

Aprimoramento de um sistema de monitoramento de PLDFT

Wolney Anjos

Sócio-Consultor na C3W CONSULT

Publicado em 6 de fevereiro de 2020

O sistema de monitoramento de que trata este artigo consiste no processo de seleção de operações para análise com foco em PLDFT por meio de regras específicas de seleção que têm o propósito de cobrir, de maneira mais ampla possível, as diversas situações que podem configurar indícios de lavagem de dinheiro ou de financiamento ao terrorismo, incluindo aquelas descritas na Carta-Circular 4.001/20.

Essa modalidade de detecção de operações suspeitas não é a única, nem a mais efetiva. A suspeição identificada por meio de pesquisa de mídia, de atendimento a demanda de reguladores, de denúncia de colaboradores ou de terceiros são exemplos de meios de identificação de atipicidades mais efetivos.

Neste artigo, o termo efetividade de método de detecção traduz-se na razão entre o número de casos, identificados por determinado método de detecção, que foram comunicados ao Coaf, e o número total de casos identificados por esse determinado método de detecção. Por exemplo, se seis dos dez clientes que tiveram suas transações analisadas em decorrência de seleção por mídia negativa foram comunicados ao Coaf, consideramos que a seleção por mídia negativa tem efetividade de 60% (6 dividido por 10).

A maior efetividade atribuída aos outros métodos em relação ao sistema de monitoramento de que trata este artigo é previsível, pois as situações identificadas pelos outros métodos, além de serem em menor número, apresentam por si só maior grau de suspeição. Por exemplo, uma denúncia de um colaborador sobre determinado cliente apresenta um grau de suspeição maior do que se esse mesmo cliente tivesse sido mais um dentre os diversos selecionados por determinada regra de detecção.

Apesar de sua reconhecida menor efetividade, a detecção por meio de sistema de monitoramento é prática comum e fundamental de monitoramento das transações. Por meio de um sistema de monitoramento desse tipo, é possível identificar clientes com transações atípicas ou suspeitas não identificadas por muitos outros métodos.

A menor efetividade do sistema de monitoramento deve-se à geração de muitos falsos positivos, isto é, a maioria dos clientes selecionados pelo sistema de monitoramento são, após submetidos a análise, considerados como não suspeitos de LDFT e, portanto, não comunicados ao Coaf. A razão disso é o fato de a lavagem de dinheiro ser considerada, estatisticamente, um evento raro.

Afirmar que a lavagem de dinheiro é um evento raro pode soar sem sentido, pois é fato que a lavagem de dinheiro ocorre muito frequentemente; sendo prova disso os constantes relatos na mídia de casos envolvendo lavagem de dinheiro. No entanto, entende-se por evento raro aquele que apresenta baixíssima proporção de ocorrência (baixo percentual de clientes lavadores associado a uma enorme quantidade de transações financeiras diárias que ocorrem numa instituição financeira).

Por exemplo, supondo que uma instituição possua 1.000.000 clientes, que 10.000 deles sejam lavadores, que a quantidade total de transações realizadas por todos os clientes em um mês seja de 100.000.000, que os lavadores sejam responsáveis por 100.000 transações com indícios de lavagem de dinheiro, teremos 0,10% das transações associadas a lavagem de dinheiro.

O elevado número de falsos positivos é uma característica comum, no Brasil ou no exterior, desse tipo de sistema de monitoramento. Nos EUA, por exemplo, 95% ou mais dos alertas gerados pelos sistemas de monitoramento são considerados falsos positivos (FRUTH, 2019[1]).

Para tornar um pouco mais complexa a discussão a respeito da efetividade de um sistema de monitoramento, é importante não associar essa efetividade à capacidade de identificar lavadores. Ou seja, o fato de um percentual significativo dos clientes selecionados por um sistema de monitoramento serem identificados como lavadores não implica a identificação de um número maior de lavadores por tal sistema. Por exemplo, um sistema de monitoramento que possui poucas regras de detecção e que demanda movimentação expressiva (valores elevados de movimentação) para a geração de alertas irá, muito provavelmente, gerar um número reduzido de alertas e identificar comportamentos muito fora do padrão (*outliers*). Dessa forma, a proporção de clientes selecionados por esse sistema que podem ser lavadores pode ser significativamente maior do que a de clientes selecionados por outros sistemas (sistemas com maior número de regras e com limites de movimentação menores para a geração de alertas).

