UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

TÉCNICAS BASEADAS EM GRAFOS PARA PRIORIZAÇÃO DE INVESTIGAÇÕES POLICIAIS DE FRAUDES BANCÁRIAS ELETRÔNICAS

ÁLEX MOREIRA DO PATROCÍNIO

ORIENTADOR: ZENILTON KLEBER GONÇALVES DO PATROCÍNIO JÚNIOR

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

PUBLICAÇÃO: PPGEE.DM - 632/2016
BRASÍLIA / DF: DEZEMBRO/2016

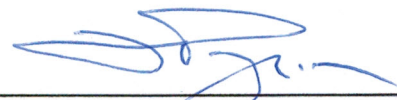
UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

TÉCNICAS BASEADAS EM GRAFOS PARA PRIORIZAÇÃO DE INVESTIGAÇÕES POLICIAIS DE FRAUDES BANCÁRIAS ELETRÔNICAS

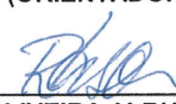
ÁLEX MOREIRA DO PATROCÍNIO

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

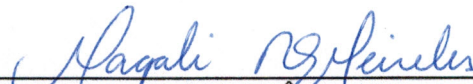
APROVADA POR:



ZENILTON KLEBER GONÇALVES DO PATROCÍNIO JÚNIOR, PUC MINAS
(ORIENTADOR)



ROBSON DE OLIVEIRA ALBUQUERQUE, UNB
(EXAMINADOR INTERNO)



MAGALI REZENDE GOUVÊA MEIRELES, PUC MINAS
(EXAMINADOR EXTERNO)

Brasília, 09 de Dezembro de 2016.

FICHA CATALOGRÁFICA

PATROCÍNIO, ÁLEX MOREIRA

Técnicas Baseadas em Grafos para Priorização de Investigações Policiais de Fraudes Bancárias Eletrônicas. [Distrito Federal] 2016.

xv, 103 p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2016).

Dissertação de Mestrado - Universidade de Brasília.

Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. Algoritmos para Grafos
2. Análise de Vínculos
3. Data Mining
4. Mobile Banking
5. Fraude Eletrônica Bancária
6. Internet Banking

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

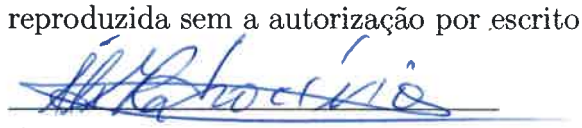
Patrocínio, Á. M. (2016). Técnicas Baseadas em Grafos para Priorização de Investigações Policiais de Fraudes Bancárias Eletrônicas. Dissertação de Mestrado em Engenharia Elétrica, Publicação PPGEE.DM - 632/2016, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 103p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Álex Moreira do Patrocínio.

TÍTULO DA DISSERTAÇÃO DE MESTRADO: Técnicas Baseadas em Grafos para Priorização de Investigações de Fraudes Bancárias Eletrônicas.

GRAU / ANO: Mestre / 2016



Álex Moreira do Patrocínio

Rua Nascimento Gurgel, 30, Gutierrez

30.430-340 Belo Horizonte - MG - Brasil.

À Deus, que me trouxe até aqui, que me guarda e me guia, e que me dá forças para honrar os que me são caros. Ao meu orientador e querido irmão Zenilton Kleber Gonçalves do Patrocínio Júnior, que sempre me apoiou na minha vida, aos meus pais Edna Moreira do Patrocínio e Zenilton Kleber Gonçalves do Patrocínio, exemplos de vida, determinação e amor, às minhas lindas filhas Bárbara Moreira do Patrocínio e Bianca Moreira do Patrocínio, que me enchem de carinho, amor, orgulho, e que fazem tudo valer a pena.

AGRADECIMENTOS

Ao meu orientador e querido irmão Prof. Dr. Zenilton Kleber Gonçalves do Patrocínio Júnior, pelo constante apoio, incentivo, amizade e dedicação, não só na minha vida acadêmica como na minha vida pessoal em família, além do seu profissionalismo essencial para o desenvolvimento deste trabalho.

Agradeço à minha família, irmãos Daniel e Zenilton Jr. e irmã Janaína, em especial aos meus pais Zenilton e Edna, que me encheram de apoio e carinho ao longo da minha vida.

Ao meu colega de trabalho e amigo o Perito Criminal Federal Stefenson Marcus Pinto Scafutto que foi o primeiro a me incentivar a concorrer a uma das vagas deste mestrado profissional da UNB.

Ao amigo Raucelio Coelho Cardoch Valdes, da DICOR/DPF, que me dava esperança nos momentos difíceis, pelo apoio fundamental nas aulas e trabalhos de estatística deste meu mestrado.

Ao meu querido amigo Erik Pereira de Siqueira, Policial Federal, que me inspirou com sua monografia de pós-graduação na UNB em 2014 - "O projeto **Tentáculos** da Polícia Federal - Da concepção à proposta de modelo aplicável na Segurança Pública Brasileira". Erik, obrigado pelas sugestões que você fez para esse trabalho e por se mostrar um amigo fiel.

Aos colegas Policiais Federais Rodrigo de Meira Navarro e Gustavo Pires de Sá, pelas conversas enriquecedoras, amizade, ajuda em diversos aspectos e colaboração na elaboração do material deste meu trabalho por meio dos relatórios de investigação que elaboraram na SRCC/DPF utilizando projeto **Tentáculos** da Polícia Federal do Brasil.

Ao meu professor do CINTEPOL e colega Policial Federal, Tarcísio José da Silva Júnior pela sua ajuda fundamental na elaboração da especificação de importação dos grafos do **Kraken** no *Analyst´s Notebook da IBM*.

Ao Delegado de Polícia Federal Marco Aurélio de Macedo Coelho, pelo seu apoio e sugestões feitas para esse trabalho.

Ao amigo e colega Policial Federal Antônio Carlos Balbino Azevedo, pelo apoio na realização deste trabalho e pelas conversas animadoras durante os meses que trabalhamos juntos.

Aos amigos e colegas desta minha turma do mestrado, pela camaradagem e apoio nos momentos difíceis que passamos juntos.

Aos colegas e amigos do serviço psicossocial da PF, representados pela pessoa da psicóloga Wênia O. Santos, meu muito obrigado por sempre me apoiarem nos momentos difíceis.

À minha filha Bárbara Moreira do Patrocínio, revisora deste meu trabalho, de quem eu muito me orgulho e admiro pela sua leveza, bondade, carinho e comprometimento.

À minha filha Bianca Moreira do Patrocínio, pelo seu amor, carinho e alegria incondicionais e contagiantes que elevam meu espírito em momentos difíceis.

À todos os meus professores do mestrado profissional em engenharia elétrica (2014/2015) da UNB, pelos seus ensinamentos, e em especial aos professores Flávio Elias Gomes de Deus e Dibio Leandro Borges, pela camaradagem nos meus momentos difíceis.

À Polícia Federal, representada na pessoa do Delegado de Polícia Federal Braulio Cezar da Silva Galloni – Coordenador Geral da Polícia Fazendária da PF e a DITEC - Diretoria Técnico-Científica da Polícia Federal. À todos, os meus sinceros agradecimentos.

O presente trabalho foi realizado com o apoio da Polícia Federal do Brasil – PF e do Instituto de Criminalística do Distrito Federal, com recursos do Programa Nacional de Segurança Pública com Cidadania – PRONASCI, do Ministério da Justiça.

Álex Moreira do Patrocínio

RESUMO

TÉCNICAS BASEADAS EM GRAFOS PARA PRIORIZAÇÃO DE INVESTIGAÇÕES POLICIAIS DE FRAUDES BANCÁRIAS ELETRÔNICAS.

Autor: Álex Moreira do Patrocínio

Orientador: Zenilton Kleber Gonçalves do Patrocínio Júnior

Programa de Pós-graduação em Engenharia Elétrica

Brasília, Dezembro de 2016

Ao longo dos últimos anos, a Polícia Federal do Brasil (**PF**) vem concentrando esforços para elucidar crimes de fraudes bancárias praticados contra a empresa pública e instituição financeira da Caixa Econômica Federal (**CAIXA**). A elucidação desses crimes é uma atribuição da PF, prevista na Constituição Federal do Brasil em seu artigo 144 § 1º. A abordagem proposta neste trabalho, denominada **Kraken**, é aprimorar o modelo de investigação existente utilizando-se de grafos e da análise de vínculos para auxiliar às técnicas de investigação da PF. O Kraken trata, em específico, de investigações contra fraudes eletrônicas relativas a transferência de valores entre contas bancárias. Para aplicar a abordagem do Kraken, foi necessária a criação de um ferramental que processe toda a informação contida na Base Nacional de Fraudes Bancárias Eletrônicas (**BNFBE**) e a transforme em grafos conexos, que representem os atores e seus relacionamentos (vínculos) na ação delitiva desse tipo de fraude. Para a realização desse trabalho tivemos acesso a um conjunto dos dados da BNFBE. O objetivo desse ferramental é servir como uma Interface Gráfica (**IG**) para que o investigador Policial Federal possa verificar os resultados do processamento da abordagem do Kraken em um formato de tabela, onde cada registro represente um grafo referente a um conjunto de atores e vínculos envolvidos na ação delitiva. É na IG do Kraken que o Policial Federal consegue reordenar a tabela que contém as investigações/grafos em ordem decrescente de métricas objetivas, como: quantidade de vítimas, pessoas beneficiadas e valor total das fraudes existentes em cada grafo gerado pelo Kraken. A principal hipótese é que com o reordenamento dos grafos, baseado nessas métricas objetivas, possa-se priorizar as investigações criminais a serem analisadas pela PF. A IG permite selecionar um grafo específico da tabela para ser exportado e visualizado no *Analyst´s Notebook* da IBM. Com a abordagem do Kraken e sua IG, espera-se diminuir a necessidade de interferência humana (investigador policial) nos relatórios de análise dos crimes e, consequentemente, acelerar as investigações da PF.

ABSTRACT

TECHNIQUES BASED ON GRAPH FOR PRIORITISATION INVESTIGATIONS FRAUD COPS BANKING ELECTRONIC.

Author: Álex Moreira do Patrocínio

Supervisor: Zenilton Kleber Gonçalves do Patrocínio Júnior

Programa de Pós-graduação em Engenharia Elétrica

Brasília, Dezembro of 2016

Over the last few years the Federal Police of Brazil (**PF**) has concentrated efforts to elucidate crimes of bank fraud committed against Caixa Econômica Federal (**CAIXA**) a public company and financial institution. The elucidation of these crimes is an attribution of the PF, provided for in the Federal Constitution of Brazil in its article 144 § 1º. The approach presented in this work, denominated **Kraken**, proposes to improve the existing research model, using graphs and link analysis, to assist in the investigation techniques of the PF. Kraken deals in particular with investigations into electronic fraud involving transfer of securities between bank accounts. In order to apply Kraken's approach it was necessary to create a tool to process all the information contained in the National Electronic Banking Fraud Database (**BNFBE**) and to transform it into related graphs that represent the actors and their relationships (links) in the criminal act of this type of fraud. For this work we had access to a set of BNFBE data. The objective of this tool is to serve as a Graphic Interface (**IG**) so that the Federal Police investigator can verify the results of the processing of the Kraken approach in a table format, where each record represents a graph referring to a set of actors and movements involved in the crime scene. In Kraken's IG the Federal Police officer can rearrange the table containing the investigations / graphs in descending order of objective metrics such as: number of victims, persons who have profited from crime and total value of frauds in each graph generated by Kraken. The main hypothesis is that with the rearrangement of the graphs, based on these objective metrics, it is possible to prioritize the criminal investigations to be analyzed by the PF. The IG allows to select a specific graph of the table to be exported and viewed on IBM's Analyst's Notebook. With the approach of Kraken and its IG, it is hoped that the need for human interference (police investigator) will be reduced in the crime analysis reports and, consequently, further accelerate investigations in the PF.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIAÇÕES

AD - Análise de Densidade

AR - Alocação de Recursos

ARS - Análise de Redes Sociais

ATM - Terminal de Auto Atendimento

BNFBE - Base Nacional de Fraudes Bancárias e Eletrônicas

CAIXA - Caixa Econômica Federal

CB - Conta Beneficiada

CGPFAZ - Coordenação Geral de Polícia Fazendária

CV - Conta Vítima

DER - Diagrama de Entidade e Relacionamento

DICOR - Diretoria de Combate ao Crime Organizado

DITEC - Diretoria Técnico-Científica da Polícia Federal

ERB - Estação Rádio Base

ETL - Extract Transform and Load

FGV - Fundação Getúlio Vargas

FEBRABAN - Federação Brasileira de Bancos

GRCC - Grupo de Repressão a Crimes Cibernéticos

GPA - Grupo Permanente de Análise

IG - Interface Gráfica

INSS - Instituto Nacional de Seguridade Social

IP - Internet Protocol

IPL - Inquérito Policial

MAFBE - Modelo de Análise de Fraudes Bancárias Eletrônicas

PB - Pessoa Beneficiada

PF - Polícia Federal / Policial Federal

RAM - Random Access Memory ou Memória de Acesso Aleatório

SGBD - Sistema de Gerenciamento de Banco de Dados

SQL - Structured Query Language, ou Linguagem de Consulta Estruturada

SRCC - Serviço de Repressão a Crimes Cibernéticos

UNB - Universidade de Brasília

URA - Unidade de Resposta Audível

VC - Vizinhos Comuns

1 INTRODUÇÃO

Hoje em dia, muitas das transações bancárias são realizadas por meio do *Internet banking*, *mobile banking* ou por meio de cartões bancários na modalidade débito ou crédito, o que faz aumentar o número de fraudes eletrônicas em desfavor das instituições financeiras. A fim de diminuir o número de procedimentos policiais para apurar estes crimes e centralizar as informações das fraudes bancárias em desfavor da CAIXA, a PF idealizou, no ano de 2007, o Projeto Tentáculos.

Desde 2008, as notícias-crime são encaminhadas, periodicamente, diretamente do órgão central da CAIXA para a PF por meio de mídia eletrônica (arquivos textos).

Por meio de um *Extract Transform and Load* **(ETL)**, a PF extrai os dados dos arquivos textos encaminhados pela CAIXA e realiza as adequações necessárias para povoar a BNFBE modelada no *IBASE* da IBM.

Na BNFBE somente existem registros de fraudes bancárias eletrônicas já constatadas pela CAIXA, portanto, não existem mais dúvidas da ilegalidade ocorrida nas transações bancárias contidas nessa base, cabendo a PF apurar criminalmente a materialidade e a autoria dessas ilicitudes.

Segundo Siqueira (2014), a BNFBE utiliza ferramentas de análise de vínculos, com modelagem própria, que buscam identificar relações existentes entre fraudes bancárias do mesmo tipo a fim de gerar um procedimento policial investigatório único, que reúna todas as vítimas de um mesmo criminoso ou organização criminosa. Desse modo, evita-se o retrabalho, resultando em um melhor aproveitamento dos recursos humanos e aprimorando a inteligência policial no combate às fraudes bancárias eletrônicas.

Os investigadores Policiais Federais, que trabalham na elaboração dos relatórios de análises dos crimes, são treinados nas técnicas de análise de vínculos dos atores envolvidos nessas fraudes bancárias, e para a visualização de suas investigações utilizam-se de grafos gerados no *software Analyst´s Notebook* que está conectado a BNFBE modelada no *IBASE* da IBM.

A PF também utiliza outros meios de investigação para complementar os dados contidos na BNFBE, como informações levantadas por meio da vigilância de Policiais Federais em locais específicos, dados fornecidos por informantes, mandados de busca e apreensão de documentos, depoimentos de vítimas, quebras de sigilos autorizados judicialmente, entre outros.

É diante de todo esse cenário, que chamamos neste trabalho de Modelo de Análise Fraudes **(MAFBE)**, praticado para a elucidação de crimes de fraudes bancárias eletrônicas, que o investigador inicia seu trabalho na elaboração dos relatórios de análise dos crimes em questão.

1.1 OBJETIVO

Este trabalho tem como objetivo geral a criação de uma abordagem denominada Kraken, capaz de varrer toda a BNFBE. Por meio da automatização da análise de vínculos, o Kraken seria capaz de agrupar em grafos especificamente os atores envolvidos na fraude eletrônica do tipo "transferência de valores entre contas bancárias" da CAIXA. Assim, o investigador Policial Federal poderá priorizar suas investigações baseado em métricas objetivas e gerar seus relatórios de análise dos crimes.

Para alcançar o objetivo geral, foram definidos os objetivos específicos, a saber:

- Criar um ferramental com uma IG para que o investigador Policial Federal possa verificar os resultados do processamento da abordagem do Kraken em um formato de tabela, onde cada registro represente os metadados de um grafo referente a um conjunto de atores e vínculos envolvidos na ação delitiva, que é foco deste trabalho;
- Identificar os componentes conexos que irão representar grupos de contas vítimas juntamente com os demais elementos vinculados às mesmas, como os telefones utilizados para acessar o saldo das contas vítimas, as contas que se beneficiaram de transferências bancárias oriundas das contas vítimas, data e valor da transação, além dos titulares das contas beneficiadas;
- Gerar um roteiro que contemple os componentes conexos de cada grafo com a cronologia e os valores, caso existam, da ação delituosa;
- Possibilitar ao investigador Policial Federal reordenar por meio da IG a tabela que contém os grafos em ordem decrescente de métricas objetivas, como: quantidade

de vítimas, pessoas beneficiadas e valor total das fraudes existentes em cada grafo gerado pelo Kraken, para que o investigador priorize as investigações a serem realizadas;

- Permitir ao investigador realizar na IG filtros nos grafos por parâmetros, como: o valor total em reais fraudado, período de datas em que ocorram as fraudes, contas vítimas de um determinado estado (UF) do Brasil, contas beneficiadas de um determinado estado (UF) do Brasil, entre outros;
- Permitir ao investigador selecionar na IG um grafo em específico da tabela para ser exportado e visualizado no *Analyst´s Notebook* da IBM.

1.2 RESULTADOS ESPERADOS

Os resultados esperados para este trabalho são:

a) Contribuir para que se atinja o objetivo estratégico estabelecido pela PF de se tornar referência entre as principais Polícias Federais do mundo;

b) Auxiliar no direcionamento e priorização do processo de investigação do MAFBE de uma forma padronizada;

c) Obter um ganho expressivo de tempo e produtividade nas investigações de crimes de fraudes bancárias eletrônicas;

d) Aumentar a qualidade na geração dos relatórios de análise das investigações da PF de fraudes bancárias eletrônicas;

e) Elevar os índices de elucidação de crimes de fraude eletrônica do tipo "transferência de valores entre contas bancárias" da CAIXA;

f) Criar uma abordagem capaz de viabilizar a análise em grande volumes de dados digitais e servir como ponto de partida na resolução de outros tipos de crimes, não só na área de fraudes bancárias eletrônicas, mas também de crimes contra o meio ambiente, a previdência social e o desvio de verbas públicas.

1.3 JUSTIFICATIVA

Desde a sua implementação em 2008, o MAFBE tem absorvido um aumento expressivo das fraudes bancárias eletrônicas. Segundo Mompean (2016), somente as transações via *Internet* na modalidade *mobile banking* cresceram 138% entre os anos de 2014 a 2015, e mais de 100 vezes de 2011 a 2015. Assim, verifica-se a necessidade imediata de se repensar as intervenções manuais dos investigadores Policiais Federais, que utilizam do modelo de análise de vínculos no MAFBE a fim de suportar a demanda crescente de combate às fraudes bancárias eletrônicas que estão por vir nos próximos anos.

A Figura 1.1 ilustra o crescimento das transações realizadas nas modalidades *Internet banking* e *mobile banking* no Brasil entre os anos de 2011 e 2015 (MOMPEAN, 2016).

Com o aumento das fraudes bancárias eletrônicas cresce em paralelo o mercado clandestino de dados, informações e artefatos maliciosos. Esse comércio é bastante intenso, organizado, interconectado e em franco processo de amadurecimento (CAVALLARO, 2014).

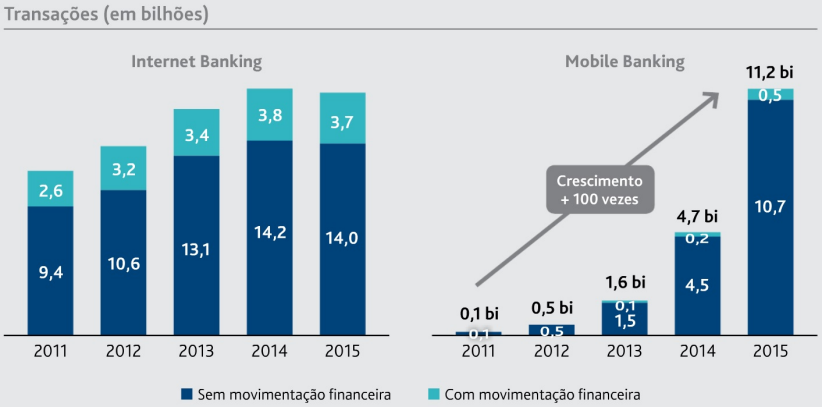


Figura 1.1: Volume de Transações Bancárias *Internet banking* e *mobile banking*.
Fonte: MOMPEAN, 2016

A necessidade da elaboração da abordagem Kraken com o seu ferramental (IG) que automatize a geração dos grafos de crimes de fraudes bancárias eletrônicas contidas na BNFBE, que permita ao investigador filtrar e priorizar essas investigações e nelas destacar os principais atores envolvidos na ação criminosa, já justificaria este trabalho, existindo ainda os seguintes fatores para serem considerados:

- A expressiva diminuição do tempo da elaboração do relatório de análise da investigação policial, uma vez que com a abordagem proposta neste trabalho será

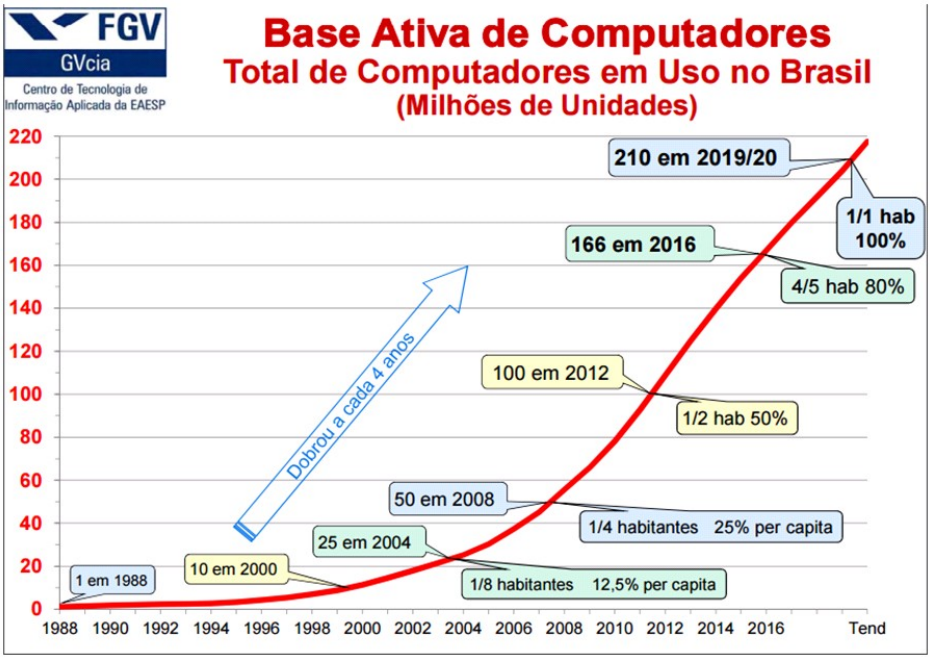


Figura 1.2: Base Ativa de Computadores em Uso no Brasil.
Fonte: FUNDAÇÃO GETÚLIO VARGAS, 2016

possível utilizar o MAFBE e promover a automação do processo de análise de vínculos já existente na BNFBE, como é demonstrado no Capítulo 3;

- Com a abordagem proposta será possível cobrir a totalidade dos registros de fraudes bancárias armazenados na BNFBE, o que é humanamente impossível de ser feito atualmente em um prazo de tempo aceitável, não só devido ao quadro reduzido de Policias Federais investigadores, mas pelas milhares de fraudes bancárias contestadas na CAIXA e enviadas periodicamente para a PF;

- O crescimento do uso de computadores no Brasil dobra a cada 4 anos (FUNDAÇÃO GETÚLIO VARGAS, 2016), aumentando a demanda de investigações policiais de fraudes bancárias eletrônicas. A Figura 1.2 mostra o crescimento da base ativa de computadores no Brasil;

- Por fim, de acordo com Mompean (2016), o número de contas de *mobile banking* no período de 2011 a 2015 apresentou um crescimento de 16 vezes, o que acarreta no aumento de fraudes bancárias nesta modalidade. A Figura 1.3 mostra o número de contas bancárias acessadas via *Internet banking* e *mobile banking* no Brasil.

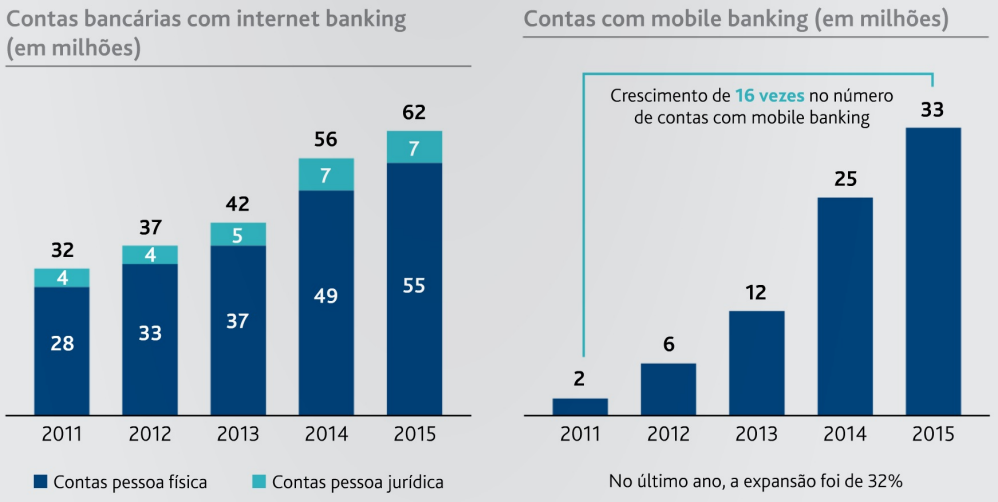


Figura 1.3: Contas Bancárias com *Internet* e *mobile banking* no Brasil.
Fonte: MOMPEAN, 2016

1.4 HIPÓTESES

Nessa seção são apresentadas as hipóteses deste trabalho.

1.4.1 DIMINUIÇÃO DO TEMPO DA INVESTIGAÇÂO CRIMINAL

Com o Kraken será possivel diminuir o tempo na elaboração dos relatórios de análise das investigações criminais relativas a fraudes bancárias eletrônicas?

1.4.2 PRIORIRAÇÃO DAS INVESTIGAÇÕES CRIMINAIS

Existindo a possibilidade, em tese, que a MAFBE ainda tenha que analisar registros na BNFBE para a elucidação de seus fatos, poderia a abordagem proposta nesse trabalho servir para priorizar quais das investigações contidas na BNFBE devem ser realizadas primeiro pela PF, baseando-se em métricas objetivas, uma vez que esse trabalho propõem varrer de uma única vez toda a BNFBE e gerar a análise dos crimes nela contidos?

1.4.3 EXPLICITAR OS PRINCIPAIS ATORES NOS GRAFOS

Por meio das análises de rede propostas nesse trabalho será possível auxiliar os investigadores na localização dos principais atores envolvidos nas fraudes bancárias eletrônicas?

1.4.4 AJUDAR NA PADRONIZAÇÃO DOS RELATÓRIOS DE ANÁLISE DE CRIMES

Com a implementação da IG do Kraken e de algoritmos e técnicas baseadas em grafos e análise de vínculos é possível criar um sistema que gere grafos conexos, que representem investigações ou parte delas, para que esses grafos sejam impressos e passem a fazer parte dos relatórios de análise dos crimes de fraudes bancárias eletrônicas utilizados pela PF na elucidação dos crimes?

1.4.5 LOCALIZAR E INCREMENTAR A QUANTIDADE DE VÍNCULOS E ATORES ENVOLVIDOS EM UMA MESMA AÇÃO CRIMINAL EM RELAÇÃO A MAFBE

Considerando, em tese, que a programação realizada na IG do Kraken é eficiente e capaz de gerar as análises dos crimes existentes na BNFBE em questão de horas ao invés de dias, semanas ou meses, seria possível diminuir o impacto causado pela entrada de dados extemporâneos (fornecidos pela CAIXA) na BNFBE, e demais **fatores que dificultam a investigação** realizada na MAFBE, tendo um acréscimo na localização e na quantidade de vínculos e atores relacionados à uma investigação envolvendo a mesma organização criminosa?

1.5 METODOLOGIA

Para alcançar os objetivos propostos foram realizados os passos descritos a seguir:

a) Revisão da literatura e de trabalhos externos a PF relacionados à investigação policial, à análise de vínculos e a grafos voltados para solução de investigações criminais;

b) Identificação dos trabalhos já existentes na PF e possíveis aprimoramentos de suas soluções e heurísticas;

c) Levantamento do MAFBE e de relatórios de análise de investigações contra fraudes bancárias eletrônicas de transferência de valores entre contas bancárias de investigações reais e já realizadas pela PF;

d) Conceber a abordagem proposta pelo Kraken, levando-se em conta a necessidade da automatização e priorização das investigações de crimes de fraudes bancárias eletrônicas contra a CAIXA;

e) Para dar suporte à metodologia do Kraken foi desenvolvido um ferramental forense especializado que auxiliará nos relatórios de análise produzidos pelos investigadores Policiais Federais de crimes de fraudes bancárias eletrônicas contra a CAIXA;

f) A metodologia do Kraken e o seu ferramental (IG) foram submetidos à prova de conceitos, e os grafos gerados pelo ferramental foram inicialmente homologados por este autor e outros Policiais Federais em confronto com relatórios de análise de investigações reais já realizadas pela PF para a extração dos resultados apresentados neste trabalho.

1.6 ORGANIZAÇÃO DO TRABALHO

No Capítulo 2 será detalhada a revisão bibliográfica com definições associadas a grafos (caminho/ciclo, árvore, componente conexo e grupamento), estudo do projeto Tentáculos que faz parte da MAFBE, definições de fraudes eletrônicas e de análise de vínculos (diretos e indiretos), além de trabalhos relacionados.

O Capítulo 3 mostra as dificuldades investigativas e como a proposta deste trabalho será implementada, assim como o modelo da BNFBE, a metodologia do Kraken, telas do ferramental (IG) criada para dar suporte a abordagem do Kraken, métricas utilizadas para mostrar ao investigador quais grafos gerados pelo Kraken devem ser priorizados, além dos filtros que a IG é capaz de realizar sobre o resultado dos grafos processados pelo Kraken.

No Capítulo 4 serão realizados vários cenários de ensaios em laboratório utilizando a própria BNFBE para comparação de resultados do MAFBE com a abordagem do Kraken.

Por fim, o Capítulo 5 entrega as conclusões deste trabalho e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Neste Capítulo é detalhada a revisão bibliográfica com definições associadas a grafos (caminho/ciclo, árvore, componente conexo e grupamento), estudo do projeto Tentáculos (que faz parte da MAFBE), definições de fraudes eletrônicas e de análise de vínculos (diretos e indiretos), além de trabalhos relacionados.

2.1 TEORIA DOS GRAFOS

> "Um grafo $G = (V, E)$ é um conjunto finito não vazio V e um conjunto E de pares não ordenados de elementos distintos de V. G é chamado trivial quando $|V| = 1$. Quando necessário, utiliza-se o termo grafo não direcionado, para designar um grafo. Os elementos de V são os vértices, e os de E são as arestas de G, respectivamente. Cada aresta $e \in E$ será denotada pelo par de vértices $e = (v, w)$ que a compõe. Nesse caso, os vértices v e w são os extremos (ou extremidades) da aresta e, sendo denominados adjacentes. A aresta e é dita incidente a ambos v e w. Duas arestas que possuem um extremo comum são chamadas de adjacentes. Utilizando a notação $n = |V|$ e $m = |E|$." (SZWARCFITER, 1986).

Ainda, segundo Szwarcfiter (1986), os grafos podem ser visualizados por meio de uma representação geométrica, na qual seus vértices correspondem a pontos distintos do plano em posições arbitrárias. Já cada aresta (v, w) é associada à uma linha arbitrária que une os pontos correspondentes a v e w, de acordo com a Figura 2.1.

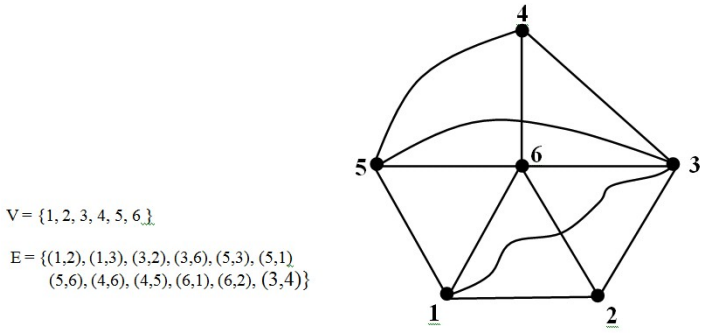


Figura 2.1: Grafo $G = (V, E)$ e uma representação geométrica do mesmo.
Fonte: SZWARCFITER, 1986

Denomina-se componentes conexos de um grafo G os subgrafos maximais de G que sejam conexos. Os componentes conexos de um grafo G são, portanto, os subgrafos de G correspondentes às porções contíguas de sua representação geométrica. O grafo da Figura 2.1 é conexo, enquanto que o da Figura 2.2 não o é.

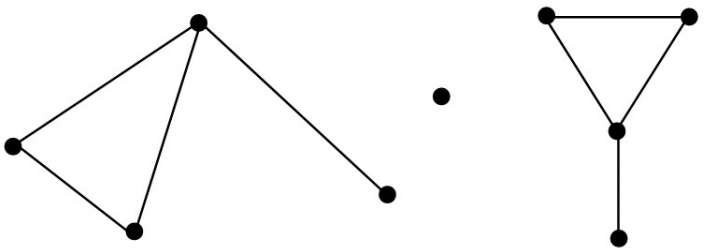


Figura 2.2: Grafo desconexo.
Fonte: SZWARCFITER, 1986

Um componente conexo é um subgrafo maximal conexo de um grafo. Cada vértice do grafo original pertence a exatamente um componente conexo, e o mesmo é valido para suas arestas. A identificação de componentes conexos do grafo de vínculos G permite o levantamento de forma automática dos grupos de criminosos. Além disso, segundo Zhang & Yang (2008), agrupar dados representados sob a forma de um grafo se resume a particionar o mesmo em subgrafos, de modo que cada subgrafo consistido de vértices interconectados é referido como um grupo. Dessa forma, os dados armazenados no BNFBE passam a ser conjuntos de vértices nos subgrafos. Assim, a abordagem proposta Kraken gera diversos subgrafos, de modo que cada um represente uma possível investigação criminal.

Segundo Siqueira (2014), o modelo de dados da BNFBE é estruturado com base nos vínculos diretos e indiretos. Portanto, a utilização de tais vínculos é importante no desenvolvimento de rotinas automatizadas de análise, nas quais os vínculos, principalmente os diretos, podem ter diversos níveis de profundidade. A Figura 2.3 mostra um exemplo dos níveis de relacionamentos a partir de uma conta que foi vítima de transferência bancária eletrônica fraudulenta.

Segundo Szwarcfiter (1986), denomina-se conectividade de vértices c_V de G à cardinalidade do menor corte de vértices de G. Já a conectividade de arestas c_E de G é igual à cardinalidade do menor corte de arestas de G. Ou seja, c_V é igual ao menor número de vértices cuja remoção desconecta G ou o transforma no grafo trivial.

Os componentes conexos são utilizados em árvores para realizar busca em profundidade. Servem para particionar a floresta em subárvore disjuntas.

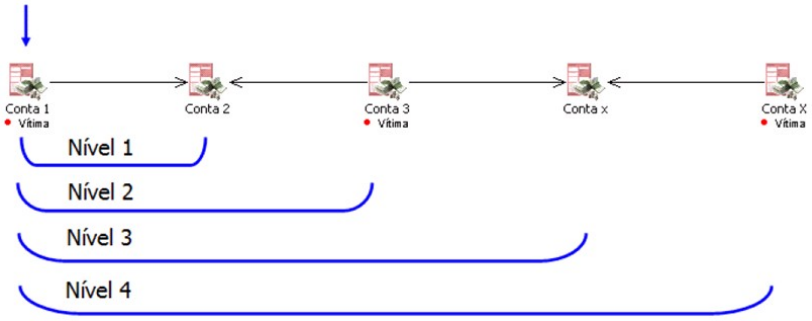


Figura 2.3: Relacionamento a partir de uma conta vítima. Software: *Analyst´s Notebook 8.*

Fonte: SIQUEIRA, 2014

> Em um grafo $G = (V, E)$, "uma sequência de vértices $v_1, \ldots, v_k$ tal que $(v_j, v_{j+1}) \in E, 1 \leq j \leq k-1$, é denominado caminho de v_1 à v_k. Diz-se então que v_1 alcança ou atinge v_k. Um caminho de k vértices é formado por $k-1$ arestas $(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)$. O valor $k-1$ é o comprimento do caminho. Se todos os vértices do caminho $v_1, \ldots, v_k$ forem distintos, a sequência recebe o nome de caminho simples ou elementar. Se as arestas forem distintas, a sequência denomina-se trajeto." (SZWARCFITER, 1986).

Um ciclo é um caminho $v_1, \ldots, v_k, v_{k+1}$ sendo $v1 = v_{k+1}$ e $k \geq 3$. Se o caminho $v_1, \ldots, v_k$ for simples, o ciclo $v1, \ldots, v_k, v_{k+1}$ também é denominado simples ou elementar. Um grafo que não possui ciclos simples é acíclico. Segundo West (2000), os grafos podem ser:

Grafo Direcionado – G é um par (V, A), em que V é um conjunto finito de vértices e A é um conjunto de arestas com uma relação assimétrica em $V \times V$.

Grafo Completo – G^k é um grafo não direcionado no qual todos os pares de vértices são adjacentes, ou seja, possui arestas ligando todos os vértices entre si.

Árvore – A representa um grafo conectado em que não existem ciclos, ou seja, que seja acíclico e conexo. Um conjunto de árvores é denominada floresta, logo todo gráfico acíclico é uma floresta.

Árvore Geradora – AG representa qualquer subárvore de G que contenha todos os vértices de G.

