

به نام خدا

تمرین چهارم

درس شبکه‌های اجتماعی

رضا منصوری خواه

810103246

:استاد درس

دکتر اسدپور

K shell:

K_shell محاسبه شده توسط کد پایتون

```
import pandas as pd
import networkx as nx

df = pd.read_csv('edg.csv')

G = nx.from_pandas_edgelist(df, source='Source', target='Target', edge_attr='Weight')

def weighted_core_number(G):
    core_numbers = {}
    for node in G.nodes():
        neighbors = list(G.neighbors(node))
        total_weight = sum([G[node][neighbor]['Weight'] for neighbor in neighbors])
        core_numbers[node] = total_weight
    return core_numbers

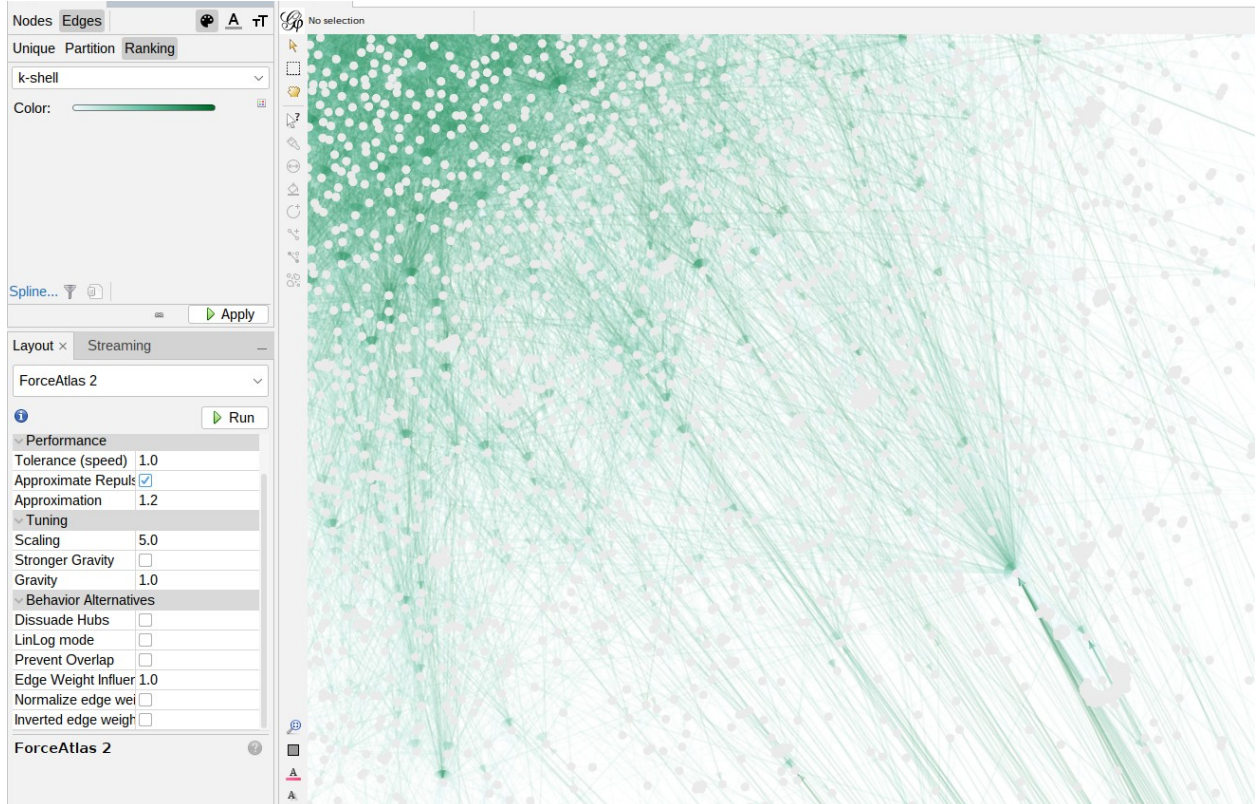
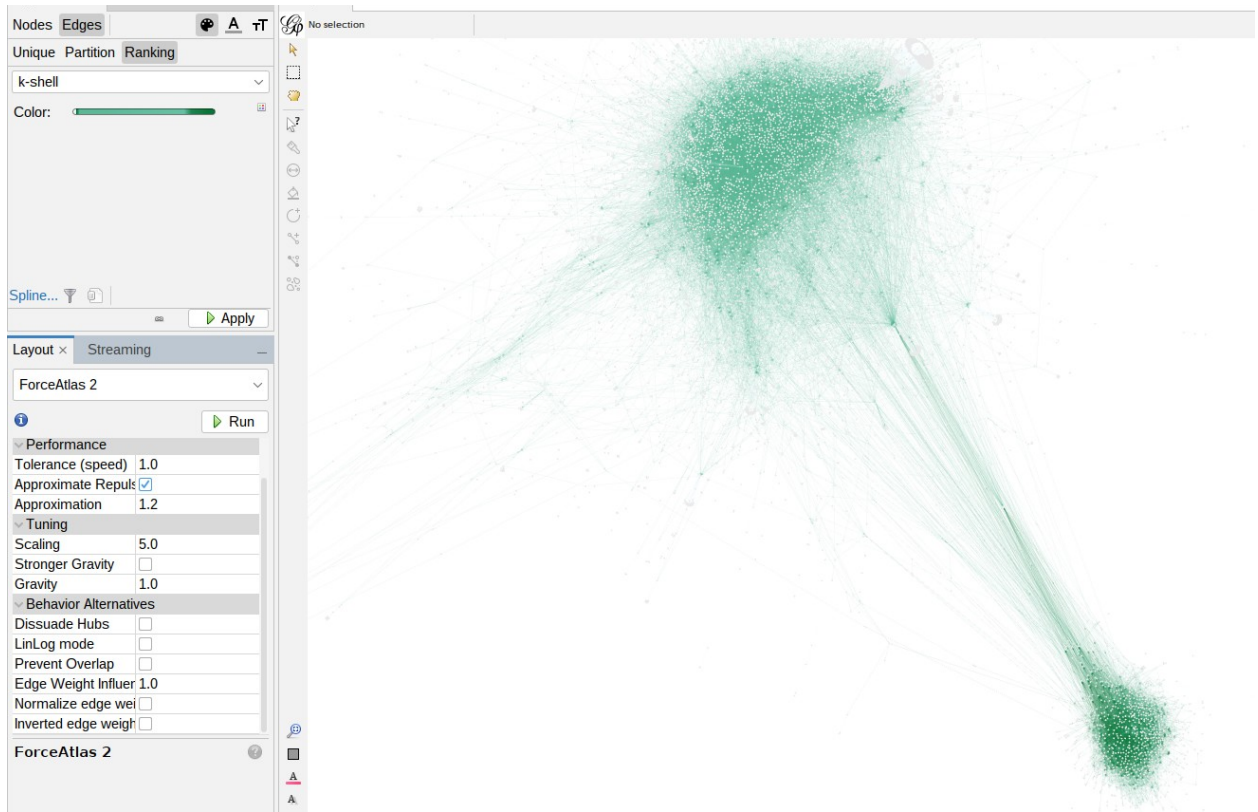
k_shell = weighted_core_number(G)

df['K-Shell'] = df['Source'].map(k_shell)

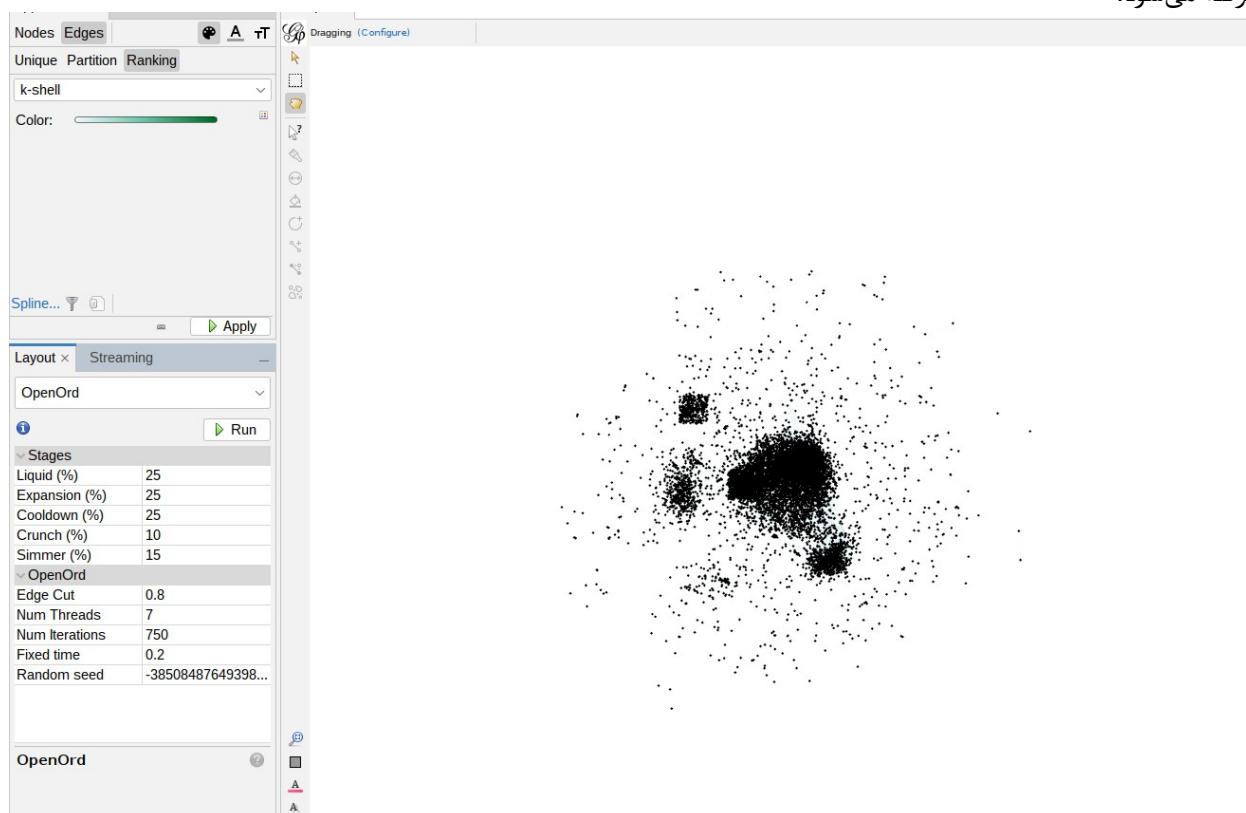
communities_by_k_shell = {}
for node, k in k_shell.items():
    if k not in communities_by_k_shell:
        communities_by_k_shell[k] = []
    communities_by_k_shell[k].append(node)

df['Community'] = df['Source'].map(lambda x: k_shell[x])

df.to_csv('your_file_with_communities_and_k_shell.csv', index=False)
print(df.head())
```



