UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

UM ESTUDO DE CASO EM AVALIAÇÃO DE RISCOS DE SEGURANÇA DA INFORMAÇÃO UTILIZANDO OTMIZAÇÃO POR COLÔNIA DE FORMIGAS E REDES BAYESIANAS

ELIAS PEREIRA SILVA

ORIENTADOR: DANIEL GUERREIRO E SILVA

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
ÁREA DE CONCENTRAÇÃO INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO

PUBLICAÇÃO: PPGENE.DM - 625A/16

BRASÍLIA / DF: 12/2016

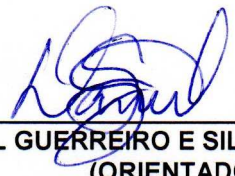
UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

UM ESTUDO DE CASO EM AVALIAÇÃO DE RISCOS DE SEGURANÇA DA INFORMAÇÃO UTILIZANDO OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS E REDES BAYESIANAS

ELIAS PEREIRA SILVA

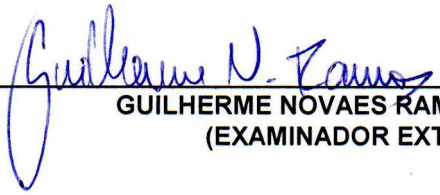
DISSERTAÇÃO DE MESTRADO PROFISSIONAL SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:



DANIEL GUERREIRO E SILVA, Dr., ENE/UNB
(ORIENTADOR)

RAFAEL TIMÓTEO DE SOUSA JÚNIOR, Dr., ENE/UNB
(EXAMINADOR INTERNO)



GUILHERME NOVAES RAMOS , CIC/UNB
(EXAMINADOR EXTERNO)

Brasília, 06 de Dezembro de 2016.

FICHA CATALOGRÁFICA

SILVA, ELIAS PEREIRA
Um estudo de caso em avaliação de Riscos de Segurança da Informação utilizando Otimização por Colônia de Formigas e Redes Bayesianas [Distrito Federal] 2016.
xvi, 55p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2016).

Dissertação de Mestrado – Universidade de Brasília, Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. AVALIAÇÃO DE RISCOS DE SEGURANÇA 2. COLÔNIA DE FORMIGAS
3. REDES BAYESIANAS

I. ENE/FT/UnB. II. Título (Série)

REFERÊNCIA BIBLIOGRÁFICA

SILVA, E. P. (2016). Um estudo de caso em avaliação de riscos de segurança da informação utilizando colônia de formigas e redes Bayesianas. Dissertação de Mestrado, Publicação PPGENE.DM - 625A/16, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 55p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Elias Pereira Silva
TÍTULO DA DISSERTAÇÃO: Um estudo de caso em avaliação de riscos de segurança da informação utilizando Otimização por Colônia de Formigas e redes Bayesianas.
GRAU/ANO: Mestre/2016.

Elias Pereira Silva
QNP 14 Conjunto D Casa 42
CEP 72.231-404 – Brasília – DF - Brasil

DEDICATÓRIA

A minha esposa e filhos, pelo amor e compreensão.

Aos meus pais e irmã, pelo apoio de sempre.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Daniel Guerreiro e Silva, pelo constante apoio, incentivo, dedicação e amizade essenciais para o desenvolvimento deste trabalho e para o meu desenvolvimento como pesquisador.

Aos Professores das disciplinas e a todos que se dedicaram a apoiar o curso.

A todos os colegas do Mestrado em Informática Forense, pela amizade.

A Sarah Costa, pelas palavras de apoio e incentivo.

A todos, os meus sinceros agradecimentos.

O presente trabalho foi realizado com o apoio do Departamento Polícia Federal – DPF, com recursos do Programa Nacional de Segurança Pública com Cidadania – PRONASCI, do Ministério da Justiça.

RESUMO
UM ESTUDO DE CASO EM AVALIAÇÃO DE RISCOS DE SEGURANÇA DA INFORMAÇÃO UTILIZANDO OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS E REDES BAYESIANAS

Autor: Elias Pereira Silva
Orientador: Daniel Guerreiro e Silva
Programa de Pós-graduação em Engenharia Elétrica
Brasília, Dezembro de 2016

O trabalho descrito nesta dissertação apresenta um estudo de caso de construção de uma Rede Bayesiana através da Otimização por Colônia de Formigas utilizando um banco de dados de treinamento formado pelos dados de controles de segurança de um sistema de informação da Polícia Federal e de especialistas em segurança da informação. Esta abordagem é baseada no Modelo de Análise de Risco de Segurança proposto por Feng *et al*. (2014), que foi ajustado para os requisitos das leis do governo brasileiro e sua principal vantagem é a possibilidade de análise do relacionamento entre as vulnerabilidades de diferentes ativos que compõem o mesmo sistema de informação.

As Redes Bayesianas são ferramentas poderosas porque fornecem respostas sobre questões probabilísticas de diferentes domínios como, por exemplo, o gerenciamento de risco de segurança para sistemas de informação. Estas estruturas foram introduzidas na Gestão de Riscos pela necessidade de monitoramento contínuo do risco, pela capacidade de demonstrar visualmente as relações entre variáveis e pela possibilidade de reavaliação frente a novas evidências. A Otimização por Colônia de Formigas é uma das diferentes técnicas de sucesso usadas para descobrir as estruturas das Redes Bayesianas a partir de um conjunto de dados de treinamento. Feng *et al*. (2014) aplicaram esses conhecimentos no domínio da segurança da informação, tendo em vista as recomendações de conformidade e operação do NIST.

Este estudo de caso compreende a primeira fase do modelo SRAM (FENG *et al*, 2014), composta pelas etapas de coletar dados de controles de segurança aplicados a todos os componentes de um sistema de informação e implementar o algoritmo ACO para identificar a relação entre esses fatores de risco e obter uma medida global do risco por sistema.

ABSTRACT
CASE STUDY IN INFORMATION SECURITY RISK ASSESMENT USING ANT COLONY OPTIMIZATION AND BAYESIAN NETWORKS

Author: Elias Pereira Silva
Supervisor: Daniel Guerreiro e Silva
Programa de Pós-graduação em Engenharia Elétrica
Brasília, December of 2016

The work described in this dissertation presents a case study of building a Bayesian Network via Ant Colony Optimization using a merged training database on security controls data from a Federal Police information system and information security experts. This approach is based on Security Risk Analysis Model proposed from Feng et al. (2014), and wich was adjusted for the brazilian government laws requirements and your main advantage are the possibility of relationship analysis between the vulnerabilities from different assets wich compound a same information system.

Bayesian Networks are powerful tools because they provide answers about probabilistic questions from different domains such as security risk management for information systems. These structures were introduced in Risk Management by the need for continuous risk monitoring, by the ability to visually demonstrate the relationships between variables and by the possibility of being reevaluated in the face of new evidence. Ant Colony Optimization it's one of different success techniques used to discover the structures of the Bayesian networks from a training data set. Feng *et al*. (2014) applied these knowledges in the information security domain, in view of NIST compliance and operation recommendations.

This case study comprises the first phase from SRAM model (FENG *et al*, 2014), composed of the steps of collecting security controls data applied to all components of an information system and implementing the ACO algorithm to identify the relationship between these risk factors and obtain an overall measure of risk by system.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIAÇÕES

ABNT	Associação Brasileira de Normas Técnicas
ACO	Do inglês *Ant Colony Optimization*, Otimização por Colônia de Formigas
EM	Algoritmo *Expectation Maximization*
GAO	Grafo Acíclico Orientado
GRSI	Gestão de Riscos de Segurança da Informação
ISACA	Do inglês *Information Systems Audit and Control Association,* Associação de Auditoria e Controle de Sistemas de Informação
ISO	Derivado do grego *isos*, que quer dizer igual, em inglês *International Organization for Standardization*, Organização Internacional para Padronização
MLE	Do inglês *Maximum Likelihood Estimation*, Estimativa por Máxima Verossimilhança
NIST	Do inglês *National Institute of Standards & Technology (US)*, Instituto Nacional de Padrões e Tecnologia dos EUA
PDCA	Do inglês *Plan-Do-Check-Act,* é um método de gestão para controle e melhoria contínua
RB	Rede Bayesiana
SANS	Do inglês *System Administration, Networking, and Security Institute*, Organização de pesquisa e treinamento em segurança da informação
SGSI	Sistema de Gestão de Segurança da Informação
SRAM	Do inglês *Security Risk Analysis Model,* Modelo de Análise de Risco de Segurança
TPC	Tabela de Probabilidade Condicional
UnbBayes	Framework para inferência em redes probabilísticas

1. INTRODUÇÃO

Independentemente do método adotado para Gestão de Riscos de Segurança da Informação (GRSI), três grandes etapas compõem este processo: Identificação dos Riscos (mapeamento de ameaças, vulnerabilidades, probabilidade de materialização, estimativa de impacto); Mitigação dos Riscos (escolha de estratégias para eliminação ou minimização dos riscos encontrados, considerando que os recursos empregados não podem ser maiores que o possível impacto da ocorrência do risco); Monitoramento do risco (os riscos devem ser continuamente analisados, visto que as ameaças vulnerabilidades e os próprios ativos alteram-se com o passar do tempo) (OLIVEIRA, 2006).

Conforme opinião de Wheeler (2016) é possível perceber a necessidade de desenvolvimento de ferramentas que tratem em uma mesma visão o conceito de GRC Digital, que corresponde às áreas de governança, risco e conformidade. Alinhado com essa direção, está o relatório do grupo Gartner (PRATAP; WHEATMAN, 2016) que estima para os próximos 2 a 5 anos a abordagem de todos os esses componentes por uma mesma ferramenta, o que trará facilidades na gestão de atividades associadas a componentes de negócios digitais, como computação em nuvem, mídias sociais, dispositivos móveis e a internet das coisas.

Considerando a necessidade de avaliar o relacionamento de riscos individuais aos quais estão sujeitos os sistemas computacionais e com o objetivo de fornecer uma forma proativa de tratá-los, foram verificados trabalhos recentes ligados ao Monitoramento Contínuo do Risco uma vez que esta etapa permite a visibilidade durante todo o processo de gestão de riscos e contribui para a identificação de situações em que uma manifestação conjunta de vulnerabilidades pudesse levar a riscos não considerados sob uma perspectiva isolada. (KOTT; ARNOLD, 2013).

Destacou-se entre tais pesquisas o modelo *Security Risk Analysis Model* (SRAM), proposto por Feng *et al.* (2014), o qual propõe uma abordagem para avaliação contínua do risco, considerando um banco de dados histórico de casos, as definições de especialistas sobre diferentes áreas de vulnerabilidades e o recebimento de novos dados, em tempo real. A partir destas fontes de dados, constrói-se uma rede Bayesiana que define os fatores de risco e seus relacionamentos causais. Um algoritmo de Otimização por Colônia de Formigas é utilizado para inferir o modelo da rede e para realizar a análise de propagação

das vulnerabilidades encontradas, determinando os caminhos com maior probabilidade de ocorrência e com maior risco estimado.

No contexto apresentado, propõe-se realizar a construção de uma rede Bayesiana que reflita o relacionamento entre os fatores de risco de um sistema computacional em produção na Polícia Federal, baseando-se no monitoramento dos controles de segurança a fim de validar se esta proposta traz benefícios relevantes em termos da gestão de segurança dos sistemas de informação, atuando como uma ferramenta de suporte à tomada de decisão.

Face à importância da utilização de dados reais frente à publicação de novas abordagens para análise de risco de segurança da informação, espera-se fornecer uma estimativa conjunta do risco de segurança e mensurar os impactos nos serviços de Tecnologia da Informação através do monitoramento de diferentes *hardware*s e *software*s que constituem sistemas computacionais utilizados pela Polícia Federal utilizando o modelo SRAM para orientação deste processo.

1.1. JUSTIFICATIVA E OBJETIVOS

É possível notar que grande parte das informações, tanto do governo quanto das empresas, passou a ser disponibilizada eletronicamente sob a forma de sistemas computacionais, em redes internas ou na rede mundial de computadores. A necessidade de proteção das informações ocorre devido à importância desses ativos para o negócio das empresas. Independente do local em que estejam armazenados ou da forma que assumam, elas devem ser sempre protegidas adequadamente a fim de garantir a continuidade dos negócios, minimizar os prejuízos e maximizar o retorno de investimentos. A defesa desses dados denomina-se Segurança da Informação e é obtida com o estabelecimento de diferentes tipos de controles – políticas, processos, procedimentos, estruturas organizacionais e funções de *hardware* e *software* – continuamente melhorados (SILVA, 2011; TCU, 2012).

Este trabalho aborda o estudo de caso de aplicação de uma técnica para monitoramento da infraestrutura de sistemas computacionais de modo que forneça subsídios para as decisões da Gestão de Riscos de Segurança da Informação. O modelo que será aplicado envolve o monitoramento dos controles de segurança descritos pela norma de segurança da informação ABNT ISO/IEC 27002, mapeia o relacionamento dessas medidas

em uma rede Bayesiana utilizando a técnica de Otimização por Colônia de Formigas com o objetivo de fornecer uma medida de risco conjunta para os componentes do sistema de informação analisado.

O objetivo geral deste trabalho é reforçar as atividades de segurança da informação em infraestruturas de sistemas computacionais críticos através do desenvolvimento de uma visão conjunta de monitoramento dos fatores de riscos dos seus componentes. E, como forma de atingir essa proposta, sugere-se o cumprimento dos seguintes pontos:

- Desenvolver as etapas necessárias para modelar o relacionamento entre os fatores de risco de segurança da informação dos componentes de um determinado sistema computacional.
- Produzir uma medida conjunta do risco ao qual um sistema computacional está submetido, considerando os diferentes *hardwares* e *softwares* que o compõem.

Além disso, as normas às quais estão sujeitos os órgãos da Administração Pública Federal e a realidade dos recursos disponíveis também devem ser levadas em consideração para a identificação das contribuições e objeções desta abordagem para uso pelas equipes de segurança da informação.

1.2. ORGANIZAÇÃO DO TEXTO

No Capítulo 2 é realizada uma revisão geral dos conceitos que envolvem o modelo de avaliação de riscos aplicado no estudo de caso. Desde as questões referentes aos problemas de segurança da informação até o detalhamento técnico do desenvolvimento das redes Bayesianas e do uso da meta-heurística por Colônia de Formigas.

O Capítulo 3 é dedicado ao detalhamento da construção da rede Bayesiana a partir dos dados dos controles de segurança da informação, conforme proposto por Feng *et al.* (2014). Além disso, descreve-se qual o fluxo de dados necessário para se chegar à estrutura de rede e quais as análises possíveis de se realizar a partir da topologia definida.

O estudo de caso foco deste trabalho é detalhado no Capítulo 4. Nele é apresentada a visão do autor para a aplicação do modelo original em um ambiente de dados reais, quais foram as adaptações realizadas com base na realidade de normas brasileiras e os resultados obtidos pela aplicação do modelo.

Por fim, no Capítulo 5, são realizadas as conclusões pertinentes ao trabalho, apresentando as situações de dificuldades e de sucesso durante as etapas de coleta de dados, reprodução do modelo, experiência obtida pelo autor durante o desenvolvimento dos trabalhos e as perspectivas futuras.

2. ESTUDO BIBLIOGRÁFICO

Como o estudo de caso aplicado neste trabalho emprega conceitos estatísticos em um problema de segurança da informação, busca-se, neste capítulo, fornecer subsídios que contribuam com a compreensão das diferentes disciplinas que envolvem tal modelo de análise de riscos de segurança. Inicialmente, é apresentado o contexto de boas práticas na gestão de riscos de segurança da informação (Seção 2.1) e como o processo de monitoramento contínuo se insere neste campo (Seção 2.2). Em seguida, procura-se expor como o raciocínio bayesiano é incorporado sob o formato visual das redes Bayesianas (Seção 2.3) e como a meta-heurística de Otimização por Colônia de Formigas pode ajudar na inferência dessas estruturas (Seção 2.4).

2.1. GESTÃO DE RISCOS DE SEGURANÇA DA INFORMAÇÃO

O processo de Gestão de Riscos de Segurança da Informação faz parte do Sistema de Gestão de Segurança da Informação (SGSI) e permite às instituições identificar o valor dos ativos de informação (elementos que suportam os processos de negócio), identificar as fragilidades que possam ser exploradas para comprometer a segurança da informação (vulnerabilidades) e identificar os eventos (ameaças) que possam explorar essas vulnerabilidades por meio de um acesso não autorizado, destruição, divulgação, modificação de informações ou da negação de serviço (ABNT, 2013a; ABNT, 2011; NIST, 2006).

Publicados por organizações bem estabelecidas e de importante papel no cenário de definição de requisitos e de padrões, existem distintos modelos para orientar a execução do processo de Gestão de Riscos de Segurança da Informação. Dos quais se podem destacar as publicações da Organização Internacional para Padronização (ISO), do Instituto Nacional de Padrões e Tecnologia dos Estados Unidos (NIST) e da Associação de Auditoria e Controle de Sistemas de Informação (ISACA) (BANERJEE, C.; BANERJEE, A., 2014).

A Organização Internacional para Padronização possui a série de normas 27000, que tem por objetivo tratar a gestão de segurança da informação e, no ano 2011, publicou a versão mais recente da norma *ABNT ISO/IEC 27005 Tecnologia da informação – Técnicas de segurança – Gestão de riscos de segurança da informação* (ABNT, 2011), que

apresenta as diretivas para condução da gestão de riscos de segurança da informação no formato do ciclo de melhoria contínua "PDCA" e com aplicação válida para todos os tipos de organizações. A Tabela 1 expõe as atividades chave da gestão de riscos segundo esta norma.

