

Projeto em Teoria da Computação (MC859)

Lista de projetos

Prof. Dr. Ruben Interian

Instituto de Computação, UNICAMP

Resumo

1 Lista de projetos

2 Informações importantes

Tema dos projetos

Tema do projeto neste semestre:

Extração de conhecimento e detecção de padrões em grafos,
usando grandes volumes de dados e datasets reais.

- Há uma lista de possibilidades com os possíveis projetos a serem desenvolvidos.
- A apresentação de novos projetos pelos alunos é **fortemente incentivada**.

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?
- 4 Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?
- 4 Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?
- 5 Quais influenciadores devem ser escolhidos para dar o maior alcance a um meme?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?
- 4 Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?
- 5 Quais influenciadores devem ser escolhidos para dar o maior alcance a um meme?
- 6 Quais são os nós de maior risco em uma rede de vínculos transacionais?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?
- 4 Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?
- 5 Quais influenciadores devem ser escolhidos para dar o maior alcance a um meme?
- 6 Quais são os nós de maior risco em uma rede de vínculos transacionais?
- 7 Quantos conflitos de interesse existem em órgãos públicos da minha cidade?

Lista de projetos

- 1 Quem dos seus contatos levará você ao emprego dos seus sonhos?
- 2 Sou amigo dos amigos dos meus amigos, e inimigo dos inimigos dos meus amigos?
- 3 Como comparar a importância institucional e midiática dos políticos no Brasil?
- 4 Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?
- 5 Quais influenciadores devem ser escolhidos para dar o maior alcance a um meme?
- 6 Quais são os nós de maior risco em uma rede de vínculos transacionais?
- 7 Quantos conflitos de interesse existem em órgãos públicos da minha cidade?
- 8 *Existe uma forma de conciliar os fenômenos *small-world* (mundo pequeno) e *efeito Matthew* (conexão preferencial) em redes?

Lista de projetos

Quem dos seus contatos levará você ao emprego dos seus sonhos?

Lista de projetos

Quem dos seus contatos levará você ao emprego dos seus sonhos?

Suponha que existem um conjunto de pessoas e um conjunto de empresas-alvo. As pessoas estão vinculadas por meio de diferentes tipos de relações: amigo, conhecido, contato. As pessoas possuem atributos: onde estudou, gênero, habilidades, idiomas. Algumas pessoas estão empregadas em empresas, incluindo algumas das empresas-alvo.

Lista de projetos

Quem dos seus contatos levará você ao emprego dos seus sonhos?

Suponha que existem um conjunto de pessoas e um conjunto de empresas-alvo. As pessoas estão vinculadas por meio de diferentes tipos de relações: amigo, conhecido, contato. As pessoas possuem atributos: onde estudou, gênero, habilidades, idiomas. Algumas pessoas estão empregadas em empresas, incluindo algumas das empresas-alvo.

O projeto consiste em estudar quais vínculos (amigos próximos, conhecidos, contatos aleatórios), e pessoas com quais atributos, podem ajudar você a chegar até as empresas-alvo. Uma forma de modelar o problema é construir um algoritmo que, partindo de qualquer pessoa no grafo, tenha mais chance de alcançar uma empresa-alvo usando um número pequeno de passos (1, 2, 3).

Lista de projetos

**Sou amigo dos amigos dos meus amigos?
Sou inimigo dos inimigos dos meus amigos?**

Lista de projetos

**Sou amigo dos amigos dos meus amigos?
Sou inimigo dos inimigos dos meus amigos?**

Em um grupo de pessoas, algumas regras lógicas são geralmente observadas. As pessoas geralmente gostam de “amigos dos seus amigos”, e geralmente não gostam dos “inimigos dos seus amigos”. Se uma rede social possui arestas positivas e negativas, e satisfaz sempre essas regras, podemos dizer que ela está em equilíbrio.

Lista de projetos

Sou amigo dos amigos dos meus amigos? Sou inimigo dos inimigos dos meus amigos?

Em um grupo de pessoas, algumas regras lógicas são geralmente observadas. As pessoas geralmente gostam de “amigos dos seus amigos”, e geralmente não gostam dos “inimigos dos seus amigos”. Se uma rede social possui arestas positivas e negativas, e satisfaz sempre essas regras, podemos dizer que ela está em equilíbrio.