در این حالت، هدف این است که نودها را بر اساس K-shell تقسیم کنیم و از K-shell به عنوان معیار برای تعیین انجمن‌ها استفاده کنیم. به عبارت دیگر، نودهایی که K-shell مشابهی دارند به عنوان یک انجمن در نظر گرفته می‌شوند. شناسایی نود مرکزی هر انجمن: برای هر انجمن، نودی که بیشترین K-shell را دارد به عنوان نود مرکزی آن انجمن در نظر گرفته می‌شود.



تراکم بالا در هسته مرکزی:

- تصویر نشان می‌دهد که تعداد زیادی از گره‌ها در هسته مرکزی قرار گرفته‌اند. این نشان‌دهنده آن است که این گره‌ها (کاربران یا حساب‌های توییتر) تعاملات بیشتری داشته‌اند و احتمالاً تأثیرگذارترین کاربران در شبکه هستند.
- این کاربران ممکن است شامل حساب‌های رسمی، سازمان‌ها، یا افراد تأثیرگذار باشند که در حمایت از اسرائیل فعالیت دارند.

ساختار کشیده در بخش‌های پایینی:

- بخش کشیده در پایین تصویر نشان‌دهنده گره‌هایی است که در K-shell پایین‌تر قرار دارند.
- این گره‌ها نقش حاشیه‌ای در شبکه ایفا می‌کنند و احتمالاً شامل کاربران عادی یا حساب‌هایی با تعامل کمتر هستند.

ارتباطات بین لایه‌ها:

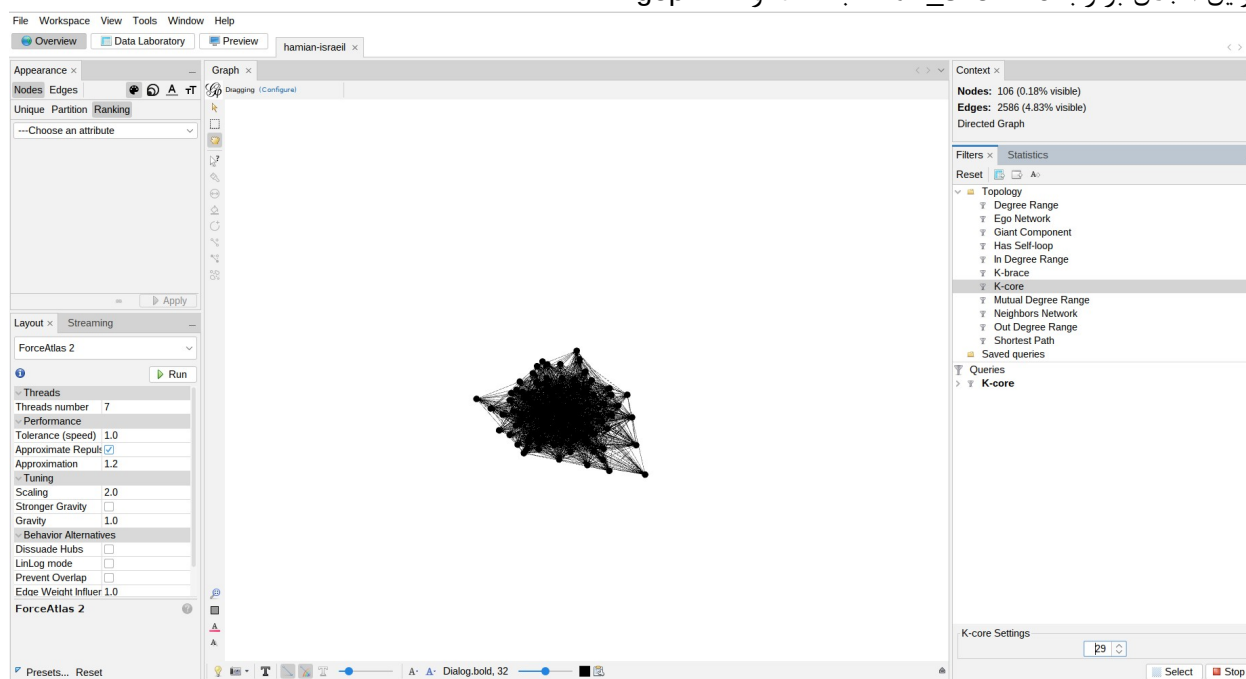
● خطوط بین هسته مرکزی و بخش‌های پایینی نشان می‌دهد که گره‌های هسته مرکزی به‌عنوان واسطه‌هایی برای انتشار اطلاعات به گره‌های کم‌اهمیت‌تر عمل می‌کنند.

تمرکز بر هسته مرکزی:

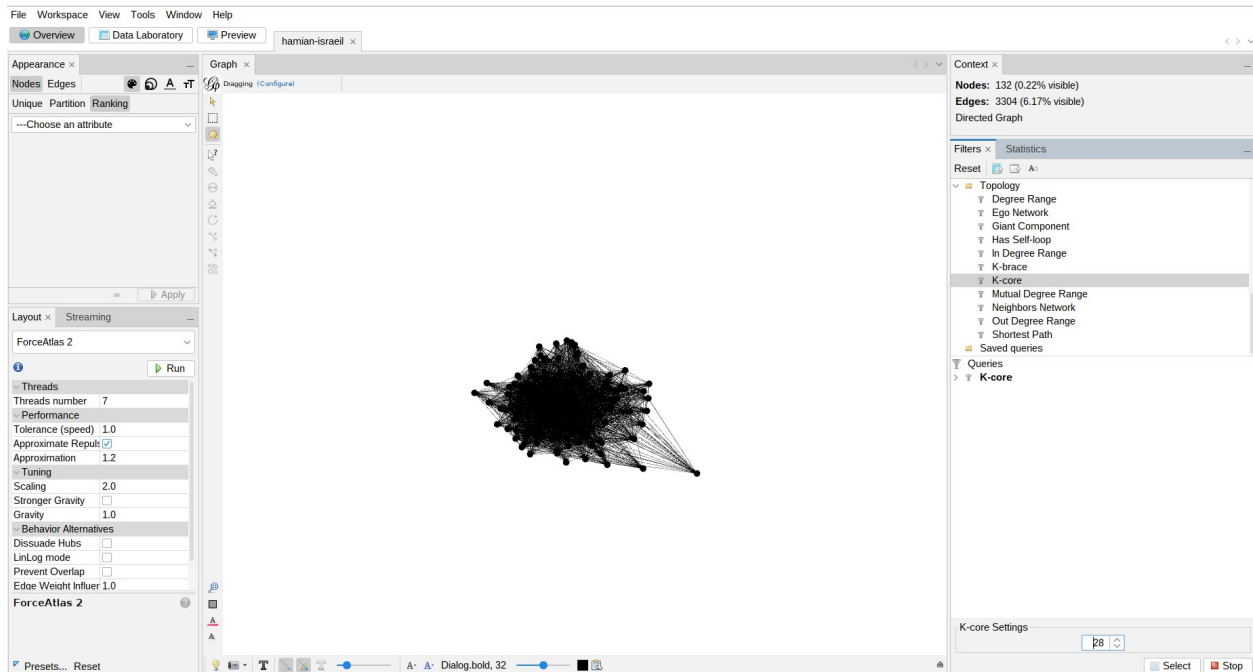
● گره‌های با **K-shell** بالا نقش مهمی در حفظ انسجام شبکه دارند. حذف این گره‌ها می‌تواند به فروپاشی بخش‌های بزرگی از شبکه منجر شود.

محاسبه توسط gephi:

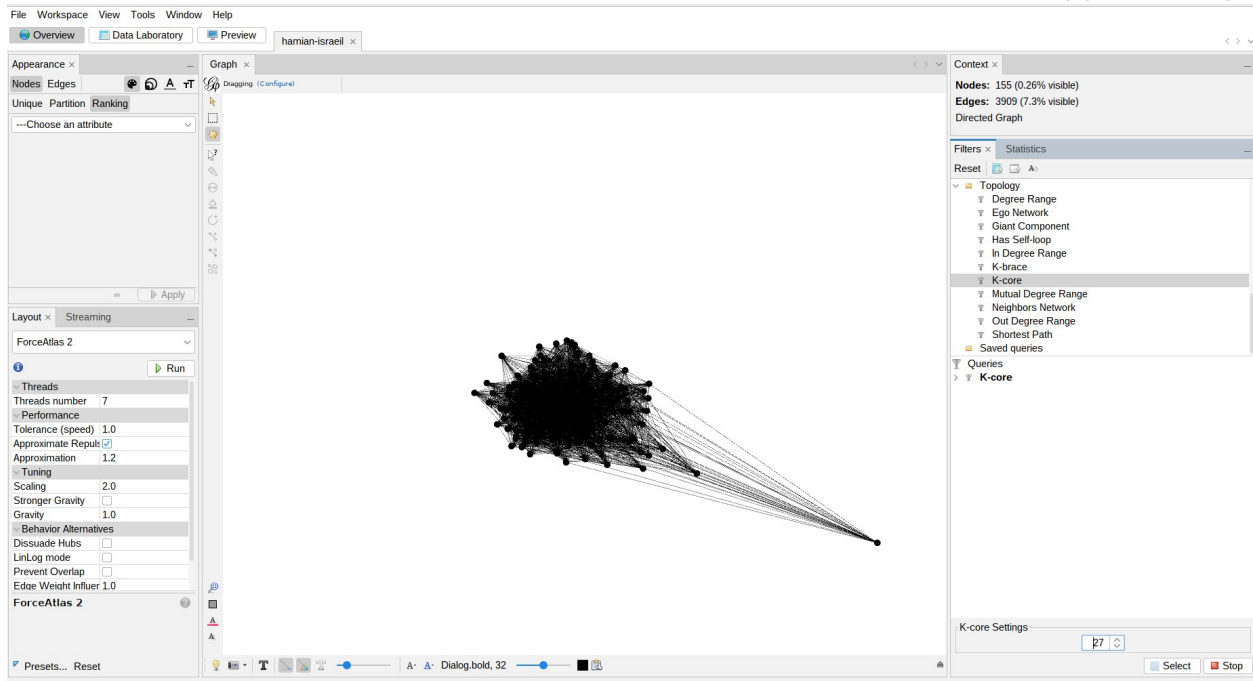
اولین انجمن برتر با $k_shell=29$ محاسبه شده توسط gephi



