



Congresso Serpro de
Tecnologia e Gestão
aplicadas a Serviços Públicos

conserpro

Análise de Redes Sociais para Inferência de Similaridade entre Fornecedores de Governo



16 a 19 de novembro de 2015



Autores do Trabalho



Autor 1 – Marcelo Rodrigo de Souza Pita (DP/CEIIG/IGBHE)

Autor 2 – Gustavo da Gama Torres (DP/CEIIG/IGBHE)

Autor 3 – Sérgio Mariano Dias (DP/CEIIG/IGBHE)



Agenda

1. **Introdução**
2. **Modelo de redes de fornecedores de Governo**
3. **Experimentos e resultados**
4. **Conclusão e trabalhos futuros**

Introdução

Motivações



conserpro

Sociedade está desenvolvendo uma nova percepção sobre a **transparência como requisito** para atuar na esfera pública.



Motivações

A ideia de que os processos de aquisição em Governo estão dominados por **associações ilegais entre fornecedores** é praticamente um senso comum em nossa sociedade.

THE WALL STREET JOURNAL.

Home **World** U.S. Politics Economy Business Tech Markets Opinion Arts Life



Russia Engaged in Sports Doping, Commission Finds



Jordan Policeman Kills Two U.S. Trainers



Russia Moving Ahead With Iran Missile Deal



S Claim Myanmar

MAKING NEWS

University of Missouri System President Tim Wolfe resigns over handling of racial issues

WORLD | LATIN AMERICA

Brazil Probe Sweeps Up Corporate Magnates

Related to Petrobras construction contract allegations probe

Compras do Governo Federal



Aquisição de produtos e serviços pelo Governo que envolve repasses financeiros para fornecedores (pessoas jurídicas e físicas).

As compras de Governo são reguladas pela lei nº 8666/1993, que estabelece normas para a realização de licitações e contratos.

Estamos interessados em investigar como fornecedores de Governo podem estar associados e se há evidências de associação ilícita.



Como conduzir essa investigação?



conserpro

1) Acesso aos dados históricos de compras de Governo e cadastro de empresas impedidas.

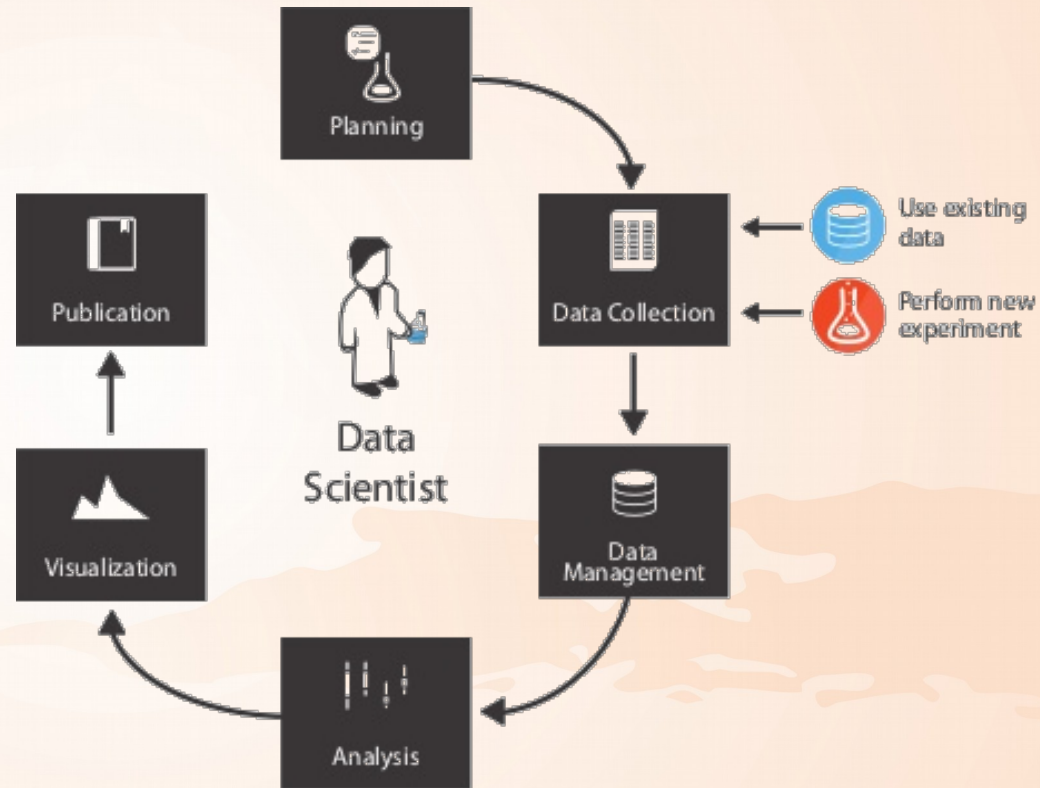
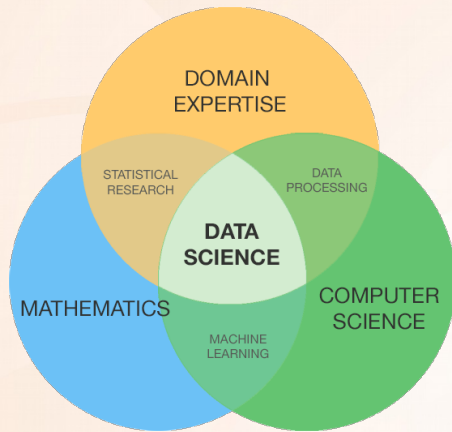


Como conduzir essa investigação?



conserpro

2) Aplicação de métodos de ciência de dados



Perguntas interessantes...



conserpro

Existem associações entre fornecedores?

Fornecedores agem sozinhos?

Existem padrões de vendas que ajudem a discriminar fornecedores idôneos de inidôneos?

Existe alguma tipo de hierarquia entre os fornecedores?



Modelos de Redes de Fornecedores de Governo

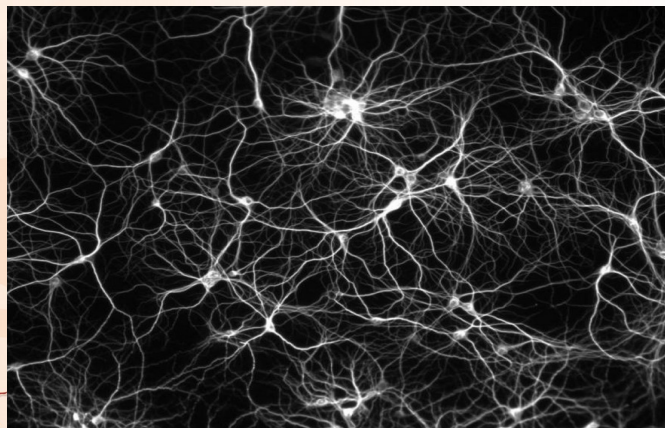
[Modelos de] redes complexas



conserpro

Redes complexas são **sistemas complexos** (grande número de componentes interdependentes, comportamento emergente, hierarquia entrelaçada, imprevisibilidade).

Redes complexas são normalmente representadas por **grafos**.

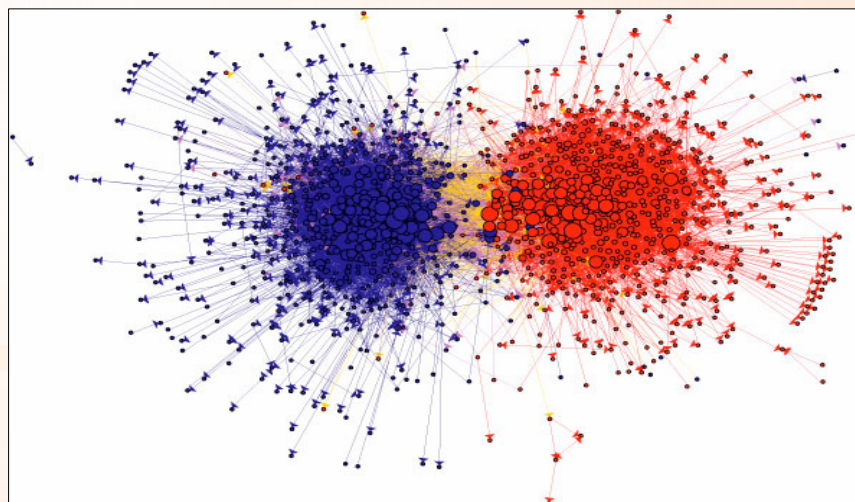
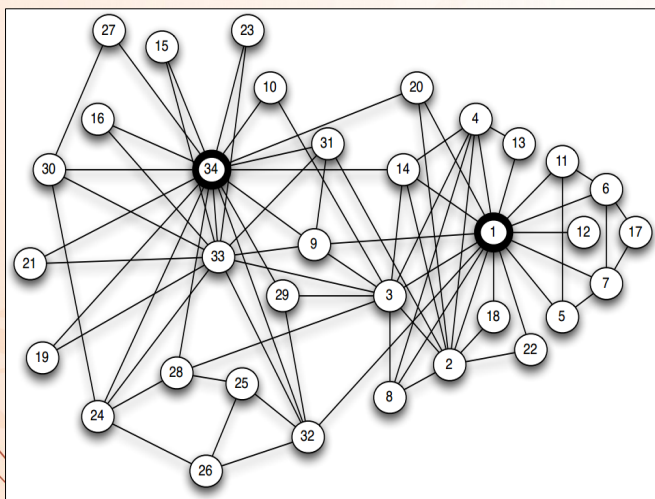


[Análise de] redes sociais



conserpro

Análise de redes sociais, ou *social network analysis* (SNA), concentra teorias e métodos para análise de redes complexas que representam interações sociais.



