

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Explorando Efeitos da Adoção do MPS-SW no Desempenho das Organizações

Matheus Rabello de Olivera

JUIZ DE FORA
JULHO, 2014

Explorando Efeitos da Adoção do MPS-SW no Desempenho das Organizações

MATHEUS RABELLO DE OLIVERA

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Departamento Ciência da Computação
Bacharelado em Ciência da Computação

Orientador: Victor Stroele de Andrade Menezes

JUIZ DE FORA

JULHO, 2014

EXPLORANDO EFEITOS DA ADOÇÃO DO MPS-SW NO DESEMPENHO DAS ORGANIZAÇÕES

Matheus Rabello de Olivera

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Victor Stroele de Andrade Menezes
D. Sc.

Marcos Kalinowski
D. Sc.

Alessandreia Marta de Oliveira
M. Sc.

JUIZ DE FORA
23 DE JULHO, 2014

Resumo

O objetivo do presente trabalho é abordar efeitos da adoção do MPS-SW no desempenho de organizações. Este tema foi escolhido por ser muito importante para a análise do impacto do modelo de maturidade MPS-SW, buscando, assim sua evolução. Para realizar o trabalho, foram utilizados os dados quantitativos da pesquisa iMPS, projeto criado para acompanhar e evidenciar a variação de desempenho nas empresas que adotaram o modelo MPS-SW. O estudo analisa a pesquisa iMPS, identificando fatores já abordados na literatura e outros ainda não discutidos. Redes Bayesianas foram utilizadas nos estudos do impacto da adoção do modelo buscando listar proposições que ajudem a compreender o uso desse modelo.

Palavras-chave: mps-sw, modelo de maturidade, redes bayesianas.

Abstract

The objective of this study is to approach effects of adopting MPS-SW in the performance of companies. This theme had been chosen because it is very important to analyze the impact of the maturity model MPS-SW, with the objective of the model be able to evolve. The kind of research that was developed is quantitative. To perform the work were used data from the search iMPS, a project that was created to monitor and demonstrate the performance variation in firms that adopted the MPS-SW. The work consisted in analyzing the research iMPS, identifying factors already discussed in the literature and others not yet addressed. To study the impact of adopting the model, Bayesian Networks was used in order to understand some relationships that have not been performed in previous studies of the iMPS project. Thus, list propositions that help to understand the impact of the adoption of the model.

Keywords: mps-sw, maturity model, bayesian networks.

Agradecimentos

À todos os meus familiares que torceram por mim durante a minha graduação, em especial aos meus pais, Nelson e Silvana, pelo grande apoio. Sem eles esse momento não seria possível.

Aos meus amigos de São Geraldo, pelo apoio de sempre. Aos amigos e parceiros da república em especial aos amigos da UFJF, pelo incentivo, força. E pelos infinitos momentos de incerteza e insegurança que passamos juntos. Vocês são demais.

Ao professor Victor pela orientação, amizade e, principalmente, pela paciência, sem a qual este trabalho não seria realizado.

Aos professores do Departamento de Ciência da Computação pelos ensinamentos. À todos, muito obrigado!

“Ninguém pode voltar atrás e fazer um novo começo. Mas qualquer um pode recomeçar e fazer um novo fim”

Chico Xavier

Sumário

Lista de Figuras	7
Lista de Tabelas	8
1 Introdução	9
1.1 Contextualização	9
1.2 Justificativa	10
1.3 Objetivos Gerais e Específicos	11
1.4 Metodologia	12
1.5 Organização do trabalho	12
2 Fundamentação Teórica	13
2.1 Modelo MPS-SW e Maturidade no Desenvolvimento de Software	13
2.2 Redes Bayesianas	16
2.3 Pesquisa iMPS	17
3 Elaboração da Rede Bayesiana	20
3.1 Preparação dos dados, Software utilizado e Algoritmo de Aprendizagem	20
3.2 Nós da Rede Bayesiana	21
3.3 Montagem da Rede Bayesiana	26
3.4 Análise Crítica	27
4 Explorando Efeitos da Adoção do MPS-SW no desempenho das organizações utilizando Redes Bayesianas	29
4.1 Identificação das proposições	29
4.2 Utilização da rede: explorando os efeitos	30
4.2.1 Organizações de nível A e C gerenciam estimativas melhor que organizações de nível G	30
4.2.2 A produtividade de organizações do Nível F é maior em relação ao nível A. Porém, a qualidade dos projetos (Def/PF) desenvolvidos por organizações nível A é melhor em relação ao nível F	31
4.2.3 Organizações sem nenhum nível MPS-SW gastam menos tempo para desenvolver projetos muito pequenos do que organizações nível G	32
4.2.4 Organizações nível F produzem mais e gastam menos do que organizações nível C	32
4.2.5 Em organizações nível C os projetos gastam mais HH/PF do que no nível F. Porém a qualidade (Def/PF) dos projetos nível C são melhores	33
4.2.6 Os projetos desenvolvidos no nível F, são maiores que os projetos desenvolvidos no nível G	34
4.2.7 Organizações nível A possuem mais funcionários que organizações nível G	35
4.2.8 Organizações nível A e C tendem a ter mais clientes no exterior que organizações nível G	35

4.3	Análise das proposições	36
5	Considerações Finais	37
	Referências Bibliográficas	39

Lista de Figuras

3.1	Tamanho dos Projetos por Pontos de Função	22
3.2	Precisão Estimativa	22
3.3	Cientes no país	23
3.4	Cientes no exterior	23
3.5	Número de funcionários	23
3.6	Número de Projetos no país	24
3.7	Número de Projetos no Exterior	24
3.8	Nível MPS	24
3.9	Momento da Pesquisa	25
3.10	Produtividade	25
3.11	Defeitos/PF	25
3.12	Tempo médio gasto nos projetos	26
3.13	Esforço médio gasto nos projetos	26
3.14	Primeira Rede Bayesiana	27
3.15	Segunda Rede Bayesiana	28
4.1	Primeira proposição	30
4.2	Segunda proposição	31
4.3	Terceira proposição	32
4.4	Quarta proposição	33
4.5	Quinta proposição	34
4.6	Sexta proposição	34
4.7	Sétima proposição	35
4.8	Oitava proposição	35

Lista de Tabelas

2.1	Níveis de Maturidade do MPS-SW, processos e capacidade(Softex, 2012)	. 14
-----	--	------

1 Introdução

Nesse capítulo será apresentado a contextualização e a justificativa do trabalho desenvolvido. Também será apresentado os objetivos gerais e específicos e a metodologia utilizada nesse trabalho.

1.1 Contextualização

A melhoria contínua da capacidade de desenvolvimento de software é fundamental para que organizações prosperem em mercados competitivos. Com o passar dos anos, modelos de referência têm surgido para ajudar na incrementação da capacidade de processos de engenharia de software. Entretanto, a melhoria baseada nesse tipo de modelo costuma ser de longo prazo e requerer grandes investimentos. Esses obstáculos podem se tornar impeditivos para que organizações melhorem seus processos, especialmente para pequenas e médias empresas (PMEs) que operam sob rígidas restrições financeiras. O programa MPS.BR representa uma iniciativa para aumentar a capacidade das empresas Brasileiras em desenvolver software.

O principal objetivo do programa MPS-BR é desenvolver e disseminar modelos de referência que atendam a indústria Brasileira de Software e Serviços de TI, visando estabelecer um caminho economicamente viável para que as organizações incluindo as pequenas e médias empresas, alcancem os benefícios da melhoria de processos e da utilização de boas práticas da engenharia de software e da prestação de serviços de TI em um intervalo de tempo razoável. Atualmente, a família de modelos é composta pelo MPS-SW para software e MPS-SV para Serviços de TI. Nesse trabalho o foco será o modelo MPS-SW(Softex, 2012).

1.2 Justificativa

Dentre os principais problemas que inibem as organizações especialmente PMEs, de adotar modelos de processos de software; tais como CMMI, ISO/IEC 12207 e ISO/IEC 15504, estão o alto investimento de implementação, manutenção e avaliação e a dificuldade de convence-las dos potenciais benefícios dos investimentos em melhoria de processos. Para superar esses problemas são necessárias diferentes abordagens que busquem a mudança do cenário de processos em organizações de software. Neste contexto, a SOFTEX decidiu realizar uma iniciativa nacional, denominada Programa MPS.BR, com o objetivo de disseminar as boas práticas da engenharia de software e tornar as organizações mais competitivas não apenas no mercado local, mas, também, no mercado global (Kalinowski et. al., 2010) (Softex, 2012).