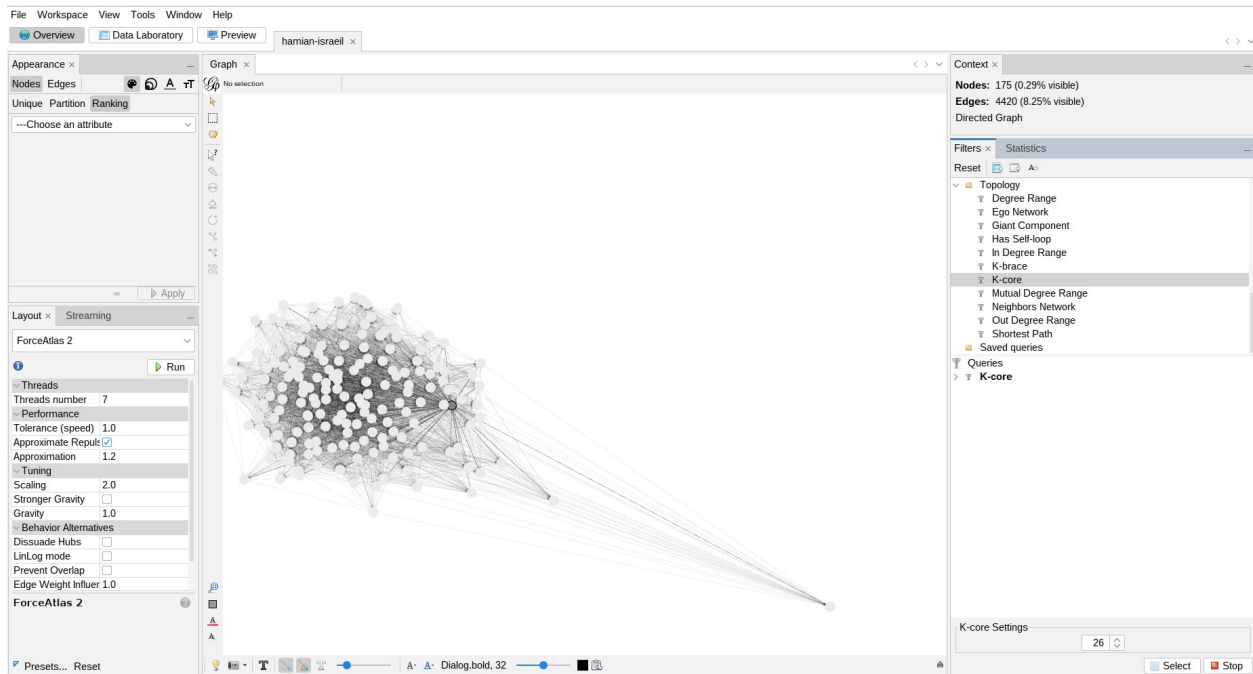
دومین انجمن برتر با $k_shell=28$ محاسبه شده توسط gephi