Dados do Portal da Transparência



Despesas - Gastos Diretos

Recursos gastos diretamente pelo Governo Federal

- [Pagamentos](#)
- [Diárias de viagens](#)
- [Cartão de Pagamento do Governo Federal \(CPGF\)](#)
- [Lista de Favorecidos – Pessoa Jurídica](#)

Despesas - Transferências

Recursos repassados a estados, municípios e entidades privadas

- [Pagamentos](#)
- [Cartão de Pagamento da Defesa Civil \(CPDC\)](#)
- [Lista de Favorecidos – Pessoa Jurídica](#)

Programas Sociais

Recursos repassados diretamente a cidadãos

- [Bolsa Família - Pagamentos](#)
- [Bolsa Família - Saques](#)
- [Pescador Artesanal](#)
- [Outros programas sociais](#)

Receitas

- [Receita prevista, realizada e lançada](#)

Convênios

Convênios firmados com o Governo Federal

- [Convênios](#)

Servidores

- [Servidores civis e militares do Executivo Federal](#)
- [Cadastro de Expulsões da Administração Federal \(CEAF\)](#)

Ceís

Cadastro de Empresas Inidôneas e Suspensas

- [Empresas Inidôneas e Suspensas](#)

Cepim

Cadastro de Entidades sem Fins Lucrativos Impedidas

- [Entidades sem Fins Lucrativos Impedidas](#)

Imóveis Funcionais

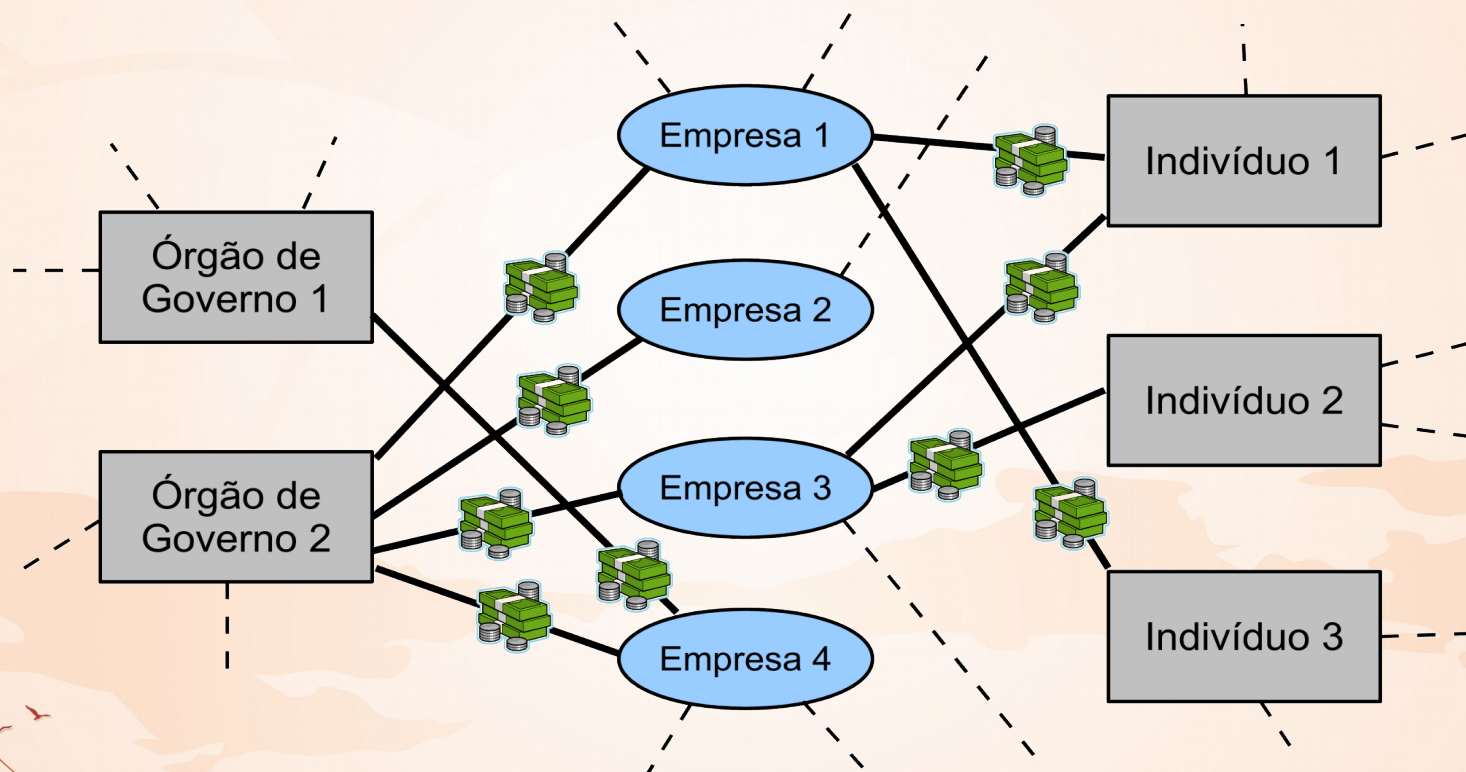
- [Imóveis Funcionais](#)

Redes de fornecedores de Governo



conserpro

Passo 1: Criação da rede de afiliações de fornecedores



Passo 2: Criação da rede de similaridades de fornecedores

Toma-se como ponto de partida a hipótese de que o grau de similaridade entre dois fornecedores é proporcional ao número de afiliações compartilhadas (órgãos e indivíduos) ponderada pela similaridade dos totais de valores transferidos.

Passo 2: Criação da rede de similaridades de fornecedores

$$J(a, b) = \frac{|N_a \cap N_b|}{|N_a \cup N_b|} \quad (1)$$

$$S(a, b) = \frac{\sum_{i \in N_a \cap N_b} s((a, i), (b, i))}{|N_a \cup N_b|} \quad (2)$$

Onde:

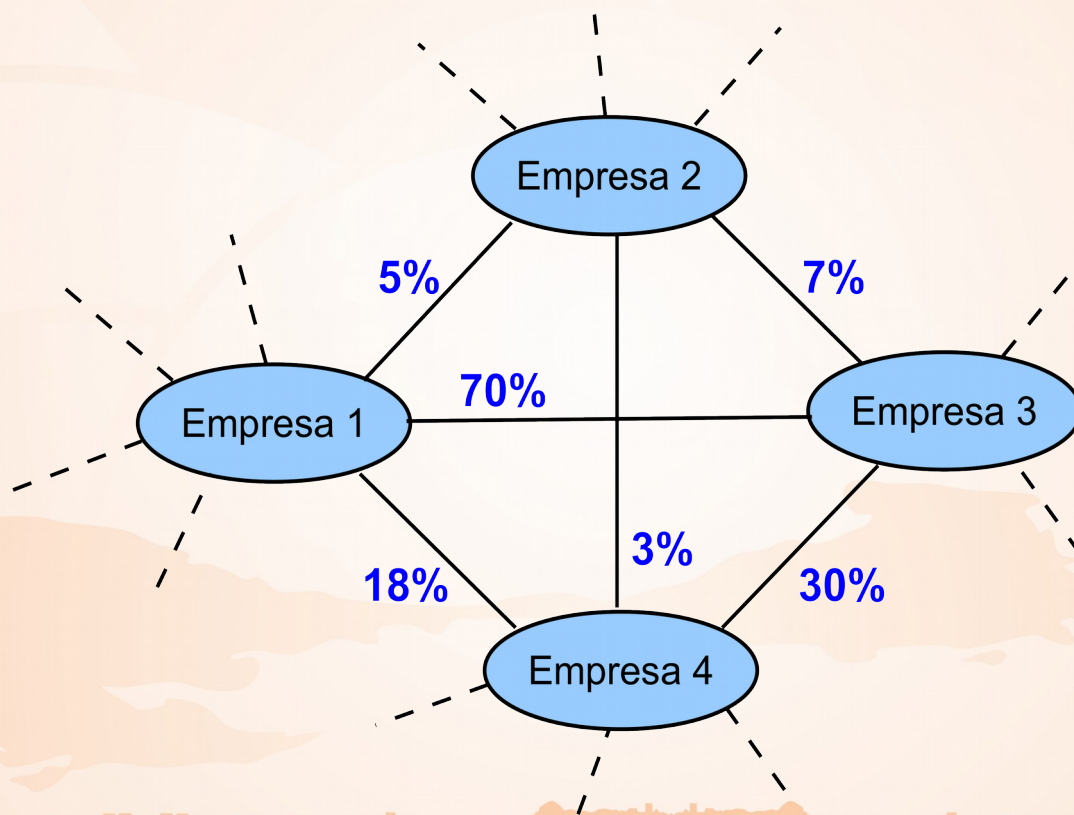
$$s((a, c), (b, c)) = 1 - \frac{|w_{(a,c)} - w_{(b,c)}|}{\max(w_{(a,c)}, w_{(b,c)})}$$