Árvore Geradora Mínima – AGM é qualquer árvore geradora do grafo que minimiza a soma do peso das arestas.

Um grafo pode ser representado computacionalmente por uma lista de adjacência de cada vértice (Figura 2.4(a)) utilizando uma representação compacta (*compact forward and reverse star representation*) gerada a partir dos registros de um banco de dados relacional, que permita determinar os vínculos/arestas entre as entidades (vértices/atores) existentes nessa base de dados.

Assim se constrói o grafo de vínculos $G = (V, E)$ em que V representa o conjunto de vértices (ou entidades) e E representa as arestas ou vínculos entre eles (Figura 2.4(b)). A Figura 2.4 mostra um exemplo de como podemos representar um grafo utilizando a *compact forward and reverse star representation*.

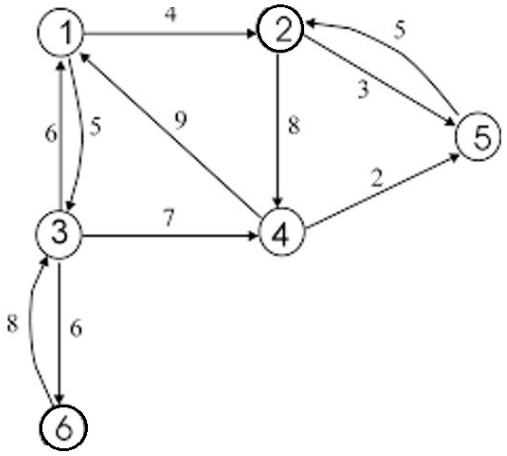
	Pointer
1	1
2	3
3	5
4	8
5	10
6	11
7	12

	Arc number	Starting node	Ending node	Arc length
1	1	2	4
2	1	3	5
3	2	4	8
4	2	5	3
5	3	1	6
6	3	4	7
7	3	6	6
8	4	5	2
9	4	1	9
10	5	2	5
11	6	3	8

	Trace
1	5
2	9
3	1
4	10
5	2
6	11
7	3
8	6
9	8
10	4
11	7

	R_Pointer
1	1
2	3
3	5
4	7
5	9
6	11
7	12

(a) Lista de adjacência de vértices com representação compacta



(b) Grafo resultante da representação compacta

Figura 2.4: Representação de um grafo pela *compact forward and reverse star representation*

Fonte: Elaborada pelo autor.

AGM é um meio muito utilizado em problemas que envolvem agrupamento e tratam de dados a serem agrupados como vértices em um grafo, em que as arestas representam relacionamentos entre vértices vizinhos cujos pesos indicam a similaridade entre eles. Quanto menor for o peso, maior será a semelhança. Desse modo, agrupar dados se resume a particionar o grafo em subgrafos, de modo que cada subgrafo consistido de vértices interconectados é referido como um **grupo** (ZHANG; YANG, 2008).

Existem vários métodos para **agrupamento** de grafos: agrupamento de complexidade global, iterativo, hierárquico, divisivo e aglomerativo. Maiores detalhes sobre estes métodos de agrupamento podem ser vistos em Schaeffer (2007) e Xu & Wunsch (2005).

Um algoritmo de **busca em profundidade** – ***Depth-First Searh*** é capaz de automatizar a análise de vínculos diretos na BNFBE realizando assim o agrupamento de vértices que estejam envolvidos em um determinado subgrafo, em tese, de uma determinada investigação policial.

Segundo Szwarcfiter (1986), uma busca é dita em profundidade quando o critério de escolha de vértice marcado (a partir do qual será realizada a próxima exploração de aresta) obedecer ao princípio: *dentre todos os vértices marcados e incidentes a alguma aresta ainda não explorada, escolher aquele mais recentemente alcançado na busca.* Esse critério garante que a escolha de vértice torna-se única e sem margem de dúvidas.

A Figura 2.5 apresenta a implementação em *ObjectPascal* do algoritmo para busca em profundidade na abordagem Kraken. Nesse algoritmo de busca em profundidade, cada aresta (v, w) é processada duas vezes. A primeira vez levando-se em conta w pertencente lista de adjacências de v, ou ainda $A(v)$, e a segunda vez em que v pertence a lista de adjacências de w, ou ainda $A(w)$. Isso nos leva a conclusão que a busca em profundidade possui uma complexidade $O(n + m)$, em que n representa o número de vértices do grafo e m, o número de aresta do mesmo.

```
procedure TMain_Form.visita(k:integer);
var
 arc, arc1 : integer;
begin
   agora  := agora + 1;
   v[k]   := agora;
   for arc := No_ORIG[k] to (No_ORIG[k+1] - 1) do
   begin
      if (v[DESTINO[arc]] = 0) then
        visita(DESTINO[arc]);
   end;
   for arc := No_DEST[k] to (No_DEST[k+1]-1) do
   begin
      arc1 := IND[arc];
      if (v[ORIGEM[arc1]] = 0) then
        visita(ORIGEM[arc1]);
   end;
end;
```

(a) Seleciona 1o. Vértice do Grafo

```
agora := 0;
for k  := 1 to no_vertice do
begin
   if (v[k] = 0) then
   begin
      visita(k);
      v[k] := -v[k];  // Marco o Início do grafo
   end;
end;
```

(b) Sequência de Visitas no Grafo

Figura 2.5: Implementação em ObjectPascal do algoritmo de busca em profundidade
Fonte: Elaborada pelo autor.

Já a Figura 2.6 destaca, na cor vermelha, números representando a sequência da ordem de visita dos vértices realizada pela pesquisa em profundidade do Kraken em um subgrafo da BNFBE.

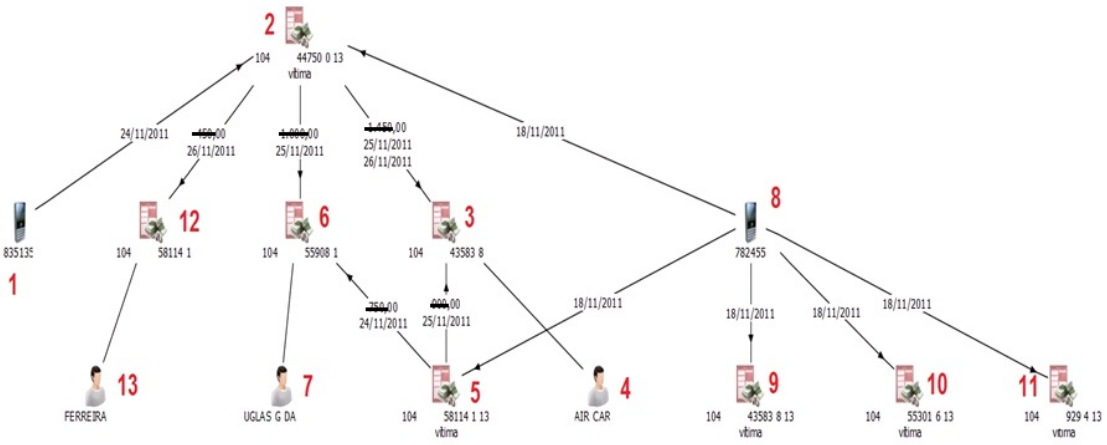


Figura 2.6: Exemplo de uma ordem de visita em vértices de um Grafo.
Fonte: Elaborada pelo autor

2.2 ANÁLISE DE VÍNCULOS

Segundo Harisson (1998), a análise de vínculos pode ser considerada uma técnica de mineração de dados com o objetivo de criar modelos baseados em padrões de relações estabelecidos pelas conexões entre registros.

Segundo Xu & Chen (2004) é indispensável a extração de informações referentes a entidades e suas associações em grande escala de dados brutos para estabelecer vínculos

em uma análise de relacionamento para gerar uma representação em rede. Na forma gráfica (grafo), os atores/entidades são representados por vértices ou nós, e as associações entre eles por uma teia ou rede. Durante as investigações criminais, os registros de dados transacionais armazenados em um banco de dados ou de documentos textuais não-estruturados são consolidados numa operação para formação dos relacionamentos.

Prosseguem os autores mencionando que a consolidação é um processo para que os dados minerados deixem de ser ambíguos, transformando-os em indivíduos específicos por meio da combinação de informação de identificação em uma chave única. Fórmulas heurísticas como: indivíduos que dividam o mesmo endereço, contas bancárias ou transações bancárias, entre outras, servem para formar os relacionamentos ou vínculos ente indivíduos consolidados (indivíduos específicos).

A análise de vínculos, se aplicada com sucesso em investigações de comportamento humano na observância de frequência de fatos e convergências que apresentam um padrão da atividade. Segundo Faccioni Filho (2013), a análise de vínculo pode ser usada em redes sociais e permite uma visualização de conexões de forma gráfica (ver Figura 2.7) por meio de sociogramas.

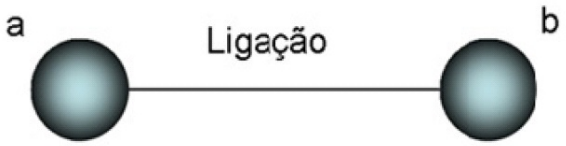


Figura 2.7: Ligação entre dois pontos.
Fonte: FACCIONI FILHO, 2013

Segundo Alves (2005), as investigações policiais contemporâneas envolvem a análise de uma enorme quantidade de dados, em múltiplos formatos, originados de três fontes básicas: (a) humanas, (b) de conteúdo e (c) de tecnologia ou tecnológicas.

As fontes humanas são oriundas de depoimentos, dados fornecidos por informantes, interrogatórios, entrevistas, denúncias e colaboradores.

As fontes de conteúdo podem ser retiradas de registros provenientes de notícias da mídia, sistemas bancários (como neste trabalho os registros oriundos da CAIXA), além de ocorrências policiais (vigilâncias, relatórios de análise), bem como de documentos de toda ordem (resultados de busca e apreensão, por exemplo), incluindo os chamados "cadastros".

Por último, as fontes de tecnologia, ou tecnológicas, têm sua expressão na telecomunicação, imagens e sinais eventualmente interceptados (interceptação telefônica e telemática), captados e devidamente analisados (por exemplo: autos-circunstanciados).

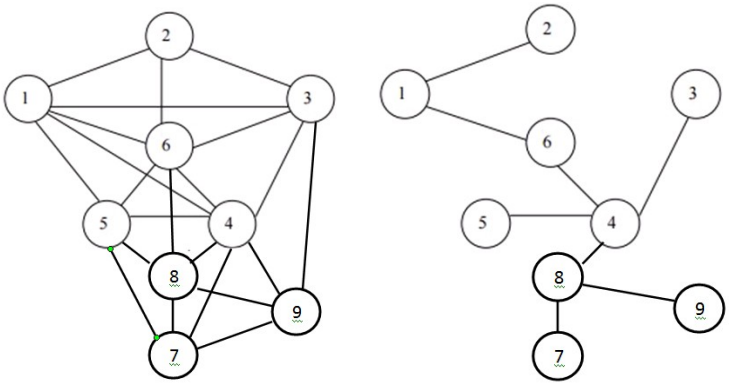
A razão de proporcionalidade entre as arestas (E) de um grafo e seus vértices (V) determina a sua densidade. Assim, grafos densos possuem muitas arestas (vínculos) interligando seus vértices, enquanto grafos esparsos têm poucas conexões por vértices, (POZZER, 2010).

A Equação 2.1 mostra a função para se calcular a densidade de um grafo.

$$densidade = \frac{|E|}{(|V| * (|V| - 1))/2} \tag{2.1}$$

Com a definição da densidade de um grafo, podemos escolher uma representação mais adequada para a estrutura de dados que irá manipular os seus dados. A implementação de um grafo esparso não é adequada para um grafo denso e a implementação usada para se representar um grafo denso não é eficiente para um grafo esparso.

A Figura 2.8 mostra exemplos de um "grafo denso" x "um grafo esparso".



(a) Grafo Denso (b) Grafo Esparso

Figura 2.8: Grafo Denso x Grafo Esparso

Fonte: Elaborada pelo autor.

Desta maneira, segundo Martins (2009), a análise de vínculos permite elaborar diferentes topologias de redes capazes de representar redes criminais de acordo com a Figura 2.9.

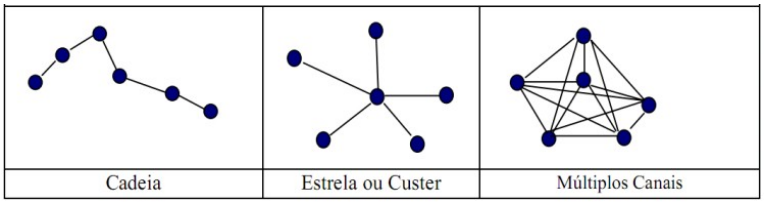


Figura 2.9: Estruturas de Redes.
Fonte: MARTINS, 2009

Conforme Dantas & Ferro Jr. (2006), a análise de vínculos é mais aplicada nas investigações de comportamento humano. Na área policial, utiliza-se de determinadas "pistas" (indícios) que estão ligadas entre si para solucionar crimes.

A análise de vínculos fornece indicações sobre a atuação de cada ator/entidade com o objetivo de apontar quais os atores mais relevantes num grafo. Podem ser realizadas as seguintes análises nos vértices:

- Grau;
- Similaridade;
- Recursos Alocados.

O **grau** de um vértice pode ser obtido somando-se o número de arestas/vínculos que saem e chegam neste vértice.

No caso específico deste trabalho a **análise de similaridade** de beneficiários é feita baseada na quantidade de contas vítimas que estes possuem em comum, de acordo com a Figura 2.10.

Segundo Berlusconi et al. (2016), existem várias medidas de similaridade entre vértices, sendo uma das mais simples aquela baseada na contagem do número de vizinhos em comum. Esta métrica pode ser adaptada para se mensurar o grau de de relacionamento entre contas beneficiárias da seguinte forma. Seja $V(B_i)$ o conjunto de contas vítimas associadas ao beneficiário B_i. Pode-se, então, definir similaridade entre as contas beneficiárias i e k baseada na interseção de contas vítimas por meio da Equação 2.2.

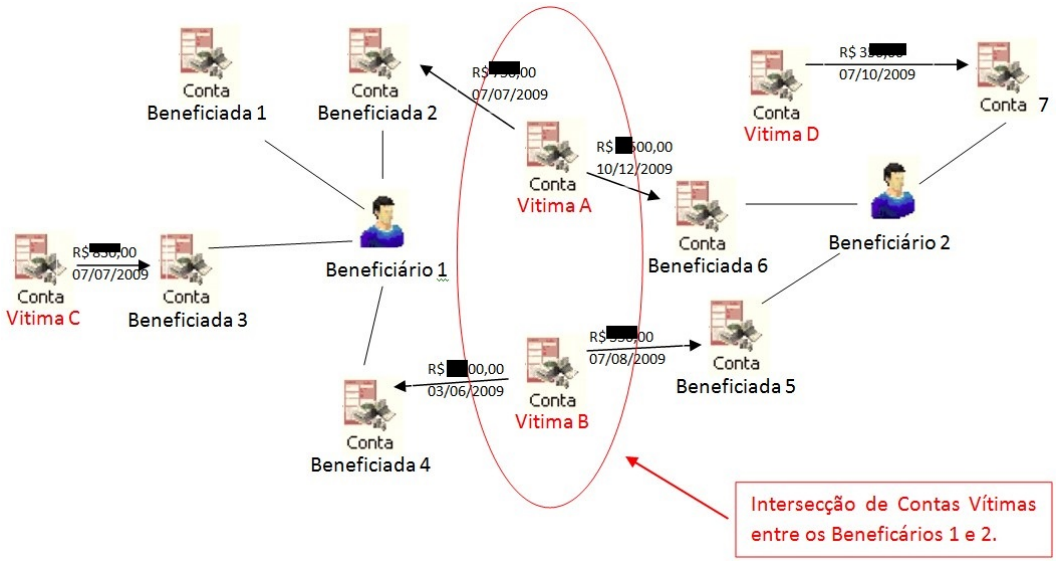


Figura 2.10: Similaridade de Beneficiários baseada na intersecção de Contas Vítimas.

Fonte: Elaborada pelo autor.

$$S_{ik} = \mid V(B_i) \cap V(B_k) \mid \tag{2.2}$$

A Figura 2.11 mostra um exemplo de como seria uma matriz de similaridade de beneficiários baseada na intersecção de contas vitimas, na qual B_1 significa beneficiário 1, B_2 significa beneficiário 2, e assim sucessivamente até o último beneficiário (B_n) do subgrafo em questão.

	B1	B2	B3	B4	...	Bn-1	Bn
B1	X	2	1	0	0	5	6
B2	X	X	0	3	4	1	1
B3	X	X	X	1	3	0	0
B4	X	X	X	X	0	1	0
|	X	X	X	X	X	2	0
Bn-1	X	X	X	X	X	X	1
Bn	X	X	X	X	X	X	X

Figura 2.11: Matriz Similaridade de Beneficiados baseada na intersecção de Contas Vítimas.

Fonte: Elaborada pelo autor.

Ainda segundo Berlusconi et al. (2016), essa medida de similaridade baseada em vizinhos em comum pode ser refinada de várias formas. Neste trabalho utilizou-se a **análise de recursos alocados** das contas vítimas em comum. Seja g_z o grau da conta vítima z. A Equação 2.3 define a similaridade entre as contas beneficiárias i e k baseada na alocação de recursos de contas vítimas em comum.

$$R_{ik} = \sum_{z \in V(B_i) \cap V(B_k)} \frac{1}{g_z} \tag{2.3}$$

O uso de softwares gráficos voltados para análise de vínculos, como: Gephi, Nexus e o Analyst´s Notebook trazem maior agilidade e, principalmente, maior compreensão na visualização e na exibição gráfica dos resultados das relações entre os diversos atores de um grafo. Contudo, para sua utilização existe a necessidade de se modelar um banco de dados com as relações entre suas entidades baseadas no modus operandi do que se deseja apurar ou investigar. Assim a análise de vínculos AV demanda o suporte conceitual da tecnologia do conhecimento.

2.3 FRAUDES ELETRÔNICAS

As fraudes bancárias eletrônicas são realizadas sem o conhecimento e a anuência dos titulares das contas bancárias.

As fraudes bancárias acontecem nos seguintes modos: *internet banking*, *mobile banking* e clonagem de cartão bancário.

As fraudes de *internet banking* acontecem online na modalidade "cartão não presente", quando o fraudador utiliza-se apenas dos dados do cartão bancário, como: nome da vítima, número do cartão, validade, código de segurança e/ou senha para efetuar as fraudes pela internet, que podem ser: transferências fraudulentas entre contas bancárias, compras em lojas da *internet*, pagamentos de boletos bancários ou créditos em celulares pré-pagos.

As fraudes de *mobile banking* são realizadas por meio de operações bancárias móveis feitas por dispositivos móveis de telecomunicações, como: *notebooks, smartphones, tablets*, entre outros.

Já na clonagem de cartão, as fraudes acontecem na modalidade "cartão presente". Neste caso, o fraudador clona fisicamente o cartão da vítima com todo os seus dados e o utiliza fisicamente para pagamentos de contas e saques em espécie.

O primeiro passo praticado pelos fraudadores para se realizar fraudes bancárias eletrônicas é conseguir os dados do titular da conta, usando para isso diversas técnicas, tais quais:

- Engenharia social, na qual o fraudador se passa por outra pessoa, por exemplo, dizendo-se ser do setor de segurança do banco da vítima e alegando que o cartão da mesma foi clonado e já encontra-se bloqueado por questão de segurança. O fraudador então orienta a vítima a cortar o cartão ao meio (sem danificar o chip) e afirma que em seguida um motoqueiro irá a sua residência pegar o cartão para que o mesmo seja substituído posteriormente. Outra técnica de engenharia social prática é induzir a vítima a entrar em um site falso do banco e recadastrar os seus dados e senha, por exemplo. Assim, existem diversas formas que o fraudador utiliza para enganar suas vítimas, obter seus dados cadastrais e até mesmo a senha do cartão bancário.

- Mensagens de *phishing scam*, com as quais os fraudadores encaminham *e-mails (spam)* para suas vítimas se passando por instituições isentas e de credibilidade como a Receita Federal, Polícia Federal, Detran, entre outras. Em todos esses *e-mails* os fraudados solicitam a vítima que acione um *link* nesse *e-mail* para responder as solicitações do suposto órgão. É nesse momento que é instalado um *malware* no computador/ celular da vítima. Cavalos de Tróia, ou *Trojan* abrem uma porta dos fundos no equipamento da vítima com o intuito de se ter acesso a segurança da máquina e permitir a entrada de outros *malwares*, *spyware* é projetado para espionar o dispositivo infectado da vítima atrás de senhas, telas de sistemas que a vítima usa, entre outros. Os *botnets* acessam o dispositivo infectado da vítima e o utilizam para práticas ilícitas pela *Internet*.

Segundo Peotta et al. (2011), os fraudadores brasileiros utilizam em conjunto o *phishing* e os *malwares* para realizarem os ataques às suas vítimas.

Existem casos em que o fraudador necessita estar presente, com uma cópia do cartão bancário (clonagem de cartão bancário) de sua vítima para realizar a fraude eletrônica, como: saques em dinheiro em caixas de auto atendimento **(ATM)** e compras em lojas físicas (supermercados, postos de gasolina, grandes magazines, etc). Nesses casos de clonagem de cartão, é comum a utilização de algum tipo de dispositivo eletrônico conhecido como chupa-cabra camuflado junto aos ATMs, ou até mesmo em máquinas de crédito adulteradas em pontos de vendas comprometidos (como postos de gasolina ou restaurantes), para que o fraudador tenha posse dos dados cadastrais e da senha bancária das suas vítimas.

Uma vez de posse desse dados cadastrais e senha, o fraudador é capaz de emitir um novo

cartão (clonado), que será utilizado mais tarde para a realização de fraudes bancárias eletrônicas contra suas vítimas.

Com as novas tecnologias de segurança implementadas pelas instituições financeiras está ficando cada vez mais difícil a clonagem de cartões bancários que utilizem *chip* para armazenamento dos dados confidenciais dos clientes.

2.4 PROJETO TENTÁCULOS

Segundo Siqueira et al. (2015), o estado da arte do projeto **Tentáculos** da Polícia Federal do Brasil está baseado no modelo divulgado no guia de acesso à BNFBE. Este guia de acesso foi elaborado pelo Grupo Permanente de Análise **(GPA)** do Serviço de Repressão a Crimes Cibernéticos **(SRCC)** da Diretoria de Combate ao Crime Organizado **(DICOR)** da PF.

> "Após a análise da contestação da fraude bancária na modalidade *internet banking*, *mobile banking* ou clonagem de cartão, a CAIXA pode negar o ressarcimento dos valores às vítimas, caso verifique que não houve fraude ou considere se tratar de auto fraude. Contudo, se for confirmada a fraude, a CAIXA ressarcirá os valores referentes às transações e a instituição financeira absorverá este prejuízo. Nesse caso, é obrigatório que a CAIXA abra um processo dito processo de contestação, com todas as informações colhidas pelo setor de segurança bancária da CAIXA referentes à fraude praticada contra aquela conta. Neste processo de contestação a CAIXA informa: os Internet Protocols **(IP)** que acessaram a conta, valores transferidos para outras contas da CAIXA, valores transferidos para contas de clientes de outros bancos, pagamentos de concessionárias de serviços públicos, pagamentos de tributos e títulos de cobrança, recargas de celulares, saques, clientes vítimas ou beneficiários de transferências, Bancos, Agências e Terminais envolvidos nas transações. Toda essa informação relativa a este tipo de fraudes é consistida e organizada no BNFBE para a verificação de pontos em comum. O objetivo dessa análise é diminuir o número de procedimentos investigatórios (uma vítima = um IPL), e otimizar recursos humanos e materiais nas investigações." (SIQUEIRA et al., 2015).

A redução do retrabalho dos policiais é possível, uma vez que uma quadrilha de fraudadores pode vitimizar inúmeras contas bancárias ao longo do tempo e do território

brasileiro. Concentrar o foco na quadrilha de fraudadores reduz significativamente os IPLs. O MAFBE usa ferramentas como o *Ibase* e *Analyst's Notebook* do i2 da IBM para o armazenamento, organização, controle e análise dos dados encaminhados pela CAIXA.

Com a ajuda dessas ferramentas, o Policial Federal realiza de forma manual a investigação das fraudes, conseguindo exibir informações complexas e realizar a análise de vínculos objetivando encontrar evidências correlatas entre as fraudes. As funcionalidades de modelação e análise do MAFBE baseiam-se nos conceitos de entidades e vínculos. As entidades são os atores que estão sendo representados, como: Processo-banco, conta, agência, cartão, IPs, terminal, pagamentos, órgão/concessionária, veículo, telefone, pessoa, terminal de conexão, identificador de máquina, local, ERB - Estação Rádio Base, operação, informação, IPL, laudo pericial, relatório de análise, notícia-crime. Os vínculos representam as relações entre estas entidades, como: transferências bancárias, saques, pagamentos, data/hora de acesso, etc. A Figura 2.12 mostra a relação gráfica entre alguns atores e seus respectivos vínculos, utilizando o *software Analyst's Notebook*.

Estabelecer vínculos pressupõe associar dados, condutas, eventos, entidades ou quaisquer outros elementos de um empreendimento criminal complexo, subsidiando a ação policial no sentido de permitir uma visão esclarecedora de um determinado comportamento ou ação delitiva, possibilitando o alcance de resultados efetivos na consecução de operações de inteligência/investigação policial (DANTAS et al., 2007).

Segundo Siqueira (2014) , os tipos de vínculos numa abordagem policial-investigativa podem ser diretos ou indiretos. Os vínculos diretos independem de complemento, como: transferência entre contas, registros de logs no servidor da instituição bancária, dispositivo (*smartphone*, *tablet*, computador, etc) utilizado na conexão com a conta vítima, conexão de *Internet*, pagamentos (transações bancárias) e recargas de celulares. A Figura 2.13 representa graficamente os vínculos diretos entre atores.

Vínculos indiretos necessitam de outra técnica investigativa para serem comprovados, como: história-cobertura, infiltração, interrogatório, busca e apreensão, etc. Logo vínculos indiretos não estão, num primeiro momento, disponíveis na BNFBE.

Segundo Siqueira (2014), os vínculos indiretos geralmente demonstram a Análise de Densidade de Ocorrência de Fenômenos **(AD)** ou *density analysis*, ou mancha criminal.

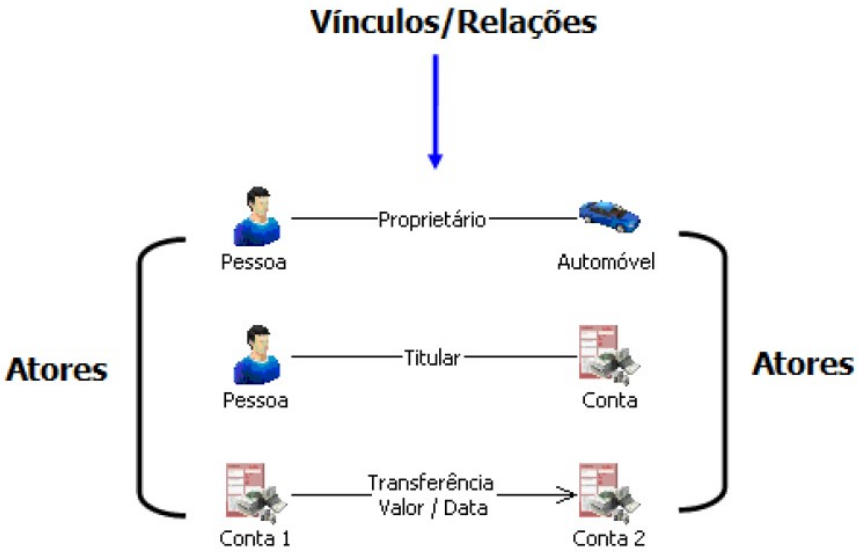


Figura 2.12: Atores e respectivos vínculos/relações. *Software Analyst's Notebook* 8.
Fonte: SIQUEIRA, 2014

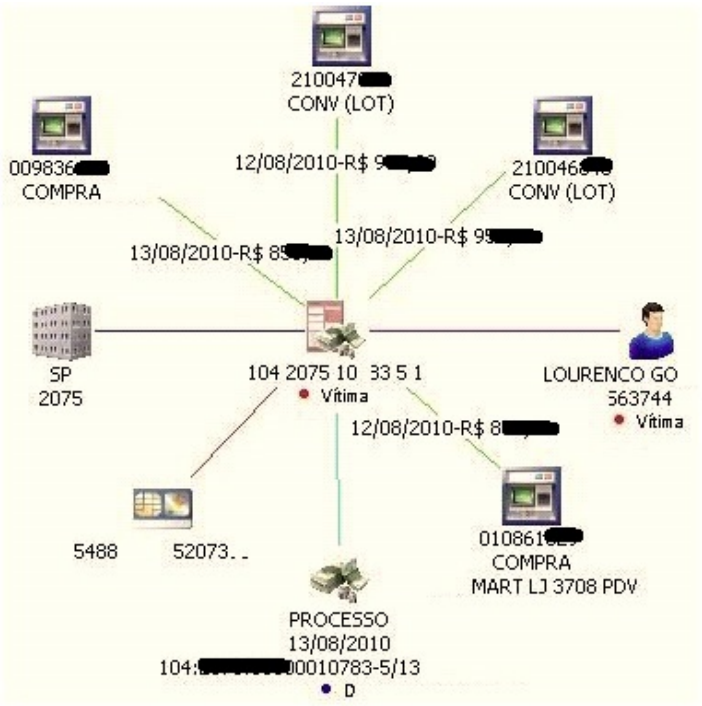


Figura 2.13: Vínculos diretos entre atores. *Software Analyst's Notebook* 8.
Fonte: SIQUEIRA, 2014

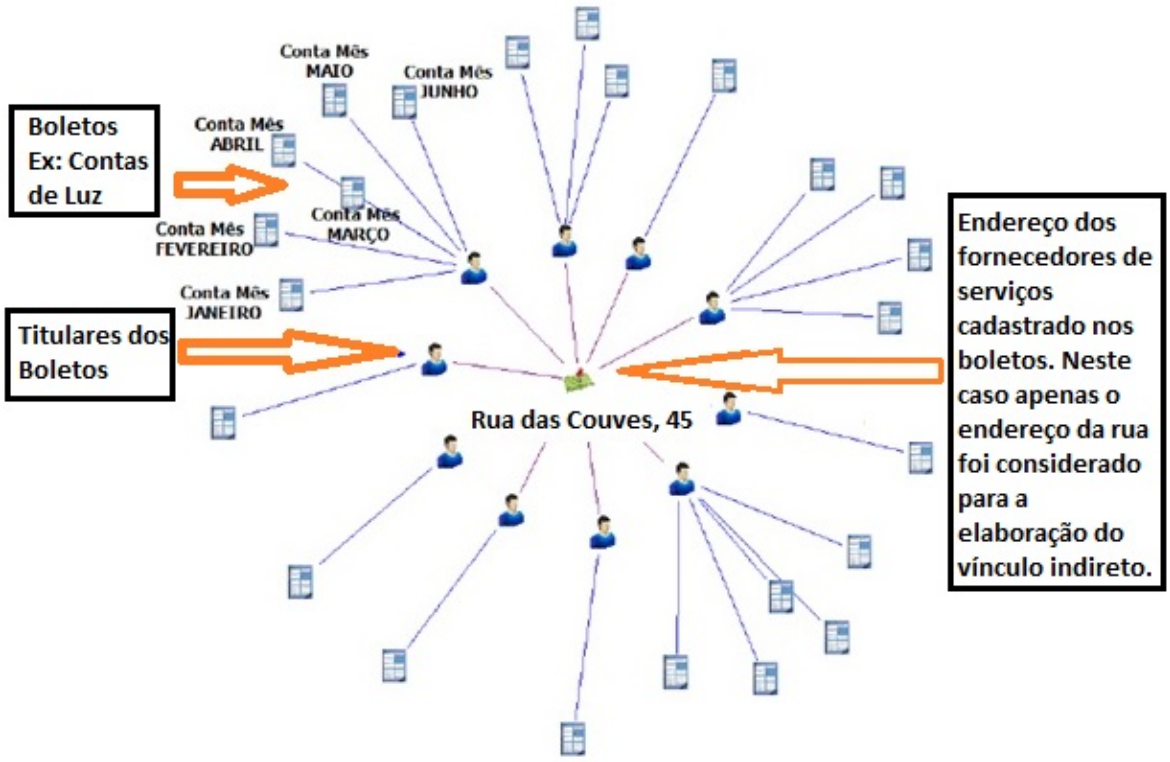


Figura 2.14: Vínculos indiretos entre atores. *Software Analyst's Notebook* 8.
Fonte: SIQUEIRA, 2014 (Adaptada pelo autor)

Os vínculos indiretos oriundos de investigações feitas em campo, por Policiais Federais, podem ser inseridos mais tarde na BNFBE. Contudo, esses não fazem parte da abordagem deste trabalho.

Segundo Dantas et al. (2007), aplica-se análise de ocorrência de fenômenos no mapeamento criminal (por intermédio de bases inter-relacionáveis de dados) tendo como finalidade mostrar a densidade ou concentração de fatos, parte de um fenômeno, em um determinado espaço e tempo.

No MAFBE, podemos citar como vínculos indiretos: localidade das agências das contas beneficiárias, endereços dos beneficiários dos pagamentos fraudulentos, entre outros.

A Figura 2.14 representa graficamente os vínculos indiretos entre atores.

No MAFBE, o Policial Federal é treinado para o uso da BNFBE. Após esse treinamento, o Policial Federal investigador está habilitado a realizar pesquisas por meio da imposição de parâmetros diretamente nas ferramentas *Analyst´s Notebook e IBASE* da IBM.

O investigador Policial Federal deve se logar no *Ibase* da IBM e abrir o banco de dados da BNFBE como mostrado na Figura 2.15.

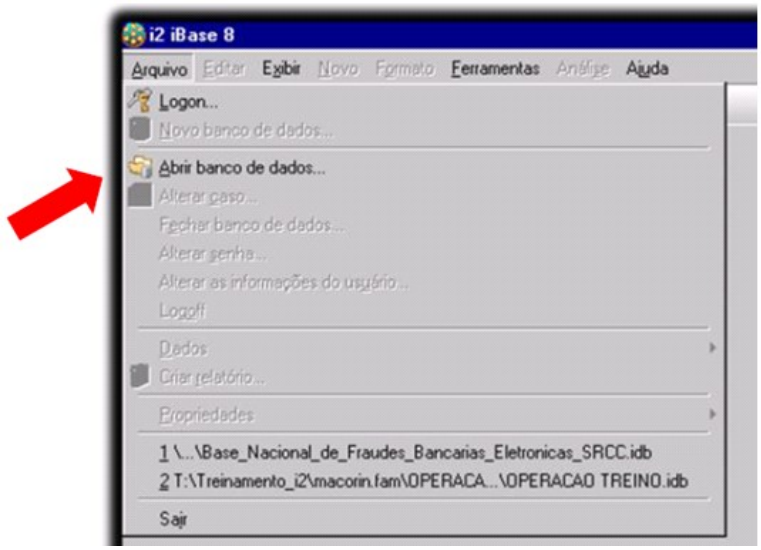


Figura 2.15: Selecionando à BNFBE. *Software Ibase* 8.
Fonte: SIQUEIRA et al., 2015

O PF, a partir de elementos investigativos obtidos em atividades externas, inicia a investigação selecionando as entidades com as quais ele deseja trabalhar na BNFBE para elucidar os crimes de fraudes bancárias eletrônicas, de acordo com a Figura 2.16.

Note-se que vínculos diretos entre os atores das fraudes bancárias eletrônicas já foram pré-estabelecidos durante a modelagem do BNFBE desde 2008. Eventuais ajustes foram e são feitos ao longo dos anos na BNFBE pela PF.

No *Analyst´s Notebook* o investigador Policial Federal realiza consultas, de forma gráfica, do tipo *Structured Query Language* (**SQL**) na BNFBE. O que o *Ibase* faz é facilitar, de uma maneira gráfica, a construção deste *SQL* para que os investigadores consigam obter resultados visuais de suas consultas (investigações) na BNFBE.

Nesse método, que utiliza a análise humana (investigadores Policiais Federais) na elaboração dos grafos que relacionam os atores de uma fraude bancária eletrônica, podemos destacar:

- O grande trabalho que o investigador tem na elaboração dos grafos;
- A necessidade de várias iterações do investigador Policial Federal com o *Ibase e*

Figura 2.16: Selecionando a entidade "PESSOA"da BNFBE. *Software: Ibase* 8.
Fonte: Elaborada pelo autor.

o Analyst's Notebook, que demandam tempo e treinamento específico do investigador;

- O Policial Federal somente consegue verificar se a investigação em curso já sofreu algum tipo de análise policial depois de realizar iterações nas ferramentas *Analyst´s Notebook e Ibase* da IBM;
- Os resultados da análise humana somente são conhecidos depois de realizadas todas as iterações de interesse da investigação na BNFBE utilizando as ferramentas da IBM.

Por óbvio, a metodologia de levantamento de informações e da própria investigação utilizada pela Policia Federal é sigilosa e não poderá ser descrita em detalhes neste trabalho.

Assim, a importância da abordagem deste trabalho é a de automatizar o método de investigação humana. E, por meio do uso da metodologia já existente no MAFBE, dar celeridade às análises das investigações e destacar os principais atores nos grafos das

fraudes bancárias eletrônicas. Para tanto, fornecendo antecipadamente aos investigadores quais registros de fraudes iniciarão investigações prioritárias, baseados em filtros pré-estabelecidos em métricas objetivas, como quantidade de contas vítimas, quantidades de contas beneficiadas, total das fraudes em reais por investigações (grafos), entre outros.

2.5 TRABALHOS RELACIONADOS

De acordo com Berlusconi et al. (2016), o problema da previsão de ligações (vínculos) tem recebido um aumento da atenção dos estudiosos da ciência da rede. Na análise das redes sociais **(ARS)**, um dos seus objetivos é recuperar as ligações que faltam, ou seja, as ligações entre os atores susceptíveis de existir, mas que não foram comunicadas porque os dados estão incompletos ou estão sujeitos a vários tipos de incertezas.

No que diz respeito as investigações criminais, os problemas de informação incompleta são encontrados quase por definição, dadas as óbvias estratégias anti-detecção criadas pelos criminosos e os limitados recursos investigativos.

Os autores desse artigo trabalharam com um conjunto de dados específico obtido a partir de uma investigação real, e se propuseram a elaborar uma estratégia para identificar os elos (vínculos) que faltam em uma rede criminosa com base na análise topológica das ligações classificadas como marginais, isto é, foram removidos durante o procedimento de investigação.

A principal hipótese é que os elos em falta devem ter características opostas em relação aos marginais.

