U.PORTO

MESTRADO
ECONOMIA E ADMINISTRAÇÃO DE EMPRESAS

Deteção de Fraude em Redes Financeiras com Modelação Baseada em Agentes

João Guilherme Pinto Pedreira de Brito

M

2018

U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

DETEÇÃO DE FRAUDE EM REDES FINANCEIRAS COM MODELAÇÃO BASEADA EM AGENTES

João Guilherme Pinto Pedreira de Brito

Dissertação

Mestrado em Economia e Administração de Empresas

Orientado por

Pedro José Ramos Moreira de Campos

Rui Manuel Santos Rodrigues Leite

2018

Agradecimentos

Em primeiro lugar, gostaria de agradecer aos Professores que me orientaram neste desafio. Ao Professor Pedro Campos por me ter guiado, apoiado, aconselhado e motivado desde a escolha do tema até à conclusão do trabalho. E ao Professor Rui Leite, cujos conselhos, observações e conhecimentos foram imprescindíveis para que os objetivos fossem cumpridos.

Desejo também agradecer aos meus pais e ao meu irmão, por sempre me apoiarem e por me terem dado uma fundação sólida a partir da qual eu posso almejar qualquer objetivo.

E por último quero agradecer à Bárbara, por me acompanhar ao longo desta e muitas outras jornadas, ajudando-me sempre a descobrir o melhor caminho.

Resumo

A fraude é um fenómeno generalizado. Frequentemente surgem novas notícias sobre fraudes utilizando as redes de computadores e em alguns casos só passado algum tempo é que a situação é detetada. O impacto da fraude é vasto, com consequências diretas para as empresas e para a economia. Metodologias para a deteção da fraude são, por isso, essenciais. Os métodos estatísticos e de *data mining* fornecem tecnologias eficazes no apoio à realização de várias atividades nas empresas. A deteção da fraude é uma delas, existindo aplicações para detetar casos como a fraude de cartões de crédito ou intrusão em sistemas de computadores. Várias destas aplicações recorrem a redes Bayesianas e/ou à modelação baseada em agentes. Neste trabalho pretende-se implementar uma abordagem análoga, mas com o objetivo de identificar suspeitas de branqueamento de capitais. Foi desenvolvido um modelo que analisa um conjunto de transações financeiras e recorre a algoritmos de classificação para avaliar se o agente de origem é propenso à fraude ou não. O modelo baseia-se na análise de uma rede de agentes que realiza transações monetárias entre si e partindo de um conjunto de dados de transações realizadas utiliza algoritmos de *Machine Learning* para produzir ilações quanto ao comportamento fraudulento dos agentes intervenientes. Os resultados obtidos demonstram que estes podem ser devidamente classificados tendo por base dados históricos, servindo como um método para a prevenção e deteção da fraude.

Abstract

Fraud is a widespread phenomenon. There are often fresh news about fraud using computer networks and in some cases only after some time the situation is detected. The impact of fraud is wide-ranging, with direct consequences for business and the economy. Methodologies for detecting fraud are therefore essential. Statistical and data mining methods provide effective technologies to support various business activities. Fraud detection is one of them, and there are applications to detect cases such as credit card fraud or intrusion into computer systems. Several of these applications rely on Bayesian networks and / or agent-based modeling. This work intends to implement an analogous approach, but with the objective of identifying suspicions of money laundering. A model has been developed that analyzes a set of financial transactions and uses classification algorithms to assess whether the source agent is fraud prone or not. The model is based on the analysis of a network of agents that perform monetary transactions with each other and from a set of performed transactions data uses Machine Learning algorithms to produce inferences about the fraudulent behavior of the intervening agents. The results show that these can be properly classified based on historical data, serving as a method for the prevention and detection of fraud.

Índice

Índice de Figuras

Índice de Tabelas

1. Introdução

A fraude é um fenómeno tão antigo como a própria humanidade e pode tomar uma variedade quase ilimitada de diferentes formatos (Bolton et al., 2002). Num ambiente competitivo a fraude pode ser um problema crítico para um negócio se se verificar sistematicamente e se os procedimentos de prevenção não forem robustos (Phua et al., 2010). Em anos recentes o desenvolvimento de novas tecnologias, que por um lado tornaram a nossa vida mais cómoda, veio ao mesmo tempo trazer novos métodos de cometer fraude. As formas tradicionais de comportamentos fraudulentos como a lavagem de dinheiro são cada vez mais fáceis de realizar, sendo agora acompanhadas de novos tipos de fraude como a intrusão em computadores e a fraude nas comunicações móveis (Bolton et al., 2002).

Segundo um artigo da OBEGEF, "*Quando falamos em fraude estamos a englobar um vastíssimo conjunto de situações, tendencialmente intencionais, em que uns cidadãos ou instituições enganam outros, causando directa ou indirectamente danos económico-sociais. Estamos, sobretudo, a considerar os processos que se inserem no tecido económico desta sociedade crescentemente mundializada."* (Pimenta, 2009). O controlo, prevenção e deteção da fraude são então cada vez mais relevantes. É importante primeiro distinguir a prevenção da deteção da fraude. A prevenção está relacionada com medidas tomadas para evitar que a fraude seja cometida. Estas podem variar conforme o contexto e podem ir desde marcas de água ou hologramas em notas, até sistemas de segurança na internet. Em contraste, a deteção da fraude envolve identificar o fenómeno assim que ele acontece. Só entra em ação quando a prevenção falha, mas na prática deve ser conduzida continuamente uma vez que por norma não existe a perceção dessa mesma falha (Bolton et al., 2002).

A deteção da fraude é uma área em evolução contínua. Sempre que se descobre que um método de deteção foi implementado, os indivíduos fraudulentos adaptam a sua estratégia para o contornar. E as instituições quando descobrem que os criminosos utilizaram um novo método de fraude, rapidamente instalam métodos de deteção apropriados. É um ciclo interminável que implica que os métodos de deteção antigos devam ser aplicados em conjunto com os mais recentes (Bolton et al., 2002).

Um outro fenómeno associado à fraude e que também atormenta as sociedades de hoje é a chamada lavagem de dinheiro. Indivíduos que se encontram na posse de fundos obtidos a

partir de atividades criminosas ou fraudulentas tentam por norma branquear os capitais para que estes possam mais tarde ser utilizados legalmente. Com o desenvolvimento da economia global, o aumento das aplicações de tecnologias *web* e os avanços de negócios altamente ligados às novas tecnologias (inclusive na área financeira), é expectável que os crimes associados à lavagem de dinheiro se tornem cada vez mais comuns, mais difíceis de investigar e mais prejudiciais ao desenvolvimento da economia e estabilização dos sistemas financeiros (Zhang et al., 2003).

Uma vez que a lavagem de dinheiro pode envolver uma grande variedade de técnicas e um número elevado de transações, é difícil para as autoridades detetar este fenómeno e indiciar os infratores. As técnicas utilizadas por estes também estão em constante alteração, o que torna esta matéria penosa de estudar. Adicionalmente, os custos para implementar controlos de prevenção e deteção são elevados e nem sempre eficazes, isto porque é necessária muita mão-de-obra para processar uma quantidade exagerada de dados, que têm de ser analisados minuciosamente. Estes processos manuais são morosos, lentos e insuficientes pois os indivíduos fraudulentos conseguem rapidamente adaptar-se aos mesmos e forçar as autoridades responsáveis a rever constantemente os processos implementados (Axelsson & Lopez-Rojas, 2012). Consequentemente, o ideal seria automatizar estes processos o máximo possível, reduzindo significativamente a mão-de-obra necessária e o tempo despendido pela mesma. Isto leva-nos à necessidade de recorrer a métodos de *Machine Learning*, pois é possível aplicar algoritmos que identificam automaticamente novos métodos de fraude ao detetar transações que são diferentes das consideradas benignas (transações *standard* que não incorporam intenções criminosas).

Muitos problemas de deteção de fraude envolvem grandes conjuntos de dados. Na indústria bancária por exemplo, é comum as bases de dados conterem informações relativas a centenas de milhões de transações (Hand et al., 2000). Ora, processar esta quantidade de dados para encontrar transações fraudulentas requer mais do que um mero modelo estatístico, necessita de algoritmos rápidos e eficientes.

Até há relativamente pouco tempo, os estudos da Inteligência Artificial eram vistos como essencialmente teóricos e com aplicações apenas em problemas específicos e de pouco valor prático (Gama et al, 2015). No entanto, com o desenvolvimento tecnológico, a crescente complexidade dos problemas e o aumento do volume de dados disponíveis, veio uma maior necessidade e capacidade de desenvolver ferramentas computacionais sofisticadas e autónomas. Nos tempos que correm, este tipo de ferramentas pode ter um vasto leque de aplicações práticas.

Podem-se utilizar técnicas que sejam capazes de criar de forma autónoma uma hipótese ou função, com base em dados históricos, que represente uma solução para problemas em análise (Gama et al, 2015).

Este processo de Extração de Conhecimento de Dados (*Data Mining*) tem tido um impacto crescente no seio empresarial e na forma como as organizações conduzem os seus negócios. Devido à facilidade em recolher dados via *web*, as empresas começaram a perceber que existe uma grande quantidade de informação nesses dados que pode ter um papel fulcral no apoio a várias decisões organizacionais. Por exemplo, a proliferação das tecnologias da informação tem transformado o modo como as empresas orientam as suas atividades de marketing e como gerem a informação dos seus clientes (Shaw et al., 2001). A disponibilidade de elevados volumes de dados acerca dos clientes tem criado oportunidades, bem como desafios, para que as empresas utilizem esses dados de modo a criar uma vantagem competitiva. O conhecimento extraído permite estabelecer uma relação mais próxima com os clientes, satisfazendo melhor as suas necessidades e exigências (Baesens et al., 2002).

Existem muitas outras aplicações onde métodos de extração de conhecimento de dados têm sido utilizados de forma a auxiliar a gestão das empresas, como a previsão de falências (Olson et al., 2012), o reconhecimento de padrões de compra (Baesens et al., 2002), a identificação de tendências nos mercados financeiros (Enke & Thawornwong, 2005), ou a gestão de redes elétricas inteligentes (Hernandez et al., 2013). Uma outra aplicação comum dos métodos de *data mining* é a deteção de fraude (Bolton et al., 2002; Dheepa & Dhanapal, 2009; Phua et al., 2010). Por norma, os métodos estatísticos não confirmam por si só a presença de fraude, apenas alertam para o facto de que uma observação é considerada anómala, ou mais provável de ser fraudulenta que outras (Bolton et al., 2002). O objetivo da análise de dados passa por devolver uma avaliação do nível de suspeita. Estas avaliações podem ser determinadas para cada registo na base de dados e devem ser atualizadas ao longo do tempo.

A motivação para este trabalho partiu do interesse em conjugar métodos de apoio à decisão, nomeadamente extração de conhecimento de dados, com uma ou mais áreas da economia e gestão. Tal como mencionado, já existem vários estudos e aplicações que aliam estes métodos de análise de dados ao Marketing, à Estratégia e às Finanças por exemplo (Baesens et al., 2002; Neil et al., 2005; Shaw et al., 2001; Sun & Shenoy, 2007). Por esta razão optou-se por desenvolver uma solução um tanto semelhante às existentes, mas direcionada a um tema que não tenha sido tão exaustivamente explorado. Assim, tendo por base métodos de deteção de

fraude utilizados em diversos ambientes, tal como o financeiro, pretende-se construir um modelo de deteção de fraude aplicado a um sistema de transações financeiras, com o intuito de identificar suspeitas de branqueamento de capitais.

Pretende-se aplicar técnicas de extração de conhecimento de dados a um conjunto de transações monetárias para depois recorrer a algoritmos de classificação que identifiquem casos suspeitos de branqueamento de capitais. Devido à inacessibilidade deste tipo de dados, parte do trabalho irá também incidir no desenvolvimento de um modelo que seja capaz de simular transações entre múltiplos agentes, fidedignos e fraudulentos, de modo a gerar dados sintéticos que possam ser analisados.

O objetivo passa então por desenhar um modelo capaz de detetar estas situações com base na análise dos dados obtidos pela simulação, da mesma forma que num sistema de controlo bancário se pretendem detetar situações de fraude na utilização de cartões de crédito (Panigrahi et al., 2009; Tuyls et al., 2002). A ideia central consiste em alimentar um algoritmo de aprendizagem computacional com um conjunto de dados de treino (extraídos de dados históricos resultantes da execução do modelo de simulação). Depois de um processo de aprendizagem, o programa deve ser capaz de classificar corretamente um conjunto de dados que nunca tenha analisado como suspeito ou não, tendo em conta as características da informação analisada.

Os métodos utilizados serão as redes Bayesianas e a modelação baseada em agentes. As redes Bayesianas são conhecidas por demonstrar bons resultados na extração de conhecimento de dados quando conjugadas com outros métodos estatísticos (Heckerman, 1997), sendo capazes de codificar relações e dependências entre variáveis. Também facilitam a combinação de conhecimento prévio de um dado problema com um conjunto de dados recentemente adquiridos. Permitem inferir acerca das causas de certos eventos, gerando crenças, bem como realizar previsões na presença de dados históricos e obtidos, revendo as suas crenças ao longo do processo (Heckerman, 1997). Aplicações destas redes na deteção de fraude têm apresentado bons resultados (Panigrahi et al., 2009; Tuyls et al., 2002).

A modelação baseada em agentes tem aplicações num diverso leque de campos, desde a modelação de comportamentos dentro de organizações até à análise de propagação de epidemias (Macal & North, 2005). Fornecem uma base que permite simular e avaliar comportamentos, decisões e interações entre vários agentes. Isto é particularmente útil para qualquer área que envolva o estudo de redes, quer estas sejam sociais, computacionais, ou outras. Esta técnica

também apresenta bons resultados quando aplicada à deteção de fraude ou de intrusos nas redes (Huang et al., 2010; Mellouli et al., 2004; Prodromidis & Stolfo, 1999; Zhu et al., 2006).

Então, o modelo deverá assentar em métodos de aprendizagem Bayesiana conjugados com modelação baseada em redes de agentes. Esta última permitirá replicar a interação entre os agentes de uma rede de transações monetária, sendo possível incorporar no modelo agentes com diferentes perfis em relação à sus propensão à fraude. Serão também modelados alguns padrões de transações que estão geralmente associados a tentativas de lavagem de dinheiro. A aprendizagem Bayesiana deverá então ser capaz de identificar estes padrões na medida em que serão díspares dos das transações credíveis. Pretende-se então desenvolver um sistema que analise as transações de inúmeros agentes para que no final seja possível classificar cada agente relativamente à probabilidade de possuírem intenções fraudulentas ou não.

Um sistema de deteção de fraude neste contexto viria complementar a literatura existente, na medida em que explora uma possível abordagem a um tema cada vez mais relevante mas que ainda comporta muitas incertezas. Pretende-se que este trabalho contribua não só para o tópico da deteção de fraude e lavagem de dinheiro, mas também para a sua modelização e simulação. Embora não utilizando dados reais, explorar de que modo se pode caracterizar o comportamento dos agentes fraudulentos pode contribuir para o estudo deste fenómeno, bem como fornecer a possibilidade de simular vários cenários hipotéticos, o que deverá ser realizado com maior frequência se as entidades responsáveis quiserem estar um passo à frente dos perpetradores. O recurso às técnicas utilizadas e os respetivos resultados positivos, pretendem também demonstrar que o *Machine Learning* pode reduzir custos e melhorar a eficiência de operações de controlo e deteção da fraude.

Quanto à estrutura deste documento, no segundo capítulo será apresentada uma revisão da literatura relevante. Este irá apresentar a Fraude (2.1) e o Branqueamento de capitais (2.2), as suas características, implicações e considerações acerca do seu controlo e deteção (2.3). Serão também discutidas as técnicas de processamento de dados (Redes Bayesinas) e a Modelação Baseada em Agentes (2.4). No terceiro capítulo será feita uma descrição detalhada do problema a abordar, da metodologia adotada e dos dados em análise. Por fim, no quarto capítulo serão descritos os resultados obtidos e no quinto serão expostas as conclusões do trabalho realizado.

2. Fraude e Empresas

2.1 Visão Geral

A ACFE (*Association of Certified Fraud Examiners*) define fraude como "*The use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets."* (Acts, I. F, 2000). O impacto económico da fraude é vasto. Estimativas do custo deste tipo de crimes para empresas Norte Americanas variam entre 200 e 600 biliões de dólares (Schnatterly, 2003). A fraude pode ter um impacto significativo na performance de uma empresa, pois pode custar a uma empresa entre 1 a 6 porcento das suas vendas anuais (Hogsett III & Radig, 1994). Tendo isto em conta, a capacidade de prevenir a fraude, ou quantificar o seu risco e as respetivas perdas, tornou-se uma potencial fonte de vantagem competitiva e de melhorias na performance financeira para as empresas (Button et al, 2012; Schnatterly, 2003).

Todas as organizações estão sujeitas a riscos de fraude, sendo que esta tem sido uma das causas de diversas falências de empresas, perdas de investimento massivas, custos legais significativos e consequente erosão da confiança nas entidades respetivas e nos mercados de capitais a que pertencem. Em reação a este tipo de eventos, hoje em dia é esperado que as organizações assumam uma atitude de tolerância zero em relação ao risco de fraude. Os bons princípios de governo da empresa exigem que os conselhos de administração, ou outros corpos de supervisão, assegurem um comportamento global com elevados níveis de ética, independentemente da sua origem pública, privada, governamental ou sem fins lucrativos, da sua dimensão ou da sua indústria (Bishop et al., 2008). O órgão de supervisão é crítico uma vez que a maioria das fraudes é realizada por gestores seniores em conluio com outros colaboradores (Bishop et al., 2008). Um controlo e gestão vigilantes dos casos de fraude numa organização enviam um sinal claro ao público, *stakeholders* e reguladores acerca da atitude dos órgãos de supervisão e gestão em relação aos riscos de fraude e à tolerância que a organização apresenta perante estes (Bishop et al., 2008).

Para além do órgão de supervisão (por norma parte constituinte do Conselho Geral de uma empresa), todo o pessoal da organização tem a responsabilidade de saber lidar com os riscos de fraude. É esperado que sejam capazes de compreender e expor de que forma é que a organização corresponde às regulações exigidas, que tipo de programa de gestão de risco de

fraude é utilizado, como é que os riscos são identificados, como a fraude é prevenida ou detetada o mais cedo possível e que processos são utilizados para investigar a fraude e tomar ações corretivas.