Redes de fornecedores de Governo



conserpro

Passo 2: Criação da rede de similaridades de fornecedores

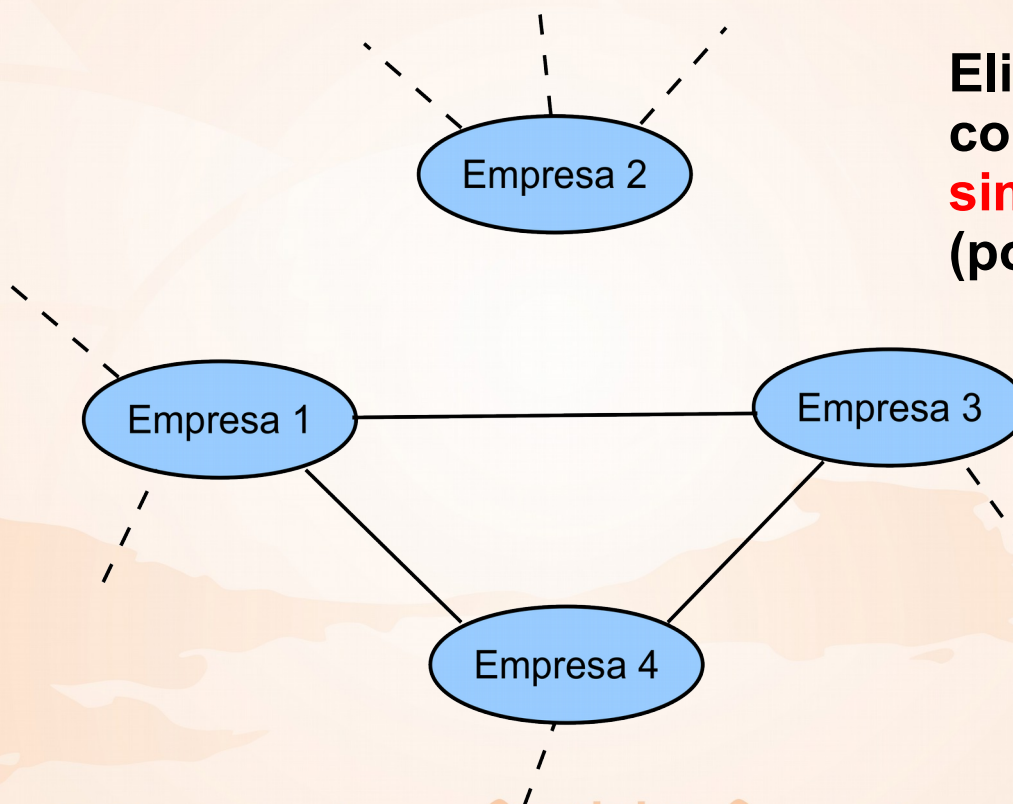


Redes de fornecedores de Governo



conserpro

Passo 2: Criação da rede de similaridades de fornecedores



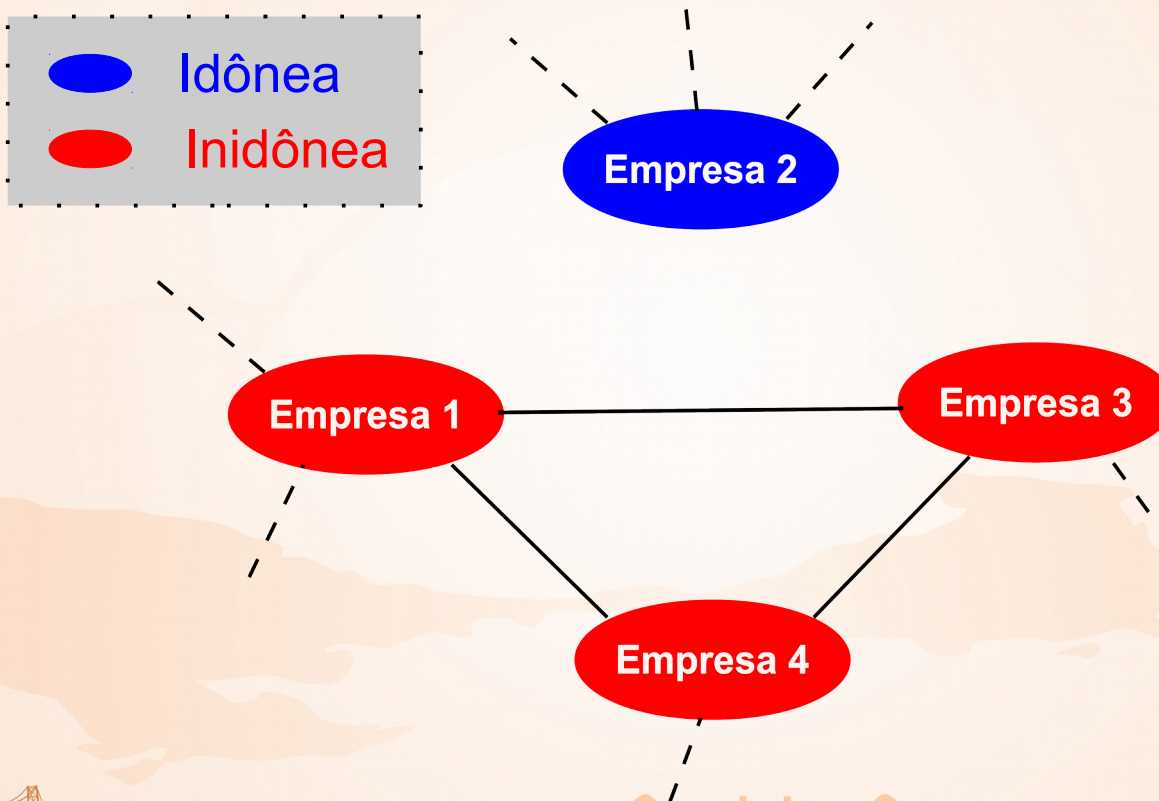
Eliminação de conexões com similaridade < 10% (por exemplo)

Redes de fornecedores de Governo



conserpro

Passo 3: Rotulação de fornecedores idôneos e inidôneos

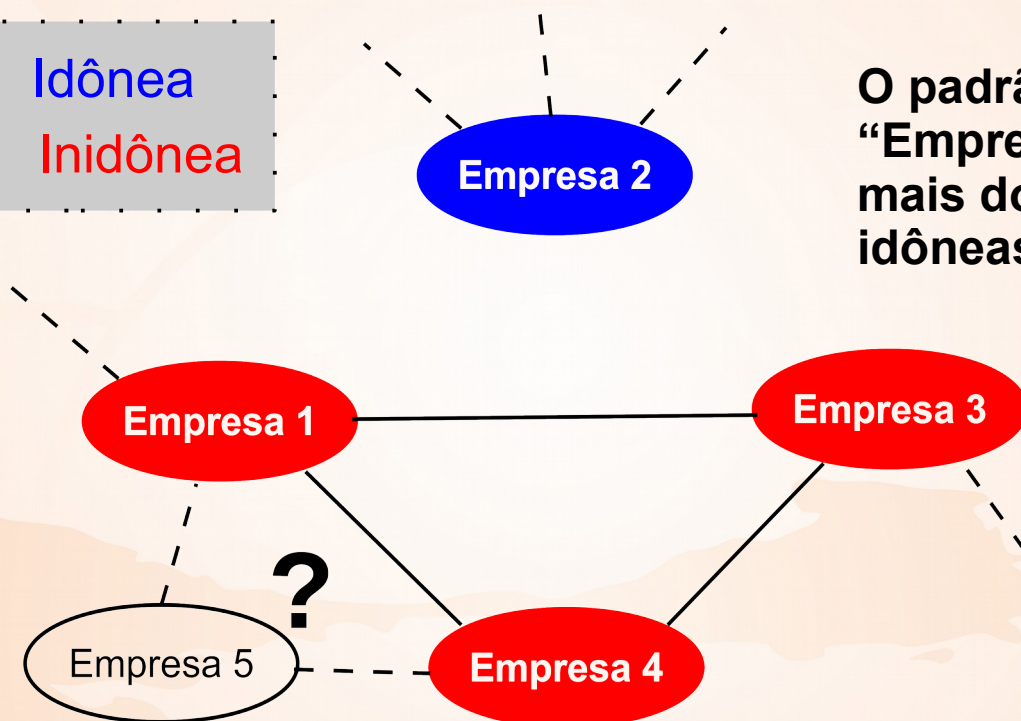
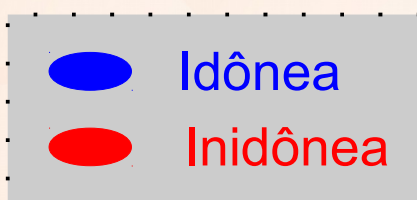


Redes de fornecedores de Governo



conserpro

Passo 3: Rotulação de fornecedores idôneos e inidôneos



O padrão de vendas da “Empresa 5” se aproxima mais do de empresas idôneas ou inidôneas?

Experimentos e Resultados



Configuração experimental



conserpro

Hipótese: Clusters formados no grafo são discriminantes em relação a idoneidade das empresas.

Foram realizados dois grupos de experimentos.

Informações sobre o primeiro grupo de experimentos:

- Pagamentos diretos coletados de Jan/2011 a Mar/2014
- CPGF de Jan/2010 a Mar/2014
- CPDC de Nov/2011 a Mar/2014
- Ceis de 09/05/2015
- Cepim de 09/05/2015



Configuração experimental

Informações sobre o segundo grupo de experimentos:

- Pagamentos diretos coletados de Jan/2011 a Mar/2014
- Ceis de 30/07/2015
- Cepim de 30/07/2015
- Balanceamento de classes de idôneos e inidôneos
- Uso do *betweenness* para eliminação de arestas (ideal para detecção de comunidades)

Hardware: Dell R710, 24 cores, 24GB de RAM (LCAD – Belo Horizonte)

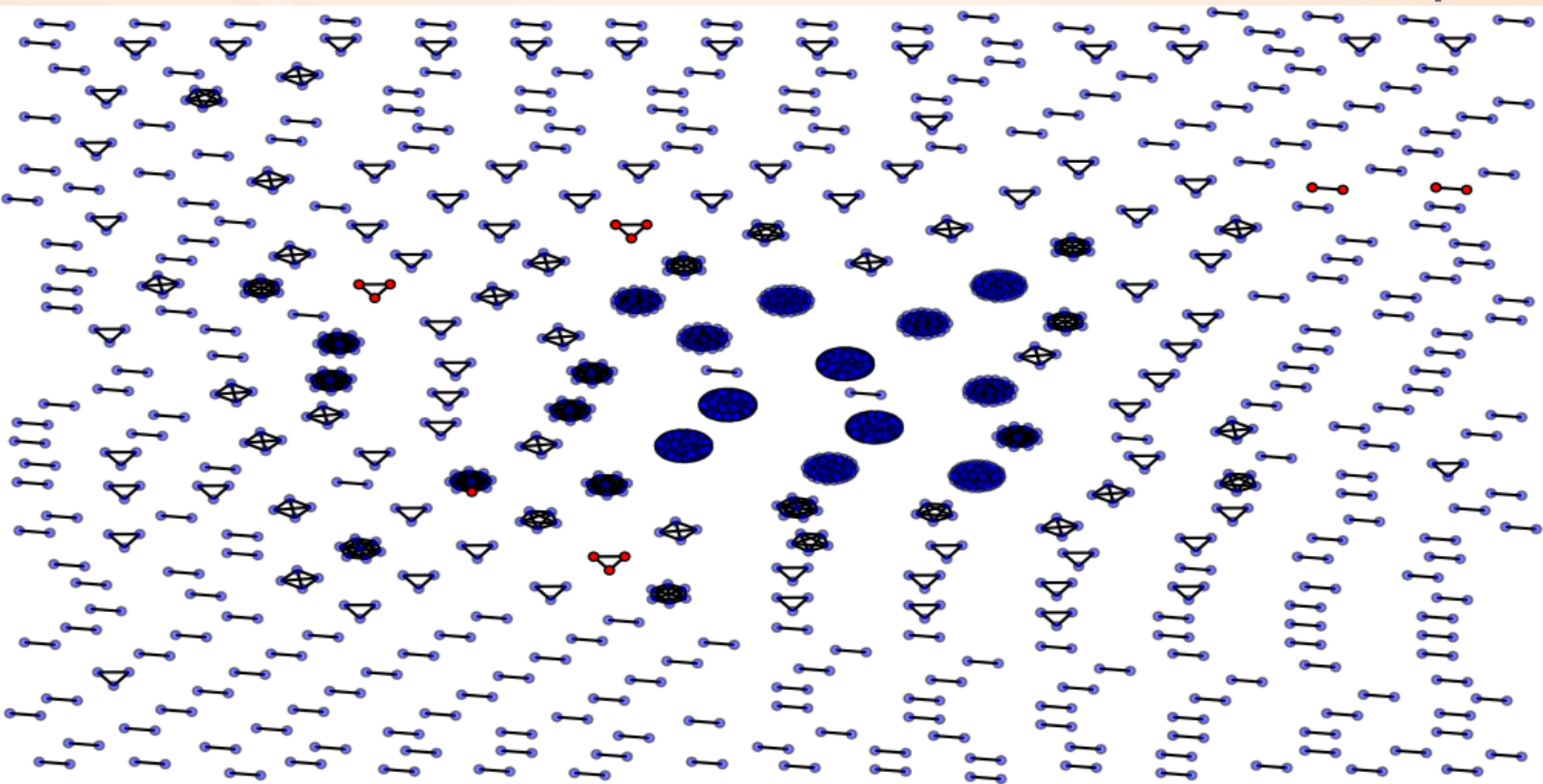
Software: R, Python (lib Networkx)

Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **100%**

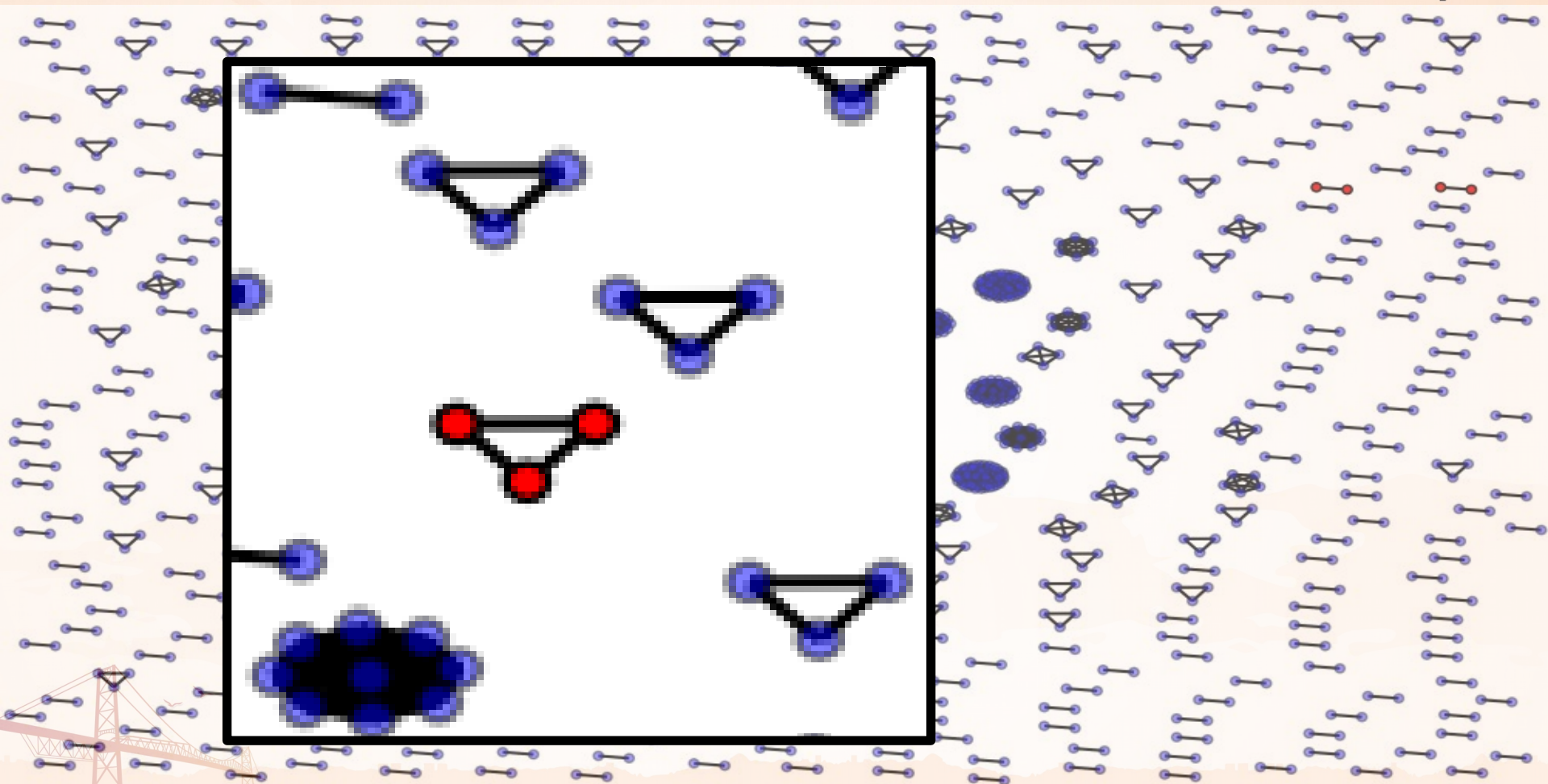


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **100%**

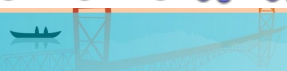
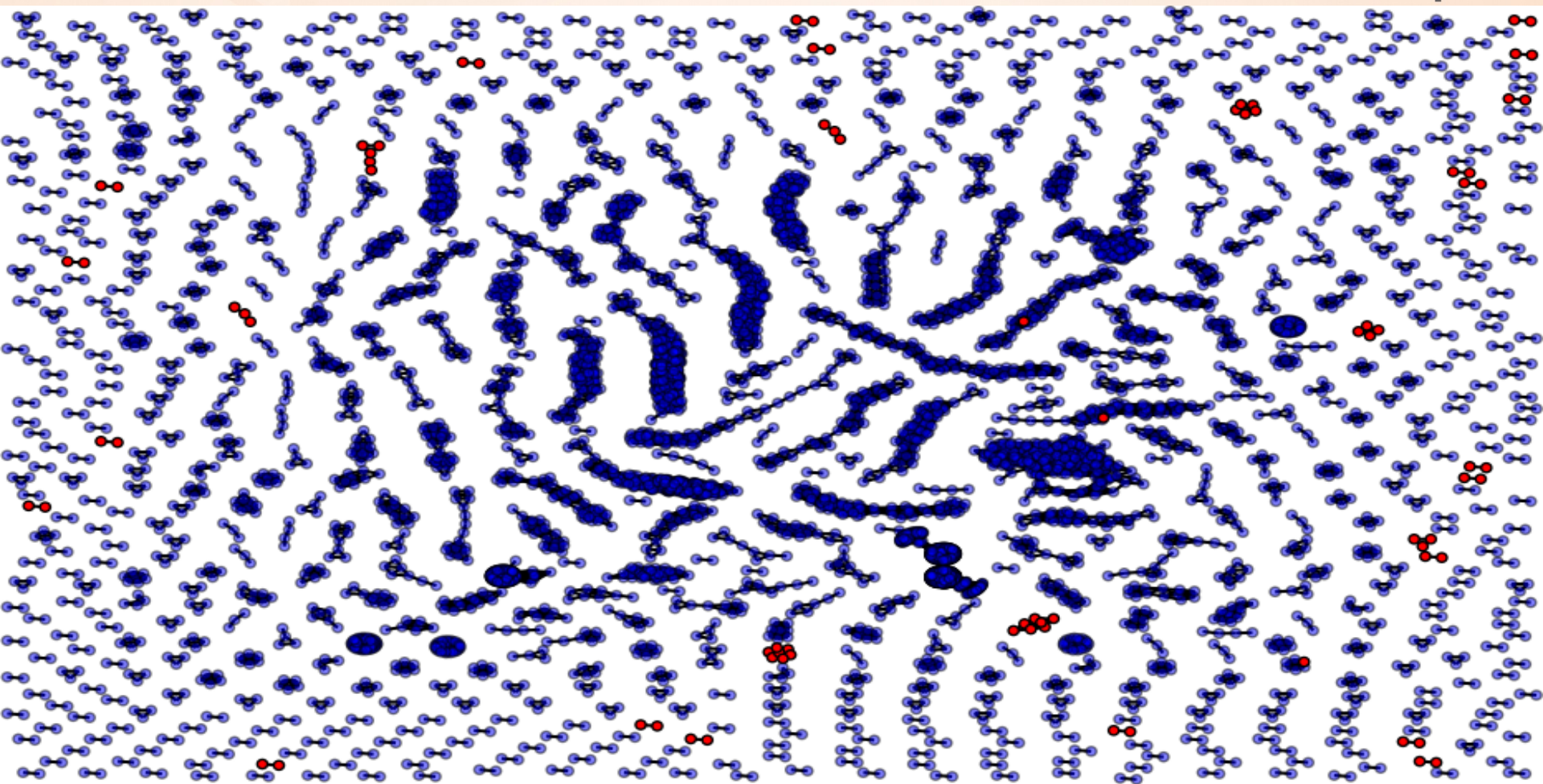