ETAPA	Atividade da gestão de riscos de segurança da informação (ISO)
P - Planejar	• Definição do Contexto • Análise/Avaliação de riscos • Tratamento do risco • Aceitação do risco
D - Executar	• Implementação do plano de tratamento do risco
C - Verificar	• Monitoramento contínuo e análise crítica do risco
A - Agir	• Manutenção e melhoria contínua do processo

Tabela 1. Atividades GRSI (ABNT ISO/IEC 27005)

No mesmo contexto, o NIST possui a série de publicações especiais SP 800, que provê materiais de referência e recomendações na área de segurança de computadores. A versão mais recente do processo macro de Gestão de Riscos de Segurança da Informação foi publicada em 2011 através do documento *NIST Special Publication 800-39: Managing Information Security Risk* (NIST, 2011a) e também prevê um formato de ciclo de atividades para as suas 6 etapas, detalhadas em diferentes publicações, conforme a Tabela 2.

ETAPA	Atividade da gestão de riscos (NIST)
Categorizar	• Classificar sistemas de informação (SP 800-60)
Selecionar	• Escolher controles de segurança (SP 800-53)
Executar	• Aplicar controles de segurança (SP 800-70)
Avaliar	• Avaliar controles de segurança (SP 800-53A)
Autorizar	• Autorizar sistemas de informação (SP 800-37)
Monitorar	• Monitoramento contínuo (SP 800-137)

Tabela 2: Atividades GRSI (NIST SP 800-39)

Por fim, a organização ISACA possui em seu portfólio de produtos de governança de TI o modelo Risk IT (ISACA, 2009), que descreve as principais atividades, os fluxos de informação e a gestão de desempenho de cada um dos três processos que compõem a gestão de riscos de TI deste modelo. Os princípios chave destes processos envolvem o alinhamento com a gestão de risco do negócio global da empresa, o funcionamento dos processos em modo contínuo, o equilíbrio entre os custos e benefícios do gerenciamento de riscos. A Tabela 3 descreve a estrutura desse modelo.

Neste cenário de gestão de riscos, os controles de segurança são empregados com o objetivo de gerenciar o nível dos riscos de segurança da informação e, nesta categoria, é admitido qualquer mecanismo administrativo, físico ou operacional que possua tal capacidade. Em geral, os controles podem fornecer um ou mais dos seguintes tipos de proteção: correção, eliminação, prevenção, minimização do impacto, dissuasão, detecção, recuperação, monitoramento e conscientização. Como controles de segurança compreendem, por exemplo, procedimentos, políticas, estruturas organizacionais, antivírus, fechaduras, cópias de segurança, entre outros (ABNT, 2011; BEZERRA, 2013).

ETAPA	Atividade da gestão de riscos de segurança da informação (Risk IT)
Governança	• Integrar com a gestão de risco • Tomar decisões de risco de negócio • Estabelecer e manter uma visão do risco comum
Avaliação	• Analisar o risco • Coletar o risco • Manter o perfil do risco
Resposta	• Gerenciar o risco • Reagir a eventos • Articular o risco

Tabela 3: Atividades GRSI (ISACA Risk IT)

O monitoramento do risco está presente de forma direta nos três modelos apresentados: ISO (Monitoramento contínuo e análise crítica do risco), NIST (Monitoramento contínuo) e ISACA (Reagir a eventos) e se mostra como uma etapa chave dentro do ciclo da gestão de riscos porque permite o acompanhamento da eficácia dos resultados frente aos controles de segurança aplicados e a percepção da evolução do risco.

2.2. MONITORAMENTO CONTÍNUO DO RISCO

O principal propósito da etapa de monitoramento do risco é observar, de modo sistemático e regular, o desenvolvimento da gestão de riscos. Após a observação do ambiente, percebe-se a necessidade de avaliar criticamente os resultados obtidos a fim de identificar problemas e pontos de melhoria.

O guia de aplicação do modelo de gestão de riscos publicado pelo NIST (NIST, 2010), por exemplo, enfatiza a importância da gestão de risco em tempo quase real (STONEBRAKER, 2005) e da autorização de sistemas através de processos fortes e

eficazes de monitoramento contínuo. Neste guia, os processos de monitoramento suportam diretamente três etapas do ciclo (Categorizar, Avaliar, Monitorar) e indiretamente contribuem com as demais etapas. Eles também permitem aos profissionais de segurança da informação manter uma visão do estado do risco de diversos componentes da infraestrutura (dados, rede, estações, aplicações, etc.) e, quando aplicado com base em um sistema de gerenciamento de eventos, é possível realizar a correlação de ocorrências em diferentes ativos, o que contribui para identificação e resposta rápida aos incidentes (NIST, 2011b).

Reconhecendo que as ameaças, vulnerabilidades, probabilidades de ocorrência ou de consequência podem mudar abruptamente, sem qualquer indicação, a norma ABNT ISO/IEC 27005 (ABNT, 2011) recomenda que a etapa de monitoramento dos riscos ocorra continuamente. Neste mesmo sentido, o centro de segurança norte-americano *SANS Institute* recomenda o uso de técnicas de tratamento automatizado de dados na execução e avaliação periódica dos 20 controles de segurança críticos, os quais são mantidos através de publicações pelo Centro para a Segurança da Internet (CIS, 2015). Essa abordagem minimiza a propagação de um ataque em uma rede comprometida e reserva o esforço de decisão humana apenas para a análise de resultados em cima de dados previamente selecionados.

Assim, considerando a afirmação de Russel e Norvig (2004) que os sistemas modernos de Inteligência Artificial para inferência probabilística são todos baseados no teorema de Bayes, a próxima Seção será dedicada a descrever como as Redes Bayesianas podem contribuir com sistemas de monitoramento e análise para fornecer suporte à tomada de decisão. Neste estudo de caso em específico, as redes Bayesianas são utilizadas na etapa de análise de riscos de segurança da informação.

2.3. REDES BAYESIANAS

As regras básicas da teoria de probabilidade fundamentam a estrutura para se quantificar e manipular os aspectos de incerteza de um conjunto de dados e, quando associadas à teoria de decisão, formam a base para se realizar inferências e predições assertivas acerca dos dados analisados, mesmo que sejam dados incompletos ou ambíguos (BISHOP, 2006).

De acordo com Duda *et al.* (2000) as redes de causalidade constituem uma forma de representação gráfica – composta por um conjunto de variáveis e um conjunto de conexões entre essas variáveis – na qual cada variável pode representar os possíveis estados de um evento, mas se apresenta em apenas um desses estados, mesmo que seja desconhecido, enquanto que as conexões traduzem os aspectos de dependência entre as variáveis. Estas estruturas são denominadas como Redes Bayesianas e possuem a capacidade de representar visualmente os conceitos da probabilidade condicional (JENSEN, 1996), o qual está no centro do tratamento Bayesiano da incerteza, e da distribuição de probabilidade conjunta total de um domínio (RUSSEL; NORVIG, 2004), o que permite estimar variáveis pela constatação de alguma evidência.

Conforme Russel e Norvig (2004), a probabilidade incondicional (*a priori*) pode ser entendida como o grau de crença acordado para uma proposição *A* na ausência de quaisquer outras informações, enquanto que a probabilidade condicional é observada quando se obtém alguma evidência relativa à variável anteriormente desconhecida.

Dado um evento X, a probabilidade condicional do evento A ocorrer será representada por $P(A|X)$. Sejam os eventos X e Y os ancestrais do evento A, então as probabilidades condicionais $P(A|X)$ e $P(A|Y)$ sozinhas não dão indício de como o impacto dos eventos X e Y interagem. Assim, será necessário considerar a especificação conjunta dessas variáveis na forma $P(A|X,Y)$ (JENSEN, 1996).

A probabilidade condicional é expressa por:

$$P(A|X) = \frac{P(A,X)}{P(X)} \quad (1)$$

Pela regra fundamental da probabilidade afirma-se:

$$P(A|X,Y) = \frac{P(A,X|Y)}{P(X|Y)} \quad (2)$$

O modelo Bayesiano de probabilidade interpreta cada proposição como um grau de crença do agente sobre a verdade da afirmação. Esse fato associa toda a probabilidade com o conhecimento do agente sobre a variável e em seu teorema permite, por exemplo, o cálculo de uma probabilidade desconhecida (*a posteriori*) a partir da coleta de evidências (*a priori*) (PIFER, 2006; CROCOMO, 2012).

A fórmula geral do teorema de Bayes é dada pela Equação 3:

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)} \quad (3)$$

Em que:

- $P(H|e)$ é a probabilidade *a posteriori* da hipótese *H* dada a evidência *e*.
- $P(e|H)$ é a probabilidade condicional (verossimilhança) da evidência *e* dada a hipótese *H*
- $P(H)$ é a probabilidade *a priori* da hipótese *H.*
- $P(e)$ é um fator de normalização.

Assumindo que o fator de normalização $P(e)$ é igual para todos os valores da hipótese *H*, a expressão a seguir demonstra a proporcionalidade entre as probabilidades *a posteriori* e *a priori.*

$$Posteriori \propto Verossimilhança \times Priori$$

Cada conexão de uma rede Bayesiana interliga duas variáveis e é direcional, indicando a influência de causalidade de uma variável na outra. Além dessa relação, as Tabelas de Probabilidade Condicional (TPC) também estão associadas às redes Bayesianas porque representam a medida de probabilidade das variáveis levando em consideração todas as configurações possíveis de seus ancestrais (DUDA et al., 2000).

Russel e Norvig (2004) demonstram que é possível responder qualquer consulta probabilística a partir da distribuição de probabilidade conjunta total e, analogamente, considerando que a estrutura de Rede Bayesiana é a representação dessa distribuição, então ela poderá ser usada para responder a essas mesmas consultas através de métodos de inferência probabilística. À vista disso, em uma Rede Bayesiana, a distribuição conjunta das probabilidades dos nós (A_i) condicionadas aos valores de todos os seus ancestrais (Pa_i), é expressa pela Equação 4.

$$P(A_1, \dots, A_n) = \prod_{i=1}^{n} P(A_i|Pa_i) \quad (4)$$

Uma restrição importante de representação para a estrutura da rede Bayesiana é que não pode haver um ciclo orientado, ou seja, uma realimentação dos nós que permitisse percorrer a rede, de nó a nó, ao longo das ligações, seguindo a direção das setas, e em

determinado momento retornasse para um nó já visitado. Um grafo que obedeça a tal restrição é classificado como Grafo Acíclico Orientado ou, simplesmente, um GAO. A Figura 1 demonstra essa diferença entre um grafo GAO (A) e um grafo não-GAO (B) (BISHOP, 2006).

(A) Grafo GAO **(A) Grafo não-GAO.**

Figura 1. Exemplos de grafos GAO e não-GAO

O exemplo apresentado a seguir foi construído com base em Crocomo (2012) e em Luna (2004), e visa melhor compreensão da Equação 3 e da representação de uma rede Bayesiana, associada à sua Tabela de Probabilidade Condicional.

Considere o uso de quatro variáveis binárias, representando o risco de diferentes domínios de segurança da informação de um sistema XYZ em produção no âmbito de um órgão do governo federal:

- $R09 = 1$. Indica que "existe um risco relativo à segurança de acesso".
- $R10 = 1$. Indica que "existe um risco relativo à segurança de física".
- $R11 = 1$. Indica que "existe um risco relativo à segurança de operações".
- $R = 1$. Indica que "existe um risco acumulado para o sistema XYZ".
- Quando qualquer das variáveis não representar risco ao sistema, o seu valor será 0.

A rede Bayesiana que representa as relações entre as variáveis R09, R10, R11 e R pode ser verificada na Figura 2.

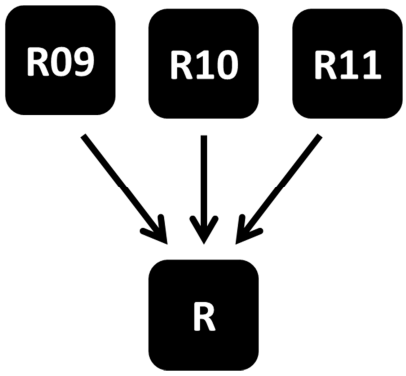


Figura 2. Rede Bayesiana que mostra a relação entre as variáveis.

A Tabela 4 reflete uma amostra de casos coerente com as bases de dados encontradas no mundo real, na qual algumas configurações de variáveis não ocorrem e, portanto, se apresenta como uma base de dados "incompleta".

A Tabela 5 demonstra uma tentativa de gerar a Tabela de Probabilidade Condicional da variável R, utilizando a amostra de casos da Tabela 4. Essa TPC corresponde ao cálculo da quantidade dos casos na situação apresentada, dividida pela quantidade total de casos na amostra. Assim, não é possível determinar a probabilidade das situações que não foram contempladas pela amostra.

ID	*R09*	*R10*	*R11*	*R*
01	0	0	1	0
02	0	1	0	0
03	0	1	0	0
04	0	1	0	0
05	1	0	0	0
06	1	0	1	1
07	1	1	0	1
08	1	1	1	1

Tabela 4. Amostra de casos incompleta.

Diante desse tipo de ocorrência, se faz necessário o uso de algum método para estimar os valores ausentes e com isso ser possível realizar as demais inferências desejadas. Um dos métodos clássicos para esta tarefa é o uso do algoritmo Esperança-Maximização, ou EM (*Expectation-Maximization*), para encontrar a estimativa por máxima verossimilhança (*MLE – Maximum Likelihood Estimation*), assumindo que a probabilidade de um valor que está faltando depende dos valores observados (*MAR – missing at random*) (DEMPSTER *et al.*; 1977).

ID	*R09*	*R10*	*R11*	*P(R=0)*	*P(R=1)*
01	0	0	0	?	?
02	0	0	1	1/8	?
03	0	1	0	3/8	?
04	0	1	1	?	?
05	1	0	0	1/8	?
06	1	0	1	?	1/8
07	1	1	0	?	1/8
08	1	1	1	?	1/8

Tabela 5. TPC da amostra de casos incompleta.

Segundo Rathie e Zörnig (2012), a estimativa por máxima verossimilhança é um método utilizado com objetivo de descrever um parâmetro qualquer de uma população. A partir de uma amostra observada $(x_1, \dots, x_n)$, com função de probabilidade $f(x|\theta_1, \dots, \theta_k)$, aquele valor do parâmetro θ que maximiza a função $L(\theta|x)$ em cada ponto amostral x, será o estimador de máxima verossimilhança. Em geral, esta é uma escolha razoável, porque corresponde ao valor do parâmetro θ para o qual a amostra observada é mais provável (CASELLA; BERGER, 2010).

Rusell e Norvig (2004) explicam que o algoritmo EM consiste em um método que utiliza os casos observados e a distribuição de probabilidade das variáveis para estimar os valores não observados (passo E) e, a partir da amostra "completada", realizar o cálculo da estimativa de máxima verossimilhança (passo M). Além disso, definem que esses dois passos do algoritmo são repetidos até que o valor da estimativa de máxima verossimilhança possa convergir para um ponto de máximo.

Uma variável X, do tipo multinomial discreta, admite estar em R possíveis estados, $x^1, x^2, \dots, x^r$, e tem sua distribuição de probabilidade dada com base nos parâmetros μ_r (Equação 5).

$$p(X|\mu) = \prod_{r=1}^{R} \mu_r^{X_r} \qquad (5)$$

.

Destaca-se o uso da distribuição de probabilidade *Dirichlet* como um conjugado *a priori* dos parâmetros μ_r da distribuição Multinomial em análises das variáveis das redes Bayesianas. A propriedade de conjugado *a priori* denota que as distribuições *a posteriori* e *a priori* dos parâmetros de interesse (neste caso μ_r) possuem a mesma classe de distribuição (BISHOP, 2006). Essa relação pode ser verificada na equação 6.

$$p(\mu|\alpha) \propto \prod_{r=1}^{R} \mu_r^{\alpha_{r-1}} \quad (6)$$

Em que:

- μ_r são os parâmetros da distribuição Multinomial.
- α_r são os parâmetros da distribuição *Dirichlet*.

Conforme Luna (2004), em uma rede Bayesiana, cada variável X_i tem r_i possíveis estados $x_i^1, x_i^2, \dots, x_i^r$, e possui probabilidade de estar no estado x_i^k expressa por $P\left(X_i = x^k \middle| Pa_i^j, S\right) = \theta_{ijk}$, dado que os seus pais Pa_i estão no *j*-ésimo estado. Exposto isto, o passo E para estimação dos parâmetros em uma rede Bayesiana consiste em determinar a frequência para cada caso possível da amostra utilizando a equação 7.

$$E_\theta\left(x_{ijk}|d^{obs}\right) = \begin{cases} 1, & se\ x_i\ e\ pa_i\ são\ observados\ e\ X_i = k\ e\ Pa_i = j \\ 0, & se\ x_i\ e\ pa_i\ são\ observados\ e\ X_i \neq k\ ou\ Pa_i \neq j \\ P(X_i = k, Pa_i = j\ |d^{obs}, \theta, S), & \text{em caso contrário} \end{cases} \quad (7)$$

Para determinar o valor da probabilidade $P(X_i = k, Pa_i = j\ |d^{obs}, \theta, S)$ a rede Bayesiana com estrutura S e conjunto de parâmetros Θ é instanciada com a evidência observada no caso d_l^{obs} e, por inferência probabilística, determina-se qual a distribuição de probabilidade.

Assim, após estimar os valores esperados de cada caso na etapa anterior, executa-se em cada caso o passo M pela Equação 8 com o objetivo calcular o valor das probabilidades (θ_{ijk}).

$$\theta_{ijk} = \frac{\sum_{l=1}^{m} E_\theta\left(x_{ijk}^l|d_l^{obs}\right)}{E\left(N_{ij}|d^{obs}\right)} \quad (8)$$

Em que:

- m é o total de casos.
- N_{ij} corresponde ao número de ocorrências de Pa_i para todos os casos encontrados da variável X_i em seus *k*-ésimos estados $\left(N_{ijk}\right)$. Por isso, pode ser obtido por $N_{ij} = \sum_k N_{ijk}$

Assim, admitindo que as variáveis possuam distribuição Dirichlet[1,1], ou seja, probabilidade de 0,5 para cada estado possível (tanto *a priori* quanto *a posteriori*) e aplicando o método EM de estimação para a amostra de casos da tabela 4, obtemos a Tabela 6, que corresponde à frequência esperada dos casos, ou seja, o passo E..