Naturalmente, é de extrema importância que as organizações desenvolvam o seu próprio programa de gestão do risco de fraude. A única forma de estas se protegerem contra atos significativos de fraude é através de um esforço contínuo e diligente para estabelecer mecanismos e processos de controlo, prevenção, deteção, investigação, reporte e correção da fraude. Para isto, segundo a ACFE, devem ser seguidos os seguintes princípios:

- Princípio 1: Como parte da estrutura de governo da empresa, deve existir um programa de gestão do risco de fraude, incluindo uma política (ou políticas) que transmita as expectativas do conselho de administração e da gestão sénior a respeito da gestão do risco de fraude.
- Princípio 2: A exposição ao risco de fraude deve ser avaliada periodicamente pela organização para que sejam identificados potenciais esquemas ou eventos que a organização necessite de mitigar.
- Princípio 3: Técnicas de prevenção para evitar potenciais eventos de fraude devem ser estabelecidas, quando realizáveis, de modo a mitigar possíveis impactos negativos na organização.
- Princípio 4: Técnicas de deteção devem ser estabelecidas para revelar eventos de fraude quando as medidas de prevenção falham ou quando são descobertos riscos por mitigar.
- Princípio 5: Um processo de registo e reporte deve ser implementado para solicitar apreciações acerca de potenciais fraudes e deve ser utilizada uma abordagem coordenada com a investigação e retificação de modo a assegurar que as potenciais fraudes são endereçadas apropriada e atempadamente.

Para que uma empresa se consiga proteger da fraude, bem como os restantes *stakeholders,* de forma eficiente e efetiva, é imperativo perceber os riscos de fraude e os riscos específicos que podem afetar a organização direta ou indiretamente. Uma avaliação estruturada do risco de fraude, adaptada à dimensão, complexidade, industria e objetivos da empresa, deverá ser

conduzida e atualizada periodicamente. Esta avaliação pode estar integrada com a avaliação global do risco na empresa ou ser realizada independentemente. No mínimo, deve incluir a identificação de riscos, a sua probabilidade, importância e quais as respostas adequadas.

Um processo de identificação de riscos de fraude eficaz deve incluir uma avaliação dos incentivos, pressões e oportunidades para cometer fraude. A avaliação dos riscos de fraude deve considerar situações em que os mecanismos de controlo são sobrepostos pela gestão, bem como as áreas onde os controlos são débeis ou inexistentes.

A velocidade, funcionalidade e acessibilidade fornecidas pela era da informação e das tecnologias têm aumentado a exposição das empresas à fraude. Por esta razão, qualquer avaliação do risco de fraude deve considerar a sobreposição do acesso aos sistemas de controlo, as ameaças internas e externas à integridade dos dados, a segurança dos sistemas e o possível furto de informação financeira ou sensível ao negócio.

Existem várias taxonomias para classificar e organizar os riscos de fraude. A ACFE classifica os riscos de fraude ocupacional em três categorias genéricas: declarações financeiras fraudulentas, apropriação indevida de ativos e corrupção. Utilizando estas categorias como ponto de partida, é possível desenvolver um modelo mais detalhado para a avaliação do risco de fraude numa organização em específico (Bishop et al., 2008).

O reporte financeiro fraudulento, segundo a ACFE, foca-se por norma em melhorar o panorama financeiro de uma organização através de sobrevalorização dos retornos, subvalorização das perdas, ou utilizando pareceres duvidosos. Contrariamente, algumas empresas minimizam os retornos para atenuar os proveitos. Qualquer distorção intencional da informação contabilística representa um reporte financeiro fraudulento. Deve ainda ser considerado o cenário onde o objetivo da fraude não passa por melhorar as declarações financeiras da organização, mas sim cobrir um buraco nas finanças deixado por utilização ou apropriação indevida dos ativos. Neste caso, a fraude também inclui reporte financeiro fraudulento.

Relativamente aos ativos de uma organização, tanto tangíveis (numerário, stock) como intangíveis (produtos proprietários ou confidenciais, informação dos clientes), podem ser apropriados indevidamente por colaboradores, clientes, fornecedores ou indivíduos externos à empresa. A empresa tem então de assegurar que existem mecanismos de controlo para proteger os seus ativos. Para iniciar o processo de avaliação de riscos de fraude é importante considerar quais são os ativos que podem ser alvo de apropriação indevida, onde se localizam e quem tem

acesso a estes. Evitar estes riscos requer não só controlos físicos de salvaguarda, mas também controlos periódicos como contagens de inventário.

A corrupção é definida operacionalmente como a utilização indevida do poder para ganhos pessoais. Formas comuns de corrupção incluem auxílios, incentivos ou subornos, a instituições reguladoras ou do governo, com o intuito de obter ou manter fontes de receita.

2.2 Economia Paralela e Lavagem de Dinheiro

Indivíduos que cometem crimes e/ou fraudes têm por norma duas grandes preocupações: evitar o encarceramento e aproveitar os frutos dos seus crimes (Levi & Reuter, 2006). Este aproveitamento é frequentemente realizado na forma de consumo imediato e percetível. No entanto, para os mais disciplinados e para aqueles que recolhem ganhos de grande dimensão, o aproveitamento pode ter que ser adiado para que seja possível aproveitar possíveis oportunidades económicas futuras. Os lucros obtidos de forma ilegal têm de ser "convertidos" de modo a poderem ser gastos sem restrições em produtos e serviços legítimos. O disfarce ou ocultação deste tipo de fundos é hoje em dia considerada uma atividade criminosa em si mesmo, com leis e normas a serem estabelecidas com o objetivo de identificar e punir estes atos de "lavagem de dinheiro".

As técnicas utilizadas para esconder estes rendimentos podem incluir o transporte do dinheiro para fora do país, compra de negócios através dos quais o dinheiro pode ser canalizado, aquisição de ativos facilmente transferíveis, ou preços de transferência. Ou, segundo Quirk (1997):

- Recurso a múltiplos depósitos em numerário, cada um menor de que o mínimo indicado para reporte.
- Declarações erradas ou falsificadas de exportações / importações, bem como de letras de crédito ou documentos alfandegários, de modo a esconder transações ilícitas.
- Bens obtidos através de furtos podem ser trocados por substâncias ilegais.
- Transferências de crédito podem ser utilizadas para evitar o olhar da economia "formal", exceto no final quando os lucros obtidos ilegalmente são utilizados para adquirir bens ou serviços comercializados legalmente.

- Transferências interbancárias podem não ser sujeitas a análises de lavagem de dinheiro. Inclusive existe a possibilidade de suborno / conluio com profissionais da banca para ocultar ou desviar as atenções de transações ilegais entre contas.

Contas *offshore* têm sido um veículo importante na lavagem de dinheiro. Existe uma necessidade para que haja cooperação internacional e uma estrutura bem estabelecida na luta contra a lavagem de dinheiro (Quirk, 1997).

Monitorização de lavagem de dinheiro foca-se em estabelecer as identidades que realizam a transação e os padrões dessas mesmas transações (Quirk, 1997).

2.2.1 Impacto Económico da Economia Paralela

Uma vez que o crime e a lavagem de dinheiro acontecem numa grande escala, os seus efeitos também têm que ser tidos em conta de um ponto de vista macroeconómico. Mas estas atividades são difíceis de avaliar, logo os dados disponíveis podem ser distorcidos o que dificulta a tarefa dos reguladores.

Para além disso, identificar o país, a moeda e a residência dos detentores das contas é um passo crucial na compreensão do comportamento monetário. A procura de moeda pode sofrer desvios devido à lavagem de dinheiro e economia paralela, que podem gerar dados monetários enganadores. Este aspeto pode ter consequências adversas, afetando a volatilidade das taxas de câmbio e de juro (Quirk, 1997).

Os efeitos da lavagem de dinheiro na distribuição da riqueza também devem ser considerados, na medida em que a atividade criminosa subjacente redireciona rendimentos de elevadas poupanças para poupanças mais pequenas, ou de investimentos sólidos para investimentos arriscados. Por exemplo, existem evidências de que proveitos obtidos por evasão fiscal nos Estados Unidos da América tendem a ser canalizados para investimentos de maior risco e maior retorno no setor das pequenas empresas, já por si um setor propenso a evasões fiscais (Quirk, 1997).

A lavagem de dinheiro também tem um impacto mais indireto na macroeconomia. Transações de valores obtidos ilegalmente podem afetar transações legais por "contaminação". Transações que envolvam participantes estrangeiros, embora totalmente legais, podem ser menos desejáveis por estarem associadas a lavagem de dinheiro. Globalmente, a confiança nos

mercados e no papel dos lucros como sinalizadores de eficiência acabam desgastadas por atividades como fraude, roubos e negócios com informações privilegiada. E dinheiro que é lavado por outras razões que não a evasão fiscal, acaba por ter esse efeito de qualquer forma, acumulando distorções económicas. Para além disto, ignorar as leis é contagiante, quebrando uma torna-se mais fácil fazê-lo com outras. Assim os balanços acumulados de ativos lavados podem atingir valores superiores a fluxos de rendimento anuais, aumentando o potencial de movimentos económicos desestabilizadores e ineficientes, quer sejam através de fronteiras ou domésticos (Quirk, 1997).

Os efeitos mencionados são até certo ponto especulativos. No entanto existem estudos que demonstram a relação entre a diminuição do crescimento do PIB com o crescimento das estimativas dos resultados de lavagem de dinheiro (Quirk, 1997).

2.2.2 Controlo e Supervisão

Os controlos contra o branqueamento de capitais começaram a ser desenvolvidos nos anos setenta nos EUA com o objetivo de combater a utilização de bancos internacionais para evasões fiscais (Levi & Reuter, 2006). Inicialmente estes controlos acabaram por ter um contributo importante na luta contra o tráfico de drogas, sendo que mais recentemente podem ser parte integrante de operações apontadas a uma grande variedade de delitos. Desde contrabando, a corrupção de funcionários com posições hierárquicas elevadas, ou até mesmo o financiamento de operações terroristas.

O leque de instituições envolvidas em operações anti lavagem de dinheiro pode ser impressionante. Nos EUA, o regime de controlo expandiu-se para além dos bancos para uma vasta extensão de negócios, como concessionários automóveis, casinos, joalharias, lojas de penhor e algumas companhias de seguros. A todos estes é requerido que participem ativamente no controlo deste crime, através de relatórios de transações suspeitas. No Reino Unido, qualquer empresa que negoceie produtos ou serviços de elevado valor tem também de reportar este tipo de suspeitas. No Canadá é pretendido que os contabilistas reportem quaisquer suspeitas de lavagem de dinheiro por parte dos seus clientes (Levi & Reuter, 2006).

Estes controlos pretendem-se cada vez mais globais, com o Fundo Monetário Internacional (FMI) e o Banco Mundial a desempenharam um papel ativo no desenvolvimento de políticas estruturais que ajudem a combater estas situações. Em conjunto com outros órgãos

regionais, estas instituições monitorizam o desempenho de cada nação principalmente em relação ao cumprimento das formalidades exigidas, mas também com uma preocupação crescente nos resultados destes controlos (Levi & Gilmore, 2002). Assim, nos dias de hoje existem centenas de organizações nacionais de Inteligência Financeira capazes de receber, analisar e processar relatórios de instituições reguladas.

No entanto, os efeitos destes sistemas nos métodos de lavagem e nos custos, bem como na disposição dos delatores em empreender atividades criminosas, são difíceis de avaliar. Os dados disponíveis sugerem debilmente que controlos contra a lavagem de dinheiro acabam por não ter um impacto significativo na diminuição dos crimes. Os controlos impostos acabam por facilitar a investigação e a acusação de alguns casos, mas acabam por ser detetados menos do que o esperado seguindo métodos que "seguem o dinheiro" (Levi & Reuter, 2006). Alguns dos controlos também almejam o financiamento terrorista, mas nos dias que correm várias operações desta índole podem ser concretizadas com custos reduzidos. Por esta razão os controlos podem não conseguir cortar o mal pela raiz, mas podem oferecer dados importantes.

É importante ter em conta que os controlos de lavagem de dinheiro impõem custos elevados às empresas e à sociedade, e por isso merecem uma análise cuidada dos efeitos.

2.3 Prevenção e Deteção da Fraude

A prevenção e a deteção da fraude estão relacionados, mas não são conceitos semelhantes. A prevenção engloba políticas, procedimentos, treino e comunicação que visam impedir que a fraude ocorra. Já a deteção foca-se nas atividades e técnicas que reconhecem de forma pronta e atempada se a fraude ocorreu ou está a decorrer. Enquanto as técnicas de prevenção não garantem que a fraude não venha a decorrer, acabam por ser a primeira linha de defesa para minimizar o risco de fraude.

Um dos maiores dissuasores da fraude é consciencialização de que estão implementados métodos de deteção eficazes. Quando combinados com controlos preventivos, a deteção pode elevar a eficácia do programa de gestão do risco de fraude, demonstrando que os controlos preventivos estão a funcionar como previsto e identificando situações de fraude quando esta ocorre. Embora os mecanismos de deteção de fraude possam dar evidências que a fraude tenha ocorrido ou está a decorrer, estes não têm o propósito de a prevenir.

Todas as organizações são suscetíveis à fraude, mas nem toda a fraude pode ser prevenida e pode nem ser economicamente viável tentar fazê-lo. Uma organização pode determinar que é preferível desenhar os seus mecanismos de deteção, em vez de tentar prevenir certos esquemas de fraude. É importante que as organizações considerem tanto a prevenção como a deteção da fraude (Bishop et al., 2008).

A deteção deste fenómeno envolve identificar atividades fraudulentas o mais depressa possível assim que forem cometidas. Métodos para esta deteção são desenvolvidos continuamente de modo a se adaptarem às estratégias dos infratores. No entanto o desenvolvimento de novos métodos enfrenta algumas dificuldades devido às limitações de partilha de informações e de ideias relacionadas com o assunto. Mesmo assim, hoje em dia já existem vários métodos de deteção de fraude implementados que recorrem a extração de conhecimento de dados, análise estatística e inteligência artificial. Estes tentam descobrir atos fraudulentos com base em anomalias e padrões em bases de dados (Bolton et al., 2002).

Se a deteção da fraude falhar, a deteção de casos de lavagem de dinheiro pode ser uma ajuda importante na identificação de ações suspeitas que, depois de analisadas detalhadamente, podem levar aos indivíduos na sua origem. Ou seja, a deteção de casos de suspeita de lavagem de dinheiro não previne que a fraude aconteça de forma direta, mas pode fazê-lo indiretamente. Ao identificar os possíveis suspeitos na origem do branqueamento de capitais, dá a oportunidade de estes serem investigados e que as respetivas atividades fraudulentas, através das quais os lucros foram ilegalmente obtidos, sejam alvo de uma examinação minuciosa. Este estudo retrospetivo pode criar informação relevante e mais tarde fomentar a criação de controlos para a prevenção e deteção de fraude mais eficazes.

Como referido, a lavagem de dinheiro prejudica as finanças de uma nação e pode contribuir para um aumento do financiamento de atividades criminais (Bartlett, 2002). Devido à grande quantidade de transações e variedade de técnicas de lavagem de dinheiro, é difícil para as entidades responsáveis detetarem este fenómeno.

A lavagem de dinheiro ganha forma num fluxo complexo que se inicia com a alocação de fundos obtidos ilegalmente, seguido de uma série de operações encadeadas de modo a esconder as suas origens até que finalmente os fundos possam ser utilizados em atividades formais e legais (Buchanan, 2004). Devido a questões tais como a elevada quantidade de transações que são realizadas diariamente nos serviços financeiros, identificar transações especificas para serem caracterizadas como suspeitas não é uma tarefa fácil. A atividade suspeita

precisa de ser suportada por fundamentos e indícios tangíveis que permitam a agências governamentais investigar com maior detalhe (Axelsson & Lopez-Rojas, 2012).

Em vários países a maioria das empresas do setor financeiro são obrigadas por lei a implementar métodos de deteção de lavagem de dinheiro. No entanto os custos de implementação desses métodos são demasiado elevados, sobretudo devido à quantidade de tarefas manuais que normalmente implicam (Magnusson, 2009). O método mais utilizado nos dias que correm para prevenir transações financeiras ilegais consiste em identificar diferentes indivíduos de acordo com o risco estimado e restringindo as suas transações através de limites (Bolton et al., 2002). Transações que ultrapassem estes limites devem despoletar uma análise profunda, exigindo aos indivíduos responsáveis que esclareçam a origem dos fundos. Estes limites são usualmente definidos por lei, sem que haja uma distinção entre diferentes setores ou atores económicos. Naturalmente, quem pratica a fraude está em constante adaptação de modo a contornar estes controlos, como executar transações de valores menores do que os limites. Deste modo, este e outros métodos similares demonstraram-se insuficientes (Magnusson, 2009).

Várias técnicas de *Machine Learning* têm sido utilizadas para a deteção de fraude, bem como para a lavagem de dinheiro (Sudjianto et al., 2010). No entanto até 2008 a maioria das aplicações focavam-se em fraudes nos seguros, em empresas e em cartões de crédito (Ngai et al., 2011), onde as técnicas de extração de conhecimento de dados mais utilizadas são redes neuronais, redes de crenças Bayesianas e árvores de decisão, sendo estas capazes de fornecer soluções aos problemas inerentes à deteção e classificação de dados fraudulentos. O uso destas técnicas para analise deste fenómeno é vantajoso devido à boa performance da classificação (elevado número de verdadeiros positivos e baixo de falsos positivos) quando comparada com métodos simples que recorrem a limites (Yue, et al., 2007; Zhang et al., 2003). O *Data Mining* (que usa algoritmos de *Machine Learning*) é frequentemente utilizado na deteção de fraude porque os algoritmos são capazes de identificar novos métodos de fraude através da deteção de transações anómalas aos padrões usuais. Algoritmos de aprendizagem supervisionada já comprovaram a sua capacidade de deteção de *outliers* em conjuntos de dados sintéticos (Abe, Zadrozny, & Langford, 2006).

2.4 Algoritmos para a Deteção da Fraude

A classificação é uma das tarefas de tomada de decisão mais frequentemente encontrada na atividade humana. Um problema de classificação ocorre quando um objeto necessita de ser assignado a um determinado grupo ou classe com base num certo número de atributos observados em relação a esse objeto (Zhang, 2000). Vários problemas em gestão, economia, industria e medicina podem ser tratados como problemas de classificação. Podemos ter exemplos como previsão de falências, avaliação de créditos, diagnósticos clínicos, controlo de qualidade, reconhecimento de textos ou discursos.

Neste trabalho serão utilizados algoritmos de classificação, nomeadamente com base em Redes Bayesianas, e adicionalmente modelação baseada em agentes para simular o comportamento de indivíduos de uma rede financeira e as transações realizadas pelos mesmos. Esta secção pretende apresentar cada um destes métodos e exemplificar de que modo é que podem ser indicados para o estudo e deteção de fraudes como o branqueamento de capitais.

2.4.1 Redes Bayesianas

Uma rede Bayesiana é um modelo gráfico para relações probabilísticas entre um conjunto de variáveis. Quando são conjugadas com outros métodos estatísticos Bayesianos, estas redes podem demonstrar bons resultados na extração de conhecimento de dados (Heckerman, 1997). A sua utilização pode ser vantajosa em vários aspetos, como por exemplo na capacidade em lidar com conjuntos de dados incompletos, uma vez que o modelo consegue codificar as relações e dependências entre as variáveis. A aprendizagem das relações causais é outro aspeto positivo das redes Bayesianas, uma vez que permite tirar inferências sobre a razão de certos eventos ou realizar previsões na presença de dados adquiridos. Uma outra vantagem é o facto de que combinar estas redes com métodos estatísticos Bayesianos facilita a combinação do conhecimento prévio do domínio do problema com o conjunto de dados em análise. Ou seja, uma vez que o modelo tem semânticas causais e probabilísticas, é ideal para combinar conhecimento prévio com dados obtidos.