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de 75%

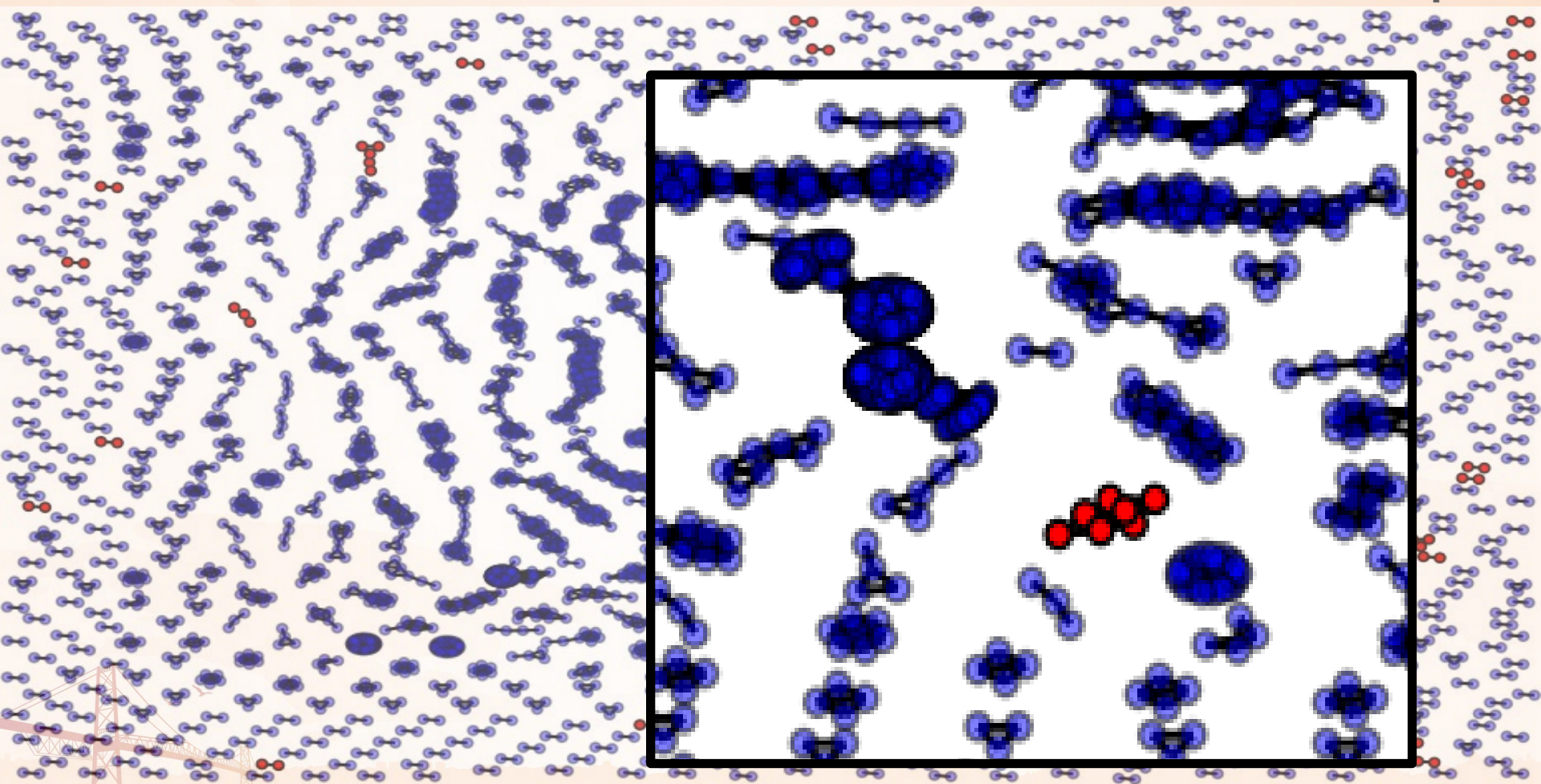


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **75%**

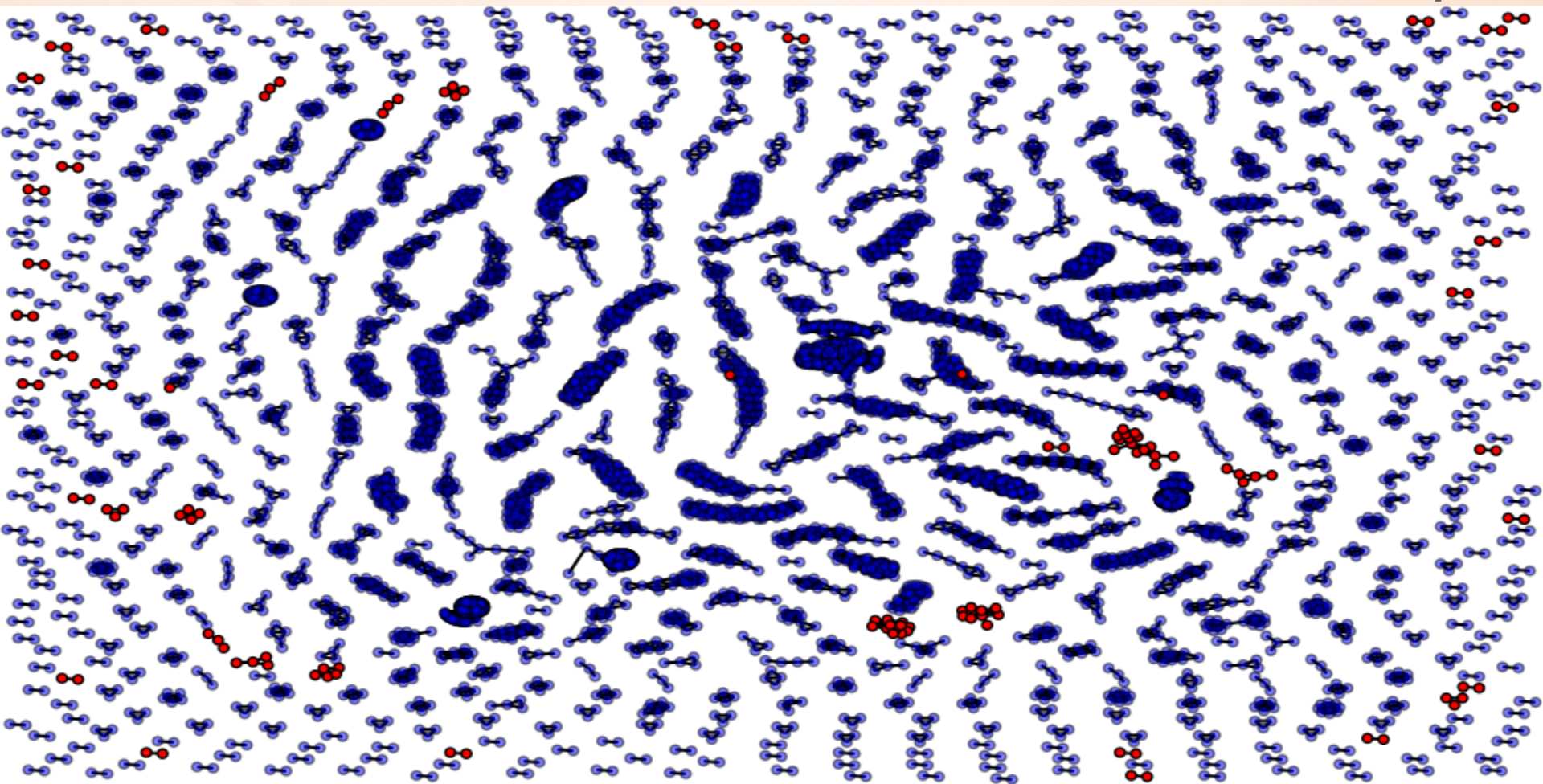


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **50%**