O modelo MPS.BR obteve, com o passar dos anos uma ampla aceitação pelas empresas de software Brasileiras e já conta com mais de 500 avaliações. É imprescindível entender as variáveis de atuação no desenvolvimento técnico e econômico dessas empresas, como prazo, produtividade, custo e qualidade o que promove a compreensão do impacto causado pelo projeto. Nesse contexto foi criado o iMPS, que tem como objetivo acompanhar e evidenciar a evolução das empresas de software que adotaram o modelo MPS (Softex, 2012).

A base de dados da pesquisa iMPS tem disponibilidade para análises adicionais, permitindo uma maior compreensão dos impactos da adoção do modelo MPS-SW. Esses dados estão disponíveis em planilhas eletrônicas, das quais muitas informações podem ser extraídas, como: se a adoção do modelo traz mais benefícios para determinados tipos de serviços ou categorias de produtos; se os impactos da adoção nas diferentes regiões do país são iguais. Trabalhar com esses dados nas planilhas será necessário para a realização deste trabalho (Simões et. al., 2010).

Para a realização dessas análises serão utilizadas Redes Bayesianas, também conhecidas como redes causais modelos gráficos para raciocínio baseado na incerteza, nos quais os nós representam as variáveis (discreta ou contínua) e os arcos a conexão direta entre eles. Redes Bayesianas vem se tornando uma metodologia padrão para a construção do conhecimento probabilístico e tem sido aplicada em diferentes áreas em todo o mundo.

A base das Redes Bayesianas é a obra do matemático Thomas Bayes, chamada Teoria da Probabilidade Bayesiana, publicada em 1763. A representação dessa rede é feita por meio de um grafo direcionado acíclico no qual os nós representam variáveis de um domínio e os arcos a dependência condicional ou informativa entre as variáveis. Para representar a força da dependência são utilizadas probabilidades associadas a cada grupo de nós pais-filhos na rede (Pearl, 2000).

1.3 Objetivos Gerais e Específicos

Apresentar a caracterização e desempenho das organizações que adotaram o MPS-SW, a partir dos dados fornecidos entre os anos 2008 e 2014 pela pesquisa iMPS. Descrever os resultados da análise global envolvendo a agregação dos dados das organizações que participaram das rodadas da pesquisa iMPS através do levantamento de proposições e da exposição dos resultados observados nas Redes Bayesianas montadas.

Analisar a importância do modelo MPS-SW através de diferentes informações ainda não exploradas na base de dados iMPS. Essa pesquisa, dentre outras informações, separa as organizações em estágios, ou seja se elas estão em processo de implantação ou se já possuem algum nível MPS-SW certificado. Assim, uma avaliação das empresas certificadas e não certificadas será útil para reafirmar a importância do modelo MPS-SW.

Analisar o impacto da implantação dos diferentes níveis do MPS-SW nas organizações. A base de dados utilizada nesse trabalho permite verificar esse impacto nas organizações. Por exemplo se o investimento para conseguir uma certificação em um nível mais alto é proporcional ao ganho em qualidade.

Finalmente, investigar o atual nível de influência do modelo de referência MPS-SW nas organizações, através da análise das diferentes métricas disponíveis na pesquisa iMPS. Para isso serão identificadas algumas proposições, que são um termo usado em lógica para descrever o conteúdo de asserções que podem ser tomadas como verdadeiras ou falsas.

1.4 Metodologia

A pesquisa a ser realizada nesse trabalho será quantitativa. Pois fará um estudo estatístico, medindo numericamente as diferentes proposições levantadas a respeito dos diversos pontos abordados na projeto iMPS.

O tipo de trabalho será documental. Para permitir a análise profunda das pesquisas envolvidas no trabalho.

No tocante ao procedimento, esse trabalho será realizado por meio de observação indireta, uma vez que os resultados e conclusões serão obtidos através da análises de diferentes fatores, que de alguma forma se relacionam. Assim é possível chegar em resultados coerentes com a pesquisa realizada.

1.5 Organização do trabalho

O restante desse trabalho se encontra dividido da seguinte forma. O capítulo 2 apresenta os aspectos teóricos referentes ao MPS-SW. O capítulo 3 apresenta a forma como as Redes Bayesianas foram desenvolvidas. O capítulo 4 apresenta como a rede foi explorada e as proposições elicitadas. Por fim, o capítulo 5 apresenta as conclusões, as limitações e os trabalhos futuros oriundos dessa pesquisa.

2 Fundamentação Teórica

Este capítulo apresenta uma fundamentos teóricos, delineando inicialmente conceitos do Modelo MPS-SW e de maturidade no desenvolvimento de software. O segundo tópico tratará de Redes Bayesianas e, o terceiro da pesquisa iMPS.

2.1 Modelo MPS-SW e Maturidade no Desenvolvimento de Software

Segundo Kalinowski (2010), um dos principais problemas que dificultam a implantação de modelos de referência em organizações especialmente em PMEs, como o CMMI, ISO/IEC 12207 e ISO/IEC 15504 são os altos custos de implementação, manutenção e avaliação e também a difícil missão de convencer as organizações da importância de investir em melhoria de processos. Para superar essas barreiras são necessárias diferentes abordagens dos modelos clássicos. Para tanto a SOFTEX criou uma iniciativa nacional para melhorar os processos de software de organizações brasileiras e disseminar boas práticas de Engenharia de Software, denominada de Programa MPS.BR. A obtenção do consenso dos interesses da indústria nacional e sua documentação em forma de modelo de processos de software foi um primeiro e importante passo o obter sucesso do programa(Viana et. al., 2010).

Um dos requisitos que o modelo MPS.BR procurou atender é foi a incorporação de práticas internacionalmente reconhecidas para a implementação e avaliação de processos de Engenharia de Software(Kalinowski et. al., 2010). Observou-se, por exemplo, que o modelo CMMI foi considerado um complemento técnico para a definição do modelo MPS-SW, e é importante para as organizações brasileiras que atuam fora do país(Kival, 2008).

Foram criados sete níveis de maturidade de processos (G, F, E, D, C, B, A) permitindo acesso gradual a melhoria de processos(Softex, 2012) cada um dos quais (do

nível G - primeiro estágio de maturidade ao nível A - mais maduro) apresenta um conjunto de processos e atributos de processos (AP) que indicam onde a unidade organizacional tem que investir esforço para melhoria.

Tabela 2.1: Níveis de Maturidade do MPS-SW, processos e capacidade(Softex, 2012)

Nível	Processos	Atributos de processo
A	Em otimização	AP1.1, AP2.1, AP2.2, AP3.1, AP3.2, AP4.1, AP4.2, AP5.1 e AP5.2
B	Gerência de Projetos-GPR(evolução)	AP1.1, AP2.1, AP2.2, AP3.1, AP3.2, AP4.1 e AP4.2
C	Gerência de Riscos - GRI Desenvolvimento para Reutilização - DRU Gerência de decisões	AP1.1, AP2.1, AP2.2 AP3.1 e AP3.2
D	Verificação - VER Validação - VAL Projeto e Construção do Projeto - PCP Integração do Produto - ITP Desenvolvimento de Requisitos - DRE	AP1.1, AP2.1, AP2.2 AP3.1 e AP3.2
E	Gerência de Projetos-GPR (evolução) Gerência de Reutilização Gerência de Recursos Humanos - GRH Definição do Processo Organizacional - DFP Avaliação e Melhoria do Proc. Organizacional - AMP	AP1.1, AP2.1, AP2.2 AP3.1 e AP3.2
F	Medição - MED Garantia da Qualidade Gerência de Portfólio de Projetos - GPP Gerência de Configuração - GCO Aquisição - AQU	AP1.1, AP2.1 e AP2.2
G	Gerência de Requisitos - GPE Gerência de Projetos - GPR	AP 1.1 e AP 2.1

Para cada nível de maturidade do modelo MPS-SW, existem diferentes Atributos de Processos(capacidades) que devem ser devidamente utilizados para a obtenção do nível, assim, como diferentes processos devem ser realizados para cada nível de maturidade. Os processos e os atributos que cada nível deve satisfazer estão listados na Tabela 2.1. Todos os níveis de maturidade herdam os processos e atributos dos níveis inferiores. Dessa forma o nível E, engloba todos os processos e atributos de processo dos níveis de maturidade G e F.