Para compreender as redes Bayesianas e as técnicas de extração de conhecimento de dados associadas, é importante compreender a abordagem Bayesiana às probabilidades e

estatística. Muito resumidamente, a probabilidade Bayesiana de um evento é o grau de crença que se tem nesse evento. Enquanto a probabilidade clássica é uma propriedade física do mundo (resultado do lançamento da moeda), uma probabilidade Bayesiana é uma propriedade da pessoa que atribui a probabilidade em si (a crença de que o lançamento da moeda vai resultar em coroa) (Heckerman, 1997).

2.4.1.1 Classificador Naive Bayes

A classificação é uma tarefa básica da análise de dados e identificação de padrões supervisionada que requer construir um classificador, ou seja, uma função que atribua uma classe a certas instâncias descritas por um conjunto de atributos. Esta indução que os classificadores realizam partindo de dados pré-classificados é um problema central do *Machine Learning*. Várias abordagens a este problema são baseadas em representações funcionais tais como árvores de decisão, redes neuronais, grafos e regras de decisão.

Um dos classificadores mais eficientes é designado por classificador *naive Bayes*. Este aprende a probabilidade condicionada de cada atributo A_i, dada a classe C, a partir de um conjunto de dados de treino (Friedman et al., 1997). A classificação é então realizada aplicando a regra de *Bayes* de forma a calcular a probabilidade de C dado uma instância particular de $A_1, \ldots., A_n$, conseguindo depois prever a classe com a maior probabilidade *a posteriori*. Esta determinação é possível devido a uma forte assunção quanto à independência probabilística dos atributos: todos os atributos de A_i são condicionalmente independentes dado o valor da classe C.

O classificador *naive Bayes* aplica-se a tarefas de aprendizagem onde cada instância x é descrita por uma conjunção de atributos e a função objetivo $f(x)$ pode tomar qualquer valor do conjunto finito V. Por norma é fornecido um conjunto de exemplos de treino da função objetivo e é apresentada uma nova instância, descrita pelo conjunto de variáveis $[a_1, a_2, \ldots a_n]$. O algoritmo de aprendizagem é então solicitado para realizar uma previsão do valor objetivo, uma classificação, para esta nova instância (Mitchell, 1997).

A abordagem Bayesiana para classificar novas instâncias passa por atribuir o valor objetivo mais provável, v_{MAP}, dado os atributos $[a_1, a_2, \ldots a_n]$ que descrevem a instância (MAP do inglês *Maximum a Posteriori*).

$$v_{MAP} = arg \max_{v_j \in V} P(v_j | a_1, a_2, \dots a_n)$$

Segundo o teorema de *Bayes*, a probabilidade *a posteriori* $P(h|D)$, de h sabendo D é

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

Então é possível reescrever a expressão de v_{MAP} da seguinte forma

$$\begin{aligned} v_{MAP} &= arg \max_{v_j \in V} \frac{P(a_1, a_2, \dots\ a_n | v_j) P(v_j)}{P(a_1, a_2, \dots\ a_n)} \\ &= arg \max_{v_j \in V} P(\ a_1, a_2, \dots a_n | v_j) P(v_j) \end{aligned} \quad (2.1)$$

Partindo da equação (2.1) é possível estimar os seus dois termos com base em dados de treino. Para estimar cada $P(v_j)$ basta apurar a frequência com que cada valor objetivo v_j ocorre nos dados de treino. Já estimar os diferentes termos de $P(a_1, a_2, \dots a_n | v_j)$ desta forma só é praticável se as dimensões dos dados de treino forem consideráveis. A questão é o número de termos em análise é igual ao número de instâncias possíveis vezes o número de valores objetivos possíveis. Portanto é necessário verificar cada instância no espaço de dados um número elevado de vezes de forma a obter estimativas fiéis.

O classificador *naive Bayes* baseia-se na assunção de que os valores dos atributos são independentes condicionalmente dado o valor da variável objetivo. Ou seja, dado o valor objetivo da instância, a probabilidade de observar a conjunção $a_1, a_2, \dots a_n$ é o produto das probabilidades dos atributos individuais: $P(a_1, a_2, \dots a_n | v_j) = \prod_i P(a_i | v_j)$. Substituindo esta expressão na equação (2.1) obtém-se a abordagem utilizada pelo classificador *naive Bayes* (Mitchell, 1997):

$$v_{NB} = arg \max_{v_j \in V} P(v_j) \prod_i P(a_i | v_j) \quad (2.2)$$

Onde v_{NB} é o resultado obtido do valor objetivo através do classificador *naive Bayes*. Aqui convém salientar que neste classificador o número de termos $P(v_j)$ distintos a estimar é apenas

o número de atributos vezes o número de valores objetivo, o que resulta num número muito menor do que se estimassem os termos $P(a_1, a_2, \ldots a_n|v_j)$ contemplados anteriormente.

Em suma, a aprendizagem *naive Bayes* envolve um passo onde os vários $P(v_j)$ e $P(a_i|v_j)$ são estimados, com base na sua frequência nos dados de treino. O conjunto destas estimativas corresponde à hipótese aprendida. Esta hipótese é depois utilizada para classificar uma nova instância aplicando a equação (2.2). Quando assunção de independência condicional é satisfeita, esta classificação v_{NB} é idêntica à classificação MAP (*Maximum a Posteriori*).

A performance do *naive Bayes* chega a ser um pouco surpreendente, uma vez que a assunção quanto à independência condicional dos atributos é irrealista. Considerando, por exemplo, um classificador para avaliar o risco em pedidos de crédito, parece contraintuitivo ignorar as correlações entre idade, nível de educação e rendimentos. Isto levanta a seguinte questão: é possível melhorar a performance destes classificadores evitando suposições injustificadas a cerca da independência?

2.4.1.2 Redes de Crenças Bayesianas

Como já foi discutido, o classificador *naive Bayes* baseia-se significativamente na suposição de que atributos $A_1, \ldots, A_n$ são condicionalmente independentes dado um valor objetivo. Esta assunção reduz drasticamente a complexidade da aprendizagem da função objetivo. Quando a assunção é cumprida, o classificador produz uma classificação de *Bayes* ótima. Contudo, em vários casos esta suposição de independência condicional é excessivamente restritiva (Mitchell, 1997).

Uma rede de crenças Bayesiana descreve a distribuição de probabilidade que guia um conjunto de variáveis, especificando um grupo de independências condicionais juntamente com as probabilidades condicionadas. Contrastando com o classificador *naive Bayes*, que assume que todas as variáveis são condicionalmente independentes dado um determinado valor da variável objetivo, as redes de crenças Bayesianas permitem definir assunções de independência condicional que se aplicam a subconjuntos das variáveis. Assim, estas redes proporcionam uma abordagem intermédia que se verifica menos restritiva do que a assunção global de independência condicional realizada pelo classificador *naive Bayes*, mas ao mesmo tempo mais flexível do que se evitassem de todo fazer assunções desse tipo (Mitchell, 1997).

As redes Bayesianas são gráficos acíclicos orientados que permitem uma representação efetiva e eficiente da distribuição da probabilidade conjunta sobre um conjunto aleatório de variáveis. Cada vértice no gráfico representa uma variável aleatória e as arestas representam correlações diretas entre variáveis. Mais precisamente, a rede codifica cada variável como independente das variáveis que não são suas descendentes. Estas independências são depois exploradas de modo a reduzir o número de parâmetros necessários para caracterizar a distribuição de probabilidade e para calcular eficientemente probabilidades posteriores com os dados obtidos. Parâmetros probabilísticos são codificados num conjunto de tabelas, uma para cada variável, na forma de distribuições locais condicionadas de uma variável dada os seus ascendentes. Com as independências codificadas na rede, a distribuição conjunta é unicamente determinada por estas distribuições locais condicionadas (Friedman et al., 1997).

Isto é uma forma de aprendizagem sem supervisão, na medida em que o algoritmo não distingue a variável de classe das variáveis de atributos contidas no conjunto de dados. O objetivo é produzir uma rede, ou um conjunto de redes, que descrevam da melhor forma a distribuição de probabilidade sobre os dados de treino. Este processo de otimização é implementado na prática utilizando técnicas de procura com heurísticas que sejam capazes de encontrar a melhor candidata num espaço de redes possíveis. O processo de procura baseia-se numa função de pontuação que avalia os méritos de cada rede candidata.

Em resumo, uma rede de crenças Bayesiana descreve a distribuição de probabilidade de um conjunto de variáveis. Considerando um conjunto arbitrário de variáveis aleatórias $Y_1 \dots Y_n$, onde cada variável Y_i pode tomar um dos valores possíveis de $V(Y_i)$. Define-se o espaço conjunto das variáveis Y como o produto cruzado $V(Y_1) \times V(Y_2) \times \dots V(Y_n)$. Ou seja, cada item no espaço conjunto corresponde a uma das possíveis atribuições de valores à cadeia de identificação das variáveis $[Y_1 \dots Y_n]$. A distribuição de probabilidade conjunta especifica a probabilidade para cada ligação de variáveis possível em $[Y_1 \dots Y_n]$. Uma rede de crenças Bayesiana descreve a distribuição de probabilidade conjunta para um conjunto de variáveis (Mitchell, 1997).

2.4.1.3 Independência Condicional

Para aprofundar mais o tema das redes de crenças Bayesianas é necessário primeiramente definir a noção de independência condicional. Se $X, Y\ e\ Z$ forem três variáveis aleatórias discretas, pode-se dizer que X é condicionalmente independente de Y dado Z se a distribuição de probabilidade de X é independente do valor de Y dado um valor para Z:

$$(\forall x_i, y_j, z_k)\, P(X = x_i \,|Y = y_j, Z = z_k) = P(X = x_i \,|Z = z_k)$$

Onde $x_i \in V(X), y_j \in V(Y), e\ z_k \in V(Z)$. A expressão acima é usualmente abreviada para $P(X|Y,Z) = P(X|Z)$. Esta definição de independência condicional pode ser alargada a conjuntos de variáveis. Um conjunto de variáveis $X_1 \ldots X_l$ é condicionalmente independente do conjunto de variáveis $Y_1 \ldots Y_m$ dado o conjunto $Z_1 \ldots Z_n$ se

$$P(X_1 \ldots X_l \,|Y_1 \ldots Y_m, Z_1 \ldots Z_n) = P(X_1 \ldots X_l \,|Z_1 \ldots Z_n)$$

De notar a correspondência entre esta definição e a utilização da independência condicional na definição do classificador *naive Bayes*. Este assume que o atributo A_1 é condicionalmente independente do atributo A_2 dado o valor objetivo V. Isto permite ao classificador calcular $P(A_1, A_2|V)$ na equação (2.2) da seguinte forma:

$$P(A_1, A_2|V) = P(A_1|A_2, V)P(A_2|V) \tag{2.3}$$

$$= P(A_1|V)P(A_2|V) \tag{2.4}$$

Chega-se à equação (2.4) uma vez que se A_1 é condicionalmente independente de A_2 dado V, então pela definição apresentada de independência condicional $P(A_1|A_2, V) = P(A_1|V)$.

2.4.1.4 Representação de Redes Bayesianas

Uma rede de crenças Bayesiana representa a distribuição de probabilidade conjunta para um conjunto de variáveis. Na figura 2.1 por exemplo, a rede Bayesiana representa a distribuição de probabilidade conjunto para as variáveis booleanas *Storm, Lightning, Thunder, ForestFire, Campfire e BusTourGroup*. No geral a rede representa a distribuição de probabilidade ao especificar um conjunto de assunções quanto à independência condicional das variáveis (através de gráfico acíclico direto), em conjunto com as probabilidades condicionadas locais. Cada variável no espaço conjunto é representada por um nodo na rede, sendo que para cada variável são especificados dois tipos de informação. Primeiro, os arcos da rede representam a asserção de que a variável é condicionalmente independente das suas não descendentes na rede, dados os seus predecessores imediatos. Diz-se que X é um descendente de Y se existe um caminho direto de Y para X. Segundo, é fornecida uma tabela para a probabilidade condicionada de cada variável, descrevendo a

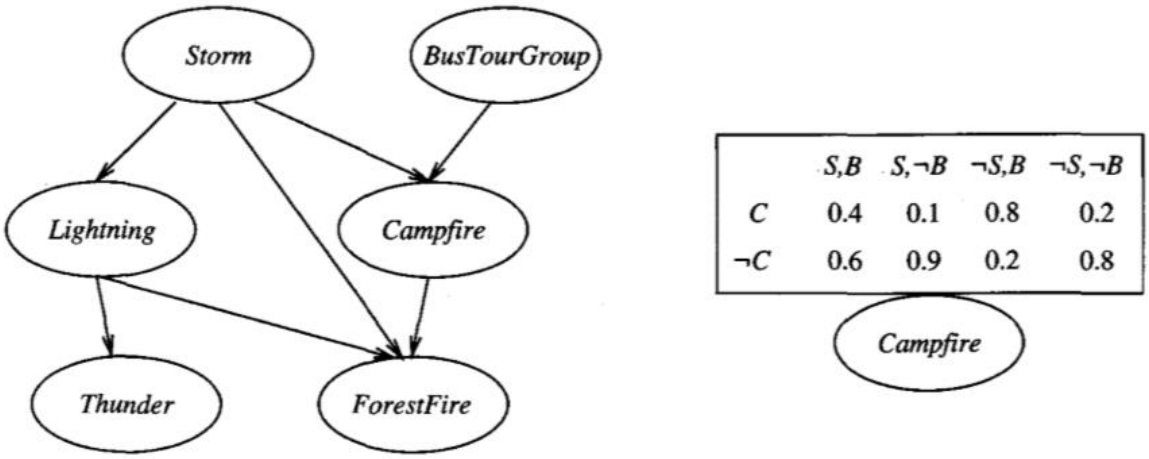

Figura 2.1 Uma rede de crenças Bayesiana (Mitchell, 1997).

distribuição de probabilidade para essa variável dado os valores dos seus predecessores imediatos. A probabilidade conjunta para qualquer atribuição desejada dos valores $\langle y_1 \dots, y_n \rangle$ para o conjunto de variáveis da rede $\langle Y_1 \dots, Y_n \rangle$ pode ser calculada a partir da fórmula

$$P(y_1 \dots, y_n) = \prod_{i=1}^{n} P(y_i | \, Pais(Y_i))$$

Onde $Pais(Y_i)$ simboliza o conjunto de predecessores imediatos de Y_i. Os valores de $P(y_i|Pais(Y_i))$ são precisamente os valores guardados na tabela de probabilidades condicionadas associada ao nodo Y_i.

Para efeitos ilustrativos, considere-se o nodo *Campfire* da rede representada na figura 2.1. Os nodos e os arcos da rede representam a asserção de que *Campfire* é condicionalmente independente dos seus não descendentes *Lightning* e *Thunder*, dado os seus pais imediatos *Storm* e *BusTourGroup*. Isto significa que uma vez conhecido o valor das variáveis *Storm* e *BusTourGroup*, as variáveis *Lightning* e *Thunder* não fornecem qualquer informação adicional acerca de *Campfire*. Na figura está presente a tabela de probabilidades condicionadas para a variável *Campfire*. O valor no canto superior esquerdo da tabela, por exemplo, expressa a seguinte asserção:

$$P(Campfire = Verdadeiro|Storm = Verdadeiro, BusTourGroup = Verdadeiro) = 0.4$$

Importa ressalvar que esta tabela representa apenas as probabilidades condicionadas de *Campfire* dadas as variáveis suas ascendentes, *Storm* e *BusTourGroup*. O conjunto de tabelas das probabilidades condicionadas locais para todas as variáveis mais as assunções de independência condicional descritas pela rede, descrevem na totalidade a distribuição de probabilidade conjunta da rede (Mitchell, 1997).

Um aspeto atrativo destas redes é que permitem representar de forma convenirente conhecimento causal, como o facto de *Lightning* causar *Thunder*. Na terminologia de independência condicional, esta relação causal é expressa pela afirmação de que *Thunder* é condicionalmente independente de outras variáveis na rede, dado o valor de *Lightning*. Esta condição está implícita nos arcos da rede.

2.4.1.5 Inferência

Uma rede Bayesiana pode ser utilizada para inferir o valor de determinada variável objetivo (*e.g.*, *ForestFire*), uma vez conhecidos os valores observados das outras variáveis. Naturalmente, se se está a lidar com variáveis aleatórias não será globalmente correto atribuir à

variável objetivo um único valor determinado. O que se pretende realmente inferir é a distribuição de probabilidade para a variável objetivo, que especifica a probabilidade que terá cada um dos seus possíveis valores dado os valores observados das outras variáveis. Este passo pode ser direto se os valores de todas as outras variáveis da rede forem conhecidos. Mas num caso mais geral o usual é tentar inferir a distribuição de probabilidade para uma variável (*e.g.*, *ForestFire*) dado os valores observados para apenas um subconjunto das outras variáveis da rede (*e.g.*, *Thunder* e *BusTourGroup* podem ser os únicos valores observados). No geral, uma rede Bayesiana pode ser utilizada para determinar a distribuição de probabilidade para qualquer subconjunto de uma rede de variáveis, dados os valores ou distribuições para qualquer subconjunto das restantes variáveis (Mitchell, 1997).

Existem vários métodos propostos para lidar com a inferência probabilística em redes Bayesianas, incluindo métodos exatos e aproximados, que sacrificam precisão para ganhar eficiência. Por exemplo, modelos de Monte Carlo fornecem soluções aproximadas recorrendo a amostragens aleatórias das distribuições das variáveis não observáveis (Pradhan & Dagum, 1996). Na teoria, tanto os métodos exatos como os aproximados podem ser *NP-hard*, mas felizmente na prática os métodos aproximados têm demonstrado a sua utilidade em várias situações (Mitchell, 1997).

2.4.1.6 Aprendizagem

Existem vários estudos que tentam encontrar uma solução, ou um algoritmo, que consiga aprender eficientemente redes de crenças Bayesianas a partir de dados de treino. Para analisar este problema é necessário primeiramente considerar alguns aspetos. A estrutura da rede pode ser conhecida previamente, ou então pode ser necessário deduzi-la através dos dados de treino. Em alguns casos todas as variáveis podem ser diretamente observáveis em cada exemplo de treino, enquanto noutros casos algumas podem não o ser.

