Modelagem analítica e técnicas avançadas para a PLD/FT

"Um modelo é sempre parcial, mas oferece recursos para o avanço do conhecimento"
Jean-Pierre Changeux[73]

Esta seção descreve algumas das tendências e práticas mais inovadoras do setor baseadas em modelagem analítica e técnicas avançadas para a identificação, gerenciamento, controle e supervisão da lavagem de dinheiro.

O contexto para a abordagem analítica da avaliação da PLD/FT

Com o surgimento de uma regulamentação mais restritiva, visando uma identificação mais rápida e melhor dos riscos, e novas tecnologias disponíveis, as instituições financeiras estão se movendo ao longo de uma nova jornada transformadora em relação à implementação da adoção de análises avançadas de PLD/FT[74]. As três principais ferramentas usadas para detectar a lavagem de dinheiro incluem a avaliação do de risco do cliente, o monitoramento das transações e as regras de *screening* de sanções.

Avaliação do risco do cliente

A avaliação do risco do cliente é um modelo baseado nos fatores de risco associados à identificação da lavagem de dinheiro, tais como país do cliente, ocupação e salário, produtos bancários etc.

Os modelos estatísticos se tornaram a prática principal para a classificação de risco do cliente, através da aplicação de diferentes técnicas para resolver o problema da detecção de anomalias. Entretanto, este problema é complexo para identificar ou reproduzir, e produz amostras desbalanceadas.

A aplicação de métodos avançados de dados nos permite superar estas limitações, melhora a precisão da avaliação do risco do cliente e fomenta sua relevância ao longo do programa de PLD/FT. A avaliação do risco do cliente evolui progressivamente para uma avaliação do risco comportamental do cliente na qual os dados contínuos são atualizados e enriquecem o processo de identificação de riscos[75]. Além disso, os próprios modelos estão incorporando o benefício do uso de técnicas de *machine learning*. Métodos supervisionados, tais como o *random forest*, são os primeiros a serem implementados para desvendar relações ocultas entre os fatores de risco em um conjunto aumentado de fatores.

À medida em que o poder computacional, a riqueza e a profundidade dos dados aumentam, estes modelos comportamentais também podem incorporar gatilhos para a estruturação de transações potenciais, ou seja, estratégias coletivas para lavar dinheiro por vários indivíduos através de pequenas quantias, para evitar a detecção por estratégias clássicas de detecção estática. A capacidade de construir algoritmos e estratégias que funcionam não com base em um cliente individual ou cliente mais transação, mas em conjuntos de clientes, permite a identificação da estruturação da transação de forma mais proativa e eficaz. Estes chamados algoritmos gráficos[76,77] aproveitam as conexões potenciais provenientes de diferentes fontes de informação[78]. Além disso, a capacidade de construir uma representação de rede abrangente de todos os clientes traz o valor adicional de racionalizar o processo de investigação de alerta, entre outros.

Monitoramento de transações

A abordagem mais comum de monitoramento de transações consiste em um sistema baseado em regras, no estilo de uma árvore de decisão. Cada regra é configurada para identificar um comportamento definido desmascarando as atividades LD potenciais dos clientes e entidades envolvidas na transação[79].

[73] Jean-Pierre Changeux (b.1936)) é um neurocientista francês conhecido por suas pesquisas em vários campos da biologia, desde a estrutura e função das proteínas, ao desenvolvimento precoce do sistema nervoso, até as funções cognitivas.

[74] Entretanto, não há uniformidade no grau de adoção destas técnicas de análise avançadas. Enquanto algumas entidades financeiras estão experimentando soluções inovadoras, aplicações simples são mais comuns no setor, e a confiança no apoio analítico está no início para outras. No entanto, o presente e o futuro dos programas PLD/FT não podem ser entendidos sem se olhar para as novas tecnologias e metodologias disponíveis.

