorrer uma falha no fornecimento de energia	3

Fonte: Desenvolvido pelo autor.

A seguir é apresentada a instância do posto técnico 2074, da supervisão operacional de Novo Hamburgo do dia 12/09/2014. Nessa instância existiam 2 equipes com disponibilidade prevista de 8 horas cada e 104 demandas pendentes. Assim, a capacidade de execução seria 16 horas de disponibilidade de equipe contra 161 horas de tempo de serviço (sem considerar o deslocamento envolvido na rota), deste modo, se as ordens não são todas iguais o processo de priorização é de extrema relevância.

Após, o processo de priorização, as ordens de serviço foram divididas em 4 níveis de prioridade: 0, 1, 2 e 3. Ficando com a distribuição de acordo com a Tabela 10:

Tabela 10 - Prioridade das ordens de serviço

Prioridade	Quantidade de ordens	Somatório tempo de serviço (h)
0	3	1,3
1	4	4,1
2	5	3,1
3	92	152,2

Fonte: Desenvolvido pelo autor.

Note que as ordens mais prioritárias se apresentam em menor número e com tempos de execução em uma proporção muito mais próxima da capacidade das equipes, esse fator faz com que o despacho seja muito mais assertivo em função da execução das ordens mais prioritárias.

Com essa distribuição prioridades de serviços, o algoritmo de roteamento retornou com as soluções de rotas contidas na Tabela 11 e na Tabela 12:

Tabela 11 - Roteamento programado para a equipe 195

Equipe	Ordem	Número OS	Prioridade	Deslocamento	Serviço	Acumulado	Percentual acumulado
195	1	20546103	0	0,58	0,92	1,49	20%
195	2	20547333	0	0,08	0,17	1,74	23%
195	3	20532353	1	0,21	1,33	3,28	43%
195	4	366569	3	0,21	0,75	4,24	56%
195	5	353495	3	0,05	0,75	5,04	66%
195	6	358893	3	0,06	0,75	5,84	77%
195	7	50185885	3	0,11	0,33	6,29	82%
195	8	20548241	3	0,22	0,17	6,67	87%
195	9	20548344	3	0,01	0,17	6,85	90%
195	10	60440475	3	0,03	0,33	7,21	94%
195	11	20548540	3	0,25	0,17	7,63	100%

Fonte: Desenvolvido pelo autor.

Tabela 12 - Roteamento programado para equipe 121

Equipe	Ordem	Número OS	Prioridade	Deslocamento	Serviço	Acumulado	Percentual acumulado
121	1	20547088	0	0,45	0,17	0,62	8%
121	2	20547717	1	0,07	0,92	1,61	21%
121	3	20548035	1	0,18	0,92	2,71	35%
121	4	20548057	1	0,04	0,92	3,67	48%
121	5	20544042	2	0,11	0,50	4,28	56%
121	6	60440476	2	0,06	0,33	4,67	61%
121	7	17604620	2	0,06	0,50	5,23	68%
121	8	20544807	2	0,06	1,33	6,62	87%
121	9	17604619	2	0,07	0,50	7,19	94%
121	10	20548134	3	0,25	0,17	7,61	100%

Fonte: Desenvolvido pelo autor.

Assim, identifica-se que as ordens mais críticas foram inseridas mais no início das rotas das equipes, deste modo, se a equipe 121 efetuar 8% da rota planejada e a equipe 195 efetuar 23%, todas as ordens com prioridade 0 serão concluídas no dia, conforme o programado. Esse processo aumenta a gerência sobre a execução das ordens pendentes, do ponto de vista que o analista tem condições de interferir no resultado da roteirização, incluindo informações que determinem a criticidade das ordens.

3.3 DESENVOLVIMENTO DO MODELO DE SIMULAÇÃO

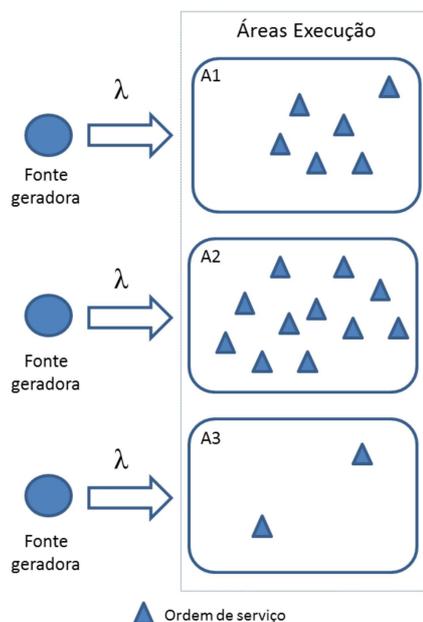
Baseado na análise dos itens anteriores propõe-se o desenvolvimento de um modelo com o viés de propiciar a simulação de ingresso e execução de serviços, baseado em um sistema de fila, com o objetivo de, ao final do trabalho, compararmos cenários com diferentes estratégias de gerenciamento de serviço.

3.3.1 Modelagem do problema

Assim, conforme descrito anteriormente, a área de execução é a menor divisão geográfica da empresa, geralmente separa regiões com dificuldade de acesso mútuo, mas também utilizada para separar bairros ou regiões distintas. Partindo desse cadastro, define-se a fonte de entrada como particular para cada área de execução, possibilitando a vinculação do serviço a ela (área de execução), conforme Figura 23.

Essa decisão permite que as ordens de serviço, uma vez vinculadas à área de execução, não restrinjam a formação dinâmica dos postos técnicos, sendo que a fila de sistema, pode ser considerada em função da área de execução ou do conjunto de áreas que formam o posto técnico.

Figura 23 – Fonte de entrada conectada a área de execução



Fonte: Desenvolvido pelo autor.

No caso da Figura 24, é demonstrado as diferentes conformações de postos técnicos. Onde: no cenário 1 existem dois postos técnicos, o primeiro formado pelas áreas A1 e A2 e o segundo formado pela área A3. Por sua vez, o cenário 2, também possui dois postos técnicos, o primeiro formado pela área A1 e o segundo formado pelas áreas A2 e A3. Já o cenário 3, possui 3 postos técnicos, o primeiro formado pela área A1, o segundo pela área A2 e o terceiro pela área A3. Assim, os clientes que se encontram na fila de atendimento no instante t pode ser definido como:

$$L_{pt} = \sum_{i=1}^I L_{A_{it}} \quad (49)$$

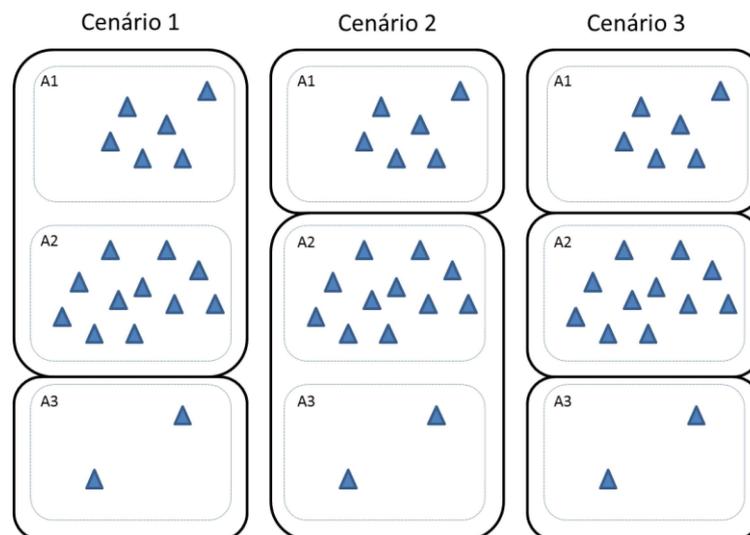
Onde:

L_{pt} : Quantidade de clientes no sistema de fila do posto p no instante t ;

$L_{A_{it}}$: Quantidade de clientes no sistema de fila da área i no instante t ;

I : Quantidade de áreas de execução que formam o posto p .