O nível G é chamado de Parcialmente gerenciado. Nele os atributos de processos

que devem ser implementados são os 1.1 e o 2.1. O primeiro 1.1 refere-se ao conhecimento da execução de um processo. Este atributo é uma medida do quanto do processo é executado. Já o 2.1 é uma medida do quanto a execução do processo é gerenciado.

O nível F é denominado de Gerenciado e é composto pelos processos do nível G mais os processos listados para o nível F. Os atributos são os mesmos apresentados no nível G acrescidos do atributo 2.2, que relaciona-se com os produtos do processo. Essa é uma medida do quanto os produtos de trabalho produzidos pelo processo são devidamente gerenciados.

O nível E chamado de Parcialmente Definido, é composto pelos processos dos níveis G e F juntamente com os processos listados para o nível E. Os atributos de processo são os mesmos do nível F acrescido dos atributos 3.1 e 3.2. O AP 3.1 refere-se a saber se o processo é definido é uma medida do quanto um processo padrão é mantido para apoiar a implementação do processo definido. O AP 3.2, por sua vez, é uma medida do quanto o processo padrão é efetivamente implementado como processo definido para a obtenção de seus resultados.

Os APs dos níveis D e C são os mesmos (3.1 e 3.2), sendo que em cada um desses níveis são acrescidos novos processos, como descrito na Tabela 2.1. O nível D é denominado Largamente Definido, e o nível C Definido.

O nível B é chamado de Gerenciado Quantitativamente. Nesse nível nenhum processo é acrescido. O processo de Gerência de Projetos sofre uma segunda evolução, com novos resultados para atender aos objetivos de gerenciamento quantitativo. E os atributos 4.1 e 4.2 são adicionados ao processo. O AP 4.2 é uma medida do quanto os resultados de medição são usados para assegurar que o desempenho do processo apoie o alcance dos objetivos relevantes para os objetivos definidos. Já o atributo 4.2 é uma medida do quanto o processo é controlado estatisticamente para produzir um processo estável, capaz e previsível dentro de limites estabelecidos.

E por fim o nível A é chamado de Em Otimização. Nesse nível não há acréscimo de processos. O AP 5.1 é uma medida do quanto as mudanças no processo são identificadas a partir da análise de causas comuns de variação do desempenho e da investigação de enfoques inovadores para a definição e implementação do processo. O AP 5.2 é uma

medida do quanto as mudanças na definição, gerência e desempenho do processo têm impacto no alcance dos objetivos relevantes melhoria do processo (Souza et. al., 2010).

2.2 Redes Bayesianas

Redes bayesianas são modelos gráficos de raciocínio baseado na incerteza, nas quais os nós representam as variáveis (discreta ou contínua) e os arcos a conexão direta entre eles. Rede Bayesiana vem se tornando uma metodologia padrão para a construção do conhecimento probabilístico e tem sido aplicada em diferentes áreas de conhecimento. A base teórica das Redes Bayesianas é a obra do matemático Thomas Bayes, chamada Teoria da Probabilidade Bayesiana, publicada em 1763. A representação da rede bayesiana é feita por meio de um grafo direcionado acíclico no qual os nós representam variáveis de um domínio e os arcos a dependência condicional ou informativa entre as variáveis. Para representar a força da dependência são utilizadas probabilidades associadas a cada grupo de nós pais-filhos na rede (Pearl, 2000).

Segundo Pearl (2000) o uso do modelo de grafos em probabilidade e estatística é dado por 3 motivos:

- Proporcionar meios convenientes de expressar suposições substantivas;
- Facilitar inferências eficientes para as observações ;
- Facilitar a economia de espaço nas representações.

Utilizar Redes Bayesianas será importante para compreender alguns relacionamentos ainda não realizados em pesquisas anteriores do projeto MPS.BR. As Redes Bayesianas possuem aplicações em diferentes áreas. Alguns exemplos são:

- Sistema para diagnóstico de problemas nas glândulas linfáticas (Hecherman, 1990).
- Map Learning: esse projeto combina problemas de diagnóstico com problemas de decisão: um robô deve percorrer um “labirinto”, procurando aprender os caminhos percorridos e, ao mesmo tempo, explorar caminhos desconhecidos (Basye, 1990).
- AutoClass: sistema de exploração e aquisição de conhecimento espacial (Nasa, 1998).

Muitos trabalhos acadêmicos são desenvolvidos utilizando Redes Bayesianas como, por exemplo, a dissertação de mestrado apresentada por Ricardo Pedreti Chagas para a Fundação Getúlio Vargas, o qual fez uso das Redes Bayesianas para construção de modelos econômicos financeiros. Esses modelos auxiliam nas tomadas de decisão em situações que envolvam grau elevado de incerteza. O trabalho de Pedreti apresenta uma metodologia para levantamento de dados e aplicações de Redes Bayesianas na obtenção de modelos de crescimento de fluxos de caixa de empresas brasileiras. Os resultados são comparados a modelos econométricos de regressão múltipla e finalmente a dados reais observados no período. O trabalho é concluído com avaliação das vantagens e desvantagens da utilização das Redes Bayesianas para as aplicações (Pedreti, 2008).

Outro exemplo é o trabalho “Bayesian Network Models for Web Effort Prediction: A Comparative Study” que teve como objetivo comparar, vários modelos de Rede Bayesiana para a estimativa de esforço Web através de um conjunto de dados empresariais. Nesse trabalho oito redes foram montadas e dessas, quatro foram estabelecidas automaticamente utilizando a ferramenta Hugin e Power Software com dois conjuntos de treinamento cada um com 130 projetos Web. As outras quatro redes foram montadas utilizando um especialista de domínio, ou seja, não foram geradas automaticamente. Dessa forma foi desenvolvido um estudo comparativo entre os diferentes tipos de redes montadas (Mendes e Mosley, 2008).

2.3 Pesquisa iMPS

Com o objetivo de compreender o impacto do modelo MPS-SW, o projeto iMPS (informações para acompanhar e evidenciar variação de desempenho nas empresas que adotaram o modelo MPS-SW) foi iniciado em 2007 junto ao Grupo de Engenharia de Software Experimental da COPPE (Softex, 2012). O objetivo desse projeto foi planejar uma “survey”, seguindo os princípios da engenharia de software experimental. Esse projeto é fortemente apoiado pela engenharia de software experimental, a qual tem se mostrado importante para evidenciar variações decorrentes da adoção de tecnologias de software. Para assegurar a validade interna do estudo e a consolidação adequada dos dados, a estratégia experimental mostrou-se válida.

As pesquisas já realizadas apontam para a necessidade de que as empresas adotem modelos de melhoria de processos (Kalinowski et. al., 2011). O MPS-SW é o principal modelo de referência para orientar iniciativas de melhoria de processos no Brasil. Dessa forma é importante compreender os impactos da adoção desse modelo na indústria.

Para melhor caracterizar as empresas que adotaram o modelo MPS-SW, Travassos e Kalinowski (2012), optaram por dividir o amplo conjunto de dados da pesquisa iMPS em 5 categorias. Quais sejam:

- Empresas Iniciando a Implementação;
- Empresas em Processo de Avaliação;
- Empresas Avaliadas em nível de maturidade G;
- Empresas Avaliadas em nível de maturidade F;
- Empresas Avaliadas em nível de maturidade E-A.

Os autores optaram por essa divisão em função da ocorrência de poucos dados de empresas nos níveis E, D, C, B, A. Assim esses níveis foram analisados em um único grupo. Já nos outros grupos essa dificuldade não ocorre, uma vez que existem muitos dados.

Além disso, foi relevante observar os dados com foco nas diferentes perspectivas tratadas no estudo, quais sejam organização, Projetos e Modelo MPS-SW (Travassos e Kalinowski, 2012).

Diferentes indicadores de desempenho foram avaliados por Travassos e Kalinowski (2012) como:

- Número de clientes no país;
- Número de funcionários;
- Tamanho médio dos projetos;
- Precisão de estimativa e;
- Produtividade.

Para analisar esses diferentes indicadores os autores fizeram uso de diferentes técnicas estatísticas, como Teste de Wilcoxon/Kruskal-Wallis, 1-way e aproximação de ChiSquare (Travassos e Kalinowski, 2012).