[75] Por exemplo, incorporação de informações de monitoramento de transações, triagem de pagamentos ou análise outlier em torno de canais, volumes, geolocalização, etc.

[76] Soltani, Reza & Nguyen, Uyen & Yang, Yang & Faghani, Mohammad & Yagoub, Alaa & An, Aijun. (2016. 1-7. 10.1109/UEMCON.2016.7777919

[77] Aprendizagem de gráficos escalonáveis para a luta contra a lavagem de dinheiro: A First Look; Weber, M; Chen, J.; Suzumura, T.; Pareja, A.; Ma, T.; Kanezashi, H., Kaler, T.; Leisersen C.E.; Schardl, Tao B

[78] Por exemplo, circuitos fechados de transacionalidade - transferências regulares -, para a propriedade conjunta de contas, endereço único, filial de escolha ou a maioria das filiais visitadas ou caixas eletrônicos, geoposicionamento via aplicativos móveis, coincidência de comerciantes, etc.

[79] Este comportamento suspeito será muito provavelmente baseado em valores anômalos em termos de localização, contagem de transações ou os valores das mesmas.

Essas regras são geralmente identificadas como "cenários". Regras e cenários mais complexos tentam abordar a identificação de contas aninhadas e relações mais sofisticadas entre as partes, mas a base da identificação anterior permanece, de modo geral, no nível da transação individual, analisando os dados recebidos durante o processo transacional. Quando um outlier é identificado, um alerta é acionado, o que posteriormente requer uma avaliação de um especialista[80].

Neste processo, o conjunto inicial de regras é dividido em uma segmentação mais profunda dos comportamentos nos quais a linha de negócios, o nível de atividade transacional e a avaliação de risco do cliente determinam os outliers comportamentais finais, ou seja, os alertas que seriam acionados.

Os métodos analíticos de dados podem ser aproveitados para detectar mais alertas de qualidade, aumentando os verdadeiros positivos e reduzindo os falsos negativos, ou seja, mais alertas verdadeiros são identificados sem aumentar o ruído na identificação. As técnicas de análise de dados e de *machine learning* são implementadas para otimizar a segmentação proporcionando uma identificação mais precisa dos padrões graças à exploração de dados históricos[81].

Entretanto, as entidades financeiras que procuram ativamente incorporar métodos avançados em seu programa PLD/FT podem decidir concentrar-se na priorização de alerta. A abordagem de regras gera grandes quantidades de alertas mesmo quando o ajuste adequado dos limites do cenário é implementado, e a segmentação foi otimizada. Para resolver isto, muitos bancos implementam métodos de aprendizagem supervisionados para classificar os alertas em termos de produtividade[82]. O aspecto chave que determina o sucesso desta abordagem é a utilização de métricas diferenciais, além das variáveis esperadas e inamovíveis disponíveis no nível da transação.

A abordagem mais disruptiva para a identificação de riscos de PLD/FT consiste em abandonar a abordagem tradicional de regras individuais para revelar uma relação oculta com análises avançadas. Entretanto, poucas instituições financeiras estão explorando a utilização de metodologias alternativas. Algumas delas são:

- A análise gráfica, que ocupa seu espaço na identificação das relações de rede e é cada vez mais determinante para as atividades de lavagem no mundo financeiro interconectado
- Técnicas de *clustering*, que ajudam a identificar os valores atípicos sem assumir comportamentos específicos; portanto, capturando com mais frequência novas atividades ilícitas potenciais.

Avançar para uma abordagem não baseada em regras não implica automaticamente o abandono de boas práticas de otimização previamente identificadas. De fato, a confiança em análises avançadas para melhorar a segmentação do cliente, combinada com a detecção de rede e outliers, e a utilização de priorização de alerta pode ser vista como uma solução integral.