Numa situação em que a estrutura da rede é conhecida previamente e onde as variáveis são totalmente observáveis nos exemplos de treino, deduzir as probabilidades condicionadas é direto. Basta estimar as entradas na tabela de probabilidades condicionadas como para um classificador *naive Bayes*. Se apenas algumas variáveis são observáveis nos dados de treino, o problema de aprendizagem complica-se. Torna-se um problema de alguma forma semelhante a aprender os pesos para as unidades ocultas numa rede neuronal artificial, onde os valores dos

nodos de entrada e saída são dados mas os valores das unidades ocultas não são especificados pelos exemplos de treino. Russell *et al* (1995) propuseram um procedimento semelhante, que procura pelas tabelas de probabilidades condicionadas num espaço de hipóteses que corresponde ao conjunto de todas as entradas possíveis. A função objetivo que é maximizada é a probabilidade $P(D|h)$, dos dados observados *D* sabendo a hipótese *h*. Por definição, isto corresponde a procurar pela hipótese de máxima verossimilhança para as entradas da tabela (Mitchell, 1997).

2.4.1.7 Aplicações das Redes Bayesianas

Em problemas reais de extração de conhecimento de dados, o objetivo passa normalmente por procurar relações entre um conjunto vasto de variáveis. A rede Bayesiana é uma representação adequada para esta tarefa. É um modelo gráfico que codifica eficientemente a distribuição da probabilidade condicionada para um grande conjunto de variáveis. Isto acaba por ser útil na abordagem a diversos problemas, existindo vários exemplos na literatura em que as redes Bayesianas são utilizadas em problemas reais. Por exemplo, Baesens *et al* (2002) utilizam redes Bayesianas para abordar a questão da retenção de clientes. O estudo foca-se no envio de *emails* a clientes existentes ou potenciais, com informação acerca de promoções e novos produtos. A questão aqui prende-se na escolha dos destinatários desses *emails*. A teoria de Bayes é aqui utilizada para analisar os dados dos clientes e das suas transações, podendo assim fazer assunções quanto à probabilidade de frequência de compras.

Neil *et al* (2005) investigaram a possibilidade de modelar distribuições de perdas estatísticas em cenários de riscos operacionais financeiros. O foco prende-se com a modelação de eventos de perda, quer inesperados quer de cauda longa, utilizando uma combinação da frequência das perdas e da distribuição da sua severidade. Estas combinações foram condicionadas por variáveis causais que modelam a capacidade do processo de controlo subjacente. Esta modelação causal pretende explorar a experiência local acerca da fiabilidade dos processos de controlo, conectando-a depois com fenómenos estatísticos resultantes do processo. Isto permite apoiar dados dispersos com juízos especializados e assim transformar conhecimento qualitativo em previsões quantitativas.

Sun e Shenoy (2007) utilizaram redes Bayesianas para prever falências. É realizada uma seleção heurística de indicadores financeiros para a previsão e com base nas correlações totais

ou parciais entre variáveis, o método procura eliminar informação redundante. Recorrendo ao classificador *naive Bayes* é então combinada uma assunção inicial quanto à falência com os dados em análise, de modo a determinar uma previsão final.

Ezawa *et al* (1996) elaboraram um método orientado por objetivos para aplicar as redes Bayesianas à gestão do risco nas telecomunicações. Os autores defendem que não basta um modelo se adaptar convenientemente aos dados em análise, sendo também essencial que o modelo deva ser treinado e aprendido tendo por base um objetivo ou aplicação específicos. O objetivo do modelo passa por classificar os clientes quanto ao seu nível de risco, uma vez que nesta indústria é comum verificarem-se perdas avolumadas devido a dívidas por saldar pelos clientes. Sendo possível identificar os clientes que não irão pagar as suas contas, ou identificar contas que não podem ser coletadas, a gestão de risco seria facilitada.

Relativamente à aplicação das redes Bayesianas na deteção de fraude, esta será explorada na secção 2.4.4.

2.4.2 Outros Algoritmos de Classificação

Para efeitos de validação e avaliação dos resultados obtidos com a classificação obtida com base em redes Bayesianas, serão utilizados outros algoritmos de classificação para que os seus desempenhos sejam comparados. Estes algoritmos são: *Random Forests* e Redes Neuronais.

As Random Forests foram introduzidas por Breiman (2001) e adicionam uma camada de aleatoriedade ao processo designado por bagging. Ou seja, em adição de construir cada árvore de decisão utilizando uma amostra diferente dos dados com bootstrapping (amostragem aleatória com reposição), as Random Forests alteram a forma como as árvores de classificação ou regressão são construídas. As árvores genéricas dividem cada nodo com base na melhor divisão possível entre todas as variáveis. Numa Random Forest cada nodo é dividido recorrendo à melhor opção entre um subconjunto de previsores escolhidos aleatoriamente para esse mesmo nodo. Esta estratégia acaba por alcançar resultados positivos quando comparada com outros classificadores, incluindo Support Vector Machines e redes neuronais, e é robusta contra a sobre amostragem (Breiman, 2001). Para além disso, é relativamente simples de utilizar pois só necessita de dois parâmetros: o número de árvores na floresta (n); e o número de variáveis a considerar no subconjunto aleatório para cada nodo (m). O algoritmo é processado da seguinte forma (Liaw & Wiener, 2002):

- Retirar amostras do tipo *bootstrap* do conjunto original de dados para n árvores.
- Para cada amostra, criar uma árvore sem poda, com a seguinte modificação: em cada nodo, em vez de escolher a melhor divisão entre todos os previsores, amostrar aleatoriamente m previsores e escolher a melhor divisão entre as variáveis possíveis.
- Realizar a previsão com base na agregação da previsão das n árvores (maioria dos votos para a classificação e média para a regressão).

Relativamente às Redes Neuronais, estas são tidas como uma ferramenta importante para efeitos de classificação. A vantagem deste método reside nos seguintes aspetos teóricos (Zhang, 2000):

- As redes neuronais são um método conduzido pelos dados e auto ajustável na medida em se ajustam aos dados sem qualquer especificação explícita.
- São aproximadores funcionais universais pois podem aproximar qualquer função com precisão arbitrária.
- São modelos não lineares, o que os torna flexíveis na modelação de relações complexas do mundo real.
- São capazes de estimar as probabilidades *a posteriori*, o que fornece a base para estabelecer uma regra de classificação e realizar análise estatística.

Os classificadores de padrões estatísticos são baseados na teoria de decisão de Bayes, onde a probabilidade *a posteriori* tem um papel central. As Redes Neuronais podem fornecer uma estimativa dessa mesma probabilidade, o que indicia uma ligação entre este método e os classificadores estatísticos. No entanto pode não ser possível compara-los diretamente uma vez que as Redes Neuronais não são lineares e são livres de modelo, enquanto os métodos estatísticos são lineares e baseados em modelos. Mas se os valores de saída forem codificados apropriadamente, as Redes Neuronais podem modelar diretamente algumas funções discriminantes (Zhang, 2000). Por exemplo, num problema de classificação com dois grupos, se o resultado desejado for 1 se o objeto pertencer à classe 1, e -1 se o objeto pertence à classe 2, então a rede irá estimar a função descriminante seguinte:

$$g(x) = P(w_1|x) - P(w_2|x) \quad (2.5)$$

A regra descriminante é simples: assignar x a $w_1$se $g(x) > 0$ ou w_2 se $g(x) < 0$. Qualquer função monótona e crescente da probabilidade *a posteriori* pode ser utilizada para substituir essa mesma probabilidade na equação 2.5 de forma a construir uma função descriminante diferente mas essencialmente com a mesma regra de classificação.

2.4.3 Sistemas Multi-Agentes

A modelação e simulação baseadas em Agentes é uma abordagem à modelação de sistemas compostos por agentes autónomos e interativos. Esta abordagem permite estudar muitos fenómenos, nomeadamente o modo como as empresas utilizam a tecnologia para suportar os processos de tomada de decisão, ou o modo como os investigadores utilizam laboratórios eletrónicos no apoio ao seu trabalho. Alguns autores chegam a argumentar de que a modelação baseada em agentes pode ser uma terceira forma de análise científica para além do raciocínio dedutivo e indutivo (Ferber, 1999).

Os avanços computacionais têm permitido um aumento das aplicações baseadas em agentes numa vasta variedade de campos. Estas aplicações vão desde a modelação do comportamento de agentes no mercado bolsista, de consumidores, ou agentes em cadeias de logística, até prever a propagação de epidemias (Macal & North, 2005).

A modelação baseada em agentes tem-se difundido bastante. Isto tem acontecido porque o mundo que nos rodeia tem vindo a tornar-se cada vez mais complexo. Os sistemas que se pretende modelar são cada vez mais complexos em termos das suas interdependências. Isto significa que as ferramentas tradicionais de modelação não são tão aplicáveis como o foram. Depois ainda existem aqueles sistemas que sempre foram demasiado complexos para se modelar adequadamente. Por exemplo a modelação de mercados económicos tem-se tradicionalmente apoiado em noções de mercados perfeitos, agentes homogéneos, e equilíbrios de longo prazo porque estas assunções tornam os problemas analiticamente e computacionalmente tratáveis. Recorrendo à modelação baseada em agentes tem sido possível abordar estes problemas de um ponto de vista mais realista. Outra explicação para a difusão deste tipo de modelação é o facto de os dados serem cada vez mais organizados, com recurso a bases de dados com níveis de

granularidade incrementalmente detalhados. Por último, a capacidade computacional tem avançado rapidamente, permitindo desenhar micro simulações de grande escala que não seriam possíveis alguns anos atrás (Macal & North, 2005).

2.4.3.1 Definição de Agente

Embora não exista nenhum acordo universal na definição precisa do termo "agente", as diferentes definições tendem a ter mais pontos em acordo do que em desacordo. Alguns autores consideram que qualquer tipo de componente independente (*software*, modelo, indivíduo, etc.) como um agente (Bonabeau, 2002). Aqui o comportamento de uma componente independente variar entre regras de decisão reativas e primitivas, até inteligência adaptativa complexa. Outros autores defendem que para uma componente ser considerada um agente o seu comportamento deve ser definitivamente adaptável, ou seja, o rótulo de agente é reservado a componentes capazes de aprender com o ambiente envolvente e alterar o seu comportamento em conformidade (Mellouli et al., 2004). Há ainda quem argumente que os agentes devem incorporar dois tipos de regras, umas de nível base para o seu comportamento geral e outras de nível elevado para reajustamento, tal como "regras para alterar as regras". As regras base fornecem respostas ao ambiente envolvente enquanto as de nível superior fornecem adaptabilidade (Casti, 1997). Já Jennings (2000) fornece um ponto de vista da ciência computacional, enfatizando a característica essencial do comportamento autónomo. O autor refere que a característica fundamental de um agente é a capacidade de este tomar decisões independentes. Isto exige que os agentes sejam ativos e não passivos.

De um ponto de vista prático, pode-se considerar que os agentes têm as seguintes características (Macal & North, 2005):

- Um agente é identificável, um indivíduo distinto com um conjunto de características e regras que governam o seu comportamento e a sua capacidade de tomada de decisão. São autónomos. A distinção implica que o agente tenha uma fronteira que permita determinar facilmente se algo faz parte do agente, se não faz parte, ou se é uma característica partilhada.
- Um agente é situado, presente num ambiente onde interage com outros agentes. Os agentes têm protocolos para interagir com outros agentes, como protocolos

de comunicação, e têm a capacidade de interagir com o ambiente. Têm a capacidade de reconhecer e distinguir as características dos outros agentes.

- Um agente é orientado por objetivos, tendo objetivos a atingir (não necessariamente maximizar) relacionados com o seu comportamento.
- Um agente á autónomo, pode funcionar independentemente no seu ambiente e nas suas interações com os outros agentes, no mínimo numa certa variedade de situações.
- Um agente é flexível e tem a capacidade de aprender e adaptar os seus comportamentos ao longo do tempo com base na experiência. Isto requer algum tipo de memória. Um agente pode incorporar regras que modificam as suas regras comportamentais.

2.4.3.2 Redes de Agentes e Topologias como Base para Interação Social

De um ponto de vista económico, as teorias neoclássicas vêm os indivíduos como decisores independentes. Os agentes transacionam nos mercados, os preços refletem os comportamentos dos outros e as pessoas afinam as suas decisões com base na informação transmitida por esses preços (Wilhite, 2006). A interação entre indivíduos é institucionalizada, impessoal e imparcial. No entanto estes mercados impessoais podem ser vistos como demasiado artificiais. Os indivíduos são influenciados por mais fatores do que apenas os preços, sendo habitualmente influenciados por outros agentes. Tipicamente recorre-se à teoria dos jogos, pois permite estudar como as ações de um individuo influenciam o comportamento de outro. Mas mesmo que a interação entre decisores seja realçada, não se diz muito em relação a quem está a jogar. Isto tem levado a que os economistas, como que observando os sociologistas que estudam vizinhanças e grupos de pares, tenham começado a reexplorar não só os efeitos da interação entre agentes mas também a identidade dos agentes envolvidos na interação. E é aqui que entram as redes, que têm sido alvo de estudos teórico e empíricos para investigar como é que estas estruturas podem influenciar o comportamento dos seus constituintes (Wilhite, 2006).

As redes têm-se tornado particularmente úteis quando agentes interagem principalmente com uma pequena parte da população, ou seja, com os seus vizinhos na rede. Mas muitas atividades locais demonstram ter efeitos em locais distantes na rede. Como é que pequenas

quantidades de informação, comportamentos ou decisões, viajam através da população? Isto tem sido estudado cada vez mais, colocando agentes numa rede e utilizando esta para explorar como é que vários indivíduos se interligam.

O paralelismo entre as redes aplicadas a estudos económicos e outros campos de aplicação é óbvio. É percetível que a metodologia de estudo das individualidades, ações, interações e comportamentos de agentes numa rede, e de que forma esses comportamentos influenciam outros agentes, de um ponto de vista económico não difere muito de uma metodologia que estude os mesmos efeitos mas numa rede de sensores ligados entre si. Estes, tal como indivíduos que atuam nos mercados, também têm as suas próprias características, os seus próprios objetivos, mas as suas ações influenciam ou são influenciadas por outros.

A modelação computacional baseada em agentes é adequada para o estudo de redes nestes campos, como em muitos outros. É possível programar redes complexas com relativa facilidade e observar milhares de interações entre agentes nestas redes. Experiências virtuais podem ser realizadas vezes sem conta, sendo possível efetuar pequenas alterações no sistema e verificar os efeitos dessas mudanças. Todas as decisões, ações e interações podem ser capturadas e analisadas. Deste modo até efeitos subtis podem ser revelados através de um fluxo sistemático de cálculos.

Quanto à estrutura ou topologia das redes, estas podem ser estáticas ou dinâmicas. A Análise de Redes Sociais (*Social Network Analysis, SNA)* estuda a caracterização da estrutura social e as suas interações através de representações de redes e tradicionalmente foca-se em redes estáticas que não alteram a sua estrutura ao longo do tempo nem como resultado de comportamentos de agentes (Macal & North, 2005). Na realidade este tipo de redes podem não existir, pois as instituições e as relações crescem, evoluem e eventualmente desvanecem. Mas se esta evolução é lenta em relação à atividade da rede, as características da rede serão muito mais influentes nas decisões da rede, do que as decisões serão na estrutura da rede (Wilhite, 2006).

No estudo de redes é usual retirar a nomenclatura da teoria dos grafos. Investiga-se uma população de n agentes numa rede, G, que é uma lista desordenada de pares de agentes ligados $\{i, j\}$ que se escreve como $i \sim j$. Deste modo $i \sim j \in G$ significa que os agentes i e j estão ligados na rede G. Nesta secção as ligações serão tidas sem direção, ou seja, se $i \sim j \in G$ então $j \sim i \in G$. As redes são simples na medida em que apenas uma aresta liga um par de agentes ligados e nenhum nodo se liga a sim mesmo. As redes são ligadas, o que significa que qualquer nodo pode chegar a outro a partir de outro nodo seguindo um caminho constituído por um

número finito de ligações. As arestas consideradas não terão pesos ou qualquer valor intrínseco (comprimento ou qualidade) que as distinga de outras arestas. Assim a importância de qualquer aresta resulta apenas da sua localização relativamente às outras. Com base nestas considerações serão apresentadas brevemente alguns tipos específicos de redes. Naturalmente que muitas mais ficaram por analisar, mas este conjunto permite endereçar as principais questões que surgem no estudo de redes.

2.4.3.3 Coordenação e Cooperação nas Redes

Supondo que os nodos de uma rede estão ocupados por agentes que interagem apenas com os seu vizinhos, como é que a topologia da rede afeta as decisões resultantes? Recorrendo a um exemplo com agentes económicos, a coordenação e a cooperação são essenciais numa sociedade, uma vez que as regras, morais e convenções são demasiado custosas de regular, monitorizar e aplicar. Em alguns casos os agentes têm fortes incentivos próprios para cooperar e aí o problema central é aprender o que outros fazem e adaptar-se em concordância. Em outros casos a cooperação pode ser mais ilusória se os agentes tiverem incentivos privados para não cooperarem mesmo que a cooperação traga um bem-estar agregado. Mais uma vez, os economistas recorrem à teoria dos jogos, desta vez para investigar até que ponto os agentes cooperam e coordenam as suas atividades (Wilhite, 2006).

Ellison (1993) investigou a cooperação comparando interações locais a tomadas de decisões globais. Primeiro emparelhou aleatoriamente agentes num jogo de coordenação e deixou o sistema encontrar o seu equilíbrio. Depois colocou os agentes num anel, restringindo a sua interação podendo cada um apenas jogar exclusivamente com os seus vizinhos. As suas descobertas, algo surpreendentes, indicam que o sistema coordena a sua atividade mais rapidamente quando os agentes se focam nos seus vizinhos. Se um agente jogar apenas com os seus vizinhos, a vizinhança, sendo pequena, rapidamente encontra a estratégia ótima. Esta rapidez replica-se através da população e o sistema converge. Alternativamente, jogando aleatoriamente leva a que os agentes tenham de encarar um certo fator surpresa nos seus oponentes, o que leva a que a estratégia predominante demore mais a ser aprendida.

Bala e Goyal (2001) estudaram difusão tecnológica incorporando a utilização de dados históricos e preferências heterogéneas de agentes localizados numa rede. Concluíram que se a informação local tiver um peso maior do que a global, tecnologias diferentes e concorrentes

podem coexistir numa rede conectada. Os mesmos autores (Bala & Goyal, 1998) estudaram agentes que utilizam uma única estratégia de jogo com os seus vizinhos. Depois de cada ronda os agentes consideram o histórico das suas decisões bem como um histórico completo das decisões dos seus vizinhos antes de atualizar as suas estratégias.

Morris (2000) utilizou um jogo de coordenação para investigar o contágio, a probabilidade de uma estratégia particular se propagar de forma a suplantar uma rede arbitrária. Mostrou que em jogos de coordenação com uma dada estrutura de proveitos, as decisões são mais propensas a espalharem-se quanto mais coesos forem as vizinhanças e mais uniformes as redes. Numa vizinhança coesa os vizinhos de um agente tendem a ser vizinhos uns dos outros. A uniformidade existe quando o padrão de vizinhança se repete através da rede. A coesão encoraja todos os agentes na vizinhança a selecionar a mesma estratégia, enquanto a uniformidade permite que esta estratégia escolhida se propague ao longo da população.