3 Elaboração da Rede Bayesiana

Esse capítulo contempla a visão geral da abordagem da pesquisa. Primeiro serão apresentadas questões sobre a preparação dos dados, o software utilizado e o algoritmo de aprendizagem. A seguir são apresentados os nós da rede essa foi montada. Por fim são elencadas algumas análises críticas com relação a montagem da Rede Bayesiana.

3.1 Preparação dos dados, Software utilizado e Algoritmo de Aprendizagem

Os dados utilizados nesse trabalho, conforme já mencionado, foram da pesquisa iMPS, na qual as organizações respondem a um formulário com várias questões relevantes ao modelo MPS-SW. Em pesquisas anteriores não foram levadas em consideração empresas que apresentavam um comportamento muito discrepante do restante dos dados, ou seja, todos os *outlines* foram removidos do conjunto de dados (Travassos e Kalinowski, 2012).

Entretanto, o objetivo aqui é descobrir novas proposições e validar alguns resultados já existentes, e portanto todos os dados são importantes. Podemos definir *outlines*, como "valores fora da curva". Ou seja, por exemplo, valores que estejam fora de um intervalo previamente definido.

Para trabalhar com Redes Bayesianas foi necessário escolher um software de apoio. Nesse cenário foi decidido trabalhar com o NETICA já que esse software possui as ferramentas necessárias para apoiar este trabalho. O software tem uma interface intuitiva e permite que os dados sejam inseridos na forma de arquivos de dados, assim, os arquivos da pesquisa iMPS podem ser incluídos diretamente.

O NETICA possui alguns algoritmos para realizar o aprendizado e encontra os valores desconhecidos para variáveis desconhecidas. Dentre essas opções escolhe trabalhar com o algoritmo Expectation Maximization (EM), que foi a unificação de uma série de trabalhos, proposto em 1977 por Dempster (Dempster et. al., 1977). O algoritmo EM,

pode ser aplicado quando se deseja estipular algum conjunto de parâmetros que descreve uma distribuição de probabilidades. É utilizado para encontrar os valores de máxima verossimilhança de um modelo estatísticos no qual as equações não podem ser diretamente resolvidas. Isto é, valores de dados faltantes são assumidos, simplificando, assim, o modelo. Pode-se também simplesmente escolher valores arbitrários para um dos dois conjuntos de incógnitas, usá-los para estimar o segundo conjunto e em seguida, usar os novos valores para encontrar uma melhor estimativa do primeiro conjunto, e depois disso é possível manter uma alternância entre os dois até que os valores resultantes converjam para pontos fixos.

3.2 Nós da Rede Bayesiana

Como base para esse trabalho foram utilizados dados da pesquisa iMPS, da Softex. O primeiro passo foi identificar quais informações da análise iriam compor a Rede Bayesiana. Além dos campos selecionados por Travassos e Kalinowski (2012) também foram criados novos nós com outras informações. Nesta seção será explicada qual informação cada nó representa e quais os critérios utilizados na construção desses.

Todos os nós da Rede Bayesiana foram divididos em 5 faixas de valores obtendo-se uma classificação de faixa do tipo “muito baixa”, “baixa”, “média”, “alta”, “muito alta”. Essa classificação proporciona uma forma de analisar as informações encontradas e trabalhar mais claramente as proposições ajudando assim a visualizar e a compreender os dados de modo mais claro.

Os nós fornecem também a média e o desvio padrão dos dados analisados em cada nó. Média nada mais é que um valor que “representa” vários outros. Todas as ocorrências de valores nos campos são somadas e divididas pelo número de vezes que foram identificadas. O desvio padrão é uma medida de dispersão usada com a média e mede a variabilidade dos valores em torno da média. O valor mínimo do desvio padrão é 0 indicando que não há variabilidade, ou seja, que todos os valores são iguais a média.

- Tamanho dos projetos/PF;

Este nó fornece o tamanho dos projetos por pontos de função. Aqui as faixas de

valores, como em todos os outros nós, tem cinco valores. A primeira, até 200 pontos de função, representa classificação “muito baixa”. E assim consecutivamente até os valores acima de 5000 pontos de função que representa a classificação “muito alta”. As duas informações na parte inferior do nó são a média (3830) e o desvio padrão (14000). Esses valores altos ocorrem porque os *outlines* não foram removidos e assim valores muito altos fazem a média e o desvio padrão aumentarem.

Tamanho Projeto/PF	
10 to 200	33.6
200 to 500	32.8
500 to 1000	14.9
1000 to 5000	12.5
5000 to 1e5	6.09
3830 ± 14000	

Figura 3.1: Tamanho dos Projetos por Pontos de Função

- Precisão de estimativa;

Este nó fornece a precisão de estimativa nos projetos. A primeira faixa de valores vai até 0.2. Representa classificação “muito baixa” e, consecutivamente até os valores acima de 0,8 até 0.99 que representam a classificação “muito alta”. Nesse campo os valores que foram informados com 1, que sempre acertam as suas estimativas foram desconsiderados, uma vez que representavam uma grande parte do total e poderiam distorcer as informações deste nó.

Precis o de Estimativa	
0.01 to 0.2	1.96
0.2 to 0.4	5.51
0.4 to 0.6	20.1
0.6 to 0.8	43.3
0.8 to 0.99	29.1
0.683 ± 0.19	

Figura 3.2: Precisão Estimativa

- Número de clientes no país;

Essa informação foi utilizada na rede como foi apresentada na pesquisa iMPS. Esse campo informa o número de clientes no país de cada organização que respondeu a pesquisa.

- Número de clientes no exterior;



Figura 3.3: Clientes no país

Essa informação foi obtida da seguinte maneira: a pesquisa relaciona em um campo o “número de clientes na América” e em um outro campo o “número de clientes fora da América latina”. O número de clientes no exterior foi obtido somando os valores dos dois campos.

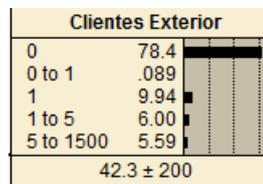


Figura 3.4: Clientes no exterior

- Número de funcionários;

Essa informação foi utilizada na rede como foi apresentada na pesquisa iMPS. O campo informa o número de funcionários de cada organização que respondeu a pesquisa iMPS.

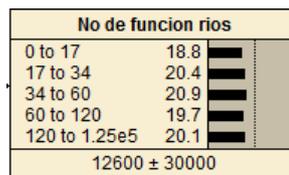


Figura 3.5: Número de funcionários

- Número de projetos no país;

Essa informação foi utilizada na rede como foi apresentada na pesquisa iMPS. O campo informa o número de projetos no país de cada organização que respondeu a pesquisa iMPS.

- Número de projetos no exterior;

Essa informação foi utilizada na rede como foi apresentada na pesquisa iMPS. O campo informa o número de projetos no exterior de cada organização que respondeu

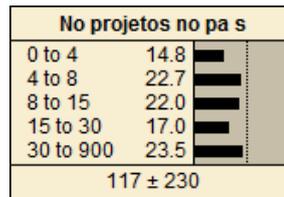


Figura 3.6: Número de Projetos no país

a pesquisa iMPS.

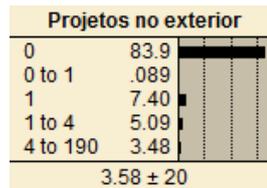


Figura 3.7: Número de Projetos no Exterior

- Nível MPS;

Nesse nó são separados os níveis MPS. Podemos observar que a grande maioria das organizações está nos níveis F e G. Existem algumas nos níveis A, C, E, mas nenhuma no nível B. Este campo é muito importante no contexto do trabalho, visto que a maioria das proposições relacionam-se com o nível MPS.

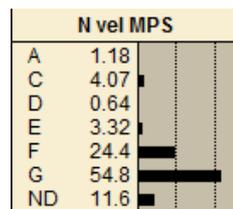


Figura 3.8: Nível MPS

- Momento da pesquisa;

Esse nó separa as organizações em momentos distintos que responderam a pesquisa iMPS. São utilizados 4 momentos. Primeiro, “Em processo de avaliação” significa que a organização está prestes a avaliar algum nível. Segundo “Iniciando a implementação” significa que a organização ainda não possui nenhum nível de maturidade MPS-SW. Terceiro “Periódico” é respondido em um momento que o nível está estabelecido nas organizações. E, por fim, “Retroativo” a organização responde a pesquisa relacionada a um momento passado. Por exemplo a um nível que ela estava

anteriormente.

