Key Methods of Hypergraph Analysis Day 3: Clique Representations of Hypergraphs

Moses Boudourides

Moses.Boudourides@gmail.com

instats Seminar

Tuesday, March 11, 2025, from 5:00 PM to 6:30 PM UTC

Table of Contents

Cliques in Graphs

Definition

- A **clique** in an undirected graph $G = (V, E)$ is a subset $C \subseteq V$ such that the subgraph induced by C is a complete graph. In other words, for any distinct vertices $u, v \in C$, there is an edge $(u, v) \in E$.
- The **size** of clique C is its cardinality, $|C|$, i.e., the number of vertices in the clique. Denoting by k-clique a clique of size k, we have the following characterizations:
  - 1-clique is a vertex,
  - 2-clique is a tuple, i.e., an edge,
  - 3-clique is a triplet, i.e., a triangle,
  - 4-clique is a quatruple,
  - 5-clique is a quintuple, etc.
- A **maximal clique** is a clique that is not a proper subset of any other clique.

- **Maximal Clique Cover**: The collection of all maximal cliques forms a cover of the vertex set, meaning that every vertex is contained in at least one maximal clique.
- **Clique Detection Complexity**:The problem of determining whether a graph contains a clique of size at least k is NP-complete.

Clique-Based Hypergraph Representation of Graphs

- The existence of the maximal clique cover of a graph implies that every graph is associated with a bipartite graph, where one partition corresponds to the vertices of the original graph and the other partition corresponds to the maximal cliques of the original graph. If we then consider the bijection between bipartite graphs and hypergraphs, this leads to the following definition:

Definition

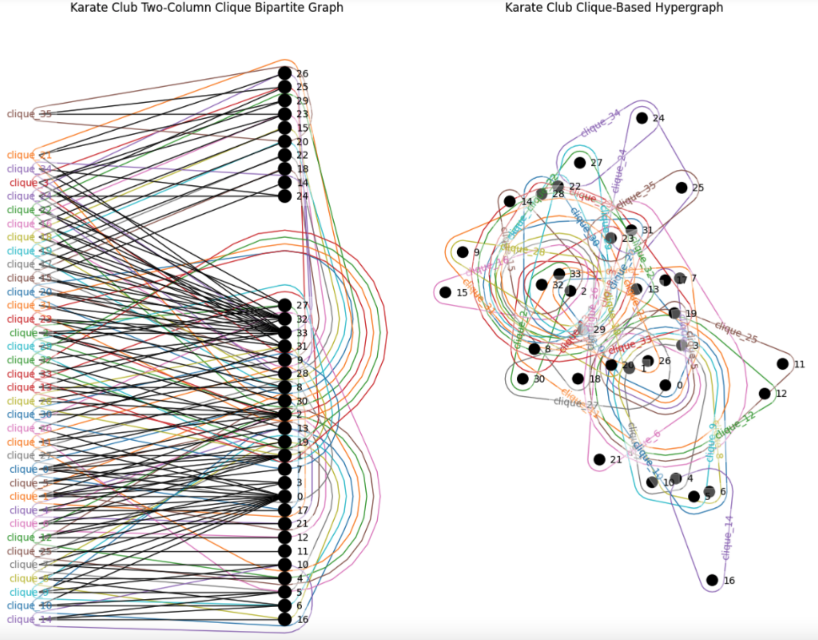
Let $G' = (V, E')$ be a graph and let V_{clique} be its set of maximal cliques. The corresponding **clique-based hypergraph** is (in general) a multiple hypergraph $G = (V, E_{\text{clique}})$ having the same vertex set as in G', i.e., V. Its hyperedge set E_{clique} consists of one hyperedge for each vertex $v \in V$, defined as:

$$E_{\text{clique}} = \{ \bigcup_{C \in V_{\text{clique}}} C : v \in C \}.$$

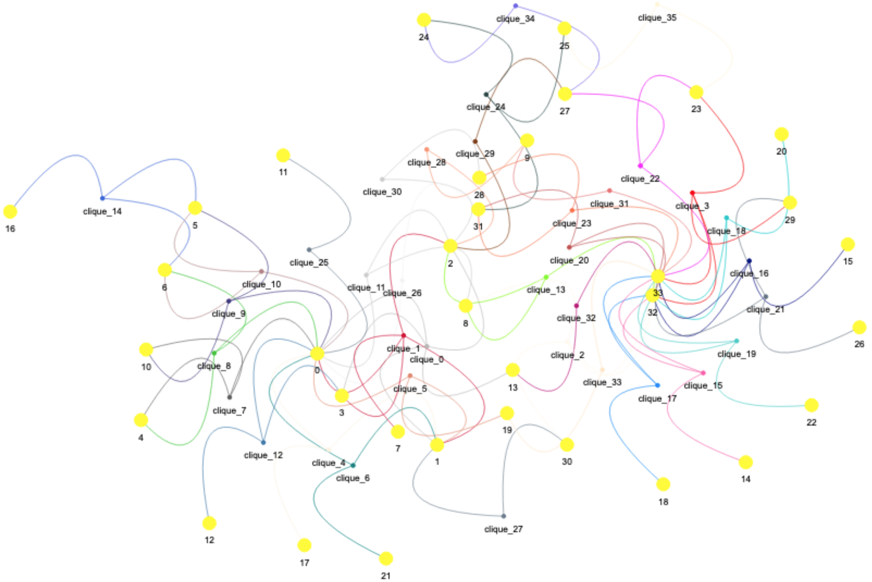
In other words, the hyperedge e_v consists of the union of all maximal cliques that contain vertex v. Stated as an incidence relation,

$$u \in e_v \Leftrightarrow \exists C \in V_{\text{clique}} \text{ such that } v, u \in C.$$