Figura 24 - Formação do posto técnico e definição da fila

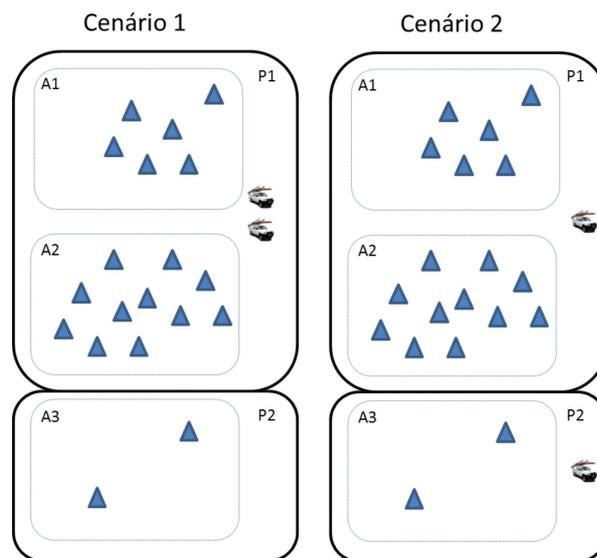


Fonte: Desenvolvido pelo autor.

Entretanto, como as equipes são vinculadas aos postos técnicos e não existe garantia que existam equipes atreladas a todos os postos, ocorre que a taxa de execução de ordens de serviço é variável tanto em função do desempenho na execução ou da característica do serviço, mas principal da existência ou não de equipes no posto. Conforme Figura 25, note

que no cenário 1 existem duas equipes vinculadas ao posto técnico P1, porém no posto P2 não existe equipe alguma. Logo, indiferente da capacidade das equipes e o seu desempenho de execução, no posto P2 a taxa de saída será igual a 0, o que não ocorre no cenário 2, onde existe equipe em ambos os postos técnicos.

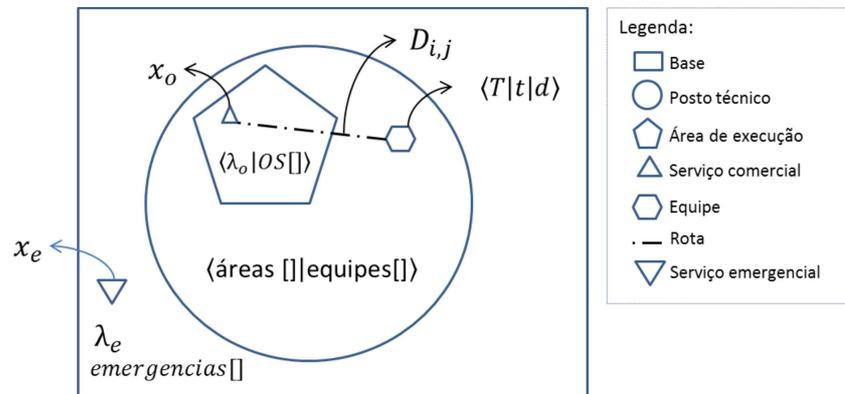
Figura 25 - Relação posto técnico, área de execução, recurso e demanda.



Fonte: Desenvolvido pelo autor.

Levando em conta os fatores levantados até o momento, pode-se definir uma distribuição de responsabilidades sobre os objetos que compõem o problema, de modo a hierarquizar o modo como o problema pode ser sistematizado, permanecendo aderente ao mundo real, conforme demonstrado na Figura 26 e detalhado na Tabela 13, onde está definida a seguinte hierarquia: os serviços comerciais ingressam com uma determinada taxa na área a qual seu cliente está vinculado gerando um conjunto de ordens pendentes (fila), as áreas de execução estão vinculadas a um posto técnico, que, por sua vez, pode ser caracterizado por um conjunto de áreas de execução e equipes. Cada equipe possui um conjunto de atributos que interfere na capacidade de execução do posto técnico. O conjunto de postos técnicos forma uma base, cada base possui uma taxa de ingresso de atendimentos emergenciais gerando uma fila de atendimentos pendentes, que serão rateados entre os postos técnicos e as equipes vinculadas a esses, que, conseqüentemente, irá afetar a disponibilidade das equipes na execução de ordens programadas, pois terão de atender a demanda emergencial.

Figura 26 - Relação entre objetos e variáveis



Fonte: Desenvolvido pelo autor.

Todo objeto que compõe o problema foi modelado em função de garantir a manutenção dos dados ao longo do processo de simulação. Sendo que a falha no processo de atribuição de responsabilidade de cada um poderia causar um comportamento inadequado do processo, dificultando ao analista a localização da falha, partindo do princípio que, no momento da análise, este não sabe se o comportamento é uma representação do mundo real, efeito da simulação ou de uma falha de semântica no processo de desenvolvimento do software.

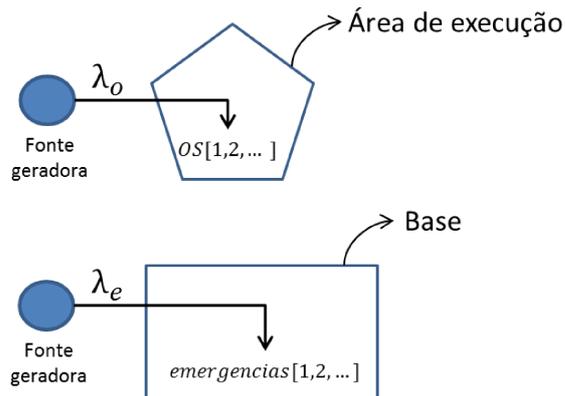
Tabela 13 - Atributos dos componentes do problema

Objeto	Variável	Descrição
Base	λ_e	Estatísticas de ingresso de emergência
	emergências []	Conjunto de emergências pendentes na base
Posto Técnico	áreas []	Conjunto de áreas de execução do posto técnico
	equipes []	Conjunto de equipes do posto técnico
Área de execução	λ_o	Estatísticas de ingresso de ordens comerciais do tipo o
	OS []	Conjunto de ordens comerciais da área de execução
Comerciais	x_o	Estatísticas de tempo de execução da ordem do tipo o
Equipe	T	Tempo de turno previsto diário
	t	Estatísticas de variação de turno previsto
	d	Estatísticas de variação de deslocamento
Rota	$D_{i,j}$	Tempo de deslocamento previsto entre o ponto i até o j
Emergências	x_e	Estatísticas de tempo de atendimento da ordem das ordens

Fonte: Desenvolvido pelo autor.

Assim, podemos desenvolver uma estrutura lógica que permite determinar duas fontes geradoras básicas: A primeira vinculada à área de execução, onde diariamente são geradas as ordens comerciais, e a segunda está relacionada com a base operacional que, por sua vez, gera as ordens emergenciais, conforme Figura 27.

Figura 27 - Fontes geradoras (entrada)



Fonte: Desenvolvido pelo autor.

Deste modo, as ordens de serviço comerciais e emergenciais acabam aguardando em filas internas aos componentes: área de execução e Base, respectivamente. Nesse caso, os vetores OS e emergências representam o conjunto de ordens em fila no sistema.

Por outro lado, a taxa de saída do sistema de filas, representada por μ no modelo apresentado por Hillier e Lieberman (2013), está vinculada diretamente ao componente posto técnico, pois é esse que possui os recursos necessários (equipes) para dar vazão à execução de serviços, cada uma com seu turno de trabalho programado (T) e sua estatística de variação de turno previsto (t), assim para definir o turno efetivo de cada equipe deve-se respeitar a equação:

$$T_{ef} = T \cdot t \quad (50)$$

Onde:

T_{ef} : Turno efetivo

T : Turno previsto

t : Variação de turno simulada

A partir da definição do turno efetivo das equipes é que podemos iniciar um processo para calcular a execução de serviços. A primeira etapa desse processo é definir a alocação dos atendimentos emergências, já que esses são prioritários em relação aos atendimentos comerciais e assim consumirão tempo de turno das equipes, reduzindo a parcela de turno disponível para a execução das ordens planejadas (comerciais). Assim, baseado nas estatísticas de ingresso da base e dia da semana, é simulado a quantidade de atendimentos emergenciais que ingressarão no dia. A Figura 28 tem por objetivo esclarecer o modo que, a partir da definição de uma quantidade de emergências simuladas para o dia da base, como o sistema irá efetuar o rateio entre os postos técnicos (cenário A) e, no momento seguinte, entre as equipes de cada posto técnico (cenário B).