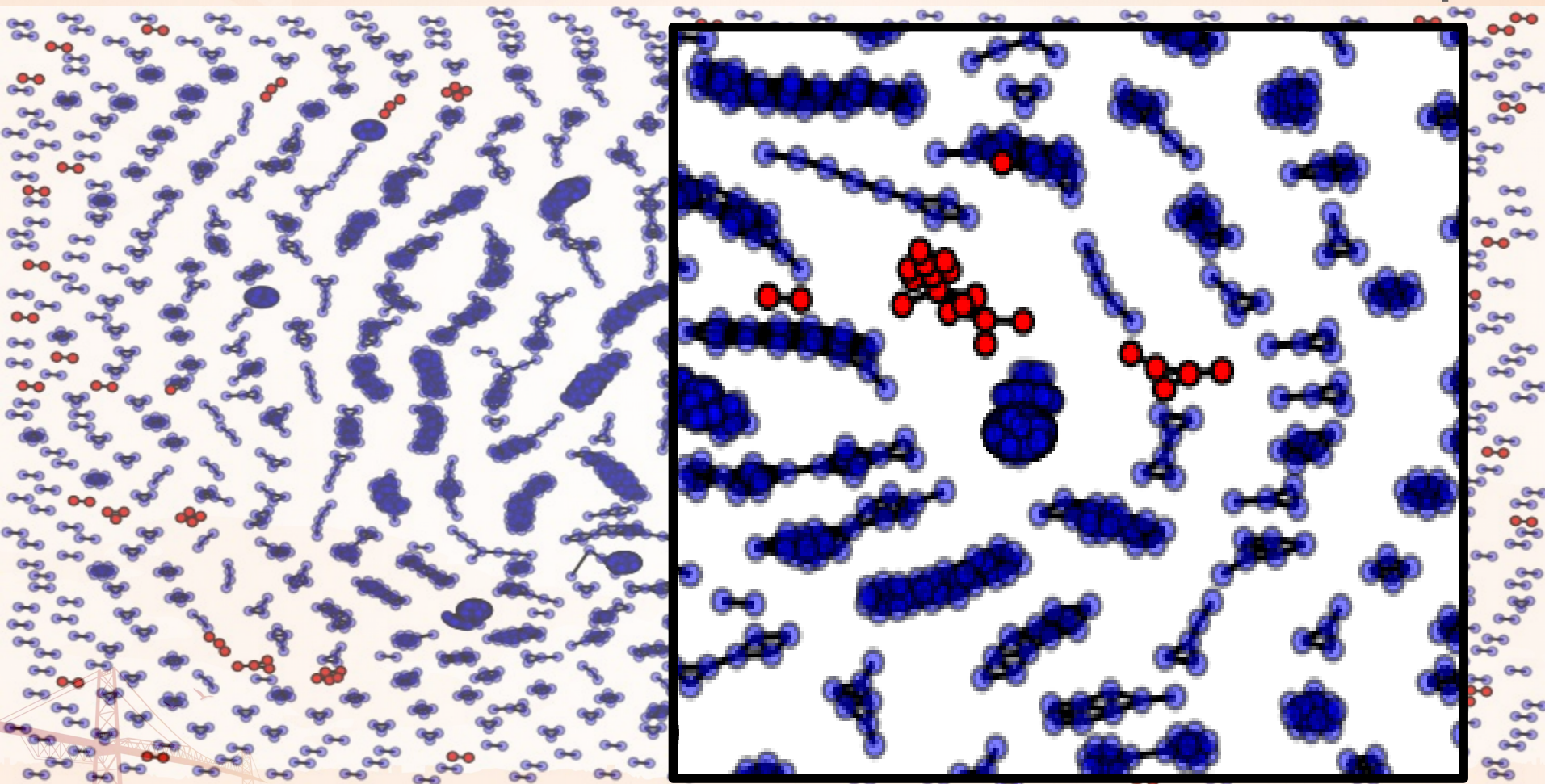


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **50%**

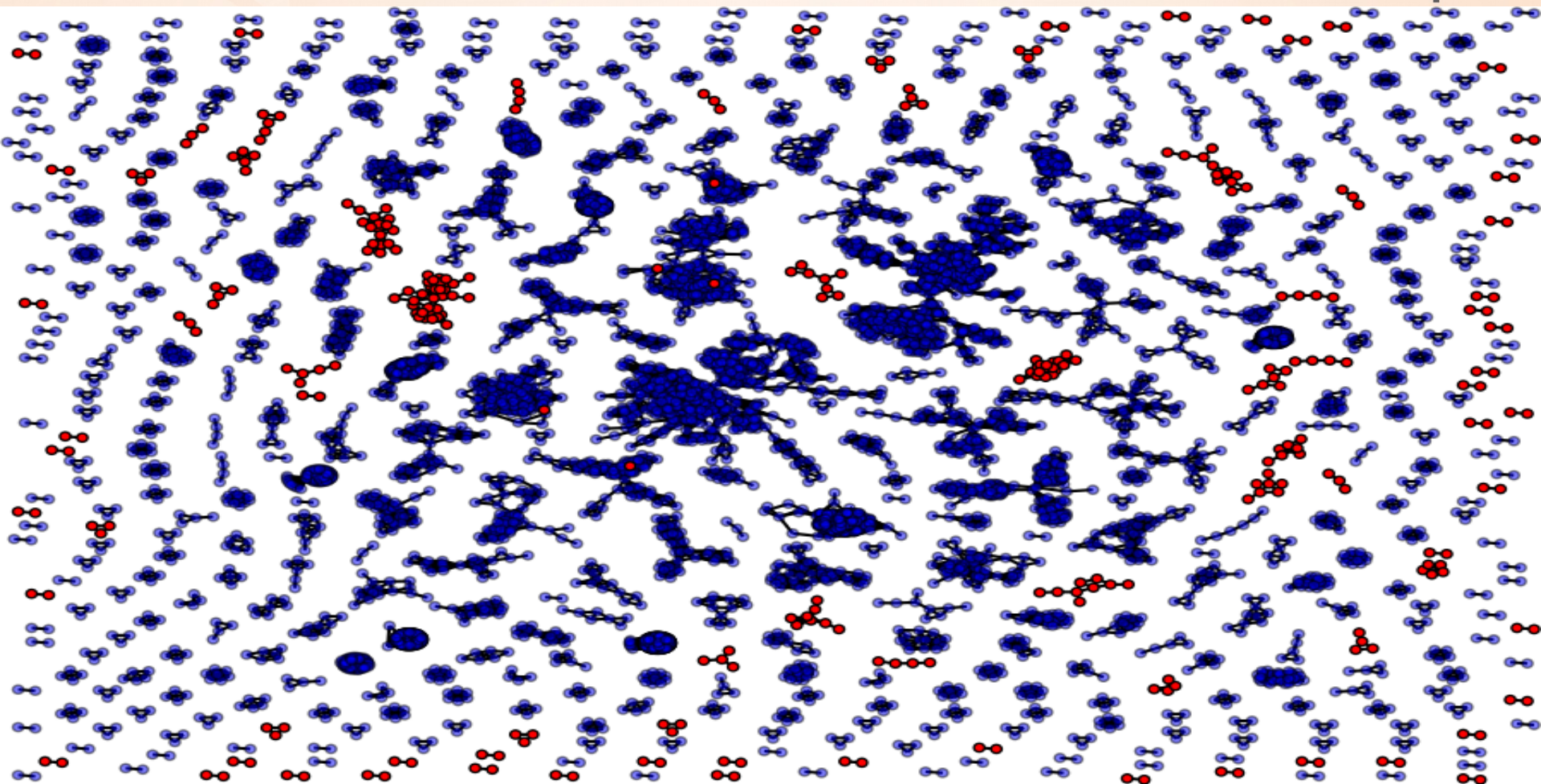


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **25%**

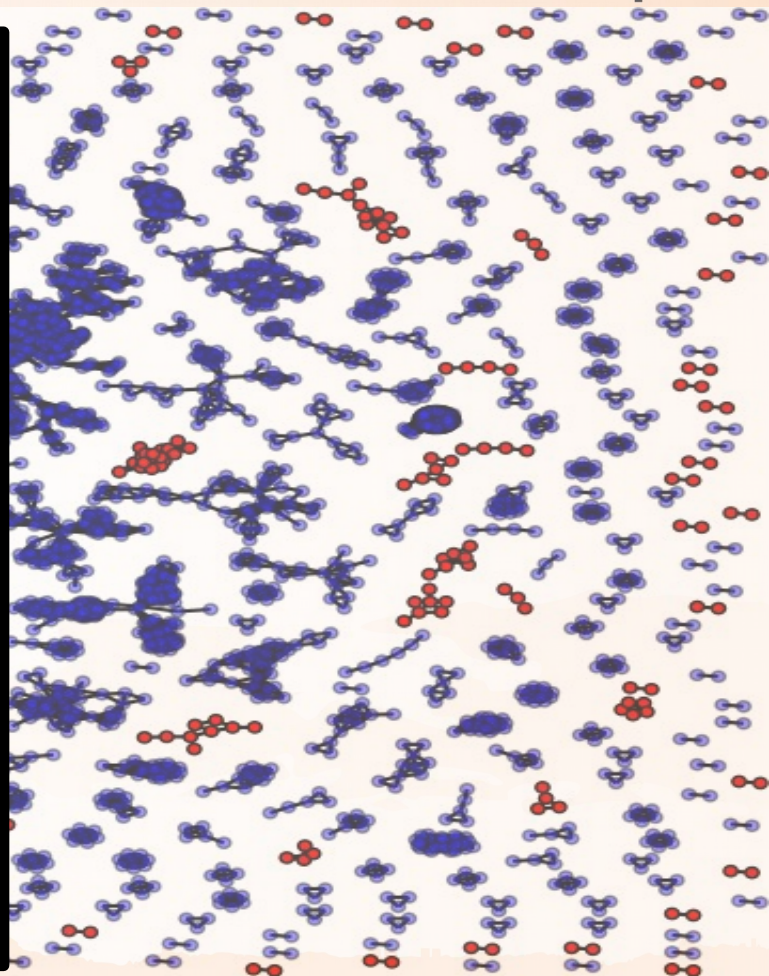
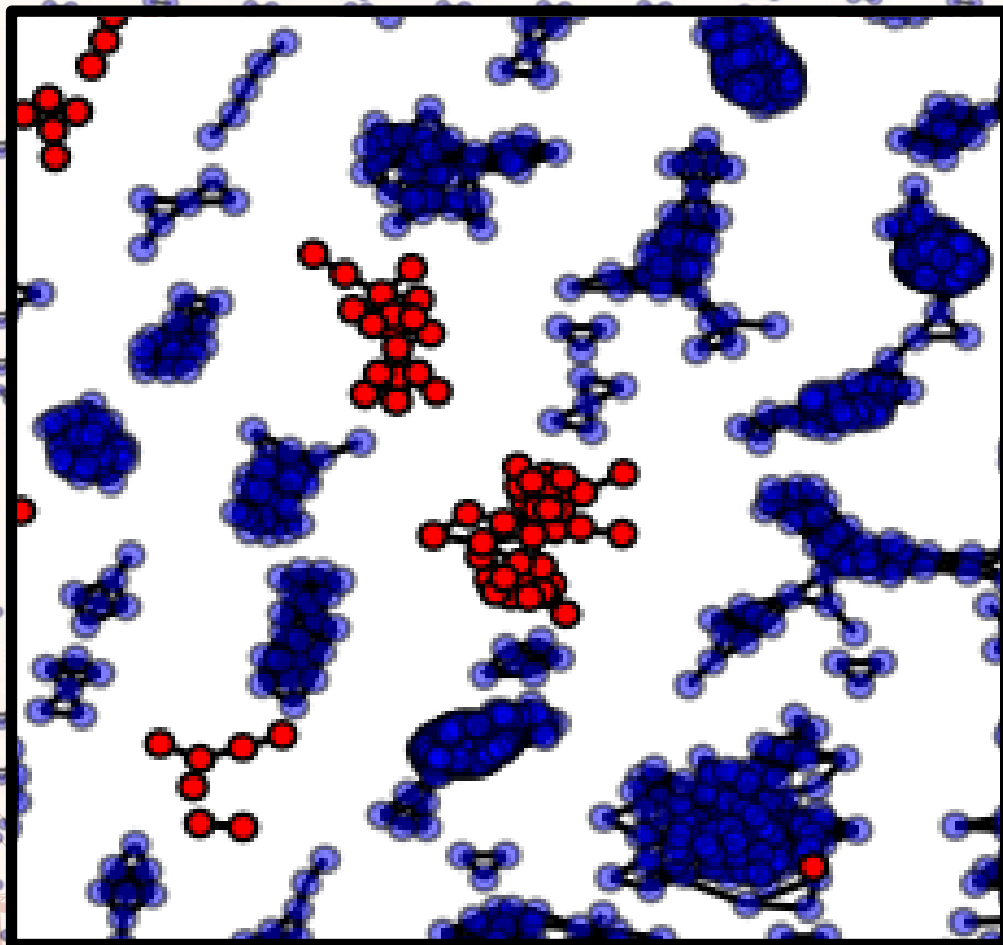