Segundo os autores, a inspeção dos documentos de fonte judicial confirma que as ligações previstas, na maioria dos casos, relacionam os atores com grande facilidade de co-participação em atividade ilícitas. Diferente dos estudos anteriores, o principal pressuposto é que os elos ausentes podem ter características contrárias àquelas das ligações marginais descartadas durante a investigação.

O vazio de dados gera problemas da incerteza da informação, potencialmente prejudicando a eficácia das investigações.

Para comprovar essa abordagem os autores utilizaram um conjunto de dados da operação

Oversize. Essa investigação durou 7 (sete) anos de 2000 a 2006 e trata-se de um caso criminal, no qual um grupo mafioso italiano foi acusado de tráfico internacional de drogas. O julgamento durou dois anos de 2007 a 2009.

Utilizou-se de análise de similaridade de vizinhança e análise de recursos alocados entre os nós (atores da investigação) para estabelecer relação entre os atores dos grafos baseado em três tipos de fontes de dados: registros de escutas telefônicas (grampos telefônicos), dados recolhidos em mandados de prisão e informações extraídas durante o julgamento dos envolvidos. Restringiu-se a análise para caso de vínculos não direcionados, negligenciando a direcionalidade dos links (se o link sai do ator "A" e vai para o ator "B" ou vice-versa).

A primeira hipótese é a de que as ligações removidas durante a investigação são caracterizadas pela baixa intermediação. Isso significa que elas são redundantes no sentido de que conectam indivíduos que já estão conectados de alguma forma na rede e não melhoram significativamente o fluxo de informações.

A segunda hipótese alternativa baseia-se na literatura sobre previsão de ligação. Vários estudos têm aplicado diferentes métodos de previsão de ligação para um número de redes. Eles mostram que os nós são mais propensos a serem conectados quando eles são semelhantes e compartilham uma série de recursos. Assim sendo, de acordo com a segunda hipótese, as ligações marginais conectam nós estruturalmente diferentes, isto é, indivíduos que ocasionalmente colaboram, mas são diferentes em termos de interesse, antecedentes e envolvimento em atividades criminosas. Portanto, essas conexões não são cruciais para as condutas criminais. A literatura propõe várias estratégias analíticas para a predição de ligação, com novos métodos constantemente adicionados, principalmente com base em medidas de similaridade nodal. Dado ao pequeno tamanho da rede da operação Oversize, tais estratégias são uma opção viável, Uma vez que o cálculo exaustivo de semelhanças para todos os pares de nós é computacionalmente viável.

As abordagens de similaridade de nós atribuem uma pontuação a todos os pares de nós (x, y) e, consequentemente, induzem uma classificação a todos os pares de nós. Dentre as muitas possíveis pontuações de similaridade, a mais simples é a contagem do número de vizinhos comuns **(VC)**.

A pontuação de similaridade baseada em VC pode ser refinada de muitas maneiras,

por exemplo, pesando não simplesmente a contagem do número de vizinhos comuns. Um desses caminhos leva à definição do escore de similaridade de Alocação de Recursos **(AR)**.

VC e AR são amplamente utilizados para quantificar a semelhança do nó. Testes extensivos sobre a capacidade de um amplo conjunto de indicadores (incluindo os dois acima) na resolução do problema de previsão de link descobriram que a VC obtém um desempenho muito bom apesar de sua extrema simplicidade, enquanto a AR classifica como um dos melhores indicadores em um grande conjunto de testes de *Bechmark* (LU; ZHOU, 2011).

No trabalho de Berlusconi et al. (2016), mostrou-se que a ARS pode apoiar a análise de inteligência criminal e as investigações em curso, identificando os elos em falta entre os suspeitos. Estes estudos demonstraram que a similaridade de nós, já aplicada em diferentes campos de predição de ligação, pode identificar possíveis elos em falta também em redes criminosas, quando a informação tem ruído (interferência) ou está incompleta quase por definição. O sistema de justiça criminal implementa uma série de garantias contra falsos positivos como acusações incorretas e interações não relacionadas a condutas criminais. Em contrapartida, estratégias eficazes para evitar falsos negativos, tais como falta de informação são escassas. Devido aos limitados recursos de coleta de dados, as agências de aplicação da lei podem, de fato, perder alguns atores e ligações, com consequências negativas nas atividades de inteligência e investigação. Isto aplica-se a redes de tráfico de drogas, como a rede Oversize, gangues de rua e grupos terroristas, entre outras. As medidas de similaridade do nó ajudaram a identificar as características dos links removidos.

De acordo com Siqueira (2014), cresce o desafio para as forças policiais combaterem os grupos criminosos especializados em práticas de fraudes bancárias pela *Internet*. O trabalho de Siqueira versa sobre a utilização da técnica de análise de vínculos para incrementar as investigações policiais. No seu trabalho, Siqueira descreve a trajetória e a aplicação do Projeto Tentáculos da PF, que centralizou as notícias-crime em um único banco de dados e utilizou-se da análise de vínculos para elucidar crimes de fraudes bancárias, praticadas pelo canal *Internet banking*, contra a CAIXA.

> "A aplicação deste modelo reduziu fortemente o número de inquéritos policiais instaurados nos últimos anos e ampliou a capacidade investigativa, mapeando os grandes grupos criminosos e selecionando o melhor local

para a atuação policial. A pesquisa detalha o modus operandi e os atores envolvidos nas fraudes realizadas no canal *internet banking*. Descreve as funções e papéis de cada um no mercado clandestino Brasileiro, propondo uma doutrina investigativa e uma nomenclatura aplicável nos processos relativos a esse tipo de fraude na persecução penal brasileira. Ao final propõe um modelo de análise de vínculos, para uso na segurança pública brasileira, aplicável nas mais diversas condutas criminosas. O resultado é uma melhor atuação investigativa, evitando a duplicidade de investigações e, principalmente, otimizando e agilizando a ação policial por meio de parcerias entre instituições públicas e privadas." (SIQUEIRA, 2014).

Schaeffer (2007), este apresenta um *survey* em que faz uma síntese das definições e dos métodos de agrupamento de grafos, ou seja, encontrar conjuntos de vértices relacionados em grafos. Neste trabalho foram apresentados algoritmos globais para produzir um agrupamento para todo o conjunto de vértices de um gráfico de entrada. Algumas ideias sobre as áreas de aplicação de algoritmos de *cluster* grafos foram dadas.

As abordagens globais existentes são capazes de lidar com até alguns milhões de vértices em grafos esrparsos (HOPCROFT et al., 2003). Nesse *survey* são tratados os métodos de agrupamento de grafos: iterativo , hierárquico, divisivo e aglomerativo.

- Em um agrupamento hierárquico, cada nível da hierarquia de agrupamento define um subconjunto diferente e geralmente os *clusters* definidos pelos níveis mais altos contêm o *cluster* do nível inferior como subgrafos. Os métodos de *cluster* que produzem agrupamentos em vários níveis são chamados de algoritmos de agrupamento hierárquico. Um agrupamento hierárquico geralmente é construído gerando uma sequência de partições, onde cada *sub-cluster* pertence a um *super-cluster* em sua entidade. O *cluster* raiz contém, no máximo, todos os dados e cada um dos *cluster* ditos folhas contém pelo menos um elemento de dados. O *cluster* semanticamente relevante geralmente aparece em níveis intermediários;
- O agrupamento iterativo acontece quando o agrupamento é feito atribuindo um elemento de cada vez a um *cluster* apropriado. No agrupamento iterativo, as atribuições de *cluster* feitas aos elementos após seu primeiro processamento podem ser consideradas imutáveis ou podem ser alteradas posteriormente para otimizar alguma propriedade do agrupamento que está sendo calculada. Se um algoritmo

de agrupamento opera um dado de cada vez, tem apenas o conhecimento de dados encontrados antes. Nesse caso, o agrupamento iterativo também é conhecido pelo nome de *on-line*;

- Os algoritmos de agrupamento divisivo são uma classe de métodos hierárquicos que trabalha de cima para baixo, particionando recursivamente o gráfico em *cluster*. A divisão em cada iteração é feita tipicamente em dois conjuntos, mas não há nenhuma razão pela qual um algoritmo de agrupamento não poderia dividir um conjunto de vértices em mais de dois conjuntos para a próxima iteração. Os vários critérios para determinar onde dividir o grafo são: cortes, fluxo máximo, métodos espectrais, intermediação, tensão e potencial, cadeias de *Markov* e passeios aleatórios;

- Naturalmente, além de dividir o gráfico de cima para baixo em *cluster*, também pode-se trabalhar de baixo para cima mesclando conjuntos de vértices iterativamente em *clusters*. Esse último é o dito agrupamento aglomerativo.

De acordo com Schaeffer (2007), existem várias aplicações para agrupamentos de grafos, entre elas:

- Transformações de dados;
- Redes de informação e uso de informações;
- Sistemas de banco de dados;
- Redes biológicas e sociológicas;
- No mundo dos negócios. empresariais.

De acordo com Ferro Jr. (2007), a complexidade do mundo moderno vem direcionando as organizações para o desenvolvimento de uma infra-estrutura tecnológica com capacidade de processamento de informações devido à distribuição do conhecimento em forma de rede.

Diante da velocidade dos acontecimentos, da conectividade das pessoas e do aumento da complexidade da criminalidade, as organizações policias exigem cada vez mais da investigação criminal.

Continua o autor mencionando que a capacidade investigativa pode ser ampliada pela inteligência organizacional, apoiada na técnica de análise de vínculos e por meio de um modelo organizacional em rede de conhecimento. Este trabalho tem como base um projeto específico da Polícia Civil do Distrito Federal de nome "Sistema Cérebro", que aplica a tecnologia de análise de vínculos à investigação criminal.

O Cérebro está integrado nas bases de informação da organização que trafegam em rede compartilhada. Esse trabalho foi concebido a partir de investigações solucionadas por analistas especializados que operam o Cérebro.

3 DESCRIÇÃO DO PROBLEMA E PROPOSTA DE SOLUÇÃO

Este Capítulo mostra as dificuldades investigativas e como a proposta deste trabalho será implementada, assim como o modelo da BNFBE, a metodologia do Kraken, telas do ferramental (IG) criada para dar suporte a abordagem do Kraken, métricas utilizadas para mostrar ao investigador quais grafos gerados pelo Kraken devem ser priorizados, além dos filtros que a IG é capaz de realizar sobre o resultado dos grafos processados pelo Kraken.

3.1 DIFICULDADES INVESTIGATIVAS

O MAFBE de investigação de crimes de fraudes eletrônicas bancárias esbarra em diversas dificuldades que vão desde a logística necessária para levantar as informações e apurar esse tipo de crime (logística essa tanto das instituições financeiras como da própria PF), passa pela utilização de canais de difícil acesso da **Internet** pelos fraudadores (*deep web*, *lan-house*, etc), até chegar no problema da distribuição geográfica da quadrilha de fraudadores que comete este tipo de crime. Muitas vezes essa quadrilha de fraudadores está localizada em estados do Brasil diferentes dos de suas vítimas e beneficiários. A seguir são elencados outros pontos que fazem com que a investigação criminal desse tipo de ilícito seja de difícil apuração:

- Escassez de recursos humanos (investigadores Policiais Federais);
- No MAFBE, o início da análise dos vínculos das fraudes eletrônicas baseia-se em parâmetros de pesquisa que são inseridos manualmente por um investigador policial, o que demanda tempo, impossibilitando que o Policial Federal verifique todos os registros fraudulentos encaminhados pela CAIXA em um prazo aceitável. Nessa metodologia, a quantidade de vítimas, os montantes envolvidos nas fraudes e a localidade dos beneficiários somente são conhecidos ao final da investigação, após várias intervenções manuais dos Policiais Federais;
- No MAFBE a elaboração do relatório de análise da investigação policial consome bastante tempo e inúmeras iterações do investigador junto ao *Analyst´s Notebook* da IBM;

- A dificuldade dos investigadores de localizarem vínculos diretos em grafos que contêm centenas de ligações, mesmo com a utilização do *Analyst´s Notebook* da IBM;

- Dados que por ventura a CAIXA tenha entregue de forma extemporânea são inseridos a posteriori na BNFBE, o que pode causar um retrabalho para os investigadores ou interferir em relatórios de análise de investigações já elaborados e entregues pelos investigadores;

- A verificação pelo investigador do peso ou de uma relação fraca dos vínculos entre os vértices (atores) do grafo, para que esses sejam desconsiderados da investigação, de forma a não unir os subgrafos, que em tese, devem ser desconexos por se tratarem de investigações distintas. Esses vínculos ou vértices (atores) são classificados como marginais pelos investigadores, isto é, são removidos pelo investigador durante o procedimento de investigação.

3.2 MODELO DE DADOS DA BNFBE

Nesta seção é abordado o modelo que representa um espelho da BNFBE e as customizações que foram feitas em algumas de suas tabelas para a implementação da IG do Kraken.

Ao receber os dados em formato de arquivo texto da CAIXA, a PF realiza um ETL para o banco de dados *MySql* 5.6.1, para um *schema* de nome "tentaculos_saida".

Depois de homologada esta importação para o schema "tentaculos_saida" do *MySql*, a PF realiza outro ETL, agora para a BNFBE modelada no *IBASE* do i2 da IBM em um servidor *SQL Server*.

Para a realização deste trabalho tivemos acesso ao "tentaculos_saida" do *MySql*. A Figura 3.1 mostra as tabelas contidas nesse banco de dados.

Para se ter uma dimensão do problema de se analisar manualmente o banco de dados "tentaculos_saida": somente em uma de suas tabelas principais, a tabela "trn" (transações bancárias), existem hoje cerca de 2 milhões de registros fraudados. A tabela de agências bancárias possui cerca de 250 mil agências cadastradas. A tabela de "comprasaque" possui cerca de 1,3 milhões de registros, e, por fim, a tabela de

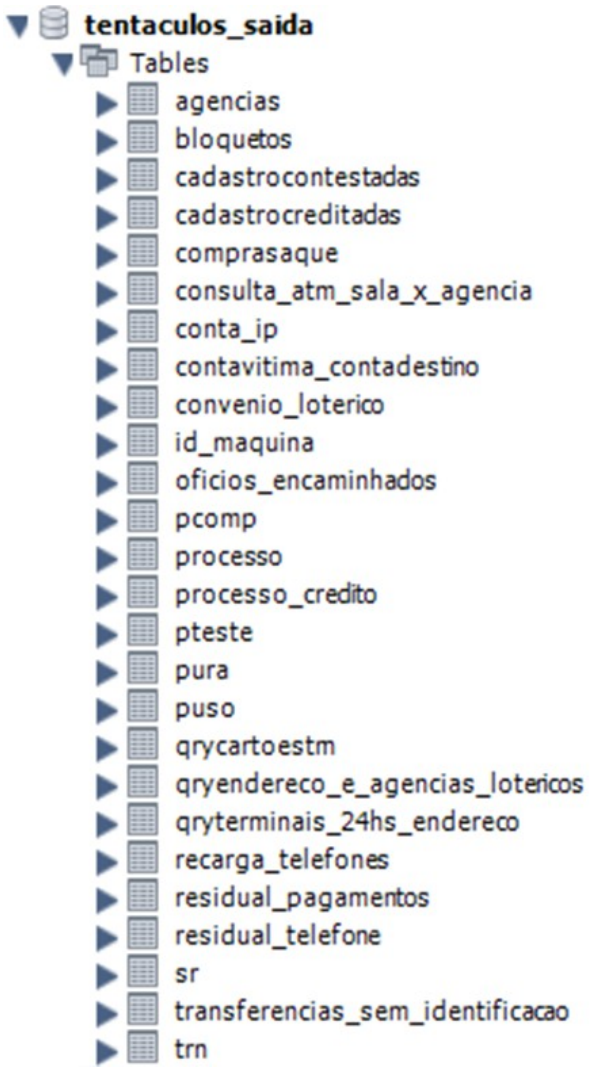


Figura 3.1: Tabelas que fazem parte do "tentaculos_saida".

Fonte: Elaborada pelo autor.

"cadastrocontestações" possui outros 1,3 milhões de registros, sem mencionar as demais tabelas. Esses dados são relativos somente a instituição financeira da CAIXA, no período de 2008 a março de 2016.

Na Figura 3.2 é mostrado o DER - Diagrama de Relacionamento entre as Entidades do "tentaculos_saida" que são utilizadas neste trabalho.

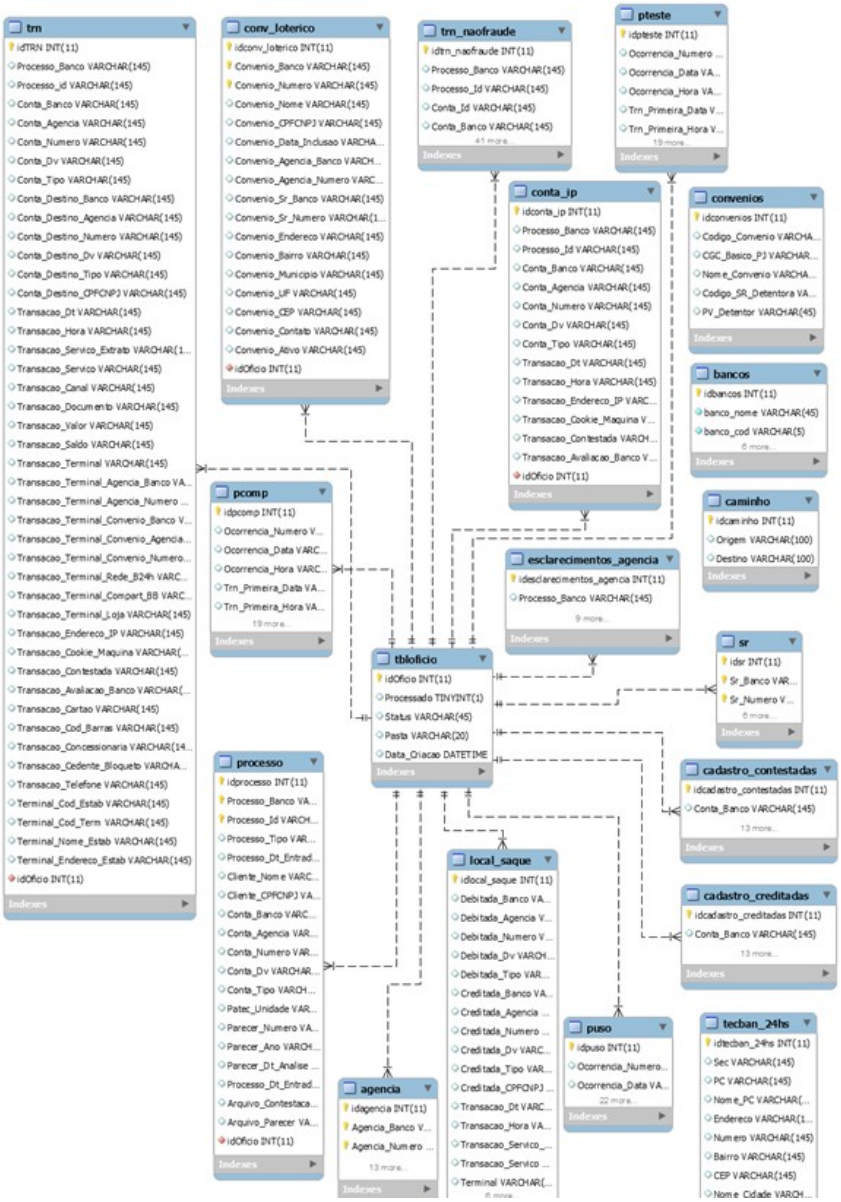


Figura 3.2: Diagrama de entidade e relacionamento do "tentaculos_saida".

Fonte: Elaborada pelo autor.

Para proceder com a criação do Kraken foi necessário fazer algumas customizações na tabela "cadastrocontestadas" para otimizar a programação da criação da representação compacta da lista de adjacência dos vértices dos grafos e na elaboração do algoritmo de busca em profundidade.

Originalmente, a tabela “cadastrocontestadas” era composta pelos campos mostrados na Figura 3.3

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
idCadastroContestadas	INT(11)	☑	☑	☐	☐	☐	☐	☐	☐	
Vitima	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Agencia	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Dv	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Tipo	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Cliente_CPFCNPJ	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Cliente_Dt_Nasc	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Cliente_Nome	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Fone	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_CEP	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Endereco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Localidade	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Uf	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Etiqueta	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
idOficio	INT(11)	☐	☐	☐	☐	☐	☐	☐	☐	NULL

Figura 3.3: Campos originais da tabela “cadastrocontestadas”.

Fonte: Elaborada pelo autor.

A Figura 3.4 mostra os campos que foram acrescentados à tabela “cadastrocontestadas”.

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
Realizou_Transferencia_Bancaria	INT(11)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Agencia	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Dv	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Tipo	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL

Figura 3.4: Campos adicionados à tabela “cadastrocontestadas”.

Fonte: Elaborada pelo autor.

Visto que a tabela “cadastrocontestadas” contém contas vítimas de vários tipo de fraudes bancárias eletrônicas, como: pagamento de boletos bancários, saques, transferências bancárias, compras pela *Internet*, entre outras, e o Kraken se propõe inicialmente a somente trabalhar com a priorização e análise de investigações voltadas para fraudes bancárias do tipo **transferências de valores entre contas bancárias**, foi necessário criar na tabela “cadastrocontestadas” o campo “Realizou_Transferencia_Bancaria”. Esse campo receberá o valor 1 (um), para diferenciar os registros desta tabela que sofreram este tipo de fraude.

O preenchimento desse campo é feito pela IG do Kraken e pode ser feito cruzando os dados da tabela “cadastrocontestadas” com a tabela “trn”. Essa última contém todas as transações fraudulentas que aconteceram com as contas vítimas da **CAIXA**. Nessa

tabela “trn” uma conta vítima pode ter sofrido “n” fraudes do mesmo tipo, ou de tipos diferentes ao longo do tempo. A Figura 3.5 mostra os campos da tabela “trn”.

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
idtrn	INT(11)	☑	☑	☐	☐	☐	☐	☑	☐	
Processo_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Processo_Id	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Agencia	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Dv	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Tipo	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Agenda	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Dv	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_Tipo	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Conta_Destino_CPFCNPJ	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Dt	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Hora	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Servico_Extrato	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Servico	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Canal	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Documento	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Valor	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Saldo	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Agencia_banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Agencia_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Convenio_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Convenio_Agencia	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Convenio_Numero	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Rede_B24h	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Compart_BB	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Terminal_Loja	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Endereco_IP	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Cookie_Maquina	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Contestada	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Avaliacao_Banco	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Cartao	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Cod_Barras	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Concessionaria	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Cedente_Bloqueto	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
Transacao_Telefone	VARCHAR(145)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
idOficio	INT(11)	☐	☐	☐	☐	☐	☐	☐	☐	NULL
pendencia	TEXT	☐	☐	☐	☐	☐	☐	☐	☐	NULL

Figura 3.5: Tabela “trn”.

Fonte: Elaborada pelo autor.

Sabe-se que se os campos: Conta_Destino_Banco, Conta_Destino_Agencia, Conta_Destino_Numero, Conta_Destino_DV e Conta_Destino_Tipo estiverem devidamente preenchidos no registro da tabela “trn”, aquele registro trata-se de uma fraude do tipo “transferência bancária”. Como os registros da tabela “trn” também possuem os campos que identificam a conta vitima, é possível relacionar este tipo de transação fraudulenta com a tabela “cadastrocontestadas”, e assim preencher o campo “Realizou_Transferencia_Bancaria” com o número 1 (um).

Para dar suporte a análise dos resultados deste trabalho foi necessário criar tabelas permanentes (estáticas) no banco de dados "Tentaculos_Saida", além das estruturas dinâmicas na memória *RAM* do computador. Essas tabelas ajudarão no processamento e na análise das fraudes, e armazenarão seus dados para serem confrontados com os dados obtidos no mesmo período de datas agora pela **análise humana** feita por Policiais Federais.

A Figura 3.6 mostra a lista das tabelas permanentes criadas no "Tentaculos_Saida" para a execução do Kraken.

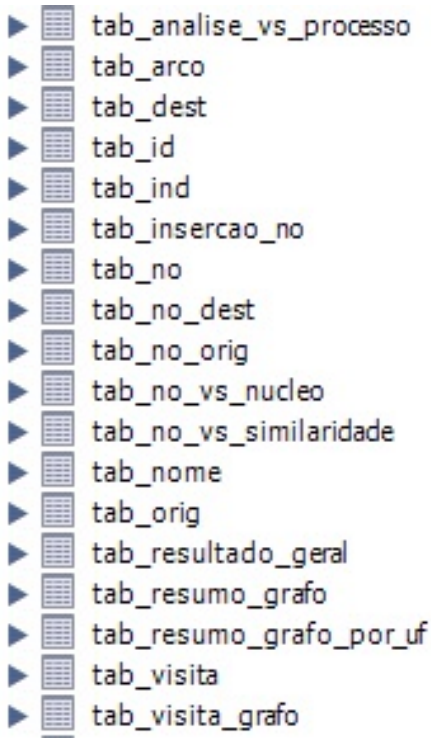


Figura 3.6: Novas tabelas acrescentadas ao banco de dados "tentaculos_saida".
Fonte: Elaborada pelo autor.

3.3 METODOLOGIA KRAKEN

Existem diversos problemas da computação nos quais podemos lançar mão dos grafos e da análise de vínculos para a sua solução. Segundo West (2000), os grafos auxiliam na representação e os seus vínculos na manipulação de conexões entre pares de objetos.

A abordagem proposta neste trabalho – denominada Kraken – é baseada na análise de vínculos entre atores (vértices) de um grafo que representa uma ou parte de uma investigação policial de fraudes bancárias eletrônicas.

A Figura 3.7 ilustra a abordagem proposta.

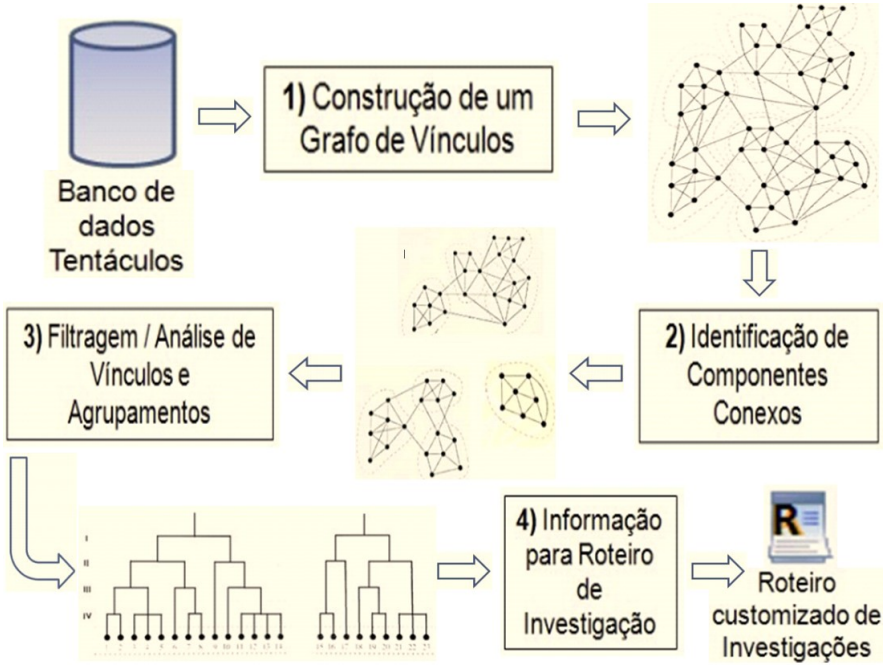


Figura 3.7: Representação gráfica da abordagem proposta – Kraken.
Fonte: Elaborada pelo autor.

Nessa abordagem, as entidades investigadas na BNFBE (contas bancárias vítimas e beneficiadas, telefones que acessaram contas vítimas, pessoas beneficiadas, etc) são representadas pelos vértices dos grafos, enquanto os seus vínculos ou relacionamentos (valores das transferências entre contas bancárias, datas destas transações bancárias, datas das ligações telefônicas para contas vítimas e a titularidade de contas beneficiadas) existentes entre as entidades são representados no Kraken como sendo as arestas que ligam os vértices dos grafos.

O resultado desta representação passa por uma análise de grupamento, na qual são separados grafos conexos através de um algoritmo de busca em profundidade – *Depth-First Search* para tanto. Para o Kraken, grafos conexos representam grupos criminosos. O Kraken possui uma IG para apresentação dos seus resultados, que permite ao Policial Federal realizar filtros do período de datas da investigação antes mesmo do processamento dos dados da BNFE.

Ao final do processamento, a IG exibe em forma de tabelas o resumo de quantidades de vítimas, de beneficiários, valor total das fraudes, e estados do Brasil envolvidos para cada grafo conexo (entre outras informações), permitindo ao investigador realizar filtros e/ou reordenar os grafos baseados nas métricas acima, **redirecionando e priorizando**

assim a sua investigação policial. Com base nos possíveis caminhos entre os vértices, o Kraken gera um roteiro com o fluxo do desvio do dinheiro desde sua origem até o destino final e comandos em SQL. Ambos auxiliarão ao Policial Federal na geração dos relatórios de análise e diagramas de elos das fraudes bancárias existentes na BNFBE, junto as ferramentas do *Ibase* e do *Analyst's Notebook* da IBM.

3.3.1 CONSTRUÇÃO DE UM GRAFO DE VÍNCULOS

O grafo é representado internamente no Kraken por uma lista de adjacência de cada vértice utilizando a representação *compact forward and reverse star representation* gerada a partir das tabelas existentes na BNFBE e que por se tratar de um banco de dados relacional permite determinar o relacionamento (vínculos) entre os diversos tipos de entidades (telefones, contas vítimas, contas beneficiárias, etc) existentes em suas tabelas.

Com este modelo pode-se atribuir pesos p_e as arestas $e \in E$, tais como o montante da fraude. Além de ser possível recuperar informações sobre período de datas, a localidade, etc. Dessa forma, por exemplo, os caminhos entre vértices do grafo podem representar o fluxo do dinheiro das contas vítimas para as contas beneficiárias e valores de pagamentos fraudulentos. De uma maneira simples pode-se definir o BNFBE como uma coleção de vários conjuntos (possivelmente disjuntos) envolvendo contas vítimas, contas beneficiárias, pessoas, localidades, transações, pagamentos, ou telefones que acessaram contas bancárias. Estes conjuntos passam então a ser representados na abordagem proposta pelo Kraken através de um grande grafo (formado, eventualmente, por vários subgrafos conexos).

3.3.2 IDENTIFICAÇÃO DE COMPONENTES CONEXOS

O grande problema na identificação de componentes conexos é como processar essa enorme e intricada massa de dados (BNFBE), utilizando a análise de vínculos, para qualificar os indivíduos (vértices/atores) e desta maneira poder traçar os relacionamentos e seus respectivos significados. Para tanto é importante conhecer o *modus operandi* do crime que se deseja modelar, o que chamou-se neste trabalho de **regra de negócio do crime**.

As regras de negócio, foco deste trabalho, foram baseadas na heurística aplicada pelos investigadores da PF, empenhados na elucidação de crimes de fraudes bancárias eletrônicas relativas a transferência de valores entre contas bancárias. Essas regras de

negócios estão sedimentadas, primordialmente, nas experiências profissionais e intuição desses investigadores.

As seguintes regras de negócios foram aplicadas pelo Kraken na BNFBE, em um determinado período de datas, para a construção de grafos e a identificação de seus componentes conexos:

a) Verificação se existiram ligações feitas por fraudadores de um telefone para uma URA da CAIXA, técnica usada pelos fraudadores para descobrir a quantidade de dinheiro que a vítima possui em sua conta bancária (identificando-se o terminal telefônico utilizado para entrar em contato com a URA);

b) Transferência de valores entre uma conta vítima e uma conta beneficiada (identificando-se a agência e o endereço da conta vítima, da conta beneficiada, a data e o valor da transferência realizada);

c) Identificação do titular da conta beneficiada (qualificação do titular da conta beneficiada, endereço).

Faz-se então a necessidade da implementação de um algoritmo de **busca em profundidade** – ***Depth-First Searh*** para se automatizar a análise de vínculos diretos na BNFBE, que leve em consideração as regras de negócio mencionadas acima, realizando assim o agrupamento de vértices que estejam envolvidos em uma determinado subgrafo, em tese, de uma determinada investigação policial.

3.3.3 FILTRAGEM / ANÁLISE DE VÍNCULOS E AGRUPAMENTOS

Os vértices de uma árvore de um subgrafo possuem geralmente vizinhos próximos (sobretudo no caso de fraudes bancárias na modalidade cartão bancário clonado), pois possuem ligações (arestas) fortes de vínculos diretos e até mesmo porque a fraude com cartões clonados acontece em uma área geográfica mais restrita.

Porém, algumas análises podem indicar elos fracos entre os vértices (principalmente no caso de fraudes bancárias nas modalidades *Internet banking* e *mobile banking*), por exemplo quando a maioria das fraudes ocorreram em um determinado estado do Brasil e em apenas um caso isolado uma fraude acontece em outra localidade bem diferente. Essa situação pode levar a eliminação desse vínculo e, por consequência, do subgrafo a ele ligado, caso este seja o único caminho de ligação entre o caso de fraude isolado com

a maioria dos outros casos. Essa eliminação segue critérios baseados na inteligência de investigação da PF, não sendo ainda foco deste trabalho.

Deve-se levar em conta que apenas a análise de vizinhança local dos vértices pode não ser suficiente, devendo-se investigar a relação entre vértices distantes e, se verificado um elo forte, estes devem integrar o subgrafo de investigação.

Assim, o filtro baseado na análise de vizinhança não deve se restringir a uma abordagem local (vértices próximos), visto que, por exemplo, existem casos nos quais as contas bancárias vítimas estão espalhadas em diferentes estados do Brasil, enquanto que as beneficiárias estão em outros estados e, por fim, o fraudador encontra-se em outra região do país. A Figura 3.8 mostra um exemplo de uma distribuição entre vítimas, beneficiários e o fraudador.

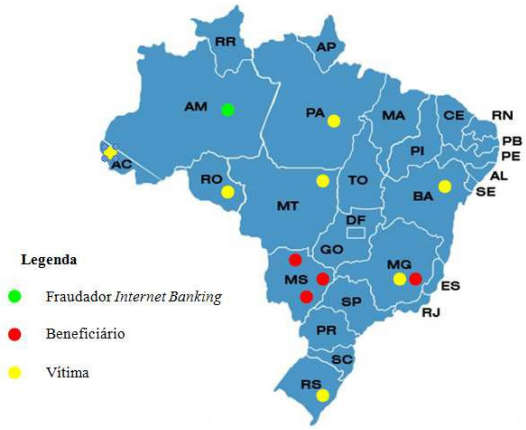


Figura 3.8: Distribuição geográfica de vítimas, beneficiários e do fraudador.
Fonte: Elaborada pelo autor.

Levando-se em conta a existência dos metadados relativos aos vínculos e vértices dos grafos pode-se aplicar, de forma objetiva, grupamentos ou filtros como: data das transações, valores em reais, tipos de atores (contas vítimas, contas beneficiárias, URAs) e localidades (UFs do Brasil), assim o investigador Policial Federal, visando o interesse da investigação, redireciona e prioriza as investigações.

3.3.4 INFORMAÇÃO PARA ROTEIRO DE INVESTIGAÇÃO

Após a etapa de filtragem e análise de vínculos e agrupamentos, o método Kraken disponibiliza as informações dos grafos (em formato de tabelas *Excel*) necessárias para o investigador elaborar os relatórios de análise das fraudes bancárias existentes na BNFBE.

Tais informações constituem um resumo de cada um dos subgrafos conexos gerados pelo Kraken com metadados das informações sumarizadas de seus atores e vínculos.

Para cada subgrafo são geradas informações, como: quantidade de contas vítimas, quantidade de contas beneficiárias, valores envolvidos nas transferências, datas das transações, telefones utilizados nas conexões com as URAs, localidades das contas (UFs do Brasil).

Essas informações servem de base para a elaboração de um roteiro contendo as transações passo a passo com a cronologia dos fatos (eventos) de como o dinheiro foi desviado da conta vítima até chegar ao destinatário final. Esse roteiro é gerado pelo Kraken com base nos possíveis caminhos entre os vértices do grafo.

Para exportar os metadados dos subgrafos gerados pelo Kraken com seus roteiros para o *software Analyst´s Notebook* da IBM, criou-se um modelo de especificação contendo a regra de negócio das transferências eletrônicas fraudulentas entre contas bancárias em um padrão aceito pelo *Analyst´s Notebook*.

3.4 INTERFACE GRÁFICA DO KRAKEN - FERRAMENTAL

Para a implantação da abordagem proposta nesse trabalho foi necessário a construção de um ferramental (uma aplicação/programa de computador) que nesse trabalho chamamos de IG. A IG deste ferramental foi implementada em *Delphi* versão 2010, com cerca de 10 mil linhas no programa fonte, escritas em ObjectPascal, para o sistema operacional *Windows 7 professional*.

A IG foi testada em uma máquina com processador i5-520M (*Dual Core*) de 2.4 GHz com memória *RAM* de 4 *GBytes* e disco de 500 *GBytes*. Foi utilizado o banco de dados *MySql* versão 5.7.12 para armazenar as tabelas da BNFBE, além daquelas que foram criadas especialmente para o funcionamento do Kraken. A Figura 3.9 exibe a tela da implementação da IG feita para o método Kraken.