ID	*R09*	*R10*	*R11*	*R=0*	*R=1*
01	0	0	1	0,5000	0,5000
02	0	1	0	2,0000	0,3333
03	0	1	0	4,0000	0,2000
04	0	1	0	0,5000	0,5000
05	1	0	0	2,0000	0,3333
06	1	0	1	0,3333	2,0000
07	1	1	0	0,3333	2,0000
08	1	1	1	0,3333	2,0000

Tabela 6. Passo E – Cálculo da frequência esperada.

Neste exemplo, foram necessárias 7 execuções dos dois passos do algoritmo EM para que os valores de probabilidades encontrados pudessem se estabilizar. Com isso, determinou-se a Tabela 7, abrangendo as probabilidades condicionais estimadas para a rede Bayesiana do exemplo.

ID	*R09*	*R10*	*R11*	*P(R=0)*	*P(R=1)*
01	0	0	0	0,5000	0,5000
02	0	0	1	0,9961	0,0039
03	0	1	0	1,0000	0,0000
04	0	1	1	0,5000	0,5000
05	1	0	0	0,9961	0,0039
06	1	0	1	0,0039	0,9961
07	1	1	0	0,0039	0,9961
08	1	1	1	0,0039	0,9961

Tabela 7. Tabela de Probabilidade Condicional estimada.

No exemplo anterior, a Rede Bayesiana utilizada para estimativa das probabilidades e para extrair as relações de dependência entre as variáveis foi fornecida previamente. Entretanto, a construção da topologia da rede Bayesiana pode não ocorrer diretamente pela definição de um especialista e sim acontecer utilizando técnicas de aprendizado da rede, nas quais de uma forma semi-automatizada, o conhecimento acerca das variáveis envolvidas é absorvido para construir ou modificar o modelo da rede (COOPER; HERSKOVITS, 1992; JENSEN, 1996).

A construção deste tipo de rede pode ser considerada como sendo um caso específico do problema geral de encontrar um modelo probabilístico que represente um

conjunto de dados (DALY *et al.*,, 2011). Neste sentido, as disciplinas de aprendizado de máquina e de estatística compartilham o objetivo comum de modelar fenômenos do mundo real. Conforme Jensen (1996), o aprendizado de redes Bayesianas pode ocorrer em duas circunstâncias distintas – a partir de uma base de dados de casos (*batch learning*) ou pela modificação de um modelo a partir da obtenção de novos casos (*adaptation*) – e é composto por duas tarefas centrais que se diferem pela abordagem – uma do tipo qualitativa (especifica a estrutura da rede) e a outra do tipo quantitativa (especifica as probabilidades condicionais).

Fundamentado nos conceitos estatísticos e na teoria de grafos, o aprendizado de redes Bayesianas foi demonstrado como sendo um problema NP-completo (CHICKERING, 1996). Desta forma, por ser inviável utilizar métodos exatos nesta tarefa, técnicas heurísticas para encontrar a rede que melhor representa uma determinada base de dados têm sido desenvolvidas de modo a obter boas soluções em tempo computacional aceitável. Neste sentido, dentre as diferentes abordagens, destacam-se os dois modelos mais utilizados: um baseado em busca e pontuação e o outro baseado em análise de dependência (DALY *et al.*, 2011; HECKERMAN *et al.*, 1995; SANTOS, 2007).

Nas abordagens por análise de dependência, o conceito de independência condicional das variáveis é utilizado como fundamento para inferir a topologia da rede Bayesiana. Nelas, o relacionamentos entre as variáveis são identificados através da exploração do critério de d-separação (JENSEN, 1996) e da execução de testes estatísticos, como o Qui-quadrado (DALY *et al.*, 2011; PIFER, 2006).

Nas abordagens por busca e pontuação, a heurística de busca recebe como entrada um conjunto de dados de treinamento e gera diferentes redes candidatas, as quais são submetidas à avaliação por uma medida padronizada e, com isso, se consegue organizar as redes descobertas em ordem de relevância. Elegendo, então, aquela que obtiver maior probabilidade de representar o conjunto de dados lidos e a informação *a priori* conhecida, transformando a aprendizagem estrutural da rede em um problema de maximização de uma função. As métricas de pontuação podem ser agrupadas em dois diferentes tipos: medidas Bayesianas (fornecem uma medida de qualidade a partir da informação *a priori* e do cálculo da probabilidade *a posteriori* das variáveis) e medidas de informação (fornecem uma medida de qualidade sem nenhuma informação sobre o modelo de distribuição *a priori*) (PIFER, 2006).

Neste aspecto, o trabalho considerado como inspirador sobre a construção de redes Bayesianas a partir de um conjunto de dados é o de Cooper e Herskovits (1992), neste trabalho é apresentado o sistema K2, no qual a métrica de pontuação das redes também ficou conhecida por este nome. Nota-se ainda que em trabalhos posteriores sobre o tema a métrica K2 é utilizada como referência de comparação (DALY *et al.*, 2011). Assim, a Equação 9 apresenta a função incorporada no algoritmo K2 de Cooper e Herskovits (1992), a qual mede a probabilidade de uma estrutura de rede Bayesiana (B_S) modelar determinado conjunto de dados (D).

$$P(B_\mathrm{S}, \mathrm{D}) = P(B_\mathrm{S}) \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \qquad (9)$$

Em que:

- $P(B_\mathrm{S})$ é um fator utilizado quando se tem uma informação prévia sobre a qualidade da rede.
- n é o número de variáveis x_i.
- q_i é o número de possíveis combinações, encontradas em D, dos valores das variáveis pais de x_i.
- r_i é o número de possíveis valores associados à variável x_i.
- N_{ij} é o número de ocorrências que possuem $\pi_i = w_{ij}$. Dado por:
    - $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$
- N_{ijk} é o número de casos em D, com $x_i = v_{ik}$ e $\pi_i = w_{ij}$, em que:
    - v_{ik} representa o k-ésimo valor possível (entre os r_i valores existentes) da variável x_i
    - w_{ij} representa a j-ésima combinação possível (entre as q_i combinações existentes em D) dos valores das variáveis π_i (pais de x_i).

É possível segmentar da pontuação total, $P(B_\mathrm{S}, \mathrm{D})$, uma função de ganho, $g(i, \pi_i)$, que modela a contribuição de cada nó i e seus pais π_i na pontuação geral da rede. Assim, a Equação 10 representa essa nova forma geral e a Equação 11 representa a função de ganho:

$$P(B_\mathrm{S}, \mathrm{D}) = P(B_\mathrm{S}) \prod_{i=1}^{n} g(i, \pi_i) \qquad (10)$$

$$g(i,\pi_i) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}! \qquad (11)$$

Diferentes algoritmos e estratégias têm sido utilizados na etapa de busca da rede, de modo que exploram o espaço de busca para chegar às redes com maiores pontuações. Pela efetividade percebida em pesquisas e estudos de caso, é possível destacar os algoritmos por busca gulosa, algoritmos genéticos e os algoritmos de Otimização por Colônia de Formigas.

Os algoritmos por busca gulosa, como o K2, foram os primeiros a obterem destaque no aprendizado de redes Bayesianas a partir de um conjunto de dados e são referências (1) de uso, pelo tempo que já se encontram disponíveis, e (2) de comparação quando novas técnicas são desenvolvidas. Eles constroem a sua solução alterando o grafo passo-a-passo, utilizando a melhor decisão em cada passo, de modo a produzir estruturas com maior pontuação em relação ao estado anterior. O grafo inicial candidato, pode ser vazio, aleatório ou sugerido (DALY *et al.*; 2011; HECKERMAN *et al.*; 1995).

Os algoritmos genéticos se baseiam em mecanismos da evolução natural para desenvolver as suas soluções. Através de critérios de seleção ao longo de diferentes gerações, as estruturas mais fortes vão sendo determinadas pela ocorrência de mutações e combinações da população de indivíduos. Na procura pelas estruturas de redes Bayesianas este tipo de algoritmo tem obtido bons resultados e com diferentes abordagens, como na redução da complexidade da rede ou no uso de sistemas híbridos (DALY *et al.*, 2011; SANTOS, 2007).

A Otimização por Colônia de Formigas também surgiu como uma alternativa ao procedimento clássico de busca de estruturas de redes Bayesianas. Essa técnica é inspirada no comportamento das colônias de formigas do mundo real, no qual elas são capazes de encontrar o menor caminho entre fontes de comida e o formigueiro. No trabalho apresentado por Campos *et al.* (2002), cada formiga constrói a sua solução para o problema e a compartilha indiretamente com as demais através da deposição de feromônio.

2.4. OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS

Conforme Castro (2006), a Computação Natural é um ramo da ciência que se destina a extrair ideias da natureza e utilizar tal inspiração para desenvolver ferramentas computacionais de resolução de problemas ou, ainda, para construir meios novos de se realizar computação. Este campo de pesquisa possui três principais enfoques: (1) a computação inspirada na natureza, na qual são desenvolvidas ferramentas (algoritmos) inspiradas no comportamento natural para solução de problemas complexos; (2) simulação de fenômenos naturais, sinteticamente trabalha com aspectos da construção de vida artificial; e (3) a computação utilizando recursos naturais, que corresponde ao uso da natureza como meio para realizar computação.

Dos campos citados, o mais antigo e consolidado é o primeiro, o qual compreende todas as soluções computacionais obtidas com base em estudos sobre o comportamento da natureza e que pode ser dividido em quatro técnicas de algoritmos: (1) computação evolucionária, que utiliza, fundamentalmente, os aspectos da teoria de evolução biológica como, por exemplo, a reprodução com herança e a seleção natural; (2) redes neurais artificiais, que consistem em sistemas de processamento de informação inspirados no comportamento do cérebro; (3) inteligência de enxames, que são técnicas de resolução de problemas utilizando agentes individuais para reproduzir um comportamento coletivo; e (4) computação imunológica, que emprega os princípios de sistemas adaptativos inspirados na teoria de imunologia.

Em termos gerais, os sistemas de computação baseados em inteligência de enxame se referem a uma população de indivíduos com habilidades limitadas, mas que interagem entre si e com o meio ambiente para exibir um comportamento coletivamente inteligente. Neste aspecto, um dos melhores exemplos é o comportamento das formigas, que são capazes de explorar uma fonte de comida sem perder a capacidade de descobrir outras características do meio ambiente, como novas fontes de comida ou caminhos menores para as mesmas. Assim, observações sobre esse equilíbrio entre "Convergência (*exploitation*) X Descoberta (*exploration*)" levaram ao desenvolvimento de algoritmos computacionais para solução de problemas complexos (CASTRO, 2006).

A técnica que desenvolve este tipo de algoritmo é a meta-heurística de Otimização por Colônia de Formigas (*ACO – Ant Colony Optimization*), a qual foi proposta por M.

Dorigo para solução de problemas de otimização combinatória (DORIGO; BLUM, 2005) e posteriormente tornou-se mais conhecida após a prova de conceito apresentada para solução do problema do Caixeiro Viajante (*TSP -Travelling Salesman Problem*). Essa técnica aplica o uso de uma trilha artificial de feromônios como meio de comunicação entre as formigas, de forma a refletir a experiência adquirida ao longo do tempo (DORIGO; STÜTZLE, 2010).

O problema do caixeiro viajante consiste em tentar determinar a menor rota para percorrer uma série de cidades e retornar à cidade de origem, visitando uma única vez cada uma delas (CASTRO, 2006). Para solução deste problema foi utilizada a seguinte sequência:

- Cada formiga inicia o seu percurso em uma cidade (nó do grafo).
- Cada formiga escolhe a próxima cidade pela lista de cidades possíveis e levando em consideração, probabilisticamente, a força do feromônio entre as duas cidades e a função que representa a distância entre as duas cidades (representando a relação entre convergência x descoberta).
- Depois que cada formiga constrói a sua solução, o feromônio é atualizado (*feedback* positivo).
- São realizadas várias rodadas de descoberta (iterações).
- Ao final, as conexões com menores distâncias possuem mais feromônio e, por isso, representam a solução coletiva para o problema.

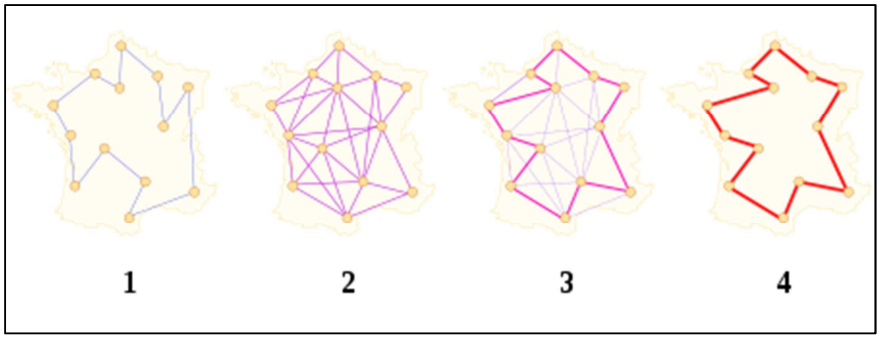


Figura 3. Construção da solução para o problema TSP por algoritmo ACO (Wikipédia).

A Figura 3 representa os estados de construção da solução do problema do caixeiro viajante, sob a ótica do efeito do algoritmo ACO ao longo das iterações, desde o estado inicial antes da execução da técnica até o estado final, quando se destaca o menor caminho pela predominância do feromônio.

Após o seu uso para solução do problema do caixeiro viajante, a meta-heurística ACO foi definida como uma metodologia para uso em diversas aplicações e, entre elas, o aprendizado de estruturas de redes Bayesianas (CAMPOS et al., 2002). Nesta proposta, cada formiga constrói a sua possível solução para o problema através de movimentos em uma sequencia finita de nós vizinhos, selecionados por uma busca estocástica influenciada pela informação local sobre o problema, pela informação compartilhada do feromônio e pelo estado interno da formiga que, diferentemente do problema do Caixeiro Viajante, quando a distância entre as cidades é fixa, no caso das redes Bayesianas a informação heurística depende da solução parcial aplicada e por isso é atualizada a cada conexão escolhida (CAMPOS *et al*., 2002).

Uma formiga k, a partir de um nó i, seleciona o próximo nó j escolhendo primeiramente qual o tipo de comportamento irá aplicar: *Exploitation* (quando dá maior importância para os conhecimentos local e coletivo sobre o problema e converge para a solução coletiva) ou *Exploration* (quando realiza uma escolha probabilística e favorece a descoberta de novos caminhos). A escolha desse comportamento ocorre através da amostra aleatória do parâmetro q no intervalo [0,1] e na sua comparação com o parâmetro q_0, que é um limiar arbitrariamente definido no início do algoritmo $(0 \leq q_0 < 1)$. Em seguida, após a escolha do comportamento a ser aplicado, o peso da informação heurística e o peso da matriz de feromônios na escolha da próxima conexão são influenciados através da definição inicial dos parâmetros α e β no algoritmo.

Assim, a escolha da próxima conexão *i,j* ocorre de acordo com a Equação 12. Caso o parâmetro q amostrado antes da escolha da conexão seja menor ou igual ao q_0, é aplicada uma busca local. Entretanto, caso o parâmetro q seja maior do que q_0, a próxima conexão será amostrada utilizando as probabilidades calculadas conforme a Equação 13.

$$i,j = \begin{cases} \arg\max_{u \in J_k(i)} \{[\tau_{iu}]^{\alpha}[\eta_{iu}]^{\beta}\} & , \quad se\ q \leq q_0 \\ p_k(i,j) & , \quad se\ q > q_0 \end{cases} \qquad (12)$$

$$p_k(i,j) = \begin{cases} \dfrac{[\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{u \in J_k(i)}[\tau_{iu}]^{\alpha}[\eta_{iu}]^{\beta}} & , \quad se\ j \in J_k(i) \\ 0 & , \quad senão \end{cases} \qquad (13)$$

Em que:

- η_{ij} é a informação heurística calculada para cada nó candidato utilizando a função de ganho K2 (definida anteriormente na Equação 11).

- $J_k(i)$ é o conjunto de nós candidatos para a formiga k.
- α, β são dois parâmetros que controlam a importância entre o feromônio e a informação heurística.
- Elemento i, j da matriz τ contém o nível de feromônio depositado no arco do nó i para o nó j.

Uma colônia de m formigas procura uma solução para o problema durante a quantidade de rodadas denominadas iterações. No caso da solução apresentada por Campos *et al.* (2002) para o problema de aprendizado de Redes Bayesianas, cada formiga da colônia constrói a sua solução e as diferentes soluções individuais são classificadas segundo a métrica K2 e aquela com maior pontuação é armazenada para comparação com as melhores soluções das demais iterações. A rede de maior pontuação dentre todas as iterações é a solução final.

A matriz de feromônios, que é o ponto fundamental da meta-heurística ACO, é atualizada em dois momentos distintos: o primeiro (atualização local) é realizado por cada formiga, sempre que transita de um nó para outro, com o objetivo de representar a evaporação do feromônio e assim favorecer a descoberta de novos caminhos. O segundo momento (atualização global) ocorre quando se encontra uma rede de maior pontuação ao final da iteração, nesta ocasião as conexões que fazem parte desta melhor solução encontrada são reforçados com o valor de feromônio inicial (feedback positivo) (CAMPOS et al., 2002; FENG *et al.*, 2014).

As Equações 14 e 15 representam, respectivamente, a regra de atualização local e a regra de atualização global do feromônio conforme o trabalho de Campos *et al.* (2002).

$$\tau_{ij} \leftarrow (1 - \psi)\tau_{ij} + \psi\tau_0 \qquad (14)$$

Em que:

- τ_0 é o nível inicial de feromônio $(0 \leq \tau_0 < 1)$.
- ψ é o parâmetro que controla a evaporação do feromônio $(0 \leq \psi < 1)$.

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij} + \rho\Delta\tau_{ij} \qquad (15)$$

Em que:

- ρ é o parâmetro que controla a evaporação do feromônio $(0 \leq \rho < 1)$.