O projeto consiste em estudar o equilíbrio em subestruturas específicas (triângulos, ciclos) do grafo de vínculos entre pessoas, construído a partir das suas interações. Uma forma de modelar o problema é detectar a ocorrência ou frequência de ciclos positivos e negativos nos grafos. Podem ser usados dados de redes sociais, e-mails, Wikipedia.

Lista de projetos

Como comparar a importância institucional e midiática dos políticos no Brasil?

Lista de projetos

Como comparar a importância institucional e midiática dos políticos no Brasil?

As Pessoas Expostas Politicamente (PEPs) são ocupantes de cargos e funções públicas. Eles possuem diferentes graus de relevância que dependem do nível de governo (federal, estadual, municipal), e do respectivo poder (executivo, legislativo e judiciário). Porém, as PEPs também possuem diferentes graus de relevância midiática.

Lista de projetos

Como comparar a importância institucional e midiática dos políticos no Brasil?

As Pessoas Expostas Politicamente (PEPs) são ocupantes de cargos e funções públicas. Eles possuem diferentes graus de relevância que dependem do nível de governo (federal, estadual, municipal), e do respectivo poder (executivo, legislativo e judiciário). Porém, as PEPs também possuem diferentes graus de relevância midiática.

O projeto consiste em calcular e comparar a importância institucional e midiática das PEPs, identificando casos atípicos de diferenças desproporcionais entre elas. Uma forma de avaliar a importância midiática é usando as medidas de centralidade (*betweenness*, *PageRank*) na rede de co-ocorrência ou no grafo de conhecimento obtido dos textos de notícias publicadas durante um determinado período de tempo.

Lista de projetos

Como comparar a importância institucional e midiática dos políticos no Brasil?

As Pessoas Expostas Politicamente (PEPs) são ocupantes de cargos e funções públicas. Eles possuem diferentes graus de relevância que dependem do nível de governo (federal, estadual, municipal), e do respectivo poder (executivo, legislativo e judiciário). Porém, as PEPs também possuem diferentes graus de relevância midiática.

O projeto consiste em calcular e comparar a importância institucional e midiática das PEPs, identificando casos atípicos de diferenças desproporcionais entre elas. Uma forma de avaliar a importância midiática é usando as medidas de centralidade (*betweenness*, *PageRank*) na rede de co-ocorrência ou no grafo de conhecimento obtido dos textos de notícias publicadas durante um determinado período de tempo.

Observação: As informações básicas sobre cada PEP no Brasil é um dado público.

Lista de projetos

Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?

Lista de projetos

Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?

Os diversos grupos de pesquisa que temos na Unicamp, e em particular no Instituto de Computação, têm forte impacto nacional e internacionalmente. O impacto é avaliado tipicamente usando a quantidade de citações recebidas pelas publicações dos professores e pesquisadores em outras publicações científicas.

Lista de projetos

Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?

Os diversos grupos de pesquisa que temos na Unicamp, e em particular no Instituto de Computação, têm forte impacto nacional e internacionalmente. O impacto é avaliado tipicamente usando a quantidade de citações recebidas pelas publicações dos professores e pesquisadores em outras publicações científicas.

O projeto consiste em avaliar o impacto desses grupos de pesquisa. Para identificar os grupos de pesquisa, será necessário construir a rede de colaboração dos pesquisadores a partir das suas publicações. Para avaliar o impacto, será necessário usar a rede de citações entre as publicações.

Lista de projetos

Qual foi o impacto científico dos grupos de pesquisa do IC nos últimos anos?

Os diversos grupos de pesquisa que temos na Unicamp, e em particular no Instituto de Computação, têm forte impacto nacional e internacionalmente. O impacto é avaliado tipicamente usando a quantidade de citações recebidas pelas publicações dos professores e pesquisadores em outras publicações científicas.

O projeto consiste em avaliar o impacto desses grupos de pesquisa. Para identificar os grupos de pesquisa, será necessário construir a rede de colaboração dos pesquisadores a partir das suas publicações. Para avaliar o impacto, será necessário usar a rede de citações entre as publicações.

Observação: Os resultados não poderão ser divulgados ou publicados fora do ambiente de ensino da disciplina sem previa autorização dos referidos professores e pesquisadores.