Screening de sanções

Os motores de *screening* de sanções comparam indivíduos ou empresas contra uma lista de sanções designada, utilizando técnicas de coincidência difusa. As abordagens mais simples são baseadas em uma ampla gama de transformações aplicadas aos "nomes" (mudança de ordem de nomes, iniciais, transliteração,

[80] Ver Scalable Graph Learning for Anti-Money Laundering: A First Look; Weber, Chen, Suzumura, Pareja, Ma, Kanezashi, Kaler, Leisersen Schardl, Tao.

[81] O ajuste do limiar de dados permite otimizar os baldes de aumento de produtividade ao longo das variáveis utilizadas nos cenários (mais verdadeiros positivos), enquanto fornece medidas do risco potencial não identificado (limitando os falsos negativos). Estas abordagens comuns se baseiam nos motores existentes baseados em regras.

[82] Esta abordagem pode ser vista como uma imitação da revisão dos alertas por analistas de nível 1; no entanto, esta poderia ser uma identificação mais complexa a ser abordada e nem todas as entidades têm sucesso neste esforço.

Um exemplo de avaliação nacional de riscos

O governo britânico publica regularmente uma avaliação nacional de riscos[1], que informa sobre os riscos de crimes financeiros enfrentados a nível nacional. Através desta avaliação nacional de riscos são incluídas referências sobre as técnicas mais habituais utilizadas na LD/FT e seu nível de implantação no país e são uma referência importante para as próprias instituições em sua avaliação do risco.

.

Uma empresa deve realizar uma avaliação de risco de Crimes Financeiros e usá-la para informar o projeto de seus controles de PLD/FT. A avaliação de risco nacional serve, portanto, como uma base sólida para construir esta avaliação, com a empresa tomando medidas extras para compreender, mais especificamente, os riscos que eles enfrentam.

Isto levaria em conta, mas não limitado à sua carteira de clientes e aos produtos que eles têm - contas correntes pessoais servem como meio de evasão fiscal para muitas pequenas empresas, bem como a introdução de exposição a muitas outras técnicas de lavagem de dinheiro devido à sua capacidade de rápidas transferências de fundos e aceitação de transações em dinheiro. Além disso, uma revisão da atividade criminal histórica pode ajudar a entender quaisquer tipologias adicionais enfrentadas pelo banco.

As transações de dinheiro que entram e saem das contas, servem como uma maneira fácil para os lavadores de dinheiro quebrarem os rastros das transações. Embora o uso de dinheiro em espécie na lavagem de dinheiro seja generalizado e esteja incluído em muitas das estratégias utilizadas, os controles em torno dos riscos de dinheiro em espécie são geralmente os mais simples, em grande parte devido a pouca informação disponível para transações em espécie.

As mulas de dinheiro são terceiros que são usadas, consciente ou inconscientemente para fazer transações adicionais em dinheiro e transferências de fundos que mascaram os rastros de transações. Isto pode ser usado em conjunto com outras estratégias, por exemplo, compra de ativos de alto valor e revenda, para remover quase completamente as suspeitas sobre a origem dos fundos, onde as contas temporárias poderiam ser as de uma rede de mula. Isto é difícil de detectar usando métodos tradicionais, pois nenhuma conta única, e nenhum cliente único, pode jamais ser usado para grandes volumes das transações usadas em qualquer etapa deste processo.