- $\Delta\tau_{ij} = \begin{cases} \frac{1}{C(S^+)} & , se\ \{ij\} \in S^+ \\ \tau_{ij} & , se\ \{ij\} \notin S^+ \end{cases}$
- S^+ é a melhor solução encontrada até o momento.
- $C(S^+)$ é o custo associado à melhor solução.

Assim, Campos *et al*. (2002) definiram a forma geral do Algoritmo ACO-B, que é responsável por aplicar a técnica de Otimização de Colônia de Formigas para aprendizado de Redes Bayesianas. Este algoritmo incorpora a técnica apresentada até aqui para seleção da rede Bayesiana de maior pontuação. E, conforme é possível verificar mais adiante, esse algoritmo também inclui uma função denominada OTIMIZACAO(), na qual são realizadas alterações por um algoritmo de busca local na solução de cada formiga (exclusão ou reversão de um arco). O uso dessas alterações a cada estrutura de rede encontrada ou ao final na iteração pode permitir escapar de um eventual local ótimo encontrado. Os autores destacam que essa função pode reaproveitar os cálculos realizados nas etapas anteriores do algoritmo ao utilizar uma métrica decomposta.

O Algoritmo ACO-B faz referência ao algoritmo FORMIGA-B, que é responsável pela definição do grafo GAO executada por cada formiga. Os passos do algoritmo FORMIGA-B permitem que cada formiga desenvolva a sua solução e eles estão relacionados com a escolha dos arcos (baseada nos cálculos das Equações 11, 12 e 13), com a atualização das restrições do grafo GAO a cada escolha de um arco e com a atualização local do feromônio (segundo a Equação 14).

```
(01) ACO-B()
     Entradas: ITERLOCAL → O número de iterações da busca local
               ITERMAX → O número máximo de iterações total
               M → Número de formigas
(02) Obter uma rede inicial pela métrica K2
(03) Calcular o feromônio inicial
(04) Definir a matriz de feromônios com o valor inicial
(05) para ITER=1 até ITERMAX faça
(06)     para K=1 até M faça
(07)         GK=FORMIGA-B()
(08)         se ( ITER mod ITERLOCAL = 0) ; então
(09)             GK=OTIMIZACAO(GK)
(10)         fim-se
(11)     fim-para
(12)     #seleciona a solução (GB) com maior pontuação entre todas as formigas (GK)
         GB=arg max(GK)
(13)     se K2(GB) >  K2(G+); então
(14)         G+ = GB;
(15)     fim-se
(16)     #atualiza o feromônio global utilizando a Equação 15
         τ_ij
(17)     para K=1 até M faça
(18)         GK0 = OTIMIZACAO(GK)
(19)     fim-para
(20)     #seleciona a solução (GB) com maior pontuação entre todas as formigas (GK0)
         GB=arg max(GK0)
(21)     se K2(GB) >  K2(G+); então
(22)         G+ = GB;
(23)     fim-se
(24)     retorna G+
```

Algoritmo ACO-B (CAMPOS et al., 2002)

```
(01) FORMIGA-B()
(02) para i=1 até N
(03)         #Limpa a informação dos nós pais de i
             Pa(i) = 0
(04) fim-para
(05) para i=1 até N
(06)         para j=1 até N
(07)                 se i≠j; então
(08)                         # Calcula a informação heurística pela Equação 11
                             N(i,j)=K2(i,j)
(09)                 fim-se
(10)         fim-para
(11) fim-para
(12) repita
(13)         # Utiliza a Equação 12 para escolher os nós
             SELECIONA(i,j)
(14)         se N(i,j) > 0; então
(15)                 Pa(i) = Pa(i) ∪ {j}  # Inclui o nó j no vetor de pais do nó i
(16)         fim-se
(17)         # Limpa a informação heurística do par (i,j)
             N(i,j) = -∞
(18)         # Limpa a informação heurística de todos possíveis pais do nó j,
             # que também sejam filhos do nó i. (evita criar um ciclo no grafo)
             ∀ a ∈ ANT(j) ∪ {j} e b ∈ DES(i) ∪ {i}, N(a,b) = -∞
(19)         para k=1 até N
(20)                 se N(i,k) > -∞; então
(21)                         #Recalcula a informação heurística pela Equação 11
                             N(i,k)=K2(i,k)
(22)                 fim-se
(23)         fim-para
(24)         #Atualiza o feromônio localmente, conforme a Equação 14
             τ(i,j)
(25) enquanto houver algum N(i,j) > 0
```

Algoritmo Formiga-B (CAMPOS et al., 2002)

2.1. SÍNTESE DO CAPÍTULO

Com este capítulo foi possível compreender os fundamentos de diferentes áreas utilizadas na construção do modelo de análise de riscos aplicado neste trabalho. Dentre tais

conceitos é importante destacar (1) a necessidade de monitoramento contínuo dos riscos frente às atuais ameaças; (2) a representação semântica das redes Bayesianas em problemas de tomada de decisão; (3) a técnica de busca dessas estruturas por Colônia de Formigas a partir de dados de treinamento; (4) as métricas de pontuação para avaliação das redes encontradas; (5) o cálculo das probabilidades condicionais após a definição da estrutura de maior pontuação, segundo a métrica utilizada. Nos próximos capítulos passa-se a discutir o modelo de análise de riscos utilizado e as questões da sua implementação.

3. CONSTRUÇÃO DA REDE BAYESIANA PARA ANÁLISE DE RISCOS DE SEGURANÇA

Este capítulo apresenta o método para construir uma rede Bayesiana que represente os fatores de riscos de segurança de um sistema de informação. Ele é baseado no modelo "*A security risk analysis model for information systems: Causal relationships of risk factors and vulnerability propagation analysis*" – um trabalho publicado por N. Feng, H. Wang e M. Li (2014). Ao longo do capítulo são detalhadas as etapas que compõem esta técnica, o fluxo de informações entre os componentes, e como os autores aplicaram a técnica de aprendizado de redes Bayesianas pela meta-heurística de colônia de formigas para apresentar uma solução na área de gestão de riscos de segurança da informação.

3.1. MOTIVAÇÃO

A principal razão para pesquisas sobre o tratamento dos riscos de segurança em sistemas de informação está relacionada com a importância destes mecanismos para as organizações, à medida que representam um domínio cada vez maior nos negócios e governos (MPOG, 2016), e também devido à complexidade dos desafios enfrentados por aqueles que trabalham para evitar violações neste tipo de estrutura, pois quebras de segurança tem representado significantes custos àqueles responsáveis por armazenar e tratar tais dados (CYBERSECURITY VENTURES, 2016).

A proposta de Feng *et al.* (2014) foi escolhida para ser aplicada neste estudo de caso principalmente pelos seus aspectos de tratamento proativo do risco e seu alinhamento com as necessidades de visão compartilhada entre governança, gestão de riscos e conformidade. Além disso, outras características importantes percebidas são a consideração conjunta da opinião de especialistas com dados históricos e a possibilidade de produzir uma medida conjunta de risco envolvendo diferentes *hardwares* e *softwares*, mas orientada por sistema de informação ou por serviço de tecnologia prestado.

Percebe-se que neste trabalho de referência foi realizado o levantamento das principais técnicas para tratamento de riscos de segurança, com diferentes enfoques – modelos matemáticos e estatísticos, valores dos ativos, estimativas dos riscos, lógica difusa – e interessantes contribuições para o desenvolvimento desta área como, por exemplo, o trabalho que apresentou a técnica de uso das redes Bayesianas como um processo contínuo de monitoramento dos riscos para suporte à tomada de decisão (FAN; YU; 2004). A partir de tais

estudos, identificou-se a oportunidade de apresentar uma nova técnica para o tratamento de riscos de segurança da informação a partir de duas questões principais.

A primeira está relacionada com o fato do processo de análise de risco clássico ser construído com base em modelos definidos por especialistas ou escolhidos a partir de estruturas-padrão conhecidas. Assim, foi proposta a construção de um modelo de avaliação de riscos de segurança com base tanto no conhecimento de especialistas como no histórico de casos conhecidos para o problema.

A segunda questão apresentada foi que as abordagens existentes são focadas na mensuração do risco considerando a sua probabilidade e severidade. Entretanto, revelou-se que a determinação de uma vulnerabilidade poderia propagar e escalar os riscos de segurança através da cadeia de fatores de riscos, dados os aspectos de causalidade entre eles, e desta forma gerar novos riscos de segurança aos sistemas de informação. Assim, um novo enfoque trazendo a análise de propagação das vulnerabilidades contribuiria para o tratamento proativo dos riscos de segurança da informação.

Assim, Feng *et al.* (2014) propõem um novo Modelo de Análise de Riscos de Segurança, baseado em redes Bayesianas e Colônia de Formigas, que ampliou os trabalhos existentes ao: (1) construir uma rede Bayesiana de fatores de risco a partir do conhecimento histórico e de especialistas, (2) utilizar os benefícios deste tipo de estrutura na tomada de decisões, realizando a avaliação de riscos sobre a estrutura encontrada a partir do recebimento de novas evidências, (3) desenvolver a análise de propagação dos riscos encontrados com objetivo de guiar os planos de tratamento dos riscos. A Figura 4 representa uma visão geral do modelo.

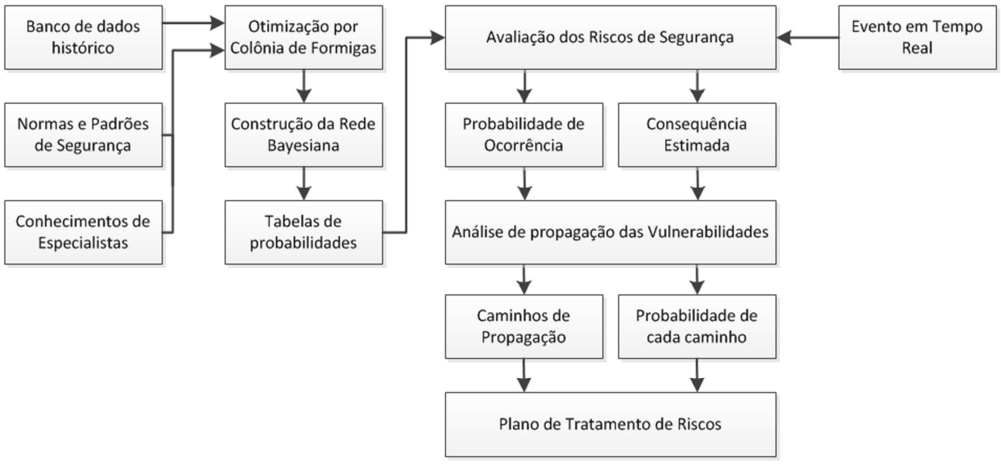


Figura 4. Modelo SRAM (adaptada de (FENG et al., 2014))

3.2. DEFINIÇÃO DA REDE BAYESIANA

No modelo proposto (SRAM), cada nó da rede Bayesiana representa um controle de segurança escolhido de acordo com a publicação *NIST Special Publication 800-53: Security and Privacy Controls for Federal Information Systems and Organizations* (NIST, 2013). O NIST foi escolhido como base para seleção dos controles de segurança por possuir uma das arquiteturas de referência na implementação da gestão de riscos de segurança. Os controles de segurança descritos por ele são divididos em seis diferentes categorias (segurança física, segurança de rede, segurança do ativo, segurança de aplicação, segurança de dados, segurança de operações). O modelo de Feng *et al.* (2014) sugere a determinação de uma rede Bayesiana diferente para cada área.

O estabelecimento da rede Bayesiana que melhor representa os aspectos de dependência entre os controles de segurança a partir da informação histórica e da contribuição de especialistas sobre tais variáveis constitui a etapa inicial deste modelo de referência. O que representa então um problema prático do aprendizado da estrutura de redes Bayesianas a partir de um conjunto de dados e, dessa forma, os autores utilizaram para solução deste problema a técnica de aprendizado de redes Bayesianas pela meta-heurística de Otimização por Colônia de Formigas (CAMPOS et al., 2002). Nesta solução, o algoritmo de aprendizado da rede recebe como entrada as especificações e o histórico de casos sobre os nós da rede e determina em sua saída uma estrutura de rede Bayesiana que melhor descreve os dados recebidos.

Nesta etapa são criados dois bancos de dados (BD1 e BD2). O primeiro contendo as especificações dos nós da rede de acordo com o conhecimento de especialistas e com a norma SP 800-53 e o segundo contemplando o histórico de casos para cada controle de segurança. Apesar do modelo não fornecer um exemplo desses bancos de dados, é relatado o uso de arquivos em formato MATLAB para armazenar os dados não tratados e a existência de uma etapa de pré-processamento, na qual os dados coletados são classificados em intervalos definidos de modo a controlar a quantidade de estados possíveis para cada variável.

O modelo de Feng *et al.* (2014) fornece uma versão do algoritmo de Otimização por Colônia de Formigas (Algoritmo 1) para representar os aspectos de causalidade entre os fatores de riscos na forma da rede Bayesiana. Assim, cada formiga constrói a sua solução para o problema adicionando um nó por vez, com base na escolha aleatória entre as duas abordagens – a primeira realiza um busca local pelo melhor movimento e a segunda é uma

seleção probabilística – nas quais ambas as opções selecionam o próximo nó considerando a informação do feromônio (τ_{ij}) e a informação heurística calculada pela métrica K2 (η_{ij}).

```
(01)  ALGORITMO-1()
(02)    Calcular o feromônio inicial e definir a matriz de feromônios com este valor
(03)    para ITER=1 até ITERMAX; faça
(04)      para K=1 até M faça
(05)          para i=1 até N faça
(06)                #Limpa a informação dos nós pais de i
                    Pa(i) = 0
(07)          fim-para
(08)          para i=1 e j =1 até N; faça
(09)                se i≠j; então
(10)                      # Calcula a informação heurística pela Equação 11
                          N(i,j) =K2(i,j)
(11)                fim-se
(12)          fim-para
(13)          Repita
(14)                # Utiliza a Equação 12 para escolher os nós (i,j)
                    SELECIONA(i,j)
(15)                se N(i,j) > 0; então
(16)                      # Inclui o nó j no vetor de pais do nó i
                          Pa(i) = Pa(i) ∪ {j}
(17)                fim-se
(18)                # Limpa a informação heurística do par (i,j)
                    N(i,j) = -∞
(19)                # Limpa a informação heurística de todos possíveis pais do nó j, que
                    também sejam filhos do nó i. (evita criar um ciclo no grafo)
                    ∀ a ∈ ANT(j) ∪ {j} e b ∈ DES(i) ∪ {i}, N(a,b) = -∞
(20)                para k=1 até N
(21)                      se N(i,k) > -∞; então
(22)                            #Recalcula a informação heurística pela Equação 11
                                N(i,k)=K2(i,k)
(23)                      fim-se
(24)                fim-para
(25)                #Atualiza o feromônio localmente, conforme a Equação 14
                    τ(i,j)
(26)          enquanto houver algum N(i,j) > 0
(27)       fim-para
(28)       #seleciona a solução (GB) com maior pontuação entre todas as formigas (GK)
           GB=arg max(GK)
(29)       se K2(GB) >  K2(G+); então
(30)           G+ = GB;
(31)       fim-se
(32)       #Atualiza o feromônio global, conforme a Equação 15
           τ(i,j)
(33)    fim-para
(34)    retorna G+
```

Algoritmo 1

São definidos arbitrariamente os seguintes parâmetros: quantidade de rodadas (N_{max}), quantidade de formigas em cada rodada (m), importância do feromônio (α), importância da informação heurística (β), peso entre busca local ou seleção probabilística (q_0), taxa de evaporação do feromônio (ρ). Conforme Feng *et al.* (2014), deve ser escolhido um número suficiente de formigas e de iterações de maneira que o feromônio nas conexões diferentes da melhor solução encontrada tenha "evaporado" ao final das rodadas. Neste tipo de situação, as conexões que fazem parte da solução de maior pontuação devem possuir maior valor de feromônio, enquanto que deve restar pouco feromônio nas demais conexões.

Cada formiga da colônia constrói a sua solução para o problema (estrutura de rede Bayesiana), a qual é avaliada segundo a métrica K2, aquela estrutura de maior pontuação dentro da mesma rodada é armazenada para comparação com as soluções das demais rodadas. Durante a construção da sua solução, cada formiga atualiza localmente o feromônio (evaporação) quando decide incluir uma conexão e, quando for encontrada uma estrutura de maior pontuação ao final de cada rodada, atualiza-se globalmente o feromônio (*feedback* positivo) através do reforço de todas as conexões participantes da solução encontrada. A solução final será aquela com maior pontuação entre todas as rodadas.

Diante do aprendizado da estrutura da rede Bayesiana, o próximo passo para a sua definição completa é estimar os parâmetros em cada nó. No caso do modelo SRAM são estimadas as Tabelas de Probabilidade Condicional segundo o método por máxima verossimilhança.

3.3. EMPREGO DA REDE BAYESIANA

A partir da definição completa da rede (estrutura e parâmetros) é possível extrair informações de gestão sobre o ambiente analisado: a estrutura de rede Bayesiana encontrada revela os aspectos de influência entre os fatores de risco e, através do processo de inferência em redes Bayesianas permite-se atualizar as probabilidades das variáveis em função das evidências percebidas.

No caso do modelo SRAM, esta é a segunda etapa e consiste em avaliar os riscos de segurança da informação a partir de novos dados. Assim, ao observar evidências do monitoramento dos controles de segurança em um terceiro banco de dados (BD3), é realizada a estimativa do risco através do cálculo da sua probabilidade posterior. Desta forma, os dados de entrada para a etapa da avaliação dos riscos são a rede Bayesiana (estrutura e parâmetros) e

as novas evidências. E, após aplicar o processo de inferência na rede, a saída desta etapa será uma medida de probabilidade estimada de ocorrência do risco.