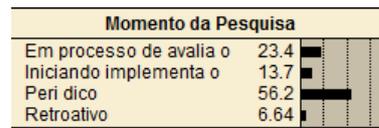


Figura 3.9: Momento da Pesquisa

- Produtividade;

Esse nó fornece uma informação importante. A produtividade é obtida dividindo o tamanho dos projetos (em pontos de função) pelo tempo gasto nos projetos. Este nó não foi fornecido diretamente na pesquisa iMPS. Foi calculado conforme descrito acima.

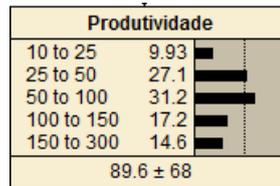


Figura 3.10: Produtividade

- Defeitos/PF;

Esse nó fornece o número de defeitos encontrados por pontos de função. Diferentes métricas foram utilizadas e não existe um padrão a seguir, assim como também não existe uma forma de converter uma métrica em outra. Dessa forma decidiu-se utilizar apenas os valores fornecidos em pontos de função.

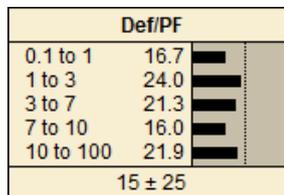


Figura 3.11: Defeitos/PF

- Tempo médio gasto nos projetos;

Esse nó da Rede Bayesiana fornece o tempo gasto nos projetos em meses. A informação foi obtida diretamente da pesquisa iMPS e foi fornecida pelas organizações.

Tempo médio gasto nos projetos	
0 to 3	32.4
3 to 6	40.2
6 to 9	16.0
9 to 12	3.85
12 to 24	7.53
5.26 ± 4.5	

Figura 3.12: Tempo médio gasto nos projetos

- Esforço médio nos projetos;

Esse nó mostra o esforço médio nos projetos em horas. Esse campo foi fornecido diretamente pelas organizações. A informação passou a fazer parte da pesquisa apenas no ano de 2010, assim não está definida nos primeiros campos da tabela da pesquisa iMPS.

Esforço médio dos projetos	
1 to 10	6.87
10 to 100	12.2
100 to 1000	39.3
1000 to 2000	19.8
2000 to 20000	21.8
2920 ± 4900	

Figura 3.13: Esforço médio gasto nos projetos

3.3 Montagem da Rede Bayesiana

Após definir os campos (nós) que fariam parte da Rede Bayesiana, o próximo passo foi a montagem da Rede. Depois de analisar os nós que comporiam a Rede e as definições de objetivos desse trabalho, decidiu-se montar duas redes separadamente. A primeira rede trará os campos que representam a caracterização das organizações; informações que não estão diretamente ligadas ao desempenho dessas. A segunda, por sua vez, é composta por nós que apresentam o desempenho das organizações através da combinação de diferentes métricas coletadas na pesquisa iMPS.

Os campos (nós) da primeira rede apresentam informações que possibilitam observar as características das organizações por nível. São eles: clientes no país, clientes no exterior, número de funcionários, projetos no país e projetos no exterior. Além desses mais 2 campos foram utilizadas nessa primeira rede: o nível MPS-SW e momento da pesquisa. Algumas ligações foram realizadas com esses campos. O nível foi ligado a todos

os outros campos, exceto ao momento da pesquisa que liga apenas ao nível MPS-SW.

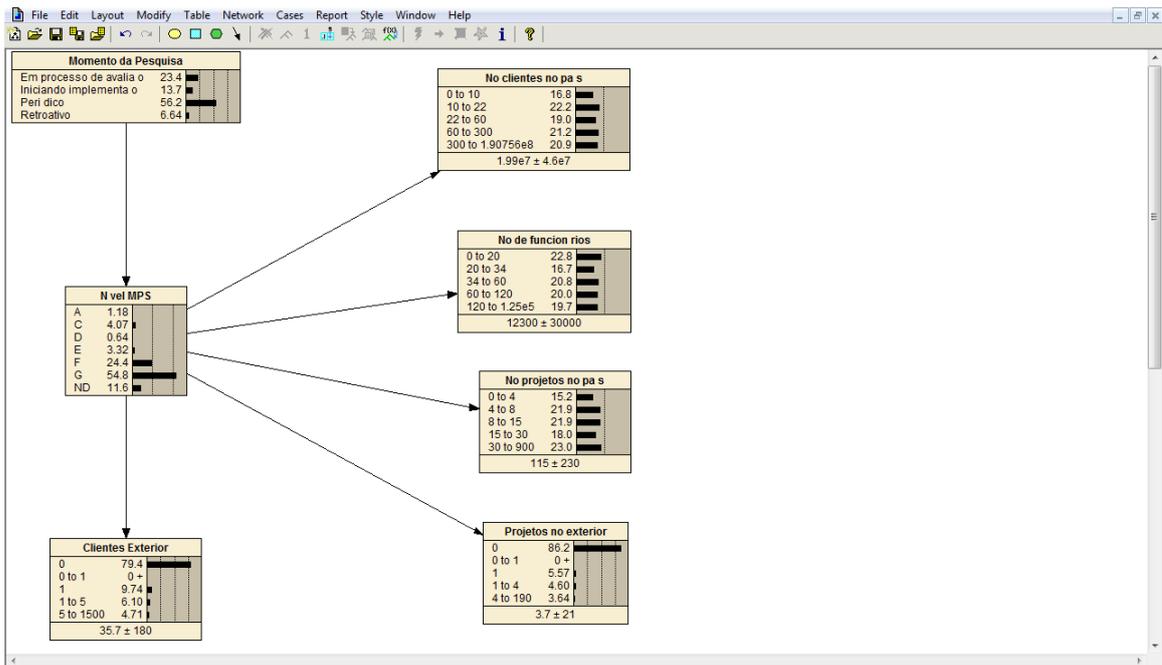


Figura 3.14: Primeira Rede Bayesiana

A segunda rede contém os campos que podem evidenciar o desempenho das organizações. Quais sejam: Tamanho dos projetos por pontos de função, precisão de estimativa, Def/PF, Tempo médio gasto nos projetos, esforço médio nos projetos, produtividade e Homem hora/ponto de função. Além desses o nível MPS-SW e o momento da pesquisa também foram incluídos nessa Rede. As proposições serão explicadas nas próximas seções.

As ligações na Rede estão diretamente relacionadas com as proposições que serão exploradas no próximo capítulo. As proposições que serão levantadas são uma tentativa de evidenciar os efeitos da adoção de um modelo de maturidade nas organizações, como o MPS-SW.

3.4 Análise Crítica

A rede Bayesiana foi elaborada através dos dados da pesquisa iMPS. Essa pesquisa é feita com o auxílio de um questionário composto por questões abertas. Assim, o funcionário responsável por inserir dados no sistema da pesquisa pode introduzir as informações como bem entender. Dessa forma, erros por descuido ou pelo não domínio do assunto podem ser adicionados aos campos da pesquisa iMPS.