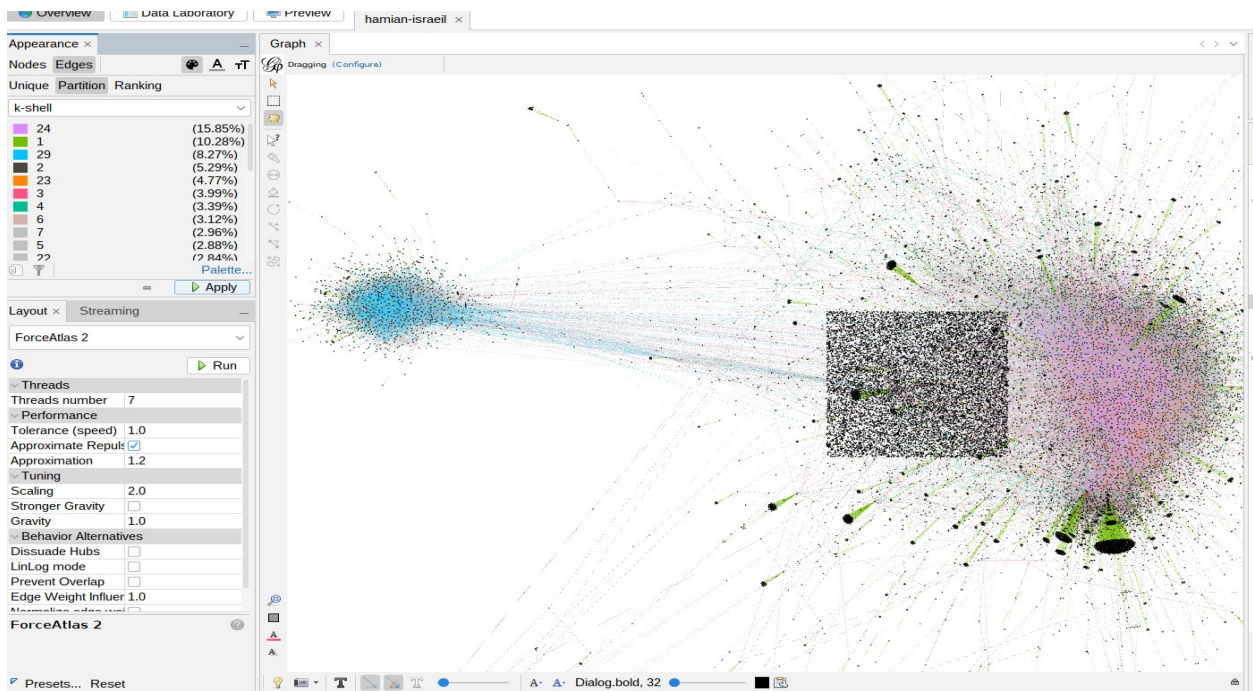
سومین انجمن برتر با $k_{shell}=27$



چهارمین انجمن برتر با $k_{shell}=26$

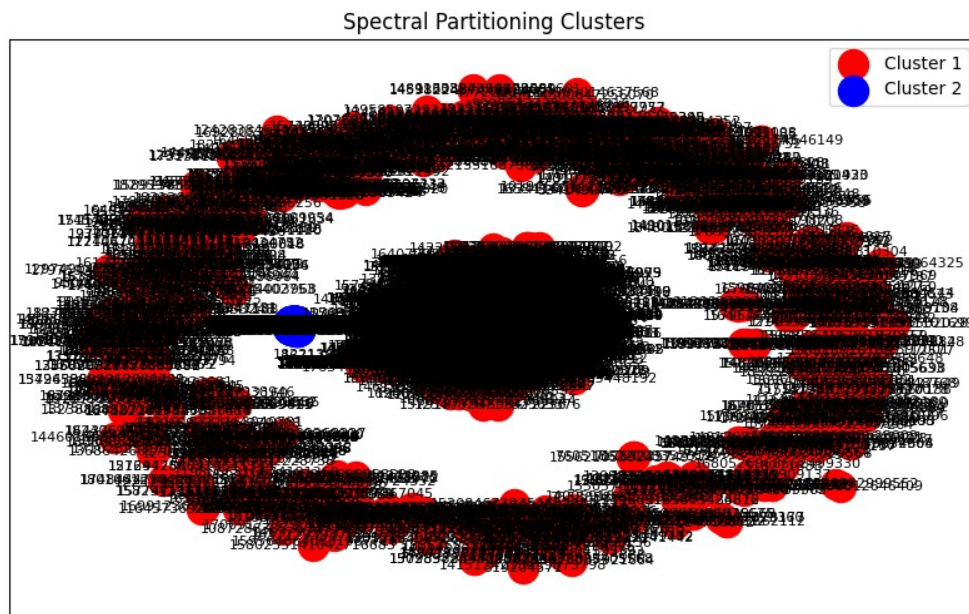


محاسبه شده با استفاده از کد



partitioning Spectral:

Spectral Partitioning (افراز طیفی گراف) یک روش ریاضی و الگوریتمی است که برای تقسیم یک گراف به دو یا چند زیربخش (خوشه یا اجتماع) استفاده می‌شود. این روش بر اساس خواص طیفی ماتریس‌های مرتبط با گراف، مانند ماتریس لاپلاسیان، عمل می‌کند.



عکس بالا نتیجه کد زیر برای محاسبه ی پارامتر است

```
import pandas as pd
import networkx as nx
import numpy as np
from scipy.sparse.linalg import eigsh
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

csv_file_path = 'edg.csv'
data = pd.read_csv(csv_file_path)
G = nx.from_pandas_edgelist(data, source='Source', target='Target')

L = nx.laplacian_matrix(G).astype(float)
k = 2
vals, vecs = eigsh(L, k=k, which='SM')

embedding = np.real(vecs)
```



```

kmeans = KMeans(n_clusters=k, random_state=42)
labels = kmeans.fit_predict(embedding)
plt.figure(figsize=(10, 6))
colors = ['red', 'blue', 'green', 'yellow', 'purple']
pos = nx.spring_layout(G)

for i in range(k):
    cluster_nodes = [n for n, label in zip(G.nodes(), labels) if label == i]
    nx.draw_networkx_nodes(G, pos, nodelist=cluster_nodes, node_color=colors[i], label=f"Cluster {i+1}")

nx.draw_networkx_edges(G, pos, alpha=0.5)
nx.draw_networkx_labels(G, pos, font_size=8)
plt.legend()
plt.title("Spectral Partitioning Clusters")
plt.show()

```

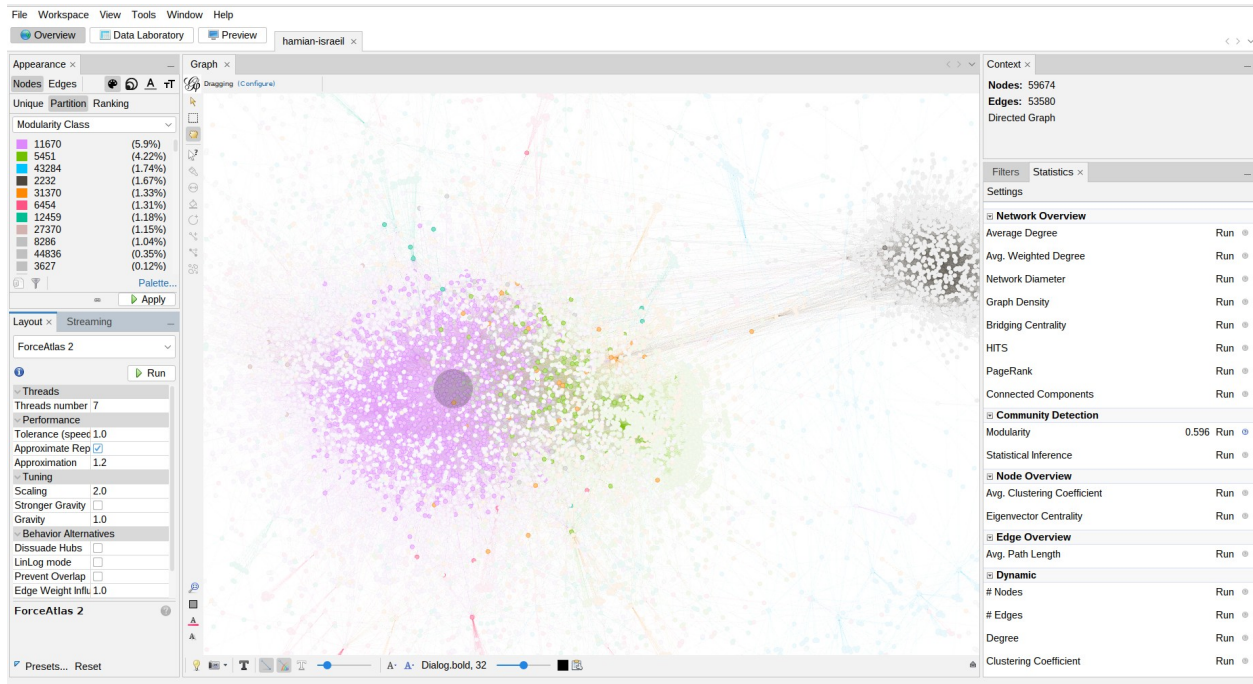
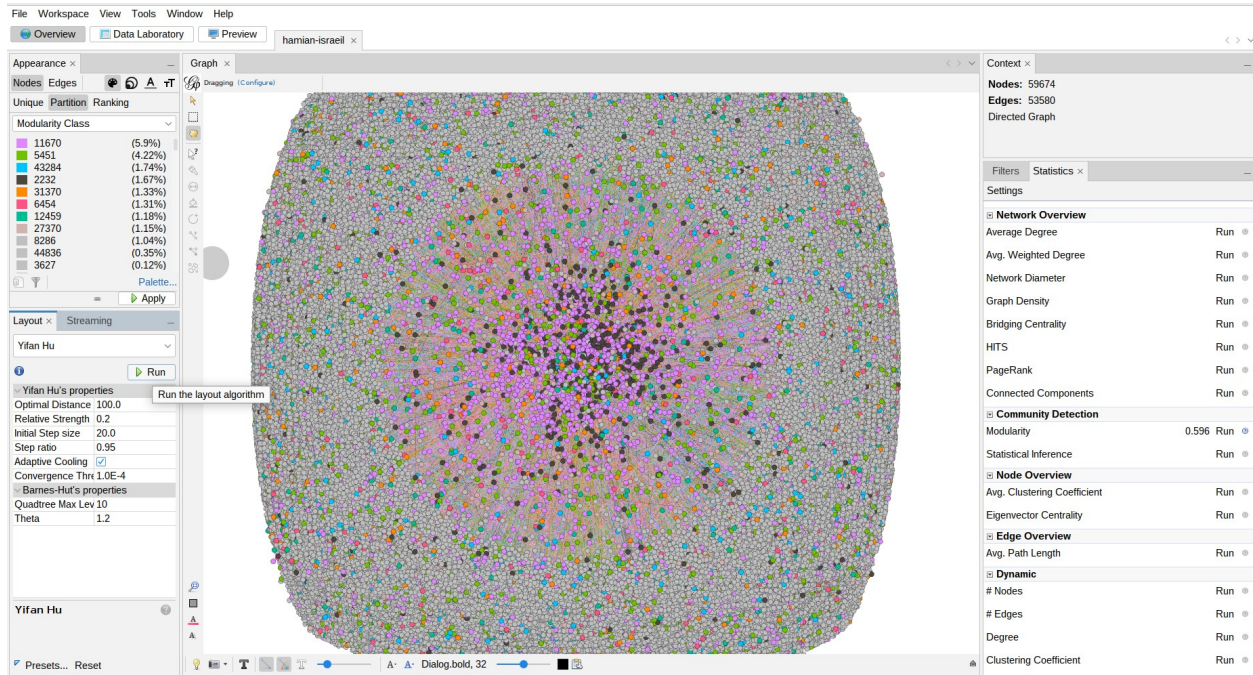
در کد بالا داریم:

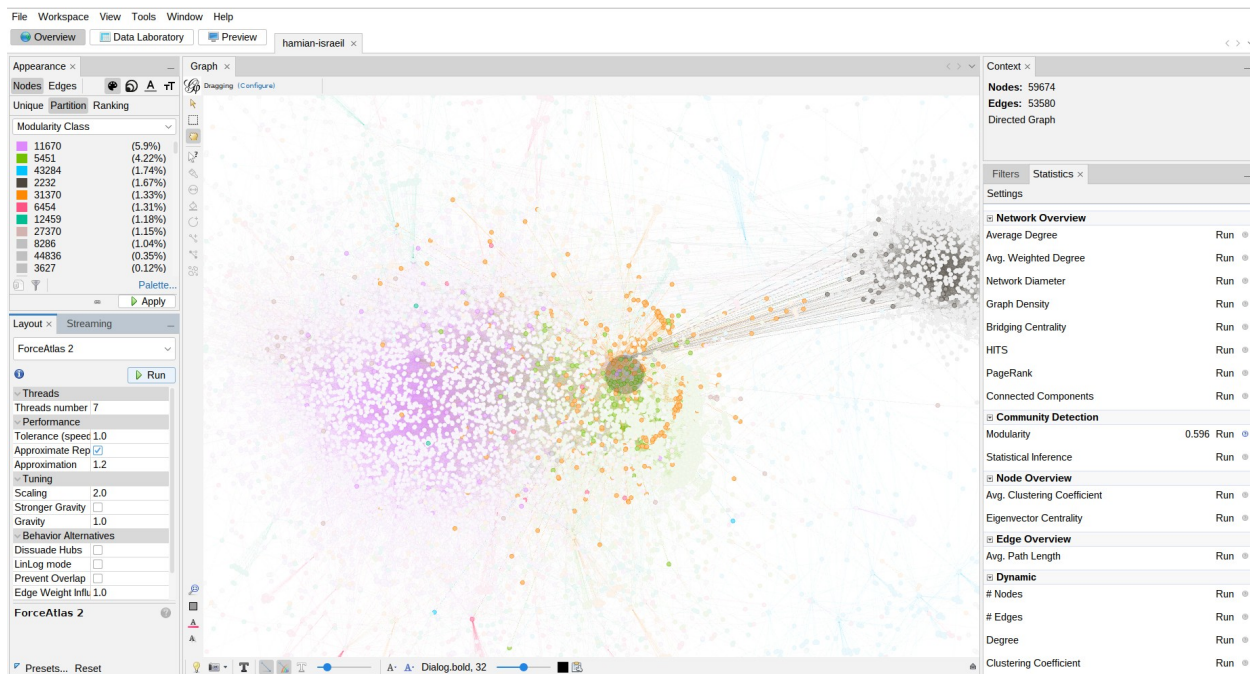
1. محاسبه ماتریس لاپلاسین:
 - ☐ کد از ماتریس لاپلاسین برای تجزیه طیفی استفاده کرده است.
 - ☐ این مرحله درست اجرا شده و مقادیر ویژه کوچکتر استخراج شده‌اند.
2. خوشه‌بندی با استفاده از K-Means:
 - ☐ از بردارهای ویژه متناظر با کوچکترین مقادیر ویژه استفاده شده است.
 - ☐ این بردارها داده‌های گره‌ها را به فضای جدیدی (embedding) نگاشت کرده‌اند و سپس K-Means خوشه‌بندی را انجام داده است.
3. رسم گراف:
 - ☐ گراف با استفاده از طرح‌بندی spring_layout نمایش داده شده است.
 - ☐ خوشه‌های دوگانه با رنگ‌های قرمز و آبی نشان داده شده‌اند.

فرضیات محتمل بر اساس موضوع:

1. خوشه قرمز (حامیان قوی‌تر):
 - ☐ کاربران در این خوشه ممکن است افرادی باشند که به‌طور فعال از اسرائیل حمایت می‌کنند و در این بحث‌ها نقش مرکزی دارند.
 - ☐ این گروه ممکن است شامل افراد تأثیرگذار، سازمان‌ها، یا حساب‌های رسمی مرتبط با موضوع باشد.
2. خوشه آبی (حامیان کم‌رنگ‌تر یا منتقدان):
 - ☐ کاربران در این خوشه ممکن است حاشیه‌ای‌تر باشند یا حتی دیدگاه‌های متضادی داشته باشند.
 - ☐ این گروه می‌تواند شامل کاربران بی‌طرف، کنجکاو، یا حتی مخالف موضوع باشد.

Louvain :





مقدار ماژوالریتی محاسبه شده برابر 0.596 است. این مقدار نشان می‌دهد که گراف به خوبی به جوامع مختلف تقسیم شده است. مقادیر نزدیک به 1 نشان‌دهنده ساختار بسیار ماژولار (تقسیم‌شده) است، و مقادیر پایین‌تر نشان‌دهنده ارتباط بیشتر بین جوامع است.