No cenário A utiliza-se a equação abaixo para definir a quantidade de emergência para cada posto técnico:

$$E_p = E_T \cdot \frac{Q_p}{Q_T} \quad (51)$$

Onde:

E_p : Quantidade de emergências do posto técnico p ;

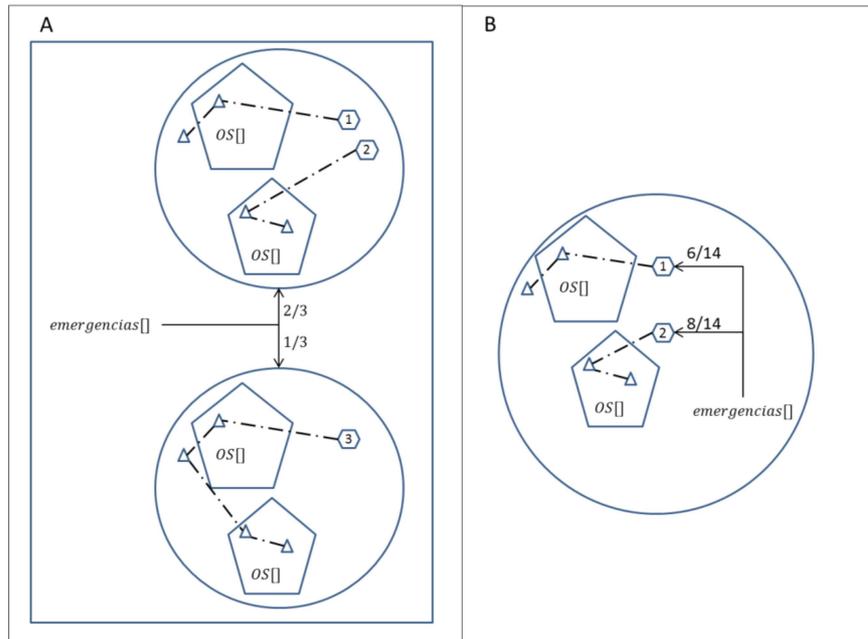
E_T : Quantidade de emergências pendentes na base;

Q_p : Quantidade de equipes no posto técnico p ;

Q_T : Quantidade de equipes da base;

Essa equação determina a parcela de atendimentos para cada posto técnico, atribuindo emergências em quantidade proporcional à quantidade de equipes do posto técnico em relação à quantidade existente na base.

Figura 28 - Atribuição de atendimentos de emergência



Fonte: Desenvolvido pelo autor.

Definida a quantidade de emergências para cada posto técnico, migra-se para o passo do cenário B (Figura 28), onde se utiliza a equação abaixo para definir a quantidade de emergências para cada equipe:

$$E_e = E_p \cdot \frac{T_e \cdot t_e}{\sum_{i=1}^N T_i \cdot t_i} \quad (52)$$

Onde:

E_e : Quantidade de emergências para a equipe e ;

E_p : Quantidade de emergências pendentes no posto técnico;

T_e : Turno previsto para equipe e ;

t_e : Variação de turno simulado para a equipe e ;

T_i : Turno previsto para equipe i ;

t_i : Variação de turno simulado para a equipe i ;

N : representa a quantidade total de equipes do posto técnico p .

No cenário B (Figura 28) existem duas equipes disponíveis, a equipe 1 possui turno efetivo de 6 horas e a equipe 2 de 8 horas. Assim, a equipe 1 receberá 43% das emergências pendentes no posto técnico e o restante caberá a outra equipe. Portanto, cada equipe recebe

uma quantidade de emergências proporcional ao seu turno efetivo em razão ao somatório dos turnos efetivos do posto técnico.

Definida a quantidade de emergências para cada equipe disponível no dia, o sistema simula o tempo utilizado para atender a cada demanda de emergência a partir das estatísticas de tempo de atendimentos efetuados na base. O somatório dos tempos dos atendimentos simulados para cada equipe irá reduzir a capacidade da equipe, assim a equação que determina o tempo utilizado em emergência para cada equipe é descrita da seguinte forma:

$$TE_e = \sum_{i=1}^I te_{i,e} \quad (53)$$

Onde:

TE_e : Tempo total em atendimento de emergência da equipe e ;

$te_{i,e}$: Tempo simulado para o atendimento da emergência i vinculado a equipe e ;

I : Total de ordens vinculadas à equipe e ;

Assim, o tempo que realmente a equipe possui para executar a rota planejada pode ser determinado como o produto entre o turno planejado e a simulação da variação do turno subtraído do tempo total em atendimento de emergência da equipe. Entretanto, os tempos de deslocamento e execução da rota planejada não são constantes, sendo que o processo de execução de rotas planejadas pode ser representado da seguinte maneira:

$$\sum_{e=1}^E \sum_{i=1}^I d_{e,(i-1,i)} + \sum_{i=1}^I x_{o_i} \leq \sum_{e=1}^E (T_e \cdot t_e - TE_e) \quad (54)$$

Onde:

$d_{e,(i-1,i)}$: Variação do tempo de deslocamento simulado da equipe e para deslocamento da ordem planejada $i-1$ até a ordem i ;

x_{o_i} : Tempo simulado para o tipo de serviço o da ordem planejada i ;

T_e : Turno previsto para equipe e ;

t_e : Variação de turno simulado para a equipe e ;

TE_e : Tempo total em atendimento de emergência da equipe e ;

E : Conjunto de equipes do posto técnico;

I : Total de ordens vinculadas à equipe e ;

A seguir é caracterizado o que cada tarefa e sub-processo da Figura 29 efetua dentro do sistema de simulação:

3.3.2.1 Define parâmetros

Esse processo é responsável por definir a configuração básica do processo de simulação. Sendo que as principais configurações são:

- i. Definição da base que ocorrerá a simulação;
- ii. Definição da quantidade de iterações;
- iii. Definição da quantidade de dias que serão simulados em sequência para cada iteração da simulação;

3.3.2.2 Simula ingresso de emergência do dia

A simulação do ingresso de emergências é baseada no histórico de ingressos de serviços emergenciais. Sendo que a cada dia simulado uma nova amostra de ingresso é gerada e somada à quantidade de ordens emergenciais já pendentes no sistema de filas.

Com a definição das ordens emergenciais pendentes, a aplicação distribui as demandas entre os postos técnicos conforme o cenário A da Figura 28.

3.3.2.3 Carrega demandas pendentes do posto

Consulta o banco de dados coletando todas as demandas comerciais e técnicas pendentes, sujeito a restrição do posto técnico que está sendo processado.

3.3.2.4 Carrega equipes com calendário do posto

Consulta o banco de dados coletando todas as equipes com calendário previsto para o dia simulado, sujeito a restrição do posto técnico que está sendo processado.

3.3.2.5 *Executa algoritmo de despacho*

O algoritmo de despacho é instanciado com o conjunto de ordens pendentes e de equipes disponíveis definidos nos processos anteriores. Assim, uma rota é criada para cada equipe seguindo as regras discutidas no item 3.2.5.

3.3.2.6 *Simula a disponibilidade da equipe*

De posse do calendário previsto para cada equipe:

- Simula-se uma amostra da variação de turno para cada equipe;
- Define-se o rateio dos serviços emergenciais pendentes do posto técnico entre as equipes vinculadas a esse posto;
- Simulam-se os tempos de atendimento para cada um dos serviços emergenciais.

3.3.2.7 *Define disponibilidade para rota planejada*

A partir das informações calculadas no processo anterior é definido o turno restante das equipes para a execução da rota planejada.

3.3.2.8 *Simula tempos de deslocamento e serviço*

São simulados os tempos de deslocamento e execução de cada ordem programada, seguindo a sequência da rota, até que o tempo acumulado de deslocamento e serviços ultrapasse o tempo restante para execução da rota planejada (definida no item anterior).

3.3.2.9 *Grava serviços executados*

Atualiza para concluído o estado dos serviços planejados e emergenciais que não “transbordaram” para fora do turno da equipe, as demais ordens continuam pendentes.

3.3.2.10 Simula ingresso de serviços comerciais

Ao fim de cada dia do processo, é simulada a geração de serviços comerciais para cada área de execução estratificada por tipo de ordem de serviço, baseado no histórico de ingresso, conforme será descrito no item 3.4.