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **25%**

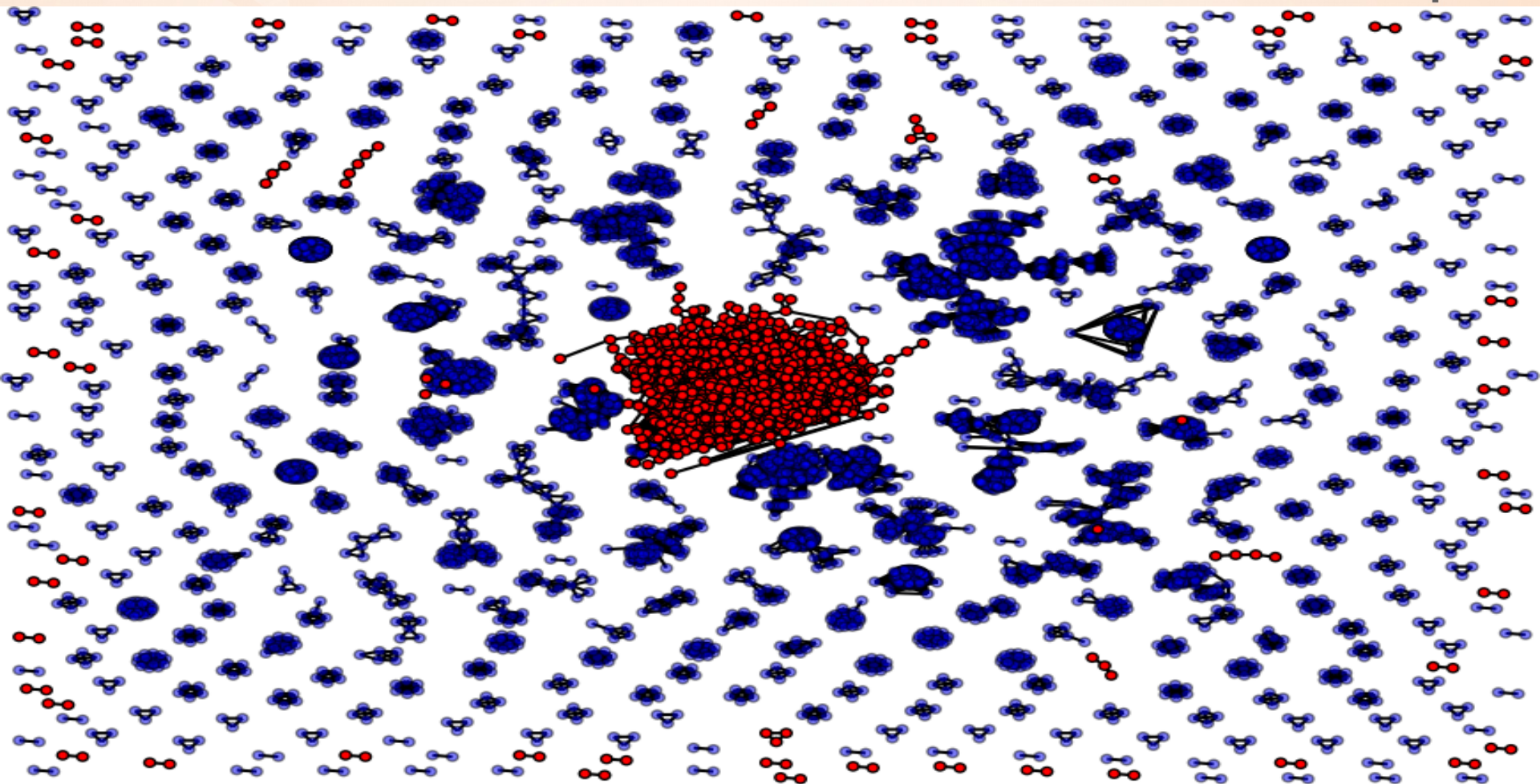


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de 5%

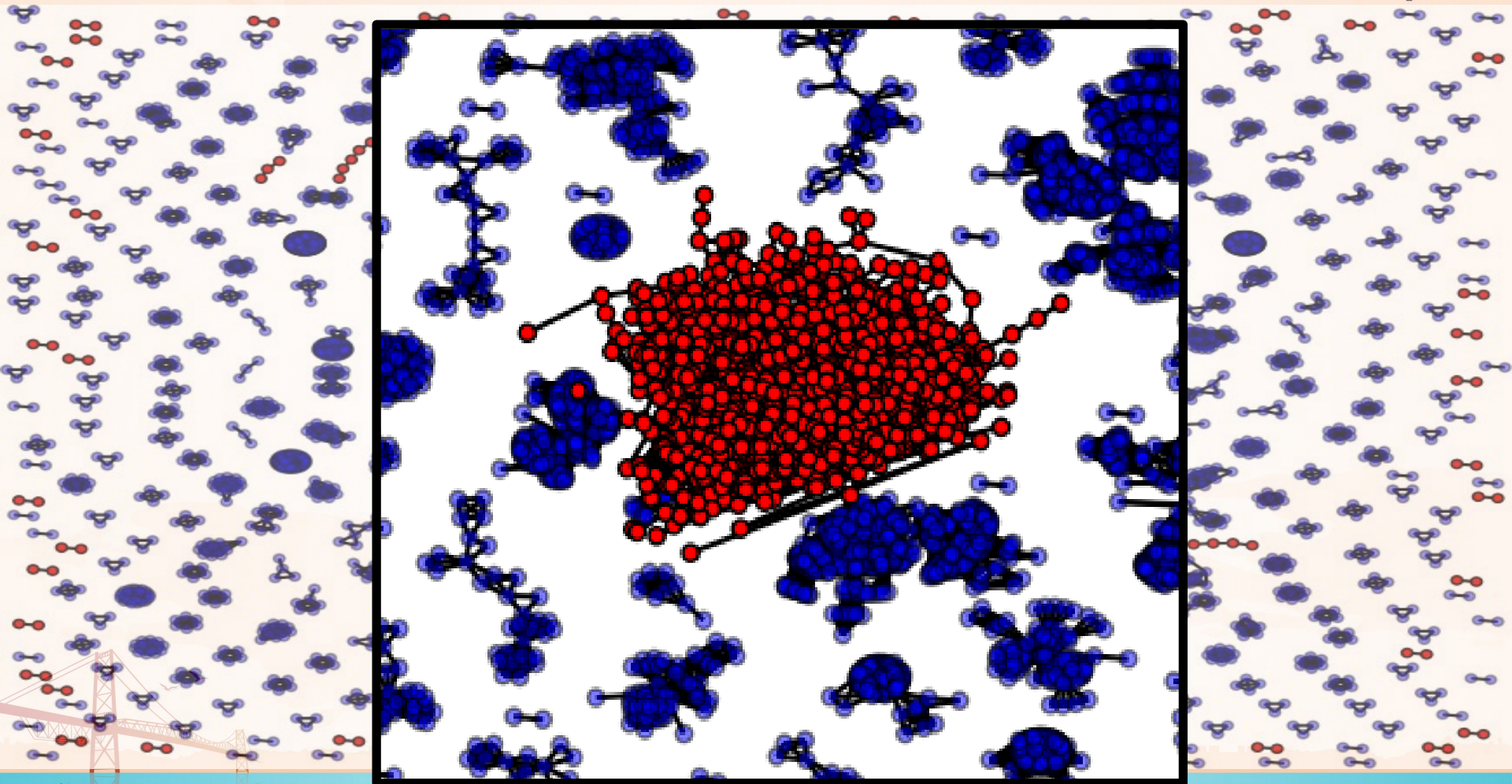


Resultados do 1º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de 5%



Resultados do 1º grupo



conserpro

Desempenho de classificação

$$\pi^i = \frac{\max(|C_{id}^i|, |C_{inid}^i|)}{|C^i|}$$

$$\Pi = \frac{\sum_{i=1}^{|C|} |C^i| \pi^i}{\sum_{i=1}^{|C|} |C^i|}$$

	$\sigma = 0,25$	$\sigma = 0,5$	$\sigma = 0,75$	$\sigma = 1,0$
Π	99,932%	99,939%	99,948%	99,954%

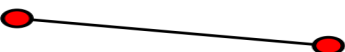
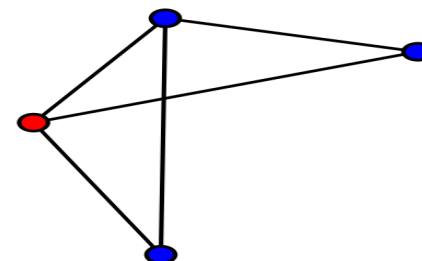
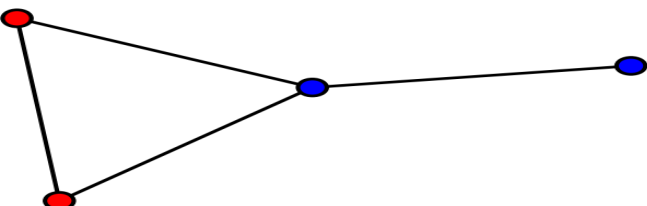
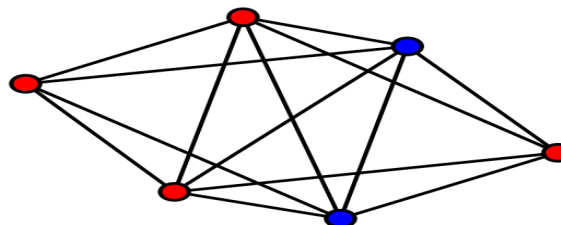
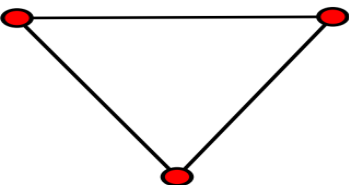