Da mesma forma, os negócios com uso intensivo de dinheiro servem como outro desafio para os métodos tradicionais de detecção. Negócios como salões de beleza, bancas de jornais e lavadoras de carros são utilizados por lavadores de dinheiro para documentar o dinheiro proveniente de atividades criminosas como receita comercial legítima, de modo que grandes volumes de fundos ilícitos das redes criminosas possam ser centralizados em uma conta. Isto se mostra difícil de detectar, pois a renda em dinheiro da empresa pode parecer consistente com sua própria história, bem como com a renda de seus pares e, portanto, pode não haver suspeitas levantadas pelas transações em dinheiro da empresa. Esses negócios, no entanto, são normalmente também ligados ao tráfico de pessoas e à escravidão moderna, que incluem seus próprios comportamentos transacionais que podem ser mais fáceis de detectar. Como no caso do uso de mulas de dinheiro, estas tipologias geralmente envolvem uma rede de terceiros aparentemente não relacionados. Estes terceiros podem ser os facilitadores ou mesmo as vítimas destes crimes e, portanto, há comportamentos específicos que se espera ver. Transações em múltiplas cidades diferentes, especialmente em cidades com centros de transporte, uso intensivo de restaurantes de fast-food, múltiplas transações no mesmo hotel no mesmo dia, múltiplos pagamentos a provedores de telefonia móvel, transferências de fundos entre contas com comportamentos similares e transações internacionais, especialmente em dinheiro e transferências de fundos, são todos fortes indicadores destas tipologias. Se estas partes puderem ser vinculadas ao negócio de caixa intensivo, então a rede completa poderá ser descoberta.

As transações internacionais são outras operações de alto risco identificadas na avaliação de risco nacional. Elas são vistas em uma variedade de técnicas de lavagem de dinheiro, bem como apresentam um risco em outros aspectos da criminalidade financeira. Isto é visto no tráfico de pessoas, que é estimado como um dos maiores geradores de receita criminal em todo o mundo. O tráfico de pessoas requer o envio para o exterior de membros da quadrilha de crime organizado associada nos países associados ao tráfico. Isto pode ser como dinheiro levantado no Reino Unido e movimentado fisicamente para o exterior ou através de mulas de dinheiro de forma semelhante ao comportamento associado aos depósitos em dinheiro descritos anteriormente.

O financiamento do terrorismo é identificado como uma tipologia de alto risco dentro do Reino Unido. A captação e movimentação de recursos não é considerada um objetivo primordial dos terroristas, especialmente porque a maioria dos recentes ataques terroristas tem sido de baixo orçamento e baixa sofisticação, frequentemente planejados, financiados e feitos por um indivíduo. O financiamento do terrorismo é comumente utilizado para mover fundos para o exterior através de métodos relativamente simples, tais como a movimentação física de dinheiro para o exterior ou o emprego de empresas de serviços monetários (MSBs). Portanto, a detecção do financiamento do terrorismo requer uma coleta de indicadores-chave da mesma forma que é requerido para o uso de empresas com uso intensivo de dinheiro na lavagem de dinheiro.

O risco associado aos criptoativos cresce ano após ano à medida que os criptoativos se tornam mais comuns e de fácil acesso, mas os controles em torno deles permanecem relativamente novos com o Reino Unido introduzindo regulamentações em torno do uso de criptoativos para lavagem de dinheiro somente em janeiro de 2020. Gangues criminosas organizadas usam criptoativos para lavagem de dinheiro comprando primeiro os criptoativos com seus fundos ilícitos, potencialmente após um estágio inicial de estratificação, antes de vender os ativos para fornecer uma fonte legal de seus fundos. Além disso, os criptoativos podem ser facilmente movimentados através das fronteiras, permitindo que os criminosos movimentem fundos significativos internacionalmente com facilidade significativa em comparação às moedas fiduciárias.

Isto serve como um exemplo de novos riscos emergentes na criminalidade financeira que representam outro desafio para as instituições desenvolverem e agirem regularmente novos controles para acompanharem as mudanças e desenvolvimentos encontrados pelos lavadores de dinheiro.

[1] HM Treasury: National risk assessment of Money laundering and terrorist financing 2020. December 2020.

erros vocais ou consonantes comuns etc.). Os nomes transformados são padronizados como cadeias e comparados com os nomes da lista de sanções, também padronizados seguindo as mesmas regras. As regras ou lógicas correspondentes medem o grau de separação entre as duas cadeias. O motor pode retornar uma pontuação da correspondência, ou um alerta baseado em uma regra pré-definida de correspondência, entretanto, a lógica subjacente é a mesma, ou seja, as duas cadeias são similares o suficiente para conceder uma revisão especializada.