3.4.1 GERANDO AUTOMATICAMENTE OS GRAFOS

O **1° passo** para o investigador Policial Federal gerar os grafos relativos as investigações contidas na BNFBE é fornecer o período de datas na IG (data inicial e data final) que ele deseja processar as análises de vínculos e atores contidos na BNFBE. Para preencher este período de datas existem as seguintes possibilidades:

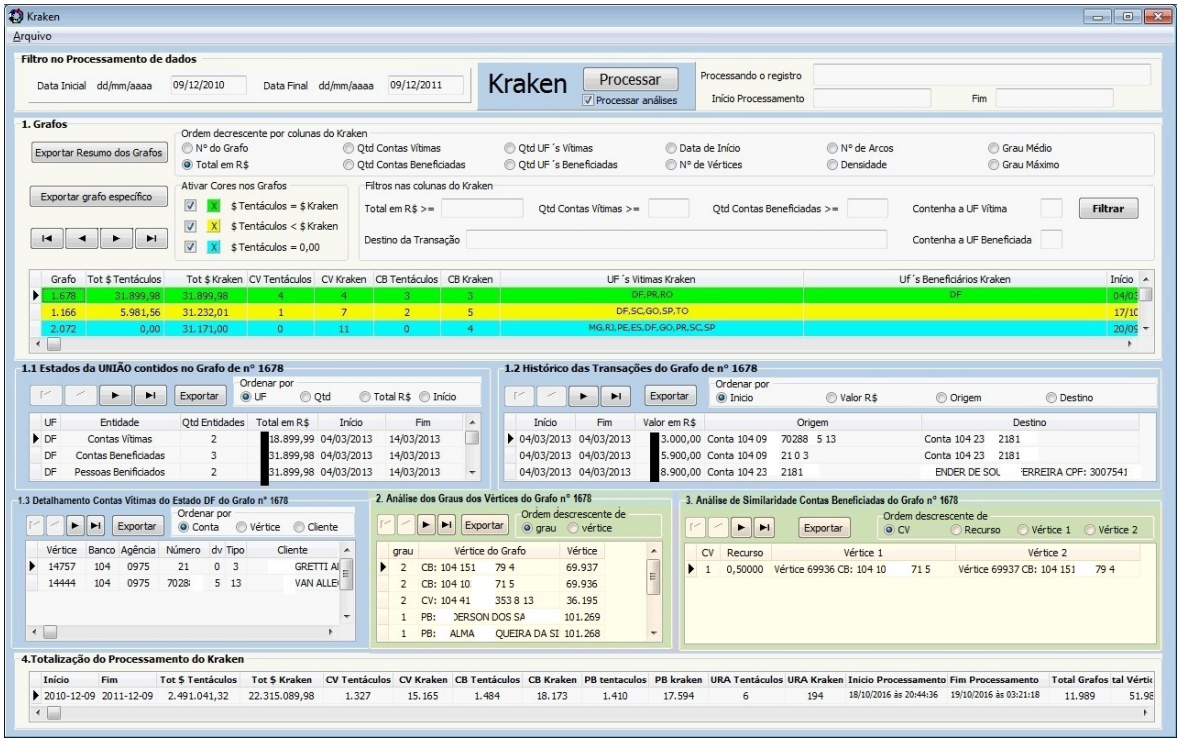


Figura 3.9: Tela da IG do Kraken.

Fonte: Elaborada pelo autor.

- Caso o investigador deseje que a abordagem do Kraken realize uma análise em toda a BNFBE ele deve informar a data inicial como sendo a data referente aos primeiros registros de transações bancárias fornecidos pela CAIXA (datados de 01/01/2008) até a data atual em que o investigador de encontra;
- Se o investigar achar desnecessário o processamento de dados da BNFBE anteriores, por exemplo, ao ano de 2010, a data inicial do processamento deve ser preenchida com 01/01/2011 e a data final com a data atual;
- A última opção é o investigador processar apenas registros de transações contidas na BNFBE de um determinado período contido entre o início da entrada dos dados na BNFBE e a data atual, por exemplo, os registros entre o ano de 2012 e 2014, nesse caso a data inicial da IG deve ser preenchida como 01/01/2012 e a data final com 31/12/2014.

O campo de preenchimento da data inicial e final de processamento está localizado no canto esquerdo superior da IG e é de preenchimento obrigatório.

O **2° passo** para o investigador dar prosseguimento na geração automática dos grafos é decidir se além de gerar os grafos a IG também irá processar as análises de grau,

similaridade e de recursos alocados. Para tomar essa decisão basta que o investigador marque ou desmarque a opção de nome "Processar análise" localizada ao centro na parte superior da IG. Caso o investigador selecione a opção "Processar análise", o tempo de processamento será aumentado em cerca de 20 a 25% no tempo total em minutos.

O **3° e último passo** na geração dos grafos pela IG do Kraken é o investigador acionar o botão de comando "Processar", localizado ao centro e na parte superior da IG, do lado do nome Kraken.

A partir desse momento a IG começa a realizar, de forma automática, as análises de vínculos e dos atores existentes na BNFBE, utilizando-se das regras de negócio pré-estabelecidas para identificar componentes conexos em um mesmo grafo e que sejam relativos à fraudes bancárias eletrônicas, foco desse trabalho. Somente são realizadas as análise de vínculos e de atores que tenham suas transações registradas na BNFBE no período de datas fornecido pelo investigador no passo de n° 2 referido anteriormente. Assim, basta o investigador Policial Federal esperar o final do tempo de processamento da IG, que irá depender do período de datas fornecido, para ter acesso as informações dos grafos gerados pela abordagem Kraken.

3.4.2 TABELA DOS GRAFOS GERADOS

Após o término do processamento, a IG mostra os grafos gerados em forma de tabela na área descrita como sendo Tabela "1. Grafos". São exibidas nessa tabela várias colunas contendo informações sumarizadas referentes a quantidade total de contas vítimas (**CV**), quantidade total de contas beneficiários (**CB**), valor total da fraude em reais (**$**), UFs vítimas, UFs beneficiadas, quantidade de pessoas beneficiadas (**PB**), quantidade de URAS, entre outras informações **relativas a cada um dos grafos gerados** pela IG. A coluna de nome "Grafo" representa, internamente para o Kraken, o número sequencial usado como uma chave de identificação única, que cada grafo recebe pela IG quando é processado. Dessa maneira, o investigador tem uma ideia de quantos grafos a IG conseguiu gerar dentro do período de datas fornecido, bastando para isso ordenar a tabela de grafos pela coluna "Grafo" e verificar qual o número sequencial atribuído pela IG ao último registro dessa tabela. Outra maneira, bem mais rápida, de se verificar a quantidade de grafos gerados pela IG é verificar a coluna de nome "Total Grafos" existente na tabela "4. Totalização do Processamento do Kraken", localizada na parte inferior da IG. Na Figura 3.9, por exemplo, pode-se verificar que foram processados 11.989 (onze mil novecentos e oitenta e nove) grafos pela IG no período de 09/12/2010

a 09/12/2011, que o processamento iniciou-sem em 18/10/2016 às 20:44:36 e terminou em 19/10/2016 às 03:21:18.

Uma vez que na BNFBE existem relações que mostram se determinada CV já foi analisada por um investigador (desde que essa informação esteja devidamente atualizada na BNFBE), a IG mostra também nas tabelas "1. Grafos" e "4. Totalização do Processamento do Kraken" colunas com a quantidade de CV, CB, PB, URA, e o total em reais da fraude que foram encontrados no MAFBE (no que diz respeito à análise humana praticada no Tentáculos) para cada um dos grafos processados pela IG. Assim o investigador tem uma base, de quais grafos processados pelo Kraken obtiveram as mesmas quantidades de CV, CB, PB, URA e total em reais da fraude, em relação ao MAFBE. Essa análise serve para a verificação da acurácia da abordagem Kraken em relação ao MAFBE.

A Tabela "1. Grafos" pode ser ordenada pelo investigador por: número do grafo, total em reais, quantidade de CV, quantidade de CB, quantidade UFs vítimas, quantidade UFs beneficiadas, data do início da transação fraudulenta no grafo, número de vértices do grafo, número de arcos do grafo, densidade do grafo, grau médio e grau máximo encontrado no grafo.

Com a IG, o investigador Policial Federal além de poder **reordenar** os dados da tabela "1. Grafos" pode aplicar **filtros** nessa tabela baseados em grafos que:

- Somarem um total de fraude em R$ maior ou igual a um determinado valor fornecido pelo investigador;
- Tiverem uma quantidade de CV maior ou igual a um determinada quantidade fornecida pelo investigador;
- Tiverem uma quantidade de CB maior ou igual a uma determinada quantidade fornecida pelo investigador;
- Contenham uma determinada UF, fornecida pelo investigador, nas UF´s vítimas do grafo;
- Contenham uma determinada UF, fornecida pelo investigador, nas UF´s beneficiadas do grafo;

- Contenham um determinado destino da transação no grafo, destino esse fornecido pelo investigador e que pode ser: CV, CB, e/ou CPFs ou nomes de pessoas beneficiárias permitindo filtrar os grafos em que esses atores aparecem.

Estes filtros foram implementados na IG, uma vez que os vínculos dos grafos do Kraken também armazenam o tipo de entidade (ator/vértice) relacionada aos seus vértices de origem e de destino (são armazenadas informações que qualificam o ator, como: o número da URA, número conta bancária, nome do titular da conta bancária, CPF, etc), além de pesos, como o valor da fraude e/ou a data em que ocorreu a transação. A Figura 3.10 destaca os filtros existentes na IG do Kraken.

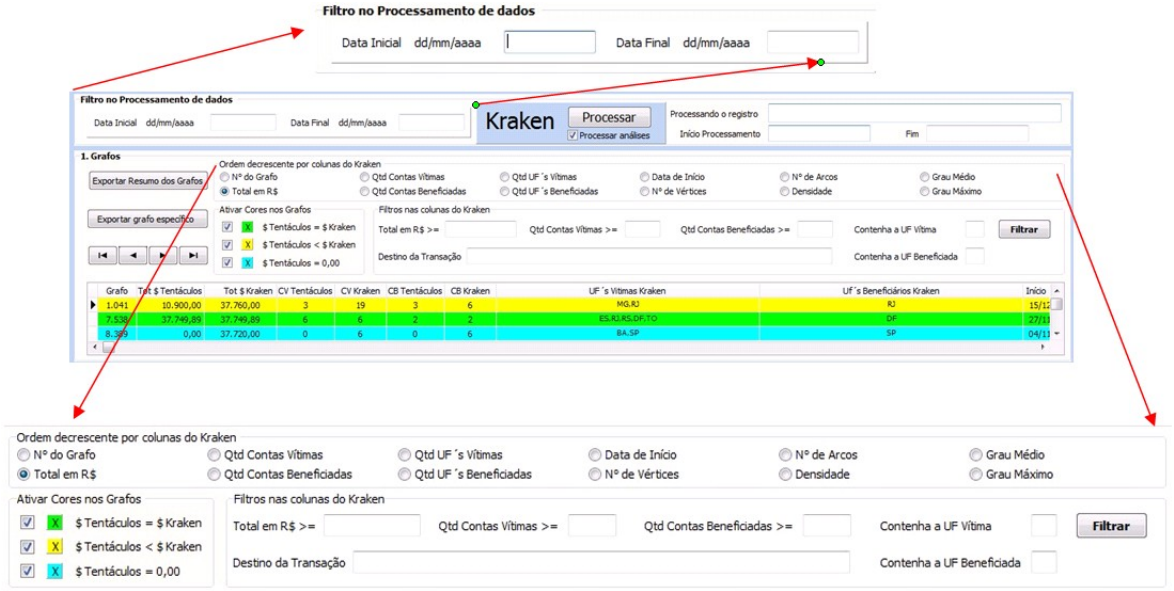


Figura 3.10: Detalhe dos Filtros da IG do Kraken.

Fonte: Elaborada pelo autor.

Por último, no que diz respeito a tabela "1. Grafos", ainda na parte superior da IG, o investigador pode realizar comparações entre os grafos gerados pelo Kraken versus o MAFBE que já foram investigados pelos Policiais Federais (Análise Humana - Tentáculos).Essas comparações são baseadas no valor total em reais (R$) fraudado em cada grafo da tabela "1. Grafos". Nesse caso, a cor verde sobre a linha de resumo do grafo nessa tabela mostra que o valor encontrado pelo Kraken é igual ao do Tentáculos, a cor amarela indica que o Kraken achou um valor de fraude maior que o Tentáculos. Já a cor azul claro indica que o Kraken gerou um grafo (investigação ou parte de uma investigação) que ainda não foi analisado pelos Policiais Federais, como o grafo de número 2.072 na Figura 3.9.

Utilizando-se da indicação das cores verde, amarela e azul (que representam respectivamente $ Tentáculos = $ Kraken, $ Tentáculos < $ Kraken e $ Tentáculos = 0,00), das opções de ordenação e de filtros (a IG permite que o investigador use mais de um filtro ao mesmo tempo) da tabela "1. Grafos", o investigador pode **priorizar suas ações de acordo com os critérios objetivos**, selecionando as maiores e melhores investigações baseado em: CV, CB, URA, PB, UFs vítimas, UFs beneficiadas, total em reais das fraudes, entre outros. Uma vez que a IG mostra na tabela "1. Grafos" os grafos (investigações ou partes de investigações) que já foram feitas (cor verde), o investigador pode evitar o retrabalho. No caso dos grafos da tabela "1. Grafos", que aparecem na cor azul claro (ainda não foram analisados pela PF) e que se enquadram nos critérios de filtros e ordenação, podem indicar ao investigador uma prioridade de investigação. Já o grafos que aparecem na cor amarela devem ser analisados pelo investigador a fim de se verificar o motivo da discrepância entre a análise automatizada do Kraken e o MAFBE.

A seguir menciona-se o conteúdo das demais tabelas exibidas na IG do Kraken e por fim é mostrado como a IG consegue exportar um determinado grafo selecionado na tabela "1. Grafos", contendo as especificações que serão utilizadas pelo *Analyst´s Notebook* do i2 para a visualização do grafo com seus atores (vértices) e vínculos (relações) propriamente ditos.

3.4.3 TABELA DAS CV, CB, PB, URA, E TOTAL DAS FRAUDES EM REAIS POR UF DO BRASIL DE UM GRAFO ESPECÍFICO

Na parte central a esquerda da IG, é exibido um detalhamento por estado do Brasil por meio da tabela de nome "1.1 Estados da UNIÃO contidos no Grafo de n°", para um dado grafo selecionado na tabela "1. Grafos".

Para que o investigador possa ver o resumo das CV, CB, PB, URAs, valor total da fraude em reais por UFs do Brasil de um grafo em específico (investigação), basta o investigador selecionar o número do grafo que deseja ver estas informações na tabela "1. Grafos", como mostrado na Figura 3.10, onde por exemplo está selecionado o grafo de número 1.041 (cor amarela). Após a seleção deste grafo, a IG automaticamente atualiza a tabela "1.1 Estados da UNIÃO contidos no Grafo de n°" com o resumo das CV, CB, URAs, total da fraude em reais por estado do Brasil. A Figura 3.11 ilustra um exemplo desta tabela levando em consideração o grafo de n° 1041, que foi previamente selecionado pelo investigador na tabela "1. Grafos".

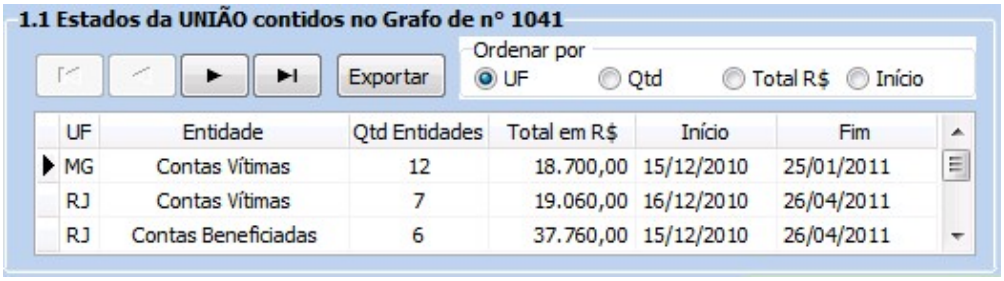


Figura 3.11: CV, CB, PB, URA e total da fraude em reais por UF do Brasil de um grafo em específico.

Fonte: Elaborada pelo autor.

Dessa forma, é possível a visualização da quantidade de CV, CB e de PB por estado juntamente com os respectivos montantes fraudados (valores em reais) e períodos de tempo de ocorrência. A tabela "1.1 Estados da UNIÃO contidos no Grafo de n°", pode ser exportada em formato de planilha *Excel* e ordenada por: sigla da UF, quantidade de CV, CB, PB ou URAs, Total em reais e pela data do início da transação fraudulenta naquele estado do Brasil para o grafo em questão.

3.4.4 TABELA DO DETALHAMENTO DAS CV, CB, PB, URA DE UMA UF DO BRASIL DE UM GRAFO ESPECÍFICO

Imediatamente abaixo da tabela de nome "1.1 Estados da UNIÃO contidos no Grafo de n°" encontra-se a tabela de nome "1.3 Detalhamento CV, CB, PB, URA do estado do Grafo n°", que também é atualizada automaticamente pela IG quando o investigador seleciona uma das UF´s contidas na tabela "1.1 Estados da UNIÃO contidos no Grafo de n°".

Para cada estado do Brasil (UF), selecionado pelo investigador na Figura 3.11 o Kraken mostra o detalhamento dos vértices (atores que podem ser: CV, CB, URA e PB) contidos naquele referido estado do brasil. A Figura 3.12 exemplifica o detalhamento das 12 (doze) CV do estado de Minas Gerais (MG) (selecionado na tabela "1.1 Estados da UNIÃO contidos no Grafo de n°") relativas ao grafo 1.041.

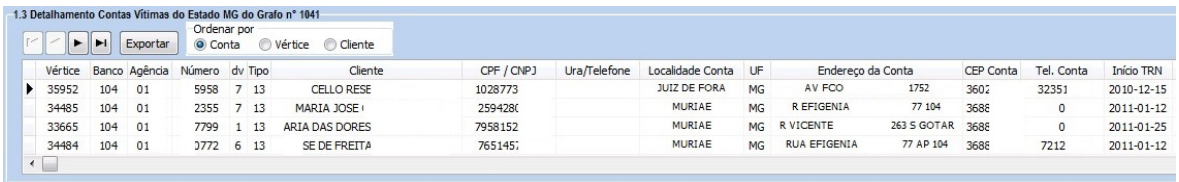


Figura 3.12: Exemplo do detalhamento das CV de MG relativas ao grafo 1.041.

Fonte: Elaborada pelo autor.

A tabela “1.3 Detalhamento CV, CB, PB, URA do estado do Grafo n°”pode ser exportada em formato de planilha *Excel* e ordenada pelo: n° da conta bancária, n° do vértice desta conta no grafo, e pelo nome do cliente.

3.4.5 TABELA DO HISTÓRICO DAS TRANSAÇÕES DE UM GRAFO ESPECÍFICO

Por fim, na parte central a direita da IG, na tabela de nome “1.2 Histórico das Transações do Grafo de n°” encontra-se uma descrição detalhada do histórico das transações fraudulentas do grafo selecionado na tabela “1. Grafos”. O conteúdo desta tabela é automaticamente atualizado quando o investigador seleciona um grafo em específico na tabela “1. Grafos”.

As informações desse roteiro são exportadas em formato *Excel*, e são aproveitadas como uma especificação de importação de dados da ferramentas *Analyst's Notebook* da IBM, para a geração e visualização gráfica do grafo elaborado pelo Kraken (diagrama de elos da investigação policial em curso). A Figura 3.13 exibe um exemplo do fluxo do desvio do dinheiro no grafo 1.041.

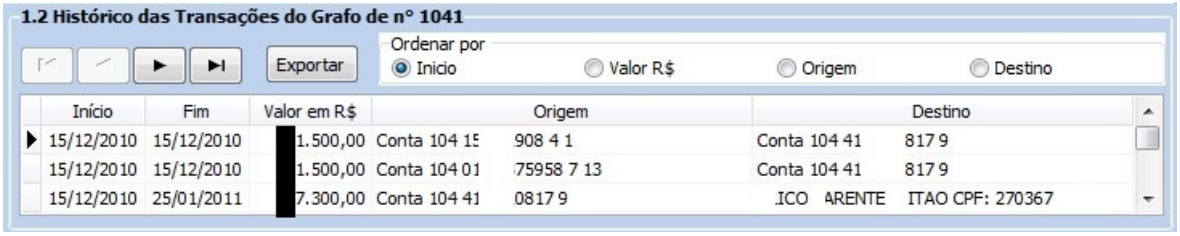


Figura 3.13: Fluxo do desvio do dinheiro da sua origem ao seu destino no grafo 1.041.
Fonte: Elaborada pelo autor.

A tabela “1.2 Histórico das Transações do Grafo de n°” ainda pode ordenada por data, valor fraudado, origem e destino da operação. A ordem cronológica permite ao investigador descrever os fatos e transações fraudulentas na medida em que ocorreram fornecendo uma descrição da atuação como um todo do grupo criminoso.

3.4.6 ANÁLISES REALIZADAS PELA IG

Durante a construção dos grafos, caso o investigador tenha selecionado a opção de “Processar análises”, a IG realizará algumas análises sobre os vértices e vínculos dos grafos, com o objetivo de apontar quais são os atores mais relevantes naquela investigação. São realizadas as seguintes análises nos vértices:

- Grau;
- Similaridade;
- Recursos Utilizados.

De posse da densidade de cada um dos subgrafos, o Kraken calcula a densidade média de todas as investigações por ele processadas, tendo assim uma dimensão da complexidade da BNFBE no período analisado.

Na parte inferior da IG, logo abaixo da Tabela "1.2 Histórico das Transações do Grafo de n°", estão localizadas as tabelas de nomes: "2. Análise dos Graus dos Vértices do Grafo n°" e "3. Análise de Similaridade Contas Beneficiadas do Grafo n°".

O conteúdo destas tabelas é automaticamente atualizado quando o investigador seleciona um grafo específico na tabela "1. Grafos". A Figura 3.14 exibe um exemplo da análise de graus feita para cada um dos vértices do grafo de n° 1.678 selecionado previamente na tabela "1. Grafos".

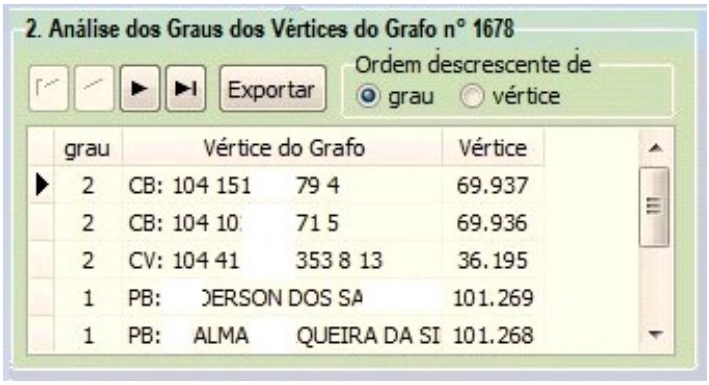


Figura 3.14: Análise de graus para cada um dos vértices do grafo n° 1.678.
Fonte: Elaborada pelo autor.

A tabela "2. Análise dos Graus dos Vértices do Grafo n°"pode ser exportada em formato de planilha *Excel* e ordenada pelo grau ou número do vértice analisado. A Figura 3.15 exibe um exemplo das análises de similaridade e recursos alocados feitas para as CB do grafo de n° 1.678, previamente selecionado na tabela "1. Grafos". A tabela "3. Análise de Similaridade Contas Beneficiadas do Grafo n°"pode ser exportada em formato de planilha *Excel* e ordenada de acordo com:

- Quantidade de CV que as CB têm em comum;

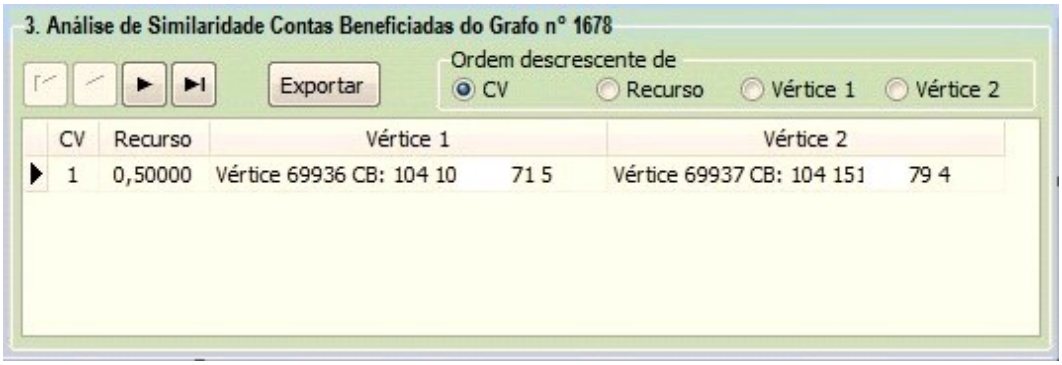


Figura 3.15: Análise de similaridade e recursos alocados do grafo n° 1.678.

Fonte: Elaborada pelo autor.

- Recurso alocados.
- Vértice n°1 (referente a conta beneficiada n°1);
- Vértice n°2 (referente a conta beneficiada n°2).

3.4.7 EXPORTANDO GRAFOS COM OS ROTEIROS CUSTOMIZADOS PARA VISUALIZAÇÃO NO *ANALYST´S NOTEBOOK* DA IBM

Para a exportação do grafo do Kraken para o *Analyst´s Notebook*, primeiro o investigador policial deve selecionar o grafo desejado na tabela "1. Grafos" e em seguida acionar o botão de comando "Exportar grafo específico" localizado na IG do Kraken conforme é mostrado na Figura 3.16, na qual vemos que o grafo de n° 408, com um valor total de R$ 253.487,00, foi selecionado pelo investigador Policial Federal para ser exportado.

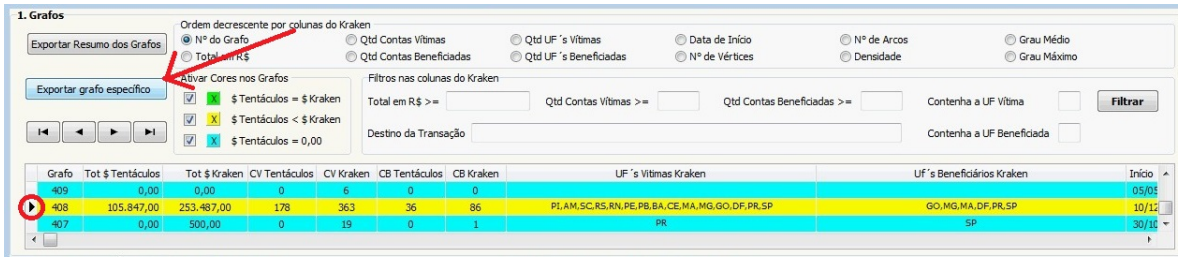


Figura 3.16: Iteração n° 1 para Exportação do grafo do Kraken para o *Analyst´s Notebook*.

Fonte: Elaborada pelo autor.

O Kraken irá solicitar ao investigador a confirmação da exportação do grafo 408 conforme é mostrado na iteração de n° 2 na Figura 3.17.

Figura 3.17: Iteração n° 2 - Confirmação da exportação do grafo 408 do Kraken.
Fonte: Elaborada pelo autor.

Na iteração de n° 3, o Policial Federal irá salvar o grafo em uma pasta que desejar do seu computador com um nome qualquer, por exemplo, “Grafo_408_Kraken”, conforme ilustra a Figura 3.18.

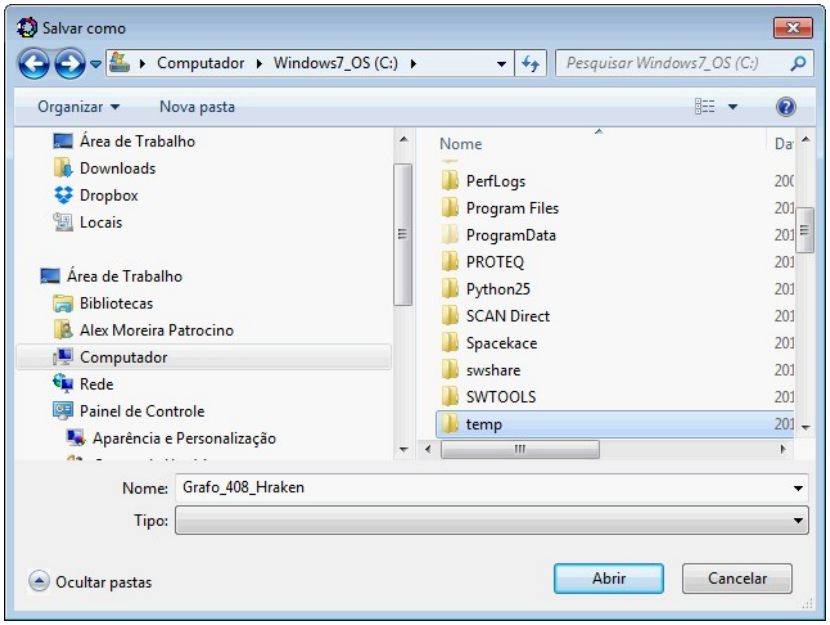


Figura 3.18: Iteração n° 3 - Gravando as especificações do grafo 408 no formato Excel.
Fonte: Elaborada pelo autor.

Em seguida, o Kraken abre a planilha em *Excel* com as especificações geradas desse grafo para a apreciação do investigador. É essa planilha em Excel que será importada pelo *Analyst´s Notebook* para montar o grafo desta investigação. A Figura 3.19 mostra a planilha em Excel das especificações do grafo 408.

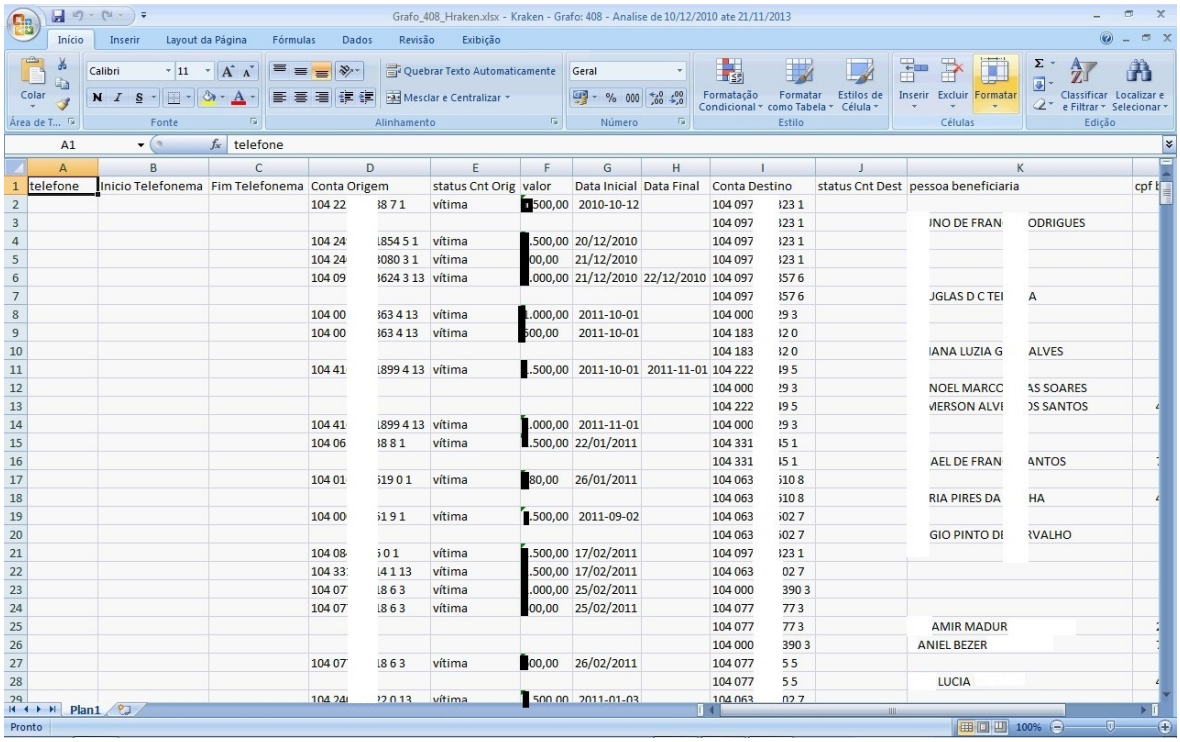


Figura 3.19: Iteração n° 4 - Planilha em *Excel* com as especificações do grafo 408.
Fonte: Elaborada pelo autor.

3.5 ESPECIFICAÇÃO DE IMPORTAÇÃO DO GRAFO NO *ANALYST´S NOTEBOOK* DA IBM

Para a exportação desse grafo do Kraken para o *Analyst´s Notebook* da IBM foi necessário criar um modelo de especificação com a descrição da regra de negócio da transferência bancária no *Analyst´s Notebook*. Esse modelo de especificação será detalhado na Seção 3.5

A Figura 3.20 mostra a iteração de n° 5, na qual o investigador, já com o *Analyt´s Notebook* aberto, seleciona o menu de "Especificações de importação salvas ..." para iniciar a importação do roteiro do grafo 408 gerado pelo Kraken.

Na próxima iteração, a de n° 6, basta o investigador selecionar o modelo já salvo de especificações do Kraken e em seguida acionar o comando "Editar" para associar esse modelo à planilha "Grafo_408_Kraken". A Figura 3.21 mostra a iteração de n° 6. A Figura 3.22 mostra a iteração de n° 7.

Após a seleção da planilha com o roteiro das especificações do grafo gerado pelo Kraken, o investigado deve acionar o comando "Importar", conforme é mostrado na Figura 3.23

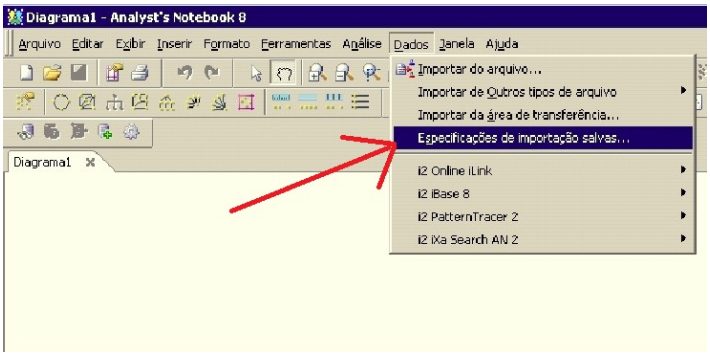


Figura 3.20: Iteração n° 5 - Importação de especificações do *Analyst´s Notebook* 8 da IBM.

Fonte: Elaborada pelo autor.

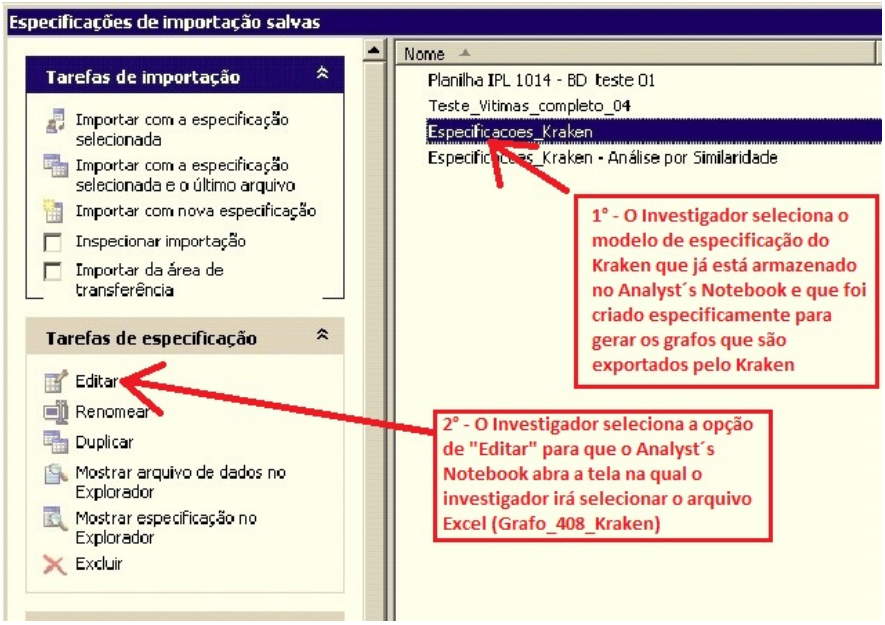
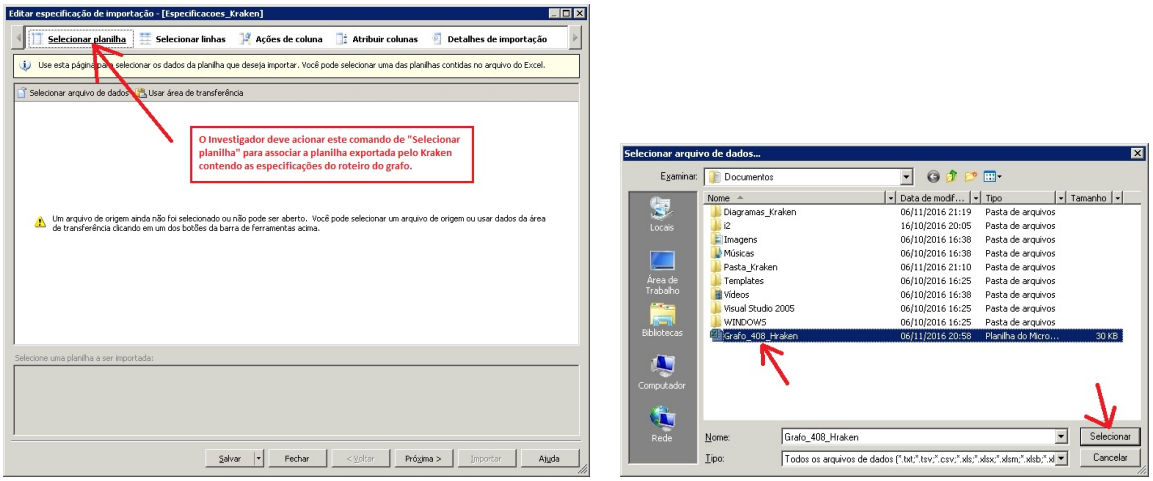


Figura 3.21: Iteração n° 6 - Seleção do modelo de importação de especificações do Kraken.

Fonte: Elaborada pelo autor.



(a) Tela de seleção da planilha com as especificações. (b) Seleção da planilha Grafo_408_Kraken.

Figura 3.22: Iteração n° 7 - Seleção da planilha com especificações do grafo do Kraken.

Fonte: Elaborada pelo autor.

(Iteração 8). A Figura 3.24 mostra a iteração de n° 9, na qual o *Analyst´s Notebook* inicia a importação do roteiro de especificações da planilha "Grafo_408_Kraken"

A Figura 3.25 mostra a importação concluída do grafo 408 gerado pelo Kraken. A Figura 3.26 mostra o grafo 408 na íntegra. A Figura 3.27 mostra uma aproximação (*zoom*) da imagem do grafo 408 gerado pela abordagem do Kraken e desenhado em poucos segundos, de forma semi-automática, pelo *Analyst´s Notebook* 8 da IBM. A Figura 3.28 mostra o passo 1 na criação desse modelo de especificação de importação dos grafos do Kraken para o *Analyst´s Notebook*.

O primeiro passo para a criação do modelo de especificação de importação é acessar o menu de dados no *Analyst´s Notebook* e nele selecionar a opção "Importar do arquivo ...". Nesse momento, o *Analyst´s Notebook* solicita um arquivo no formato Excel para servir como base da regra de negócio da especificação. Em seguida, é selecionado um arquivo de nome "Grafo_408_Kraken", que foi gerado na Seção 3.3.4, conforme mostra a Figura 3.29

O *Analyst´s Notebook* exibe a seguinte tela mostrada na Figura 3.30 com os dados do arquivo em *Excel* gerado pelo Kraken com o roteiro do fluxo da fraude bancária eletrônica do grafo 408.