3.3.2.11 Calcula indicadores

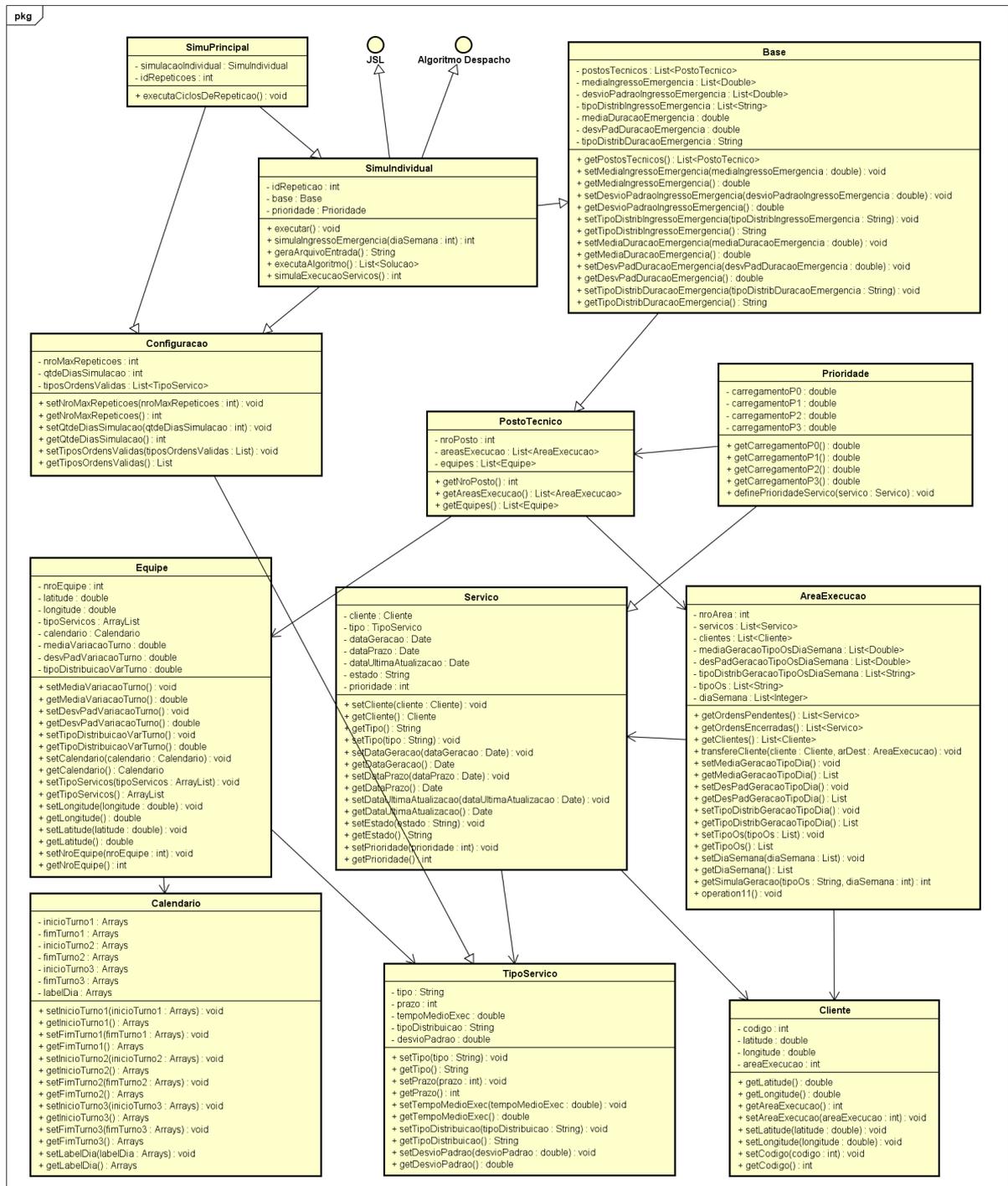
Para cada processo de simulação os indicadores elencados são calculados a partir dos dados gravados ao longo do processo de simulação.

As inter-relações entre os componentes do sistema desenvolvido podem ser visualizadas no diagrama UML contido na Figura 30. Nesse diagrama, buscou-se transcrever os componentes do problema (objetos), as suas principais funções (métodos) e a responsabilidades de cada componente com o processo, tal qual ocorre no problema real. Assim, o fluxograma contido na Figura 29 permeia as classes contidas nesse UML.

A estrutura UML foi desenvolvida em uma linguagem estrutura com o conceito de orientação a objetos (Java), buscando desenvolver a solução de forma mais próxima ao comportamento real, facilitando a evolução do aprendizado e a manutenção da aplicação.

Entretanto, essa estrutura é instanciada a partir de informações contidas em um banco de dados. Esse, por sua vez, possui informações de cadastro, configuração, dados históricos do processo e dados que serão persistidos ao longo da execução da simulação, possibilitando a coleta do desempenho da simulação.

Figura 30 - Diagrama UML do sistema



Fonte: Desenvolvido pelo autor.

O desenvolvimento da aplicação de simulação foi arquitetado para possibilitar qualquer tipo de dado em função do seu comportamento estatístico, sendo que indiferente do tipo de comportamento dos dados, o processo de simulação vai se comportar conforme a

configuração de tipo de distribuição e parâmetros calculados no processo de coleta e análise de dados.

3.4 ESTUDO DE CASO

Esse estudo de caso foi desenvolvido na empresa AES Sul Distribuidora Gaúcha de energia, concessionária de distribuição de energia elétrica no estado do Rio Grande do Sul. A sua área de concessão compreende 1/3 do estado, possuindo uma área de mais de 99.000 km², atendendo a 118 municípios e mais de 1.300.000 clientes.

O estudo e coleta de dados restringiram-se a supervisão operacional de Canoas, que fica localizada na região metropolitana de Porto Alegre. Essa supervisão possui 16% dos clientes e 10% das equipes de toda empresa.

3.4.1 Coleta e tratamento dos dados

Os dados coletados são do período entre 01/fevereiro/2016 a 31/maio/2016, totalizando 121 dias, esse período foi escolhido devido a não conter grandes eventos climáticos, o que poderia alterar o comportamento dos dados de forma acentuada. Os seguintes dados foram coletados:

- a) Histórico de ingresso de ordens comerciais por tipo de serviço e área de execução;
- b) Histórico de ingresso de atendimentos emergenciais por base;
- c) Histórico de turno previsto por equipe;
- d) Histórico de turnos efetivos por equipe;
- e) Histórico de tempos de deslocamento por equipe;
- f) Histórico de tempos de execução por tipo de OS e base;
- g) Histórico de tempos de atendimento de serviços emergenciais por base;
- h) Conformação dos postos técnicos atuais;
- i) Ordens de serviços pendentes no dia 31/maio/2016.

Os dados foram coletados diretamente do sistema de banco de dados legado da empresa, pois como todas as equipes recebem serviços de modo automatizado, através de smartphones, todos os dados ficam registrados em tempo real, facilitando a coleta e aumentando a veracidade da informação.

Ao longo do trabalho foram apresentados diversos atributos e comportamentos prováveis dos componentes do sistema, que serão testados para identificar o seu

comportamento real. Assim, as estatísticas que foram definidas como relevantes ao longo do trabalho estão definidas na Tabela 14, foram calculadas a partir dos históricos mencionados acima.

Tabela 14 - Estatísticas relevantes

Variável	Descrição	Tipo
$\lambda_{e,d}$	Estatísticas de ingresso de emergência por dia da semana	Discreto
$\lambda_{o,d}$	Estatísticas de ingresso de ordens comerciais do tipo o por dia da semana	Discreto
x_o	Estatísticas de tempo de execução da ordem do tipo o	Contínuo
t_q	Estatísticas de variação de turno previsto da equipe q	Contínuo
d_q	Estatísticas de tempo de deslocamento da equipe q	Contínuo
x_e	Estatísticas de tempo de atendimento da ordem das ordens de emergências da base	Contínuo

Fonte: Desenvolvido pelo autor.