Lista de projetos

Quais influenciadores escolher para dar o maior alcance a um meme?

Lista de projetos

Quais influenciadores escolher para dar o maior alcance a um meme?

Os memes, notícias, e epidemias são transmitidas através de interações entre pessoas no ambiente virtual e real. Se quisermos disseminar rapidamente uma informação entre usuários de uma rede social, precisaríamos escolher da melhor forma possível um grupo de influenciadores que darão início ao processo de divulgação.

Lista de projetos

Quais influenciadores escolher para dar o maior alcance a um meme?

Os memes, notícias, e epidemias são transmitidas através de interações entre pessoas no ambiente virtual e real. Se quisermos disseminar rapidamente uma informação entre usuários de uma rede social, precisaríamos escolher da melhor forma possível um grupo de influenciadores que darão início ao processo de divulgação.

O projeto consiste em criar um algoritmo que possa identificar um conjunto pequeno de influenciadores de forma que, após uma quantidade fixa de passos, a maior proporção possível da rede esteja infectada (tenha visto a notícia/meme divulgada pelos influenciadores). Em cada passo, uma pessoa pode ser infectada com uma probabilidade proporcional ao número de vizinhos já infectados.

Lista de projetos

Quais são os nós de maior risco em uma rede de vínculos transacionais?

Lista de projetos

Quais são os nós de maior risco em uma rede de vínculos transacionais?

Em uma rede de vínculos transacionais, as arestas ou arcos representam transações financeiras entre pessoas físicas ou jurídicas. Alguns nós ou vértices dessa rede são de risco quando possuem comportamento considerado atípico ou previamente identificado como de risco para outros clientes.

Lista de projetos

Quais são os nós de maior risco em uma rede de vínculos transacionais?

Em uma rede de vínculos transacionais, as arestas ou arcos representam transações financeiras entre pessoas físicas ou jurídicas. Alguns nós ou vértices dessa rede são de risco quando possuem comportamento considerado atípico ou previamente identificado como de risco para outros clientes.

O projeto consiste em identificar ao menos 5 comportamentos ou estruturas consideradas de risco, e implementar algoritmos que permitam detectar os nós da rede que participam dessas estruturas. Existem diversos comportamentos e estruturas possíveis, levando a diferentes formas de modelar e resolver o problema.

Lista de projetos

Quantos conflitos de interesse existem em órgãos públicos da minha cidade?

Lista de projetos

Quantos conflitos de interesse existem em órgãos públicos da minha cidade?

Um conflito de interesse ocorre quando uma pessoa fortemente ligada à administração de alguma instituição pública ou privada é beneficiária de uma empresa contratada por essa instituição. Conflitos de interesse oferecem riscos pois podem provocar o uso ineficiente ou indevido de recursos, sejam públicos ou privados.

Lista de projetos

Quantos conflitos de interesse existem em órgãos públicos da minha cidade?

Um conflito de interesse ocorre quando uma pessoa fortemente ligada à administração de alguma instituição pública ou privada é beneficiária de uma empresa contratada por essa instituição. Conflitos de interesse oferecem riscos pois podem provocar o uso ineficiente ou indevido de recursos, sejam públicos ou privados.

O projeto consiste em escolher um estado ou município, e identificar casos de conflitos de interesse de gestores públicos (pessoas que participam da administração). Exemplo simples de conflito de interesse: um secretário estadual de saúde contrata uma empresa da qual é sócio para o fornecimento de insumos para hospitais no seu estado. Há vários tipos de conflitos de interesse: o secretário poderia ser irmão do sócio da empresa.

Lista de projetos

Apresente o seu projeto!

Lista de projetos

Apresente o seu projeto!

Faça algo que você considere interessante, útil, relevante!

Resumo

1 Lista de projetos

2 Informações importantes

Lista de projetos

Os alunos/grupos precisarão capturar **datasets** que permitam obter e interpretar os resultados, e alcançar os objetivos propostos. Os datasets não podem ser inferiores a dezenas de milhares de vértices e arestas. O tamanho do dataset capturado será um dos critérios de avaliação da fase de coleta de dados e criação da instâncias.