O quarto passo para a geração da especificação é eliminar a primeira linha da planilha

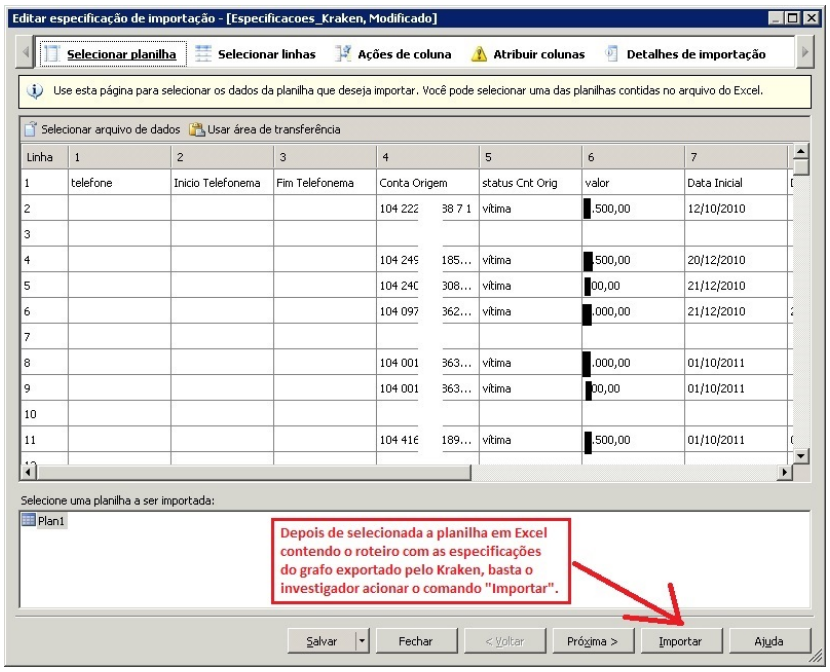


Figura 3.23: Iteração n° 8 - Importação da planilha com o roteiro do grafo gerada no Kraken.

Fonte: Elaborada pelo autor.

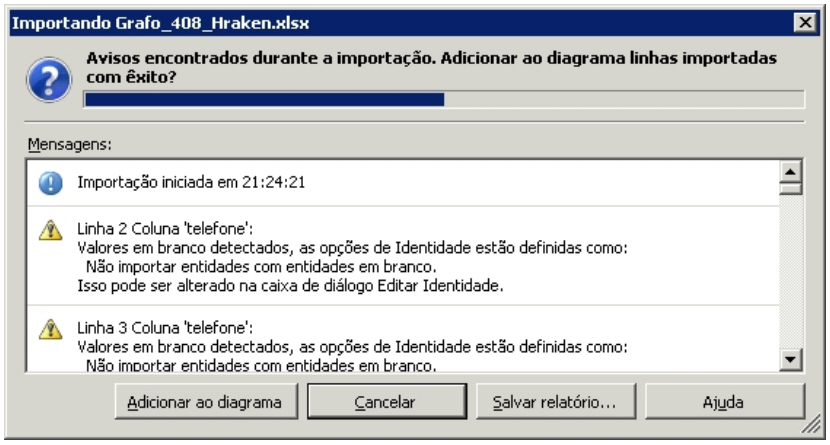


Figura 3.24: Iteração n° 9 - Importação do roteiro com as especificações contidas na planilha “Grafo_408_Kraken”. Fonte: Elaborada pelo autor.

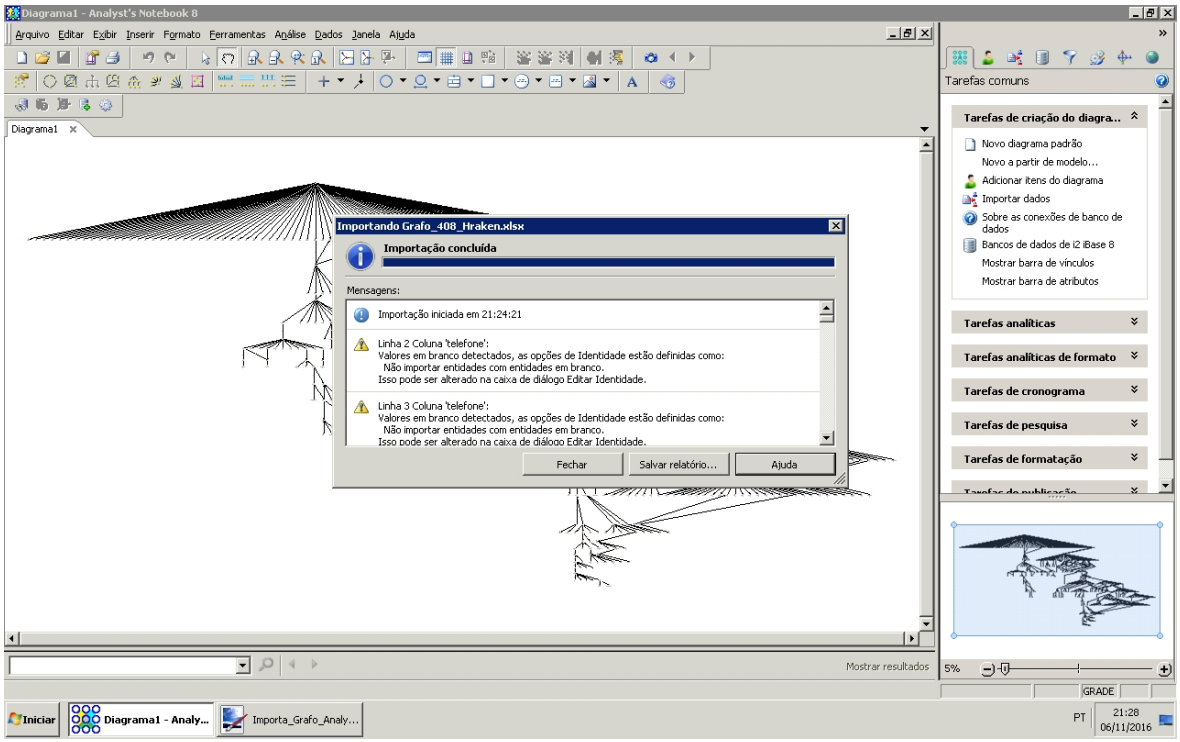


Figura 3.25: Fim da importação do grafo 408 pelo *Analyst´s Notebook*.

Fonte: Elaborada pelo autor.

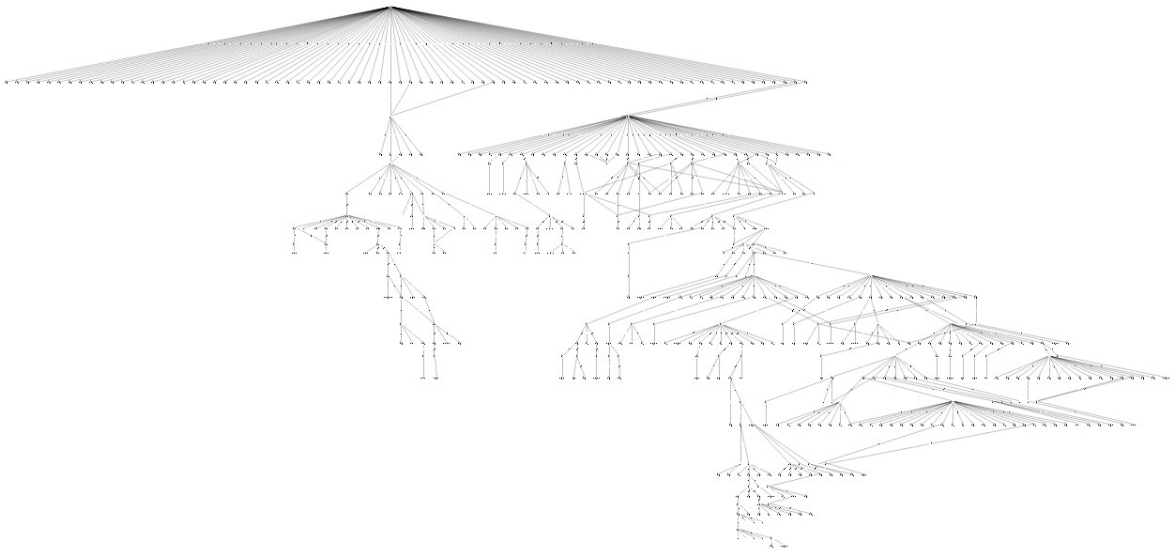


Figura 3.26: Grafo 408 gerado no *Analyst´s Notebook*.

Fonte: Elaborada pelo autor.

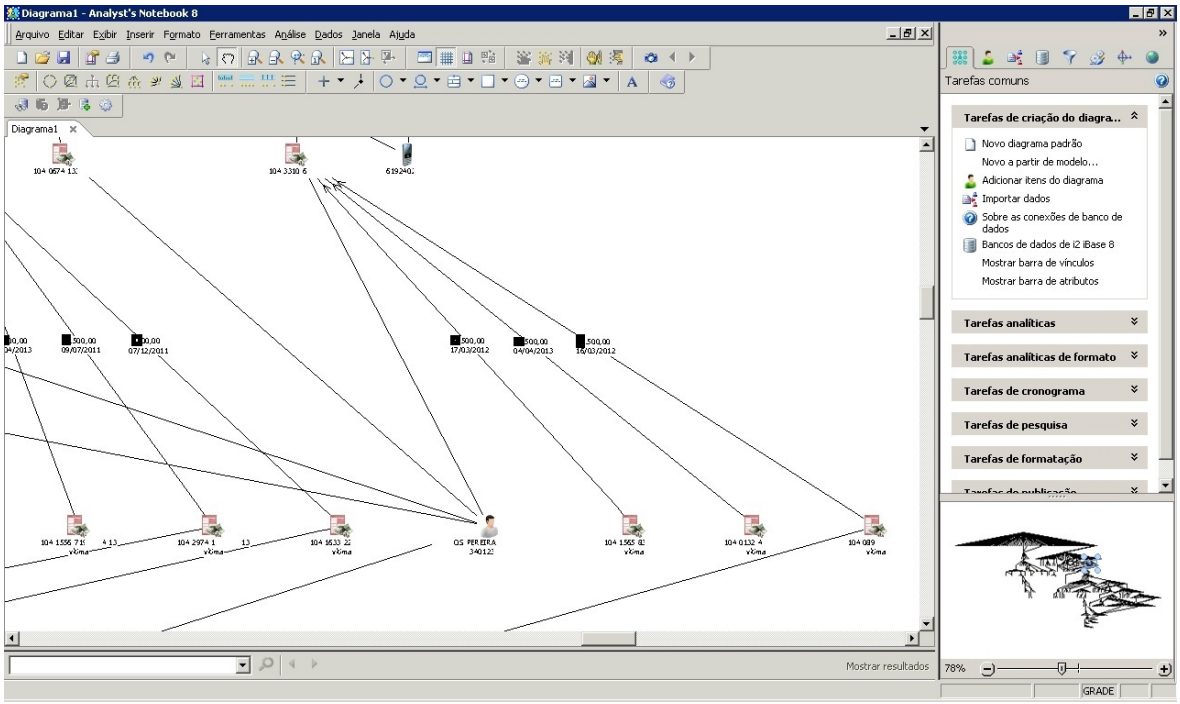


Figura 3.27: Exemplo de uma aproximação de alguns dos atores (vértices) do grafo 408 realizada pelo *Analyst´s Notebook*. Fonte: Elaborada pelo autor.

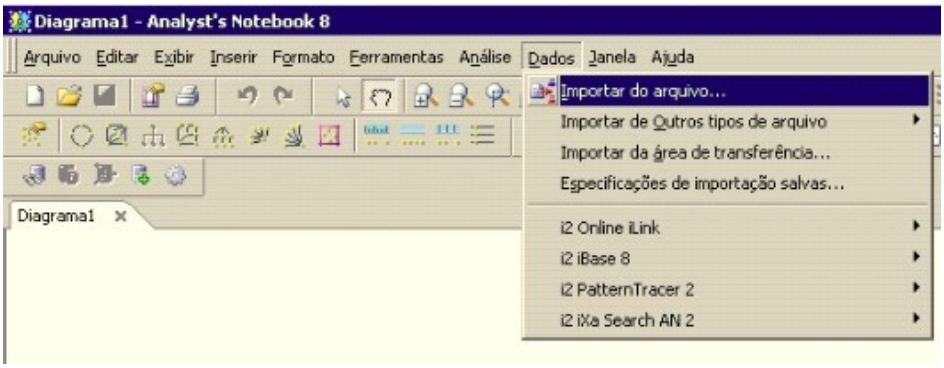


Figura 3.28: Passo 1 da criação da especificação - Selecionar menu de dados e escolher a opção importar arquivo. Fonte: Elaborada pelo autor.

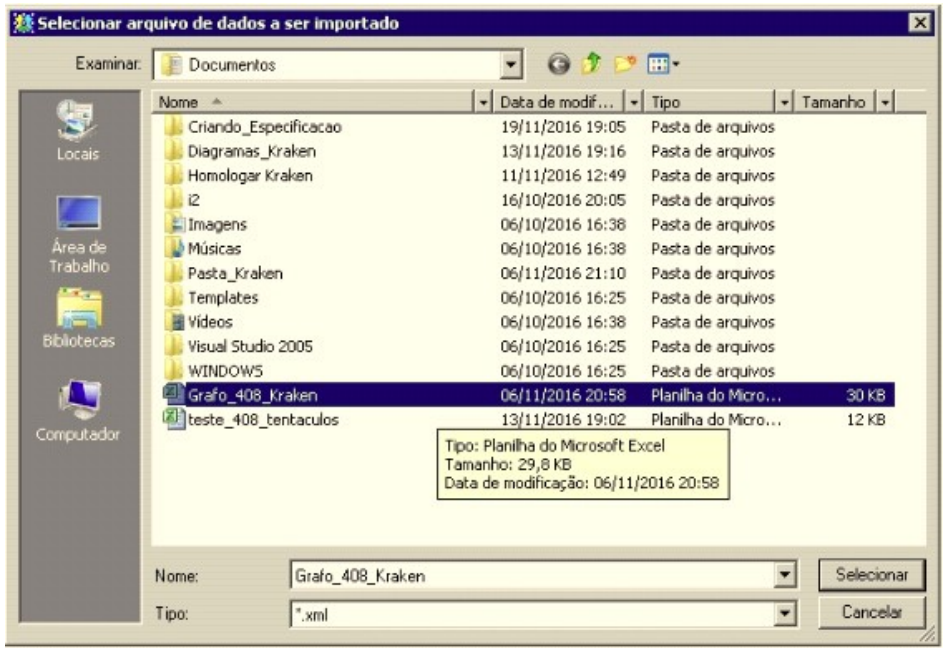


Figura 3.29: Passo 2 da criação da especificação - escolher o arquivo Excel gerado na etapa roteiro do Kraken. Fonte: Elaborada pelo autor.

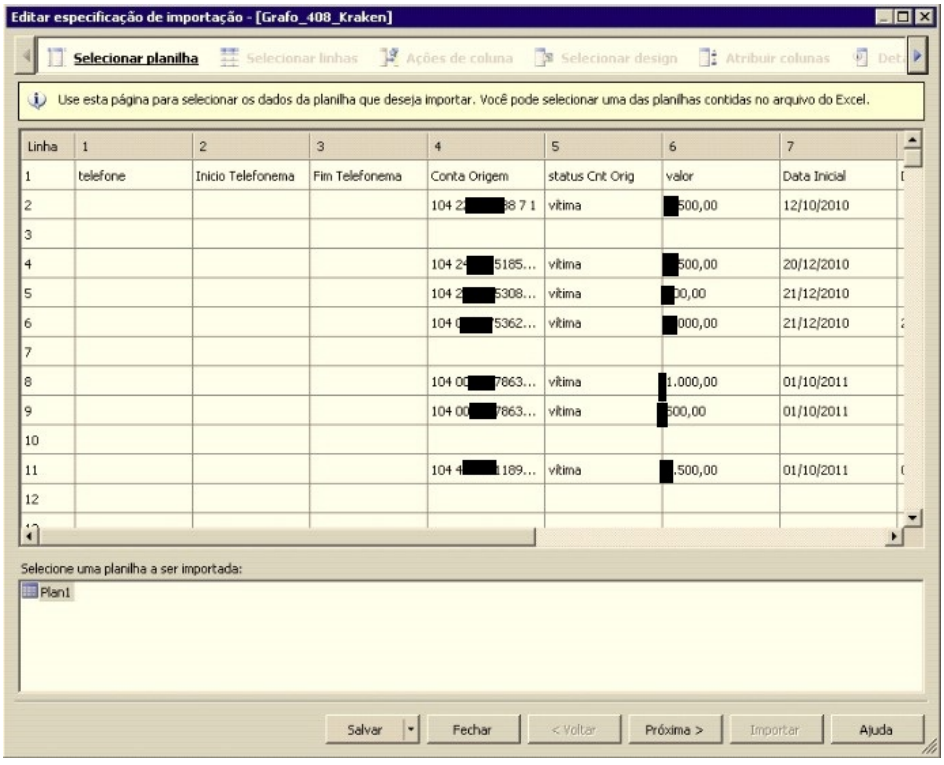


Figura 3.30: *Analyst´s Notebook* abre a planilha do *Excel* gerada no Kraken. Fonte: Elaborada pelo autor.

Excel que contém o cabeçalho das colunas do roteiro da fraude de transferências entre contas bancárias, conforme mostra a Figura 3.31

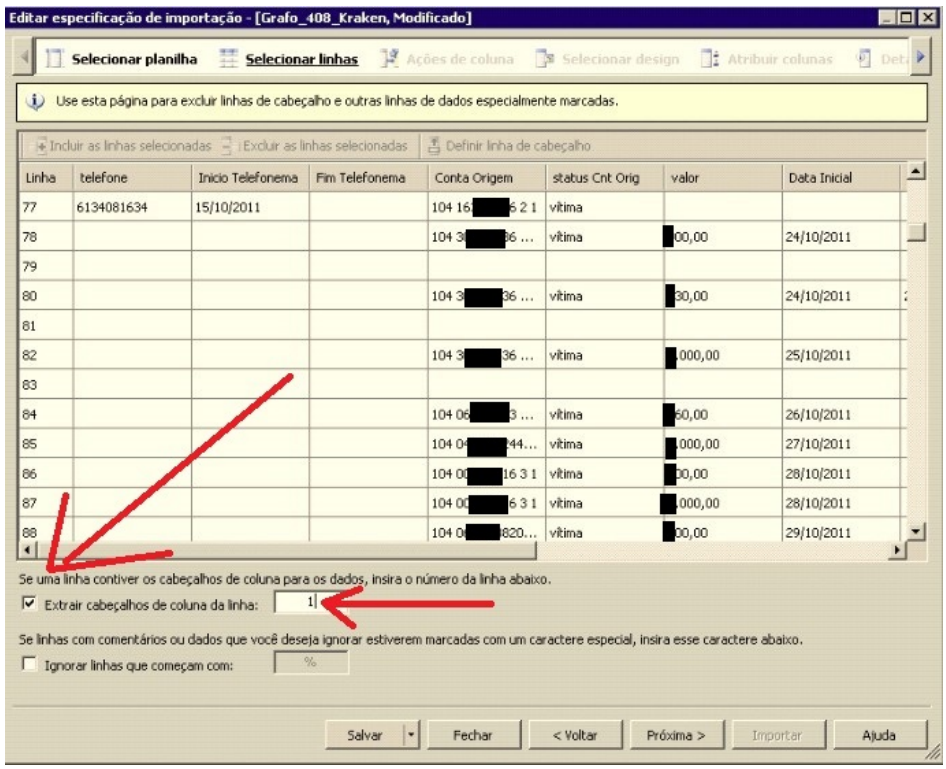


Figura 3.31: *Analyst´s Notebook* abre a planilha do *Excel* gerada no Kraken.
Fonte: Elaborada pelo autor.

O quinto passo para a geração da especificação seria especificar ações que o *Analyst´s Notebook* poderia realizar em cada uma das colunas da planilha em *Excel* gerada pelo Kraken. Ações como: adicionar sufixo, extrair parte do texto, entre outras. Como nenhuma dessas ações são necessárias para a regra de negócio de transferências bancárias, o procedimento nesse passo foi simplesmente acionar o botão próximo, como mostrado na Figura 3.32

O sexto passo para a geração da especificação de importação da planilha do Kraken no *Analyst´s Notebook* é selecionar o tipo de *design* que a regra de negócio das transferências bancárias fraudulentas deve se enquadrar. Nesse momento foi escolhida a opção de *design* de "Diagrama de associação mais complexo", conforme mostra a Figura 3.33

O passo n° 7 é determinar quais colunas da planilha em *Excel* devem ser atribuídas para cada entidade do "Diagrama de associação mais complexo" escolhido como *design* para modelar a regra de negócio de transferências bancárias fraudulentas. A Figura 3.34,

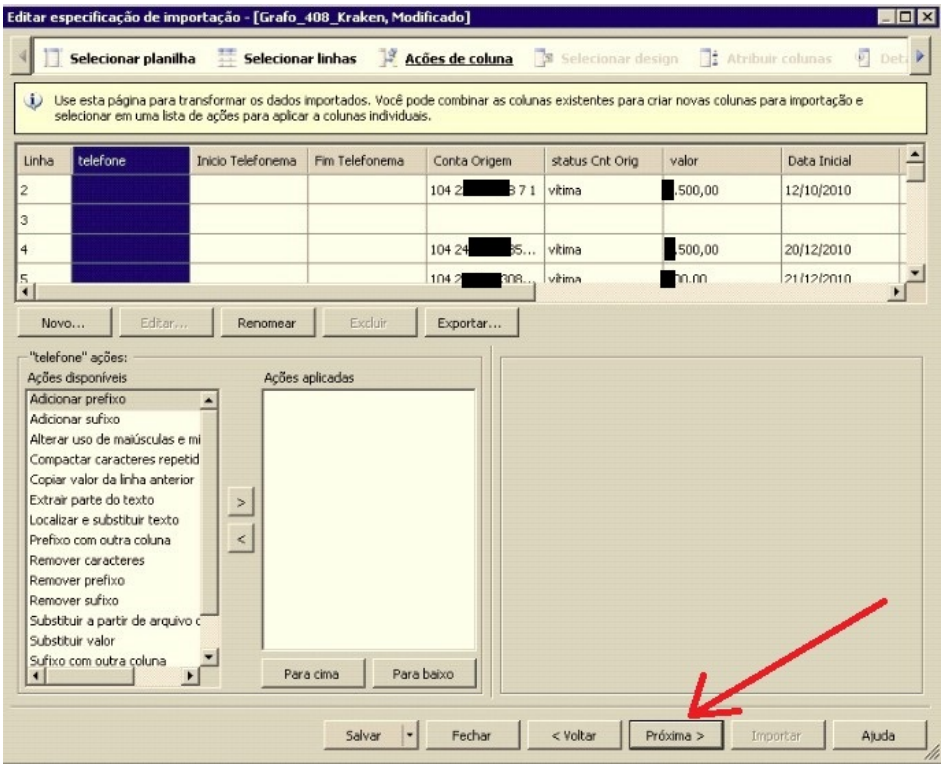


Figura 3.32: *Analyst´s Notebook* permite criar ações para cada coluna da planilha do *Excel* gerada no Kraken. Fonte: Elaborada pelo autor.

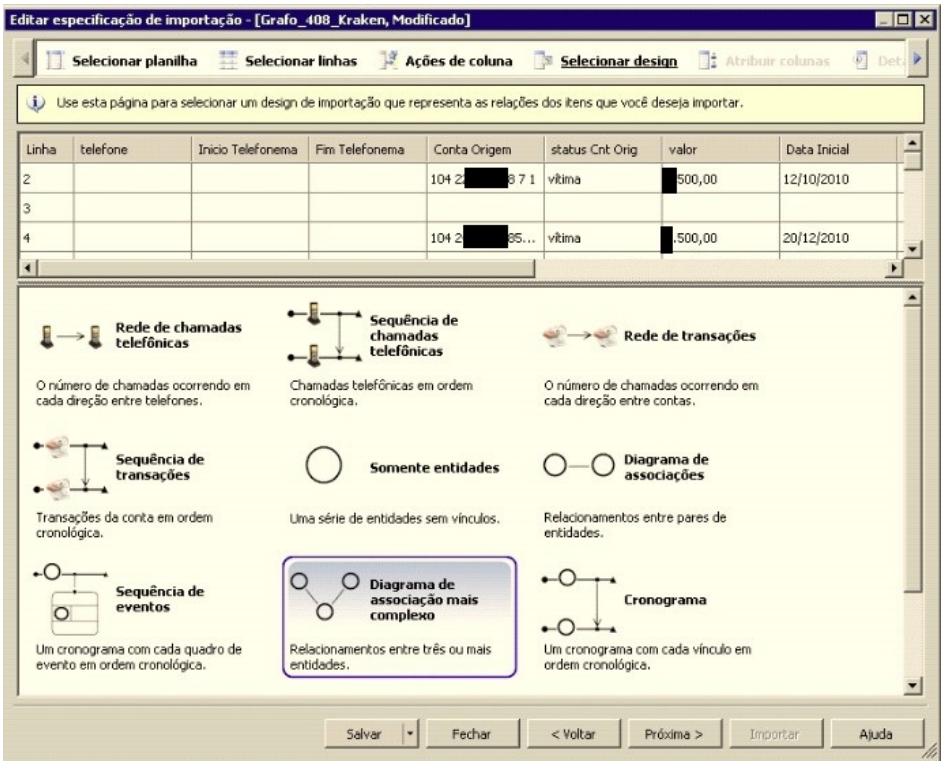


Figura 3.33: Escolha do tipo de *design* da regra de negócio das transferências bancárias fraudulentas. Fonte: Elaborada pelo autor.

mostra esta etapa da geração da especificação de importação dos grafos do Kraken.

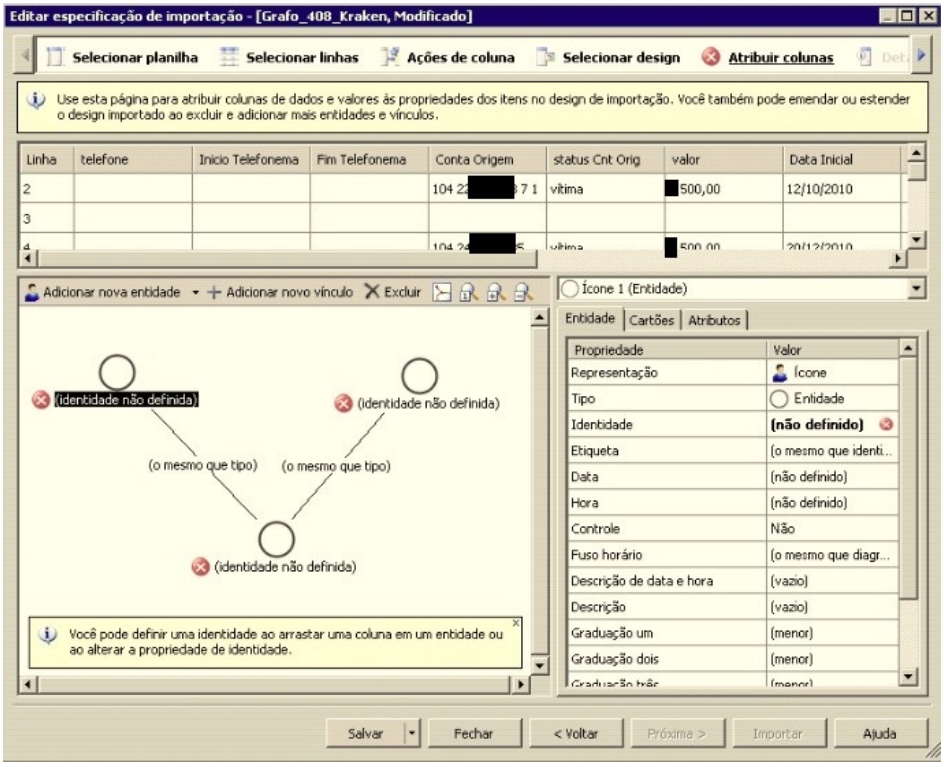


Figura 3.34: Atribuição das colunas da planilha *Excel* ao diagrama de associação mais complexo do *Analyst´s Notebook*. Fonte: Elaborada pelo autor.

A Figura 3.35 mostra como atribuir para cada entidade do diagrama de associação mais complexo um dos atores envolvidos na fraude de transferências bancárias (URA, CV, CB, PB). No exemplo da Figura 3.35 é mostrado como foi atribuído à 1a. entidade do diagrama de associação mais complexo do *Analyst´s Notebook*, o ator URA da planilha *Excel* gerada pelo Kraken. Note-se que a coluna da planilha *Excel* denominada "telefone" representa as URAs da regra de negócio das transferências bancárias do Kraken.

Para os demais atores (vértices) envolvidos nas fraudes bancárias eletrônicas (CV, CB, PB), o processo acima foi repetido na respectiva entidade do diagrama de associação mais complexo do *Analyst´s Notebook* da IBM.

Além de atribuir os atores das transferências bancárias fraudulentas geradas pelo Kraken aos respectivos vértices (entidades) do diagrama de associação mais complexo do *Analyst´s Notebook*, é necessário atribuir às arestas deste diagrama os respectivos vínculos que relacionam os atores desse tipo de fraude.

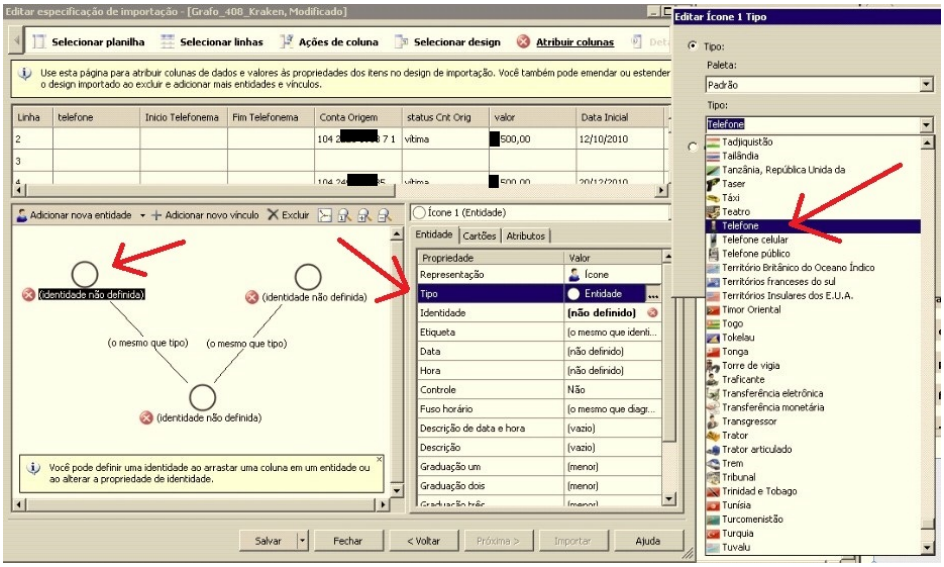


Figura 3.35: Atribuição da coluna telefone da planilha *Excel* ao diagrama de associação mais complexo do *Analyst´s Notebook*. Fonte: Elaborada pelo autor.

Assim, para uma aresta entre uma conta vítima e uma conta beneficiária foi atribuído o valor em reais (R$) da transferência com o período da data em que ela aconteceu.

Toda essa informação veio de maneira automática da planilha *Excel* gerada pela exportação do grafo 408 pela IG do Kraken.

Ainda foi atribuído nas arestas o sentido do deslocamento do dinheiro, por meio de setas.

O resultado final dessas atribuições das colunas da planilha *Excel* gerada pelo Kraken para o diagrama de associação mais complexo do *Analyst´s Notebook* da IBM está representado na Figura 3.36

O último passo, o passo de n° 8, é indicar ao *Analyst´s Notebook* com qual formato ele deve representar o grafo importado do Kraken entre os tipos possíveis: leque, organização criminosa, leque compacto, entre outros.

Foi escolhido o formato de organização criminosa para exibir o grafo do Kraken no *Analyst´s Notebook*, por atender de forma mais precisa a regra de negócio das fraudes de transferências bancárias. Esta última etapa na geração do modelo de especificação de importação de grafos do Kraken no *Analyst´s Notebook* é mostrada na Figura 3.37

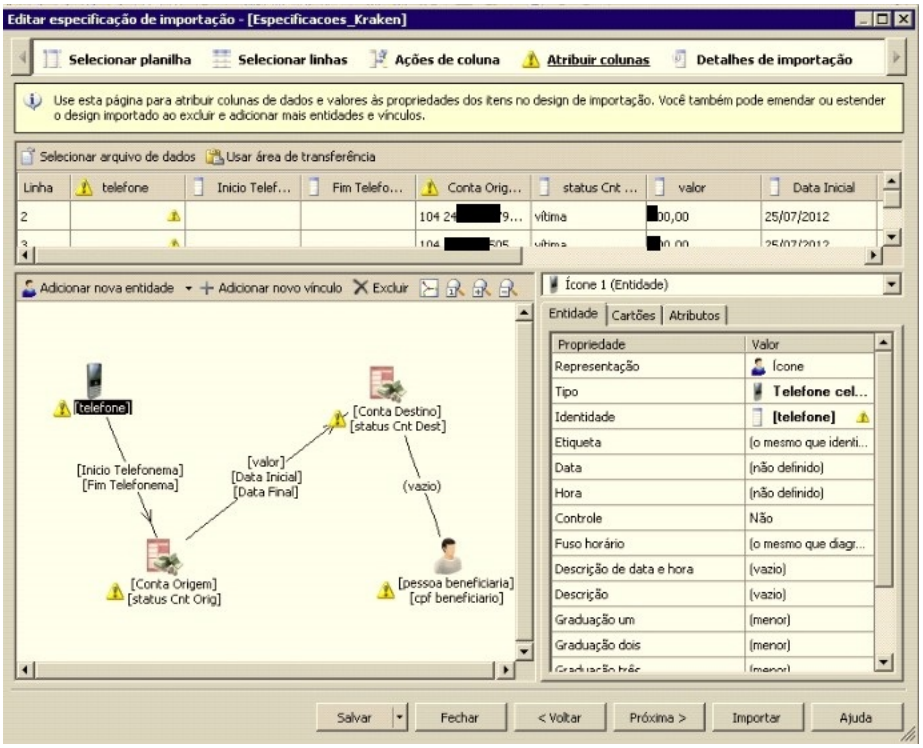


Figura 3.36: Atribuições das colunas da planilha *Excel* no diagrama de associação mais complexo do *Analyst´s Notebook*. Fonte: Elaborada pelo autor.

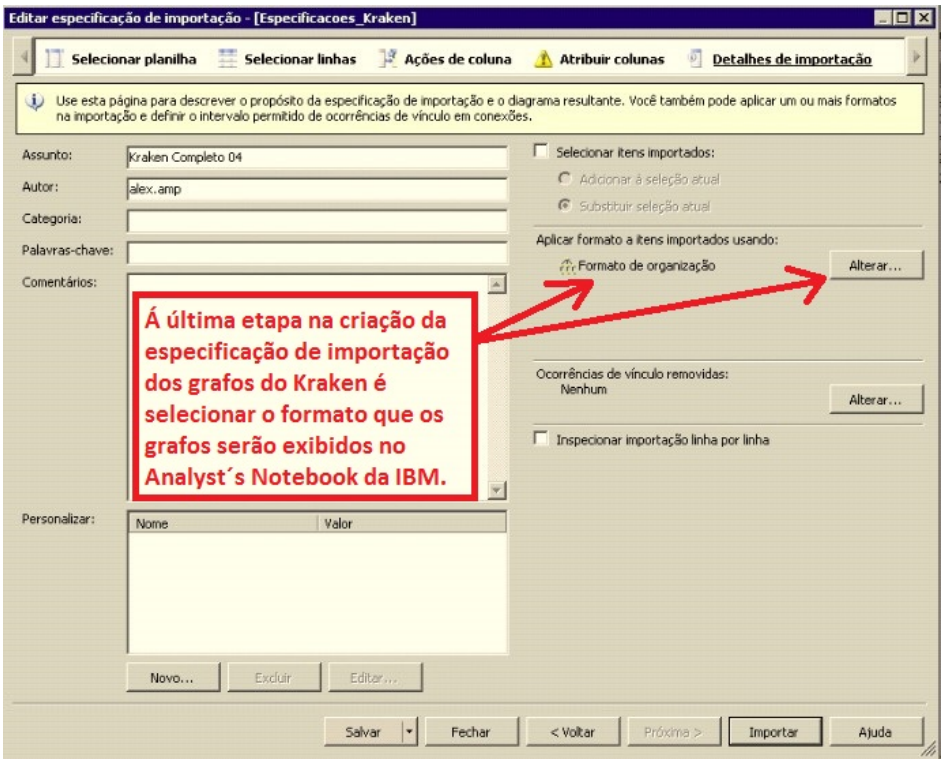


Figura 3.37: Escolha do formato de organização criminosa para mostrar os grafos do Kraken no *Analyst´s Notebook* da IBM. Fonte: Elaborada pelo autor.

Note-se que neste exemplo, depois da geração do grafo do Kraken de n° 408, o investigador poderá a qualquer momento reorganizar o formato do grafo entre os tipos existentes (leque, leque compacto, organização criminosa, entre outros), além de poder à mão livre realocar os vértices (atores) e as arestas (vínculos) do grafo.

Para salvar este modelo de especificação de importação, basta salvá-lo, por exemplo, com o nome de "Especificacoes_Kraken", acionando o comando "Salvar" da Figura 3.37.

É esta especificação de importação que será utilizada para automatizar as próximas importações dos grafos gerados pelo Kraken no *Analyst´s Notebook* da IBM, sem a necessidade do investigador Policial Federal ter que refazer todos os oitos passos mencionados acima como foi demostrado na Seção 3.3.4. Voltando a Figura 3.37, basta o investigador acionar o comando "Importar"para que o grafo 408 do Kraken seja mostrado no *Analyst´s Notebook*. A Figura 3.38 ilustra o exemplo do grafo 408 produzido pelo Kraken e desenhado no *Analyst´s Notebook* da IBM.