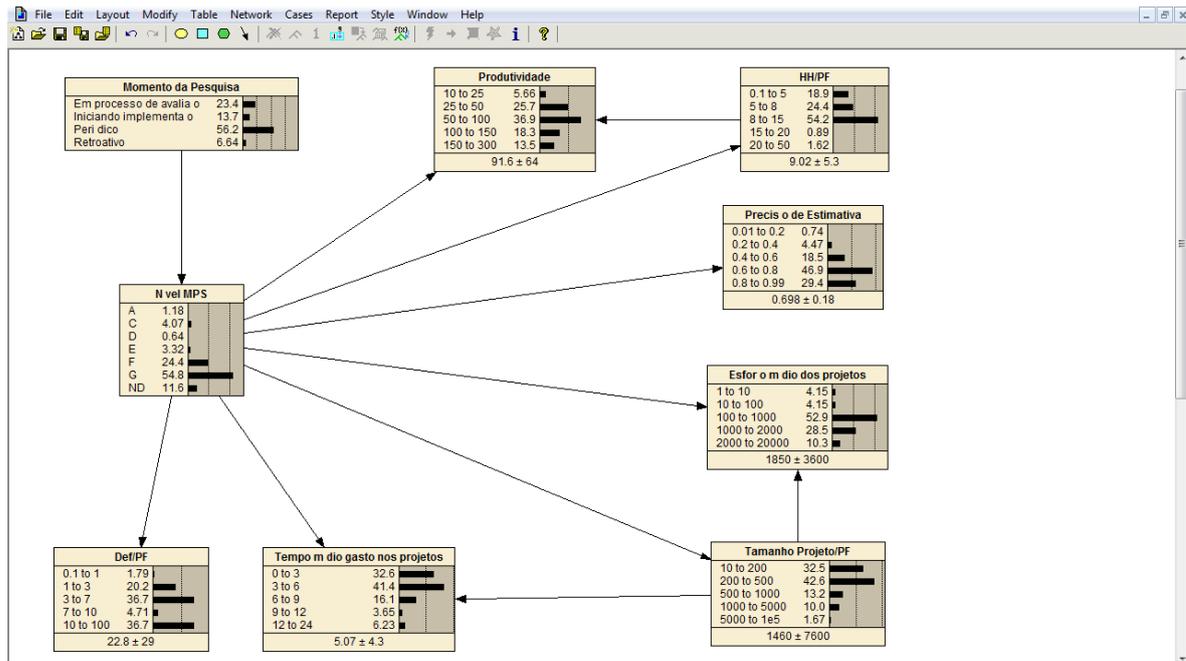


Figura 3.15: Segunda Rede Bayesiana

Por fim, um ponto que se deve ter muita atenção: a grande maioria dos dados está concentrada nos níveis de maturidade G e F. É prudente frisar essa questão, uma vez que algumas proposições são levantadas relacionando organizações de níveis G e F com outras de nível A, por exemplo. Apesar dos dados dos níveis E-A não serem maioria, tendem a ser muito representativos, pois são informados por organizações maduras, com grande experiência em desenvolvimento de software.

4 Explorando Efeitos da Adoção do MPS-SW no desempenho das organizações utilizando Redes Bayesianas

4.1 Identificação das proposições

Explorar os efeitos da adoção do modelo MPS-SW utilizando Rede Bayesiana foi um processo lento e gradativo. Após todos os nós serem escolhidos e a rede devidamente montada, o próximo passo foi explorar as Redes Bayesianas e identificar algumas proposições. As Redes montadas, principalmente que trata do desempenho, pode fornecer uma grande quantidade de combinações de nós.

O processo de identificação das proposições foi feito com testes sucessivos com base tentativa e erro de se achar questões interessantes. A ideia foi combinar os nós da rede e montar proposições relevantes que evidenciem efeitos da adoção do modelo MPS-SW nos processos das organizações.

Antes de mais nada é prudente definir o que é uma proposição. Na lógica matemática uma proposição respeita duas regras: a primeira é o princípio do terceiro excluído, ou seja, a proposição é uma sentença declarativa que pode somente ser julgada como verdadeira (V) ou falsa (F), e a segunda é o princípio da não contradição na qual ela nunca é classificada como verdadeira e falsa simultaneamente. Em síntese, ela nunca pode ser considerada "meio verdadeira e meio falsa".

Todas as proposições descobertas relacionam-se com o nível MPS-SW o que permite evidenciar a importância da maturidade de software para as organizações que adotaram o modelo. Assim, o nó "nível" foi combinado de diferentes maneiras com os outros nós da rede. Após o processo de utilização da rede para a descoberta das proposições serão apresentadas algumas análises dos resultados obtidos.

4.2 Utilização da rede: explorando os efeitos

4.2.1 Organizações de nível A e C gerenciam estimativas melhor que organizações de nível G

Essa proposição foi obtida selecionando o nível G no nó nível MPS e depois analisando as faixas de valores no nó precisão de estimativa conforme representado na figura 4.1. É possível observar que, no nível G, 28,9% das organizações possuem acerto nas estimativas muito alto, ou seja, estão na faixa mais alta de 0,8-0,99 e 45,6% possuem acerto alto, ficando na faixa de 0,6-0,8. Já o restante, 25,8%, está distribuído nas faixas de valores inferior a 0,6. No nível A, apesar dos poucos dados, todos estão distribuídos na faixa alta, ou seja de 0,6-0,8. No nível C, 52,9% está distribuído na faixa muito alta, ou seja precisão de estimativa entre 0,8-0,99 e 29,4% está distribuído na faixa de valores alta, entre 0,6 e 0,8, com o restante dos valores nas faixas mais baixa.

Com esses valores foi possível obter a proposição acima. Tais informações revelam que estar em níveis de maturidade mais altos levam as organizações a estimarem o prazo com mais precisão.

No nó “precisão de estimativa“, é importante lembrar que na montagem da rede bayesiana, os valores 1 e 0 não foram considerados. sendo que o valor 0 significa que a organização não faz qualquer tipo de estimativa e o valor 1 significa que a organização possui 100% de acerto em todos os seus projetos o que é improvável.

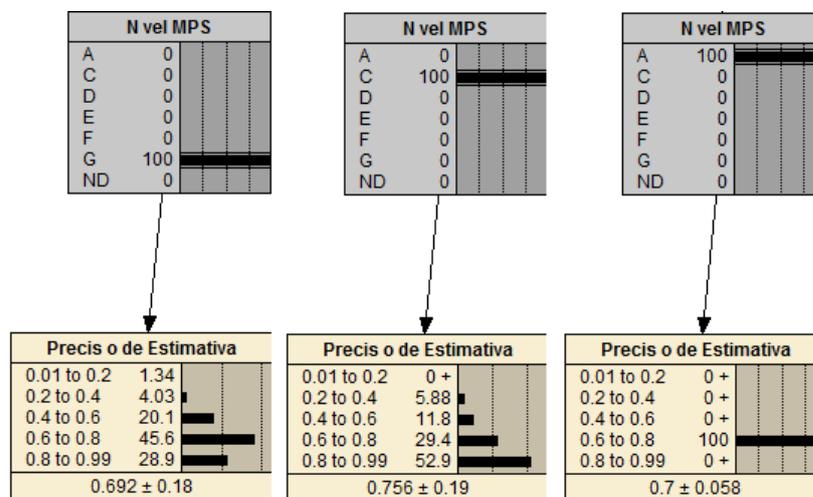


Figura 4.1: Primeira proposição

4.2.2 A produtividade de organizações do Nível F é maior em relação ao nível A. Porém, a qualidade dos projetos (Def/PF) desenvolvidos por organizações nível A é melhor em relação ao nível F

Essa proposição mostra que a produtividade é melhor no nível F, comparado ao nível A o que não significa que o nível F seja melhor. Evidencia, ainda, que o nível F é mais objetivo, uma vez que não existem tantos processos a ser executados e, dessa forma, a produção ganha em velocidade. Porém a velocidade tem um preço, já que as organizações nível A produzem menos defeitos por pontos de função. A figura 4.2 ilustra quantitativamente a proposição.

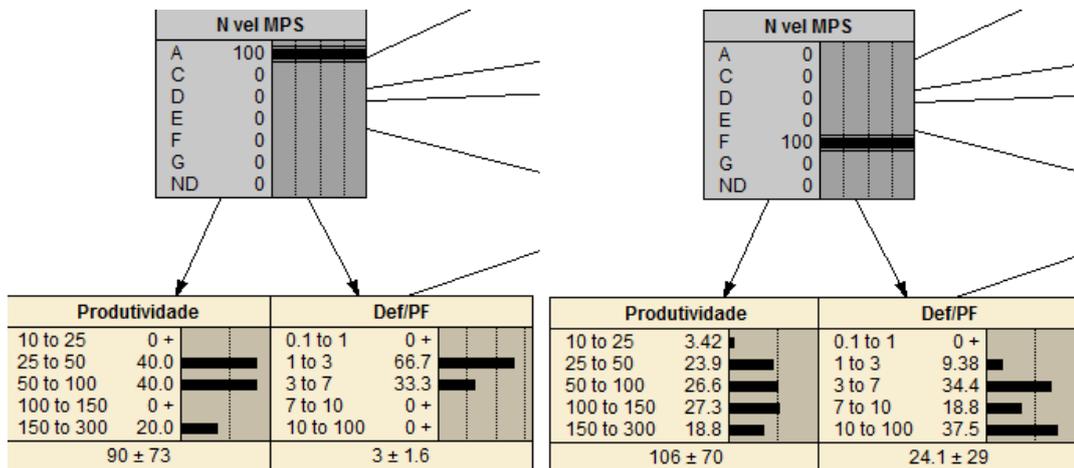


Figura 4.2: Segunda proposição

Para observar essa proposição na Rede Bayesiana, basta selecionar o nível A e analisar o comportamento dos nós de produtividade e Def/PF, o que também deve ser realizado para o nível F.

