

Optimization Problem Description

■ Informal Description:

– Given:

- A **set T of collaborative** (or shared) **tasks** to be performed, where each task is described by a vector of attributes (or requirements) to be fulfilled
- A **set P of people**, where each person is **skilled** with a vector of attributes (expertise levels), expressing the **capabilities** that each person has to work on tasks in T.
- **Capacity constraints**, stating the maximum (minimum) number of tasks that may be assigned to a person.
- An amount to be paid (**cost**) for engaging (**hiring**) people to work on collaborative tasks in T.
- A **budget B**, equivalent to the money available for hiring people.
- A social network represented by an **undirected graph** $G=(V,U,w)$, where **vertices** in V denote **people** and **edges** (u,v) in E correspond to **social connections** between \underline{u} and \underline{v} , and $\mathbf{w}: E \rightarrow \mathbb{R}^+$ is a cost function indicating the level of **collaboration** between \underline{u} and \underline{v} computed from historical data concerning past interactions.

– On is required to **find an assignment of persons to tasks** aiming at maximizing the (potential) total collaboration level of the team members.

Recently a very similar problem was proven to be NP-hard, by [LLT-2009]



Data sources

1. The DBLP Computer Science Bibliography:
 - Data: bibliographic information on major computer science journals and proceedings.
 - Task: form teams of scholars to write (multidisciplinary) papers
2. IBM Connections Collaboration Platform:
 - Data: work network reports like records in: wikis, forums, e-mails, shared and co-authored documents, common memberships, functional hierarchies; historical records of projects' engagements.
 - Task: construct successful project teams where members are expected to work in collaborative activities
3. **Marvel comic book database:**
 - Data: records of Marvel comic books.
 - Task: given a team of villains, find a best team of heroes that are able to beat the given villains' team.

For all these three scenarios, we are interested in identifying teams that maximize collaboration among their members.

Data sources: Marvel comic book database

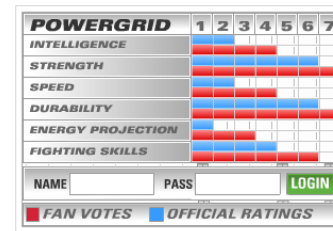
Given:

a) a **set of villains** to be defeated, chosen at random from nodes of a same connected component $C[v \in \text{Villains}]$, i.e., C is the subgraph induced by the partition of villains.

b) for each **character** :

1. It is know a set of **power grids** (according to the official ratings of Marvel site) where each power grid is composed by a set of six **abilities**, namely:

- Intelligence
- Strength
- Speed
- Durability
- Energy projection
- Fighting Skills



2. It is associated a **cost** proportional to its average power grid and popularity.

c) a (bipartite) weighted graph where:

- Each **node** represents a **character**;
- Exists the **edge** (u, v) if the pair of characters \underline{u} and \underline{v} appeared in a same comic book;
- The **collaboration level** between \underline{u} and \underline{v} is given by the **number of times** that these **characters** appeared in a same comic book.

d) a **minimum AVG ability**: A team H is capable to beat an opponent team V if the average power grid of H , for all six abilities, is at least equal to the average power grid for V .

e) a value B corresponding to the **budget** available to form a team of heroes.