Young (1993, 1998) estudou a emergência de convenções sociais. Começou por sugerir que em várias situações a híper racionalidade dos agentes nas teorias económicas neoclássicas são uma representação pouco razoável da natureza. No seu lugar, considera agentes que tomam decisões sensíveis, não necessariamente ótimas, e que tipicamente interagem com uma pequena porção da população utilizando a informação limitada mais prontamente disponível. Na examinação da difusão de inovações através de redes sociais, Young (2005) descobriu que agentes organizados em pequenos grupos de malha-fechada adotam uma tecnologia particular mais rapidamente do que agentes organizados aleatoriamente.

Em jogos de cooperação é possível observar que a cooperação sobrevive se o sucesso social for incluído nas funções utilidade dos indivíduos, tal como a sobrevivência das espécies beneficia um individuo ao aumentar as suas probabilidades de passar o seu código genético para as gerações seguintes (Wilhite, 2006). Os agentes começam com a sua própria estratégia num jogo entre eles e os seus vizinhos. Após recolherem os seus resultados, os agentes observam as estratégias e resultados dos seus vizinhos e ajustam a sua estratégia copiando a utilizada pelo agente mais bem-sucedido na sua vizinhança. Com a sua estratégia atualizada, os agentes voltam a jogar, a inspecionar os vizinhos, a atualizar, e assim por diante. O jogo continua até a distribuição agregada das estratégias se tornar relativamente estável, o chamado estado estacionário. O facto de os agentes imitarem os vencedores em vez de mapear uma estratégia ótima é particularmente importante. De acordo com a literatura de jogos cooperativos em redes, a imitação pode tomar diversas formas, como copiar o melhor retorno médio, copiar a melhor

estratégia de uma amostra de vizinhos, copiar a estratégia copiada mais frequentemente, etc. (Wilhite, 2006).

2.4.3.4 Aplicações da Modelação Baseada em Agentes

Tal como para as redes Bayesianas, também existem várias aplicações reais para a modelação baseada em agentes, desde simulações de comportamento humano até à gestão de redes elétricas inteligentes.

Segundo Bonabeau (2002) a capacidade dos agentes em executar e simular vários comportamentos permite representar de forma apropriada vários sistemas. Num contexto de economia e gestão, os sistemas de interesse poderão ser classificados como fluxos, mercados, organizações ou difusões. Ou seja, a modelação baseada em agentes pode ser útil na gestão de fluxos como o trânsito ou cadeias logísticas, no estudo do nível de impacto que algumas alterações em mercados financeiros podem ter nos seus agentes, na análise de comportamentos coletivos emergentes em certas organizações, ou mesmo na avaliação do grau de influência a que os indivíduos estão sujeitos nos seus contextos sociais.

Jiao *et al* (2006) utilizaram a modelação baseada em agentes para estudar a negociação colaborativa numa rede de logística global. Uma vez que estas redes por norma incluem atividades dispersas e executadas em diversas localizações, as decisões coordenadas são cruciais para a implementação da estratégia logística. Numa rede de logística podem estar inseridas várias entidades autónomas e semiautónomas, que são responsáveis por adquirir, produzir ou distribuir um ou mais produtos. O desempenho destas entidades depende da performance de outras, bem como da sua vontade e capacidade em colaborar e coordenar as suas atividades no seio da cadeia logística.

Modelar um sistema desta natureza pode ser extremamente complexo. Uma tecnologia baseada em agentes facilita a integração de toda a cadeia logística como uma rede de escalões independentes, onde cada um utiliza o seu próprio processo de tomada de decisão. Num sistema baseado em agentes, um número de agentes heterogéneos trabalham de forma independente ou em conjunto com o intuito de resolver problemas num ambiente descentralizado. Através deste paradigma, é possível atingir o objetivo global do sistema agregando os objetivos locais através de várias rondas de negociação (Jiao et al., 2006).

Garcia (2005) estudou a metodologia da modelação baseada em agentes e a sua aplicação na inovação e desenvolvimento de produtos. Segundo o seu trabalho, esta modelação traz grandes benefícios na aprendizagem e compreensão da difusão de inovações, bem como em fluxos de conhecimento e na estratégia organizacional. Traduzir um protótipo de um modelo de difusão para simulações dinâmicas com redes de agentes é de certa forma simples. Os consumidores são modelados como agentes que tomam decisões de adoção ou compra de produtos, tendo por base influências provenientes da interação com outros agentes (consumidores), como o passa-palavra. Outros agentes podem ser modelados como consumidores influenciadores, empresas com campanhas publicitárias, ou outros tipos de comunicação de massas.

Hernandez *et al* (2013) constatam que os avanços tecnológicos, tanto na produção energética como nas tecnologias de informação, tem ajudado a alterar o paradigma da logística na eletricidade de modo a atingir novos níveis de eficiência e sustentabilidade. As *smart grids* são uma tendência que introduz inteligência na rede energética de modo a otimizar a sua utilização. Para que esta inteligência seja eficaz, é necessário recolher informação acerca das operações da rede em conjunto com outros dados do contexto, como fatores ambientais, e modificar o comportamento dos elementos da rede de acordo com as necessidades. Os autores apresentam então um modelo baseado em agentes para um novo conceito de centrais elétricas, onde a geração de energia deixa de ocorrer em instalações de grande dimensão, e passa a ser o resultado da cooperação de elementos mais pequenos e inteligentes. O modelo proposto foca-se não só na gestão dos diferentes elementos, mas inclui também uma série de agentes integrados com redes neuronais artificias para previsão colaborativa da procura desagregada de energia por parte dos utilizadores domésticos.

É possível constatar o vasto leque de aplicações reais para a modelação baseada em agentes, sendo que a deteção de fraude é uma delas. Isto será revisto em maior detalhe na secção 2.4.4.

2.4.4 Redes de Agentes e Aprendizagem Bayesiana na Deteção da Fraude

Relativamente ao fenómeno da fraude, este é um problema que sempre afetou vários setores, nomeadamente o financeiro. No entanto, com a evolução tecnológica têm vindo a surgir

formas mais sistemáticas e organizadas de recolher grandes quantidades de dados, estimulando iniciativas baseadas em análise de dados que pretendem modelar as relações entre combinações de indicadores de fraude, de forma a atualizar, ou melhorar, sistemas automáticos de deteção de práticas fraudulentas. Soluções de aprendizagem automática e de inteligência artificial são cada vez mais exploradas com o objetivo de prevenirem ou diagnosticarem a fraude em vários setores.

Da mesma forma que existem vários tipos fraude e diferentes formas de as cometer, também existem múltiplos métodos e modelos para a deteção desta atividade ilícita. Na literatura existente é possível encontrar vários exemplos onde a análise de redes baseadas em agentes e a aprendizagem de redes Bayesianas são parte integral desses modelos.

Dheepa & Dhanapal (2009) recorreram às redes Bayesianas para descrever o comportamento dos utilizadores envolvidos em fraudes de cartões de crédito. Foi construída uma rede para modelar o comportamento do utilizador fraudulento e outra para modelar o utilizador legítimo. A rede da fraude foi configurada recorrendo a dados fornecidos por especialistas em fraude. A rede dos utilizadores legítimos foi formada com base em dados históricos de utilizadores não fraudulentos. Durante a execução do algoritmo a rede é adaptada a um utilizador específico baseando-se em dados emergentes. Ao inserir evidência nestes modelos e propagando-a através das redes, obtém-se uma avaliação do grau de ajustamento do comportamento do utilizador ao comportamento típico de um utilizador fraudulento ou não.

Num outro estudo, Tuyls *et al* (2002) também recorrem às redes Bayesianas na sua abordagem ao problema da deteção de fraude nos cartões de crédito. A ideia central é fornecer ao algoritmo de aprendizagem um conjunto de dados com informação relativa a transações financeiras inerentes ao sistema onde se pretende detetar a fraude. Depois de um processo de aprendizagem é esperado que o programa seja capaz de classificar corretamente uma transação nova como fraudulenta ou não, com base nas características da transação. Neste caso as redes Bayesianas são encarregues de construir a rede automaticamente e de atuar como o esquema base de representação do conhecimento. O processo de aprendizagem é dividido em duas partes. A primeira diz respeito à identificação da topologia da rede, especificando as ligações em falta e a direção dos arcos da rede. Aqui foi utilizada a métrica MDL (*Minimum Description Length*) pois foi considerado que apresenta um bom compromisso entre a maximização da precisão do ajustamento e minimização da complexidade do modelo. Na segunda procede-se à determinação dos parâmetros numéricos, ou seja, as probabilidades condicionadas para uma certa topologia de rede. Os autores comparam o desempenho das redes Bayesianas com redes neuronais

artificiais e concluíram que as primeiras demonstram melhores resultados na deteção da fraude, bem como um período de treino menor, enquanto as redes neuronais artificiais apresentaram um processo de deteção mais rápido.

Panigrahi *et al* (2009) recorreram à aprendizagem Bayesiana aliada à teoria Dempster-Shafer para propor um modelo de deteção de fraude em cartões de crédito, combinando dados de comportamentos atuais com dados históricos. De acordo com os autores, o aumento da utilização de ferramentas tecnológicas tem levado ao avanço do comércio eletrónico, o que por sua vez incrementa significativamente a utilização de cartões de crédito. Isto no entanto representa também um aumento das oportunidades disponíveis para as tentativas de fraude e o avanço tecnológico também leva a que surjam métodos fraudulentos cada vez mais complexos e sofisticados. Ao longo do tempo várias soluções foram adotadas pelas instituições afetadas por este problema. Mas quem comete fraudes por norma é capaz de se adaptar e com tempo as medidas tomadas pelas instituições acabam por ser contornadas. O objetivo de um sistema de deteção fiável deve então passar por aprender os comportamentos dos indivíduos de forma dinâmica. Um sistema que não tenha capacidade de evoluir ou aprender, poderá rapidamente tornar-se obsoleto. Um individuo fraudulento pode variar os seus métodos e comportamentos de forma propositada para despistar as instituições no seu encalce. Existe assim a necessidade de sistemas de deteção de fraude capazes de integrar múltiplas evidências de diversos comportamentos, incluindo utilizadores fidedignos e fraudulentos.

Para responder a estas exigências, Panigrahi *et al* (2009) propuseram um modelo que combina diferentes tipos de evidência através de combinação de dados, recorrendo ao agregador Dempster-Shafer para esse efeito. O propósito da agregação é conseguir sumariar e simplificar significativamente quantidades massivas de dados que podem provir de inúmeras fontes. Para além desta combinação de dados, foi integrada a aprendizagem Bayesiana no modelo, incorporando conhecimentos prévios e dados observados em cartões suspeitos. Os autores justificam a utilização da aprendizagem Bayesiana pelo facto de esta fornecer uma estrutura que permite construir sistemas capazes de tomar decisões racionais quando confrontados com incerteza. São ainda métodos que simulam a intuição humana de uma forma rigorosa, dando azo a modelações promissoras para os processos neurológicos. O modelo proposto tem então quatro componentes. Um filtro baseado em regras, um agregador Dempster-Shafer, uma base de dados do histórico de transações e o algoritmo de aprendizagem Bayesiano. No filtro é determinado o nível de suspeita de cada transação com base no seu desvio em relação um padrão fidedigno. O

agregador é utilizado para combinar múltiplos dados das transações e determinar uma crença inicial. A transação é classificada como normal, anormal ou suspeita dependendo da comparação com essa crença. Assim que uma transação é tida como suspeita, a crença é posteriormente fortalecida ou enfraquecida de acordo com a sua semelhança em relação a transações fraudulentas ou genuínas presentes no histórico, recorrendo à aprendizagem Bayesiana.

Num outro estudo, Prodromidis e Stolfo (1999) utilizam aprendizagem baseada em agentes para detetar fraude e intrusos em sistemas de redes de informação, nomeadamente em aplicações para a prevenção de fraude em cartões de crédito. A sua abordagem incluiu duas componentes principais, agentes deteção de fraude locais e um mecanismo integrado que combina o conhecimento coletivo adquirido individualmente pelos agentes locais. Estes consistem em modelos de classificação calculados através de aprendizagem automática num ou mais locais. Depois, com recurso à meta-aprendizagem, procede-se à combinação dos resultados dos vários classificadores. Segundo os autores, conduzindo o treino dos classificadores através de bases de dados distribuídas, a meta-aprendizagem pode reduzir significativamente o tempo de aprendizagem total, uma vez que é realizada uma aprendizagem paralela sobre conjuntos de dados mais pequenos. A combinação final dos detetores de fraude distribuídos pode ser usada como sentinelas prevenindo possíveis fraudes, inspecionando e classificando cada transação nova.

Colladon e Remondi (2017) exploraram a aplicação do estudo das redes sociais na prevenção da lavagem de dinheiro. Este fenómeno é uma prática ilícita comum e passa por tentar transformar lucros obtidos ilegalmente em ativos legítimos. Dinheiro obtido ilegalmente é geralmente "lavado" através de transações que envolvem bancos ou outro tipo de instituições financeiras. Pode ter um forte impacto numa economia, aumentando o risco operacional de transações financeiras e ameaçando a estabilidade das instituições financeiras.

Neste estudo os autores analisaram a base de dados interna das transações de uma empresa de factoring, cuja atividade envolve adquirir créditos de curto prazo aos seus clientes, sendo que estes obtiveram esses créditos através do fornecimento de bens ou serviços que ainda não tinham sido pagos. Este tipo de negócio tem sido historicamente associado a transações para lavagem de dinheiro. É então essencial monitorizar estas transações de modo a identificar indivíduos ou entidades suspeitas, dos quais os perfis são de alto risco. Para isto os autores desenvolveram um modelo que pode ser replicado para qualquer empresa de factoring, com o objetivo de criar perfis dos seus clientes e de terceiros que estejam envolvidos, atribuindo-lhes

uma classe de risco. Recorrendo a métricas baseadas na análise de redes sociais, foi possível avaliar e prever o nível de risco dos perfis. Conclui-se que os indivíduos mais ameaçadores lidam com transações maiores e mais frequentes, e são de certa forma mais periféricos na rede de transações. Medeiam transações através de diferentes setores e operam em regiões de elevado risco. Os autores sugerem também uma forma de identificar possíveis núcleos de criminosos, com base numa análise visual das ligações implícitas entre diferentes empresas com o mesmo representante, salientando a importância da abordagem baseada em redes de agentes na procura de operações suspeitas.

Axelsson e Lopez-Rojas (2012) depararam-se com a dificuldade em obter dados reais para este tipo de estudos, devido à natureza sensível das transações financeiras. Recorreram então à simulação baseada em sistemas Multiagentes para gerar dados de transações sintéticos. Assim foram capazes de obter registos de transações financeiras e utiliza-los para alimentar o estudo de diferentes cenários de deteção de branqueamento de capitais através de métodos de *Machine Learning*. Lopez-Rojas (2014) também recorreu a sistemas baseados em agentes para representar e simular o comportamento de utilizadores de aplicações *mobile* para realização de transações monetárias, bem como clientes e vendedores de lojas de retalho *online*. O comportamento *standard* foi modelado com base em dados observados no campo, sendo codificado nos agentes na forma de regras de transações e interações entre os mesmos. Alguns destes agentes foram desenhados intencionalmente para agirem de modo fraudulento, de forma a replicar padrões observados em casos de fraude reais. Indícios de fraude conhecidos foram introduzidos nas simulações para testar e avaliar os resultados da deteção de fraude. Dados sintéticos já foram utilizados anteriormente para propósitos semelhantes (Barse et al., 2003), onde proteção dos dados privados dos clientes é uma vantagem.

Shamshirband *et al* (2013) analisaram vários sistemas de deteção de intrusos em redes móveis com base em sistemas multiagentes. Os autores fazem notar que a implementação de redes de sensores *wireless* para aplicações como serviços de emergência, vigilância ou monitorização, inclui sempre a ameaça permanente de múltiplos riscos, intrusões e ataques cibernéticos. Uma das maiores dificuldades na segurança destas regras é a deteção de intrusos, que procura identificar o mau uso e os comportamentos anormais de modo a assegurar operações na rede que sejam seguras e de confiança. Depois de analisar um vasto leque de implementações para a deteção de intrusos, os autores afirmam que esta é conseguida de uma forma mais eficiente quando se recorre a sistemas multiagente.

Huang *et al* (2010) elaboraram um exemplo deste tipo de sistemas. A rede implementada incorpora três tipos de agentes periféricos e um agente central. Os agentes de recolha de dados filtram e reorganizam os dados obtidos e depois transmitem a informação filtrada para os agentes de análise de dados. Estes são a chave para o processo de deteção de intrusos, uma vez que são responsáveis por analisar compreensivamente os dados de forma a identificar padrões fora do comum. Existem ainda os agentes responsáveis pela comunicação, que não tratam nem analisam informação, apenas se responsabilizam pela difusão da mesma. O agente central monitoriza todo o sistema, assegurando-se que o processo é conduzido conforme o pretendido.

Este sistema recolhe informação em vários pontos-chave da rede de sensores e computadores, para depois a processar e analisar. Ao partilhar e aglomerar a informação recolhida, é possível descobrir sinais de ataque à rede de uma forma eficiente. A rede de agentes permite uma deteção de intrusos em tempo real e está preparada para tomar ações imediatas para mitigar o problema.

Zhu *et al* (2006) abordam o mesmo problema com dois tipos de agentes. Um que incorpora um módulo de aprendizagem, que é capaz de se auto ajustar aos dados e à rede recorrendo a uma ou mais técnicas de *data mining*. Este é capaz de criar regras e o outro tipo de agentes, os detetores, utilizam essas regras para detetar a presença de anomalias.

3. Problema, Modelo e Dados

3.1 Problema

Depois de cometido um ato fraudulento, por norma os agentes envolvidos encontram-se na posse dos respetivos ganhos, mas uma vez que estes foram obtidos de foram ilegal, é necessário que haja um processo de transformação que permita a sua utilização no mercado regulado. Isto envolve uma série de transações onde são por vezes adquiridos bens que possam servir como ativos "limpos". Também envolve uma elevada quantidade de transferências monetárias de valores consideravelmente mais baixos do que o montante total que se pretende "lavar". Ao distribuir este montante em várias transações, por diferentes contas, agentes e aquisição de ativos, o agente fraudulento dificulta a deteção da utilização dos fundos obtidos ilegalmente.

A quantidade de operações envolvidas, a realização de transações por canais onde não há registo de informação e a dimensão da própria rede transações, torna a análise, prevenção e deteção destas situações extremamente complexa. É um processo moroso e árduo. Tendo isto em conta, esta secção pretende explorar uma metodologia que possa agilizar em certa medida esse processo. O foco recairá na fase de deteção do branqueamento de capitais, sem se preocupar para já sobre o tipo de fraude que deu origem aos fundos obtidos.