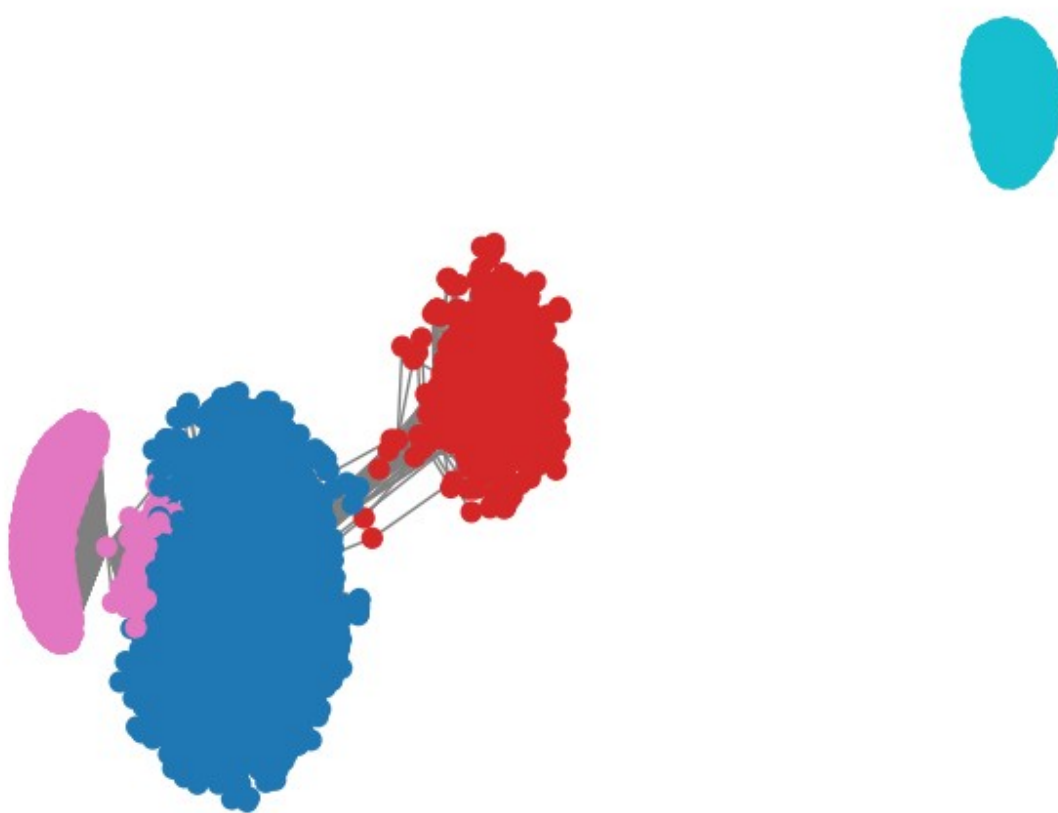
جوامع شناسایی شده به تفکیک رنگ در گراف نمایش داده شده‌اند. هر رنگ نشان‌دهنده یک جامعه (community) است که احتمالاً گروه‌های مختلف از کاربران یا گروه‌های مرتبط به یکدیگر را نمایندگی می‌کنند.

تراکم زیاد در مرکز گراف نشان‌دهنده وجود یک شبکه به هم پیوسته است که کاربران نزدیک به مرکز معمولاً فعال‌ترین و تأثیرگذارترین افراد هستند.

مناطق پراکنده‌تر در حاشیه‌ها احتمالاً نشان‌دهنده کاربران کم‌فعالیت‌تر یا گروه‌هایی است که ارتباط کمتری با مرکز دارند.

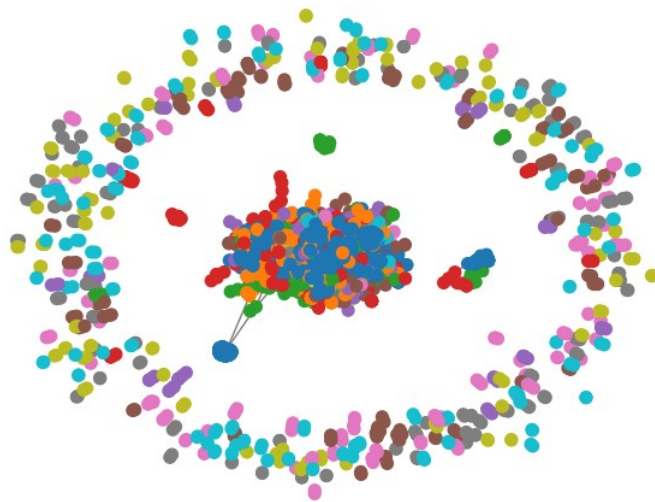
InfoMap:

InfoMap یک الگوریتم شناسایی جوامع (Community Detection) است که بر پایه تئوری اطلاعات کار می‌کند. هدف آن کمینه‌سازی اطلاعات لازم برای توصیف مسیرهای تصادفی (random walks) در گراف است. این الگوریتم گره‌ها را به جوامعی تقسیم می‌کند که مسیرهای تصادفی درون هر جامعه کوتاه‌تر و متمرکزتر هستند. InfoMap برای گراف‌های بزرگ و پیچیده بسیار مؤثر است و می‌تواند ساختارهای سلسله‌مراتبی را نیز شناسایی کند. این روش به دلیل دقت و کارایی بالای خود، به‌ویژه در تحلیل شبکه‌های اجتماعی، زیستی، و ارتباطی کاربرد گسترده‌ای دارد.



به علت سنگین تر بودن محاسبه این قسمت برای رسم انجمن‌ها با پایتون فقط ۴ انجمن برتر را رسم کردم
توی google colab تونستم کامل تر رسم کنم

Graph with Infomap Communities



```
import pandas as pd
import networkx as nx
from cdlib import algorithms
import matplotlib.pyplot as plt

csv_file_path = 'edg.csv'
data = pd.read_csv(csv_file_path)

G = nx.Graph()
edges = list(zip(data['Source'], data['Target']))
G.add_edges_from(edges)

infomap_communities = algorithms.infomap(G)

sorted_communities = sorted(infomap_communities.communities, key=len, reverse=True)[:4]

nodes_to_draw = set(node for community in sorted_communities for node in community)
subgraph = G.subgraph(nodes_to_draw)

colors = plt.cm.get_cmap('tab10', len(sorted_communities)).colors
node_colors = {}
for i, community in enumerate(sorted_communities):
    for node in community:
        node_colors[node] = colors[i % len(colors)]

pos = nx.spring_layout(subgraph)
nx.draw(
    subgraph, pos,
```



```
node_color=[node_colors[node] for node in subgraph.nodes()],
with_labels=False, node_size=50, edge_color="gray"
)
plt.title("Graph with Top 4 Infomap Communities")
plt.show()
```

ساختار مرکزی متراکم:

- مرکز گراف بسیار متراکم است و تعداد زیادی از گره‌ها به صورت فشرده در آنجا قرار گرفته‌اند.
- این نشان می‌دهد که گروهی از کاربران (احتمالاً حساب‌های کاربری تأثیرگذار یا پرتعامل) به شدت در حال بازتویییت یا تعامل با یکدیگر هستند.