O modelo SRAM ainda vai além da estimativa do risco porque prevê uma terceira etapa de funcionamento, que consiste em um passo adicional executado apenas quando o risco estimado na etapa anterior ultrapassa o limite definido previamente. Nesta situação, busca-se determinar através da análise de propagação dos riscos qual o caminho entre um nó de origem e um nó de destino apresenta o maior risco estimado. Esta terceira etapa também se baseia em um algoritmo de Colônia de Formigas.

Considerando que este trabalho faz um estudo de caso somente da primeira etapa do modelo SRAM, i.e. a descoberta da estrutura da Rede Bayesiana e o cálculo das tabelas de probabilidade condicional, ressalta-se que a etapa de avaliação dos riscos de segurança da informação, a qual é fundamentada na teoria de inferência probabilística, pode ser executada por uma ferramenta externa como é o caso do *software* UnBBayes (MATSUMOTO, 2011) ou outras soluções comerciais.

3.4. SÍNTESE DO CAPÍTULO

Como foi possível verificar neste capítulo, o aprendizado da rede Bayesiana é o ponto de partida para a aplicação do modelo de análise de riscos de segurança proposto por Feng *et al*. (2014). O uso dos conhecimentos de especialistas e dos dados históricos procura fornecer a oportunidade para que a estimativa dos riscos seja mais próxima da realidade de cada sistema analisado. Além disso, foi possível verificar como as demais etapas do modelo analisado se interligam. Em seguida, serão apresentados os resultados obtidos por esse estudo de caso na aplicação da primeira etapa do modelo.

4. ESTUDO DE CASO

Este capítulo apresenta um estudo de caso para definição da rede Bayesiana que modela os fatores de risco de um sistema de informação em produção na Polícia Federal. Estas ações são fundamentadas na aplicação da primeira etapa do Modelo de Análise de Riscos de Segurança para sistemas de informação proposto por Feng *et al*. (2014) e apresentado anteriormente. Através da coleta de dados acerca dos controles de segurança desse sistema são discutidos os aspectos de implementação e os resultados obtidos a partir desta técnica de tratamento proativo dos riscos de segurança da informação a fim de contribuir com a evolução do modelo e identificar a sua validade como ferramenta de suporte à tomada de decisão.

4.1. APRESENTAÇÃO

Diante da oportunidade de analisar o comportamento do Modelo de Análise de Riscos de Segurança sob uma ótica de dados reais, foi selecionado um sistema em produção no ambiente da Polícia Federal. E, a partir do levantamento dos componentes de *software* e *hardware* envolvidos com a disponibilidade de tal sistema, foram cedidos os dados dos controles de segurança (Seção 2.1) aplicados a estes ativos de informação para que fosse possível aplicar a metodologia proposta pelos autores do modelo.

Uma vez que os órgãos da Administração Pública Federal seguem as orientações de Gestão de Segurança da Informação conforme a definição do Gabinete de Segurança Institucional da Presidência da República (GSI/PR) e tais orientações são baseadas na família de normas da ABNT ISO/IEC 27000, os controles de segurança considerados neste estudo de caso foram modificados em relação ao modelo original da norma NIST SP 800-53 para a norma ABNT ISO/IEC 27002. Como a própria norma NIST SP 800-53 sugere em seu apêndice H uma relação de correspondência entre as duas publicações, os controles de segurança foram analisados e selecionados conforme detalhamento a seguir.

O conhecimento de especialistas foi obtido a partir da opinião de profissionais da área de segurança da informação em relação à classificação dos níveis de risco que representam as evidências trazidas pelos controles de segurança observados. Neste estudo, foram analisados os dados relativos ao grupo “Controle de acesso” da norma ABNT ISO/IEC 27002 (ABNT, 2013b), o qual corresponde à Seção 9 desta norma e possui quatro objetivos específicos: (9.1) limitar o acesso à informação e aos recursos de processamento da informação; (9.2) assegurar

acesso de usuário autorizado e prevenir acesso não autorizado a sistemas e serviços; (9.3) tornar os usuários responsáveis pela proteção das suas informações de autenticação; (9.4) prevenir o acesso não autorizado aos sistemas e aplicações. A Tabela 8 detalha os controles de segurança previstos na norma para alcançar os objetivos específicos 9.1, 9.2, 9.3 e 9.4.

Objetivo	*Controle*	*Descrição*
9.1	9.1.1	Estabelecer uma política de controles de acesso físico e lógico
	9.1.2	Autorizar previamente o acesso à rede e aos serviços de rede
9.2	9.2.1	Estabelecer processo formal para registro e cancelamento de usuário
	9.2.2	Estabelecer processo formal para concessão e cancelamento de privilégios dos usuários
	9.2.3	Estabelecer processo formal para concessão e cancelamento de acesso administrativo nos ativos
	9.2.4	Gerenciar a informação de autenticação secreta dos usuários
	9.2.5	Analisar periodicamente os direitos de acesso ao ativo de informação pelos seus proprietários
	9.2.6	Remover permissões no ativo de informação quando o vínculo do usuário com a organização for interrompido ou modificado
9.3	9.3.1	Orientar os usuários sobre a utilização de suas senhas individuais para acesso aos ativos de informação da organização
9.4	9.4.1	Restringir o acesso aos dados e funções do ativo de informação de acordo como perfil do usuário
	9.4.2	Estabelecer um procedimento seguro para acesso ao ativo de informação
	9.4.3	Assegurar o uso de senhas de qualidade para acesso ao ativo de informação
	9.4.4	Restringir o uso de programas capazes de sobrepor os controles dos ativos de informação

Tabela 8. Controles de segurança do grupo Controle de Acesso (ABNT, 2013b).

Neste estudo de caso, sob a visão dos ativos de informação que compõem o sistema da Polícia Federal, foi monitorado o controle 9.1.2 no sentido de identificar para cada ativo qual o nível de segurança de autorização aplicado: Alto (autorização individual requerida para acesso ao ativo), Médio (autorização compartilhada requerida para acesso ao ativo) e Baixo (nenhuma autorização requerida para acesso ao ativo). Também foi monitorado o controle 9.2.4, no qual se relacionou o nível de segurança com o tipo de autenticação aplicado no ativo de informação: Alto (autenticação por dois fatores), Médio (autenticação centralizada), Regular (autenticação local), Fraco (sem autenticação). Além desses, foram monitorados os controles 9.4.2 e 9.4.3, nos quais se pautou o nível de segurança com a quantidade de diretrizes de segurança aplicada em cada controle.

A norma ABNT ISO/IEC 27002 (ABNT, 2013b) lista as diretrizes de implementação sugeridas para cada controle. Assim, as Tabelas 9 e 10 descrevem tais diretrizes para os controles 9.4.2 e 9.4.3, respectivamente.

Id	*Diretrizes*
01	Não mostrar identificadores de sistema até que o processo tenha sido concluído com sucesso;
02	Mostrar um aviso geral informando que o acesso é somente para usuários autorizados;
03	Não fornecer mensagens de ajuda que possam auxiliar um usuário não autorizado;
04	Validar as informações de entrada no sistema somente quando todos os dados de entrada estiverem completos.
05	No caso de erro, não indicar qual parte do dado de entrada está correta ou incorreta;
06	Proteger contra tentativas forçadas de entrada no sistema;
07	Registrar tentativas de acesso ao sistema, sem sucesso e bem-sucedida;
08	Comunicar um evento de segurança quando detectar uma tentativa potencial ou uma violação bem-sucedida de entrada no sistema;
09	Quando o procedimento de entrada no sistema finalizar com sucesso: - Mostrar data e hora da última entrada no sistema com sucesso; - Mostrar registros de tentativas sem sucesso de entrada no sistema, desde o último acesso com sucesso;
10	Não mostrar a senha que está sendo informada;
11	Não transmitir senhas em texto claro pela rede;
12	Encerrar sessões inativas após um período definido de inatividade;
13	Restringir os tempos de conexão para fornecer segurança adicional nas aplicações de alto risco e para reduzir a janela de oportunidade para acesso não autorizado;

Tabela 9. Diretrizes de implementação do controle 9.4.2 (ABNT, 2013b).

Id	*Diretrizes*
01	obrigar o uso individual de ID de usuário e senha para manter responsabilidades;
02	permitir que os usuários selecionem e modifi quem suas próprias senhas, incluindo um procedimento de confirmação para evitar erros;
03	obriguar a escolha de senhas de qualidade;
04	obrigar os usuários a mudarem as suas senhas temporárias no primeiro acesso ao sistema;
05	forçar as mudanças de senha a intervalos regulares, conforme necessário;
06	manter um registro das senhas anteriores utilizadas e bloqueie a reutilização;
07	não mostrar as senhas na tela quando forem digitadas;
08	armazenar os arquivos de senha separadamente dos dados do sistema da aplicação;
09	armazenar e transmitr as senhas de forma protegida

Tabela 10. Diretrizes de implementação do controle 9.4.3 (ABNT, 2013b).

4.2. AMOSTRA DE DADOS

Assim como no modelo original, os controles de segurança selecionados para coleta de dados foram escolhidos porque tinham a capacidade de representar o nível de exposição ao

risco da aplicação XYZ da Polícia Federal com relação ao grupo “Controle de acesso” da norma ABNT ISO/IEC 27002 (ABNT, 2013b). Foi realizado o trabalho de determinação dos valores possíveis para cada risco (espaço de estados) e a classificação do risco estimado (conhecimento de especialistas) frente às possibilidades de estado dos demais controles. A Tabela 11 apresenta quais foram os controles de segurança utilizados neste estudo de caso e os seus possíveis estados.

ID	*Norma/Controle*	*Estados*
9.1.2	Nível de segurança de autorização	Alto, Médio, Baixo.
9.2.4	Nível de segurança de autenticação	Efetivo, Médio, Regular, Fraco.
9.4.2	Nível de segurança do processo de logon	Efetivo, Médio, Regular, Fraco.
9.4.3	Nível de qualidade da senha	Efetivo, Médio, Regular, Fraco.
9.0.0	Nível de segurança estimado	Alto, Médio, Baixo.

Tabela 11. Controles de acesso (segurança lógica).

Para armazenamento no banco de dados BD1, os estados foram transformados em dados discretos no intervalo de 01 até 05, em que 01 representa o nível mais seguro e 05 representa o nível menos seguro. Assim, dependendo da quantidade de estados possíveis para o controle de segurança, o valor discreto da variável foi classificado conforme a Tabela 12.

Quantidade de estados possíveis	*Valores discretos*
(04) Quatro	Efetivo (1), Médio (2), Regular (4), Fraco (5).
(03) Três	Alto (1), Médio (3), Baixo (5).
(02) Dois	Presente (1), Ausente (5)

Tabela 12. Valores discretos dos estados das variáveis.

Nos casos dos controles de segurança que se baseiam em uma medida do nível de proteção do ativo (9.4.2 e 9.4.3), como não existe um peso em relação à existência de diretrizes mais importantes, definiu-se que os estados das variáveis seriam classificados conforme a quantidade de diretrizes aplicadas em cada controle. A Tabela 13 representa esse relacionamento.

Nível de segurança	*Quantidade relativa de diretrizes implementadas*
Efetivo	Mais de 90% das diretrizes aplicadas
Médio	De 70% até 90% das diretrizes aplicadas
Regular	De 50% até 70% das diretrizes aplicadas
Fraco	Menos de 50% das diretrizes aplicadas

Tabela 13. Diretrizes de implementação do controle 9.4.2 (ABNT, 2013b).

De posse dessas padronizações e da realização da classificação do risco pelos especialistas em segurança da informação o banco de dados BD1 foi completamente

preenchido. Para o grupo analisado observa-se a possibilidade das variáveis se apresentarem em uma das 192 possibilidades de configurações dos estados das variáveis deste grupo. Como um exemplo de classificação, a instância de valores {9.1.2=”Alto”; 9.2.4=”Médio”, 9.4.2=”Médio”, 9.4.3=”Baixo”} teve a classificação de risco estimado (9.0.0) no estado de risco “Médio”. Assim, a Tabela 14 representa a frequência de estados definidos pelo conhecimento dos especialistas para a variável “Risco Estimado”.

Nível de segurança estimado do controle 9.0.0	*Total de ocorrências*
Alto	12
Médio	50
Baixo	130

Tabela 14. Frequências observadas da classificação dos especialistas

Após a determinação do banco de dados BD1, o próximo passo foi identificar os ativos de informação (Seção 2.1) envolvidos com a aplicação XYZ, em produção na Polícia Federal, para que fosse possível iniciar a coleta dos dados relativos aos controles de segurança desses ativos e assim fornecer subsídios para construção do banco de dados BD2.

Considerando o uso de uma aplicação com modelo de multicamadas (MICROSOFT, 2009), na qual os ativos de informação da camada de apresentação são responsáveis pelo tratamento das requisições do usuário, os ativos da camada de aplicação têm acesso aos dados de negócio e respondem às requisições da camada superior, enquanto que os ativos da camada de dados fornecem a interface com o *hardware* de armazenamento. Além desses, por se tratar de componentes da infraestrutura que suportam o sistema, também foram considerados os ativos de informação da camada de virtualização, que suportam a instalação dos ativos das camadas superiores, e os ativos da camada de rede, que fornecem a comunicação entre as diferentes camadas. Assim, a Tabela 15, reproduz os ativos de informação avaliados neste estudo de caso, classificados por função que exercem na aplicação XYZ.

Função	*Ativos de informação*
Camada de Apresentação	SRV010, SRV011, SRV057, SRV058, SRV149, SRV150
Camada de Aplicações	SRV103, SRV065, SRV294, SRV295, SRV296, SRV298
Camada de Dados	SRV059, STO001, STO002
Camada de Virtualização	VMP001, VMP002, VMP004, VMP005
Camada de Rede	SWI001, SWI002, FIR001, FIR002, BAL001, BAL002

Tabela 15. Ativos de informação da aplicação XYZ.

Utilizando as informações desse levantamento, os dados dos ativos da Tabela 15 com relação aos controles de segurança descritos nas Tabelas 10 e 11 foram coletados durante um período de 30 dias para a definição do banco de dados BD2.

Foi observada uma padronização nas ocorrências dos casos coletados para o grupo de controles de segurança analisado e também a repetição de tais casos ao longo do período de observação. A Tabela 16 apresenta a frequência dos casos armazenados em BD2.

Controle 9.1.2	*Controle 9.2.4*	*Controle 9.4.2*	*Controle 9.4.3*	*Total de Casos*
Alto	Regular	Regular	Regular	330
Alto	Médio	Regular	Alto	120
Alto	Médio	Regular	Médio	150
Médio	Regular	Regular	Fraco	120

Tabela 16. Casos observados para o grupo "Controles de acesso".

Diante da uniformidade dos casos existentes em BD2, uma perspectiva importante a ser discutida na aplicação deste modelo de análise de riscos de segurança é o uso de uma abordagem de aprendizado da rede Bayesiana considerando uma base de dados completa ou considerando uma base de dados incompleta. Essa discussão se mostra necessária porque quando considerados apenas os dados históricos (BD2) percebe-se claramente uma situação do mundo real para o problema de dados ausentes, na qual vários casos não são amostrados e por isso não se tem uma perspectiva da frequência com que ocorrem. Entretanto, como o modelo prevê a combinação dos dados históricos com o conhecimento dos especialistas (BD1), ao se agregar as duas bases de dados para construção da rede Bayesiana, nota-se que todos os casos passam a ter ao menos uma ocorrência e consequentemente uma classificação válida. Mas, apesar disso, não transmitem a relação de frequência de tais ocorrências para o modelo.

A despeito dessa última discussão, a qual será retomada mais à frente, os bancos de dados BD1 e BD2 foram preparados em arquivos MATLAB na forma de matrizes nas quais cada coluna recebia dados de uma variável (controle de segurança) e as linhas representavam os diferentes casos do banco de dados de treinamento. Foi preparado o algoritmo de Otimização por Colônia de Formigas e como seus dados de entrada foram utilizados os valores armazenados em BD1 (conhecimentos de especialistas) e em BD2 (registro histórico). Esse algoritmo foi elaborado de modo a aplicar os princípios de comportamento coletivo da técnica e assim conseguir chegar à rede Bayesiana que represente as dependências entre as variáveis.

Com o objetivo de avaliar o comportamento do algoritmo de Otimização por Colônia de Formigas implementado neste estudo de caso à luz da pontuação das redes Bayesianas apresentadas como solução, foram realizados 100 diferentes ensaios utilizando os parâmetros definidos no modelo SRAM (FENG et al., 2014). Esses parâmetros têm relação direta com a convergência do algoritmo para determinação de uma solução, pois dizem respeito ao tamanho da colônia de formigas ($m = 40$), à importância do feromônio na escolha dos caminhos ($\alpha = \beta = 1$), à taxa de evaporação do feromônio ($\rho = 0{,}5$) e à quantidade de iterações do algoritmo ($itermax = 400$).

4.3. RESULTADOS

Conforme detalhado nos Capítulos 2 e 3, o processo de aprendizagem da rede Bayesiana dos controles de segurança a partir dos dados é dividido em duas fases. A primeira fase é a determinação da estrutura da rede, a qual ocorre pela meta-heurística de Otimização por Colônia de Formigas com base nos bancos de dados BD1 e BD2. A segunda fase é a determinação dos parâmetros da rede Bayesiana, representados pelas Tabelas de Probabilidade Condicional, segundo a estimativa de máxima verossimilhança.

4.3.1. ESTRUTURA DA REDE

Conforme Seção 2.4, em cada rodada (iteração) do algoritmo de Otimização por Colônia de Formigas, cada formiga constrói a sua rede Bayesiana para solução do problema e a rede com maior pontuação (métrica K2) entre todas as formigas de uma mesma rodada é selecionada para comparação com as redes selecionadas nas demais iterações. Assim, a rede com maior pontuação entre todas as rodadas será a rede selecionada. Os dados apresentados em seguida correspondem às soluções finais obtidas para busca da rede Bayesiana durante a realização de 100 ensaios com 400 iterações por ensaio.

