The background features a complex network graph with red lines connecting green nodes, overlaid on a light blue and white geometric pattern. A small inset image in the top left shows a cluster of orange and red nodes.

Session 3. gSpan: A Pattern Growth Approach

Pattern-Growth Approach

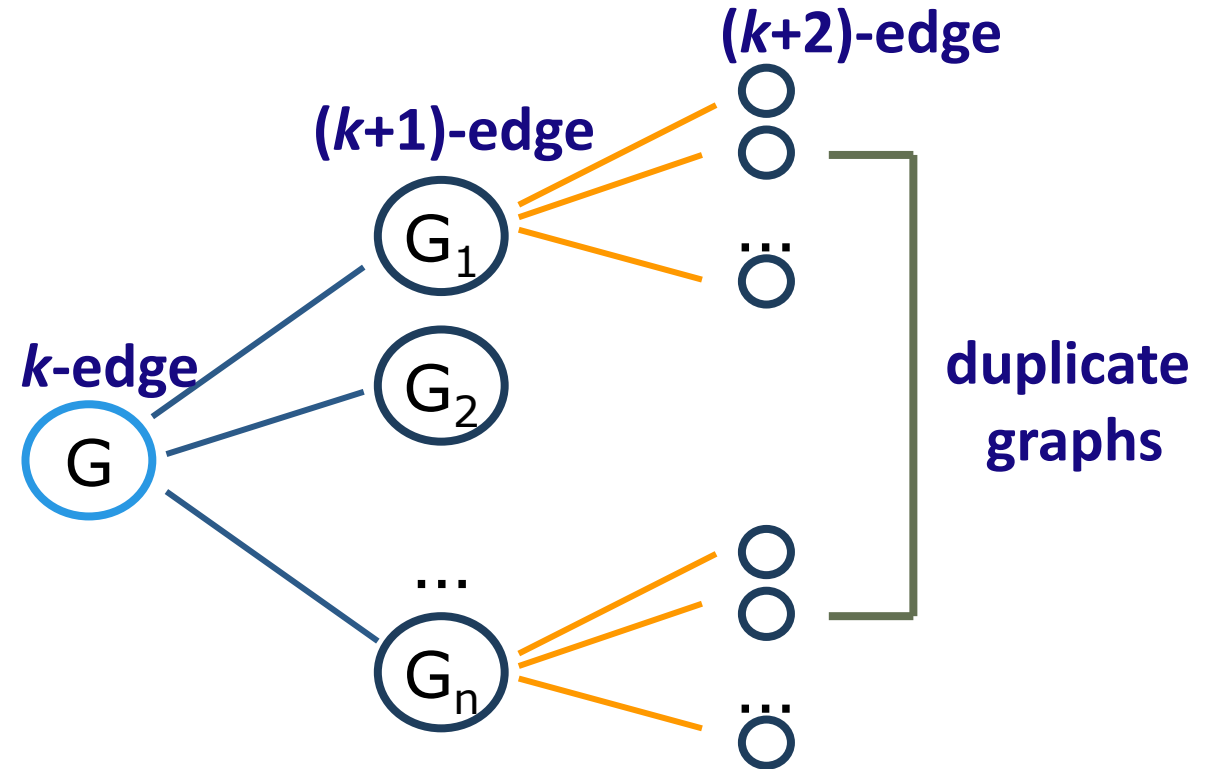
- Depth-first growth of subgraphs from k -edge to $(k+1)$ -edge, then $(k+2)$ -edge subgraphs

- Major challenge

- Generating many duplicate subgraphs

- Major idea to solve the problem

- Define an order to generate subgraphs
 - DFS spanning tree: Flatten a graph into a sequence using depth-first search
 - gSpan (Yan & Han: ICDM'02)



gSPAN: Graph Pattern Growth in Order

- ❑ **Right-most path extension** in subgraph pattern growth
 - ❑ Right-most path: The path from root to the right-most leaf (choose the vertex w. the smallest index at each step)
 - ❑ Reduce generation of duplicate subgraphs
- ❑ **Completeness:** The Enumeration of graphs using right-most path extension is complete
- ❑ DFS Code: Flatten a graph into a sequence using depth-first search