Após a definição das variáveis e tipos de dados que elas possuem, foram analisados os comportamentos dos dados para identificar o tipo de distribuição para a qual pertencem, possibilitando uma aderência adequada no processo de simulação:

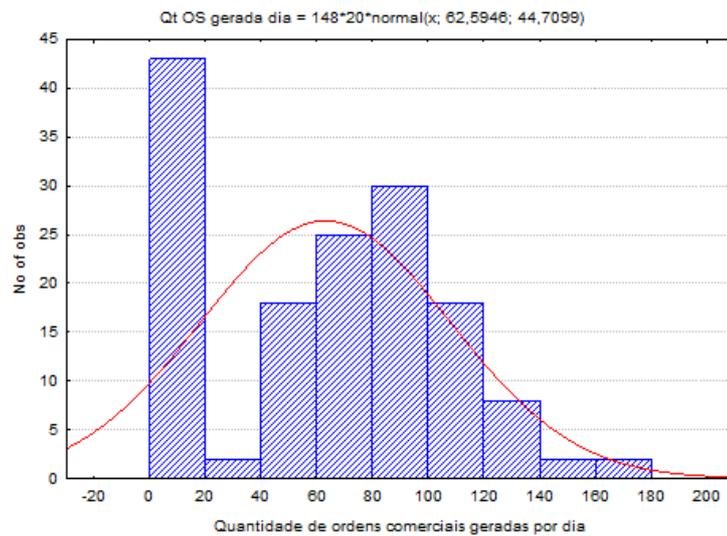
3.4.1.1 Estatísticas de ingresso de ordens comerciais por tipo e dia da semana ($\lambda_{o,d}$)

Os dados coletados sobre ingresso de atendimentos comerciais, quando agrupados por completo, apresentam uma distribuição com forma anormal, não se adequando a nenhum tipo de distribuição, conforme pode ser verificado na Figura 31.

Deste modo, identificou-se a necessidade de estratificação mais adequada da variável, assim devendo segregar os dados por atributos selecionados como relevantes. Nesse caso, o teste foi efetuado estratificando o ingresso por dia da semana e tipo de ordem, e assim correspondendo de forma mais adequada ao comportamento mencionado por Hillier e Lieberman (2013), que dados discretos de ingresso de clientes em um sistema de filas tendem a se comportar com uma distribuição de Poisson, como demonstrado no gráfico da Figura 32. Nesse gráfico está apresentada a distribuição dos ingressos de ordens do tipo ligação e instalação da medição que ocorreram em quartas-feiras.

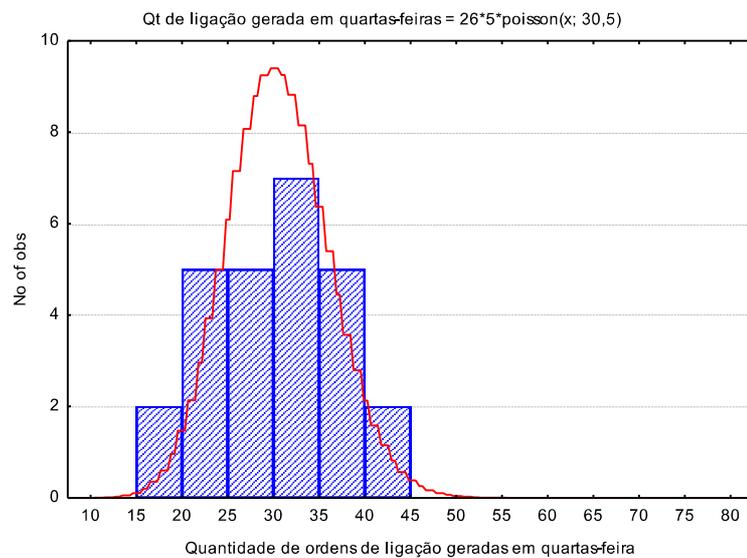
Assim, a variável $\lambda_{o,d}$ fora considerada como possuindo a distribuição de Poisson para todas as áreas de execução, e para cada processo de simulação apenas a média histórica será utilizada para definir o valor da amostra seguinte. Sempre retirando as amostras (dos dados coletados) que estiverem com uma variação superior a 2,5 desvios padrão de distância da média, classificando-as como um outlier.

Figura 31 - Histograma dos dados de ingresso de ordens comerciais



Fonte: Desenvolvido pelo autor.

Figura 32 - Histograma distribuição do ingresso de ordens de ligação em quartas-feiras

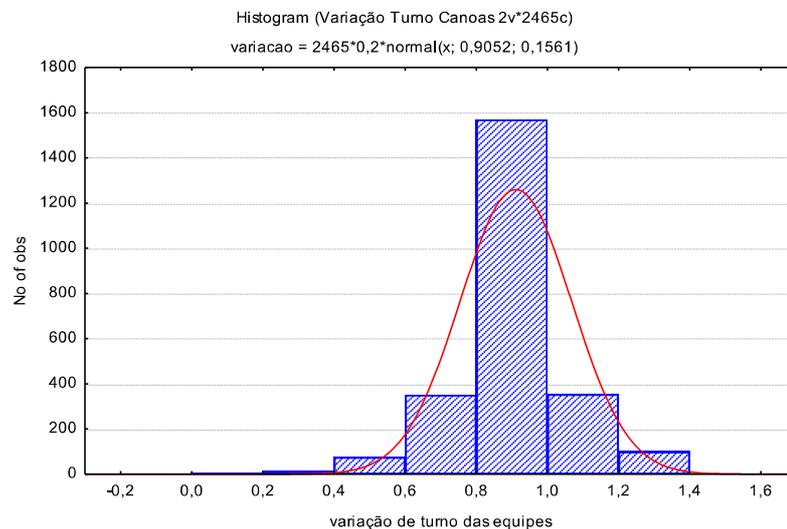


Fonte: Desenvolvido pelo autor.

3.4.1.2 Variação de turno das equipes

A variação de turno das equipes, que corresponde a razão entre o turno executado e o turno previsto, apresenta um comportamento normal quando os dados estão agrupados, conforme pode ser visualizado na Figura 33. Porém, os dados da variação dos turnos fazem parte dos números racionais positivos e não podem tomar valores negativos. Deste modo, a decisão no processo de simulação continuará sendo pela utilização da distribuição normal, porém para todo valor menor que zero será considerado zero.

Figura 33 - Histograma da variação de turno



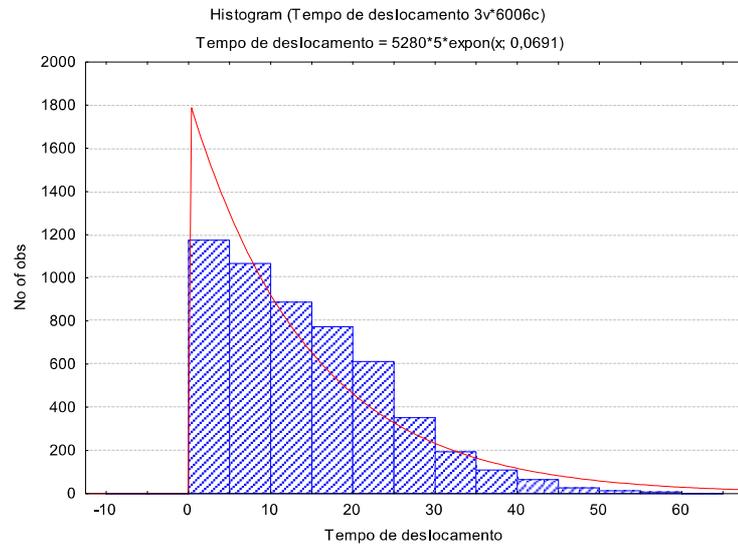
Fonte: Desenvolvida pelo autor.

Assim, cada equipe terá um comportamento igual à média da sua variação de turno, excluindo dados coletados com variação superior a 2,5 desvios padrões de distância da média, com uma distribuição de frequência do tipo normal.

3.4.1.3 Tempo de deslocamento

A distribuição do tempo de deslocamento se comporta conforme uma distribuição exponencial, conforme Figura 34. Essa distribuição será mantida para o processo de simulação. Sendo assim, cada equipe terá os parâmetros de média de deslocamento, calculado a partir do histórico excluindo dados com variação superior a 2,5 desvios padrão de distância da média e distribuição de frequência exponencial.