Para o grupo de controles de segurança “Controle de Acesso” foi possível descobrir 6 diferentes redes Bayesianas distribuídas em 7 diferentes pontuações pela métrica K2. Conforme será demonstrado em seguida, o intervalo de pontuações encontradas foi de 1050 pontos, ficando entre $4{,}3605 \times 10^5$ (menor valor encontrado) e $4{,}3710 \times 10^5$ (maior valor encontrado). A Figura 5 apresenta as duas redes que se destacaram como soluções ao longo dos 100 ensaios realizados, a rede Bayesiana A foi descoberta em 74% dos ensaios, enquanto que a rede Bayesiana B foi solução em 22% dos ensaios, compreendendo assim, de forma conjunta, ao total de 96% das redes descobertas pelo algoritmo de Colônia de Formigas.

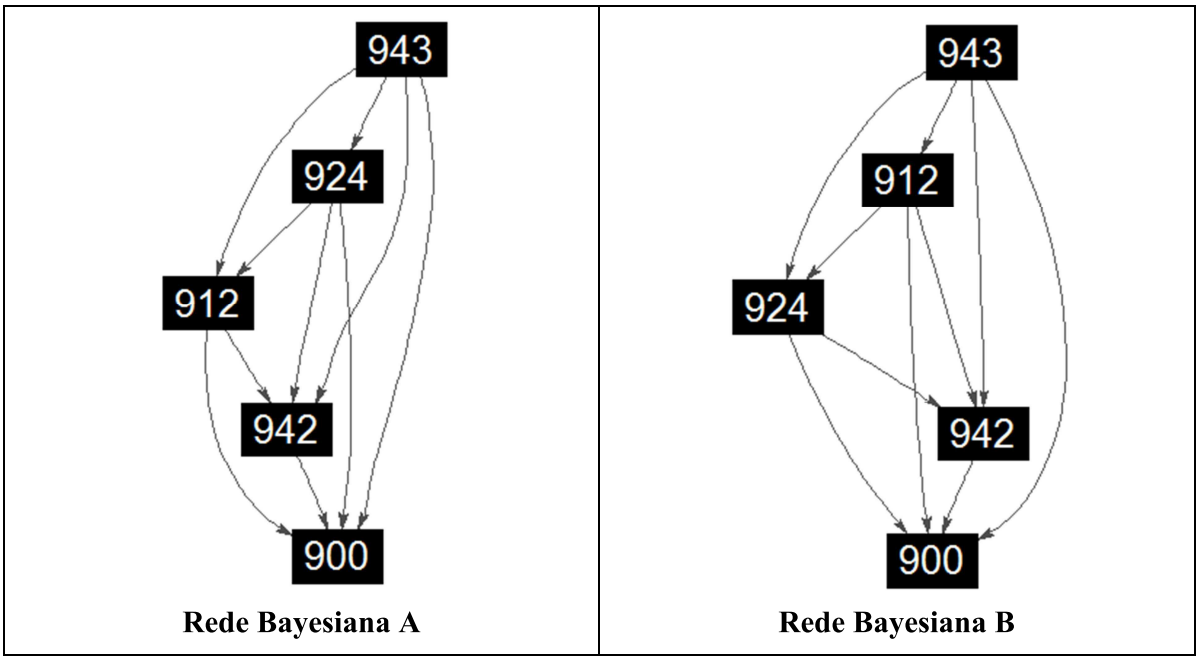


Figura 5. Redes Bayesianas de maior frequência ao longo dos ensaios.

A diferença verificada nessas duas redes de destaque está na inversão de dependência dos nós que representam os controles de segurança 9.2.4 e 9.1.2. Na primeira rede o nó 9.1.2 é dependente do valor do nó 9.2.4, enquanto que na segunda estrutura o nó 9.2.4 passa a depender da variável 9.1.2. Em termos de segurança da informação, essa alteração de dependência entre os controles 9.1.2 e 9.2.4 da norma ABNT ISO/IEC 27002 (2013) reflete a influência dos estados dos controles de segurança aplicados nas atividades de autorização e de autenticação durante o acesso aos ativos de informação do sistema analisado.

Na Rede Bayesiana A, a quantidade de diretrizes de segurança aplicada durante a entrada de um usuário no sistema (Controle 9.2.4 – Autenticação) forneceria influência sobre a forma com que é concedida a permissão de acesso aos ativos de informação deste mesmo sistema (Controle 9.1.2 – Autorização). Em contrapartida, caso seja utilizada a Rede Bayesiana B, os aspectos de dependência e causalidade ocorreriam da forma inversa, com o tipo de autorização requerida para acesso aos ativos (Individual, Compartilhada ou Nenhuma) influenciando o nível de segurança aplicado na entrada ao sistema (Efetivo, Médio, Regular ou Fraco).

No universo do total de ensaios realizados, outras 4 estruturas de redes Bayesianas foram fornecidas como soluções pelo algoritmo de Colônia de Formigas, mas obtiveram frequência de apenas 1 (uma) ocorrência para cada rede. A Figura 6 apresenta a disposição visual de tais redes e, apesar de não representarem uma frequência de soluções relevante

dentro do total de ensaios realizados, nota-se a mesma característica de alternância de posição entre os controles de segurança 9.2.4 e 9.1.2, primeiramente entre a rede Bayesiana C e a rede Bayesiana D e, posteriormente, na comparação da rede Bayesiana E com a rede Bayesiana F.

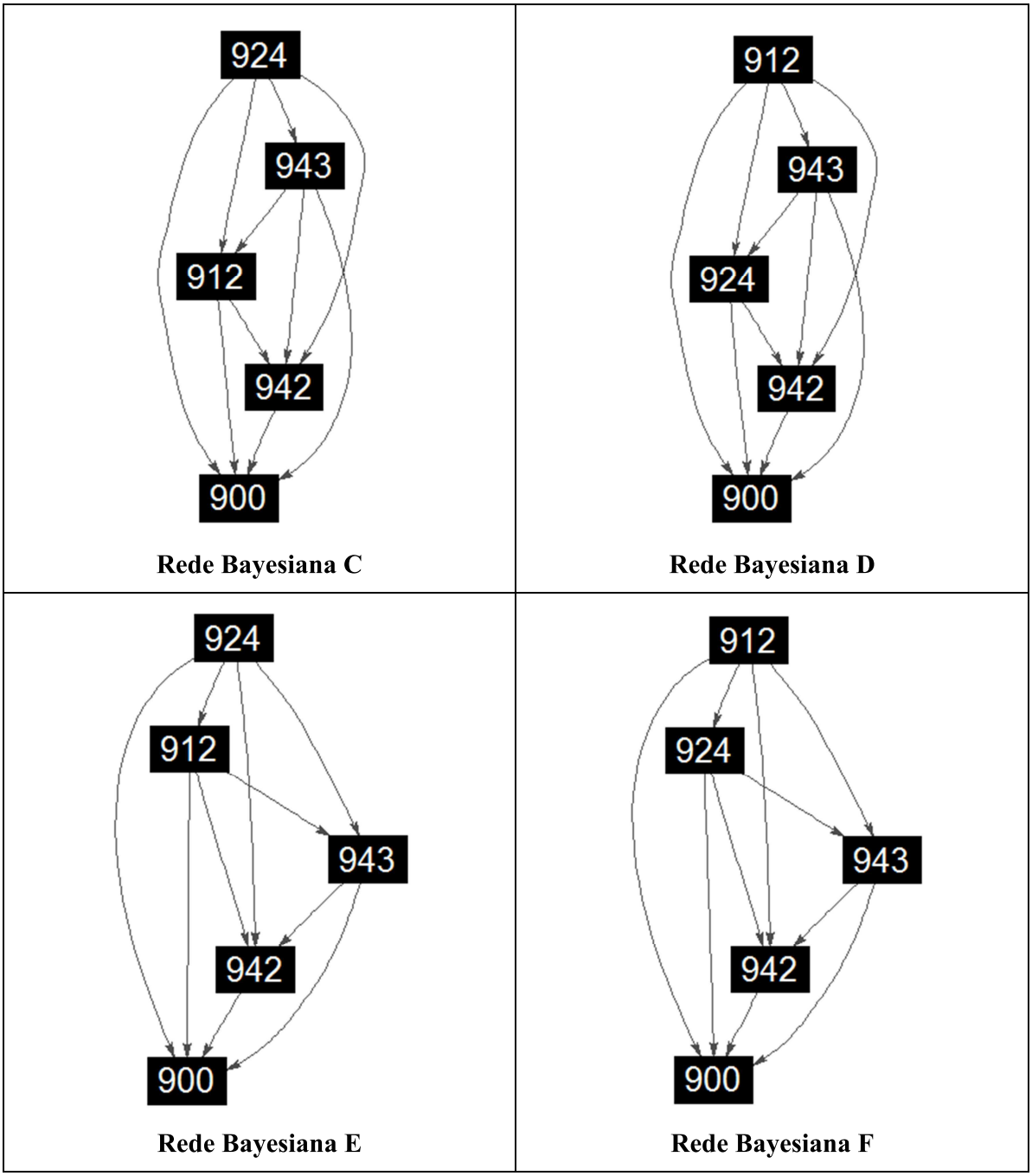
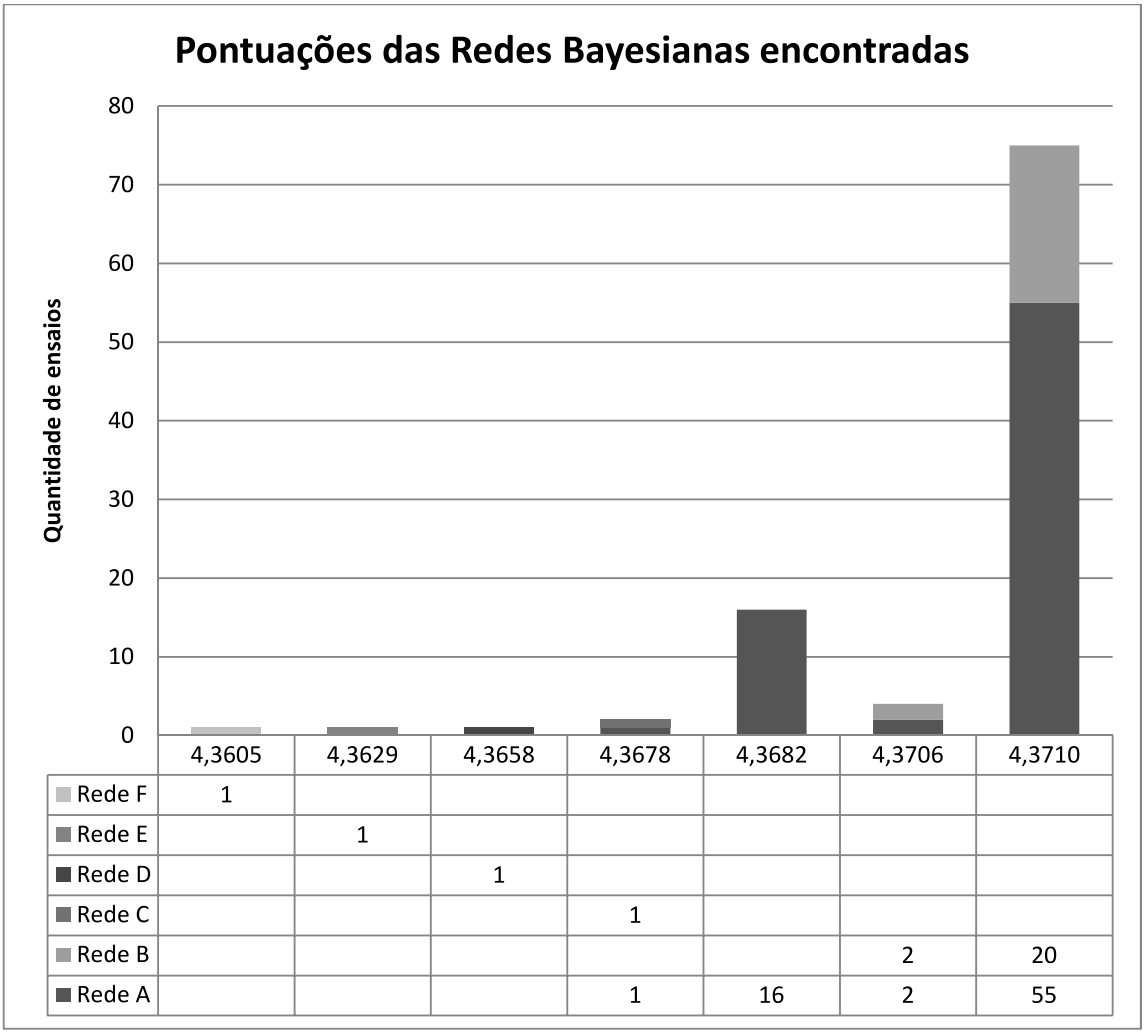


Figura 6. Redes Bayesianas de menor frequência ao longo dos ensaios.

Após coletar as estruturas de redes Bayesianas encontradas pelo algoritmo de Colônia de Formigas foram analisadas as pontuações dessas mesmas redes para que fosse possível comparar as diferentes soluções encontradas nos ensaios. A Figura 7 apresenta a distribuição

de pontuações para os 100 ensaios realizados e nela é possível verificar que o algoritmo de colônia de formigas utilizado conseguiu chegar a uma mesma maior pontuação em 75% dos ensaios.



	4,3605	4,3629	4,3658	4,3678	4,3682	4,3706	4,3710
Rede F	1						
Rede E		1					
Rede D			1				
Rede C				1			
Rede B						2	20
Rede A				1	16	2	55

Figura 7. Histograma da pontuação das redes Bayesianas encontradas nos ensaios.

Assim, foi possível verificar a descoberta de duas redes diferentes, mas que alcançaram a mesma pontuação máxima e, como o modelo original não menciona outro critério para escolha das redes senão a pontuação das estruturas pela métrica K2, a execução repetida dos experimentos permitiu perceber que a estrutura de rede Bayesiana A ocorreu com mais frequência do que a estrutura de rede Bayesiana B. As Figuras 8 e 9 demonstram as pontuações recebidas por cada uma dessas estruturas de destaque ao longo dos ensaios realizados.

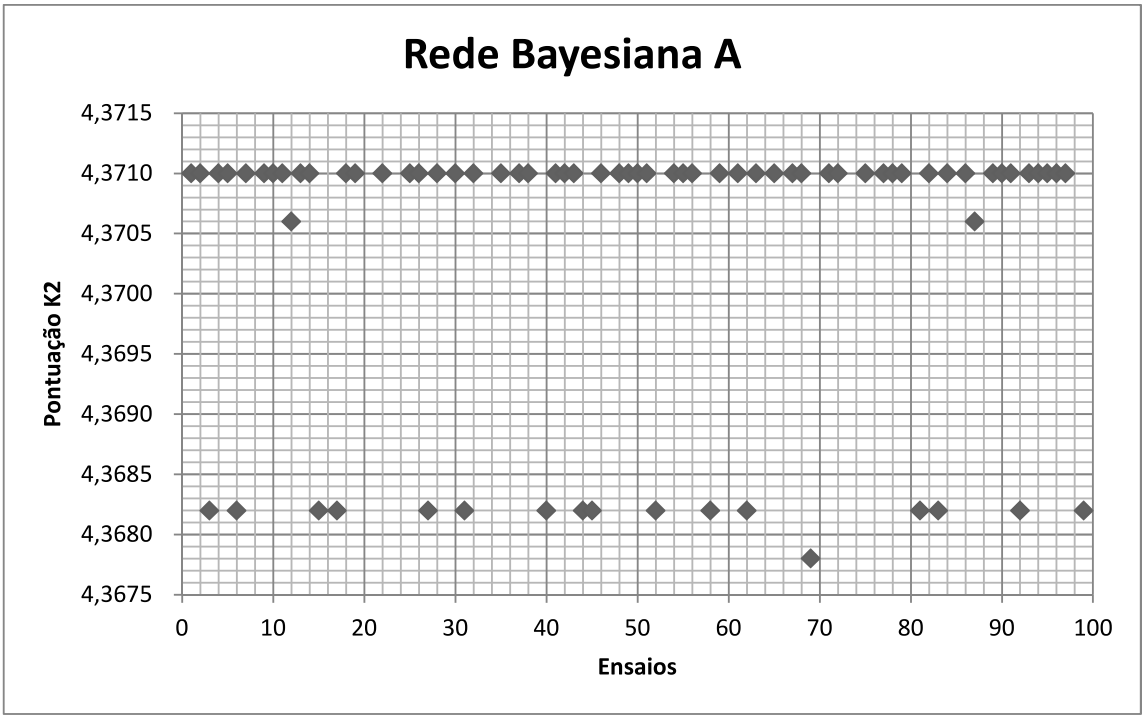


Figura 8. Pontuação da Rede Bayesiana A ao longo dos ensaios.

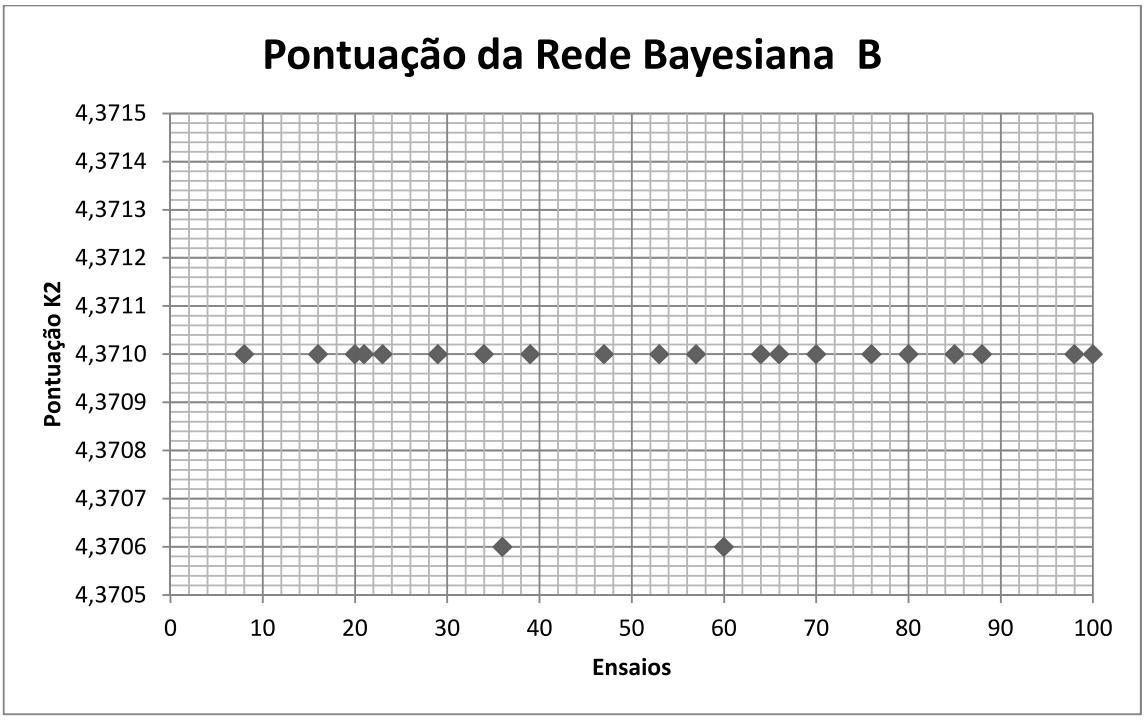


Figura 9. Pontuação da Rede Bayesiana B ao longo dos ensaios.