Resultados do 2º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **100%**

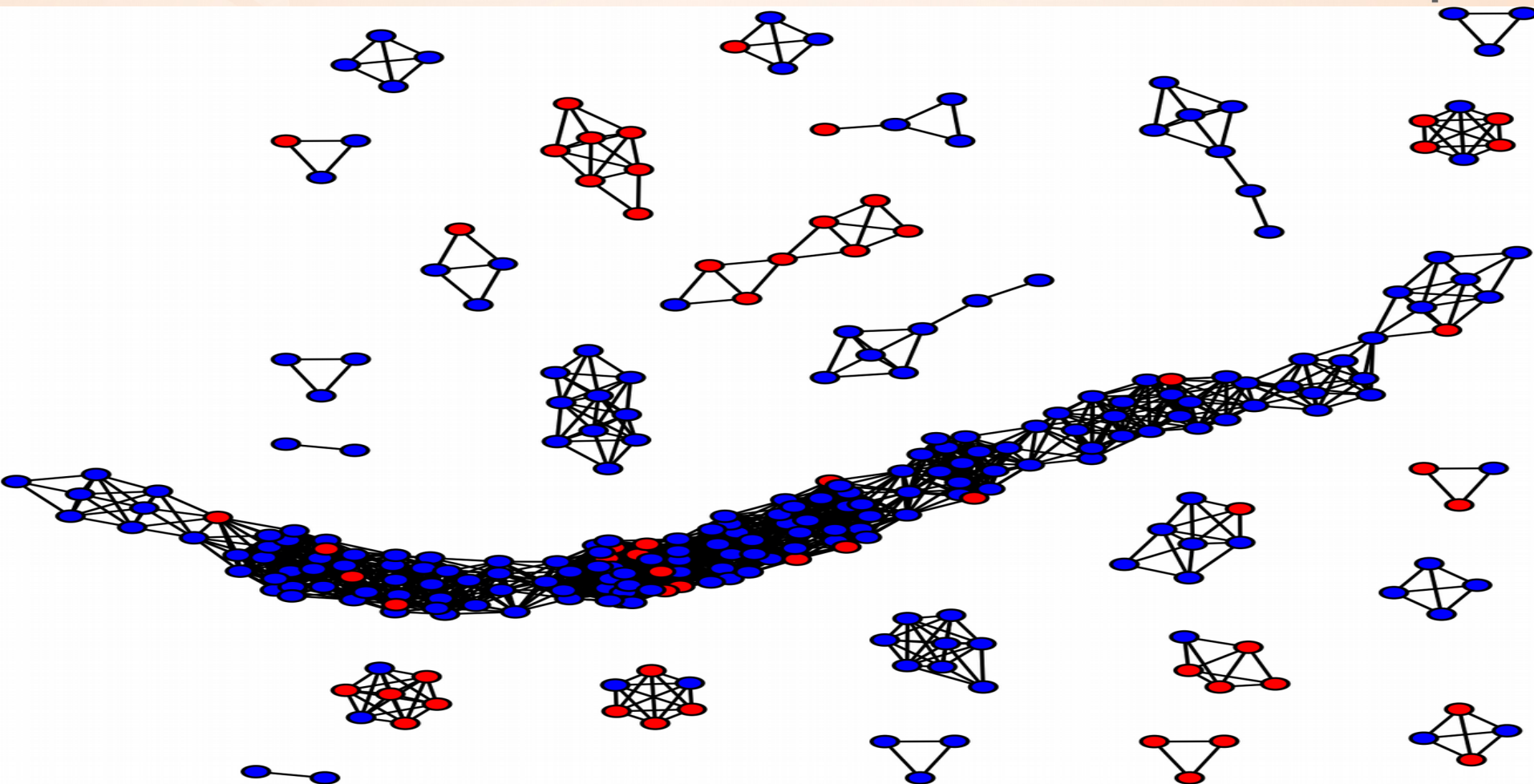


Resultados do 2º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **75%**

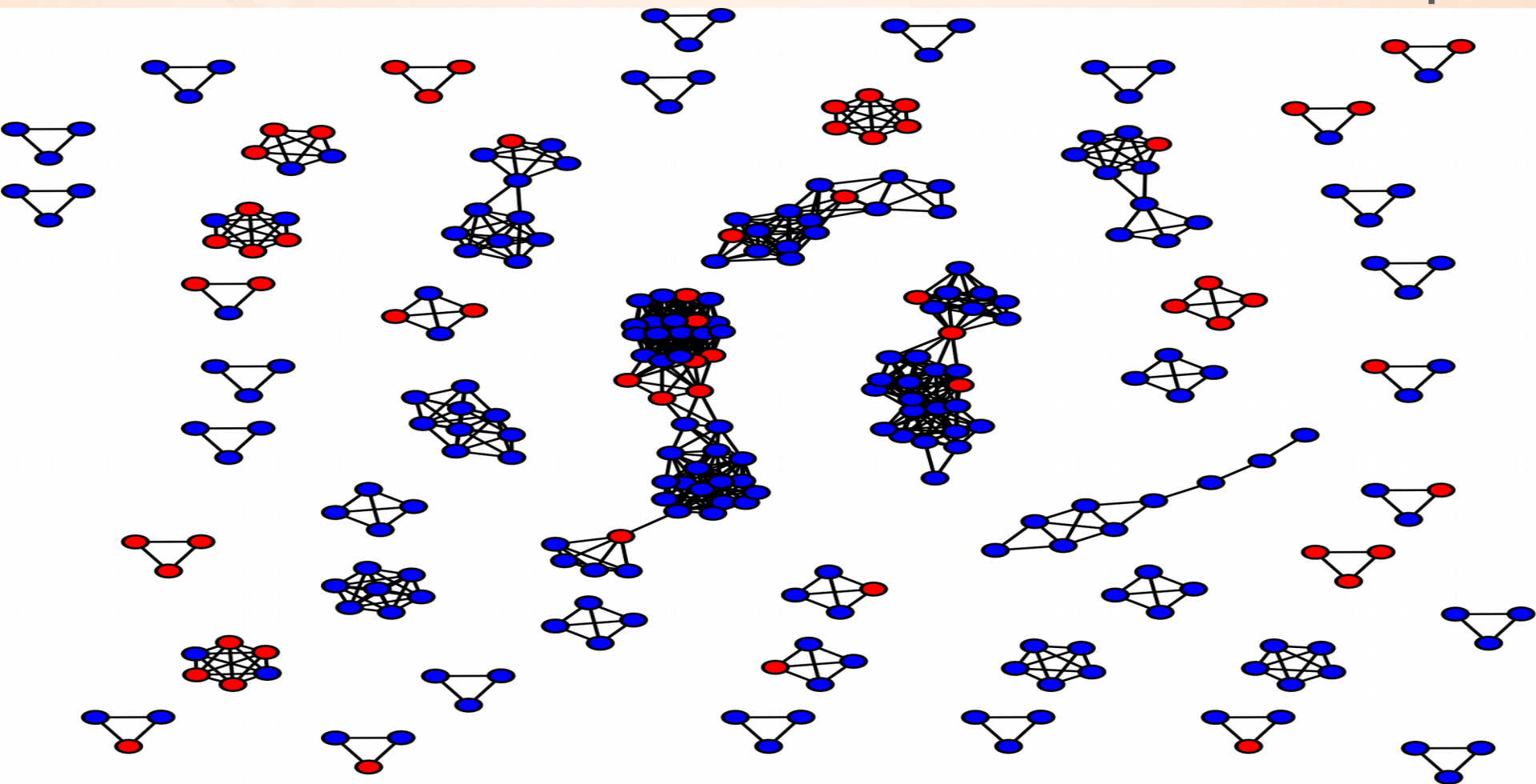


Resultados do 2º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **50%**

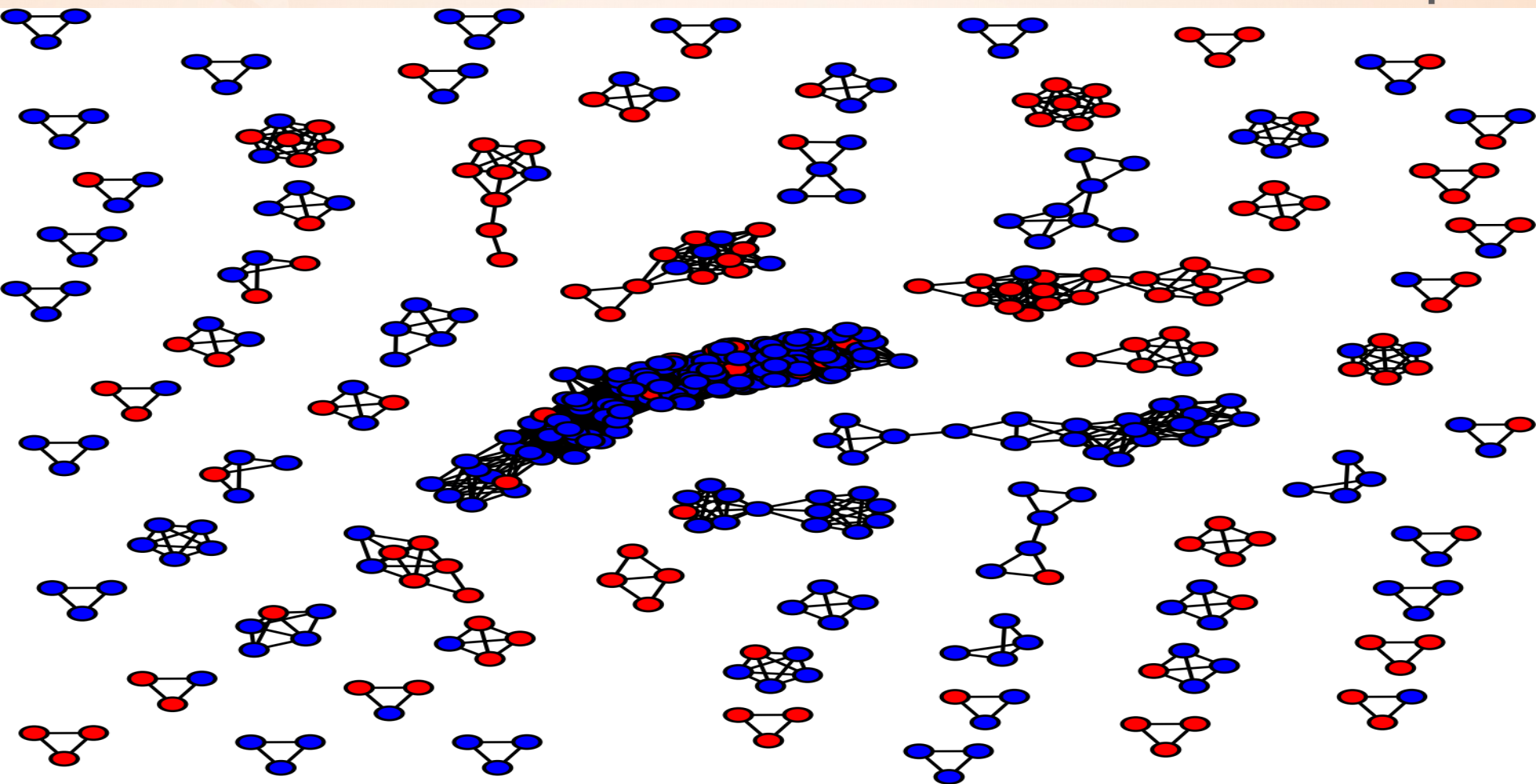


Resultados do 2º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **25%**

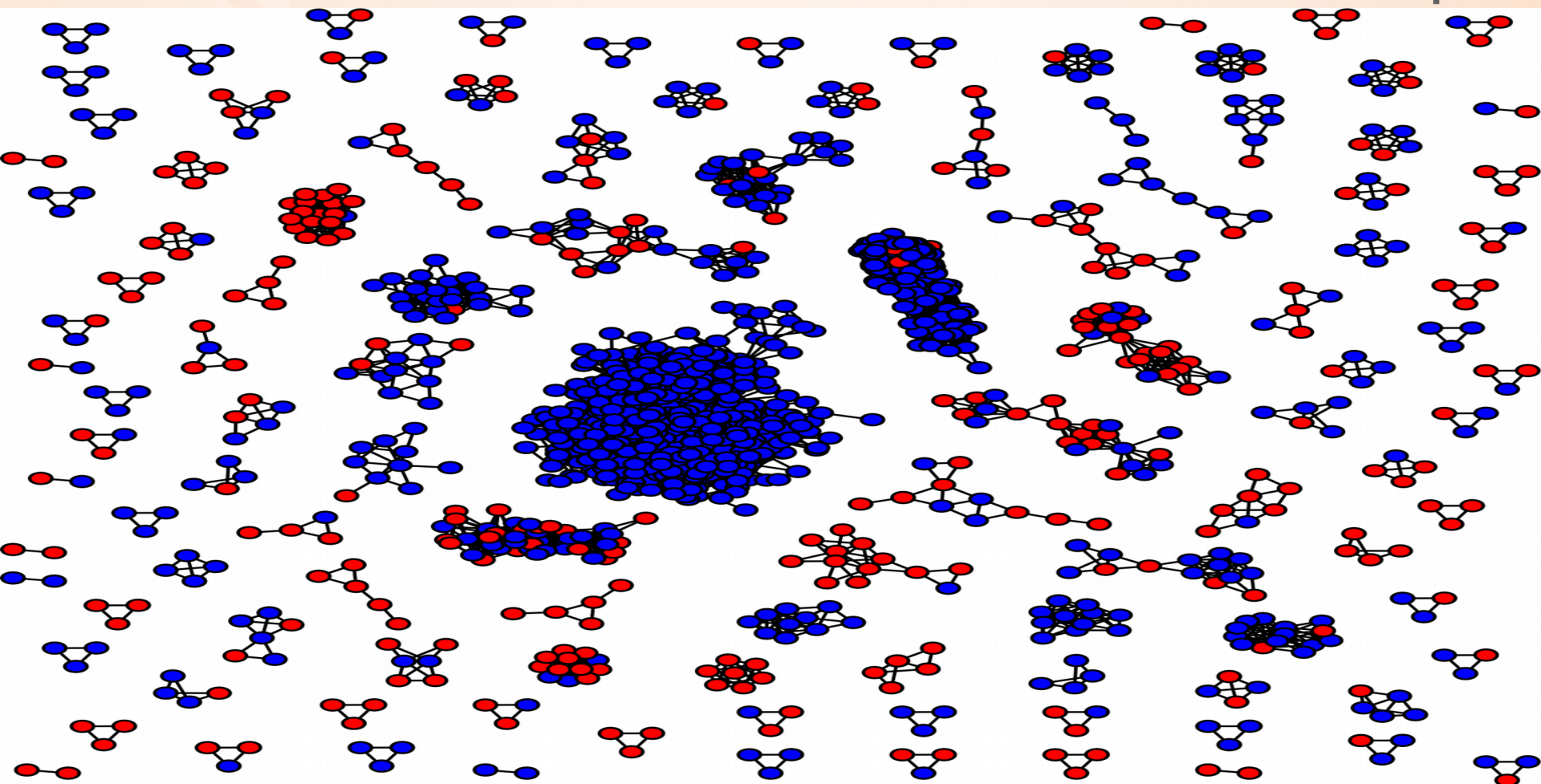


Resultados do 2º grupo



conserpro

Empresas **idôneas** e **inidôneas** para similaridade mínima de **10%**



Resultados do 2º grupo

Desempenho de classificação

	$\sigma = 0,1$	$\sigma = 0,25$	$\sigma = 0,5$	$\sigma = 0,75$	$\sigma = 1,0$
Π	89,06%	86,78%	88,26%	88,62%	75%

Conclusão e Trabalhos Futuros



Conclusão

- Os resultados apresentados, tanto numéricos quanto visuais, mostram que os modelos de redes apresentados geram *clusters* que são discriminativos em relação à idoneidade de empresas, permitindo seu uso como modelo de classificação **(acerto de aproximadamente 9 em cada 10 fornecedores)**.
- Poder discriminativo dos *clusters* reforça a hipótese de que existem associações entre fornecedores, com potencial aprofundamento para **detecção de cartéis**.
- Sendo o ponto de partida as afiliações de Governo, fortalece-se numericamente a hipótese de que **agentes públicos desempenham um papel central da legalidade das transações financeiras** entre Governo e seus fornecedores.



- Melhorar o modelo de rede com **modelos econômicos mais completos**, visando tornar ainda maior o desempenho em classificação.
- Investigar se/como o trabalho pode ser adaptado e aplicado como uma ferramenta de ciência de dados na redução do espaço de busca por hipóteses plausíveis em **órgãos competentes de investigação de crimes financeiros (ex: COAF, PF, MPF, CGU, TCU) – todos potenciais clientes do SERPRO** nesta linha de negócio.
- Disseminar o conhecimento em SNA para **aplicações em outros domínios de interesse**, não apenas dados públicos (Portal da Transparência, dados.gov.br, etc), mas também dados do SERPRO e seus clientes.

Obrigado!

Marcelo Pita

E-mail: marcelo.pita@serpro.gov.br

Telefone: (31) 3311-6891

Website: <http://marcelopita.com>

Twitter: @marcelo_pita