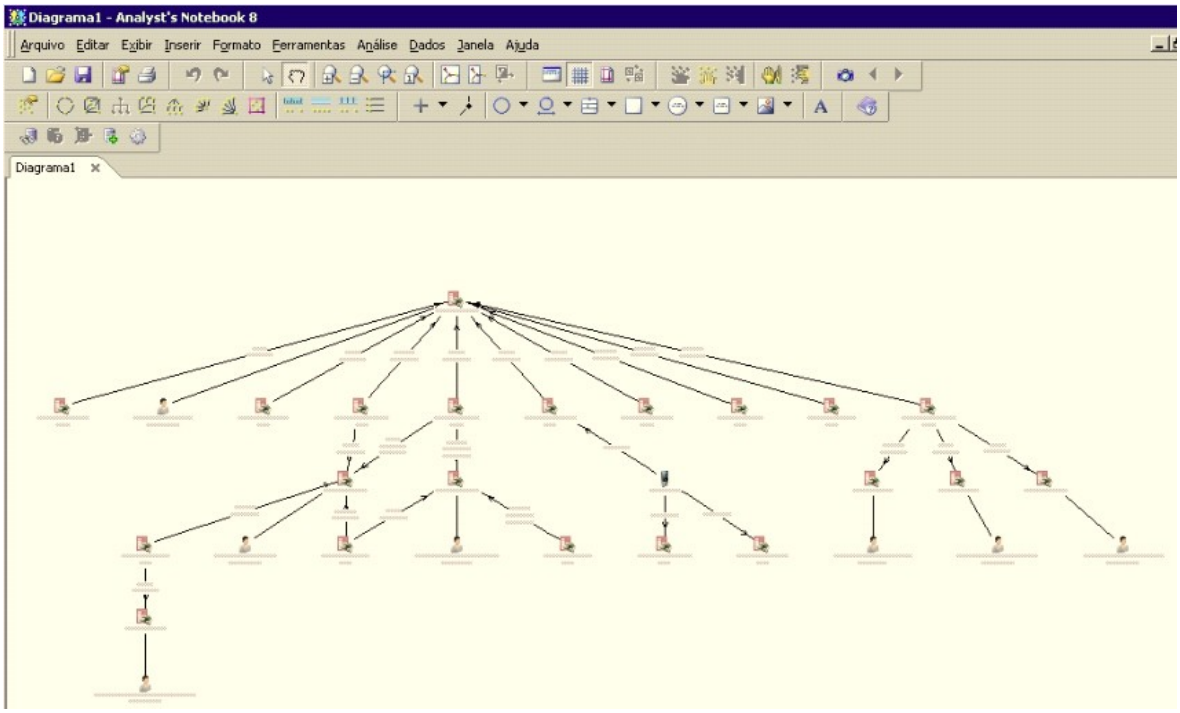


Figura 3.38: Grafo 408 produzido pelo Kraken e desenhado automaticamente no *Analyst´s Notebook* da IBM. Fonte: Elaborada pelo autor.

A Figura 3.39 mostra um *zoom* aplicado na Figura 3.38. Por meio das diversas ferramentas disponíveis no *Analyst´s Notebook* e com a ajuda da sua IG, a análise do crime de transferências bancárias fraudulentas pelo investigador Policial Federal é realizada de uma maneira muito mais rápida, visando a elucidação dos fatos delituosos e objetivando chegar aos autores e beneficiários da organização criminosa.

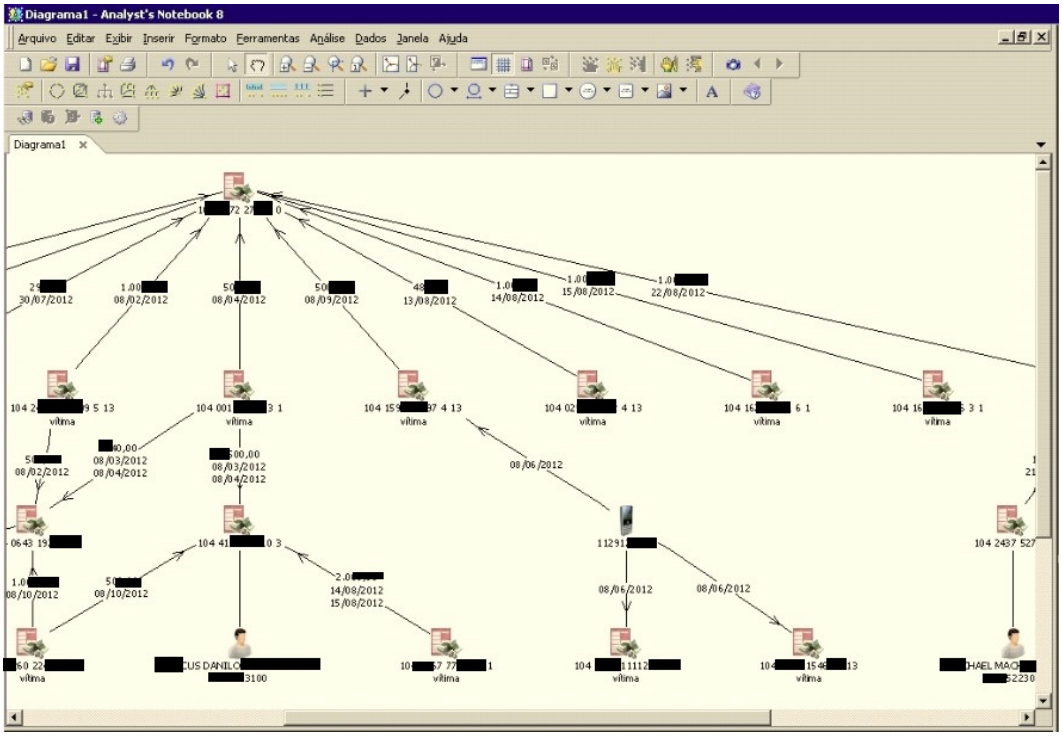


Figura 3.39: Zoom aplicado ao grafo 408 gerado no Kraken e mostrado no *Analyst´s Notebook* da IBM. Fonte: Elaborada pelo autor.

3.6 MÉTRICAS UTILIZADAS

As métricas utilizadas na abordagem do Kraken são divididas em:

- Ordenação dos grafos;
- Filtros aplicados nos grafos;
- Análises de redes realizadas nos grafos.

As métricas de ordenação (CV, CB e total em reais "R$"), juntamente com todas as métricas utilizadas nos filtros aplicados aos grafos servem, em conjunto ou individualmente, para auxiliar o investigador Policial Federal de forma discricionária a priorizar as diversas investigações (grafos) existentes na BNFBE, como foi demonstrado no Capítulo 3, Seção 3.4. Já as métricas utilizadas nas análises de grau, similaridade e recursos alocados servem para destacar os principais atores nos grafos gerados pela abordagem Kraken.

4 RESULTADOS E ANÁLISES

Visando a avaliação do método Kraken foram carregados, processados e analisados dados extraídos da BNFBE no período de dezembro de 2010 a dezembro de 2013. As análises automatizadas do Kraken foram realizadas sobre os vínculos e atores existentes na BNFBE nesse período de datas e, em seguida, confrontadas com os relatórios de análise já feitos pelo MAFBE no mesmo período de datas. A confrontação dos resultados dessas análises é possível, uma vez que a BNFBE possui uma relação dos relatórios de análise de investigações policiais já elaborados pelo MAFBE (por meio da análise humana dos investigadores). Em ambas as abordagens (MAFBE e Kraken) foram extraídos os grafos (investigações) contidos no BNFBE no período supracitado. Utilizando-se as métricas de quantidades de CV, CB, PB, URAs e o valor total em reais das fraudes desses grafos, é possível realizar a comparação dos resultados entre as abordagens mencionadas. Para tanto, foram considerados quatro cenários diferentes:

a) Cenário n° 1: escolheu-se aleatoriamente uma única investigação em específico já realizada pelo MAFBE e comparou-se quantidade de CV, CB, PB, URAs e valor total da fraude em reais dessa investigação com as respectivas quantidades de CV, CB, PB, URAs e valor total da fraude em reais encontrados pelo Kraken sobre a mesma investigação;

b) Cenário n° 2: foram levantadas todas as análises feitas pelo MAFBE na BNFBE, mês a mês, desde dezembro de 2010 até dezembro de 2013, e comparadas com as análises geradas pelo Kraken nos respectivos meses. Nesse cenário, a cada mês foram totalizadas as quantidades de CV, CB, PB, URAs, e o total das fraudes em reais encontradas nos grafos pelo MAFBE e pelo Kraken, sendo acumulados estes valores mês a mês até completarem os 36 meses do período analisado. Desse modo, procedeu-se a comparação quantitativa, a cada mês das CV, CB, URAs, PB, e do total das fraudes em reais levantados pelo MAFBE e pelo Kraken. Em seguida, verificou-se o percentual, mês a mês, de grafos da MAFBE que:

   - Encontraram o mesmo valor total de fraudes em reais, CV, CB, PB e URAs que o Kraken;
   - Encontraram um valor total de fraudes em reais ou quantidades de CV, CB, PB e URAs, porém menor que aquele encontrado pelo Kraken;

– Não foram realizadas análises em grafos (valor total em reais das fraudes igual a 0), mas que o Kraken conseguiu localizar e gerar grafos adicionais (para mais em relação a quantidade de grafos gerados pelo MAFBE) no mesmo período analisado.

c) Cenário n° 3: verificou-se a possível existência de um ganho real na quantidade de CV, CB, PB, URAs e no valor total em reais dos grafos, mês a mês, se for aplicada a abordagem do Kraken sobre os mesmo grafos gerados pelo MAFBE;

d) Cenário n° 4: levantou-se os 100 maiores grafos do MAFBE gerados no período de dezembro de 2010 a dezembro de 2013, no que diz respeito a quantidade de CV, CB, e no valor total das fraudes em reais e se comparou com os 100 maiores grafos gerados no mesmo período pelo Kraken, levando em consideração as mesmas métricas.

A seguir é apresentado o resultado de cada um dos cenários mencionados acima.

4.1 CENÁRIO N° 1 - APLICANDO O KRAKEN A UMA INVESTIGAÇÃO ESPECÍFICA FEITA NO MAFBE

Nesta cenário o foco é a análise dos resultados obtidos aplicando-se o MAFBE e o Kraken em uma investigação específica. A investigação selecionada foi a relativa ao grafo de número 408 (o número 408 foi atribuído de forma sequencial pela heurística do Kraken, durante a geração deste grafo.) Este número 408 representa o nome fictício dado a uma investigação existente na BNFBE. A Figura 4.1 ilustra os grafos resultantes da análise do MAFBE da investigação de n° 408, coletados na BNFBE. Nota-se que existem vários subgrafos desconexos na investigação 408 realizada pela MAFBE. Em entrevistas realizadas com alguns Policiais Federais que investigaram casos similares, foram levantadas algumas hipóteses que podem explicar o porquê da MAFBE encontrar grafos desconexos na investigação de n° 408. Entre esses motivos pode-se elencar:

- O ser humano (investigador) teve dificuldades em localizar vínculos diretos (entre as dezenas/centenas de vínculos diretos existentes entre os atores) na investigação, mesmo realizando iterações na BNFBE por meio do *software Analyst´s Notebook*;
- Poderiam existir dados que por ventura a CAIXA tenha entregue de forma extemporânea, ou seja, depois da análise do MAFBE ter sido feita. Assim, é possível

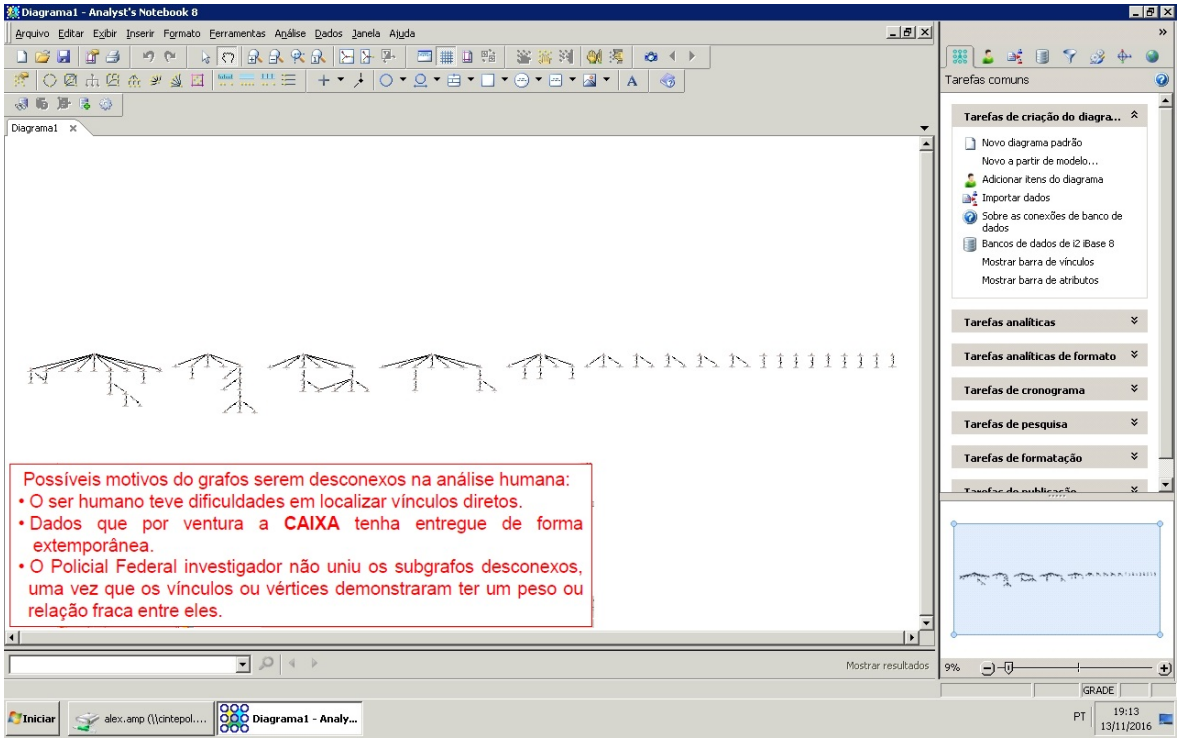


Figura 4.1: Grafos da investigação de n° 408 elaborado MAFBE - *Software Analyst´s Notebook*.

Fonte: Elaborada pelo autor.

que vínculos ou atores não estivessem presentes no momento da elaboração dos grafos pelo MAFBE;

- Uma vez que esta análise do MAFBE usa de forma acentuada as iterações humanas do investigador no *Analyst´s Notebook* junto a BNFBE e, portanto, demanda tempo para a elaboração do relatório de análise, após a entrega desse relatório pode existir, novamente, a entrada de dados extemporâneos pela CAIXA;

- O investigador não uniu os subgrafos desconexos, uma vez que os vínculos ou vértices que poderiam fazer esta união demonstraram ter um peso ou relação muito fraca entre eles, como uma transferência bancária entre uma conta vítima (de um subgrafo A) e conta beneficiária (de um subgrafo B) que teria acontecido em uma data muito distante da data de contestação da vítima junto a CAIXA (que seria, na verdade, o motivo neste caso da investigação 408) e/ou com um valor de transferência de dinheiro irrisório.

Nota-se que essa última situação (o peso dos vínculos de transferências bancárias com datas distantes da data de contestação junto a CAIXA) ainda não é levada em consideração pela abordagem Kraken, e deve ser alvo de trabalhos futuros.

Já a Figura 4.2 ilustra o resultado obtido pela abordagem Kraken levando-se em consideração o mesmo período de datas que originou o grafo 408 do MAFBE, e o mesmo tipo de fraude (transferência bancária).

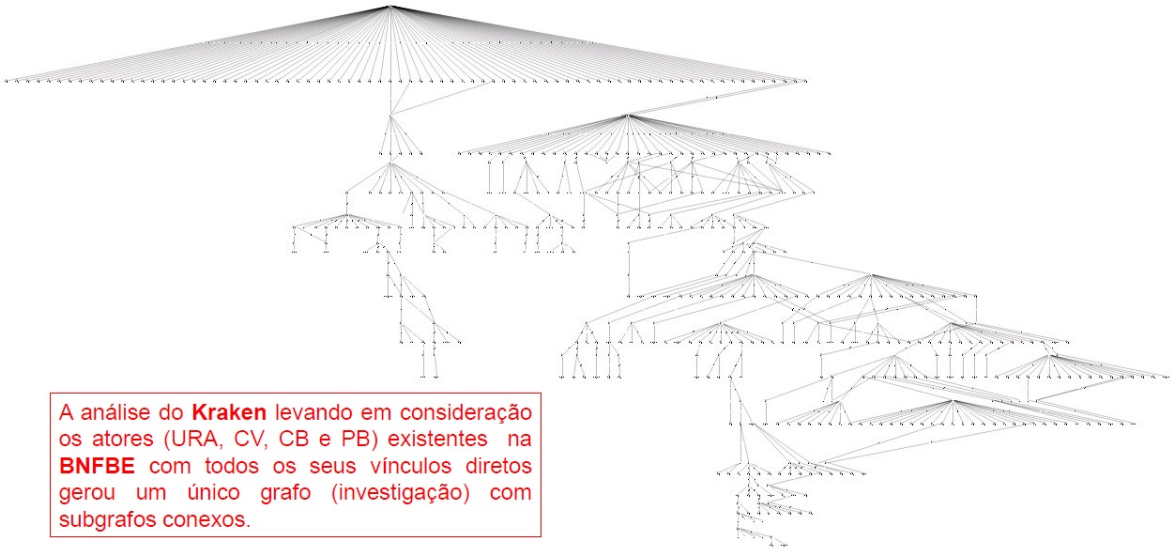


Figura 4.2: Exemplo do grafo de n° 408 elaborado pela abordagem Kraken.
Fonte: Elaborada pelo autor.

A Figura 4.3 ilustra de maneira específica um dos subgrafos da investigação de n° 408 resultado do MAFBE.

A Figura 4.4 mostra como a abordagem Kraken localizou um maior número de vínculos e vértices do que a versão original da investigação 408 concebida pelo MAFBE. Desse modo, o subgrafo inicialmente do MAFBE tornou-se uma parte (subgrafo) menor, do grafo 408 produzido pelo Kraken.

4.1.1 RESULTADOS OBTIDOS NO CENÁRIO N° 1

A seguir são mostradas tabelas com seus respectivos gráficos relativos as quantidades de contas vítimas (CV), contas beneficiárias (CB), pessoas beneficiadas (PB) e valor da fraude (R$), da “investigação 408 do MAFBE” x “a investigação 408 da abordagem do Kraken”. A Tabela 4.1 mostra a comparação de CV, CB e Total em reais do grafo 408. A Tabela 4.2 mostra a comparação de PB e URAs do grafo 408. A Tabela 4.3 mostra a comparação de vértices e arcos do grafo 408. A Tabela 4.4 mostra a comparação de densidade, grau médio e grau máximo do grafo 408. A Figura 4.5 mostra os gráficos da comparação de contas vítimas, e total em R$ do grafo 408. A Figura 4.6 mostra os gráficos da comparação de contas beneficiadas e pessoas beneficiadas do grafo 408.

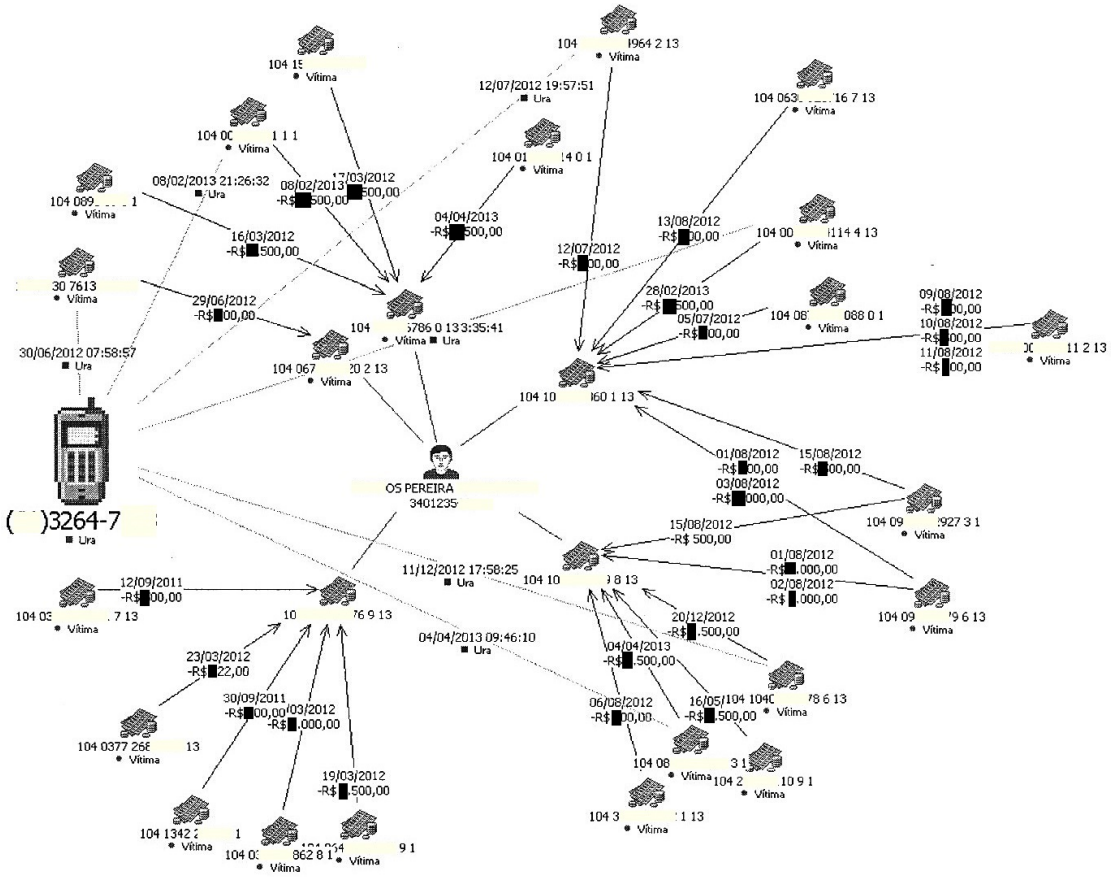


Figura 4.3: Subgrafo da investigação de n° 408 elaborado pelo MAFBE.

Fonte: Elaborada pelo autor.

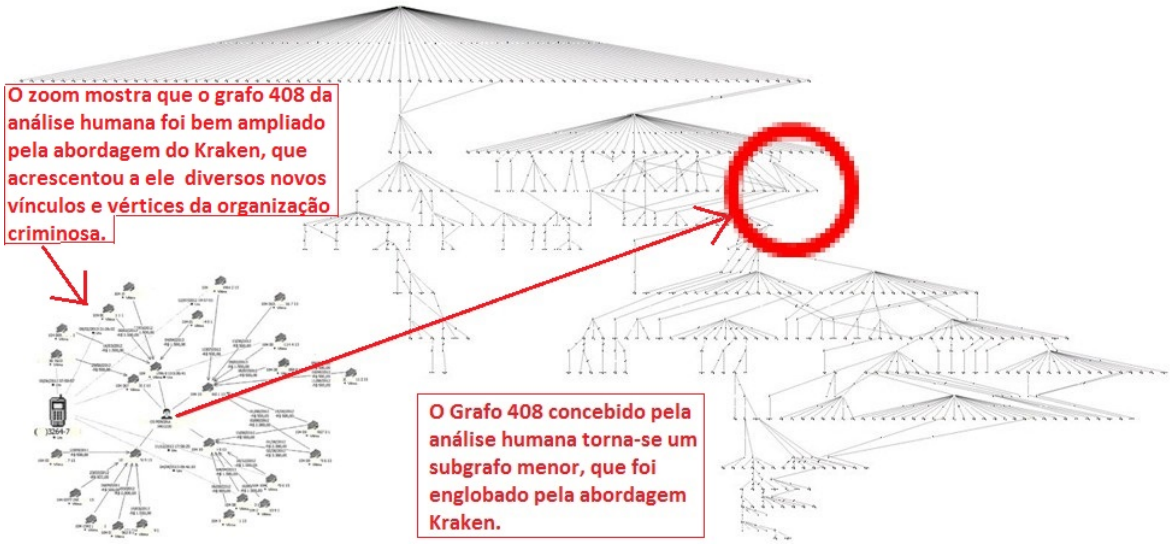


Figura 4.4: Exemplo do grafo de n° 408 elaborado pela abordagem Kraken.

Fonte: Elaborada pelo autor.

Tabela 4.1: 1a. Tabela de comparação grafo 408 - MAFBE x Kraken.

Total em R$ MAFBE	Total em R$ Kraken	% a mais de R$ Kraken	Contas Vítimas (CV) MAFBE	Contas Vítimas (CV) Kraken	% a mais de CV Kraken	Contas Beneficiadas (CB) MAFBE	Contas Beneficiadas (CB) Kraken	% a mais de CB Kraken
105.847,00	253.487,00	139,48%	178	363	103,93%	36	86	138,89%

Tabela 4.2: 2a. Tabela de comparação grafo 408 - MAFBE x Kraken.

Pessoas Beneficiadas (PB) MAFBE	Pessoas Beneficiadas (PB) Kraken	% a mais de PB Kraken	URAs MAFBE	URAs Kraken	% a mais de URAs Kraken
25	65	160,00%	9	40	344,44%

Tabela 4.3: 3a. Tabela de comparação grafo 408 - MAFBE x Kraken.

Vértices MAFBE	Vértices Kraken	% a mais de Vértices Kraken	Arcos MAFBE	Arcos Kraken	% a mais de Arcos Kraken
248	554	123,39%	86	585	580,23%

Tabela 4.4: 4a. Tabela de comparação grafo 408 - MAFBE x Kraken.

Densidade MAFBE	Densidade Kraken	Grau Médio MAFBE	Grau Médio Kraken	Grau Máximo MAFBE	Grau Máximo Kraken
0,00281	0,00382	2,266	2,112	78	78

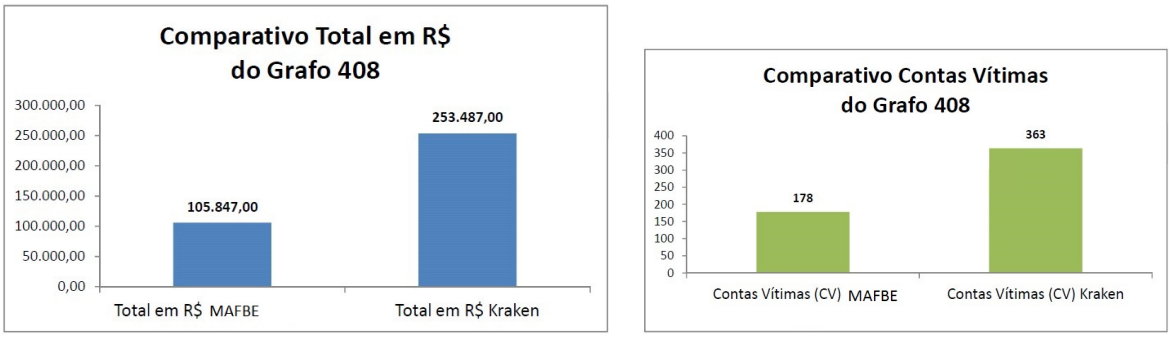


Figura 4.5: Comparação do total em reais e CV do grafo 408 - MAFBE x Kraken.
Fonte: Elaborada pelo autor.

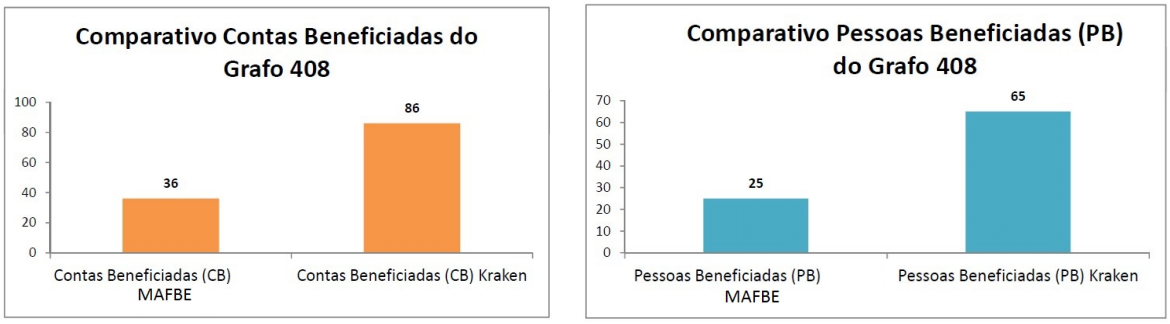


Figura 4.6: Comparação de CB e PB do grafo 408 - MAFBE x Kraken.
Fonte: Elaborada pelo autor.

A Figura 4.7 mostra os gráficos da comparação de URAs e vértices do grafo 408.

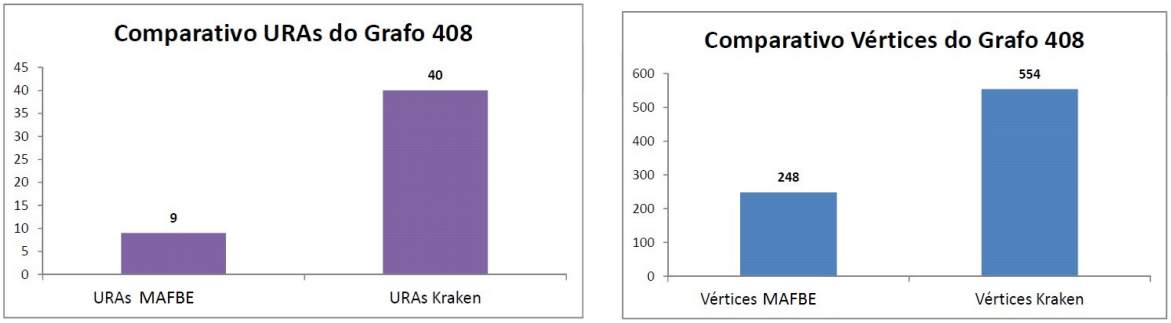


Figura 4.7: Comparação de URAs e vértices do grafo 408 - MAFBE x Kraken.
Fonte: Elaborada pelo autor.

A Figura 4.8 mostra os gráficos da comparação de arcos e densidade do grafo 408.

A Figura 4.9 mostra o gráfico da comparação de grau médio do grafo 408.

Por meio da **IG** do **Kraken** ainda é possível realizar análise de rede no grafo 408, para destacar os autores principais deste grafo de quatro maneiras diferentes:

- Análise de grau dos vértices;
- Análise de similaridade entre CB levando-se em conta a quantidade de contas vítimas que essas têm em comum;
- Análise de recursos alocados onde pondera-se o custo/peso na relação entre a CB e as suas CV.

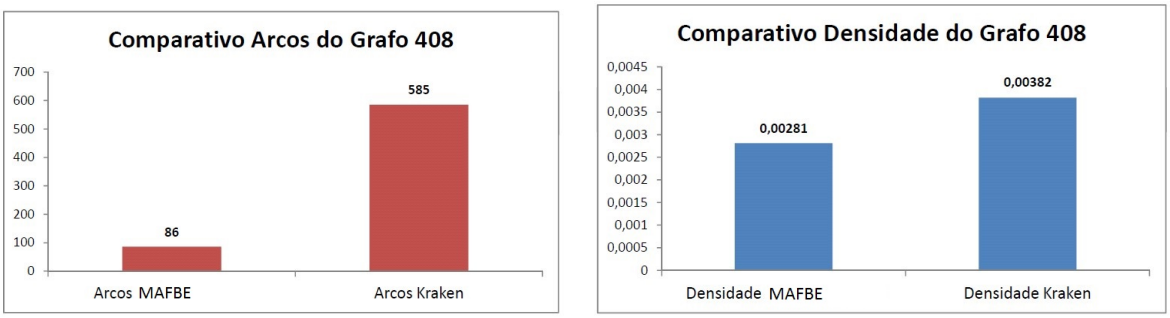


Figura 4.8: Comparação de arcos e densidades do grafo 408 - MAFBE x Kraken.
Fonte: Elaborada pelo autor.

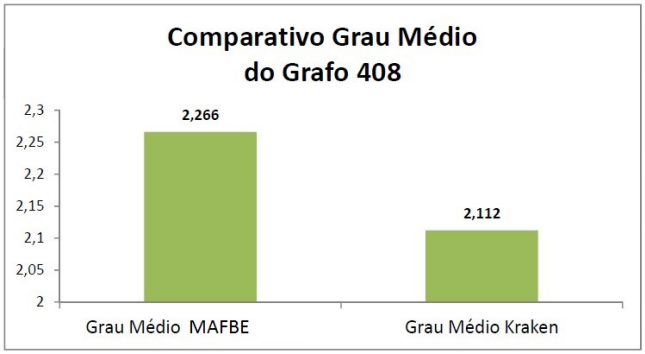


Figura 4.9: Comparativo de grau médio do grafo 408 - MAFBE x Kraken.
Fonte: Elaborada pelo autor.

O estudo de análise de rede tem o propósito de destacar no grafo os principais atores das respectivas análises.

A análise de similaridade entre contas beneficiadas leva em conta as pessoas beneficiadas que tiveram mais contas vítimas em comum, o que indicaria, em tese, um vínculo indireto estre estas pessoas na organização criminosa.

A Figura 4.10 mostra a análise de similaridade e de recursos que o **Kraken** produziu na sua análise sobre o grafo 408 O primeiro registro da Figura 4.10 mostra que as

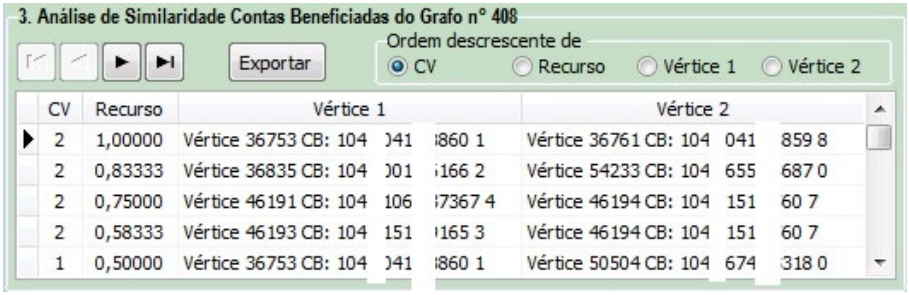


Figura 4.10: Análise de similaridade e de recursos de CB do grafo 408.
Fonte: Elaborada pelo autor.

contas correntes 104 041 860 1 (conta fictícia) e 104 041 859 8 (conta fictícia) possuem 2 (duas) contas vítimas (CV) em comum. E se não bastasse essa similaridade, foi constatado ainda que os titulares dessas contas beneficiadas (CB) utilizaram 100% dos seus recursos nos vínculos entre suas contas beneficiárias e suas contas vítimas. Em tese, isso fortalece a existência de um vínculo indireto na organização criminosa entre os titulares dessas contas beneficiadas do grafo 408.

4.1.2 VANTAGENS DO KRAKEN NO CENÁRIO N° 1

Sem desconsiderar nenhum vínculo direto ou vértice, o Kraken conseguiu unir os demais grafos da investigação de n° 408 ao grafo da Figura 4.3, juntando todos em uma única investigação de maior porte de forma automatizada e em tempo bem reduzido.

Desse modo, o Kraken cria a oportunidade da PF agregar em uma única investigação um maior número de CV, CB, valor total em reais (R$) das fraudes e pessoas envolvidas ou afetadas por um criminoso ou organização criminosa, o que reduz ainda mais o número de IPLs na elucidação de fraudes eletrônicas bancárias.

4.2 CENÁRIO N° 2 - COMPARAÇÃO DA QUANTIDADE DE GRAFOS, CV, CB, PB, URAS E TOTAL DAS FRAUDES EM REAIS LEVANTADAS PELO MAFBE E PELO KRAKEN NO MESMO PERÍODO DE 3 ANOS DA BNFBE

A análise aplicada à este cenário consistiu em comparar a quantidade de grafos, CV, CB, PB, URAs e o total em reais que o MAFBE conseguiu elaborar para os períodos de: 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 18, 24 e 36 meses com as respectivas quantidade dessas métricas geradas pelo Kraken no mesmo período.

4.2.1 RESULTADOS OBTIDOS NO CENÁRIO N° 2

O Kraken varre toda a BNFBE, de uma única vez e em um tempo bastante curto se comparado com a análise humana feita no MAFBE. Assim o MAFBE por demandar um tempo superior nas suas análises de vínculos, realiza um percentual menor de grafos no mesmo período.

Para se ter uma ideia a Tabela 4.5 mostra a quantidade de grafos gerados pelo MAFBE versus a quantidade de grafos gerados pelo Kraken e a comparação do valor total das fraudes em reais (R$) levantado por cada método, ambos aplicados a toda a BNFBE nos períodos de 01 a 36 meses.

Tabela 4.5: Comparação da quantidade de grafos e o total das fraudes em reais levantado pela MAFBE x Kraken aplicado a toda BNFBE.

N° de Meses	Total Grafos Kraken	Total Grafos MAFBE	% Total Grafos MAFBE	Total Grafos MAFBE Valor = Kraken	% Análise MAFBE = Kraken	Total Grafos MAFBE Valor < Kraken	% MAFBE < Kraken	Total Grafos Análise MAFBE = 0	% Análise MAFBE = 0
1	1.594	294	18,44%	276	17,31%	18	1,13%	1.300	81,56%
2	3.324	406	12,21%	382	11,49%	24	0,72%	2.918	87,79%
3	4.969	540	10,87%	504	10,14%	36	0,72%	4.429	89,13%
4	6.336	622	9,82%	567	8,95%	55	0,87%	5.714	90,18%
5	7.519	683	9,08%	622	8,27%	61	0,81%	6.836	90,92%
6	8.548	774	9,05%	700	8,19%	74	0,87%	7.774	90,95%
7	9.610	833	8,67%	745	7,75%	88	0,92%	8.777	91,33%
8	10.375	887	8,55%	788	7,60%	99	0,95%	9.488	91,45%
9	10.837	909	8,39%	803	7,41%	106	0,98%	9.928	91,61%
10	11.269	913	8,10%	802	7,12%	111	0,99%	10.356	91,90%
11	11.656	918	7,88%	200	1,72%	301	2,58%	10.738	92,12%
12	11.989	926	7,72%	812	6,77%	114	0,95%	11.063	92,28%
18	13.638	934	6,85%	283	2,08%	252	1,85%	12.704	93,15%
24	16.052	1.015	6,32%	432	2,69%	190	1,18%	15.037	93,68%
36	22.622	1.264	5,59%	1.071	4,73%	193	0,85%	21.358	94,41%
			9,17%		**7,48%**		**1,09%**		**90,83%**

No período de 01 (um) mês de análise dos registros da BNFBE o Kraken, gerou 1.594 grafos enquanto o MAFBE gerou 294 grafos, ou seja, 18,44% dos grafos em relação ao Kraken no mesmo período. Note-se que na geração dos grafos através do MAFBE, não se pode afirmar com certeza, que estes 294 grafos gerados, foram os 294 grafos que

teriam, em tese, as maiores quantidades de CV, CB, PB, URAS e valor total da fraude em reais (R$) no período supracitado.