Com isso foi vencida a primeira fase de definição da rede Bayesiana, na qual houve destaque de duas estruturas de rede a partir da execução do algoritmo de Colônia de Formigas sob os dados de treinamento disponibilizados. Dessa maneira, como duas redes obtiveram a

mesma pontuação K2, propôs-se, então, prosseguir com o estudo de caso avaliando o comportamento de ambas as redes encontradas na etapa seguinte, que corresponde à determinação das Tabelas de Probabilidade Condicional, dadas as redes Bayesianas encontradas.

4.3.2. PARÂMETROS DA REDE

A técnica de estimativa por máxima verossimilhança foi utilizada para calcular as Tabelas de Probabilidade Condicional para inferência do nível de risco do grupo de controles de segurança avaliado. Como foram obtidas diferentes redes Bayesianas na etapa de descoberta da estrutura pelo algoritmo de Otimização por Colônia de Formigas, foram determinadas as Tabelas de Probabilidade Condicional para cada uma dessas redes.

Como nas duas redes obtidas o nó de risco estimado (9.0.0) é influenciado pelas demais variáveis analisadas (9.1.2, 9.2.4, 9.4.2, 9.4.3) então as Tabelas de Probabilidade Condicional para as redes Bayesianas A e B são idênticas porque ambas consideram a contribuição de cada nó no nível de segurança estimado e elas possuem, desta maneira, um total de 576 possibilidades, valor que pode ser calculado através da multiplicação da quantidade de estados possíveis para cada variável:

$$(9.1.2) \times (9.2.4) \times (9.4.2) \times (9.4.3) \times (9.0.0) = 3 \times 4 \times 4 \times 4 \times 3 = 576$$

Do mesmo modo, para verificar como ficou a distribuição das probabilidades da Tabela de Probabilidade Condicional do nó 9.0.0, que traduz o nível de segurança estimado do grupo “Controle de Acesso”, o qual é inversamente proporcional ao risco (quanto menor o nível de segurança, maior o risco estimado), foram gerados 4 gráficos de 4 dimensões, os quais representam o valor do controle 9.0.0 a partir dos controles 9.2.4 (eixo X), 9.1.2 (eixo Y), 9.4.2 (eixo Z) e com um valor fixo em cada gráfico para o controle 9.4.3. Assim, as Figura 10, 11, 12 e 13 correspondem à distribuição das probabilidades quando o nível de segurança do controle 9.4.3 é Efetivo, Médio, Regular e Fraco, respectivamente. Esta classificação do controle está registrada na Tabela 11 (Seção 4.2).

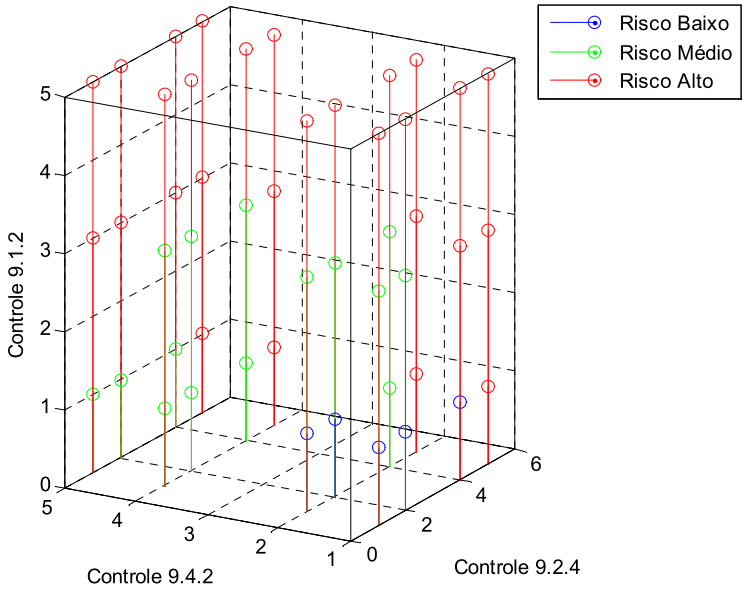


Figura 10. Nível de segurança do nó 9.0.0 quando o controle 9.4.3 é Efetivo.

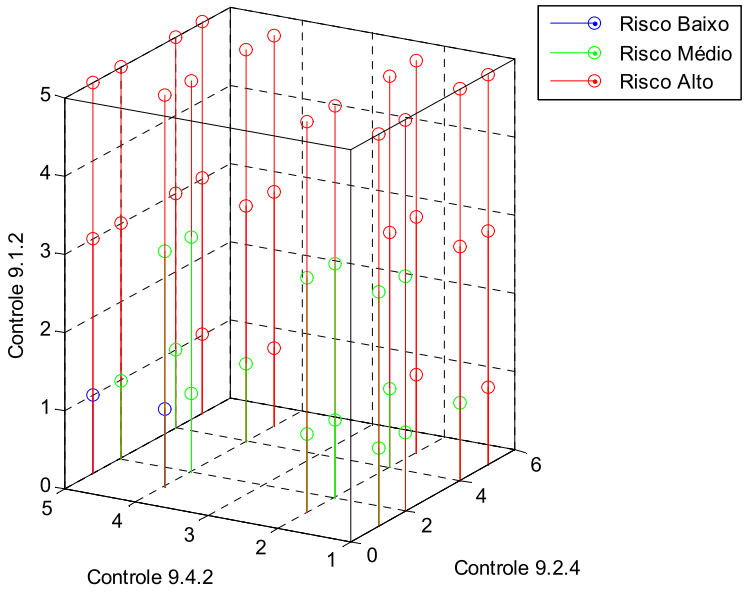


Figura 11. Nível de segurança do nó 9.0.0 quando o controle 9.4.3 é Médio.

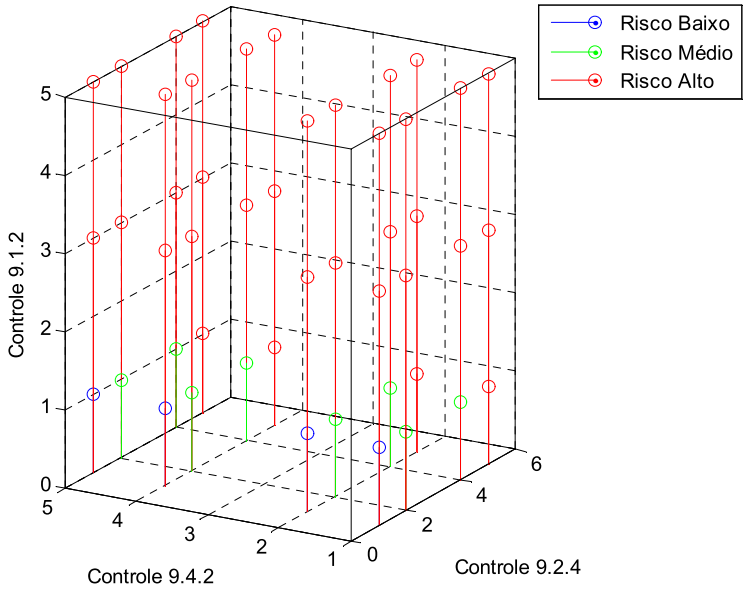


Figura 12. Nível de segurança do nó 9.0.0 quando o controle 9.4.3 é Regular.

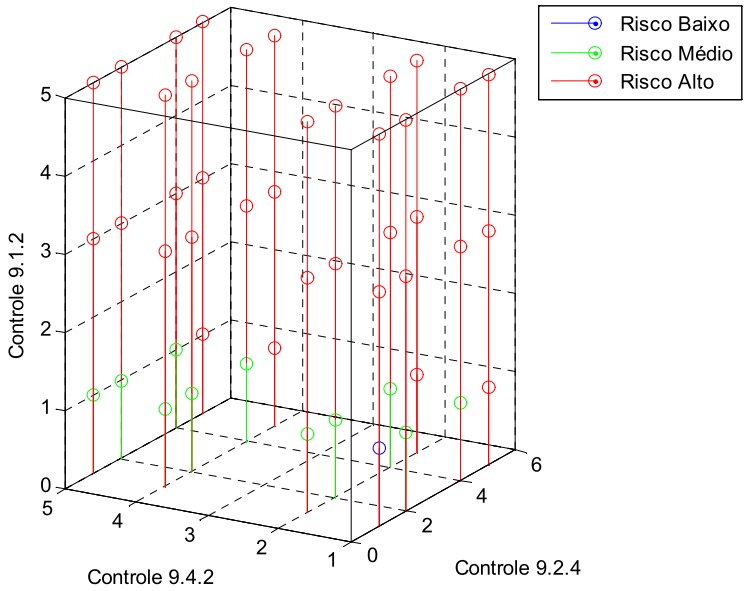


Figura 13. Nível de segurança do nó 9.0.0 quando o controle 9.4.3 é Baixo.

Com isso, foi possível verificar a predominância da probabilidade do Grupo 9 possuir um maior risco quando os controles de segurança observados também apresentaram maior risco. Observando as Figuras 11 e 12, quando a qualidade da senha é melhor (9.4.3 = Efetivo ou 9.4.3 = Médio) percebe-se maior ocorrência de configurações nas quais o nível do risco estimado para o grupo é Baixo ou Médio. Entretanto, observando as Figuras 12 e 13, quando

a qualidade da senha é menor (9.4.3 = Regular ou 9.4.3 = Fraco) o risco estimado é predominantemente Alto, ou seja, o nível de segurança para o sistema analisado será Baixo.

Acerca das duas redes Bayesianas descobertas pelo algoritmo de colônia de formigas a partir do conjunto de dados de treinamento, foi possível observar que ambas obtiveram a mesma pontuação máxima e a mesma relação de causalidade entre os controles de segurança na formação do nível de risco estimado para o sistema de informação da Polícia Federal.

Entretanto, analisando a diferença particular dessas redes encontradas (Figura 5), sob a perspectiva dos parâmetros calculados para os controles de segurança envolvidos nessa distinção (9.1.2 e 9.2.4), observa-se a ocorrência de diferentes Tabelas de Probabilidade Condicional, a depender da estrutura de rede utilizada. As Tabelas 17 e 18 representam os parâmetros do controle de segurança 9.1.2 utilizando as duas estruturas de rede.

943	*924*	*P(912= Alto)*	*P(912=Médio)*	*P(912=Baixo)*
Efetivo	Efetivo	33,33%	33,33%	33,33%
Efetivo	Médio	93,94%	3,03%	3,03%
Efetivo	Regular	33,33%	33,33%	33,33%
Efetivo	Fraco	33,33%	33,33%	33,33%
Médio	Efetivo	33,33%	33,33%	33,33%
Médio	Médio	95,06%	2,47%	2,47%
Médio	Regular	33,33%	33,33%	33,33%
Médio	Fraco	33,33%	33,33%	33,33%
Regular	Efetivo	33,33%	33,33%	33,33%
Regular	Médio	33,33%	33,33%	33,33%
Regular	Regular	97,66%	1,17%	1,17%
Regular	Fraco	33,33%	33,33%	33,33%
Fraco	Efetivo	33,33%	33,33%	33,33%
Fraco	Médio	33,33%	33,33%	33,33%
Fraco	Regular	3,03%	93,94%	3,03%
Fraco	Fraco	33,33%	33,33%	33,33%

Tabela 17. TPC do nó 9.1.2 para a Rede Bayesiana A da Figura 5.

943	*P(912= Alto)*	*P(912=Médio)*	*P(912=Baixo)*
Efetivo	80,95%	9,52%	9,52%
Médio	83,84%	8,08%	8,08%
Regular	91,53%	4,23%	4,23%
Fraco	9,52%	80,95%	9,52%

Tabela 18. TPC do nó 9.1.2 para a Rede Bayesiana B da Figura 5.

Assim, conforme estrutura e tabelas obtidas, na primeira configuração de rede o controle de segurança que quantifica o nível do risco pelo tipo de autorização aplicado para acesso ao ativo de informação (controle 9.1.2) é influenciado pelos controles de segurança

que representam o tipo de autenticação (controle 9.2.4) e a qualidade da senha (controle 9.4.3), ambos aplicados no acesso ao ativo de informação.

De acordo com a definição dos controles de segurança utilizados (Seção 4.1) e a partir desses dados coletados podemos analisar um caso específico, no qual quando uma senha de qualidade fraca (9.4.3 = Fraco) é utilizada durante o acesso a um ativo de autenticação local (9.2.4 = Regular), extraímos da Tabela 17 a probabilidade de 93,94% da autorização de acesso ao ativo ser compartilhada (9.1.2 = Médio).

Agora, utilizando a segunda configuração de rede e o cálculo da Tabela 18 podemos verificar que esse mesmo controle de segurança (9.1.2) passa a depender apenas do controle de segurança que representa a qualidade da senha (9.4.3). E, mantendo o exemplo anterior, a probabilidade da autorização de acesso ao ativo ser compartilhada (9.1.2 = Médio) é de 80,95% para o uso de uma senha de qualidade fraca (9.4.3 = Fraco).

A fim de analisar o comportamento do nó 9.2.4, foram construídas as Tabelas 19 e 20 de representação dos parâmetros deste controle de segurança nas duas redes encontradas. Na rede Bayesiana A, o nível de segurança de autenticação de acesso ao ativo de informação está em função apenas da qualidade da senha (controle 9.4.3) enquanto que, na rede Bayesiana B, ele passa a ser influenciado também pelo tipo de autorização aplicado no acesso ao ativo de informação (controle 9.1.2).

Mantendo o exemplo anterior para fins de comparação, uma senha de qualidade fraca (9.4.3 = Fraco) produz a probabilidade 78,57% da autenticação no acesso ao ativo ser do tipo local (9.2.4 = Regular) na rede Bayesiana A e, caso esteja utilizando a rede Bayesiana B e o nível de autorização for compartilhado (9.1.2 = Médio), temos pela Tabela 20 que a probabilidade da autenticação no acesso ao ativo ser do tipo local passa a ser 91,18%.

943	*P(924= Efetivo)*	*P(924= Médio)*	*P(924=Regular)*	*P(924=Fraco)*
Efetivo	7,14%	78,57%	7,14%	7,14%
Médio	6,06%	81,82%	6,06%	6,06%
Regular	3,18%	3,18%	90,48%	3,18%
Fraco	7,14%	7,14%	78,57%	7,14%

Tabela 19. TPC do nó 9.2.4 para a Rede Bayesiana A da Figura 5.

943	*912*	*P(924= Efetivo)*	*P(924= Médio)*	*P(924=Regular)*	*P(924=Fraco)*
Efetivo	Alto	2,94%	91,18%	2,94%	2,94%
Efetivo	Médio	25,00%	25,00%	25,00%	25,00%
Efetivo	Baixo	25,00%	25,00%	25,00%	25,00%
Médio	Alto	2,41%	92,77%	2,41%	2,41%
Médio	Médio	25,00%	25,00%	25,00%	25,00%
Médio	Baixo	25,00%	25,00%	25,00%	25,00%
Regular	Alto	1,16%	1,16%	96,53%	1,16%
Regular	Médio	25,00%	25,00%	25,00%	25,00%
Regular	Baixo	25,00%	25,00%	25,00%	25,00%
Fraco	Alto	25,00%	25,00%	25,00%	25,00%
Fraco	Médio	2,94%	2,94%	91,18%	2,94%
Fraco	Baixo	25,00%	25,00%	25,00%	25,00%

Tabela 20. TPC do nó 9.2.4 para a Rede Bayesiana B da Figura 5.

Após a definição das Tabelas de Probabilidade Condicional para cada nó da rede, obteve-se a especificação completa da rede Bayesiana que representa o relacionamento conjunto dos controles de segurança aplicados ao sistema de informação XYZ da Polícia Federal. Com isso, a rede pode ser utilizada, por exemplo, como dado de entrada em um algoritmo de inferência probabilística para estimar os níveis de segurança das variáveis da rede pela probabilidade *a posteriori* quando forem detectadas novas evidências.

Em comparação ao trabalho de Feng *et al* (2004), o estudo de caso realizado até aqui corresponde à primeira etapa do Modelo de Análise de Riscos de Segurança e a próxima fase segundo o modelo SRAM envolveria o monitoramento do banco de dados DB3, que recebe as informações em tempo real dos controles de segurança analisados para inserção como evidências na rede.

4.4. SÍNTESE DO CAPÍTULO

Neste capítulo foi possível examinar a aplicação da primeira etapa do modelo de análise de riscos de segurança em uma infraestrutura de sistemas computacionais do ambiente de produção da Polícia Federal. Inicialmente, verificou-se a questão de escolha dos controles de segurança baseados na norma ABNT ISO/IEC 27002, foram definidos os estados possíveis para cada variável e os padrões para armazenamento nos bancos de dados BD1 e BD2. A aplicação XYZ da Polícia Federal teve seus componentes mapeados e os dados dos controles de segurança escolhidos foram armazenados para aplicação da técnica.

