

بررسی نمونه‌ای از گراف وب

مهدی حسین‌زاده

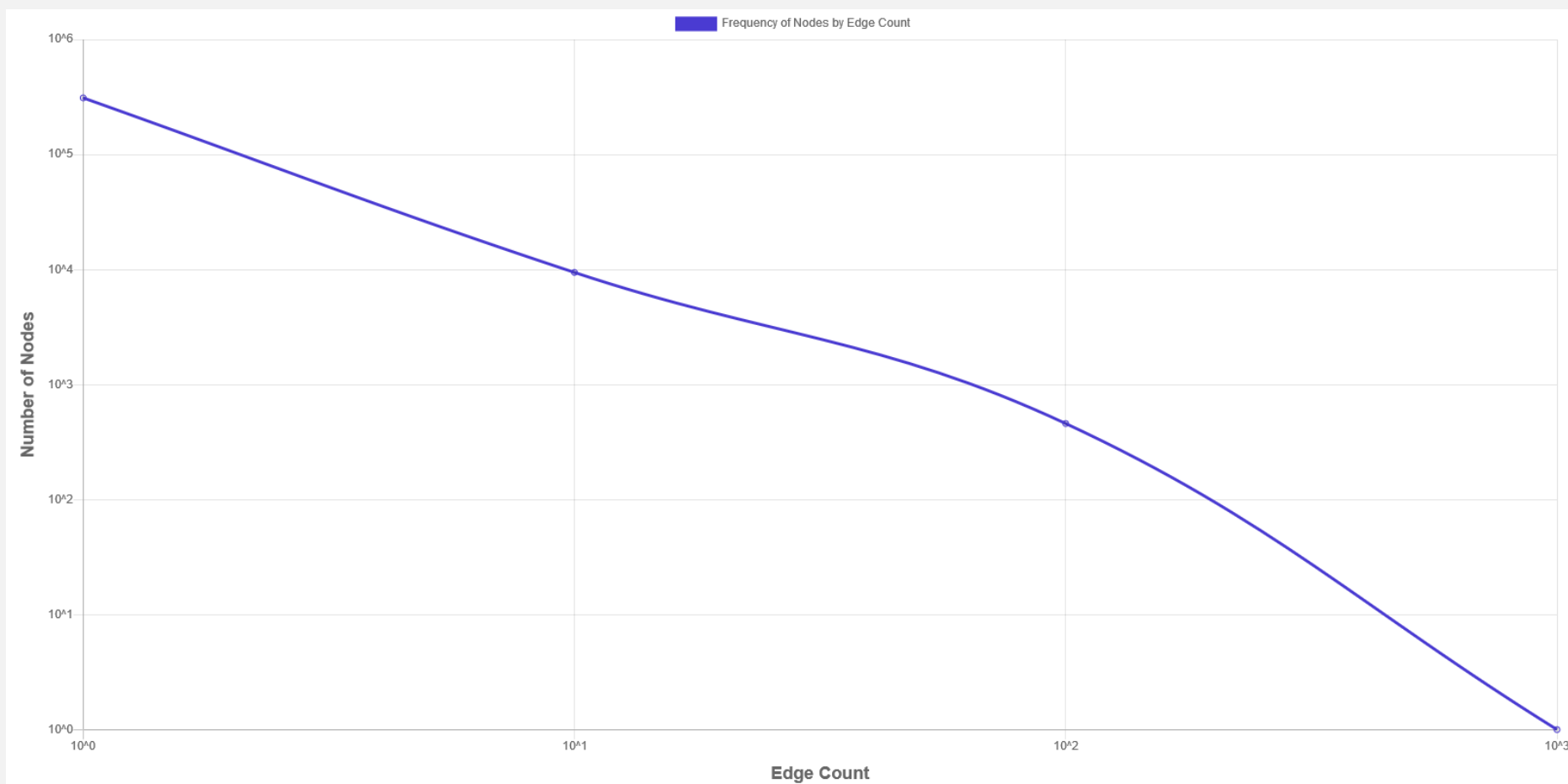
I. POWER LAW

- لیست تعداد یال‌های ورودی هر نود ایجاد می‌شود
- تعداد نودهایی که تعداد یال یکسان دارند مشخص می‌شود
- نودها بر اساس تعداد یال دسته‌بندی می‌شوند
- مقدار آلفا و گاما بر طبق زیر به دست می‌آید:

$$\gamma = 1 + \frac{n}{\sum_{i=1}^n \ln \frac{x_i}{(x_{min} - \frac{1}{2})}} \xrightarrow{x_{min}=1} \gamma = 1 + \frac{n}{\sum_{i=1}^n 2 \ln x_i}$$

مقدار α با توجه به γ و فرمول $P(x)$ محاسبه می‌شود.