Como no caso do monitoramento de transações, estas regras produzem muitos falsos positivos[83]. Além disso, o potencial de otimização com base no ajuste é menor do que no caso do monitoramento de transações.

Por esta razão, as instituições estão explorando métodos alternativos para melhorar a qualidade da identificação baseada em tecnologias de tradução e transliteração, e a aplicação de técnicas de processamento de linguagem natural (NLP) para melhorar a correspondência do nome. A melhoria dos métodos analíticos de *screening* de sanções ocorre paralelamente à exploração dessas técnicas na identificação de notícias negativas.

Os próximos passos em abordagens analíticas para a avaliação de PLD/FT

A aplicação de métodos e tecnologias inovadoras não se restringe aos destacados acima. O processamento de linguagem de natureza ampliada e o aprendizado profundo, aplicações em cadeia de bloqueio, verificação eletrônica de identidade, reconhecimento de voz e fala, biometria ou geolocalização são outras tecnologias que podem contribuir para a identificação de atividades ilícitas.

Subjacentes a todas estas abordagens potenciais, podem ser encontradas várias tendências na análise PLD/FT:

- Uma análise mais profunda dos dados existentes tanto no momento da transação, quanto do cliente e seus relacionamentos são implementados. Algumas das opções analíticas descritas acima tornam-se impotentes se os dados diferenciais não estiverem disponíveis e incorporados à análise.

- Dados suplementares das fontes internas e das diferentes dimensões do programa PLD/FT (isto é, classificação de risco do cliente, due diligence, identificação de sanções, transações) e fontes externas (dados públicos sobre PEP, relações de propriedade, fontes reputacionais, buscas abertas) são necessários para criar uma abordagem holística para a identificação de risco PLD/FT.

- As tecnologias e métodos podem ser tão complexos quanto a inovação permite, porém o dimensionamento dos mais adequados à natureza do negócio e a avaliação de risco da instituição é fundamental para otimizar o uso de recursos tecnológicos e humanos e, ao mesmo tempo, garantir a conformidade regulatória.

Os supervisores e reguladores estão em geral relutantes a mudanças repentinas e favorecem metodologias bem estabelecidas antes de abraçar completamente as mudanças revolucionárias. Entretanto, para as instituições dispostas a

[83] Os motores podem ser mais ou menos complexos na incorporação de transformações inovadoras aplicadas a nomes, ou incorporar mais fontes de sanções de qualidade melhoradas com informações PEP, no entanto, todos eles apresentam os mesmos pontos fracos.

embarcar em um programa completo de transformação da análise de PLD/FT, uma série de avanços tem ocorrido nos últimos anos[84]: desde desenvolvimentos específicos de aplicações de coincidências difusas ou detecção de PEP em colaborações conjuntas, até a constituição de centros de inovação e *sandboxes*.

Na jornada rumo a uma identificação de risco mais sofisticada, a interpretabilidade e o controle de risco apropriado permanecem no centro das preocupações do regulador (e das instituições).

O uso de análises avançadas no programa de PLD/FT está vinculado a que as regras implementadas seja consideradas modelos e estejam, portanto, sujeitos às práticas de identificação, monitoramento e controle que as instituições implementaram sob a função de gestão do risco de modelo (MRM). Embora a distinção para a classificação de risco do cliente seja clara, pois cumpre todas as condições tipicamente estabelecidas no *framework* de gestão do risco de modelo (MRM) para ser um modelo ou pelo menos uma ferramenta do usuário que deve ser monitorada, motores de regras de PLD/FT não foram inicialmente vistos como modelos. A assimilação dos motores de PLD/FT na disciplina de gerenciamento de risco do modelo não aconteceu uniformemente entre jurisdições e principais atores querem evitar o peso de um escrutínio incremental dos programas de PLD/FT[85].