Figura 34 - Histograma do tempo de deslocamento das equipes

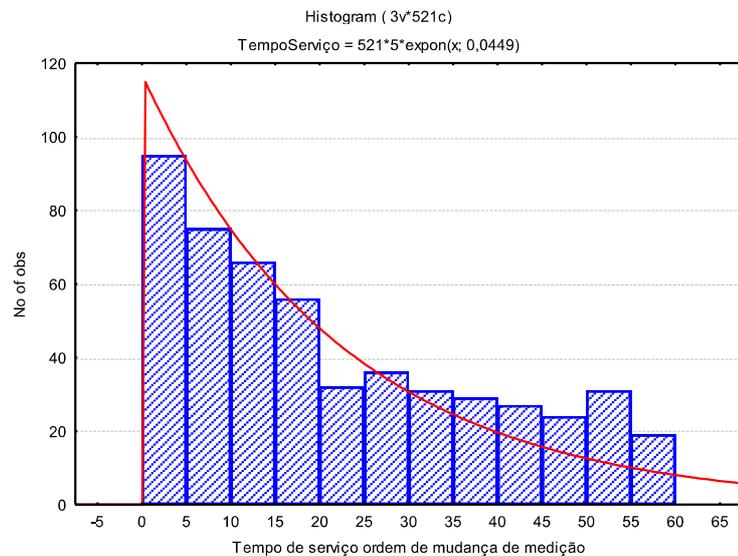


Fonte: Desenvolvido pelo autor.

3.4.1.4 Duração do serviço comercial

O tempo de execução de serviços comerciais, quando segregado por tipo de serviço apresenta uma distribuição muito próxima a exponencial conforme pode ser visualizado no gráfico da Figura 35, possibilitando seguir a orientação de Hillier e Lieberman (2013).

Figura 35 - Histograma com os tempos de serviço da ordem de mudança de local

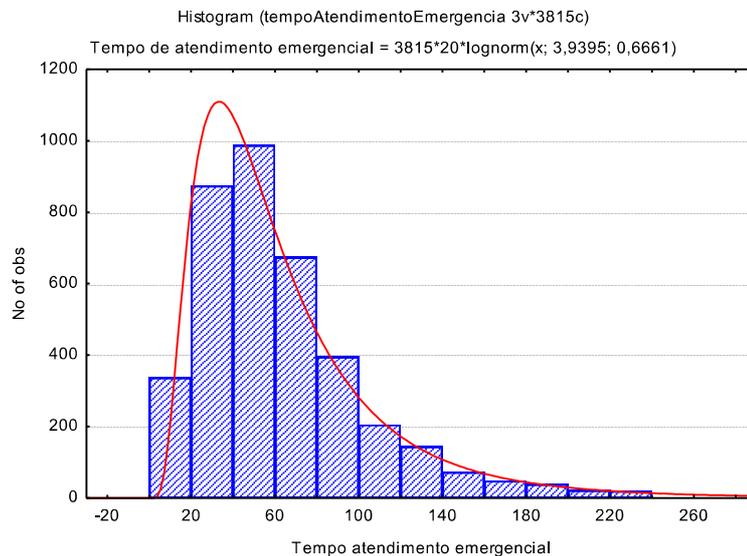


Fonte: Desenvolvido pelo autor.

3.4.1.5 Duração do atendimento de emergência

Os dados de tempo de atendimento de emergências se comportam com uma distribuição de frequência do tipo log-normal, conforme pode ser visualizado na Figura 36. Nesse caso, outro ponto que torna essa distribuição interessante é a garantia de representar a não negatividade dos dados. Sendo assim, a duração do atendimento de emergência será representada por uma distribuição do tipo log-normal, possuindo os parâmetros: média e desvio padrão. Sempre respeitando a exclusão de outliers.

Figura 36 - Histograma com os tempos de atendimentos de emergência



Fonte: Desenvolvido pelo autor.

Deste modo, a Tabela 14 pode ser editada, incluindo informações mais precisas sobre como cada tipo de dado será considerado dentro do processo de simulação. Assim, a Tabela 15 resume as definições da etapa de tratamento de dados.

Tabela 15 - Tipos de dados

Variável	Tipo	Distribuição	Parâmetros	Estratificação
$\lambda_{o,d}$	Discreto	Poisson	Média	Por tipo de ordem, área de execução e dia da semana
$\lambda_{e,d}$	Discreto	Poisson	Média	Por base e dia da semana
x_o	Contínuo	Exponencial	Média	Por tipo de ordem e base
t_q	Contínuo	Exponencial	Média	Por equipe
d_q	Contínuo	Exponencial	Média	Por equipe
x_e	Contínuo	Log-normal	Média e desvio padrão	Por base

Fonte: Desenvolvido pelo autor.

3.4.2 Cenários simulados

Para o estudo de caso em questão foram simulados 4 cenários diferentes com o propósito de analisar o impacto no nível de serviços para cada panorama, os indicadores escolhidos para a análise são:

- Quantidade média de ordens efetuadas fora de prazo, por grupo de ordem de serviço, conforme a Tabela 8;
- Quantidade média de ordens de serviços fora de prazo sem ser executadas;
- Ociosidade de equipes média;
- Quantidade de ordens executadas entre os grupos definidos na Tabela 8;
- Tempo médio de espera de ordens de serviço por grupo.

A partir da definição dos cenários, o processo de simulação foi executado, onde para cada cenário, executou-se 100 processos de repetição e, para cada um deles, se evoluiu por um período de 15 dias o processo de geração e execução em cada um dos postos técnicos. A seguir é explicado a configuração dos cenários e os resultados obtidos na simulação.

3.4.2.1 Cenário A

Esse cenário possui 31 equipes, segregadas em 4 postos técnicos, conforme Tabela 17. Os postos técnicos são formados por 13 áreas de execução, conforme Tabela 16. As prioridades respeitam a evolução dos vencimentos de ordens, porém uma ordem de grupo

inferior nunca atinge o nível de prioridade máxima de uma ordem de prioridade superior, conforme pode ser visualizado na Tabela 18.

Tabela 16 - Relação área de execução e posto técnico

Posto técnico	Áreas de execução
34	1, 2, 3
38	7, 8, 9
39	4, 5, 6
2042	11, 12, 17, 18

Fonte: Desenvolvido pelo autor.

Tabela 17 - Relação equipes e posto técnico

Posto técnico	Equipes
34	213, 233, 204, 221, 299, 223, 234, 242, 243, 201, 205, 274
38	240, 241, 258, 269, 217, 254, 270, 265, 289
39	200, 203, 263, 271, 272, 268
2042	202, 284, 267, 273

Fonte: Desenvolvido pelo autor.

Tabela 18 - Matriz de prioridades

Prazo	Prioridades		
<1000	2	3	3
<500	2	2	3
<72	2	2	3
<48	1	1	2
<24	0	1	2
<0	0	1	2
Grupo	Reguladas	Cientes	AES

Fonte: Desenvolvido pelo autor.

O processo de simulação consumiu um tempo de 54 horas para efetuar todo o ciclo de repetições. Após o processo de simulação estar concluído os indicadores de desempenho do processo foram calculados e estão apresentados na Tabela 19. As equipes apresentaram uma ociosidade média total de 0,8hs por dia.

Tabela 19 - Resultado Cenário A

Indicador	Reguladas	Cliente	AES
Média do tempo de espera (dias)	2,7	5,8	3,6
Média de ordens executadas	330	222	523
Média de ordens vencidas executadas	2,5	2,85	0
Média de ordens vencidas pendentes	1,6	47	30

Fonte: Desenvolvido pelo autor.

Os resultados da Tabela 19 indicam uma priorização elevada das ordens de serviços reguladas, sendo que do total de ordens executadas e pendentes vencidas, somente 1,2% das ordens acabaram ultrapassando o prazo de vencimento, sendo que o valor máximo da meta permitiria até 5%.

Porém, esse comportamento não se refletiu nas demais ordens solicitadas pelos clientes, quando 18% das ordens foram executadas ou estão pendentes fora do prazo estipulado. Um percentual bem maior que o máximo definido para a meta desse tipo de ordem que seria 10%. Esse desempenho fica aquém inclusive das ordens geradas pela AES, que possuiu um desempenho de 5,4%.