انجمن‌های محیطی:

- انجمن‌های کوچک‌تر و پراکنده‌تر در اطراف گراف مشاهده می‌شوند. این‌ها احتمالاً کاربران یا گروه‌هایی هستند که ارتباط کمتری با مرکز دارند و به صورت محدود در گفتگوهای مربوط به حمایت از اسرائیل مشارکت می‌کنند.

تفکیک واضح انجمن‌ها:

- رنگ‌های مختلف برای انجمن‌ها نشان‌دهنده تفکیک مناسبی بین گروه‌های اجتماعی مختلف است. این ممکن است نشان‌دهنده زیرگروه‌هایی باشد که روی موضوعات خاص یا جنبه‌های متفاوت حمایت از اسرائیل تمرکز دارند.

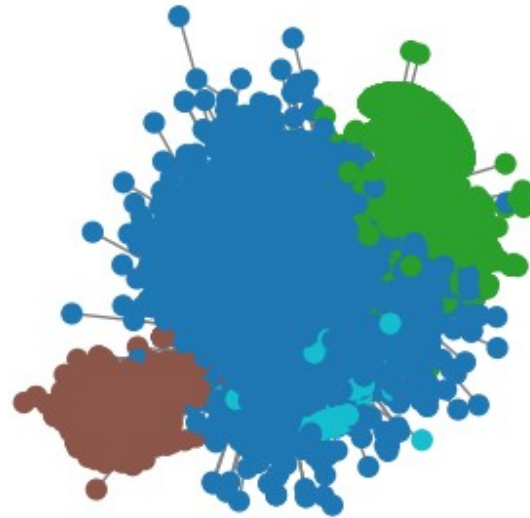
یال‌های متقاطع:

- یال‌هایی که بین گره‌های انجمن‌های مختلف متصل شده‌اند نشان‌دهنده تعاملات محدود بین این انجمن‌هاست. این تعاملات می‌تواند بازتویییت، پاسخ، یا منشن باشد.

مشارکت گسترده:

- توزیع گره‌ها در اطراف گراف نشان‌دهنده این است که کاربران زیادی با درجات مختلف مشارکت در بحث حضور دارند.

Propagation Label:



```
import pandas as pd
import networkx as nx
import matplotlib.pyplot as plt

csv_file_path = 'edg.csv'
data = pd.read_csv(csv_file_path)

if 'Source' not in data.columns or 'Target' not in data.columns:
    raise ValueError("The CSV file must contain 'Source' and 'Target' columns.")

G = nx.Graph()
edges = list(zip(data['Source'], data['Target']))
G.add_edges_from(edges)
```

```

communities = nx.algorithms.community.label_propagation_communities(G)

communities = list(communities)

sorted_communities = sorted(communities, key=len, reverse=True)[:5]

nodes_to_draw = set(node for community in sorted_communities for node in community)
subgraph = G.subgraph(nodes_to_draw)

colors = plt.cm.get_cmap('tab10', len(sorted_communities)).colors
node_colors = {}
for i, community in enumerate(sorted_communities):
    for node in community:
        node_colors[node] = colors[i % len(colors)]

pos = nx.spring_layout(subgraph)
nx.draw(
    subgraph, pos,
    node_color=[node_colors[node] for node in subgraph.nodes()],
    with_labels=False, node_size=50, edge_color="gray"
)
plt.title("Graph with Top 5 Label Propagation Communities")
plt.show()

print(f"Number of communities detected: {len(sorted_communities)}")
for i, community in enumerate(sorted_communities):
    print(f"Community {i+1}: {list(community)}")

```

یک خوشه مرکزی متراکم (آبی و سبز):

- بخش اصلی گراف شامل گره‌هایی است که بسیار به هم متصل هستند.
- این خوشه احتمالاً نشان‌دهنده حساب‌های تأثیرگذار (influencers)، سازمان‌ها یا گروه‌هایی است که ارتباطات و تعاملات گسترده‌ای با دیگر کاربران دارند.
- خوشه‌های آبی و سبز ممکن است نشان‌دهنده دو زیرگروه اصلی در میان حامیان اسرائیل باشند که اهداف یا سبک تعامل متفاوتی دارند.

خوشه‌های کوچک‌تر (قهوه‌ای و خاکستری):

- خوشه قهوه‌ای در نزدیکی خوشه مرکزی قرار دارد. این نشان‌دهنده یک زیرگروه فرعی است که با بخش مرکزی تعامل بیشتری دارد، اما به عنوان یک اجتماع مستقل عمل می‌کند.

● خوشه خاکستری که کاملاً جدا از سایر بخش‌ها قرار دارد، احتمالاً شامل گروهی از کاربران است که مشارکت محدود یا تعاملات بسیار خاصی با دیگر بخش‌های گراف دارند. این گروه ممکن است شامل کاربران بی‌طرف یا کاربران مخالف بحث باشد.

تمرکز و تفکیک:

● بخش مرکزی (آبی و سبز) تراکم بیشتری دارد و تعاملات میان‌گروه‌ها در این بخش بسیار زیاد است.
● تفکیک خوشه‌ها به‌وضوح توسط الگوریتم شناسایی شده، به طوری که هر رنگ نشان‌دهنده یک انجمن مستقل با تعاملات داخلی قوی است.

تعاملات بین انجمن‌ها:

● تعداد محدودی از یال‌هایی که خوشه‌ها را به هم متصل می‌کنند نشان می‌دهد که تعاملات میان‌گروهی (بین انجمن‌ها) در این گراف بسیار محدود است. این موضوع می‌تواند حاکی از جدا بودن گروه‌های مختلف در بحث‌های حمایتی باشد.

اهمیت خوشه‌ها در زمینه حمایت از اسرائیل:

● خوشه آبی و سبز به احتمال زیاد نشان‌دهنده حامیان قوی‌تر اسرائیل هستند، زیرا تعاملات داخلی آن‌ها بسیار زیاد است.
● خوشه قهوه‌ای ممکن است یک گروه متمرکز باشد که نقش حمایتی کمتری دارد یا روی جنبه خاصی از حمایت تمرکز کرده است.
● خوشه خاکستری می‌تواند نشان‌دهنده گروهی از کاربران با تعاملات حاشیه‌ای‌تر باشد یا حتی گروهی باشد که در موضوعات مرتبط با مخالفت یا بی‌طرفی درگیر هستند.

نود ها با importance بالا در همه این انجمن ها:

Iranazadi1395

Iranema2017

esmaili8sahar

simayazaditv