No entanto, as tecnologias de *machine learning* para melhorar a identificação de riscos estão ampliando a concepção do que se entende como modelo sujeito ao MRM. Apesar de sua vontade de fomentar sua aplicação aos programas PLD/FT, os supervisores deixam clara a necessidade de garantir um grau adequado de compreensão e interpretabilidade das metodologias implementadas e dos resultados obtidos. Os modelos de caixa preta devem ser evitados. Os modelos de *machine learning* podem sofrer com a falta de transparência na seleção e explicabilidade dos recursos, avaliação do desempenho dos modelos etc. Uma documentação apropriada, teste do modelo, módulos de interpretabilidade; os princípios básicos de um *framework* robusto de MRM apoiarão a adequação desses modelos para o uso de PLD/FT.

Estudo de caso: melhoria a detecção de padrões suspeitos através da análise de redes

Uma das técnicas aplicadas com sucesso para detectar fraudes é a chamada análise de rede. Esta técnica pode ajudar a identificar, detectar e caracterizar comportamentos suspeitos usando métricas, técnicas de *machine learning* e algoritmos fuzzy.

Para desenvolver a análise da rede, há três etapas relevantes a serem consideradas (i) coletar dados relevantes e construir um gráfico que represente os relacionamentos entre as entidades; (ii) decidir sobre a estratégia de identificação que irá identificar o cluster de entidades e relacionamentos suspeitos; e (iii) caracterizar estes clusters através de métricas apropriadas a serem usadas como características dos modelos de detecção (ver figura 4).

Etapa 1. Representação da rede

Uma rede permite o exame de relacionamentos complexos entre entidades relacionadas, seja através de vínculos de dados internos, tais como transações, ou externos, tais como

Figura 4. Etapas para a detecção através da análise de redes.

[84] Nas palavras do recente documento emitido pelo GAFI, "as novas tecnologias têm o potencial de tornar as medidas de PDL/FT sejame mais rápidas, mais baratas e mais eficazes". Além disso, o GAFI enumera as múltiplas iniciativas de supervisores e instituições de todo o mundo que constituem a vanguarda da evolução do setor: Opportunities and challenges of new technologies for AML/CTF, disponível em https://www.fatf-gafi.org/media/fatf/documents/reports/Opportunities-Challenges-of-New-Technologies-for-AML-CTF.pdf

[85] Uma declaração conjunta da FRS, FDIC e OCC abordou questões da indústria sobre como a orientação de MRM deve ser aplicada aos modelos de compliance de BSA/AML. Os supervisores consideram que nem todos os sistemas devem ser classificados como modelos, e o próprio banco pode categorizar os modelos como entenderem. Mais importante ainda, eles declararam que os bancos não são obrigados a ter processos duplicados ou a conduzir atividades de testes duplicados para cumprir com a regulaão da BSA. Embora fornecendo certo grau de manobra às instituições financeiras, a declaração reforça a visão do banco abordando os riscos associados aos sistemas de PLD (com ou sem modelos).

endereços e titularidades (ver Figura 5). A construção de redes requer computação de pontos de dados granulares suficientes que possam conectar entidades a diferentes objetos, tais como empresas, endereços, usos digitais, etc., e para considerar a força destes relacionamentos (por exemplo, conexão transacional). Esta rede pode ser estruturada como um grafo (tanto direcionado como não direcionado, e ponderado ou não ponderado). A rede construída e as informações nela contidas determinarão a idoneidade de determinadas técnicas (por exemplo, um grafo não direcionado ponderado poderia ser tratado nas seguintes etapas usando técnicas de agrupamento, tais como o agrupamento espectral).

Etapa 2. Estratégia de identificação

É necessária uma estratégia de identificação para desvendar possíveis padrões de lavagem de dinheiro ou outras atividades ilícitas dentro da rede identificada. Há diferentes estratégias que podem ser usadas, por exemplo:

- Abordagens heurísticas baseadas na proximidade de casos ou entidades suspeitas confirmadas.
- Abordagens probabilísticas e reconhecimento de padrões.
- Abordagem de detecção de comunidades baseada em técnicas de *machine learning*.