Exemplificando melhor, um sistema que identificou, como lavadores, dez em 100 clientes selecionados, e deixou de selecionar 10 outros lavadores para análise não apresenta de fato resultado melhor do que um sistema que identificou, como lavadores, 15 em 200 clientes (menor efetividade), assumindo ser 20 o total de lavadores.

O que se espera de um sistema de monitoramento é que seja capaz de selecionar a totalidade, idealmente, ou a maior parte dos clientes que usam a instituição para a lavagem de dinheiro ou financiamento ao terrorismo. Ocorre que a efetividade calculada como a razão entre os casos selecionados que foram identificados como lavagem de dinheiro sobre o total de casos selecionados é sinônimo de precisão e não de cobertura.

A precisão mede o quanto o seu modelo (no caso, o sistema de monitoramento) é capaz de identificar o que se propõe. Se o modelo gerou 100 alertas (100 clientes) e, após análise, concluiu-se que 5 são lavadores, a precisão do modelo é de 5%. No entanto, se supormos que o modelo deixou de gerar alertas para outros 45 clientes que realizam operações de lavagem de dinheiro, a cobertura do modelo foi de 10% [2].

Se por outro lado, o modelo gerou dez alertas (dez clientes) e, após análise, concluiu-se que três são lavadores, a precisão do modelo é de 30%. Supondo que o modelo deixou gerar alertas para outros 47 clientes que realizam operações de lavagem de dinheiro, a cobertura do modelo foi de 6% [3]. Nesse caso, a efetividade, apesar de ser maior do que no caso anterior, não se traduziu de fato numa maior capacidade de identificar lavadores.

Ou seja, precisão e cobertura tendem, na prática, a caminhar em sentidos opostos, apesar de isso não ser a regra. O modelo ideal de sistema de monitoramento é aquele gerador de um número reduzido alertas e capaz de capturar o maior número possível de lavadores.

Toda essa discussão no campo teórico não pode ser dirimida na prática, pois não se pode afirmar que um sistema de monitoramento mais preciso é de fato melhor do que um sistema de monitoramento menos preciso. A razão disso é o fato de as seguintes premissas adotadas nunca ou raramente ocorrerem na prática:

- assumir, após a análise dos alertas, que os clientes identificados como lavadores e comunicados ao Coaf são de fato lavadores;
- assumir, após análise dos alertas, que os clientes não identificados como lavadores são de fato não lavadores;
- assumir que seja possível conhecer quais clientes não selecionados pelo sistema de monitoramento (clientes cujas transações/movimentações não geraram alertas) são de fato lavadores.

Ou seja, as métricas precisão e cobertura não são confiáveis o suficiente para serem consideradas na avaliação de um sistema de monitoramento de PLDFT. Aliás, a métrica cobertura sequer é passível de mensuração no caso de lavagem de dinheiro.

Uma das soluções que tem sido apresentada para aumento da efetividade do sistema de monitoramento é o emprego de técnicas estatísticas e de aprendizagem de máquina (*machine learning*), seja por meio de técnicas de aprendizagem supervisionada ou por meio de técnicas de aprendizagem não supervisionada.

O emprego de técnicas de aprendizagem supervisionada implica associar diversas variáveis independentes qualitativas e quantitativas a uma variável resposta. No caso de LDFT, a variável resposta seria o indicador de suspeição de LDFT, isto é, indicar se é ou não lavagem de dinheiro.

Ocorre que essa variável resposta não é confiável o suficiente para o emprego de aprendizagem supervisionada. A razão disso é que a instituição não tem como afirmar, na esmagadora maioria das vezes, que o caso é ou não de fato lavagem de dinheiro. Prova dessa incapacidade é evidenciada pelo fato de, conforme Murilo Portugal (presidente da Febraban), apenas cerca de 16% das comunicações feitas pelos bancos (setor que mais e melhor comunica) ao Coaf terem sido consideradas como de fato possuidoras de atipicidades passíveis de envio ao MP para investigação [\[4\]](#).

Ou seja, se um banco comunica 100 clientes como suspeitos de lavagem de dinheiro e utiliza esses 100 como resposta positiva em uma aprendizagem supervisionada, esse banco associará erroneamente 84 resultados como lavagem (se considerarmos efetividade de 16%) e, além disso, não saberá – pois o Coaf não divulga tal informação - quais são os 16 resultados corretos.

