Graph theory

Victor Eijkhout

Fall 2020

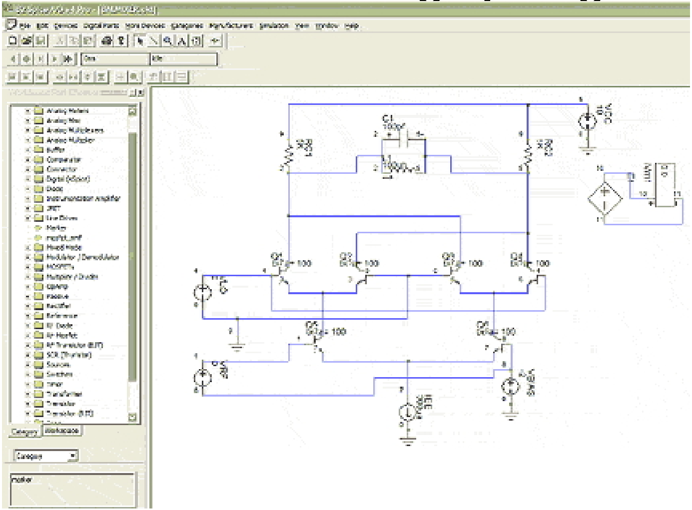
Justification

Graph theory has many applications in computational math. Here we focus on the equivalence with sparse matrices.

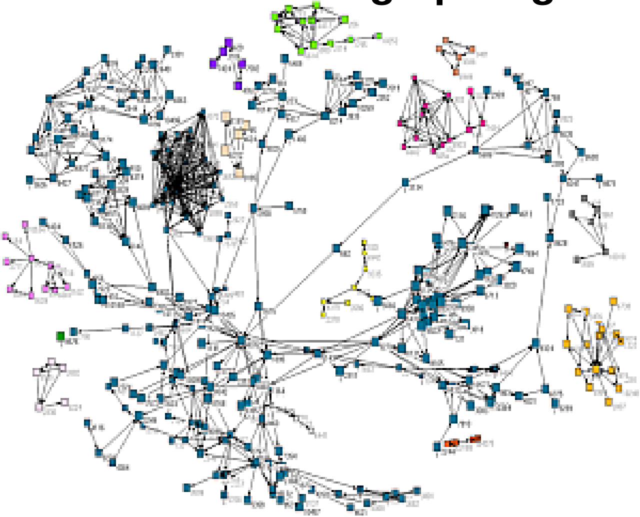
Graph algorithms

- Traditional: search, shortest path, connected components
- New: centrality

Traditional use of graph algorithms

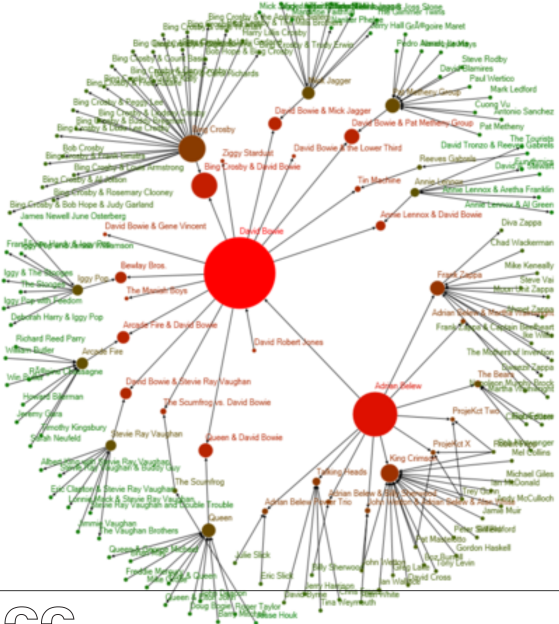


1990s use of graph algorithms



2010 use of graph algorithms

2010 use of graph algorithms



Traditional graph algorithm

Input : A graph, and a starting node s

Output: A function $d(v)$ that measures the distance from s to v

Let s be given, and set $d(s) = 0$

Initialize the finished set as $U = \{s\}$

Set $c = 1$

while *not finished* **do**

  Let V the neighbours of U that are not themselves in U

  if $V = \emptyset$ **then**

    We're done

  else

    Set $d(v) = c + 1$ for all $v \in V$.

    $U \leftarrow U \cup V$

    Increase $c \leftarrow c + 1$

Computational characteristics

- Uses a queue: central storage
- Parallelism not self-evident
- Flexible assignment of work to processors, so no locality

Matrix formulation

Let

$$x_i = \begin{cases} 1 & i = s \\ \infty & \text{otherwise} \end{cases}$$

Let x zero except in i,
then $x^t G$ nonzero in j if there is an edge (i,j)

Matrix algorithm

Define a product as

$$y^t = x^t G \equiv \forall_i \colon (y^t)_j = \min_{i\colon G_{ij} \neq 0} x_i + 1,$$

Iterate

$$x, x^t G, x^t G^2, \ldots$$

After k (diameter) iterations $(x^t G^k)_i$ is the distance $d(s,i)$.

Single Source Shortest Path

Similar to previous, but non-unit edge weights

Let s be given, and set $d(s) = 0$
Set $d(v) = \infty$ for all other nodes v
for $|E| - 1$ *times* **do**
  for *all edges* $e = (u, v)$ **do**
    Relax: **if** $d(u) + w_{uv} < d(v)$ **then**
      Set $d(v) \leftarrow d(u) + w_{uv}$

$$y^t = x^t G \equiv \forall_i \colon y_j = \min\left\{x_j, \min_{i \colon G_{ij} \neq 0} \{x_i + g_{ij}\}\right\},$$

All-pairs shortest path

$$\Delta_{k+1}(u,v) = \min\left\{\Delta_k(u,v), \Delta_k(u,k) + \Delta_k(k,v)\right\}. \qquad (1)$$

Algebraically:

for k *from zero to* $|V|$ **do**
$\quad D \leftarrow D._{\min}\left[D(:,k) \min\cdot_{+} D(k,:)\right]$

Similarity to Gaussian elimination

Pagerank

T stochastic: all rowsums are 1.

Prove $x^t e = 1 \Rightarrow x^t T = 1$

Pagerank is essentially a power method: $x^t, x^t T, x^t T^2, \ldots$ modeling page transitions.

Prevent getting stuck with random jump:

$$x^t \leftarrow s x^t T + (1-s)e^t$$

Solution of linear system:

$$x^t(I - sT) = (1-s)e^t$$

Observe

$$(I - sT)^{-1} = I + sT + s^2 T^2 + \cdots$$

‘Real world’ graphs

- Graphs imply sparse matrix vector product
- …but the graphs are unlike PDE graphs
- differences:
  - low diameter
  - high degree
  - power law
- treat as random sparse: use dense techniques
- 2D matrix partitioning: each block non-null, but sparse

Parallel treatment

- Intuitive approach: partitioning of nodes
- equivalent to 1D matrix distribution
- not scalable $\Rightarrow$ 2D distribution
- equivalent to distribution of edges
- unlike with PDE graphs, random placement may actually be good