É válido verificar que no nível F, 50,5% dos valores de produtividade estão distribuídos nas faixas de valores baixa e média, e 44% estão nas faixas alta e muito alta. Já no nível A, a produtividade é de 80% nas faixas de valores baixa e média e os Def/PF no nível A estão distribuídos nas faixa de 1-4 e 4-7 Def/Pf. Já no nível F, 56,3% dos valores estão nas faixas mais altas, ou seja, possuem mais de 7 Def/Pf.

4.2.3 Organizações sem nenhum nível MPS-SW gastam menos tempo para desenvolver projetos muito pequenos do que organizações nível G

Essa proposição é importante para comparar o momento anterior a obtenção de um nível MPS-SW com organizações que já têm um nível, no caso, o nível G. Organizações nível G executam dois processos: a gerência de projetos e a gerência de requisitos. Por outro lado, organizações sem nível desenvolvem projetos sem processo definido formalmente pelo modelo MPS-SW.

Escolhemos essa proposição pois o número de projetos muito pequenos desenvolvidos por esses grupos são equivalentes. Do total de projetos desenvolvidos 34,5% são muito pequenos. A figura 4.3 ilustra quantitativamente a proposição.

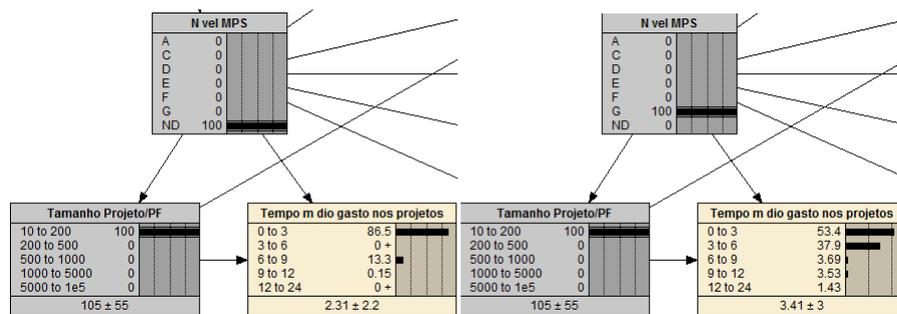


Figura 4.3: Terceira proposição

Para observar essa proposição na Rede Bayesiana, primeiro deve-se selecionar no nível a opção desejada, nesse caso ND ou G. Dessa forma é possível verificar que 34,5% dos projetos estão na primeira faixa de valores. A seguir, no nó Tamanho projeto/PF, deve-se selecionar a primeira faixa de valores, correspondente aos projetos muito pequenos.

4.2.4 Organizações nível F produzem mais e gastam menos do que organizações nível C

Essa proposição foi obtida observando os nós produtividade e HH/PF. A produtividade no nível F é maior. O número de HH/PF necessário no nível F é menor do que no nível C. A figura 4.4 ilustra quantitativamente esta proposição.

Para observar essa proposição na Rede Bayesiana, primeiro deve-se selecionar no

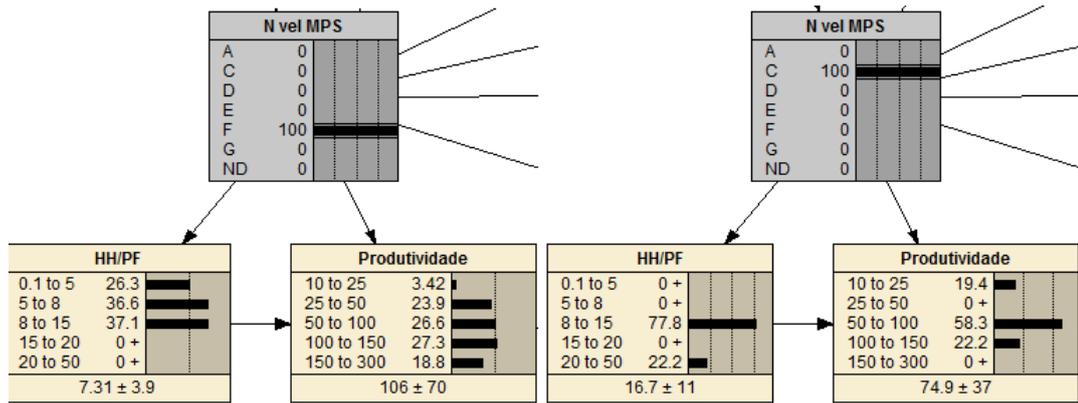


Figura 4.4: Quarta proposição

nó nível os campos F ou C. E depois observar o comportamento dos nós produtividade e HH/PF.

A produtividade no nível C é de 22,2% nas duas faixas de valores mais altas. Como mostrado na proposição 4.2.2, a produtividade no nível F é superior aos valores apresentados no nível C. Em relação ao nó HH/PF podemos observar que, no nível F, 62,9% utiliza até 8 homens por hora para desenvolver um ponto de função. Já no nível C, todos os projetos utilizam mais de 8 homens por hora para desenvolver um ponto de função.

4.2.5 Em organizações nível C os projetos gastam mais HH/PF do que no nível F. Porém a qualidade (Def/PF) dos projetos nível C são melhores

Essa proposição é um complemento da anterior. A primeira parte refere-se aos HH/PF. No nível C gasta-se mais HH/PF comparado ao nível F, porém, como consequência, no nível C o número de defeitos por pontos de função é menor, o que leva a um ganho de qualidade. A figura 4.5 ilustra quantitativamente a proposição.

Podemos observar que no nível F o número de homem hora por ponto de função que utilizam até 8 homens para desenvolver um ponto de função é 62,9%, ou seja, as duas faixas de valores mais baixas. Já no nível C, todos os projetos utilizam mais de 8 homens por hora para desenvolver um ponto de função. Porém, organizações nível C produzem no máximo 7 defeitos por ponto de função. Já 56,2% das organizações nível F produzem

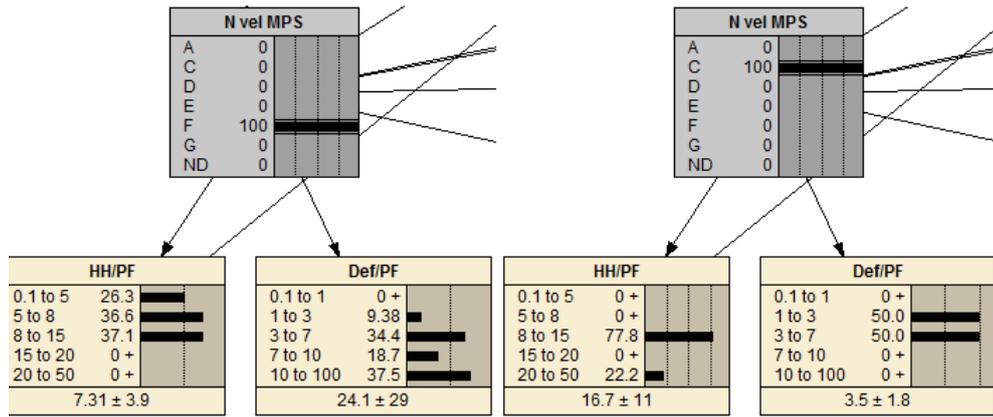


Figura 4.5: Quinta proposição

mais de 7 defeitos por ponto de função.

4.2.6 Os projetos desenvolvidos no nível F, são maiores que os projetos desenvolvidos no nível G

Nessa proposição podemos observar que os projetos desenvolvidos por organizações nível F são maiores que os desenvolvidos por organizações nível G. A figura 4.6 ilustra quantitativamente a proposição.

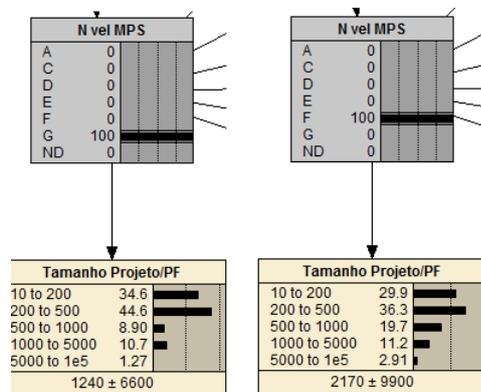


Figura 4.6: Sexta proposição

Nas três faixas de valores mais elevados, os projetos do nível F correspondem a 33,9% do total e são maiores que os projetos nível G, os quais para esse conjunto de valores possuem 20,9% dos valores. Em outras palavras, há uma tendência das organizações que atingem o nível F de maturidade de trabalhar com projetos maiores.