Note-se que este percentual de grafos elaborados pelo MAFBE, decai a medida em que o período de análise das fraudes aumentam, chegando a ser de 5,59% em relação ao Kraken quando o período de análise da investigação atinge 36 (trinta e seis) meses.

Uma análise interessante retirada deste experimento é que em alguns casos os grafos gerados pelo MAFBE chegou na mesma quantidade de CV, CB e no mesmo total do valor da fraude (R$) ao dos respectivos grafos gerados pelo Kraken, esse percentual de grafos que possuem métricas iguais ao do Kraken gerados pelo MAFBE chegou a 17,31% para o período de 01 (um) mês e decaiu para 4,73% no período de 36 meses.

Este experimento mostra que existe um percentual de 1,13% de grafos que o MAFBE elaborou no período de 01 (um) mês que possuem grafos correspondentes gerados pelo Kraken neste mesmo período, contudo o MAFBE não atingiu a mesma quantidade CV, CB, ou o valor total da fraude (R$) levantados pelo Kraken, ou seja, o grafo gerado pelo MAFBE tornou-se um subgrafo (uma parte menor) contida dentro de um grafo maior gerado pelo Kraken no mesmo período de 01 (um) mês.

Essa última situação demonstra que o Kraken localizou mais vínculos diretos e/ou vértices (atores), relativos a fraudes de transferências bancárias conexas num mesmo grafo, aumentando assim o tamanho do grafo originalmente gerado pelo MAFBE e consequentemente integrando em uma única investigação policial mais atores, diminuindo significativamente o número de procedimentos investigativos a serem instaurados para a elucidação das fraudes de transferência bancárias da BNFBE, uma das premissas do projeto Tentáculos.

Contudo, é vital que a equipe de Policiais investigadores analisem os vínculos desses grafos gerados pelo Kraken a fim de verificar se existem vértices ou arestas fracas o suficiente para serem eliminadas dentro da metodologia de investigação da PF, o que resultaria na diminuição do tamanho do grafo gerado pelo Kraken e de suas métricas mencionadas acima. O corte de vértices ou arestas fracas de forma automática nos grafos ainda não faz parte da abordagem deste trabalho.

A seguir são mostrados os gráficos que representam as comparações existentes na Tabela 4.5.

A Figura 4.11 mostra gráficos com a comparação dos grafos elaborados pelo MAFBE versus Kraken aplicado a toda a BNFBE

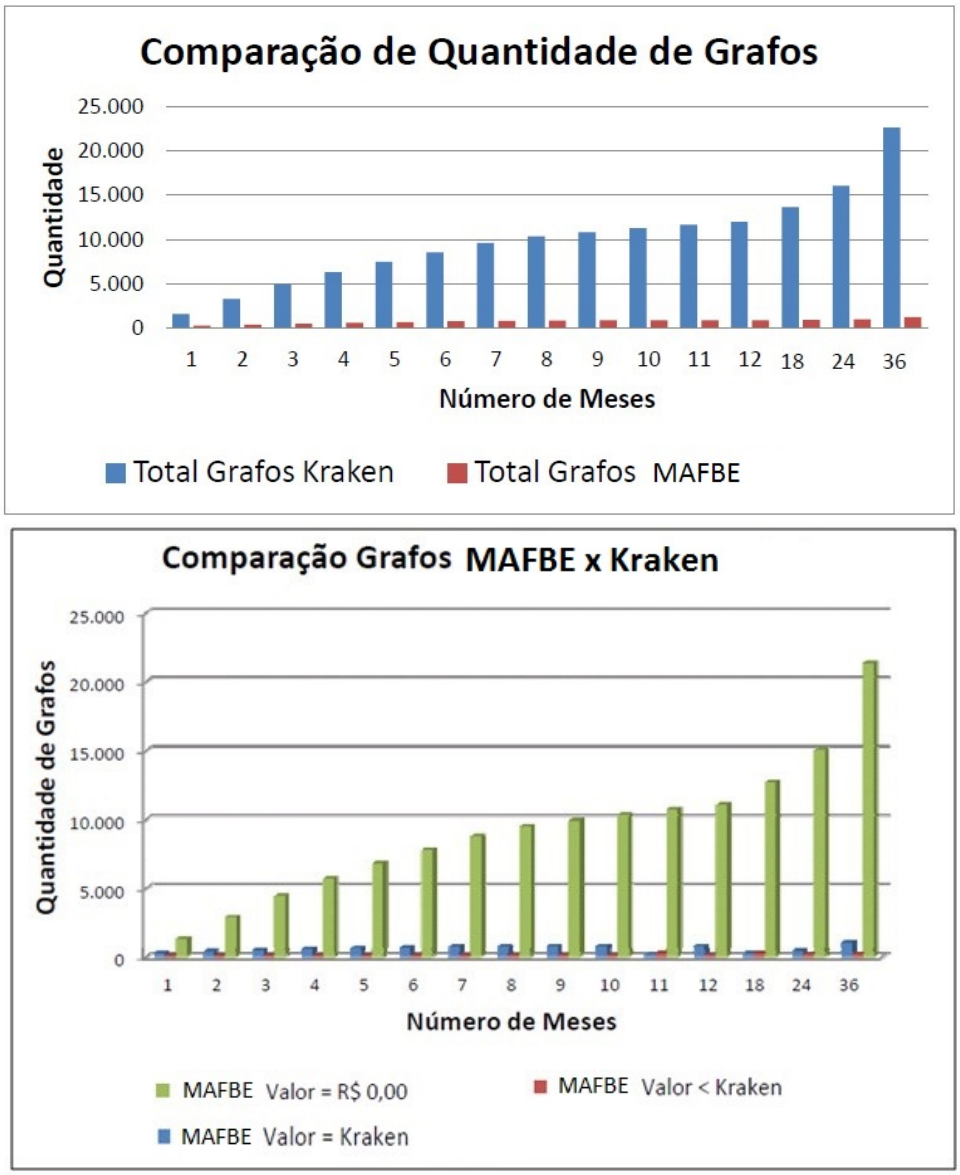


Figura 4.11: Grafos gerados pelo MAFBE x Kraken aplicado a toda BNFBE.
Fonte: Elaborada pelo autor.

A Tabela 4.6 mostra o resultados das comparações entre as CV, CB e o Total em reais levantados pelo MAFBE versus Kraken aplicado a toda BNFBE nos períodos de 01 a 36 meses.

A Figura 4.12 mostra gráficos com a comparação das contas vítimas e contas beneficiadas dos grafos gerados pelo MAFBE versus Kraken aplicado a toda a BNFBE nos períodos supracitados.

Tabela 4.6: Resultado do experimento aplicado a toda BNFBE - MAFBE x Kraken.

N° de Meses	Contas Vítimas (CV) - Kraken	Contas Vítimas (CV) - MAFBE	% CV MAFBE	Contas Beneficiadas (CB) - Kraken	Contas Benenficiadas (CB) - MAFBE	% CB MAFBE	Fraudes Kraken R$	Fraudes MAFBE R$	% R$ MAFBE
1	2.059	388	18,84%	2.786	491	17,62%	3.038.467,59	659.961,04	21,72%
2	4.351	575	13,22%	5.612	696	12,40%	6.221.564,58	1.036.935,21	16,67%
3	6.422	755	11,76%	8.262	930	11,26%	9.043.906,38	1.366.498,14	15,11%
4	8.139	877	10,78%	10.322	1.047	10,14%	11.683.436,23	1.590.553,36	13,61%
5	9.613	986	10,26%	12.066	1.133	9,39%	13.861.562,58	1.837.827,89	13,26%
6	10.894	1.119	10,27%	13.498	1.251	9,27%	15.959.725,84	2.125.556,94	13,32%
7	12.175	1.198	9,84%	14.945	1.333	8,92%	18.036.903,64	2.267.405,21	12,57%
8	13.122	1.260	9,60%	16.003	1.410	8,81%	19.472.249,34	2.369.362,09	12,17%
9	13.800	1.287	9,33%	16.695	1.439	8,62%	20.444.433,03	2.409.685,32	11,79%
10	14.416	1.296	8,99%	17.316	1.453	8,39%	21.159.519,21	2.426.435,32	11,47%
11	15.310	1.308	8,54%	17.762	1.467	8,26%	21.761.080,68	2.451.205,32	11,26%
12	16.026	1.334	8,32%	18.173	1.484	8,17%	22.315.089,98	2.491.041,32	11,16%
18	19.006	1.441	7,58%	20.136	1.510	7,50%	25.491.591,20	2.550.748,08	10,01%
24	24.407	1.622	6,65%	24.290	1.634	6,73%	35.736.329,48	3.014.989,32	8,44%
36	35.102	2.188	6,23%	33.787	2.031	6,01%	67.211.558,96	5.088.715,81	7,57%
			10,01%			**9,43%**			**12,67%**

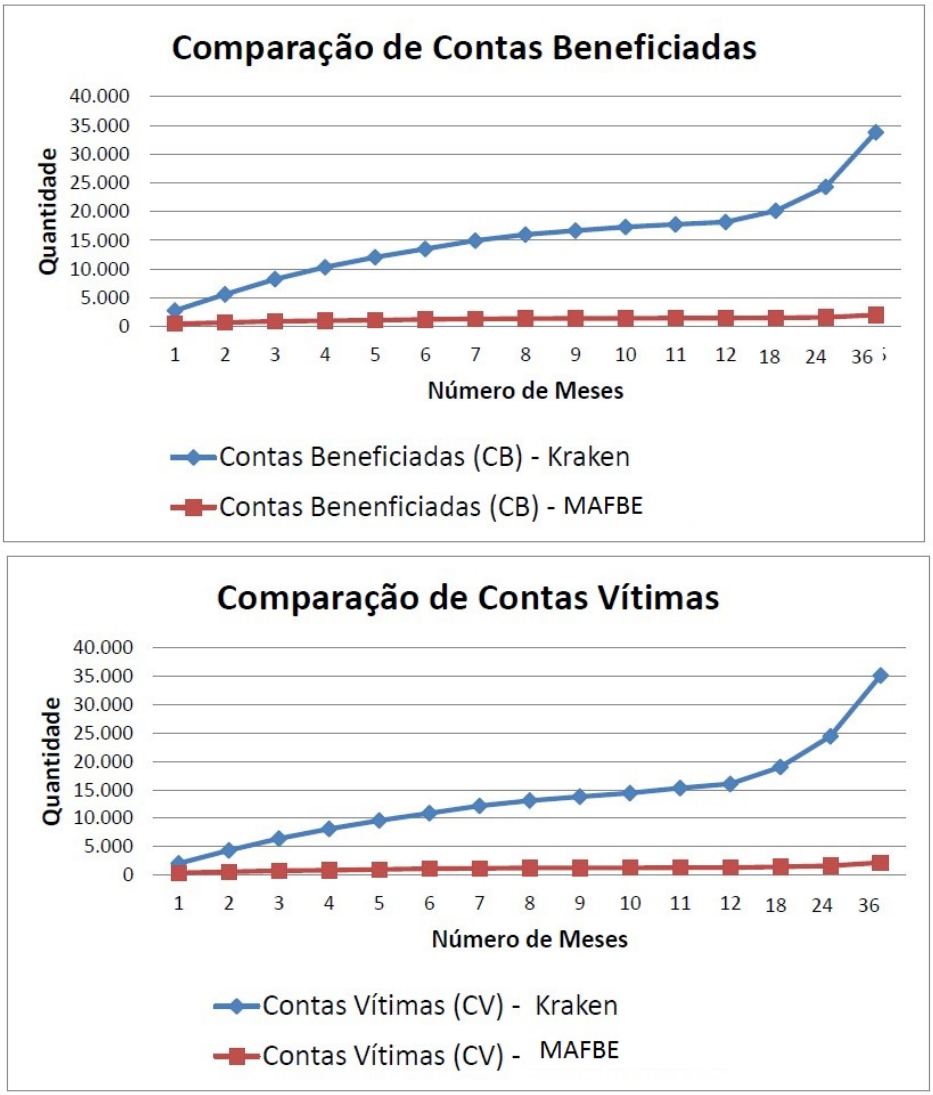


Figura 4.12: Contas vítimas e contas beneficiadas dos grafos gerados pelo MAFBE x Kraken aplicado a toda BNFBE. Fonte: Elaborada pelo autor.

A Figura 4.13 mostra gráficos com a comparação do valor da fraude em reais (R$) dos grafos gerados pelo MAFBE versus Kraken aplicado a toda a BNFBE nos períodos supracitados.

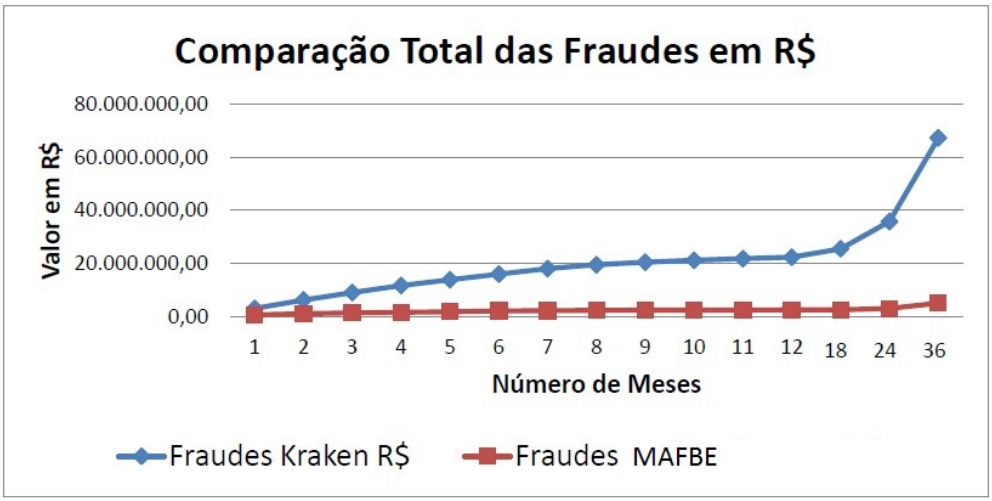


Figura 4.13: Valor das fraudes em R$ dos grafos gerados pelo MAFBE x Kraken aplicado a toda BNFBE. Fonte: Elaborada pelo autor.

A Tabela 4.7 mostra o resultados das comparações entre as PB e URAs levantados pelo MAFBE versus Kraken aplicado a toda BNFBE nos períodos de 01 a 36 meses.

Destaque-se que os valores nas colunas de URAS (telefones qe ligaram para contas vítimas) igual a zero que aparecem na Tabela 4.7 nos meses de 01 a 09, se deve ao fato que somente depois de alguns meses da implantação do projeto Tentáculos, a CAIXA começou a encaminhar para a PF registros contento os telefones que os fraudadores utilizavam para acessar as URAs da CAIXA, a fim de verificar o saldo das contas vítimas.

A Figura 4.14 mostra gráficos com a comparação das URAs e PB dos grafos elaborados pelo MAFBE versus Kraken aplicado a toda a BNFBE nos períodos supracitados.

A Tabela 4.8 mostra o resultados das comparações das quantidades de vértices, arcos e da densidade média dos grafos levantados pelo MAFBE versus Kraken aplicado a toda BNFBE nos períodos de 01 a 36 meses.

A Tabela 4.8 revela que os grafos gerados pelo Kraken são mais esparsos que os gerados pelo MAFBE.

Tabela 4.7: Comparação das pessoas beneficiadas e URAs levantadas pela MAFBE x Kraken aplicado a toda BNFBE.

N° de Meses	Pessoas Beneficiadas (PB) - kraken	Pessoas Beneficiadas (PB) - MAFBE	% PB MAFBE	URA Kraken	URA MAFBE	% Ura MAFBE
1	2.744	479	17,46%	0	0	0,00%
2	5.534	676	12,22%	0	0	0,00%
3	8.134	901	11,08%	0	0	0,00%
4	10.141	1.013	9,99%	0	0	0,00%
5	11.827	1.087	9,19%	0	0	0,00%
6	13.188	1.186	8,99%	0	0	0,00%
7	14.587	1.264	8,67%	0	0	0,00%
8	15.611	1.339	8,58%	0	0	0,00%
9	16.268	1.367	8,40%	1	0	0,00%
10	16.820	1.380	8,20%	15	1	6,67%
11	17.213	1.393	8,09%	123	3	2,44%
12	17.594	1.410	8,01%	194	6	3,09%
18	19.443	1.431	7,36%	508	7	1,38%
24	23.258	1.543	6,63%	923	19	2,06%
36	31.325	1.879	6,00%	1.295	37	2,86%
			9,26%			**1,23%**

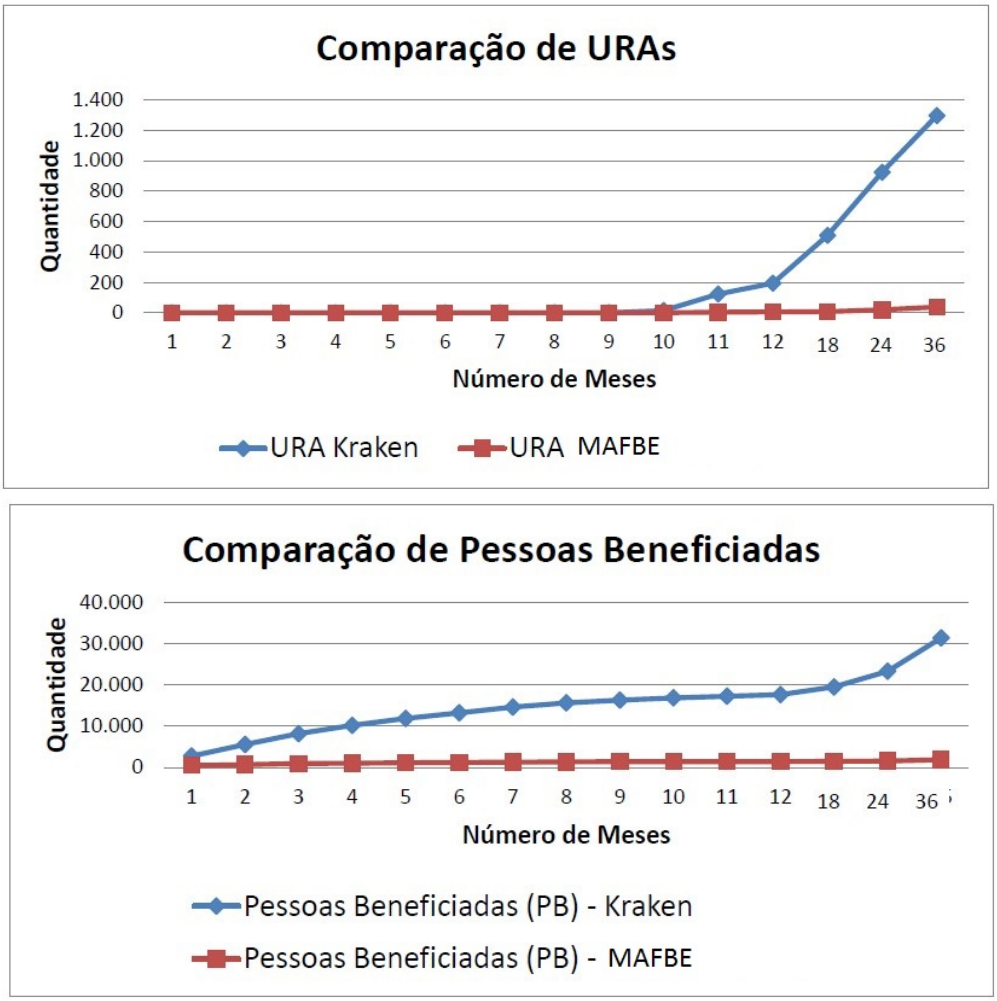


Figura 4.14: URAs e pessoas beneficiadas dos grafos gerados pelo MAFBE x Kraken aplicado a toda BNFBE. Fonte: Elaborada pelo autor.

Tabela 4.8: Comparação de vértices, arcos e densidade média levantados pela MAFBE x Kraken aplicado a toda BNFBE.

N° de Meses	Total Grafos Kraken	Total Grafos MAFBE	% Total Grafos MAFBE	Vértices MAFBE	Vértices Kraken	% a mais de Vértices do Kraken	Arcos MAFBE	Arcos Kraken	% a mais de Arcos do Kraken	Densidade Média MAFBE	Densidade Média Kraken	% para menos na Densidade Média do Kraken
1	1.594	294	18,44%	1.358	7.589	458,84%	585	6.042	932,82%	0,00063	0,00021	-66,67%
2	3.324	406	12,21%	1.947	15.497	695,94%	867	12.253	1313,26%	0,00046	0,00010	-78,26%
3	4.969	540	10,87%	2.586	22.818	782,37%	1.146	17.952	1466,49%	0,00034	0,00007	-79,41%
4	6.336	622	9,82%	2.937	28.602	873,85%	1.305	22.387	1615,48%	0,00030	0,00005	-83,33%
5	7.519	683	9,08%	3.206	33.506	945,10%	1.432	26.121	1724,09%	0,00028	0,00005	-82,14%
6	8.548	774	9,05%	3.556	37.580	956,81%	1.588	29.185	1737,85%	0,00025	0,00004	-84,00%
7	9.610	833	8,67%	3.795	41.707	999,00%	1.689	32.255	1809,71%	0,00023	0,00004	-82,61%
8	10.375	887	8,55%	4.009	44.736	1015,89%	1.773	34.532	1847,66%	0,00022	0,00003	-86,36%
9	10.837	909	8,39%	4.093	46.764	1042,54%	1.807	36.105	1898,06%	0,00022	0,00003	-86,36%
10	11.269	913	8,10%	4.130	48.567	1075,96%	1.824	37.491	1955,43%	0,00021	0,00003	-85,71%
11	11.656	918	7,88%	4.171	50.408	1108,54%	1.845	38.984	2012,95%	0,00021	0,00003	-85,71%
12	11.989	926	7,72%	4.234	51.987	1127,85%	1.876	40.266	2046,38%	0,00021	0,00003	-85,71%
13	13.638	934	6,85%	4.389	59.093	1246,39%	1.917	45.790	2288,63%	0,00020	0,00003	-85,00%
14	16.052	1.015	6,32%	4.818	72.878	1412,62%	2.098	57.638	2647,28%	0,00018	0,00002	-88,89%
15	22.622	1.264	5,59%	6.135	101.509	1554,59%	2.744	80.000	2815,45%	0,00015	0,00002	-86,67%
			9,17%			**1019,75%**			**1874,10%**			**-83,12%**

A Equação 2.1 explica por que os grafos gerados pelo Kraken são mais esparsos que os gerados pelo MAFBE, isso se deve ao fato da abordagem do Kraken localizar um número maior de vértices (CB, CV, URA, PB) em relação ao MAFBE, o que proporciona grafos conexos com menos arcos que os do MAFBE.

4.2.2 VANTAGENS DO KRAKEN NO CENÁRIO N° 2

A complexidade da investigação aumenta a medida em que a quantidade das análises de vínculos são acumuladas mês a mês ao longos desses 36 meses analisados na BNFBE.

Assim com o passar dos meses fica mais difícil para o MAFBE, que utiliza das iterações do ser humano (investigador) junto o *Analyst´s Notebook* para visualizar na BNFBE os vínculos diretos que relacionam os atores das investigações em curso.

Neste cenário de n° 2 verifica-se que o Kraken, por ser automatizado, ter uma regra de negócio bem definida, que não despreza vínculos fracos entre os atores e sofre menos interferência de dados extemporâneos inseridos na BNFBE, conseguiu gerar um número maior de grafos e de CV, CB, PB, URAs e um total em reais de fraude maior em relação o MAFBE e que a diferença a favor do Kraken aumenta consideravelmente a medida em que a análise das fraudes de transferências entre contas bancárias engloba mais meses.

4.3 CENÁRIO N° 3 - APLICANDO A ABORDAGEM DO KRAKEN SOMENTE SOBRE OS GRAFOS GERADOS PELO MAFBE

Este cenário de n° 3 vislumbra os experimentos realizados com o Kraken somente sobre os mesmos grafos inicialmente elaborados pelo MAFBE. Assim, teremos a ideia de como seriam, em tese, os resultados dos grafos elaborados pelo MAFBE se os investigadores tivessem utilizado a abordagem do Kraken para automatizar as análise dos vínculos dos atores durante suas investigações.

4.3.1 RESULTADOS OBTIDOS NO CENÁRIO N° 3

A Tabela 4.9 mostra a quantidade de grafos gerados pelo MAFBE (investigador Policial Federal) no periodo de 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 18, 24 e 36 meses com a comparação da quantidade de CV e CB que a o MAFBE encontrou versus o que foi levantado pela abordagem do Kraken.

Tabela 4.9: Método Kraken aplicado nos mesmos grafos produzidos pela MAFBE na comparação de CV e CB.

N° de Meses	Total Grafos MAFBE	Contas Vitimas (CV) - MAFBE	Contas Vitimas - Kraken	% de CV realizado a mais pelo Kraken	Contas Beneficiadas (CB) - MAFBE	Contas Beneficiadas - Kraken	% de CB realizado a mais pelo Kraken
1	294	388	426	9,79%	491	511	4,07%
2	406	575	631	9,74%	696	725	4,17%
3	540	755	843	11,66%	930	987	6,13%
4	622	877	1.011	15,28%	1.047	1.138	8,69%
5	683	986	1.139	15,52%	1.133	1.235	9,00%
6	774	1.119	1.298	16,00%	1.251	1.375	9,91%
7	833	1.198	1.405	17,28%	1.333	1.464	9,83%
8	887	1.260	1.489	18,17%	1.410	1.555	10,28%
9	909	1.287	1.527	18,65%	1.439	1.591	10,56%
10	913	1.296	1.543	19,06%	1.453	1.611	10,87%
11	918	1.308	1.561	19,34%	1.467	1.629	11,04%
12	926	1.334	1.595	19,57%	1.484	1.647	10,98%
18	934	1.441	1.772	22,97%	1.510	1.686	11,66%
24	1.015	1.622	2.734	68,56%	1.634	1.925	17,81%
36	1.264	2.188	4.428	102,38%	2.031	2.718	33,83%
				25,60%			**11,26%**

Através desses experimentos é possível verificar que no período de 01 (um) mês o MAFBE levantou 294 grafos, encontrou 388 contas vítimas e 491 contas beneficiadas. O Kraken localizou 426 contas vítimas e 511 contas beneficiadas nos mesmos 295 grafos. Isso demonstra um ganho, em tese, na utilização da abordagem do Kraken de 9,79% em relação as contas vítimas e de 4,07% em relação as contas beneficiadas do MAFBE.

Destaque-se que por óbvio, podem existir vínculos ou vértices fracos na análise levantada pelo Kraken, que podem vir a serem descartados, baseados nos métodos do MAFBE, a posteriori, o que diminuiria as quantidades dos atores mencionados acima gerados pelo Kraken.

A Figura 4.15 mostra gráficos com resultados de CV e CB aplicando-se o Kraken nos mesmos grafos produzidos pelo MAFBE nos períodos de 01 a 36 meses.

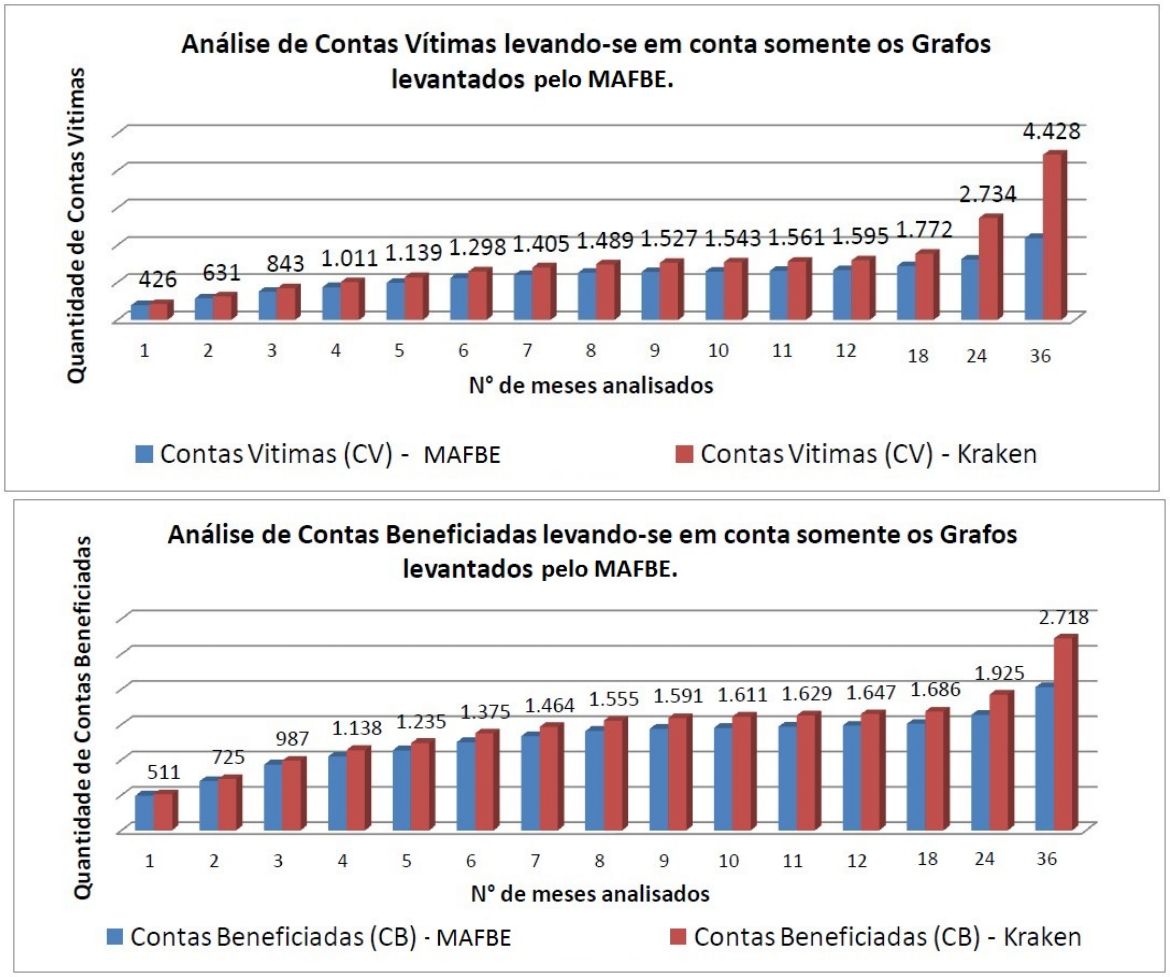


Figura 4.15: Resultado de CV e CB aplicando-se o Kraken nos mesmos grafos produzidos pelo MAFBE. Fonte: Elaborada pelo autor.

A Tabela 4.10 mostra a quantidade de grafos produzidos pelo MAFBE (investigador Policial Federal) nos respectivos períodos supracitados além da comparação do total das fraudes em reais e de pessoas beneficiadas que o investigador encontrou versus o que foi levantado pela abordagem do Kraken, nos mesmos grafos durante os períodos de 01 a 36 meses de análise na BNFBE.

A Figura 4.16 mostra gráficos com resultado do total de fraudes em reais e PB aplicando-se o Kraken nos mesmos grafos produzidos pelo MAFBE humana nos períodos de 01 a 36 meses.

Através desses experimentos é possível verificar que no período de 01 (um) mês o

Tabela 4.10: Método Kraken aplicado nos mesmos grafos produzidos pelo MAFBE na comparação do total da fraude em reais e PB.

N° de Meses	Total Grafos MAFBE	Total em R$ das Fraudes MAFBE	Total em R$ em Fraudes Kraken	% de Fraudes (R$) realizado a mais pelo Kraken	Pessoas Beneficiadas (PB) - MAFBE	Pessoas Beneficiadas (PB) - Kraken	% de PB realizado a mais pelo Kraken
1	294	659.961,04	728.736,04	10,42%	479	497	3,76%
2	406	1.036.935,21	1.137.792,21	9,73%	676	701	3,70%
3	540	1.366.498,14	1.532.480,86	12,15%	901	950	5,44%
4	622	1.590.553,36	1.888.759,17	18,75%	1.013	1.091	7,70%
5	683	1.837.827,89	2.178.853,70	18,56%	1.087	1.175	8,10%
6	774	2.125.556,94	2.530.801,48	19,07%	1.186	1.288	8,60%
7	833	2.267.405,21	2.731.785,40	20,48%	1.264	1.368	8,23%
8	887	2.369.362,09	2.878.427,74	21,49%	1.339	1.454	8,59%
9	909	2.409.685,32	2.940.595,55	22,03%	1.367	1.484	8,56%
10	913	2.426.435,32	2.966.495,55	22,26%	1.380	1.499	8,62%
11	918	2.451.205,32	3.005.053,55	22,59%	1.393	1.513	8,61%
12	926	2.491.041,32	3.046.789,55	22,31%	1.410	1.530	8,51%
18	934	2.550.748,08	3.247.240,47	27,31%	1.431	1.561	9,08%
24	1.015	3.014.989,32	4.088.147,96	35,59%	1.543	1.775	15,04%
36	1.264	5.088.715,81	7.551.696,81	48,40%	1.879	2.444	30,07%
				22,07%			**9,51%**

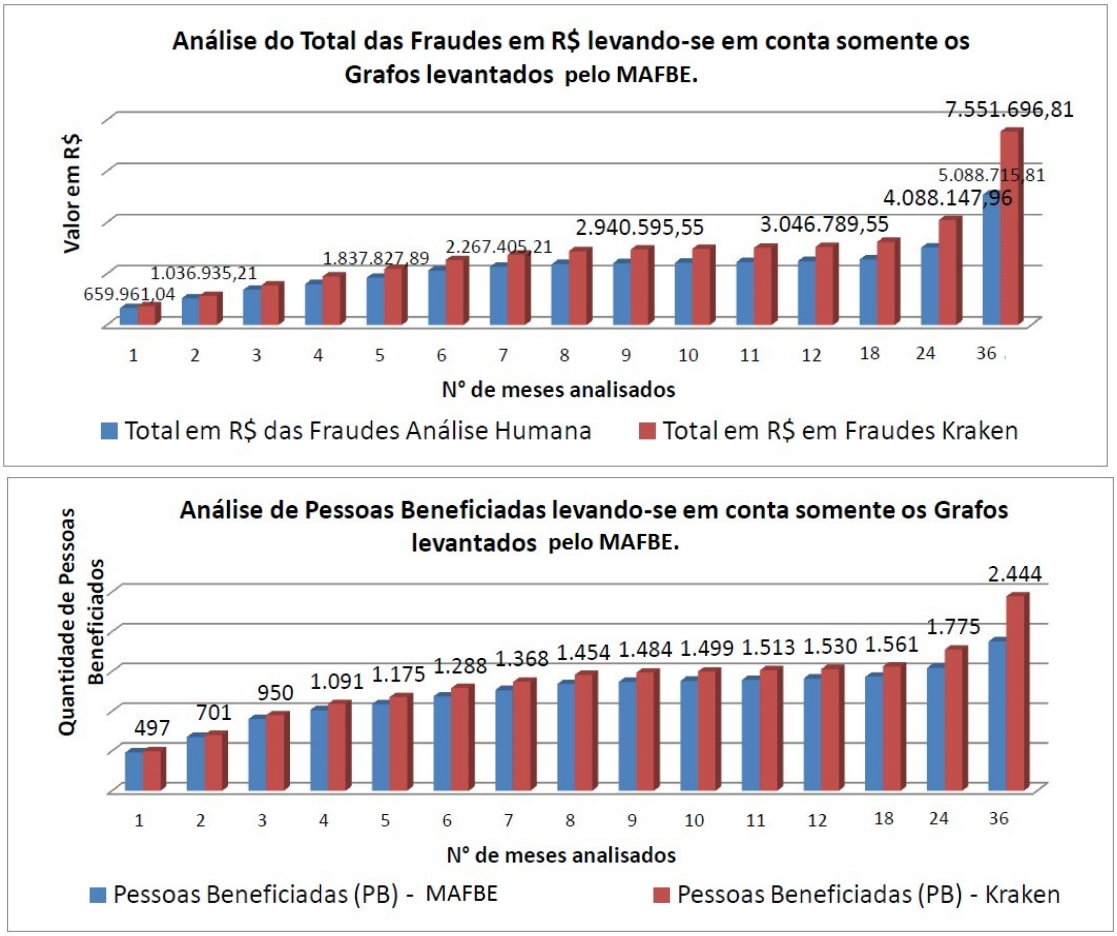


Figura 4.16: Resultado do total da fraude em reais e PB aplicando-se o Kraken nos mesmos grafos produzidos pelo MAFBE. Fonte: Elaborada pelo autor.

MAFBE levantou 294 grafos encontrando um total de fraudes no valor de R$ 659.961,04 e 479 pessoas beneficiadas (PB) contra R$ 728.736,04 de fraudes e 497 pessoas beneficiadas do Kraken. Fica assim demonstrando um ganho na abordagem do Kraken de 10,42% no total das fraudes em reais e de 3,76% na quantidade de pessoas beneficiadas em relação ao MAFBE.

A medida em que a investigação aumenta seu período de meses, a fim de englobar um número maior de transações bancárias fraudulentas, em tese praticadas pelo mesmo criminoso ou organização criminosa, a diferença a favor do Kraken em relação ao MAFBE aumenta consideravelmente.

No período de 36 meses a diferença a favor da abordagem do Kraken foi de 48,40% a mais no total das fraudes em reais e de 30,07% de pessoas beneficiadas.

A Tabela 4.11 mostra a quantidade de grafos produzidos pelo MAFBE (investigador Policial Federal) nos períodos de 01 a 36 meses, além da comparação de URAs que o MAFBE encontrou versus o que foi levantado pelo Kraken, nos mesmos grafos durante a análise na BNFBE.

Tabela 4.11: Método Kraken aplicado nos mesmos grafos produzidos pelo MAFBE na comparação de URAs.

N° de Meses	Total Grafos MAFBE	URA MAFBE	URA Kraken
1	294	0	0
2	406	0	0
3	540	0	0
4	622	0	0
5	683	0	0
6	774	0	0
7	833	0	0
8	887	0	0
9	909	0	0
10	913	1	1
11	918	3	3
12	926	6	7
18	934	7	15
24	1.015	19	95
36	1.264	37	282

A Figura 4.17 mostra gráficos com resultado de URAs aplicando-se o Kraken nos mesmos grafos produzidos pela análise humana nos períodos de 01 a 36 meses.