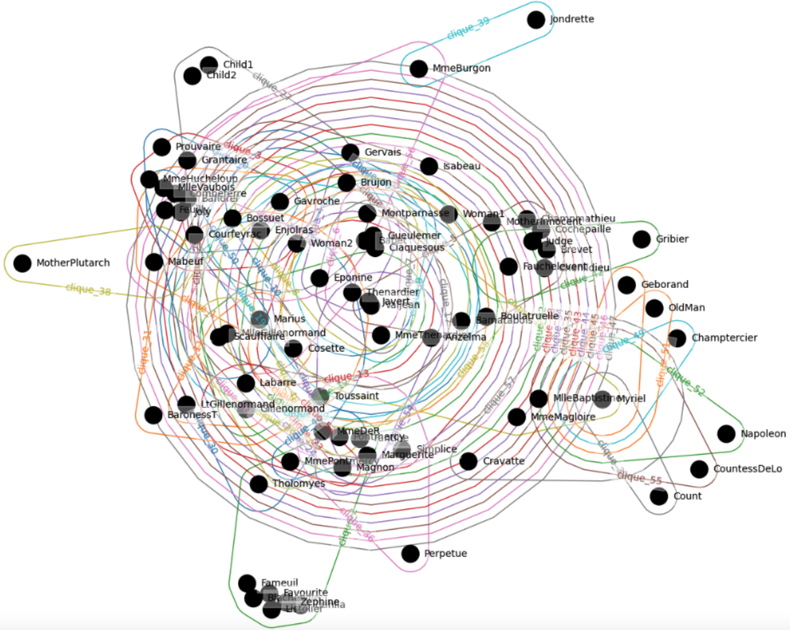
Karate Club Clique-Based Hypergraph



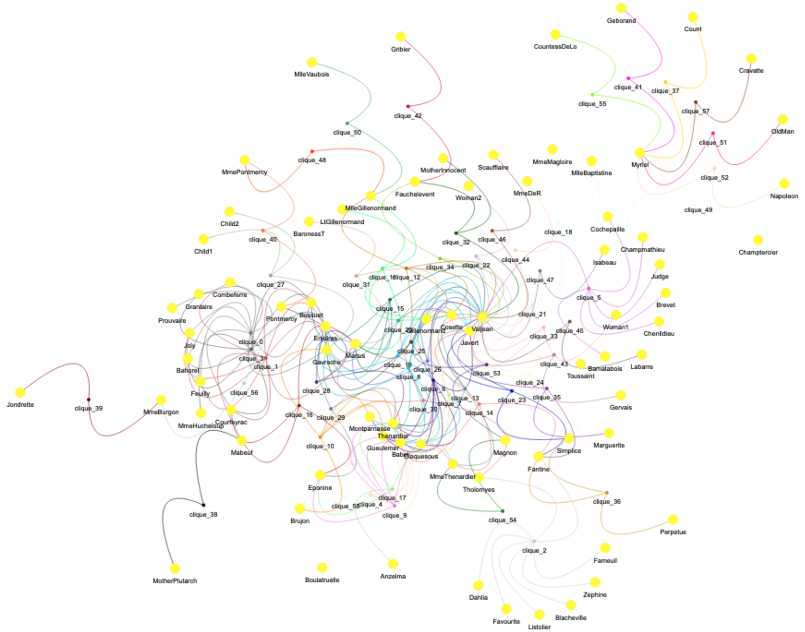
Karate Club Clique-Based Bipartite Plot of Hypergraph



Les Miserables Clique-Based Hypergraph



Les Miserables Clique-Based Bipartite Hypergraph Plot



(Multi–)Ego-Centered Graphs

Definition

Let $G' = (V, E')$ be a graph. If there exists a set of verices denoted as egos such that

$$\bigcup_{\text{ego} \in \text{egos}} \overline{N}(\text{ego}) = V,$$

then G' is called a **ego-centered graph**. The closed neighborhood of a vertex v is denoted and defined as $\overline{N}(v) = \{u \in V : (u, v) \in E'\} \cup \{v\}$. The neighbors of the vertices egos are called **alters**. Note that an ego can also be an `alter` of another ego. Furthermore, for each ego,

$$\overline{N}(\text{ego}) = \bigcup_{\substack{C \in V_{\text{clique}} \\ \text{ego} \in C}} C.$$

In other words, all the `alters` of an ego are contained in the set of maximal cliques that include the ego.

Ego-Clique-Based Hypergraph Representation of Ego-Centered Graphs

Definition

Let $G' = (V, E')$ be an ego-centered graph with a distinguished subset of vertices, called $\texttt{egos} \subset V$,[c] and let V_{clique} be its set of maximal cliques. Since the union of the closed neighborhoods of all egos covers V, we define the simple **ego-clique-based hypergraph** $G = (V_{\text{ego}}, E_{\text{clique}})$, where

- The vertex set is $V_{\text{ego}} = \texttt{egos}$.
- Each ego forms a hyperedge e_{ego} corresponding to its closed neighborhood, i.e., $e_{\text{ego}} = \overline{N}(\text{ego})$.

By construction, each hyperedge consists of all vertices that belong to every maximal clique containing the corresponding ego. The incidence relation is given by:

$$\texttt{alter} \in e_{\text{ego}} \Leftrightarrow \exists C \in V_{\text{clique}} \text{ such that } \text{ego}, \texttt{alter} \in C.$$

[c] Typically, $|\texttt{egos}| < \frac{n}{2}$, where $|V| = n$ is the order of the graph.

Dominating Ego-Clique-Based Hypergraphs

Definition

Let $G = (V, E')$ be an undirected graph. A **decomposition of G into dominating ego-centered subgraphs** is a family of ego-centered subgraphs of G defined by a partition of $V = V_{\texttt{egos}} \cup V_{\texttt{alters}}$ (where vertices in $V_{\texttt{egos}}$ are called egos and vertices in $V_{\texttt{alters}}$ are called alters) such that the set of edges in all these ego-centered subgraphs covers V. In other words, for any $e \in E$, (at least[d]) one of the following five conditions should hold:

- $e = (\text{ego}, \texttt{alter})$, where $\texttt{alter} \in N(\text{ego})$,
- $e = (\texttt{alter}_i, \texttt{alter}_j)$, where $\texttt{alter}_i, \texttt{alter}_j \in N(\text{ego})$,
- $e = (\texttt{alter}_i, \texttt{alter}_j)$, where $\texttt{alter}_i \in N(\text{ego}_r), \texttt{alter}_j \in N(\text{ego}_s)$,
- $e = (\text{ego}_r, \text{ego}_s)$.

[d] Since an ego can also be an alter; note that the partition into egos and alters remains intact, as an alter–ego is always encoded in $V_{\texttt{egos}}$.

