Labeled graphs' vertices and edges sets clustering

Andrey Kotenko, Maksim Bukarenko
Samara State Technical University (Applied Mathematics and Information Technologies Dept.), Samara State Aerospace University (Math. Methods in Economics Dept.), Samara, Russia

Let there be given a connected indirected loop-free graph $G(V, R)$ with vertices $\nu_i \in V$, $i \in \overline{1, n}$, $n := |V| < \infty$, and edges $r_i \in R$, $i \in \overline{1, m}$, $m := |R| < \infty$, labeled with nonnegative labels $|\nu_i| \geq 0$ and $|r_i| \geq 0$ correspondently.

Let us consider the problem of partition of V set of vertices of G graph into disjoint clusters $U_i \subset V$, $i \in \overline{1, k}$, $k < \infty$ with fixed centers $u_i \sim U_i$, $u_i \in V$ and minimality condition for the distance between the vertex and the center of corresponding cluster. The distance $\rho^V(\nu_i, \nu_j)$ between the vertices $\nu_i, \nu_j \in V$ is defined as the minimal labels sum for the edges which constitute the path connecting these vertices. Note that introduced distance ρ^V satisfies metric separation axiom if and only if there are no edges with zero labels. Let us call conformal the introduced clustering criterion for G graph vertices and ρ^V metric.

Inverse problem is given as follows: to label the edges $r \in R$ of connected indirected loop-free finite graph $G(V, R)$ (with a given V vertexes set partition into disjoint clusters) with numeric nonnegative labels $|r| \geq 0$ generating conformal ρ^V metric. It has trivial solution: zero labels of the edges, incident with one cluster vertices, and unit labels of the rest of the edges generate conformal metric without separation axiom. The solution with separation axiom is also possible: it is enough to label the edges incident with one cluster vertices with sufficiently small labels and the rest of the edges with sufficient large ones. Accurate estimates of these labels depend on the structure of the clusters and V set of vertexes.

Thus partition $G(V, R)$ graph vertices set into clusters is equivalent to labels identification for the edges generating ρ^V conformal metric.

Transposing V set of vertices and R set of edges of $G(V, R)$ graph in the presented rule for clusters conforming we come to the similar conclusion regarding R set of edges partition into disjoint clusters: it is equivalent to labels identification for the edges generating ρ^V conformal metric.

Genuinely, we shall carry out the partition of R set of edges into disjoint clusters $W_i \subset R$, $i \in \overline{1, l}$, $l < \infty$, with fixed centers $w_i \sim W_i$, $w_i \in R$ and minimality condition for the distance between the edge and the center of corresponding cluster, according to the given labels $|\nu_i| \geq 0$ of vertices $\nu_i \in V$. The distance $\rho^R(r_i, r_j)$ between the edges $r_i, r_j \in R$ is defined as a minimal sum of the labels of the vertices on the path connecting these edges. Note that introduced distance ρ^R satisfies metric separation axiom if and only if there are no vertices with zero labels. Let us call conformal the introduced clustering criterion for G graph edges and ρ^R metric. The inverse problem is the identification of the vertices labels set generating the conformal metric ρ^R according to the given partition of R set of edges into clusters. It is solved in a similar way.

The fact that vertices clustering problem with edges labeled and edges clustering problem with vertices labeled are symmetric ones requires to develop a universal algorithm for corresponding ρ^V and ρ^R metrics calculation in vertexes and edges sets of $G(V, R)$ graph. Matrix modification of Bellman-Moore algorithm [1] which enables simple software implementation can be suggested for this purpose.

References

[1] A.P. Kotenko. Bellman-Moore matrix algorithm [in Russian]. *Business Systems Management* **10** (2013) 33-37.