Ao aplicar a abordagem de detecção de comunidades, é necessário descobrir diferentes comunidades. Uma comunidade é um subgrafo da rede com um número maior e uma relação mais intensa entre os membros da comunidade, em comparação com subgrafos aleatórios e pouco informativos (ver figura 6). A detecção de comunidades é uma abordagem útil para detectar e caracterizar as estruturas-alvo, o que pode exigir o uso de algoritmos como k-means, clustering hierárquico, clustering espectral, algoritmos evolutivos ou otimização da modularidade[86].

Para encontrar as comunidades ótimas, uma função específica é otimizada: a função de modularidade. Dada uma rede representada como um grafo ponderado e dividido em comunidades ou módulos, esta esta fórmula depende da estrutura específica da representação gráfica, e expressa a definição matemática de modularidade em termos de pesos:

$$Q = \frac{1}{2w}\sum_i\sum_j\left(w_{ij} - \frac{w_i w_j}{2w}\right)\delta(C_i, C_j)$$

Onde Ci é a comunidade para a qual o nó i é atribuído, w_{ij} representa o valor do peso na ligação entre os nós i e j (0 se não houver ligação), $w_i = \sum j\ w_{ij}$, e $w=\sum i\ w_i$. Por último, a função δ corresponde à função delta de kronecker delta: δ(i,j) toma o valor 1 se os nós i e j estiverem no mesmo módulo e 0 caso contrário.

Etapa 3. Uso de funções

Uma vez identificadas as comunidades objetivo dentro da rede, métricas ou características específicas podem ser definidas para avaliar a profundidade e a importância dos relacionamentos ou o risco das conexões entre entidades. Estas características podem ser usadas em regras ou algoritmos de *machine learning* para melhorar a capacidade de previsão dos modelos, reduzindo os falsos positivos e identificando melhor os padrões suspeitos. A abordagem baseada em regras com incorporação de características "enriquecidas" pode ser útil para produzir alertas qualitativos, já que incorporam novas

[86] Vários autores desenvolveram algoritmos ótimos para a detecção de padrõesão de padrões. Ver L. Alsedà, A. Awasthi, Jörg Lässig (2012).

Figura 5. Representação do algoritmo de detecção de comunidades (etapas (a) a (d) usando a otimização da modularidade.
A comunidade superior contém um comportamento suspeito (objetivo) que corresponde, neste caso de uso, à rede do Cliente A da Etapa 1 (representado na figura 7 na próxima página).

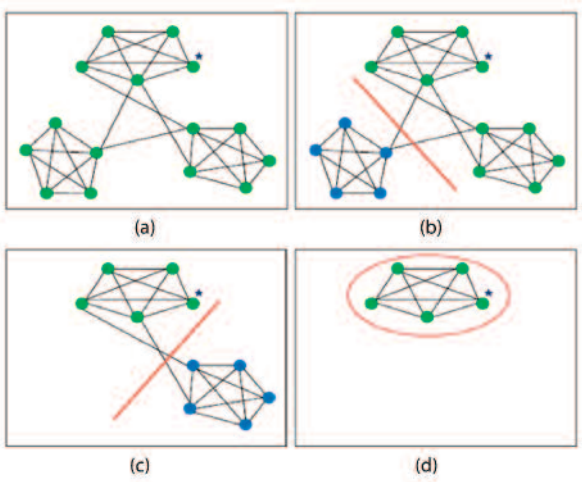
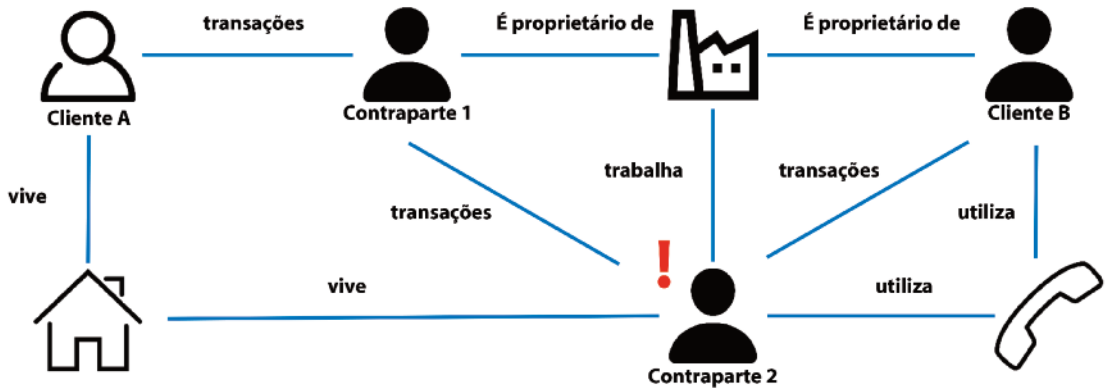


