Figure 5 - Pseudo code for Pro-CLARANS algorithm

**Input:** A training set $D$, $D = \{O_h\}_{h=1..n}$; $n$ is the size of $D$

**Initialize:** $f(V)_{max} = 0$; iteration = 0;

**Repeat**

1. Set $C$ an arbitrary node from $D$; ($C = [R_1, R_2, ..., R_k]$)

2. Set $j = 1$;

3. Repeat
   - Consider a random neighbor $C^*$ of $C$;
   - Compute $TS_{ih}$ of $C^*$ and $TS'_{ih}$ of $C$;
   - If $TS_{ih}$ > $TS'_{ih}$ then
     - $C = C^*$;
     - $j = 1$;
   - Else
     - $j = j + 1$;

4. For each object $O_h \in D$ do
   - Compute the similarity score of $O_h$ with each medoid $R_i$ ($i \in [1..K]$), using Smith Waterman algorithm;
   - Assign $O_h$ to the cluster with the nearest $R_i$;

5. Compute $f(V)$;

6. If $f(V) \leq f(V)_{max}$ then
   - iteration = iteration + 1;
   - Else
     - $f(V)_{max} = f(V)$;
     - BestSets = CurrentSets;
     - Go back to Step3;

**Until** iteration = $q$;

**End**

**Output:** BestSets; BestSets is the best partition of $D$ into $K$ clusters; each cluster is defined by a medoid $R_i$