Um aspeto que por vezes dificulta a análise deste problema é o facto de que as transações são realizadas em várias contas e em diferentes instituições financeiras. Por essa razão, o problema é abordado de um ponto de vista de uma suposta entidade reguladora que tem acesso a informação de transações financeiras de várias instituições, sem restrições. Pretende-se então desenvolver um sistema que analise as transações de vários agentes e que no final seja capaz de os classificar em relação à possibilidade de terem intenções fraudulentas ou não. Na prática, tal sistema poderia contribuir significativamente para uma melhor eficiência e eficácia na análise e deteção da fraude. Se o resultado obtido fosse uma lista de agentes suspeitos, com um valor de probabilidade associado, as instituições poderiam iniciar as suas análises com base nesses dados, começando por investigar os agentes com maior probabilidade de serem fraudulentos, podendo assim poupar tempo e recursos valiosos.

3.2 Metodologia Adotada

Para estudar, desenvolver e testar tal sistema, são naturalmente necessários dados que possam ser analisados. Neste caso, são necessários dados de transações financeiras, nomeadamente quais os agentes de origem e destino ou o montante envolvido. Este tipo de informação é sensível e privada, e por conseguinte não é de fácil acesso. As instituições financeiras que estão na sua posse tentam sempre garantir a privacidade dos dados dos seus clientes, logo, não têm por hábito divulgá-las, mesmo para propósitos académicos. Por esta razão, surgiu a necessidade de desenvolver um modelo baseado em agentes que seja capaz de gerar dados sintéticos de transações financeiras.

O modelo pretende representar uma rede de transações financeiras. O espaço físico não é modelado explicitamente, bem como o movimento dos agentes que constituem a rede. Os agentes são criados com um conjunto de características que representam o seu perfil, como a localização, nacionalidade e predisposição para encetar atividades ilegais. Recorrendo a estas características os agentes criam as suas próprias redes de transações financeiras, que serão utilizadas para realizar transações ao longo da execução do modelo. O desenho da rede financeira acaba por ser enviesada pelas características individuais de cada agente.

Durante a execução do modelo se um agente decidir realizar uma transação, este pode escolher faze-lo legalmente ou ilegalmente. Se a transação for legal o agente irá apenas movimentar o montante pretendido da sua conta para a conta destino. Se a transação tiver intenções ilegais, o agente decide utilizar um esquema por camadas, dividindo o montante pretendido em múltiplas transações de quantias mais pequenas para diferentes contas, evitando assim alguns reportes requisitados pelos bancos.

O modelo irá produzir uma lista sumária de todas as transações entre agentes (data, origem, destino, montante, etc). Essa lista deverá representar o conjunto de dados a ser analisado pelos algoritmos de *machine learning* para que seja realizada a deteção de fraude. Os algoritmos utilizados serão as já mencionadas Rede Bayesianas e ainda Redes Neuronais e *Random Forests* para efeitos de comparação de performance.

Para o desenvolvimento do modelo foi utilizada a ferramenta Netlogo (Wilensky, 1999). É uma ferramenta amplamente utilizada estudos que envolvam a modelação de sistemas baseados em redes de agentes, principalmente para quem está a dar os primeiros passos nesta área, mas não só. Isto porque possui uma linguagem de programação simples e uma interface

"amigável" para o utilizador. Permite modelar sistemas complexos, definindo regras que irão ditar o comportamento dos agentes. Com base nas regras e nos parâmetros pré-definidos, os agentes tomam decisões ao longo do tempo, permitindo mais tarde observar e analisar os resultados de modo a tirar conclusões quanto às suas interações. Assim, é possível modelar o comportamento dos agentes a um nível local, para depois observar como estes se comportam e quais serão as consequências das suas ações a um nível global.

Para o processamento de dados foram utilizadas duas aplicações, o RStudio (RStudio Team, 2015) e o Knime (Berthold et al., 2009). Ambas são ferramentas apropriadas para a aplicação de algoritmos de *Machine Learning* e métodos de análise e tratamento de dados. São de fácil utilização e a curva de aprendizagem é rápida, sendo que possuem uma vasta coleção de módulos e/ou funções previamente desenvolvidas. O processamento de dados neste trabalho poderia ter sido realizado unicamente em qualquer uma das ferramentas, mas o desejo de explorar e conhecer ferramentas diferentes levou a que ambas fossem utilizadas.

3.2.1 Assunções e Funcionamento do Modelo

Os agentes presentes neste modelo podem representar três tipos de entidades, individuais, negócios e intermediários financeiros. Os negócios subdividem-se em atividades com fins lucrativos, sem fins lucrativos, *trust funds* ou empresas do tipo *shell*. Os intermediários podem ser divididos em formais (bancos, fundos) ou informais (transferências que não passam pelo sistema bancário, de mão para mão). Os indivíduos não são divididos em nenhuma outra subcategoria.

No processo de inicialização todos os agentes definem a sua localização e nacionalidade. Neste protótipo, estas características podem apenas assumir o valor de "Portugal" ou "Estrangeiro". Ambos são definidos por uma percentagem definida pelo utilizador, que pode determinar qual a percentagem de agentes é de nacionalidade estrangeira e se encontra localizada no estrangeiro.

O parâmetro relacionado com a propensão de cada agente em participar em transações fraudulentas pode ir de 0 a 100, sendo que quanto mais elevado o valor, maior a propensão. Este valor é atribuído aleatoriamente, sendo que o utilizador também pode definir a percentagem de agentes com elevada propensão à fraude.

Existe uma outra variável relacionada com a inclusão do agente numa lista "negra" de potenciais suspeitos, designada por *watch list*. São selecionados aleatoriamente alguns agentes com elevadas predisposições para serem incluídos na lista. Para introduzir algum ruido, são também selecionados aleatoriamente alguns agentes correspondentes a uma pequena percentagem da população total.

Para criar a rede de ligações entre agentes foram utilizadas dois tipos de ligações. Um relativo a ligações entre intermediários financeiros e outros agentes (linhas de transação), e outro entre indivíduos e/ou negócios (conexões). Primeiro assumiu-se que cada intermediário financeiro estava ligado a 40% dos restantes intermediários (as instituições financeiras por norma pertencem a uma ou mais redes com protocolos de transações, como TARGET 2 ou SEPA, onde possuem ligações com vários bancos da mesma rede, de modo a facilitar e agilizar a realização de transações monetárias). Depois cada indivíduo e negócio cria uma ligação com um certo número de intermediários, sendo que esse número é aleatório entre 1 e 4 (máximo de 4 contas por exemplo).

De seguida os indivíduos e negócios criam conexões entre si. Cada agente deste grupo determina primeiro o seu número de vizinhos retirando um número aleatório de uma distribuição exponencial de média 2 (está demonstrado que em diversas tipologias o número de vizinhos é bem representado por esta distribuição). Os indivíduos e negócios criam então ligações a esse número de agentes (exceto intermediários financeiros), sem duplicação. São ainda criadas algumas conexões adicionais. Um indivíduo ou negócio que tenha uma predisposição maior que 70% cria uma ligação com um negócio do tipo *trust* ou *shell* (são usualmente associados a tentativas de fraude e de fugas ao fisco). Se tiverem uma predisposição acima de 90%, criam uma outra ligação a negócios de tipo trust ou Shell e ainda a um outro agente que tenha uma predisposição acima de 70%. Isto pretende simular a tendência de que indivíduos com elevada propensão a procurar transações fraudulentas estão por norma altamente conectados a agentes com propensões semelhantes.

Quando o modelo é executado, 5% do total de indivíduos e negócios iniciam um plano de transação (num cenário real, é improvável que todos os agentes realizem transações todos os dias), a cada intervalo de tempo. Estes executam o seguinte:

1. Os agentes limpam os valores de transações prévias.
2. Definem o montante a transferir (até um máximo de 100 000).

3. Definem o destino (de entre o conjunto de vizinhos). Em algumas ocasiões (10% de probabilidade) é escolhido um agente fora da sua rede, para introduzir algum ruido na rede.
4. O agente de origem verifica se tem ligações com intermediários financeiros que possam permitir fazer o dinheiro chegar ao destino. Se não tiverem ligações em comum, é criada uma nesses moldes.
5. O agente de origem decide como vai estruturar a transação. Se a origem e destino tiverem um predisposição baixa (<60%) então a transação é realizada de uma forma simples. Se a predisposição de ambos for elevada mas o montante for menor que 10 000, também se realiza uma transferência simples. A utilização deste valor deve-se ao facto de ser o limite a partir do qual as instituições financeiras serem obrigadas por lei a reportar a respetiva transação às entidades reguladoras. Se for maior que esse valor, o agente de origem decide estruturar a transação e dividir o montante em várias transações de valores menores. Escolhe um ou mais agentes na sua rede que tenham uma ligação com o agente de destino e desencadeia uma série de transações entre eles, de modo a que o montante total chegue ao destino sem que seja facilmente detetável a origem inicial do dinheiro.

O agente cria então um agendamento para as suas transações. Se a transação for simples, o montante é transferido diretamente da origem para o destino. Se não, calendariza as suas várias transações, de montante heterogéneo ou não (sempre menor que 10000), a distribuir por um mais agentes da sua rede para fazer chegar o valor pretendido ao nó de destino.

Cada transação é registada numa tabela, sendo que esta será depois analisada via *data mining*. As entradas da tabela comportam os seguintes campos:

X01 – Origem (Agente que executa a transação)

X02 – Tipo da Origem (Indivíduo, Negócio com fins lucrativos, Negócio sem fins lucrativos, Trust ou Shell)

X03 – Nacionalidade da origem (Portuguesa ou Estrangeira)

X04 – Localização da origem (Portugal ou Outra)

X05 – Situação quanto à presença da origem na *watch list* (Sim ou Não)

X06 – Destino (Agente ao qual será transferido o montante)

X07 – Tipo do destino (Indivíduo, Negócio com fins lucrativos, Negócio sem fins lucrativos, Trust ou Shell)

X08 – Nacionalidade do destino (Portuguesa ou Estrangeira)

X09 – Localização do destino (Portugal ou Outra)

X10 – Situação quanto à presença do destino na *watch list* (Sim ou Não)

X11 – Montante da transação

X12 – Dia da transação

X13 – Classificação ao nível da presença de Fraude (Sim ou Não)

3.2.2 Processamento de dados

O objetivo da análise de dados é conseguir classificar um conjunto de transações de um dado agente, num determinado período, relativamente à probabilidade constituírem uma tentativa de lavagem de dinheiro. Aqui, como referido anteriormente, assume-se que um agente que pretenda realizar este ato fraudulento irá dividir o montante alvo em quantias menores e várias transações, para um ou mais agentes da sua rede. Deste modo, não faria sentido analisar os dados transação a transação, uma vez que isto não permitiria realizar uma inferência robusta. Tendo isto em conta, numa primeira fase de tratamento de dados as transferências são ordenadas, primeiro por agente na origem da transação, depois por ordem cronológica. Desta forma o conjunto de dados fica dividido em blocos, sendo que cada bloco dirá respeito a todas as transações iniciadas por um dado agente.

Estando as transações ordenadas por ordem cronológica, isto permite aplicar uma janela deslizante a cada bloco de transações. Esta janela pretende agregar a informação de várias linhas (transações) numa só, para que mais tarde o algoritmo de classificação possa avaliar a probabilidade de existência de fraude num determinado período de uma forma mais consistente. Mais uma vez, analisando uma transação de cada vez não permitirá tirar este tipo de conclusões. É necessário analisar cada situação de um ponto de vista mais abrangente, que inclua várias transações, de modo a que seja possível detetar comportamentos e tendências indicativas da presença de fraude. Na figura 3.1 está ilustrado o diagrama de fluxo do modelo de simulação e do processamento de dados. Inicialmente o utilizador define os parâmetros de entrada do sistema (total de empresas e indivíduos, percentagem de agentes incluídos na *watch list,* e percentagem de agentes fraudulentos). De seguida, as restantes variáveis associadas aos agentes

são atribuídas automaticamente com base nas assunções do modelo e nos parâmetros introduzidos pelo utilizador. Assim que esta fase está completa, os agentes formam a sua rede de vizinhos e iniciam a realização de transações. Os dados das transações serão guardados num ficheiro (formato Excel .csv), e este será tratado e processado, incluindo seleção de variáveis e a aplicação de uma janela temporal deslizante para agregar informação de um determinado período de tempo. Depois deste processamento, os dados estarão prontos para serem divididos em duas partes (treino e teste) para que sejam aplicados os algoritmos de *Machine Learning* e por último se realize a classificação dos agentes relativamente à sua propensão para a fraude.

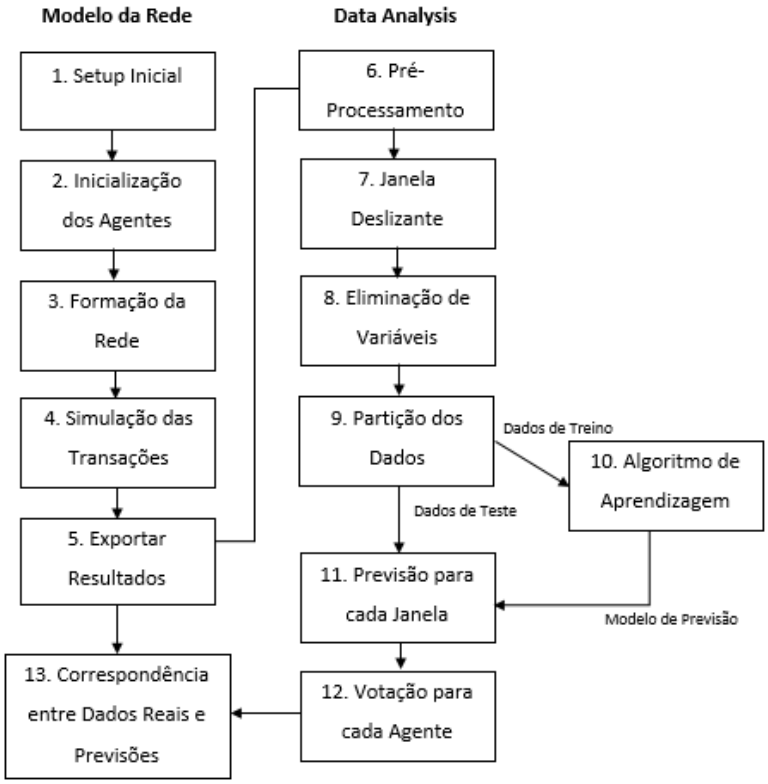

Figura 3.1 Diagrama de Fluxo do Modelo e do Processamento de Dados.

4. Simulação e Resultados

A interface da aplicação NetLogo foi utilizada para definir o número de empresas e indivíduos, a percentagem de agentes incluídos na *watch list* e a percentagem de agentes propensos à fraude. Uma vez definidos estes valores, a restante configuração dos agentes e da rede é completada pelo próprio modelo, de acordo com as assunções mencionadas previamente. A figura 4.1 demonstra um exemplo de configuração de uma simulação e a ilustração da rede de agentes resultante.

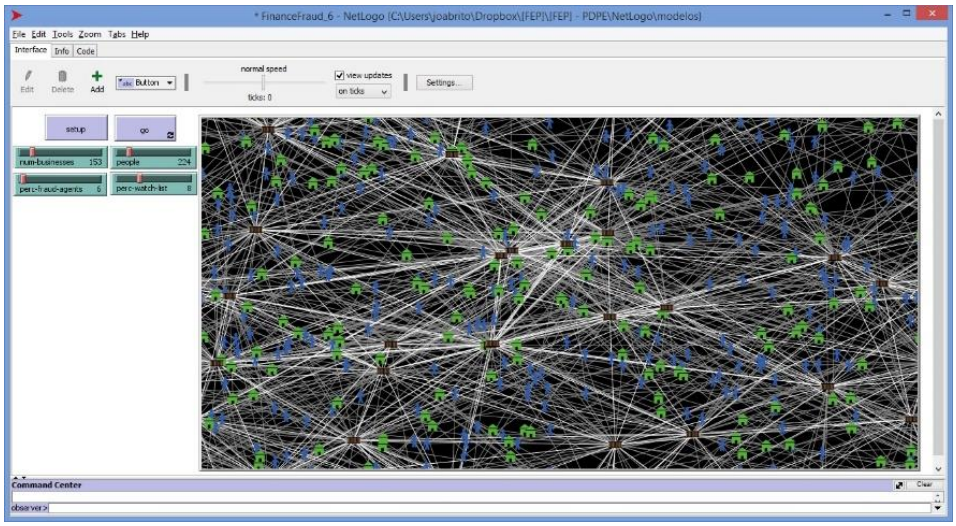

Figura 4.1 Exemplo de configuração e rede resultante no NetLogo. *Nota: linhas brancas representam ligações entre nodos (linhas para transações e ligações a vizinhos); nodos são indivíduos (azul), empresas (verde) e intermediários financeiros (castanho).*

De modo a analisar o comportamento global do modelo, 30 configurações diferente foram executadas. Os valores introduzidos nessas simulações encontram-se entre os seguintes intervalos: número de indivíduos [400 a 1200]; número de empresas [150 a 900]; percentagem de utilizadores na *watch list* [2% a 11%]; e percentagem de agentes com predisposição à fraude [4% a 21%]. Na tabela 4.1 é possível visualizar um resumo dos dados inseridos para as 30 simulações. O objetivo passa por detetar transações fraudulentas de um determinado agente e utilizar essa classificação para avaliar se um agente é ou não propenso a cometer fraude.

Um certo intervalo de tempo é considerado de acordo com a assunção de que um agente fraudulento irá tentar dividir um valor avultado em várias transações para diversos vizinhos.

	Nº de Indivíduos	Nº de Empresas	% de Agentes na Watch List	% Agentes Propensos à Fraude
Média	946,7	627,2	6,8	9,8
Desvio Padrão	210,0	212,6	2,3	3,9

Tabela 4.1 Resumo dos dados de entrada para a simulação no Netlogo.

Assim, não faria sentido analisar as transações uma a uma. Por esta razão, o primeiro passo do processamento de dados passa por ordenar as transações por agente de origem e de seguida por ordem cronológica. Deste modo é possível dividir os dados em blocos, em que cada um agrega transações iniciadas por um dado agente. Assim que as transações estejam ordenadas por ordem cronológica, isto abre a possibilidade de aplicar uma janela deslizante a cada bloco de transações.

A janela deslizante procura agregar a informação relativa a diversas transações numa só linha do conjunto de dados. Esta agregação permitirá ao algoritmo de classificação analisar os eventos de um determinado período de tempo, numa só linha.

Cada linha irá conter informação relativas às transações dos últimos 30 dias e o intervalo entre a data de início de cada janela é de 15 dias. Isto significa que as janelas se irão sobrepor por 15 dias, permitindo uma análise mais robusta.

Para cada janela, o processo deverá determinar os seguintes valores: W1 – Agente de origem; W2 – Tipo da Origem; W3 – Nacionalidade da Origem; W4 – Localização da Origem; W5 – Estado da Origem na Watch List; W6 – Nº de Transações na Janela; W7 – Nº de Transações na Janela com Trust/Shells; W8 – Nº Transações com Agentes Estrangeiros; W9 – Nº de Transações com Agentes na Watch List; W10 – Montante Médio por Transação da Janela; W11 – Coeficiente de Variância; W12 – Centralidade; W13 – Suspeita de Fraude.