I. POWER LAW



نتیجه اجرا:

Alpha: 5.49

Gamma: 1.56

Time: 11s

2. BOW TIE

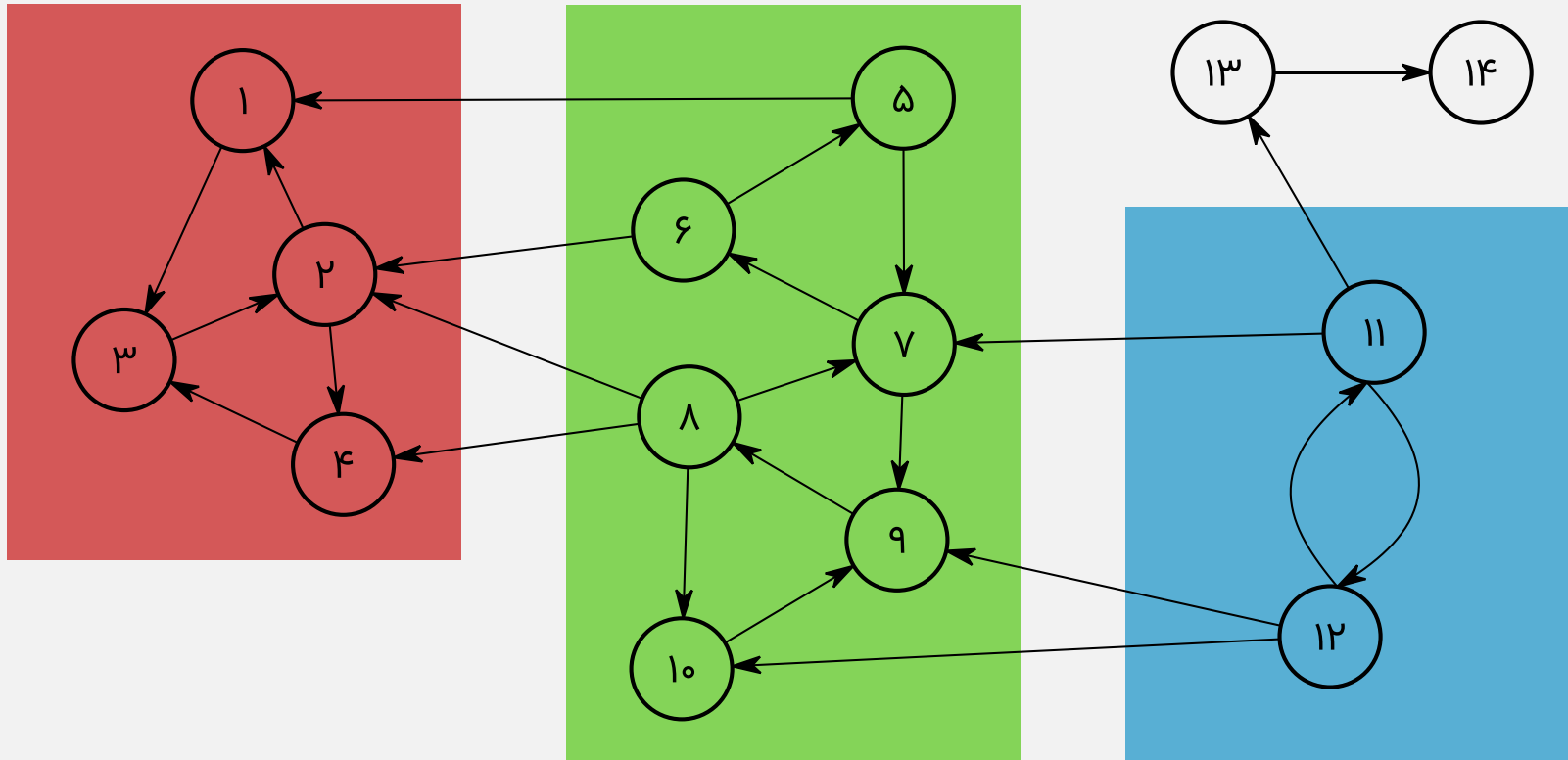
شناسایی بخش‌های گراف با استفاده از الگوریتم ارائه شده در فایل PDF

(۱) خود گراف و معکوس آن خوانده می‌شود و در حافظه به صورت map نگهداری می‌شود

(۲) استخراج قسمت core گراف

- بخش همبندی که اندازه‌ی آن از همه بزرگتر است
- استخراج یک بخش همبند با شروع از نود n :
- به دست آوردن لیست نودهای قابل دسترسی از n و دارای دسترسی به n (با الگوریتم DFS)
- بخش همبند اشتراک دو لیست بالا است (شامل خود n)
- نکته‌ی مهم حذف نودهای همبند تشخیص داده شده از گراف داخل حافظه است

2. BOW TIE



2. BOW TIE

(۳) به دست آوردن بخش **Out** گراف:

- انتخاب یک نود نمونه از **core** مانند n سپس
- نودهای **core** – نودهای قابل دسترس از $n = \text{Out}$

(۴) به دست آوردن بخش **In** گراف:

- انتخاب یک نود نمونه از **core** مانند n سپس
- نودهای **core** – نودهای دارای دسترسی به $n = \text{In}$

2. BOW TIE

• نتیجه اجرا:

- Core size: 504,930 (58%)
- Out size: 108,472 (13%)
- In size: 242,507 (28%)
- Other: 722 (1%)
- Time: 58s

3. AVERAGE DISTANCE

- انتخاب ۲۰۰۰ عدد (۱۰۰۰ جفت) نود متمایز به صورت تصادفی
 - به دست آوردن فاصله‌ی جفت نودها (فاصله‌ی نود اول از دوم)
 - استفاده از الگوریتم BFS
 - محاسبات با چند ترد انجام می‌شود
 - نتیجه یک اجرا:
- Average distance: 4.22
 - Time: 25m

با تشكر