4.2.7 Organizações nível A possuem mais funcionários que organizações nível G

Organizações nível A possuem alto número de funcionários e nenhuma delas possuem menos de 20 funcionários. Em organizações nível G, aproximadamente 25% das organizações possuem menos de 20 funcionários.

Essa proposição foi obtida da rede de caracterização das organizações (primeira rede). Para observá-la, basta selecionar no nó nível as faixas A e G, e observar as informações no nó “número de funcionários”. A figura 4.7 ilustra quantitativamente essa proposição.

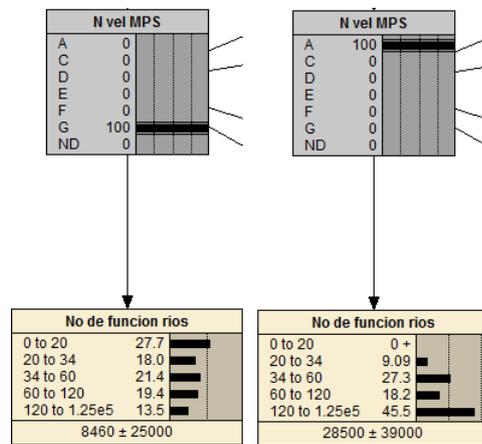


Figura 4.7: Sétima proposição

4.2.8 Organizações nível A e C tendem a ter mais clientes no exterior que organizações nível G

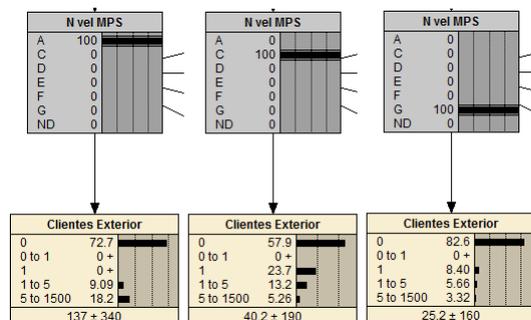


Figura 4.8: Oitava proposição

A figura 4.8 ilustra quantitativamente esta proposição, ela também foi obtida

da rede de caracterização das organizações (primeira rede). Evidencia que organizações com nível de maturidade mais alto tendem a ter mais clientes no exterior. Elas, além de possuírem mais clientes no exterior, também apresentam um número menor de 0 clientes, ou seja, existem menos organizações níveis A e C que não tem clientes no exterior, comparadas a organizações nível G.

4.3 Análise das proposições

Podemos observar nas proposições listadas algumas questões:

- O nível F possui resultados excelentes em métricas como produtividade e homem hora por ponto de função. Esse nível apresenta bons resultados para as métricas porque os processos nele desenvolvidos nele garantem a boa execução do projeto. Com base na análise das proposições é possível observar que os níveis mais elevados apresentam uma queda nessas questões, pois a inclusão de novos processos pode fazer com que todos os processos fiquem mais lentos. Como nos níveis de maturidade mais elevados muitos processos são executados, é necessário utilizar mais funcionários para executar todos os processos e, possivelmente é, nesse ponto que a empresa perde em produtividade.
- O mesmo nível F apresenta uma queda em questões de qualidade comparado a níveis mais elevados de maturidade. Nessas organizações, o desenvolvimento do software é realizado sem, por exemplo, os processos de validação e verificação, os quais estão presentes no nível D e impactam diretamente na qualidade.
- Essas proposições mostram que adotar o modelo MPS-SW impacta nas diferentes métricas adotadas no desenvolvimento de software. Logo, estar em níveis de maturidade mais elevados não está diretamente relacionado a obter os melhores resultados para todas as métricas relacionadas neste trabalho.

5 Considerações Finais

O objetivo desse trabalho foi identificar, organizar e analisar resultados da adoção do modelo MPS-SW referentes a dados da pesquisa iMPS, respondida anualmente por diferentes organizações, utilizando Redes Bayesianas. Para isso, apresentou-se uma breve revisão bibliográfica sobre a importância dos modelos de maturidade nas organizações de software, as Redes Bayesianas e a pesquisa iMPS.

Posteriormente, foi feita uma explicação da montagem da Rede Bayesiana e os critérios para a sua montagem. Esse trabalho fornece proposições que tentam ajudar a compreender os fatores relevantes da adoção do modelo de maturidade MPS-SW.

É necessário reafirmar que os dados da pesquisa iMPS não foram validados pois, para isso, seria necessário pessoalmente realizar uma entrevista com todos os envolvidos na pesquisa iMPS. Dessa forma, utilizamos os dados da maneira da maneira que foram fornecidos pela pesquisa e com base neles desenvolvemos esse trabalho.

Para a futura evolução desse trabalho, as seguintes sugestões podem ser seguidas:

- Fazer uma análise por ano em que a pesquisa iMPS foi desenvolvida e analisar a evolução do modelo ano a ano;
- Adicionar à pesquisa o nó categorias de serviços, presente na pesquisa iMPS. Existem várias categorias presentes na pesquisa, o que possibilita uma análise a respeito dessa informação e qual o impacto dela no modelo MPS-SW;
- A pesquisa iMPS poderia evoluir. Algumas informações geram dificuldades de entendimento, principalmente para as organizações que respondem o questionário da pesquisa iMPS. Em relação ao questionário, alguns campos contêm informações claramente incorretas. Esses campos poderiam utilizar questões fechadas. Por exemplo, se a pergunta é: Quanto tempo gasta em média para desenvolver os seus projetos? A resposta deveria ser um número (com base na quantidade de meses) e esse campo não deveria aceitar letras, apenas números. Dessa forma os erros poderiam ser diminuídos. E esse é apenas um exemplo, de vários tratamentos que poderiam

ser realizados no formulário dessa pesquisa.

Referências Bibliográficas

- Basye, K. **Map learning**. 1990.
- Dempster, A.; Laird, N. ; Rubin, D. **Maximum likelihood from incomplete data via the em algorithm**. volume 39, p. 1–38, 1977.
- Hecherman, P. **Sistema para diagnósticos de problemas nas glândulas linfáticas**. 1990.
- Kalinowski, M.; Santos, G.; Reinehr, S.; Montoni, M.; Rocha, A.; Weber, K. ; Travassos, G. **Mps.br: Promovendo a adoção de boas práticas de engenharia de software pela indústria brasileira**. 2010.
- Kalinowski, M.; Santos, G.; Prikladnicki, R.; Rocha, A. R.; Weber, K. C. ; Antonioni, J. A. **From software engineering research to brazilian software quality improvement**. In: SBES, p. 120–125. IEEE, 2011.
- Kalinowski, M.; Weber, K. C. ; Travassos, H. G. **Imps: An experimentation based investigation of a nationwide software development reference model**. 2008.
- Mendes, E.; Mosley, N. **Bayesian network models for web effort prediction: A comparative study**. volume 34, p. 723–737, 2008.
- Center, N. A. R. **Autoclass**. 1998.
- Pearl, J. **Causality: Models, reasoning and inference**. cambridge university press, isbn 978-0-521-77362-1. 2000.
- Pedreti Chagas, R. **Aplicação de redes bayesianas na previsão de fluxo de caixa**. 2008.
- Simões, C.; Lasmar, C. ; Santos, C. **Implantação dos processos gerência de projeto e medição com auxílio de ferramenta baseada em planilhas**. 2010.
- Softex. **Mr mps sw guia geral mps de software**. In: www.softex.br/mpsbr, disponível em <http://www.softex.br/mpsbr/>, último acesso em 20/07/2014, 2012.
- Souza, W.; Ramasco, M.; Mattos, A. ; Pinheiro, E. **Mps.br nível a: Experiência da stefanini**. 2010.
- Travassos, G. H.; Kalinowski, M. **imps 2013 evidências sobre o desempenho das empresas que adotaram o modelo mps-sw**. softex, isbn 978-85-99334-75-1. 2013.
- Viana, D.; Rabelo, J.; Mar, C.; Aguiar, E.; Wagner, P.; Vilela, D. ; Conte, T. **Resultados de um estudo qualitativo sobre a implementação do modelo mps em empresas do programa amazonsoft**. 2010.