	Número de janelas com fraude	Média de transações por janela	Média de transações por janela com fraude	Média da centralidade dos agentes de origem (por janela)	Média da centralidade doas agentes de origem (em janelas com fraude)
Média	1597.1	3.2	6.0	0.01	0.09
Desvio Padrão	1081.7	0.8	1.0	0.01	0.04

Tabela 4.2 Resumo das transações por janela deslizante.

Comparando esta lista de atributos com os dados resultantes do modelo, é possível identificar dois novos campos a analisar, a centralidade e o coeficiente de variância. O coeficiente de variância é utilizado pois o valor médio por transação pode por vezes ser enganador quando utilizado de forma isolada. Uma vez que se está a analisar um conjunto de transações, uma medida que normalize esse valor pode ser útil. Assim, o coeficiente de variância é calculado dividindo o valor médio das transações de uma dada janela pelo desvio padrão dessas transações. Este valor permitirá avaliar o nível de dispersão dos montantes transferidos sem recorrer a uma escala.

A centralidade irá contribuir na avaliação da relevância de cada agente na sua própria rede de vizinhos. Este deverá ser um fator importante a ter em conta pois um agente com predisposição à fraude por norma é um ponto central para um número elevado de vizinhos.

Assim que todos estes valores forem determinados para todas as janelas, este conjunto de dados servirá com entrada para os algoritmos de classificação. Classificação essa que será conduzida segundo uma configuração padrão, dividindo o conjunto de dados em dois blocos, um com dados de teste e outro para alimentar os algoritmos de aprendizagem. Estes construirão o modelo de previsão que será então aplicado aos dados de teste, obtendo assim a classificação para cada janela. Por fim, a classificação será avaliada com recurso à análise às curvas de ROC, nomeadamente à medida da *Area Under Curve* (AUC).

No entanto, isto apenas avalia a probabilidade de uma janela conter uma ou mais transações fraudulentas. Seria também interessante estimar se um agente é propenso a cometer este tipo de fraude ou não. Com este propósito, os resultados da classificação podem ser utilizados para construir uma "opinião" acerca de cada agente, através de uma simples votação. Por cada janela classificada com classe positiva, o agente de origem das transações englobadas pela janela terá um voto a favor da possibilidade de este ser propenso à fraude. Agentes com mais de dois votos serão classificados como suspeitos de fraude. Esta classificação será então comparada com os dados "reais" resultados do modelo da rede de agentes, onde basta uma transação com intenções fraudulentas para considerar o agente como fraudulento. A eficácia da previsão poderá então ser medida através da taxa de acertos quando comparando as duas listas. A lista de agentes que obteve mais de dois votos e como tal são suspeitos de fraude, com a lista de agentes que realizou pelo menos uma transação com intenções fraudulentas no decorrer da simulação no Netlogo.

	ROC AUC para a previsão de fraude das janelas deslizantes			Precisão da classificação por votação das janelas		
	Redes Bayesianas	**Redes Neuronais**	**Random Forests**	**Redes Bayesianas**	**Redes Neuronais**	**Random Forests**
Média	0,9619	0,8988	0,9712	0,8972	0,8818	0,8988
Desvio Padrão	0,0115	0,0232	0,0101	0,0349	0,0343	0,0311

Tabela 4.3 Resultados agregados da classificação das janelas deslizantes para 30 cenários diferentes.

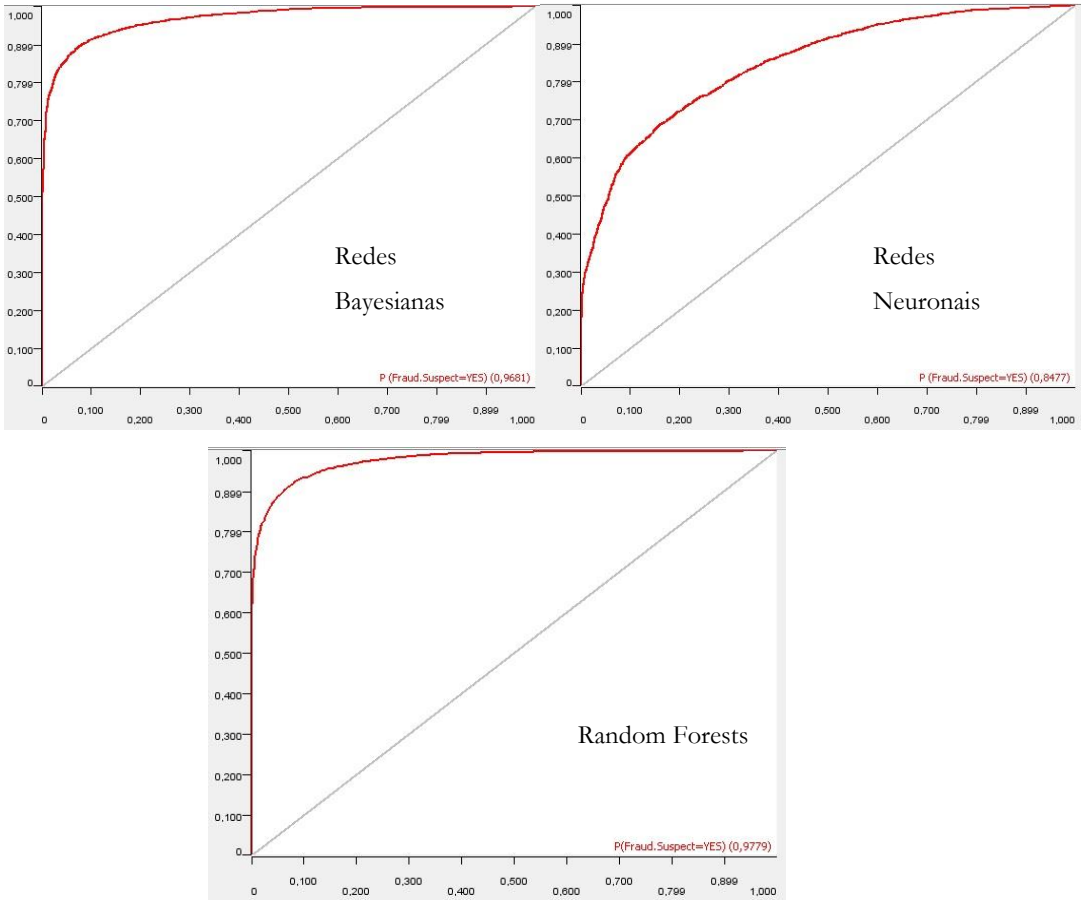

Figura 4.2 Representação gráfica das curvas de ROC obtidas para os três algoritmos de classificação utilizados. *Nota: Resultados relativos a uma das 30 simulações realizadas, escolhida aleatoriamente para efeitos de ilustração.*

Com os resultados da janela deslizante utilizados como dados de entrada, três algoritmos de classificação foram utilizados para detetar/prever a presença de fraude em cada janela

analisada: Redes Bayesianas, Redes Neuronais e *Random Forests*. Os algoritmos foram avaliados de acordo com a AUC (*Area Under Curve*) da respetiva curva de ROC (tabela 4.3), uma medida comum para avaliar a qualidade de algoritmos de extração de conhecimento de dados.

No final da classificação de todas as janelas, os resultados são utilizados para realizar uma votação na possibilidade de cada agente ser propenso à fraude. O resultado da votação é então comparado os dados do modelo BOND (os dados "reais").

Tendo em conta os resultados da Tabela 4.3, é possível concluir que os três algoritmos tiveram performances bem-sucedidas. Os resultados apresentados são relativos à média de 25 simulações. O baixo desvio padrão das AUC obtidas indica a estabilidade do procedimento utilizado, reforçando assim a consistência dos resultados.

No entanto, as Redes Neuronais tiveram resultados ligeiramente inferiores, quer na análise da curva de ROC ou na taxa de acerto da votação. Também é de notar que obteve um desvio padrão mais elevado nos seus resultados e teve um tempo de processamento significativamente mais longo quando comparado com os restantes.

Relativamente aos resultados das Redes Bayesianas e das *Random Forests*, tiveram performances semelhantes em ambos os critérios, mas as últimas tiveram um desempenho tenuemente superior e um desvio padrão mais reduzido. É também importante analisar os resultados da etapa de seleção de atributos durante a extração de conhecimento de dados. Antes de executar o algoritmo de aprendizagem, é realizada uma eliminação de atributos para trás, com recurso ao mesmo algoritmo de aprendizagem. De acordo com esta etapa, é possível proceder com os restantes passos da análise de dados considerando apenas cinco variáveis da janela deslizante, à custa de um erro de 0,076 para as Redes Bayesianas. Essas variáveis são: Nº de Transações; Nº de transações com Trust/Shell; Montante Médio por Transação; Coeficiente de Variação; Centralidade. A mesma análise para as *Random Forests* dita que para um erro de 0,035, as cinco variáveis são: Tipo de Origem; Nº de Transações; Montante Médio por Transação; Coeficiente de Variação; Centralidade.

5. Notas Finais

Neste trabalho foi desenvolvido um modelo de deteção de transações fraudulentas de um determinado agente, com o objetivo de classificar esse agente como sendo propenso ou não, à fraude. O modelo é baseado na análise de uma rede financeira de agentes (criada e simulada através de modelação baseada em agentes). Foram definidos três tipos de agentes: indivíduos, negócios e intermediários financeiros. As ligações podem ser de dois tipos: entre intermediários financeiros e entre indivíduos e/ou negócios. Partindo do conjunto de dados de transações, uma janela deslizante percorre as linhas previamente agregadas por agente e fornece os dados de entrada para os algoritmos de classificação. Os resultados demonstram que é possível prever de forma eficiente os comportamentos dos agentes, com base em transações anteriores. Cinco variáveis da janela deslizante foram identificadas com sendo especialmente relevantes.

O objetivo passou também por conseguir que a simulação da rede de agentes replicasse o comportamento de agentes comuns e fraudulentos aquando da execução de transações financeiras. Para os fraudulentos, a intenção foi associar uma situação de lavagem de dinheiro a ações como dividir um montante elevado por várias transações de valores menores, recorrendo a instituições financeiras frequentemente associadas a este tipo de atividades, ou realizar um elevado número de transações num determinado espaço temporal. Isto acabou por ser refletido com sucesso pelo modelo e foi identificado pelos algoritmos de classificação, comprovado não só pelos bons resultados das previsões, mas também pela seleção de características, onde as variáveis selecionadas estão alinhadas com as assunções acerca do comportamento fraudulento que constitui a base do sistema. Embora as *Random Forests* e as Redes Bayesianas tenham demonstrado performances semelhantes, ambos apresentam elevados valores de falsos positivos, algo que deve ser explorado no futuro de modo a melhorar os resultados globais.

Outros aspetos importantes a desenvolver no futuro passam por integrar no modelo a componente geográfica dos agentes (incluindo paraísos fiscais por exemplo) e atribuir características adicionais aos agentes de forma a modelar um perfil mais próximo da realidade. Isto porque de momento as transações baseiam-se demasiado no fator de predisposição à fraude de cada agente, o que acaba por ser uma visão simplificada da realidade. Isto serve o propósito de simulação e validação inicial dos pressupostos, mas no futuro seria relevante conceder aos agentes características mais complexas para que este ajam de acordo com as suas preferências

(saldo de conta, destinatários frequentes, setor profissional). Outro aspeto importante seria basear o processo de construção da rede de agentes em topologias de rede mais complexas, que se assemelhem mais ao mundo real. Por último, seria também interessante comparar a performance dos modelos de classificação utilizados com modelos mais simples.

É essencial referir que os resultados obtidos, pese embora estarem assentes em visões simplificadas da realidade, demonstram a possibilidade de aplicar com sucesso métodos de *Machine Learning* à deteção de casos de fraude em ambientes financeiros. Podem também ser aplicados no seu estudo e simulação, contribuindo para uma melhor compreensão e prevenção deste fenómeno. Recorrendo a estas metodologias, é possível simplificar e automatizar a análise de práticas fraudulentas, permitindo assim um controlo mais eficaz da fraude, o que deverá eventualmente levar a uma maior segurança nas empresas e uma maior estabilidade nos mercados e na economia no geral.

Atualmente muitos dos processos de controlo da fraude são manuais e exigem um elevado número de peritos e de horas para alcançar os objetivos. Neste trabalho demonstrou-se que a utilização da modelação baseada em agentes permite estudar e simular os comportamentos de agentes fraudulentos, impondo diferentes condições de modo a averiguar inúmeros cenários, contribuindo para uma melhor compreensão deste fenómeno, das suas causas e efeitos. Já os algoritmos de *Machine Learning*, demonstraram serem capazes de detetar automaticamente comportamentos suspeitos no meio de um conjunto de dados de dimensão considerável. Isto comprova que a utilização destas metodologias pode contribuir significativamente para melhorar a eficácia e eficiência dos sistemas de controlo, prevenção e deteção do branqueamento de capitais e da fraude em geral.

O branqueamento de capitais é um problema cada vez mais prevalente, e várias instituições e órgãos governamentais têm incutido a necessidade de estabelecer normas e politicas que forcem as instituições financeiras a monitorizar proativamente este fenómeno, de forma a prevenir a lavagem de dinheiro, a fraude e a corrupção. No entanto, ainda existem demasiados casos de instituições que se referem a este tipo de políticas e processos como demasiado custosos e pouco eficientes. Daí a necessidade de que continuem a existir estudos e trabalhos a esplorar possibilidades que possam contribuir para um melhor rendimento dos controlos impostos. É também necessário estudar métodos que sejam capazes de ser flexíveis e acompanhar, se não mesmo prever, as novas tendências comportamentais dos infratores, para que sejam mitigadas o mais cedo possível. Este tem de ser um tema presente na gestão das

instituições financeiras, sendo urgente uma renovação tecnológica na abordagem ao tema, o que torna importantes trabalhos como este, pois ao demonstrar os resultados possíveis com base em metodologias que recorram a simulação e tratamento de dados automáticos, com recurso a modelos computacionais e inteligência artificial, estão a contribuir para essa mudança necessária.

6. Bibliografia

Abe, N., Zadrozny, B., & Langford, J. (2006). Outlier detection by active learning. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 504–509). ACM.

Acts, I. F. (2000). University of Houston System Administrative Memorandum.

Axelsson, S., & Lopez-Rojas, E. A. (2012). Multi agent based simulation (MABS) of financial transactions for anti money laundering (AML). In *Nordic Conference on Secure IT Systems*. Blekinge Institute of Technology.

Baesens, B., Viaene, S., Van Den Poel, D., Vanthienen, J., & Dedene, G. (2002). Bayesian neural network learning for repeat purchase modelling in direct marketing. *European Journal of Operational Research*, *138*(1), 191–211.

Bala, V., & Goyal, S. (1998). Learning from Neighbours. *The Review of Economic Studies*, *65*(3), 595–621.

Bala, V., & Goyal, S. (2001). Conformism and diversity under social learning. *Economic Theory*, *17*(1), 101–120.

Barse, E. L., Kvarnstrom, H., & Johnson, E. (2003). Synthesizing test data for fraud detection systems. In IEEE (Ed.), *19th Annual Computer Security Applications Conference, 2003. Proceedings.* (pp. 384–394).

Bartlett, B. L. (2002). The negative effects of money laundering on economic development. *Asian Development Bank Regional Technical Assistance Project*, *5967*.

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., … Wiswedel, B. (2009). KNIME - the Konstanz information miner. *ACM SIGKDD Explorations Newsletter*, *11(1)*, 26–31.

Bishop, T. J., Bloom, C. A., Carcello, J. V., & Cotton, D. L. (2008). Managing the business risk of fraud: a practical guide. *The Institute of Internal Auditors, The American Institute of Certified Public Accountants & Association of Certified Fraud Examiners*, *5–7*.

Bolton, R. J., Hand, D. J., Provost, F., Breiman, L., Bolton, R. J., & Hand, D. J. (2002). Statistical Fraud Detection: A Review. *Statistical Science*, *17*(3), 235–249.

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*(suppl. 3), 7280–7287.

Breiman, L. (2001). Random forests. *Machine Learning*, (45(1)), 5–32.

Buchanan, B. (2004). Money laundering - A global obstacle. *Research in International Business and Finance*.

Button, M., Gee, J., & Brooks, G. (2012). Measuring the cost of fraud: an opportunity for the new competitive advantage. *Journal of Financial Crime*, *19*(1), 65–75.

Casti, J. L. (1997). *Would-be worlds: how simulation is changing the world of science*. New York: Wiley.

Colladon, A. F., & Remondi, E. (2017). Using Social Network Analysis to Prevent Money Laundering. *Expert Systems with Applications*, *67*, 49–58.

Dheepa, V., & Dhanapal, R. (2009). Analysis of Credit Card Fraud Detection Methods. *International Journal of Recent Trends in Engineering*, *2*(3), 126–128.

Ellison, G. (1993). Learning, Local Interaction, and Coordination. *Econometrica: Journal of the Econometric Society*, *61*(5), 1047–1071.

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications*, *29*(4), 927–940.

Ezawa, K., Singh, M., & Norton, S. (1996). Learning Goal-Oriented Bayesian Networks for Telecommunications Risk Management. *13th International Conference on Machine Learning*, 139–147.

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. *JasssThe Journal Of Artificial Societies And Social Simulation*.

Friedman, N., Geiger, D., Goldszmidt, M., Provan, G., Langley, P., & Smyth, P. (1997). Bayesian Network Classifiers. *Machine Learning*, *29*, 131–163.

Gama, J., Carvalho, A. P. D. L., Faceli, K., Lorena, A. C., & Oliveira, M. (2015). *Extração de Conhecimento de Dados: Data Mining* (2nd ed.). Lisbon: Edições Sílabo.

Garcia, R. (2005). Uses of agent-based modeling in innovation/new product development research. *Journal of Product Innovation Management*, *22*(5), 380–398.

Hand, D. J., Blunt, G., Kelly, M. G., & Adams, N. M. (2000). Data Mining for Fun and Profit. *Statistical Science*, *15*(2), 111–131.

Heckerman, D. (1997). Bayesian Networks for Data Mining. *Data Mining and Knowledge Discovery*, *1*(1), 79–119.

Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., Sanchez-Esguevillas, A., Lloret, J., … Cook, D. (2013). A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants. *IEEE Communications Magazine*.

Hogsett III, R. M., & Radig, W. J. (1994). Employee rime: the cost and some control measures. *Review of Business*, *16*(2), 9.

Huang, W., An, Y., & Du, W. (2010). A Multi-Agent-based Distributed Intrusion Detection System. *Advanced Computer Theory and Engineering (ICACTE)*, *3*, 141–143.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence*, *117*(2), 277–296.

Jiao, J. R., You, X., & Kumar, A. (2006). An agent-based framework for collaborative negotiation in the global manufacturing supply chain network. *Robotics and Computer-Integrated Manufacturing*, *22*(3), 239–255.

Levi, M., & Gilmore, W. (2002). Terrorist finance, money laundering and the rise and rise of mutual evaluation: a new paradigm for crime control? In *Financing terrorism* (pp. 87–114). Springer Netherlands.