Ao verificar o motivo dessa diferença entre o desempenho das ordens do tipo Cliente em relação às ordens do tipo AES, identifica-se que, dos 10 dias de prazo das ordens do tipo cliente, 8 deles as ordens estão com prioridade 2, as mesmas que as ordens do tipo AES nos seus dois últimos dias de prazo. Como as ordens do tipo AES são em maior número e consequentemente uma parcela delas está localizada em regiões centrais, elas se tornam atrativas para o processo de roteirização que tende a minimizar o tempo de espera das ordens. Sendo que o mesmo indicador de roteirização, também tende a dar preferência para ordens mais rápidas (com menor tempo de serviço), pois quanto menor o tempo de serviço da ordem predecessora, menor será o tempo de espera da ordem sucessora. Assim o algoritmo de roteirização termina por priorizar ordens localizadas em regiões com maior densidade de serviços e que possuam um menor tempo de serviço, exatamente as características das ordens do tipo AES. De modo que, como por um período as ordens ficam concorrendo dentro de uma mesma prioridade, as ordens do tipo AES levaram vantagem na roteirização.

3.4.2.2 Cenário B

A configuração do cenário de despacho é um reflexo do cenário A, porém com 4 equipes a menos, sendo que foi reduzida exatamente uma equipe de cada posto técnico, conforme Tabela 20:

Tabela 20 - Relação equipes e posto técnico

Posto técnico	Equipes
34	233, 204, 221, 299, 223, 234, 242, 243, 201, 205, 274
38	240, 241, 258, 269, 254, 270, 265, 289
39	200, 203, 263, 272, 268
2042	202, 284, 267

Fonte: Desenvolvido pelo autor.

Para esse cenário o processo de simulação durou 52 horas para percorrer todo o ciclo de repetições. Os dados resultantes da simulação estão apresentados na Tabela 21. A ociosidade média das equipes foi menor que o cenário A, com 0,02 horas. Por sua vez, o cenário A possuiu indicadores melhores em todos os indicadores de ordens reguladas (mais relevantes), porém mesmo com menos equipes o cenário B executou mais ordens de serviço totais.

Tabela 21 - Resultado Cenário B

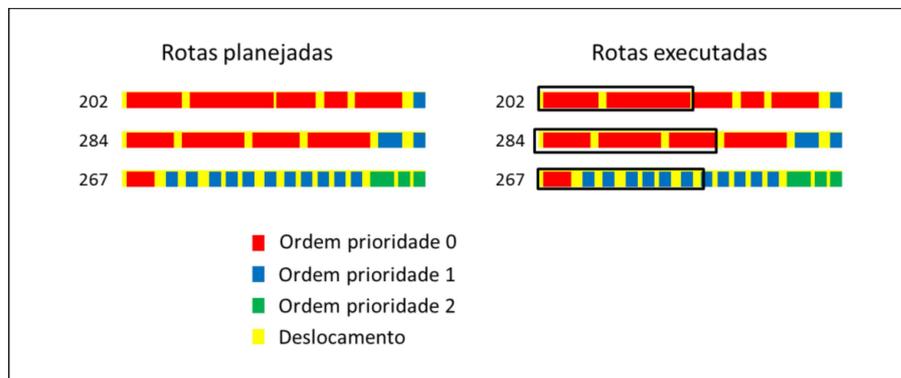
Indicador	Reguladas	Cliente	AES
Média do tempo de espera	3,1	6,2	4,3
Média de ordens executadas	294	210	618
Média de ordens vencidas executadas	5,3	1,68	0
Média de ordens vencidas pendentes	23,3	52,5	37

Fonte: Desenvolvido pelo autor.

Inicialmente, ao analisar a comparação do cenário A com o cenário B, onde se reduz a quantidade de equipes e como consequência aumenta-se a quantidade de ordens executadas, aparenta existir uma incostência no processo de simulação. Porém, ao analisarmos o que ocorreu realmente, identifica-se que, com a diminuição na quantidade de equipes a relação entre demanda e capacidade aumentou, passando do limite especificado no algoritmo de roteirização (55%), conforme explicado no item 3.2.5, isso fez com que ocorresse um

desequilíbrio de prioridades entre as rotas. Como pode ser visualizado na Figura 37, onde no posto 2042 o algoritmo construiu rotas desequilibradas entre si em relação a prioridade, possibilitando que mais ordens de prioridade 2 fossem alocadas em posições iniciais de uma das rotas. Coincidentemente, nessa prioridade as ordens de serviço do tipo AES são em maior número e possuem um tempo de execução previsto menor que as demais ordens de serviço. Essas situações, somadas a função objetivo do algoritmo ser baseada na minimização do tempo de espera, fez com que fossem executadas mais ordens do tipo AES.

Figura 37 - Rotas desequilibradas em função da prioridade



Fonte: Desenvolvido pelo autor.

A Figura 37 é a reprodução de uma amostra do processo de simulação de despacho e execução. Onde, ao lado esquerdo é apresentado uma rota planejada e, à direita, a rota executada (evidenciada pela borda preta).

Desse modo, o controle sobre a relação de demanda e capacidade é vital para o bom funcionamento do algoritmo atual.

3.4.2.3 Cenário C

A configuração do cenário de despacho é um reflexo do cenário B, porém alterando a prioridade das ordens de serviço, conforme pode ser visualizado na Tabela 22. Nesse caso, cada grupo de ordem de serviço (Reguladas, Clientes e AES) possui um valor de prioridade distinta, de modo que uma ordem do tipo Cliente recebe o valor 1 no momento em que é gerada e permanece com essa prioridade até o momento em que é executada.

A estratégia da alteração da matriz, visa obrigar que o processo de roteirização sequeencie respeitando somente os tipos de ordens sem levar em razão o prazo delas.

Tabela 22 - Matriz de prioridade

Prazo	Prioridades		
<1000	0	1	2
<500	0	1	2
<72	0	1	2
<48	0	1	2
<24	0	1	2
<0	0	1	2
Grupo	Reguladas	Clientes	AES

Fonte: Desenvolvido pelo autor.

Assim, no cenário C era esperado uma piora nos resultados relacionados ao atendimento do prazo, em relação ao cenário B. Pois, como todas as ordens de um mesmo grupo são definidas com a mesma prioridade que ordens mais antigas, não evoluindo em prioridade com o passar do tempo, o processo de roteamento entende que todas as ordens novas são iguais às antigas e assim ordens que estão em regiões marginais podem acabar sendo atrasadas frequentemente para não onerar o deslocamento.

Tabela 23 - Resultado Cenário C

Indicador	Reguladas	Cliente	AES
Média do tempo de espera	3,2	6,4	4,3
Média de ordens executadas	297	206	621
Média de ordens vencidas executadas	5,13	2,4	0
Média de ordens vencidas pendentes	21,3	56,5	38,1

Fonte: Desenvolvido pelo autor.

Porém, o resultado não foi muito diferente do cenário B, pois acabaram sendo praticamente iguais. Esse cenário consumiu 53 horas de tempo de execução e as equipes possuíram um tempo médio de ociosidade em torno de 0,02 horas.

Existe a possibilidade da hipótese não ter se concretizado, sobre o cenário C ser pior que o cenário B em razão da dificuldade de atender ordens periféricas, devido ao estudo de caso ter sido efetuado em uma região urbana (Canoas) com uma elevada densidade geográfica de ordens pendentes, não existindo ordens periféricas muito custosas (distantes). Como proposta para estudo posterior, pode ser analisado o comportamento desses dois cenários em uma região que possua uma menor densidade de clientes, com áreas rurais maiores.

3.4.2.4 Cenário D

A configuração do cenário de despacho é um reflexo do cenário A, porém alterando a configuração dos postos técnicos. De modo que, foi criado um para cada área de execução e distribuído as equipes entre os postos técnicos na mesma proporção que a quantidade de ordens pendentes se distribui entre os postos técnicos.

Essa estratégia visa restringir a área de atuação das equipes, restringido a autonomia do algoritmo de roteirização. Nesse cenário, as equipes poderão estar executando ordens de menor prioridade em alguns postos técnicos devido à segregação do planejamento, mas por outro lado, minimiza-se o deslocamento.