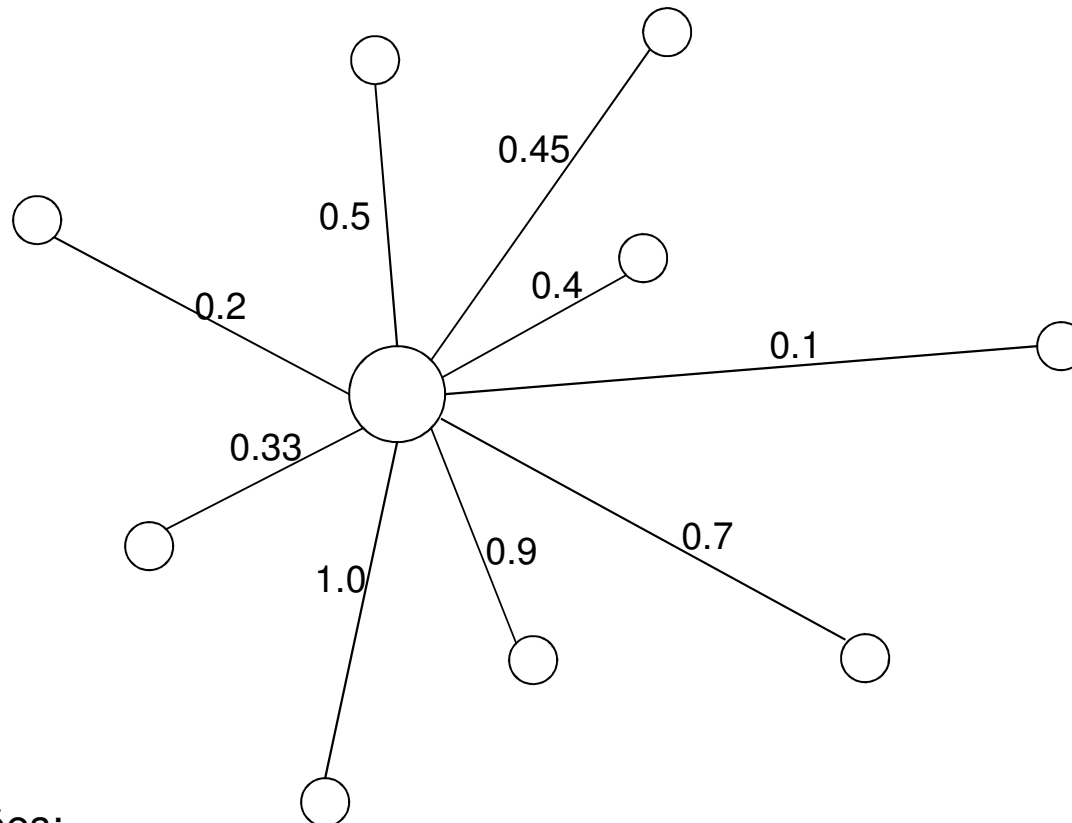
f) upper and/or lower **bound(s)** on the heroes' team size.

Trabalho Implementação & Seminários

• Implementação

- Uma solução heurística (ou híbrida) baseada em algoritmos de otimização aplicáveis ao problema de (Social) Team Formation.
- **Dados:** grafo construído em sala de aula e super-heróis
- **Algoritmos heurísticos** sugeridos:
 - Simulated Annealing; Tabu Search; Genetic Algorithms; Scatter Search; GRASP; Iterated Local Search; Multi-star methods; Asynchronous Teams; Ant Colony ou Neural Networks para problemas de otimização.
- **Fontes:** cópias digitais dos livros que serão disponibilizadas (para uso na disciplina); os alunos podem complementar seus trabalhos com artigos disponíveis na literatura de cada algoritmo.
- **Entrega:**
 - Prazo: **01/12/2014 (08/12/2014** para trabalhos limitados ao valor de **50% do inicial)**
 - Conteúdo: texto descritivo (explicando o algoritmo e especificidades da implementação) e códigos (fonte e executável).
- **Avaliação:**
 - Implementações incorretas ou incompletas: nota máxima: 40% do valor inicial
 - Implementações corretas:
 - a) Comparação com implementação GRASP atual (valor de função objetivo)
 - b) Resultados superiores ou iguais em pelo menos 50% dos casos (pelo menos um deles SUPERIOR ao do GRASP já disponível): nota = 100% do valor inicial!
 - c) Demais resultados: nota acima de 50% e abaixo de 100% do valor inicial.

Grafo representando a “rede social” da turma



Sugestões:

- 1 - Só por zero (sem aresta) se não souber quem é a outra pessoa.
- 2 – Valores com 2 casas decimais, no máximo.

Trabalho Implementação & Seminários

- **Tópicos Seminários:**

- Os alunos deverão apresentar o método heurístico de solução de escolhido para implementação, incluindo uma descrição da implementação feita.
- Cada grupo tem liberdade de propor uma variação de um dos algoritmos citados acima como opção de implementação.
- **Tempo:** 30 minutos para apresentação (até 3 por aula)
- **Prazo** para definição de grupos e tópicos: 29/10.
- **Tamanho dos grupos:** 4 a 6 pessoas
- **Datas:** 01/12, 03/12 e 10/12
 - Nota: Se mais que 3 dias forem necessários, dia 15/12 será usado para as últimas apresentações.

Lista de Exercícios sobre ML

- **Lista de Exercícios:**

- Trabalho individual ou em dupla
- Quatro exercícios do primeiro capítulo do livro *Learning from Data*:
 - Obrigatório: Exercício 1.3
 - Opcional: escolher 3 exercícios dentre 1.1, 1.2, 1.4, 1.5 e 1.6.
- Prazo para entrega (sem perda de pontos por atraso): 29/10.