Outro fator em desfavor do emprego da aprendizagem supervisionada é o fato de não necessariamente a totalidade dos casos arquivados (não comunicados ao Coaf) pelo banco corresponder a casos de não lavagem de dinheiro, ou seja, parte dos casos arquivados

podem ser falsos negativos – deveriam ter sido considerados como lavagem de dinheiro e comunicados ao Coaf, mas foram arquivados. Uma forma de a instituição identificar essa falha é quando o cliente tem seu nome divulgado na mídia associado a ilícitos financeiros e apresenta movimentação criminosa semelhante à que realizou nesse banco.

Ou seja, o uso de aprendizagem supervisionada, amplamente utilizado para fraude, não se aplica adequadamente para lavagem de dinheiro.

As técnicas de aprendizagem não supervisionada, por sua vez, não têm por propósito identificar casos de lavagem de dinheiro diretamente, mas sim, por meio de diversas técnicas (*clustering*, *outlier detection*, por exemplo) identificar padrões, similaridades entre dados. O analista pode então, mediante análise dos grupos, identificar aquele que reúne conjunto de atributos com maior probabilidade de estar associado a lavagem de dinheiro.

Por exemplo, a técnica de *clustering* poderia ter agrupado um certo número de empresas que se distinguiram, dentre outros fatores, por realizarem um percentual significativo de transações em espécie no período. Ao se isolar esse grupo para análise, verifica-se que muitas dessas empresas são empresas que de fato operam com valores em espécie (transporte, alimentação, posto de combustível, etc.), mas que há entre as empresas agrupadas uma empresa de consultoria, que não tem, de forma alguma, por característica movimentar recursos em espécie.

A aprendizagem não supervisionada permite, portanto, aprimorar as regras de seleção ou contribuir para a criação de regras mais efetivas.

Com a edição da Circular no 3.978, de 23 de janeiro de 2020, o aprimoramento do sistema de monitoramento, principalmente no sentido de redução de falsos positivos, passa a ser uma prioridade para as instituições financeiras, uma vez que essa norma - em seus artigos 39, parágrafo único, e 43, parágrafo 1º - inovou ao estabelecer prazos para a seleção da transação para análise (45 dias contados da ocorrência da transação) e para a conclusão da análise (45 dias contados da seleção para análise).

Esse aprimoramento pode-se dar com o emprego de técnicas de aprendizagem não supervisionada para a redução de falsos positivos e criação de regras mais eficazes de detecção.

No entanto, uma série de ações podem ser realizadas com o propósito de redução do tempo de análise, tais como:

- uso de sistema de monitoramento que traga o máximo de informações que subsidiarão a análise do alerta^[5];
- uso de sistema de monitoramento que permita o registro da análise realizada e a inclusão de documentos utilizados na análise;
- uso de sistema que passe a incluir informações sobre o tempo, contado da data da seleção, que um alerta permanece pendente de análise;
- uso de sistema de monitoramento de fácil navegação e visualização; e

- capacitação dos funcionários responsáveis pela análise dos alertas.

Esses exemplos não exaustivos de ações que contribuem para agilizar a análise de alertas devem ser seriamente considerados pelas instituições financeiras, principalmente por aquelas que atualmente apresentam prazos de seleção e análise superiores ao que dispõe a Circular n. 3.978/20.

Em resumo, há necessidade de as instituições avaliarem imediatamente seus sistemas de monitoramento quanto aos prazos de seleção e análise praticados, bem como aprimorar as regras de detecção de transações suspeitas. Essas tarefas estão, inclusive, previstas no artigo 63, inciso II, alínea b, da Circular n. 3.978/20, que trata da avaliação de efetividade.

[1] FRUTH, Joshua. ***Anti-money laundering controls failing to detect terrorists, cartels, and sanctioned states***. Reuters: March 14, 2018. Disponível em: <https://www.reuters.com/article/bc-finreg-laundering-detecting/anti-money-laundering-controls-failing-to-detect-terrorists-cartels-and-sanctioned-states-idUSKCN1GP2NV> Acesso em: 14 de setembro de 2019.

[2] $[5/(5+45)] = 10\%$

[3] $[3/(3+47)] = 6\%$

[4] Vide: <https://twitter.com/FEBRABAN/status/1171400179637002241>

[5] Há sistemas que demandam a busca de informações em outros sistemas (cadastro, movimentação financeira, por exemplo), o que acarreta um maior tempo de análise.