QAP-based heuristics can be used off-the-shelf as examples with quasimetric edit costs and attributes $a, \beta \in \mathbb{R}_{\geq 0}$:

- edit path cost: $c(P) = \sum_{i=1}^{m} c(a_i)$
- example with quasimetric edit costs and attributes

**Baseline Transformation**

- edit path $\equiv$ perfect matching:

- as assignment matrix: $X = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}$

**Compact Transformation for Quasimetric Edit Costs**

- key property for quasimetric edit costs:
  - exactly $\max(0, |V_C| - |V_H|)$ node removals in optimal edit path
  - exactly $\max(0, |V_H| - |V_C|)$ node insertions in optimal edit path
- edit path $\equiv$ maximal matching:

**Contributions**

- graph edit distance (GED): minimal cost of transforming one graph into another by substituting, removing, and inserting nodes and edges
- widely used in Pattern Recognition community but NP-hard to compute
- one state of the art approach [1–3]:
  1. transform GED to instance of quadratic assignment problem (QAP)
  2. use well-performing heuristics for QAP for approximating GED
- our assumption: edit costs are quasimetric, i.e., satisfy triangle inequality
- our contributions:
  1. reduce size of QAP-instance constructed by the transformation
  2. speed up QAP-based heuristics by using the smaller instances

**Quadratic Assignment Problem (QAP)**

- QAP($C$) := $\min_{C \in \mathbb{R}^{M \times N}} \sum_{(i,j) \in E} c_{ij}$
- cost matrix $C \in \mathbb{R}^{M \times N}$, assignment matrix $X \in \{0, 1\}^{M \times N}$
- sizes of instances constructed by transformations from GED:
  - baseline transformation [1]: $N = M = |V_C| + |V_H|$,
    - first improvement [2]: $N = |V_C| + 1, M = |V_H| + 1$;
    - non-standard version of QAP $\neq$ QAP-based heuristics must be adapted
  - our transformation: $N = |V_C|, M = |V_H|$; uses standard version of QAP like baseline $\sim$ QAP-based heuristics can be used off-the-shelf

**Graph Edit Distance (GED)**

- GED($G, H$) := $\min \{c(P) | P \text{ edit path between } G \text{ and } H\}$
- $G = (V_C, E_C)$ and $H = (V_H, E_H)$ are attributed graphs
- $P = (a_1, \ldots, a_m)$ is sequence of edit operations transforming $G$ into $H$
- edit operations and edit costs ($i \in V_C, k \in V_H, (i, j) \in V_C, (k, l) \in E_H$):
  - node substitution: $c_{i,j} = c_{i,j}$
  - node insertion: $c_{i,j} = c_{i,j}$
  - node removal: $c_{i,j} = c_{i,j}$
  - edge insertion: $c_{i,j} = c_{i,j}$
  - forbidden assignment cost: $c_{i,j} = \infty$
  - free assignment cost: $c_{i,j} = 0$

**Experiments**

- tested QAP-based heuristic: mlPFPP (conditional gradient descent for QAP) [3], best available QAP-based heuristic for GED
- compared transformations: baseline [1], non-standard [2], transformation proposed in this paper
- metrics: computed distance ($d$), error ($e$), runtime in seconds ($t$)

References


Presented at ICPR 2018, Beijing, China, August 20–24, 2018