No experimento das URAs verifica-se que por meses o valor foi igual a zero. Isso se deve ao fato de que nos primeiros 09 (nove) meses do projeto Tentáculos a CAIXA não enviou a informação a respeito dos telefones que entraram em contato com URAs

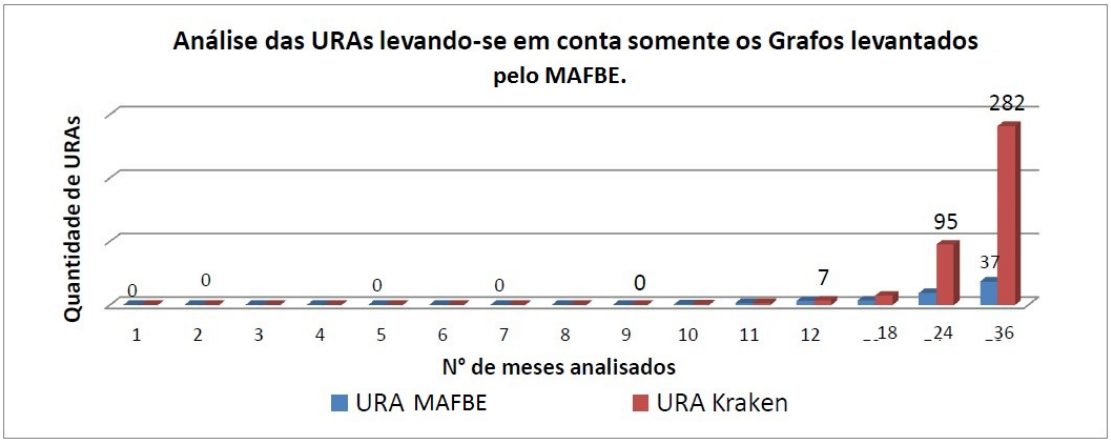


Figura 4.17: Resultado do total de URAS aplicando-se o Kraken nos mesmos grafos produzidos pelo MAFBE. Fonte: Elaborada pelo autor.

da CAIXA. Essa informação só passou a ser fornecida em setembro de 2011. Por esse motivo, nem o MAFBE e nem o Kraken conseguiram processar essa informação nos meses anteriores.

4.3.2 VANTAGENS DO KRAKEN NO CENÁRIO N° 3

Para todos os períodos desse experimento, a abordagem do Kraken mostrou um ganho crescente a medida em que a investigação aumenta seu período de meses, intensificando assim a sua complexidade. Em 36 meses de investigações de transferências bancárias, a diferença a favor do Kraken em relação ao MAFBE foi de 102,38% a mais nas contas vítimas e de 33,83% nas contas beneficiadas.

Assim, a utilização da abordagem desse trabalho, baseado na análise de vínculos proposta originalmente pelo MAFBE, demonstrou acelerar a investigação de forma substancial, como veremos mais adiante nos tempos de processamento do Kraken para os períodos supracitados. Além disso, conseguiu resultados expressivos no que diz respeito a quantidade de contas vítimas, contas beneficiadas, pessoas beneficiadas, URAs e total dos valores das fraudes em reais (R$) em relação ao MAFBE no mesmo período de tempo e nas mesmas investigações.

4.4 CENÁRIO N° 4 - COMPARAÇÃO DOS 100 MAIORES GRAFOS GERADOS PELO MAFBE COM OS 100 MAIORES GRAFOS GERADOS PELO KRAKEN

No cenário de n° 4 é mostrado o resultado da comparação dos experimentos realizados nos 100 maiores grafos elaborados pelo MAFBE contra os 100 maiores grafos gerados

pela abordagem do Kraken. A ideia é verificar se a metodologia do MAFBE consegue varrer a BNFBE e retirar dela as maiores investigações em relação ao Kraken.

4.4.1 RESULTADOS OBTIDOS NO CENÁRIO N° 4

A Tabela 4.12 mostra a comparação dos 100 maiores grafos elaborados pelo MAFBE versus os 100 maiores grafos gerados pelo Kraken.

Tabela 4.12: Comparação entre os 100 maiores grafos do MAFBE x 100 maiores grafos do Kraken.

Comparação 100 maiores Investigações	Total em R$ Fraudes	Contas Vítimas	Contas Beneficiadas	Pessoas Beneficiadas	URAs	Vértices	Arcos	Densidade	Grau Médio	Grau Máximo
MAFBE	2.006.099,30	687	427	360	20	1.494	842	0,17325	1,890	78
Kraken	7.636.747,23	2.999	1.701	1.388	231	6.319	6.724	0,10987	1,993	78
% a mais Kraken	**280,68%**	**336,54%**	**298,36%**	**285,56%**	**1055,00%**	**322,96%**	**698,57%**			

Para se conseguir os valores mostrados na Tabela 4.12 relativos ao MAFBE, foi realizado um *SQL* na "tabela_ resumo_grafo" da base do "tentaculos_saida" com a cláusula "*order by Total_Fraudes_Tentaculos desc limit 100*", o que garante que foram retornados os 100 maiores grafos baseados no Total de Fraudes em reais feitos pelo MAFBE. Após o retorno desta consulta, procedemos com a totalização das colunas de interesse desse experimento. O *SQL* com os dados que foram retornados dessa consulta pode ser visualizado na Figura 4.18

Da mesma maneira, para se conseguir os valores mostrados na Tabela 4.12 relativos a abordagem Kraken, foi realizado um *SQL* na "tabela_ resumo_grafo" da base do "tentaculos_saida" com a cláusula "*order by Total_Fraude_em_Reais desc limit 100*", o que garante que foram retornados os 100 maiores grafos baseados no Total de Fraudes em reais feitos pela abordagem do Kraken. Após o retorno dessa consulta procedemos com a totalização das colunas de interesse desse experimento. O *SQL* e o retorno dessa consulta podem ser visualizados na Figura 4.19

A seguir são mostrados gráficos que foram extraídos da Tabela 4.12

A Figura 4.20 mostra os gráficos da comparação do total em reais e CV dos 100 maiores grafos.

A Figura 4.21 mostra os gráficos da comparação de CB e de PB dos 100 maiores grafos.

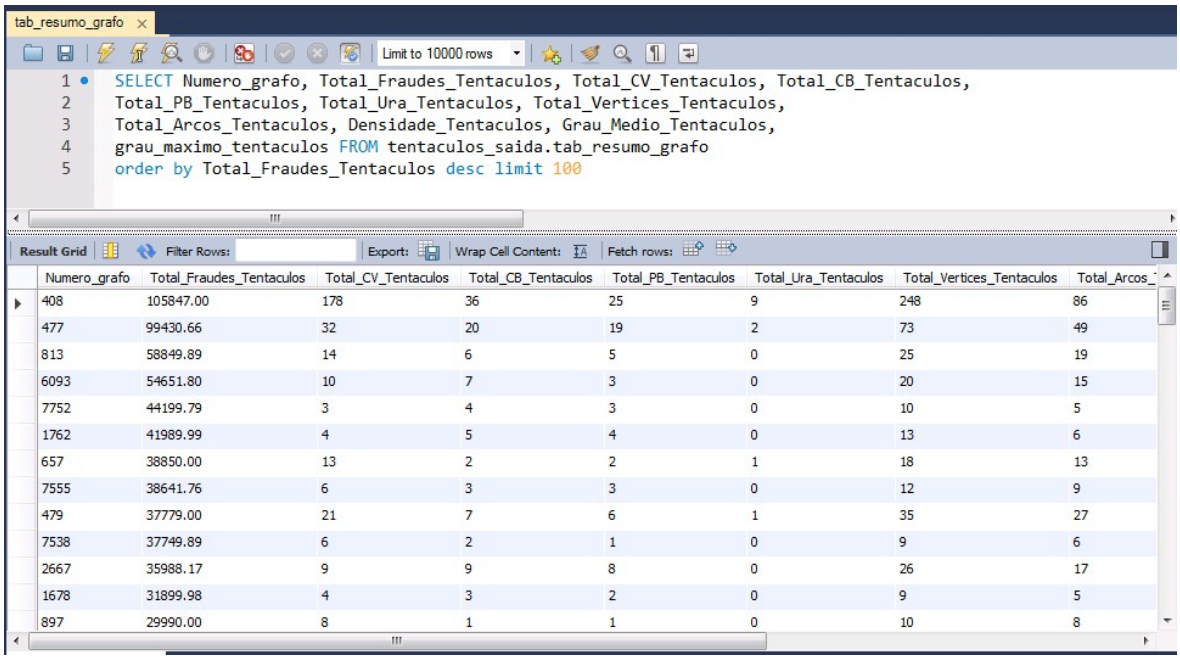


Figura 4.18: *SQL* realizado na base do "tentaculos_saida" para a extração dos 100 maiores grafos feitos pelo MAFBE. Fonte: Elaborada pelo autor.

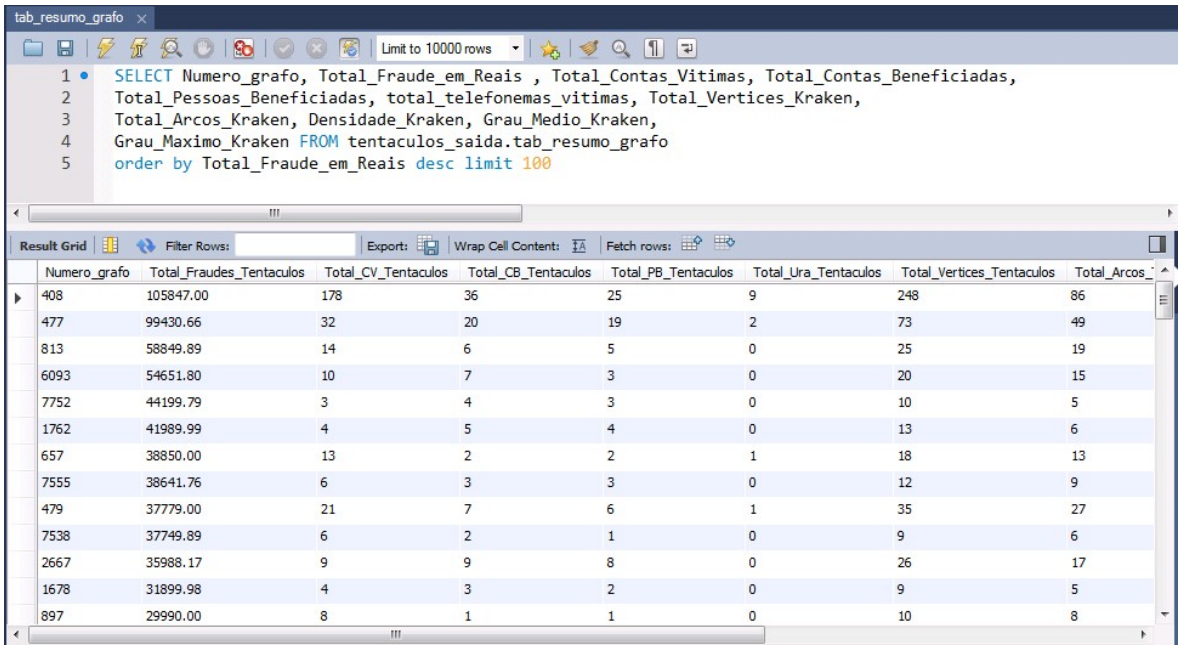


Figura 4.19: *SQL* realizado na base do "tentaculos_saida" para a extração dos 100 maiores grafos feitos pela abordagem do Kraken. Fonte: Elaborada pelo autor.

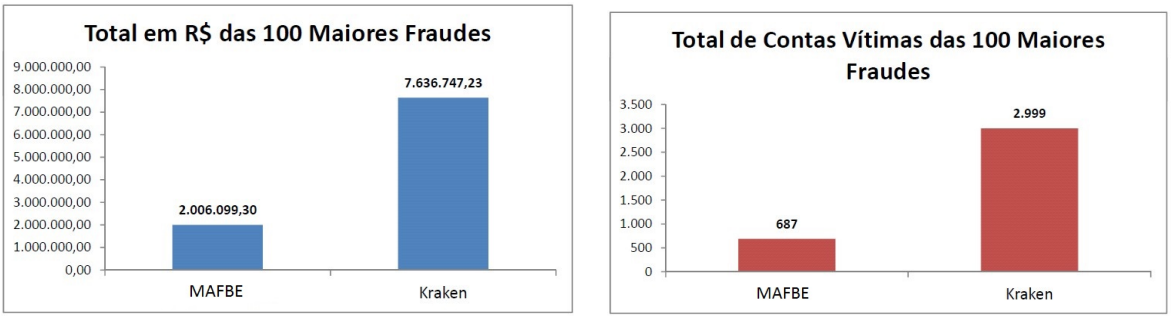


Figura 4.20: Comparação do total em fraudes (R$) e CV dos 100 maiores grafos - MAFBE x Kraken. Fonte: Elaborada pelo autor.

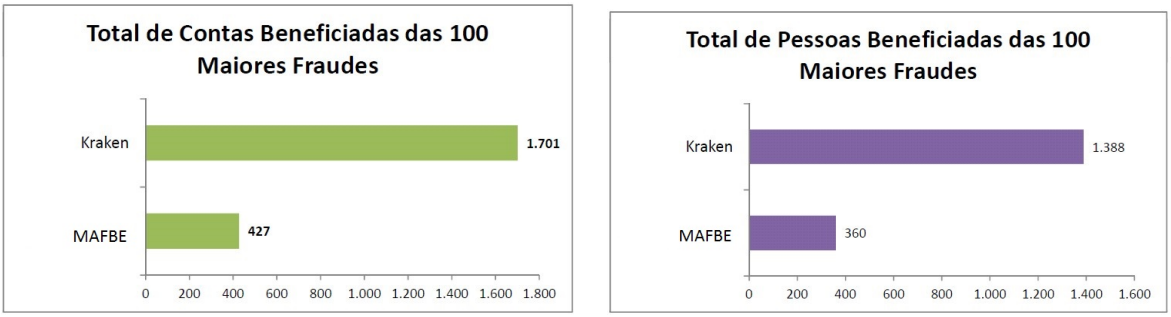


Figura 4.21: Comparação das CB e PB dos 100 maiores grafos - análise humana x Kraken.

Fonte: Elaborada pelo autor.

A Figura 4.22 mostra os gráficos da comparação de URAs e vértices dos 100 maiores grafos.

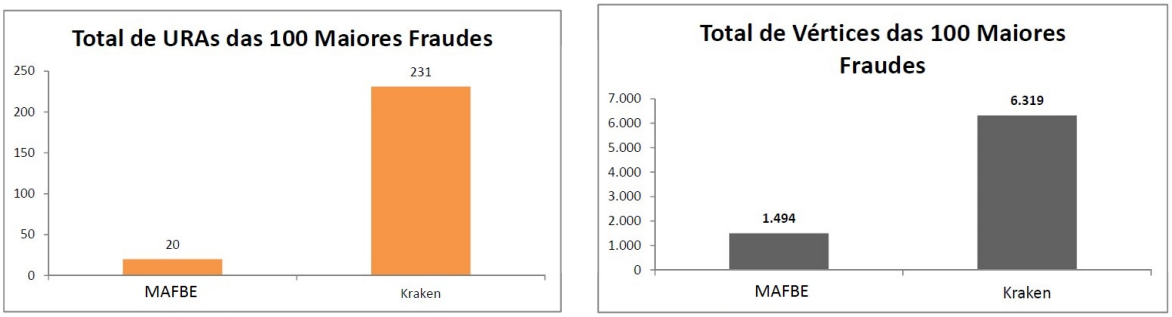


Figura 4.22: Comparação URAs e vértices dos 100 maiores grafos - MAFBE x Kraken.

Fonte: Elaborada pelo autor.

A Figura 4.23 mostra os gráficos da comparação de arcos dos 100 maiores grafos.

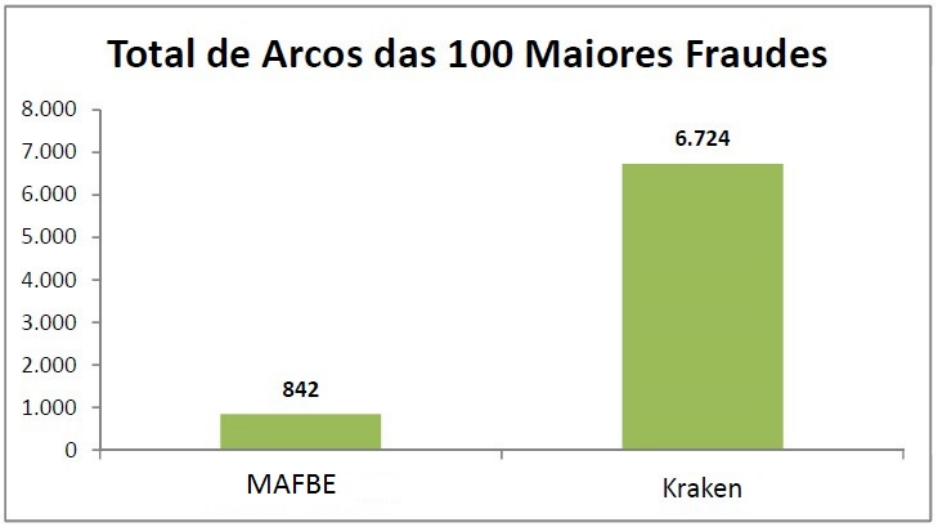


Figura 4.23: Comparação dos arcos dos 100 maiores grafos - MAFBE x Kraken.

Fonte: Elaborada pelo autor.

4.4.2 VANTAGENS DO KRAKEN NO CENÁRIO N° 4

A partir dos resultados mostrados no experimentos dos 100 maiores grafos, é possível afirmar que apesar da análise do MAFBE gerar inúmeras investigações (grafos), a abordagem automatizada do Kraken consegue gerar um número bem superior de grafos. Por esse motivo, consegue-se uma amostra maior das investigações contidas na BNFBE e, a partir dessa amostra, selecionam-se as maiores investigações baseadas nas métricas elencadas na Seção 4.4.1

Devem ser considerados os dados que por ventura a CAIXA tenha entregue de forma extemporânea, depois que o MAFBE tenha elaborado suas análises. Assim, o seu reprocessamento antes da entrega dos relatórios de análises é inviabilizado.

A abordagem proposta pelo Kraken leva em conta todo o banco de dados da BNFBE, sofrendo menos influência dos dados entregues pela CAIXA de forma extemporânea. Isso se deve ao fato de que a análise de vínculos do Kraken consegue gerar em poucas horas os grafos com um número elevado de vértices (atores) relacionados entre si (arcos).

A IG do Kraken permite, de forma segura, ordenar os grafos por ordem decrescente do valor total das fraudes em reais (R$), CV e CB. Assim, o investigador pode selecionar com uma maior precisão os 100 maiores grafos (investigações) existentes naquele período de datas na BNFBE para serem exportados para o *Analyst´s Notebook*.

4.5 RESTRIÇÕES DA ABORDAGEM KRAKEN EM ANÁLISES REAIS

A abordagem proposta pelo Kraken neste trabalho possui atualmente a restrição de só trabalhar com a regra de negócio modelada para fraudes eletrônicas de transferências entre contas bancárias da CAIXA. Em análises reais será necessário não só a modelagem de regras de negócio adicionais para serem incorporadas ao ferramental do Kraken (IG), mas também a remodelação das tabelas de apoio do Kraken para o processamento de dados de novas instituições financeiras que por ventura venham a integrar a BNFBE no futuro. Entre as novas regras de negócios que devem ser modeladas para a utilização do Kraken em toda a BNFBE, destacam-se: compras fraudulentas realizadas pela *Internet*, saques em ATMs, compras feitas com cartões clonados em lojas físicas (postos de gasolinas, grandes magazines, etc), pagamentos de boletos bancários, transferência de recargas para celulares, entre outras. O custo de processamento irá aumentar, e como a *compact forward and reverse star representation* opera em memória *RAM*, será necessária a aquisição de uma máquina servidora com uma maior capacidade de processamento.

4.5.1 TEMPO DE PROCESSAMENTO DOS EXPERIMENTOS

O tempo de processamento dos experimentos mencionados neste capítulo foi obtido utilizando-se uma máquina com processador i5-520M (Dual Core) de 2.4 GHz, com memória RAM de 4 GBytes e disco de 500 GBytes, e do banco de dados *MySql* versão

5.7.12 para armazenar a base teste da BNFBE com as alterações necessárias citadas na seção 3.2

Nota-se que os percentuais mostrados em todos os experimentos acima, no que diz respeito a análise MAFBE, podem sofrer algum tipo de variação para mais, uma vez que a base contendo os processos de fraudes bancárias eletrônicas já analisados pelos investigadores da PF somente é atualizada de tempos em tempos. Sendo assim, isso pode ter levado a uma variação a respeito da data da coleta dessa informação junto a BNFBE para comparação com o método Kraken.

A Tabela 4.13 mostra o tempo gasto pelo método Kraken no processamento dos períodos de 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 18, 24 e 36 meses da BNFBE, a fim de gerar os resultados dos experimentos mencionados neste trabalho.

Tabela 4.13: Tempo de processamento do Kraken por meses analisados da BNFBE.

N° de Meses	Data Inicio Transferências Bancárias	Data Fim Transferências Bancárias	Data do Início do Processamento do Kraken	Data do Fim do Processamento do Kraken	Tempo de Duração do Processamento do Kraken (H:M:S)	Total Grafos Kraken	Total Grafos MAFBE
1	2010-12-09	2011-01-09	17/10/2016 às 10:19:53	17/10/2016 às 11:08:00	0:48:07	1.594	294
2	2010-12-09	2011-02-09	17/10/2016 às 11:33:17	17/10/2016 às 13:09:44	1:36:27	3.324	406
3	2010-12-09	2011-03-09	17/10/2016 às 15:31:18	17/10/2016 às 17:32:13	2:00:55	4.969	540
4	2010-12-09	2011-04-09	17/10/2016 às 19:11:15	17/10/2016 às 21:43:39	2:32:24	6.336	622
5	2010-12-09	2011-05-09	08/10/2016 às 20:43:56	08/10/2016 às 23:54:06	3:10:10	7.519	683
6	2010-12-09	2011-06-09	09/10/2016 às 01:17:43	09/10/2016 às 05:43:01	4:29:18	8.548	774
7	2010-12-09	2011-07-09	09/10/2016 às 07:15:04	09/10/2016 às 12:39:30	5:24:26	9.610	833
8	2010-12-09	2011-08-09	09/10/2016 às 15:34:15	09/10/2016 às 21:15:43	5:41:28	10.375	887
9	2010-12-09	2011-09-09	09/10/2016 às 20:59:59	10/10/2016 às 04:59:56	6:00:56	10.837	909
10	2010-12-09	2011-10-09	17/10/2016 às 00:28:11	17/10/2016 às 06:48:08	6:19:57	11.269	913
11	2010-12-09	2011-11-09	18/10/2016 às 12:31:31	18/10/2016 às 19:27:30	6:55:59	11.656	918
12	2010-12-09	2011-12-09	18/10/2016 às 20:44:36	19/10/2016 às 03:21:18	7:21:18	11.989	926
18	2010-12-09	2012-06-09	17/10/2016 às 21:59:20	18/10/2016 às 08:56:00	10:56:40	13.638	934
24	2010-12-09	2012-12-09	21/10/2016 às 16:49:59	22/10/2016 às 08:32:40	15:42:41	16.052	1.015
36	2010-12-09	2013-12-09	22/10/2016 às 09:02:49	23/10/2016 às 19:22:28	34:19:39	22.622	1.264

A Figura 4.24 mostra o gráfico do tempo de processamento (horas, minutos e segundos) gasto pelo Kraken para realizar as análises nos períodos de 01 a 36 meses da BNFBE.

Com base na Tabela 4.13, podemos afirmar que a ferramenta Kraken trouxe um ganho de tempo expressivo (de semanas, e até meses) na elaboração e confecção dos grafos voltados para as investigações de transferências bancárias e seus vínculos diretos da BNFBE. Sabe-se que a elaboração desses mesmos grafos, ainda que em poucas quantidades (algumas dezenas), exigem mão de obra treinada nas ferramentas *Analyst´s Notebook e Ibase* da IBM. Isso demanda muitas iterações humanas do investigador policial, resultando em dias, semanas ou meses para a elaboração dos diversos grafos (algumas dezenas) que podem vir a compor uma investigação de fraudes bancárias eletrônicas.

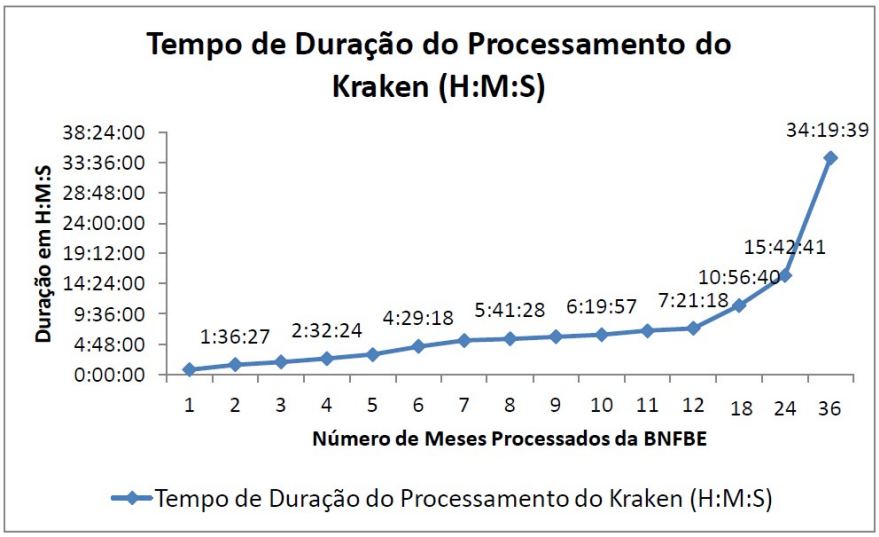


Figura 4.24: Tempo de processamento do Kraken por meses analisados na BNFBE.
Fonte: Elaborada pelo autor.

Portanto, é lógico afirmar ser humanamente impossível a geração de todos os grafos contidos na BNFBE no tempo de processamento executado pelo Kraken. Não apenas devido ao fato da necessidade de treinamento específico do quadro de investigadores Policiais Federais nas ferramentas da IBM utilizadas no MAFBE e suas diversas iterações, mas também pelo próprio tamanho da BNFBE (com milhões de registros) e que tende, em tese, a aumentar, conforme demonstrado no início desse trabalho.

5 CONCLUSÕES

Parafraseando Mayer-Schonberger & Cukier (2013), a era do big-data desafia a maneira como vivemos e interagimos com o mundo, existindo trabalhos em grande escala que não podem ser feitos em escala menor, para extrair novas ideias e criar novas formas de valor de maneiras que alterem as organizações e por que não a forma como investigamos crimes de fraudes bancárias eletrônicas.

5.1 RESULTADOS ENCONTRADOS

A visão da abordagem do Kraken segue esta linha de pensamento ao realizar a análise dos crimes contidos na BNFBE baseada na metodologia original do projeto Tentáculos, mas agora de uma única vez, podendo considerar todo o período da sua existência. Não haverá, portanto, a necessidade de limitar as análises à períodos mais curtos.

Por meio de uma lista de adjacência utilizando uma representação compacta (*compact forward and reverse star representation*), foi criado um modelo computacional da representação de grafos de fraudes bancárias eletrônicas do tipo transferência de valores entre contas bancárias existentes na BNFBE.

Com a ajuda de um algoritmo do tipo *Depth-First Search*, que se mostrou o mais adequado para a regra de negócio de transferências bancárias fraudulentas, foi elaborado um ferramental capaz de agrupar subgrafos conexos por meio de seus vínculos diretos e vértices (atores como contas vítimas, contas beneficiadas, URAs e pessoas beneficiadas). Esse agrupamento ocorre de forma ultra-rápida, otimizando em muito os resultados das investigações de transferências bancárias fraudulentas em relação a tradicional análise humana, conforme demonstrado nos resultados dos vários experimentos desse trabalho.

A Tabela 4.13 mostra o quão rápido a abordagem do Kraken processa dezenas de milhares de atores envolvidos nas fraudes de transferências bancárias e milhões de transações (vínculos) relacionadas a esses atores (CV, CB, PB, URAs). Desse modo, fica ilustrado o quão veloz e eficiente é o modelo de lista de adjacência compacta adotado nesse trabalho.

Com o Kraken foi possível um ganho espantoso de tempo na investigação de transferências bancárias fraudulentas, visto que com a geração do roteiro em formato *Excel*, a intervenção manual do policial investigador na geração de grafos juntos as ferramentas *Analyst´s Notebook e Ibase* da IBM foi reduzida ao máximo. Desse modo, o Kraken dispensa a necessidade das diversas iterações do ser humano para a geração dos grafos, como relatado no Capítulo 2 desse trabalho, com exceção daquelas intervenções essenciais para eventuais cortes de vértices ou arestas que possam ser considerados como fracos pela metodologia atual de investigação da PF, uma vez que estes cortes feitos pelo MAFBE não foram foco deste trabalho.

Assim, diante dos resultados exibidos, esse trabalho obteve êxito, de maneira forense, na elaboração de um ferramental que automatize as análise de vínculos. Essa automatização foi proposta originalmente no projeto Tentáculos, entre os atores de fraudes eletrônicas do tipo transferências bancárias, conseguindo gerar de forma automática os grafos que representam as investigações contidas na base BNFBE da PF. Com a ajuda da sua IG, o Kraken prioriza quais dessas investigações devem ser executadas baseadas em métricas objetivas como a quantidade de vítimas, quantidades de beneficiários e o valor total das fraudes em reais, entre outras, não existindo assim mais surpresas das variáveis envolvidas na investigação como na análise humana, na qual os resultados só aparecem no final das diversas iterações junto as ferramentas da IBM.

Logo, a abordagem do Kraken permitiu o direcionamento do processo de investigação de uma forma padronizada, a fim de obter um ganho expressivo de tempo, produtividade e de qualidade na geração dos relatórios de análise das fraudes bancárias do tipo transferências bancárias, em relação a análise humana. Além disso, demonstrou ainda ser um ferramental voltado para a gestão de negócio, que neste caso é a apuração de fraudes bancárias eletrônicas em desfavor da CAIXA pela PF.

As análises dos atores da rede dos grafos feitas pelo Kraken foi um desafio a parte, conforme demonstrado na Seção 4.1. Essas análises de rede envolvendo a similaridade entre contas beneficiadas, a utilização de recursos gastos pelos atores trazem um aprimoramento na elucidação dos crimes de fraudes bancárias eletrônicas. Ao permitir ao Policial Federal investigador destacar nos grafos gerados pelo Kraken quais são os atores (CV, CB, URAs, PB) mais relevantes em cada uma dessas análises, facilita a sua visualização principalmente quando se trata de grafos que possuam dezenas, centenas de atores envolvidos na rede criminosa.

Com a abordagem desse trabalho, a PF poderá enfrentar de maneira forense a crescente demanda de fraudes bancárias eletrônicas no Brasil em desfavor da CAIXA e demais instituições financeiras que por acaso vierem a fazer parte no futuro do projeto Tentáculos. Com alguns ajustes, a regra de negócio do Kraken pode ser adaptada para investigações de crimes contra previdência social (INSS), desvio de verbas públicas entre outros.

5.2 PUBLICAÇÃO

Este trabalho gerou a publicação do artigo "Técnicas baseadas em Grafos para Priorização de Investigações Policiais de Fraudes Bancárias Eletrônicas", no V Workshop de Forense Computacional (WFC) do XVI Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais (SBSeg 2016), que foi apresentado na Universidade Federal Fluminense (UFF) em Niterói/RJ em 10/11/2016 pelo autor deste trabalho.

5.3 TRABALHOS FUTUROS

Trabalhos futuros devem levar em conta outros tipos de fraudes bancárias eletrônicas no seu modelo computacional, tais quais: saques em ATMs, pagamento de boletos fraudados, compras realizadas pela *Internet*, entre outros. Sugere-se a inserção de pesos nos vínculos que relacionam os atores da investigação no algoritmo computacional de agrupamento dos subgrafos conexos, a fim de realizar filtros que possam vir a excluir automaticamente vértices ou arestas considerados fracos. Desse modo, subgrafos que representem melhor o *modus operandi* de um fraudador ou organização criminosa serão isolados.

Investir ainda mais em métodos de análise de redes para destacar de forma automática os principais atores dentro de um grafo quando este for desenhado por meio de uma ferramenta como o *Analyt´s Notebook* da IBM, por exemplo.

Por fim, há de se pensar em criar um padrão de relatório de análise de forma automática, uma vez que durante a análise dos vínculos dos atores dos grafos de crimes de fraudes bancárias eletrônicas, a regra desse negócio já expõe o fluxo do dinheiro (roteiro do crime) de sua origem (CV) até seu destino (CB).

Este relatório serviria, no futuro, como um template para iterações do investigador Policial Federal durante sua investigação. Isso diminuiria ainda mais o tempo da elu-

cidação do crime de fraude bancária eletrônica, desde sua contestação administrativa junto a entidade financeira até a sua elucidação criminal pela Polícia Federal do Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

BERLUSCONI, G. et al. Link prediction in criminal networks: A tool for criminal intelligence analysis. *PLOS ONE*, Public Library of Science, v. 11, n. 4, p. 1–21, abr. 2016. Disponível em: <http://dx.doi.org/10.1371/journal.pone.0154244>.

CAVALLARO, L. *Malicious Software and its Underground Economy Two Sides to Every Story*. 2014. Curso online. Disponível em: *https://pt.coursera.org/learn/malsoftware*. Acesso em março, 2016.

DANTAS, G. F. D. L. et al. *Análise Criminal: Novas Tendências em Proveito da Análise Criminal Avançada e da Inteligência de Segurança Pública*. 2007. FENEME – Federação Nacional de Entidades de Oficiais Militares Estaduais. Disponível em: *http://www.feneme.org.br/pagina/810/analise-criminalnovas-tendencias-em-proveito-da-analise-criminal-avancada-e-da-inteligencia-de-seguranca-publica*. Acesso em dezembro, 2016.

DANTAS, G. F. L.; FERRO JR., C. M. *Descoberta e a Análise de Vínculos na Complexidade da Investigação Criminal Moderna*. 2006. Disponível em: *http://www.egov.ufsc.br/portal/sites/default/files/anexos/13124-13125-1-PB.pdf*. Acesso em dezembro, 2016.

FACCIONI FILHO, M. *Análise de Redes Sociais*. 2013. Disponível em: *http://www.open.edu/openlearnworks/course/view.php?id=1436*. Acesso em dezembro, 2016.

FERRO JR., C. M. *Inteligencia Organizacional, Análise de Vínculos e a Investigação Criminal: Um Estudo de Caso na Polícia Civil do Distrito Federal*. Mestrado em Gestão do Conhecimento e Tecnologia da Informação — Universidade Católica de Brasília, 2007.

FUNDAÇÃO GETÚLIO VARGAS. *27a Pesquisa anual de uso de TI*. 2016. Pesquisa anual realizada pelo Centro de Tecnologia de Informação Aplicada da FGV-EAESP, Disponível em: *http://eaesp.fgvsp.br/sites/eaesp.fgvsp.br/files/pesti2016gvciappt.pdf*. Acesso em março, 2016.

HARISSON, T. H. *Intranet data warehouse: Ferramentas e Técnicas para Utilização do Data Warehouse na Intranet*. São Paulo: Bekerley Brasil, 1998.

HOPCROFT, J. et al. Natural communities in large linked networks. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2003. (KDD '03), p. 541–546. ISBN 1-58113-737-0. Disponível em: <http://doi.acm.org/10.1145/956750.956816>.

LU, L.; ZHOU, T. Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and its Applications*, v. 390, n. 6, p. 1150 – 1170, 2011. ISSN 0378-4371. Disponível em: <http://www.sciencedirect.com/science/article/pii/S037843711000991X>.

MARTINS, I. T. *Descoberta de conhecimento em históricos criminais: algoritmos e sistemas*. Doutorado em Engenharia Industrial — Pontifíia Universidade Católica do Rio de Janeiro, 2009.

MAYER-SCHONBERGER, V.; CUKIER, K. *Big Data: Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana*. 1. ed. Rio de Janeiro: Elsevier, 2013. ISBN 8535270907.

MOMPEAN, A. Transações com mobile banking crescem 138% em um ano. *Revista Ciab FEBRABAN*, p. 16–23, 2016. Disponível em: *https://issuu.com/revistaciab/docs/revista_ciab_63_jun16/1*. Acesso em março, 2016.

PEOTTA, L. et al. *Análise de Malware: Investigação de Códigos Maliciosos Através de uma Abordagem Prática*. 2011. Disponível em: *http://www.peotta.com/sbseg2011/resources/downloads/minicursos/90650.pdf*. Acesso em dezembro, 2016.

POZZER, C. T. *Teoria dos Grafos*. 2010. Disponível em: *http://www-usr.inf.ufsm.br/∼pozzer/disciplinas/ed_7_grafos.pdf*. Acesso em dezembro, 2016.

SCHAEFFER, S. E. Graph clustering. *Computer Science Review*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, v. 1, n. 1, p. 27–64, 2007. ISSN 1574-0137. Disponível em: <http://dx.doi.org/10.1016/j.cosrev.2007.05.001>.

SIQUEIRA, E. P. *O Projeto Tentáculos da Polícia Federal: Da concepção à Proposta de Modelo Aplicável na Segurança Pública Brasileira*. Monografia de Especialização em Gestão da Segurança da Informação e Comunicações — Universidade de Brasília, 2014.

SIQUEIRA, E. P. et al. *Base Nacional de Fraudes Bancáias Eletrônicas – BNFBE – Projeto Tentáculos – Guia de Acesso à Base*. 2015. Polícia Federal. GPA/SRCC/DPF.

SZWARCFITER, J. *Grafos e algoritmos computacionais*. Rio de Janeiro: Campus, 1986. ISBN 8570013418.

WEST, D. B. *Introduction to Graph Theory*. 2. ed. Upper Saddle River: Prentice Hall, 2000. ISBN 0130144002.

XU, J. J.; CHEN, H. Fighting organized crimes: Using shortest-path algorithms to identify associations in criminal networks. *Decis. Support Syst.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, v. 38, n. 3, p. 473–487, dez. 2004. ISSN 0167-9236. Disponível em: <http://dx.doi.org/10.1016/S0167-9236(03)00117-9>.

XU, R.; WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, IEEE Press, Piscataway, NJ, USA, v. 16, n. 3, p. 645–678, May 2005. ISSN 1045-9227. Disponível em: <http://dx.doi.org/10.1109/TNN.2005.845141>.

ZHANG, X. X.; YANG, Y. M. Minimum spanning tree and color image segmentation. In: *2008 IEEE International Conference on Networking, Sensing and Control*. Hainan, China: [s.n.], 2008. p. 900–904. Disponível em: <http://doi.org/10.1109/ICNSC.2008.4525344>.