Destaca-se a discussão sobre como compreender o banco de dados de treinamento (BD1 + BD2) para aprendizado da rede Bayesiana, uma vez que pode ser encarado como uma amostra de dados completa ou uma amostra de dados incompleta. Nesse tópico, Russell e Norvig (2004) relatam este tipo de situação como uma aprendizagem com dados completos, mas com uma amostra de dados muito pequena, na qual algumas situações ainda não tenham sido observadas, e o artificio utilizado para contornar o cálculo das estimativas de máxima probabilidade é iniciar a contagem para cada evento com o valor um ao invés de zero.

A partir do desenvolvimento do Algoritmo 1 (Seção 3.2), que aplica a meta-heurística de colônia de formigas para descoberta da rede Bayesiana utilizando o banco de dados de treinamento formado por BD1 e BD2, foram obtidas duas estruturas de redes Bayesianas que alcançaram a mesma pontuação pela métrica K2. Cabe ressaltar que o modelo original se baseia nessa métrica para definição da melhor rede, motivo pelo qual não há previsão para uma situação de conflito entre redes igualmente pontuadas.

No sentido de validar o funcionamento do algoritmo ACO para descoberta da Rede Bayesiana que representasse os relacionamentos de dependência entre os controles de segurança aplicados ao sistema de informação XYZ, foram realizados repetidos ensaios do algoritmo desenvolvido sob o conjunto de dados de treinamento. A partir desses ensaios foi possível perceber destaque para estrutura de Rede Bayesiana A, superando em 2,5 vezes a quantidade de estruturas Rede Bayesiana B descobertas. A etapa seguinte foi realizada para as duas estruturas de redes Bayesianas de destaque e trata da estimação dos parâmetros das redes através do cálculo das Tabelas de Probabilidade Condicional pela estimativa por máxima verossimilhança.

Notou-se ao calcular as Tabelas de Probabilidade Condicional do nó 9.0.0, o qual representa o nível de segurança do grupo analisado e também reproduz a medida de risco estimado, que em todas as configurações de rede a tabela deste nó era igual. O que se explica pela sua dependência direta das mesmas variáveis nas duas redes.

Em relação à segurança da informação, as redes Bayesianas encontradas a partir dos dados analisados demonstram que os riscos estimados para o grupo “Controle de Acesso” da aplicação XYZ são decorrentes do tipo de autorização requerida para acesso ao ativo (individual, compartilhada, nenhuma), do tipo de autenticação aplicada (dois fatores, centralizada, local, nenhuma), e da quantidade de diretrizes aplicadas durante o processo de entrada no sistema e na qualidade exigida durante a formação da senha individual.

Apesar do nível de segurança geral do grupo não ter se alterado em relação às estruturas encontradas, foram encontradas diferenças nos relacionamentos individuais entre as variáveis. Sobre essa situação, foram coletadas as Tabelas de Probabilidade Condicional dos nós que sofreram alterações de uma rede para a outra. Como observado no capítulo, a principal discussão de mudança envolve os nós relacionados com os controles relativos ao nível de segurança de autorização (9.1.2) e de autenticação (9.2.4). Na rede Bayesiana A da Figura 5, o nível de autorização requerida é dependente do nível de autenticação aplicado ao ativo de informação, na rede Bayesiana B ocorre exatamente o contrário.

Logo, não houve uma situação em que um dos controles de segurança não influenciasse diretamente o risco conjunto da aplicação (nó 9.0.0) e pudesse dessa maneira ser considerado apenas no cálculo de outras variáveis da rede. Caso essa situação tivesse ocorrido, em termos de tomada de decisão, seria possível priorizar os controles de segurança diretamente ligados ao risco estimado e deixar os controles intermediários para uma segunda etapa.

Mas como neste estudo todas as variáveis se fizeram importantes na medida do risco estimado, foi possível destacar o fato do controle de segurança 9.4.3, referente ao nível de qualidade exigida na formação da senha, ser a variável que exerce influência sobre todos os nós nas duas redes Bayesianas selecionadas.

Por fim, foi possível definir completamente a estrutura da rede Bayesiana a partir dos dados históricos e do conhecimento de especialistas sobre os controles de segurança avaliados e com isso finalizar a primeira etapa do modelo de analise de riscos de segurança escolhido.

5. CONCLUSÕES

O monitoramento contínuo é reconhecidamente uma etapa fundamental nos modelos de referência para tratamento dos Riscos de Segurança da Informação e, frente às ameaças avançadas, o uso de técnicas de tratamento automatizado de dados deve ser aplicado com o objetivo principal de reservar ao esforço humano a incumbência da tomada de decisões ou a análise de dados previamente tratados.

O Modelo de Análise de Riscos de Segurança proposto por Feng *et al* (2014) está inserido neste conjunto de ferramentas que procuram fornecer uma visão gerencial para o problema de análise de riscos em sistemas computacionais. Do modo como foi proposto, a sua principal contribuição é permitir aliar o conhecimento dos especialistas da área de segurança da informação com os dados históricos dos controles de segurança, que podem ser obtidos por diferentes ferramentas existentes no mercado. Este modelo consegue trazer para uma mesma perspectiva a visão sobre o risco conjunto de diferentes *hardwares* e *softwares* que compõem determinado serviço de TI.

Outro ponto forte do modelo é a sólida fundamentação teórica para cada uma das diferentes etapas que compõem o modelo, na qual é relevante o uso de redes Bayesianas para representar as relações de dependências entre os controles do modelo, o emprego da meta-heurística de Otimização por Colônia de Formigas para aprendizado da rede e para determinação dos caminhos com maior exposição ao risco, a estimação das probabilidades a partir de novas evidências.

Por outro lado, diante da variedade de componentes existentes (servidores Linux e Windows, ativos de rede, ativos de armazenamento, etc.), destaca-se o esforço necessário para padronizar a saída de dados dessas diferentes ferramentas para uma mesma medida, assim como foi feito na Seção 4.2, quando definiu valores discretos para os controles de segurança.

Em relação à preparação dos dados destaca-se a situação mencionada na Seção 4.2, na qual a uniformidade dos registros históricos apresentou poucos casos únicos e isso levou ao questionamento sobre que tipo de abordagem fosse necessária tomar para realizar as tarefas de aprendizado da estrutura e dos parâmetros da Rede Bayesiana. O modelo original relata o emprego da abordagem com dados completos porque a empresa observada tinha um histórico considerável dos controles. No caso da Polícia Federal, foram considerados 30 dias de histórico de casos, mas devido aos ativos envolvidos na coleta estarem em produção e neste

período não ter ocorrido mudanças nesses componentes, houve um comportamento fixo no histórico dos controles de segurança avaliados.

Outra discussão importante foi sobre qual a rede Bayesiana utilizar, porque apesar de duas estruturas de mesma pontuação, uma delas se destacou ao realizar o ensaio de descoberta da rede Bayesiana repetidas vezes. Diante das variações obtidas no cálculo das probabilidades dos nós utilizando-se de uma estrutura ou de outra, na visão de segurança da informação, não foi possível quantificar qual delas poderia melhor representar os dados coletados do sistema XYZ, da Polícia Federal.

Entretanto, como dado de gestão obtido pela aplicação desta etapa do modelo, é admissível afirmar que melhorias no controle de segurança relativo à qualidade da senha exigida para acesso aos ativos de informação permitem que o nível de segurança de todas as variáveis do grupo "Controle de acesso" seja elevado e com isso seja possível reduzir os riscos. Isto porque nas duas Redes Bayesianas de destaque, obtidas pelo algoritmo de Colônia de Formigas, o controle de segurança 9.4.3, que é relativo à qualidade da senha, foi fator determinante para o estado das demais variáveis analisadas.

Cabe ressaltar que esta situação está em consonância com os fatos recentes da prática de segurança da informação, quando vazamentos de senhas de grandes portais de serviços da Internet como Dropbox, Yahoo e LinkedIn revelaram o uso massivo de senhas de baixa qualidade. Deste modo, forçar o cumprimento de políticas de senhas por parte dos usuários tende a minimizar os impactos de ataques de força bruta ou engenharia social em ambiente corporativo.

Assim, este estudo de caso aplicou com sucesso a primeira etapa do Modelo de Análise de Riscos de Segurança, desenvolvendo a rede Bayesiana que é a base para a realização das demais etapas. Como não foram aplicadas as demais etapas do modelo, não foi possível perceber o benefício do uso de uma abordagem automatizada no processo de monitoramento contínuo dos riscos de segurança da informação. Mas, diante dos dados de gestão extraídos, quando se destacou a influência da qualidade da senha sobre as outras variáveis e também se conseguiu agrupar em uma mesma medida os riscos de diferentes ativos de informação, é possível caracterizar esta técnica como promissora, principalmente diante das necessidades atuais do mercado, que está em busca de ferramentas de consolidação do conhecimento.

À medida que este trabalho trouxe novos questionamentos durante a aplicação da técnica em um ambiente de dados reais, ele contribuiu com desenvolvimento de abordagens futuras. E, como perspectivas, destaca-se a possibilidade de aplicar as demais etapas do modelo e com isso permitir ampliar as discussões sobre a consideração de um conjunto de dados completo ou incompleto, provocar o surgimento de novos parâmetros para seleção da melhor rede dentre aquelas encontradas pelo algoritmo de colônia de formigas, aplicar esta mesma etapa em outro conjunto de controles de segurança e aumentar a visão sobre a contribuição do modelo avaliado para a gestão de segurança da informação.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ABNT (2011) NBR ISO/IEC 27005. Tecnologia da informação – Técnicas de segurança – Gestão de riscos de segurança da informação. Rio de Janeiro: ABNT, 2011.

ABNT (2013a) NBR ISO/IEC 27002. Tecnologia da informação – Técnicas de segurança – Código de prática para gestão da segurança da informação. Rio de Janeiro: ABNT, 2013.

ABNT (2013b) NBR ISO/IEC 27001. Tecnologia da informação – Técnicas de segurança – Sistemas de gestão de segurança da informação – Requisitos. Rio de Janeiro: ABNT, 2013.

Banerjee, C., Banerjee, A. (2014). IT Security Practices in an Organization: Balancing Technology and Management Perspective. Editorial Board Chief Bebefactor, v. 495, p. 506. 2014.

Bezerra, Edson K. (2013) "Gestão de riscos de TI: NBR 27005". Rio de Janeiro: RNP/ESP, 2013.

Bishop, Christopher M. (2006). "Pattern Recognition and Machine Learning". New York: Springer-Verlag, 2006.

Casella, G., Berger, Roger L. (2010). "Inferência Estatistica". São Paulo: Cengage Learning, 2010.

Castro, Leandro N. (2006). "Fundamentals of natural computing: basic concepts, algorithms, and applications". Boca Raton: CRC Press, 2006.

Campos, L., Fernández-Luna, J. Gaméz, J. Puerta, J. (2002) "Ant colony optimization for learning Bayesian networks". International Journal of Approximate Reasoning. 31, 291-311, 2002.

Center for Internet Security. (2015). "Critical Security Controls for Effective Cyber Defense". CIS. 2015. Disponível em: <https://www.cisecurity.org/critical-controls.cfm>. Acesso em 24 de Out 2016

Chickering, D. M. (1996). "Learning Bayesian networks is NP-complete". In Learning from data (pp. 121-130). New York: Springer, 1996.

Cooper, Gregory F., Herskovits, Edward. (1992). "A Bayesian method for the induction of probabilistic networks from data". Machine Learning. v9, p. 309-347, 1992.

Crocomo, Marcio K. (2012) "Algoritmo de otimização bayesiano com detecção de comunidades". Tese (Doutorado em Ciências de Computação e Matemática Computacional). São Paulo: Universidade de São Paulo, 2012.

Cybersecurity Ventures. (2016). "Cybersecurity Marker Report Q3 2016". 2016. Disponível em: < http://cybersecurityventures.com/cybersecurity-market-report/ >; Acesso em 25 de out. 2016.

Daly, R., Shen, Q., Aitken, S. (2011). "Learning Bayesian networks: approaches and issues". The knowledge engineering review, v. 26(02), p.99-157.

Dempster, Arthur P., Laird, Nan M., Rubin, Donald B. (1977). “Maximum Likelihood from Incomplete Data via the EM Algorithm”. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1. (1977), p.1-38.

Dorigo, M., Blum, C. (2005). “Ant colony optimization theory: A survey”. Theoretical computer science, Vol. 344(2), p. 243-278.

Dorigo, M., Stützle, T. (2010). “Ant colony optimization: overview and recent advances”. Handbook of metaheuristics. Brussels: Springer US, p. 227-263, 2010.

Duda, Richard O., Hart, Peter E., Stork, David G. (2000). “Pattern Classification”, 2nd Edition. New York: Wiley-Interscience, 2000,

Fan, C., Yu, Y. (2004). “BBN-based *software* project risk management”, Journal of Systems and Software, v. 73(2) p. 193–203, 2004.

Feng, N., Wang, H. J., Li, M. (2014) “A security risk analysis model for information systems: Causal relationships of risk factors and vulnerability propagation analysis”. Information Sciences. v. 256, p. 57–73, jan. 2014.

Pratap K., Wheatman, J. (2016) “Magic Quadrant for IT Risk Management Solutions”. GARTNER GROUP. 2016.

Heckerman, D., Geiger, D., Chickering, D. M. (1995). “Learning Bayesian networks: The combination of knowledge and statistical data”. Machine learning, v. 20(3), p. 197-243. 2004.

ISACA. (2009). “The Risk IT Framework”. ISACA. Rolling Meadows, 2009.

Jensen, Finn V. (1996) “An introduction for Bayesian Networks”. London: UCL Press, 1996.

Kavanagh, K., Rochford, O. (2015). “Magic Quadrant for Security Information and Event Management”. Gartner database, ID G00267505, jul. 2015.

Kott, A., Arnold, C. (2013). “The promises and challenges of continuous monitoring and risk scoring”. IEEE Security & Privacy, v. 11(1), p. 90-93. 2013.

Luna, José E. O. (2004). “Algoritmos EM para aprendizagem de Redes Bayesianas a partir de dados incompletos”. Dissertação de Mestrado, UFMS, Campo Grande, Brasil, 2004.

Matsumoto, S., Carvalho, R. N., Ladeira, M., da Costa, P. C. G., Santos, L. L., Silva, D., Cai, K. (2011). “UnBBayes: a java framework for probabilistic models in AI”. Java in Academia and Research, 34.

Microsoft (2009). “Microsoft Application Architecture Guide, 2nd Edition”. 2009. Disponível em: <https://msdn.microsoft.com/en-us/library/ee658109.aspx>. Acesso em 24 de out. 2016

Ministério do Planejamento, Orçamento e Gestão. (2016). “Estratégia de Governança Digital da Administração Pública Federal 2016-19”. Brasília, MPOG. 2016.

National Institute of Standards and Technology (2006). “Federal Information Processing Standards Publication 200: Minimum Security Requirements for Federal Information and Information Systems”. Gaithersbug: NIST, 2006.

National Institute of Standards and Technology (2010). “NIST Special Publication 800-37: Guide for Applying the Risk Management Framework to Federal Information Systems”. Gaithersbug: NIST, 2010.

National Institute of Standards and Technology (2011a). “NIST Special Publication 800-39: Managing Information Security Risk”. Gaithersbug: NIST, 2011.

National Institute of Standards and Technology (2011b). “NIST Special Publication 800-137: Information Security Continuous Monitoring (ISCM) for Federal Information Systems and Organizations”. Gaithersbug: NIST, 2011.

National Institute of Standards and Technology (2013). “NIST Special Publication 800-53: Security and Privacy Controls for Federal Information Systems and Organizations”. Gaithersbug: NIST, 2013.

Oliveira, V. L. (2006) “Uma análise comparativa das metodologias de gerenciamento de risco FIRM, NIST SP 800-30 e OCTAVE”. Dissertação de Mestrado, UNICAMP, Campinas, Brasil, 2006.

Olhar Digital, Vazamento do LinkedIn mostra que pessoas continuam escolhendo senhas fracas. Disponível em < http://olhardigital.uol.com.br/fique_seguro/noticia/vazamento-do-linkedin-mostra-que-pessoas-continuam-escolhendo-senhas-fracas/58532>. Acesso em: 10 de Novembro de 2016.

Pearl, J. (2011). Bayesian networks. Department of Statistics, UCLA.

Pifer, C. A., (2006) “Estudo Comparativo de Métricas de Pontuação para Aprendizagem Estrutural de Redes Bayesianas”. Dissertação de Mestrado, UFRN, Natal, Brasil, 2006.

Rathie, Pushpa N., Zörnig, P. (2012). “Teoria da Probabilidade”. Brasília: Editora Unb, 2012.

Russel, S., Norvig, P. (2004) “Inteligência Artificial: tradução da segunda edição”. Rio de Janeiro: Elsevier, 2004.

Santos, Edimilson B. (2007) “A ordenação das variáveis no processo de otimização de Classificadores Bayesianos: uma abordagem evolutiva”. Dissertação de Mestrado, UFSCar, São Carlos, Brasil, 2007.

Silva, Wagner L. (2011). Segurança da Informação: um estudo sobre a percepção do usuário da informação contábil. Dissertação de Mestrado. Universidade Presbiteriana Mackenzie, São Paulo, 2011.

Stonebraker, M., Çetintemel, U. Zdonik, S. (2005). “The 8 requirements of real-time stream processing”. ACM SIGMOD Record, v. 34(4), p. 42-47, 2005.

Tribunal de Contas da União. (2012). Boas práticas em segurança da informação. 4. ed. Brasília: TCU, Secretaria de Fiscalização de Tecnologia da Informação, 2012.

Wheeler, John A. (2016). “Digital GRC: The Dawn of a New Era”. 2016. Disponível em <http://blogs.gartner.com/john-wheeler/digital-grc-the-dawn-of-a-new-era/>. Acesso em 24 de out. 2016.