Levi, M., & Reuter, P. (2006). Money Laundering. In *Crime and Justice* (Vol. 34(1), pp. 289–375). Chicago: The University of Chicago Press Books.

Liaw, a, & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, (2(3)), 18–22.

Lopez-Rojas, E. A. (2014). *On the Simulation of Financial Transactions for Fraud Detection Research*. Bleekinge Institute of Technology.

Macal, C. M., & North, M. J. (2005). Tutorial on agent-based modeling and simulation. In *Proceedings of the 37th conference on Winter simulation* (pp. 2–15). Winter Simulation Conference.

Magnusson, D. (2009). The costs of implementing the anti-money laundering regulations in Sweden. *Journal of Money Laundering Control*, (12(2)), 101–112.

Mellouli, S., Mellouli, S., Moulin, B., Moulin, B., Mineau, G., & Mineau, G. (2004). Laying Down the Foundations of an Agent Modelling Methodology for Fault-Tolerant Multi-agent Systems. In *International Workshop on Engineering Societies in the Agents World* (pp. 275–293). Springer Berlin Heidelberg.

Mitchell, T. M. (1997). *Machine Learning. McGraw-Hill.*

Morris, S. (2000). Contagion. *The Review of Economic Studies*, *67*(1), 57–78.

Neil, M., Fenton, N., & Tailor, M. (2005). Using Bayesian networks to model expected and unexpected operational losses. *Risk Analysis*, *25*(4), 963–972.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review

of literature. *Decision Support Systems*, *50(3)*, 559–569.

Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, *52*(2), 464–473.

Panigrahi, S., Kundu, A., Sural, S., & Majumdar, A. K. (2009). Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. *Information Fusion*, *10*(4), 354–363.

Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A Comprehensive Survey of Data Mining-based Fraud Detection Research. *Monash University*.

Pimenta, C. (2009). *Esboço de Quantificação da Fraude em Portugal. Working Papers - OBEGEF*.

Pradhan, M., & Dagum, P. (1996). Optimal Monte Carlo Estimation of Belief Network Inference. In *Proceedings of the Twelfth international conference on Uncertainty in artificial intelligence* (pp. 446–453). Morgan Kaufmann Publishers Inc.

Prodromidis, A. L., & Stolfo, S. J. (1999). Agent-based Distributed Learning Applied to Fraud Detection. *Sixteenth National Conference on Artificial Intelligence*.

Quirk, P. J. (1997). Money laundering: Muddying the macroeconomy. *Finance & Development*, *34*(1), 7–9.

RStudio Team. (2015). RStudio: Integrated Development for R. *RStudio, Inc., Boston, MA*. Retrieved from http://www.rstudio.com/

Russell, S., Binder, J., & Koller, D. (1995). Local learning in probabilistic networks with hidden variables. *IJCAI*, *95*, 1146–1152.

Schnatterly, K. (2003). Increasing firm value through detection and prevention of white-collar crime. *Strategic Management Journal*, *24*(7), 587–614.

Shamshirband, S., Anuar, N. B., Kiah, M. L. M., & Patel, A. (2013). An appraisal and design of a multi-agent system based cooperative wireless intrusion detection computational intelligence technique. *Engineering Applications of Artificial Intelligence*, *26*(9), 2105–2127.

Shaw, M. J., Subramaniam, C., Tan, G. W., & Welge, M. E. (2001). Knowledge management and data mining for marketing. *Decision Support Systems*, *31*(1), 127–137.

Sudjianto, A., Nair, S., Yuan, M., Zhang, A., Kern, D., & Cela-Díaz, F. (2010). Statistical methods for fighting financial crimes. *Technometrics*, (52(1)), 5–19.

Sun, L., & Shenoy, P. P. (2007). Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operational Research*, *180*(2), 738–753.

Tuyls, K., Maes, S., & Vanschoenwinkel, B. (2002). Credit card fraud detection using Bayesian

and neural networks. In *Proceedings of the 1st international NAISO congress on euro fuzzy technologies* (pp. 261–270).

Wilensky, U. (1999). NetLogo. *Center for Connected Learning and Computer Based Modeling*. Evanston, IL: Northwestern University. Retrieved from http://ccl.northwestern.edu/netlogo/

Wilhite, A. (2006). Economic Activity on Fixed Networks. In L. Tesfatsion & K. L. Judd (Eds.), *Handbook of Computational Economics* (Vol. 2, pp. 1013–1045). Amsterdam: Elsevier.

Young, H. P. (1993). The Evolution of Conventions. *Econometrica*, *61*(1), 57–84.

Young, H. P. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. *Princeton University Press* (Vol. 110).

Young, H. P. (2005). The Diffusion of Innovations in Social Networks. *The Economy as an Evolving Complex System*, *III*(1966), 267–282.

Yue, D., Wu, X., Wang, Y., Li, Y., & Chu, C. H. (2007). A review of data mining-based financial fraud detection research. In *International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2007* (pp. 5519–5522). IEEE.

Zhang, G. P. (2000). Neural networks for classification: a survey. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, (30(4)), 451–462.

Zhang, Z. (Mark), Salerno, J. J., & Yu, P. S. (2003). Applying data mining in investigating money laundering crimes. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 747–752). ACM.

Zhu, X. D., Huang, Z. Q., & Zhou, H. (2006). Design of a multi-agent based intelligent intrusion detection system. *2006 1st International Symposium on Pervasive Computing and Applications, Proceedings*, 290–295.

7. Anexos

7.1 Código do Modelo de Simulação em Netlogo

Os nomes das variáveis, os comentários e outros tipos de texto encontram-se na língua Inglesa, para facilitar a partilha de conteúdos na comunidade "NetLogo User Community Models".

```
extensions [nw csv]

globals [
  counter
  list-of-recipients
]

breed [ individuals individual ]
breed [ businesses business ]
breed [ containers container ]

undirected-link-breed [connections connection]
undirected-link-breed [translines transline]

containers-own [
  transfer
]

turtles-own [
  predisposition
  gender ;; 0-individual  1-profit  2-non profit  3-shell  4-trust
  location
  nationality
  watch-list
```

```
  neighbs
  trans-destination
  trans-type
  total-trans
  n-trans

]

;;;;;;;;;;;;;;;;;;;;;;;; Initialization ;;;;;;;;;;;;;;;;;;;;;;;;

to setup
  clear-all
  initialize-variables
  ask patches [set pcolor black]
  set-default-shape individuals "person"
  set-default-shape businesses "house"
  set-default-shape containers "container"
  create-turtles 1 [set color yellow]
  create-individuals people [setup-individuals]
  create-businesses num-businesses [setup-businesses]
  create-containers num-containers [setup-containers]
  ask turtle 0 [die]
  setup-connections

  reset-ticks

  if (file-exists? "TransactionData.csv" = true)
  [
    file-close-all
    file-delete "TransactionData.csv"
  ]
  file-open "TransactionData.csv"
```

```
  file-print (word "SOURCE,SOURCE-TYPE,SRC-NATIONALITY,SRC-LOCATION,SRC-
WATCH-LIST,DESTINATION,DESTINATION-TYPE,DT-NAT,DT-LOC,DT-WATCH-
LIST,TRANSACTION-AMOUNT,DAY-OF-TRANS,FRAUD-SUSPECT")
  file-flush

end

to setup-individuals
  set color blue
  setxy random-xcor random-ycor
  ifelse (random-float 100 < perc-fraud-agents)
  [set predisposition (random-float 40) + 60]
  [set predisposition (random-float 70)]
  set gender 0
  set neighbs (random-exponential 2) + 1              ;; number of neighbors is defined through a
exponencial distribution with mean 2
  ifelse (random-float 100 < perc-watch-list or predisposition > 95)              ;; Agents of the
defined percentage are on the watch list
    [set watch-list "Yes"]
    [set watch-list "No"]
  ifelse (random-float 100 < 60)
    [set nationality "Portuguese"]              ;; Portuguese
    [set nationality "Foreign"]              ;; Foreign
  ifelse (random-float 100 < 60)
    [set location "Portugal"]              ;; Portugal
    [set location "Other"]              ;; Other

  set total-trans 0
  set n-trans 0

end
```

```
to setup-businesses
  set color green
  setxy random-xcor random-ycor
  ifelse (random-float 100 < perc-fraud-agents)
  [set predisposition (random-float 40) + 60]
  [set predisposition (random-float 70)]
  set gender (random 3) + 1
  set neighbs (random-exponential 2) + 1          ;; number of neighbors is defined through a exponencial distribution with mean 2
  ifelse (random-float 100 < perc-watch-list or predisposition > 95)          ;; Agents of the defined percentage are on the watch list
    [set watch-list "Yes"]
    [set watch-list "No"]
  ifelse (random-float 100 < 40)
    [set nationality "Portuguese"]          ;; Portuguese
    [set nationality "Foreign"]          ;; Foreign
  ifelse (random-float 100 < 40)
    [set location "Portugal"]          ;; Portugal
    [set location "Other"]          ;; Other

   set total-trans 0
   set n-trans 0
end

to setup-containers
  set color brown
  setxy random-xcor random-ycor
  ifelse (random-float 100 < perc-fraud-agents)
  [set predisposition (random-float 40) + 60]
  [set predisposition (random-float 70)]
  set transfer (random 1)
```

```
  set neighbs (random-exponential 2) + 1          ;; number of neighbors is defined through a
exponencial distribution with mean 2
  ifelse (random-float 100 < perc-watch-list or predisposition > 95)
    [set watch-list 1]
    [set watch-list 0]
  ifelse (random-float 100 < 30)
    [set nationality 0]                ;; Portuguese
    [set nationality 1]                ;; Foreign
      ifelse (random-float 100 < 30)
    [set location 0]                   ;; Portugal
    [set location 1]                   ;; Other
   ifelse (random-float 100 < perc-informal-cont)
    [set transfer 0]                   ;; Bank transfer
    [set transfer 1]                   ;; Informal transfer
end

to setup-connections
  ask containers [
    create-translines-with n-of (0.4 * num-containers) other containers
  ]
  ask translines [set color white]
  ask turtles with [breed != containers]
  [
    create-translines-with n-of ((random 4) + 1) other containers                    ;; businesses and
individuals have 1 to 4 means of transfer
    if (predisposition > 70)
    [
      if(count transline-neighbors with [ predisposition > 70] < 1)          ;; if presdisposition is
higher than 70 -> has a connection with another agent with that kind of predisposition
      [
        ask self [create-transline-with one-of other containers with [predisposition > 70] ]
      ]
```

```
    ]
  ]
  ask individuals [create-connections-with n-of neighbs other (turtles with [breed != containers])]
  ask businesses [create-connections-with n-of neighbs other (turtles with [breed != containers])]
  ask individuals with [predisposition > 70] [create-connection-with one-of other businesses with
[gender > 2]]    ;; gender 3 and 4 are often related with fraud
  ask businesses with [predisposition > 70] [create-connection-with one-of other businesses with
[gender > 2]]
  ask individuals with [predisposition > 90] [                                    ;; an agent
with an extremely high predisposition normally has
    create-connection-with one-of other businesses with [gender > 2]                ;;
more connections with the same type of fraudulent agents
    create-connection-with one-of other individuals with [predisposition > 60]
  ]
  ask businesses with [predisposition > 90] [
    create-connection-with one-of other businesses with [gender > 2]
    create-connection-with one-of other individuals with [predisposition > 60]
  ]

 ask connections [set color black]

end

to initialize-variables
  set counter 0
end

to go
  do-business
  tick
end
```

```
;;;;;;;;;;;;;;;;;;;;;; Do Transactions ;;;;;;;;;;;;;;;;;;;;

to do-business  ;;turtle procedure
 nw:set-context turtles links
 ask turtles
 [
  set trans-destination 0
  set trans-type 0
 ]
 let %t 0.05                    ;; 5% of the agents are asked to make a transfer
 ask (n-of (%t * count turtles with [breed != containers]) turtles with [breed != containers])
 [
  pick-destination
  plan-transaction

 ]

end

to pick-destination
 ifelse (random-float 100 < 10)     ;; 10% of the time the destination chosen is not part of the agents network (to introduce some randomness)
 [ifelse (random-float 100 < 50)
  [
   set trans-destination n-of 1 (other turtles with [breed != containers and not connection-neighbor? myself])
  ]
  [
   ifelse([predisposition] of self > 80)
    [
```

```
      set trans-destination n-of 1 (other turtles with [breed != containers and not connection-neighbor? myself])
     ]
     [
      set trans-destination n-of 1 (other turtles with [breed != containers and predisposition > [predisposition] of myself  and not connection-neighbor? myself])
     ]
    ]
  ]
  [set trans-destination n-of 1 connection-neighbors]
end

to plan-transaction
  let amount precision ((random 100000) + 1) 3            ;; transfer amount between 1 and 100000
  let idd [who] of trans-destination     ;; id of destination
  let ids [who] of self                  ;; id of source
  set trans-type 0

  if (([predisposition] of turtle (item 0 idd) < 60) or ([predisposition] of turtle ids < 60))       ;; if predisposition is low -> normal transfer
  [
    set trans-type 0 ;;basic
    ask turtle ids
    [
     if(count connection-neighbors with [ breed = containers and transfer < 1] < 1)
     [
      ask turtle ids [create-connection-with one-of other containers with [transfer < 1] ] ;; if it does not have a connection with a container, it gets one
     ]

    ]
    nw:set-context turtles links
```

```
  let id1 [who] of one-of connection-neighbors with [breed = containers and transfer < 1]
  set list-of-recipients lput id1 idd
  set total-trans total-trans + amount
  set n-trans n-trans + 1

  show (word "Regular transfer from " ids word " to " (item 0 list-of-recipients) word ", with container " (item 1 list-of-recipients))
  file-print (word ids "," [gender] of turtle ids "," [nationality] of turtle ids "," [location] of turtle ids "," [watch-list] of turtle ids "," (item 0 idd) ","
    [gender] of turtle (item 0 idd) "," [nationality] of turtle (item 0 idd) "," [location] of turtle (item 0 idd) "," [watch-list] of turtle (item 0 idd) ","
    amount "," (random 365 + 1) ",NO")

 ]
 if ([predisposition] of turtle (item 0 idd)) > 70 and ([predisposition] of turtle ids > 70)    ;; both have high predisposition
 [
  ifelse (amount < 10000)     ;; low amount -> Informal transfer
  [
    set trans-type 1
    ask turtle ids
    [
     if(count connection-neighbors with [ breed = containers and transfer > 0] < 1)
     [
       ask turtle ids [create-connection-with one-of other containers with [transfer > 0] ] ;; if it does not have a connection with a container, it gets one
      ]

    ]

    let id1 [who] of one-of connection-neighbors with [ breed = containers and transfer > 0 ]
    set list-of-recipients lput id1 idd
```

```
    set total-trans total-trans + amount
    set n-trans n-trans + 1

    show (word "Informal Transfer from " ids word " to " (item 0 list-of-recipients) word ", via
" (item 1 list-of-recipients))
    file-print (word ids "," [gender] of turtle ids "," [nationality] of turtle ids "," [location] of turtle
ids "," [watch-list] of turtle ids "," (item 0 idd) ","
    [gender] of turtle (item 0 idd) "," [nationality] of turtle (item 0 idd) "," [location] of turtle
(item 0 idd) "," [watch-list] of turtle (item 0 idd) ","
    amount "," (random 365 + 1) ",NO")
  ]
  [
    set trans-type 2  ;; structered
    ask turtle ids
    [
     let n-otherneigh count (connection-neighbors with [ who != (item 0 idd) ])
     if(n-otherneigh < 1)
     [
       ask turtle ids [create-connection-with one-of other turtles with [breed != containers and
predisposition > 50 and not connection-neighbor? myself] ]
       ask turtle ids [create-connection-with one-of other turtles with [breed != containers and
predisposition > 50 and not connection-neighbor? myself] ]
     ]
     let n-recipients (random n-otherneigh) + 1

     set list-of-recipients [who] of n-of n-recipients (connection-neighbors with [who != (item
0 idd)])
     show (word "Structured transfer initiated from: " ids word " to " (item 0 idd) word ", for "
amount word ", through one or more of the following: " list-of-recipients)

    nw:set-context turtles connections
```

```
let partial-amount 0
let amount-left amount
set counter 1
let first_day random 365 + 1

while [amount-left > 0]
[
  set partial-amount (precision ( 10000 * ((random-float 0.9) + 0.1) ) 0)
  set amount-left ( amount-left - partial-amount )
  set total-trans total-trans + partial-amount
  set n-trans n-trans + 1
  let transfer_day first_day + (random 60)

  let target (item 0 list-of-recipients)
  set list-of-recipients shuffle list-of-recipients
  let source-breed [breed] of turtle ids

  show (word "Structured transfer -  part ( " counter word ") - " partial-amount word " from
" ids word " to " target word ", to reach " (item 0 idd) )

  file-print (word ids "," [gender] of turtle ids "," [nationality] of turtle ids "," [location] of
turtle ids "," [watch-list] of turtle ids "," target ","
  [gender] of turtle target "," [nationality] of turtle target "," [location] of turtle target ","
[watch-list] of turtle target ","
  partial-amount "," transfer_day ",YES")
   ask turtle target
   [
    ifelse ( (length nw:turtles-on-path-to turtle (item 0 idd)) <= 2 )
    [
      set total-trans total-trans + partial-amount
      set n-trans n-trans + 1
      set transfer_day first_day + (random 60)
```

```
            show nw:turtles-on-path-to turtle (item 0 idd)
            show (length nw:turtles-on-path-to turtle (item 0 idd))
            show (word "[" counter word "] Further transfer - " partial-amount word " from " target word " to " (item 0 idd) )

            file-print (word target "," [gender] of turtle target "," [nationality] of turtle target "," [location] of turtle target "," [watch-list] of turtle target "," (item 0 idd) ","
            [gender] of turtle (item 0 idd) "," [nationality] of turtle (item 0 idd) "," [location] of turtle (item 0 idd) "," [watch-list] of turtle (item 0 idd) ","
            partial-amount "," transfer_day ",YES")
          ]
          [
            let path nw:turtles-on-path-to turtle (item 0 idd)
            let idm [who] of (item 0 path)

            set total-trans total-trans + partial-amount
            set n-trans n-trans + 1
            set transfer_day first_day + (random 60)
            show (word "[" counter word "] Further transfer - " partial-amount word " from " idm word " to " (item 0 idd) )
            file-print (word idm "," [gender] of turtle idm "," [nationality] of turtle idm "," [location] of turtle idm "," [watch-list] of turtle idm "," (item 0 idd) ","
            [gender] of turtle (item 0 idd) "," [nationality] of turtle (item 0 idd) "," [location] of turtle (item 0 idd) "," [watch-list] of turtle (item 0 idd) ","
            partial-amount "," transfer_day ",YES")
          ]
         ]

        set counter counter + 1
       ]
     ]
   ]
```

```
  ]

  set counter 0
  file-flush

end
```

FACULDADE DE **ECONOMIA**