Properties of Dominating Ego-Centered Decompositions

- A **minimum dominating set** of graph G is the smallest set of vertices such that every vertex is either in the set or adjacent to a vertex in the set. The **domination number**, denoted $\gamma(G)$, is the size of a minimum dominating set.
- Depending on the algorithm used to construct a dominating ego-centered decomposition on $G = (V, E)$, the size of V_{egos} (i.e., the number of egos in the decomposition) serves as an approximation of the domination number $\gamma(G)$.
- Any two dominating ego-centered subgraphs are either overlapping (i.e., sharing some common vertices) or disjoint (although their vertices–egos and alters–might be connected), but they are never nested (as one being a subset of the other).
- For a given graph, multiple dominating ego-centered decompositions generally exist.
- The following three slides detail our algorithm for constructing a dominating ego-centered decomposition.

Algorithm 1: Dominating Ego-Centered Graph Decomposition (Part 1)

Input : Undirected graph $G = (V, E)$
Output: Tuple (E, A), where E is the list of Egos and A is the list of Alters

Initialize empty lists $E \leftarrow []$, $A \leftarrow []$
Set NL as the list of nodes in V sorted by degree in descending order
while *NL is not empty* **do**
    Shuffle NL randomly
    Select node $u \leftarrow NL[0]$ with the highest degree and remove it from NL
    if $degree(u) > 0$ **then**
        Add u to E
        $N_u \leftarrow$ list of neighbors of u
        $v \leftarrow$ neighbor of u with the highest degree $\leq$ degree of u

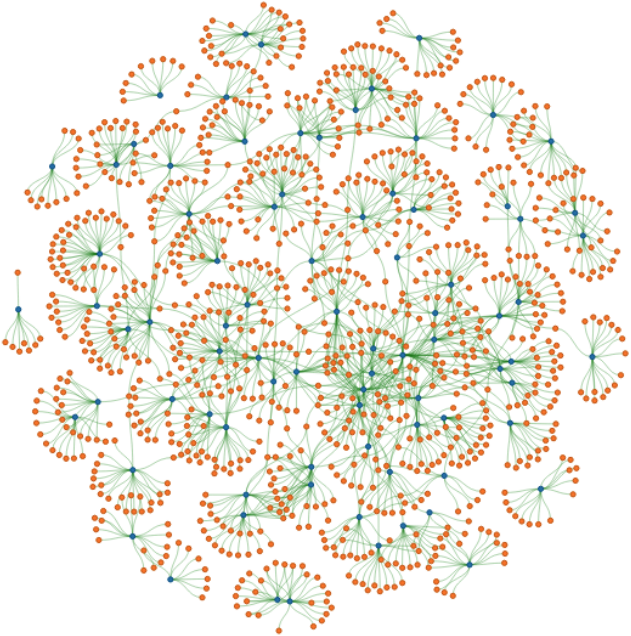
Algorithm 2: Dominating Ego-Centered Graph Decomposition (Part 2)

if *v exists* **then**
    Add all $N_u \setminus \{v\}$ to A
else
    Add all N_u to A
Remove all nodes in A from NL
foreach $u \in A$ **do**
    $N_u \leftarrow$ neighbors of u
    if *all* N_u *are in* A *and none in* E **then**
        Move u from A to E

Algorithm 3: Dominating Ego-Centered Graph Decomposition (Part 3)

```
foreach u ∈ E do
    foreach v ∈ E with u ≠ v and (u, v) ∈ G do
        N_u ← neighbors of u excluding v
        N_v ← neighbors of v excluding u
        if some w ∈ N_u or w ∈ N_v is in E then
            Randomly choose z ∈ {u, v}
            Move z from E to A
        else
            Compute two-hop neighbors T_u and T_v
            if some w ∈ T_u or w ∈ T_v is in E then
                Randomly choose z ∈ {u, v}
                Move z from E to A
        break
foreach u ∈ A do
    if u is a leaf node and its only neighbor is in E then
        continue (keep u in A)
return (E, A)
```

The LivingLab Ego-Clique-Based Hypergraph at Wave 1



The LivingLab Ego-Clique-Based Hypergraph at Wave 2