Lista de projetos

Os alunos/grupos precisarão capturar **datasets** que permitam obter e interpretar os resultados, e alcançar os objetivos propostos. Os datasets não podem ser inferiores a dezenas de milhares de vértices e arestas. O tamanho do dataset capturado será um dos critérios de avaliação da fase de coleta de dados e criação da instâncias.

Os alunos/grupos precisarão apresentar uma **interface gráfica** que visualize as estruturas e padrões que foram estudados.

- Em alguns casos, pode ser útil ter um dashboard com os resultados.
- Em quase todos os casos, vai ser útil visualizar o grafo ou rede construída.

Lista de projetos

Existem diversos **datasets** disponíveis publicamente que podem ser essenciais para o seu projeto ou enriquecer as análises a serem realizadas. **Exemplos:**

Lista de projetos

Existem diversos **datasets** disponíveis publicamente que podem ser essenciais para o seu projeto ou enriquecer as análises a serem realizadas. **Exemplos:**

- **Publicações e citações – Scopus.** Periodicidade: poucos dias.
<https://www.scopus.com/search/form.uri>

Lista de projetos

Existem diversos **datasets** disponíveis publicamente que podem ser essenciais para o seu projeto ou enriquecer as análises a serem realizadas. **Exemplos:**

- **Publicações e citações – Scopus.** Periodicidade: poucos dias.
`https://www.scopus.com/search/form.uri`
- **Empresas no Brasil.** Periodicidade: mensal. Atualização: 13/07/2024.
URL: `https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj`
Mais recursos: `https://github.com/turicas/socios-brasil`

Lista de projetos

Existem diversos **datasets** disponíveis publicamente que podem ser essenciais para o seu projeto ou enriquecer as análises a serem realizadas. **Exemplos:**

- **Publicações e citações – Scopus.** Periodicidade: poucos dias.
<https://www.scopus.com/search/form.uri>
- **Empresas no Brasil.** Periodicidade: mensal. Atualização: 13/07/2024.
URL: <https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>
Mais recursos: <https://github.com/turicas/socios-brasil>
- **Pessoas Expostas Politicamente.** Periodicidade: mensal. Atualização: 06/2024.
<https://portaldatransparencia.gov.br/download-de-dados/pep>

Lista de projetos

Existem diversos **datasets** disponíveis publicamente que podem ser essenciais para o seu projeto ou enriquecer as análises a serem realizadas. **Exemplos:**

- **Publicações e citações – Scopus.** Periodicidade: poucos dias.
<https://www.scopus.com/search/form.uri>
- **Empresas no Brasil.** Periodicidade: mensal. Atualização: 13/07/2024.
URL: <https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>
Mais recursos: <https://github.com/turicas/socios-brasil>
- **Pessoas Expostas Politicamente.** Periodicidade: mensal. Atualização: 06/2024.
<https://portaldatransparencia.gov.br/download-de-dados/pep>
- **E muitos outros!**

Lista de projetos

Observação crucial: NUNCA coletar ou extrair informações que são **identificáveis**.

Lista de projetos

Observação crucial: **NUNCA** coletar ou extrair informações que são **identificáveis**.

Exemplos:

- 1 Não coletar o nome completo.

Lista de projetos

Observação crucial: **NUNCA** coletar ou extrair informações que são **identificáveis**.

Exemplos:

- 1 Não coletar o nome completo.
- 2 Não coletar CPF, número de documento de identidade, CNH.

Lista de projetos

Observação crucial: **NUNCA** coletar ou extrair informações que são **identificáveis**.

Exemplos:

- 1 Não coletar o nome completo.
- 2 Não coletar CPF, número de documento de identidade, CNH.
- 3 Não coletar endereço de PF (pode coletar o estado, ou no máximo a cidade).

Lista de projetos

Observação crucial: **NUNCA** coletar ou extrair informações que são **identificáveis**.

Exemplos:

- 1 Não coletar o nome completo.
- 2 Não coletar CPF, número de documento de identidade, CNH.
- 3 Não coletar endereço de PF (pode coletar o estado, ou no máximo a cidade).
- 4 Não coletar fotos ou alguma outra informação que permita identificar o indivíduo.

Datas importantes

30 de agosto – Último dia para a entrega da Proposta de Projeto a ser implementado.

Não espere até o último dia!