A Tabela 24 apresenta as alterações na conformação dos posto técnicos e na Tabela 25 a vinculação das equipes aos postos técnicos.

Tabela 24 - Relação áreas de execução e posto técnico

Posto técnico	Área de execução
34	1
35	2
36	3
37	4
38	5
39	6
40	7
41	8
42	9
44	11
45	12
48	17
49	18

Fonte: Desenvolvido pelo autor.

Tabela 25 - Relação equipes e posto técnico

Posto técnico	Equipes
34	274, 284, 289, 299
35	217, 221
36	203, 204
37	263, 265, 267
38	202
39	223, 233, 234
40	205, 213
41	240, 241, 242
42	268, 269, 270
44	271, 272, 273
45	243, 254, 258
48	201
49	49

Fonte: Desenvolvido pelo autor.

O cenário D teve um desempenho melhor em função do tempo de execução, consumindo não mais de 35 horas de processamento. Essa redução é devida ao aumento na quantidade de instâncias, aumentando a segregação das ordens de serviço e equipes, portanto diminuindo a quantidade de ordens em cada uma das instâncias, e conseqüentemente reduzindo as possibilidades combinatórias no processo iterativo de criação e melhoramento de rota pelo algoritmo.

Porém, os resultados foram os piores entre os quatro cenários simulados. Seja pelas equipes terem possuído um tempo médio de ociosidade em torno de 0,3 horas, ou pela execução de mais ordens reguladas fora de prazo, mas principalmente por postergar outras 73 ordens reguladas para além do prazo de vencimento, deixando-as pendentes de execução. Assim, totalizando 26% das ordens reguladas fora do prazo. Outro ponto, que ficou muito aquém dos demais, foi a quantidade de serviços executados, 122 ordens a menos que o cenário A, como pode ser visualizado na Tabela 26.

Tabela 26 - Resultado Cenário D

Indicador	Reguladas	Cliente	AES
Média do tempo de espera	3,3	6,4	4,2
Média de ordens executadas	226	233	494
Média de ordens vencidas executadas	6,2	1,61	0
Média de ordens vencidas pendentes	73	24,5	27,8

Fonte: Desenvolvido pelo autor.

O desempenho inadequado foi ocasionado pela demasiada segregação de áreas e equipes com a conseqüente impossibilidade de remanejamento dessas ao longo do processo de simulação. Pois, conforme o processo de geração e execução ocorre, o ambiente do posto técnico se altera, como exemplo: se um pico de demanda é seguido por uma redução de disponibilidade de equipe, a demanda pendente irá aumentar e se nesse posto técnico contiver uma quantidade de equipes muito próxima ao limiar da capacidade de atender a demanda média, o tempo para a fila de demandas pendentes voltar a um tamanho normal pode demorar mais do que o tempo até o vencimento de algumas ordens pendentes, aumentando a quantidade de ordens vencidas e assim caindo no ciclo reforçador apresentado na Figura 20.

O mesmo ocorre em relação à ociosidade das equipes, pois se ocorrer uma diminuição no ingresso de serviços e/ou uma elevação no desempenho das equipes, a quantidade de serviços pode reduzir a um nível que não existam serviços pendentes para execução.

Porém, apesar de tudo, esse cenário permite um controle maior sobre a localização das equipes, que ficam com as rotas em locais delimitados menores, garantindo a existência de equipes em locais estratégicos durante todo o turno de trabalho. Contudo, o monitoramento dessa alocação de equipes deve ser constante, pois como no estudo em questão, as equipes ficaram alocadas a uma mesma área durante 15 dias sem nenhuma ação de gestão, maximizando os resultados negativos relatados acima.

4 CONCLUSÃO

O processo de gerenciamento de serviços em uma empresa de distribuição de energia elétrica foi mapeado de forma minuciosa, buscando identificar todas as variáveis relevantes do problema, resultando em um modelo robusto de simulação e assim criando uma plataforma de estudos do comportamento desse sistema.

A modelagem apresentada, além de atender a necessidade da aplicação de simulação representa um documento formal para a empresa, auxiliando também no treinamento de funcionários e no entendimento geral sobre o problema.

Os cenários que foram testados no estudo de caso apresentaram um comportamento muito próximo ao esperado, quando das suas escolhas. Porém, é uma pequena parcela das possibilidades de conformação do ambiente, de modo que essa pesquisa somente iniciou um processo de análise e melhoria contínua do problema de gerenciamento de serviços. Assim, os resultados, mesmo que satisfatórios, não são suficientes para melhorar o nível de serviço de forma consistente, sendo que, se encontrou um melhor resultado dentre quatro. Muito distante de ter encontrado o melhor resultado para o sistema.

Em relação ao desempenho do processo de simulação, os tempos de execução são muito elevados, chegaram a 54 horas mesmo testando somente 100 repetições, o que é bem abaixo dos valores recomendados em um processo de simulação. Esse tempo elevado se dá devido à necessidade de execução do algoritmo de roteamento para cada iteração do método, sendo que, o tempo de execução desse pode variar dependendo da quantidade de equipes e demandas pendentes. Porém, em relação à qualidade da solução, o simulador se comportou de forma adequada mostrando aderência ao problema mapeado.

4.1 SUGESTÕES DE TRABALHOS FUTUROS

Como sugestão de trabalhos futuros, uma possibilidade seria desenvolver uma metodologia para equilibrar de forma automática a demanda e capacidade, seja redistribuindo a alocação das equipes nos postos técnicos, alterando prioridades das ordens ou a conformação de postos técnicos.

REFERÊNCIAS

- ANDRADE, E. L. **Introdução à Pesquisa Operacional** - Métodos e Modelos para Análise de Decisões. 4ª edição. Editora LTC, 2009. Vital Book file, disponível em: <<http://integrada.minhabiblioteca.com.br/books/978-85-216-1919-2>>. Acesso em: 04 fev. 2016.
- CHUN-FENG, F.; YUN, M.; SHAN, J. **Design and Implement of the Rapid Dispatching System of Power Failure Repair Vehicles**. Queue, pp1388-1391. 2011.
- CURCIC, S.; OZVEREN, C. S.; CROWE, L.; LO, P. K. L. **Electric power distribution network restoration: a survey of papers and a review of the restoration problem**. Electric Power Systems Research, 35, pp73-86. 1996.
- FITZSIMMONS, J. A. e FITZSIMMONS, M. J. **Administração de Serviços**, 6ª Edição. Editora Bookman, 2010.
- HILLIER, F. S. e LIEBERMAN, G. J. **Introdução à Pesquisa Operacional**. 9ª Edição. AMGH, 2013. Vital Book file, disponível em: <<http://integrada.minhabiblioteca.com.br/books/9788580551198>>. Acesso em: 04 fev. 2016.
- HUANG, S.; AXSÄTER, S.; DOU, Y.; CHEN, J. **A real-time decision rule for an inventory system with committed service time and emergency orders**. European Journal of Operational Research 215, pp. 70 – 79. 2011.
- PERRIER, N., AGARD, B., BAPTISTE, P., FRAYRET, J., PELLERIN, R., RIOPEL, D., TRE, M. **Computers & Operations Research A survey of models and algorithms for emergency response logistics in electric distribution systems**. Part I: Reliability planning with fault considerations, 40, pp1895-1906. 2013a
- _____. _____. Part II: Contingency planning level, 40, pp1907-1922. 2013b
- TAYLOR, RW. **A Linear Programming Model to Manage the Maintenance Backlog**. Omega, Int. J. Mgmt Sci. Vol. 24, N° 2, pp. 217-227. 1996
- WANG, D.; TANG, O.; HUO, J. **A heuristic for rationing inventory in two demand classes with backlog costs and a service constraint**. Computers & Operations Research 40, pp. 2826-2835. 2013
- YAN, T., RABINOVICH, E., DOOLEY, K., EVERS, P. T. **Managing backlog variation in order fulfillment: The case of Internet retailers**. Int. J. Production Economics 128, pp. 261-268. 2010
- ZHU, J. **Optimization of Power System Operation**. Wiley – IEEE Press. 2009.