Efficient PageRank Tracking in Evolving Networks

Алмаз Насибуллин

План

- Задача
- Предложенный алгоритм
- Сравнительный анализ
  - Теоретическое сравнение
  - Эксперимент

Задача

- Подсчет важности каждой вершины в режиме реального времени
- Personalized PageRank

Метрики качества

- Время работы алгоритма
- Точность решения
- Пространственная сложность

Актуальность задачи

- Громадные сети
- Быстрое развитие сетей

Name	Total	Growth
Web pages	60.0T	+600K/sec
Facebook users	1.30B	+5/sec
Google+ users	700M	+19/sec
Twitter users	300M	+3/sec

Таблица 1: Развитие крупных сетей Интернет, 1.11.2014.

Personalized PageRank(PPR)

- $G = (V, E)$ - не взвешенный ориентированный граф
- $P_{ij} = \begin{cases} 1/d_j, & (j,i) \in E \\ 0, & \text{otherwise} \end{cases}$ - матрица переходов
- b - вектор предпочтений, причем $b_i \geq 0$ и $\sum_{i \in V} b_i = 1$
- $\alpha \in (0, 1)$ - фактор ослабления($\alpha = 0.85$)

$$x = \alpha P x + (1 - \alpha) b$$

Метод Gauss–Southwell

$$x = \alpha P x + (1 - \alpha) b$$

- ✔ $x^{(\nu)}$ - ?
- ✔ $r^{(\nu)} = (1 - \alpha) b - (I - \alpha P) x^{(\nu)}, \quad r^{(\nu)} \to 0$

- $$x^{(\nu)} = x^{(\nu-1)} + r_i^{(\nu-1)} e_i,$$
  $$r^{(\nu)} = r^{(\nu-1)} - r_i^{(\nu-1)} e_i + \alpha r^{(\nu-1)} P e_i$$

Метод Gauss–Southwell

Algorithm 1 Gauss–Southwell algorithm for PPR. Here $x^{(0)}$ and $r^{(0)}$ are initial solution and the corresponding residual such that $r^{(0)} = (1-\alpha)b - (I - \alpha P)x^{(0)}$.

```
procedure GAUSS–SOUTHWELL(x^(0), r^(0))
    for ν = 1, 2, ... do
        Pick the largest entry r_i^(ν)
        if |r_i^(ν)| < ε then
            terminate
        end if
        Update x^(ν) and r^(ν) according to Eq. (6)
    end for
end procedure
```

$$x^{(\nu)} = x^{(\nu-1)} + r_i^{(\nu-1)} e_i, \tag{6}$$

$$r^{(\nu)} = r^{(\nu-1)} - r_i^{(\nu-1)} e_i + \alpha r^{(\nu-1)} P e_i. \tag{7}$$

Алмаз Насибуллин

Алмаз Насибуллин