Figura 6. Lógica simplificada para um cenário baseado em características obtidas através da análise da rede. Outras características, tais como valores de transações, podem ser incorporadas de acordo com o padrão ou combinação de padrões a serem detectados. Este exemplo é meramente ilustrativo.

Lógica e parametrização das regras de detecção com características de rede:

- Se tipo de entidade = **Individual**
- Se comunidade suspeita identificada: conexões suspeitas identificadas = **Sim**
- Se o número de rotas para o nó suspeito >= **num_rotas_para_no_suspeito**
- Se a distância até o nó suspeito for <= **distancia_maxima_ate_no_suspeito**
- Se as conexões primárias com o nó suspeito = **tipo_conexao_principal**

Figura 7. Representação dos relacionamentos de rede relacionados ao Cliente A incluindo um nó suspeito (a Contraparte 2 está na blacklist) e possíveis entidades sintéticas (nós relacionados à Companhia)



informações além da base transacional tradicional relacionada com o cliente (ver figura 8). No entanto, as técnicas de *machine learning* podem desvendar relações mais sólidas que permitem separar os verdadeiros alertas positivos dos falsos.

No exemplo da figura 7, cuja infomação de redes é apresentada na figura 8, o cliente A e o cliente B pertencem ao mesmo cluster suspeito com conexões com a entidade suspeita (Contraparte 2), mas o cliente B tem a relação mais forte, tanto pessoal quanto profissionalmente com a Contraparte 2. Com base nesse cenário, se os limites forem calibrados para ser num_rotas_para_no_suspeito = 1, distancia_maxima_ate_no_suspeito = 5 e tipo_conexao_principal = "all" (seja transacional, pessoal ou de qualquer outro tipo), então tanto o Cliente A como o B serão marcados como entidades suspeitas (ou suas transações relacionadas, etc.). No entanto, considerando uma abordagem mais tradicional, sem a análise de redes, somente o cliente B seria marcado como tal; o cliente A não tem conexões transacionais com a Contraparte 2.

Características complexas podem ser avaliadas e diferentes tipos de algoritmos de *machine learning* podem ser treinados, permitindo atribuir um risco maior ao Cliente B e às transações associadas. A adição de novas características aos modelos também permite maior precisão e maior detecção de comportamentos potencialmente arriscados (reduzindo falsos alertas negativos), enquanto se discrimina melhor o risco entre os comportamentos identificados (reduzindo os falsos alertas positivos).

Figura 8. Informação sobre os clientes para a identificação de conexões suspeitas

Empresa	Distância mínima para o nó suspeito	Conexão primária com o nó suspeito	Conexão de dados pessoais	Número de rotas para o nó suspeito	Cluster identificado	Conexões suspeitas identificadas
Cliente A	2	Transacional	Sim	2	1	Sim
Cliente B	1	Transacional	Sim	4	1	Sim