社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2006－AL－108 (7)
2006／9／27

データ結合問題で現れる多次元割当問題の近似解法

黒木 裕介[1] 松井 知己[2]

[1] 東京大学大学院 情報理工学系研究科
[2] 中央大学 理工学部 情報工学科

概要. 完全 k 部グラフ $G=(V_1,V_2,\ldots,V_k;E)$ で $|V_1|=|V_2|=\cdots=|V_k|=n$ を満たすものと, G の k クリークすべてに対して重みが与えられたとき, k 次元割当問題とは, G の点集合の (互いに交わりのない) n 個の k クリークへの分割で, 分割を構成するクリークの重みの総和が最小になるものを求める問題である. 本稿では, 各クリークの重みはクリークで誘導される辺の重みの総和で表される場合を考える. さらに, G の点は d 次元空間 $\mathbb{Q}^d$ に埋め込まれているものとし, 辺の重みは辺の両端点間のユークリッド距離の平方で定めるとする. 問題例は, データ結合問題の多次元ガウシアンモデルにおいて現れる.
本問題に対する二次錐計画問題緩和と多項式時間乱択丸め手続きを提案する. 提案するアルゴリズムにより得られる目的関数値の期待値は, 最適値の $(5/2-3/k)$ 倍で抑えられることを示す. この結果は, いままで知られていた近似率 $(4-6/k)$ を改善した.

Approximation Algorithm for Multidimensional Assignment Problem Arising from Data Association Problem

Yusuke Kuroki[1] and Tomomi Matsui[2]

[1] Graduate School of Information Science and Technology, the University of Tokyo
[2] Department of Information and System Engineering, Faculty of Science and Engineering, Chuo University

Abstract. Given a complete k-partite graph $G=(V_1,V_2,\ldots,V_k;E)$ satisfying $|V_1|=|V_2|=\cdots=|V_k|=n$ and weights of all k-cliques of G, the k-dimensional assignment problem finds a partition of vertices of G into a set of (pairwise disjoint) n k-cliques which minimizes the total sum of weights of the chosen cliques. In this paper, we consider a case that the weight of a clique is defined by the sum of given weights of edges induced by the clique. Additionally, we assume that vertices of G are embedded in the d-dimensional space $\mathbb{Q}^d$ and a weight of an edge is defined by the square of the Euclidean distance between two end vertices. We describe that these problem instances arise from a multidimensional Gaussian model of a data association problem.
We propose a second-order cone programming relaxation of the problem and a polynomial time randomized rounding procedure. We show that the expected objective value obtained by our algorithm is bounded by $(5/2-3/k)$ times the optimal value. Our result improves the previously known bound $(4-6/k)$ of the approximation ratio.

1 Introduction

Let $\mathcal{F}=\{V_1,V_2,\ldots,V_k\}$ be a family of vertex sets satisfying $|V_1|=|V_2|=\cdots=|V_k|=n$. A complete k-partite graph $G=(V_1,V_2,\ldots,V_k;E)$ is defined by vertex sets $V_1,V_2,\ldots,V_k$

and an edge set $E = \bigcup_{\{U,V\}\in\binom{\mathcal{F}}{2}}\{\{u,v\} \mid u \in U,\ v \in V\}$. A vertex subset Q is called a *clique* (*q-clique*) of G if and only if the complete graph induced by Q is a subgraph of G (and $q = |Q|$). Given weights of all k-cliques of G, the k-dimensional assignment problem finds a partition of vertices of G into a set of (pairwise disjoint) n k-cliques which minimizes the total sum of weights of the chosen n k-cliques.

We introduce following definitions and assumptions. For any clique Q of G, every edge connecting two vertices in Q is called *clique edge* of Q. Given an edge weight vector $w \in \mathbb{R}^E$, we define a weight of a clique Q by the sum of weights of clique edges of Q. Additionally, we assume that vertices of G are embedded in the d-dimensional space and the weight of an edge is defined by the square of the Euclidean distance between two end vertices. In the rest of this paper, we assume that the input of the problem is k n-sets $V_1, V_2, \ldots, V_k$ of rational d-dimensional vectors (i.e., $V_1, V_2, \ldots, V_k \subseteq \mathbb{Q}^d$ and $|V_1| = |V_2| = \cdots = |V_k| = n$). These problem instances arise from a multidimensional Gaussian model of a data association problem, described in the next section.

In this paper, we propose a second-order cone programming relaxation of the problem and a polynomial time randomized rounding procedure. We show that the expected objective value obtained by our algorithm is bounded by $(\frac{5}{2} - \frac{3}{k})$ times the optimal value. Our result improves the previously known bound $(4 - \frac{6}{k})$ of the approximation ratio, which is obtained by Bandelt, Crama and Spieksma in [2].

When $k = 2$, the k-dimensional assignment problem is the well-known assignment problem and can be solved by Hungarian method. The 3-dimensional assignment problem has been actively investigated. When weights of all the 3-cliques are arbitrary, the problem is a generalization of 3-dimensional matching (3DM) and thus NP-hard [8]. NP-hardness of some subclasses is discussed in papers [3, 6, 17]. When edge weights satisfy triangle inequalities, Crama and Spieksma [6] showed that a simple heuristic gives a 4/3-approximation algorithm. For values $k \geq 4$, the k-dimensional assignment problem has been less studied. Early mention of the problem can be found in Haley [9] and Pierskalla [12]. Bandelt, Crama and Spieksma [2] considered cases where the weights of cliques are not arbitrary, but given as a function of edge weights. When edge weights satisfy triangle inequalities and the weight of a clique is defined by the sum of weights of edges induced by the clique, they showed that there exists $(2 - 2/k)$-approximation algorithm. We briefly describe (a modified version of) their algorithm and its approximation ratio in Section 3.1. For more detained references, see recent survey papers [4] by Burkard and Çela and [18] by Spieksma.

Multidimensional assignment problem arises from many application areas. Pierskalla [11, 12] mentioned some application settings: capital investment, dynamic facility location, satellite launching. Other applications appear in Frieze and Yadegar [7] and Crama et al. [5]. Recently, multidimensional assignment problem have found applications as a technique to solve data association problems. For example, in multitarget multisensor surveillance systems, we need to associate reports from multisensor to enhance target identification and state estimation. General classes of these problems can be formulated as multidimensional assignment problems [13, 14]. Another example is the integration of market databases. When there is no single source database available about all the information of interest, techniques of integrating different databases are often applied. By integrating multiple source market survey data, the obtained single data-set will have answers to all the questions in original surveys. A class of integration methods is known as data-fusion procedures or statistical matching [15]. In [16], Soong and de Montigny studied a problem instance of fusing three databases.

2 Data Association Problem

In this section, we show that our model of multidimensional assignment problem arises from a simple probabilistic framework of the data association problem. Assume that there are n objects (targets, randomly chosen customers, etc.) and k data-sets (observations obtained by radar or global positioning system, results of questionnaires, etc.) such that each data-set consists of n reports (observations) corresponding to n objects. For fusing k data-sets, we need to find a partition of all the reports into (pairwise disjoint) n k-sets such that each subset of reports meets every data-set in exactly one report, because we do not know the correspondence (matching) between reports for any pair of data-sets. We assume that each report of object i may be independently and identically distributed from d-dimensional normal distribution $\mathrm{N}(\boldsymbol{\theta}_i, \Sigma)$. In the following, we assume that Σ is the d-dimensional identity matrix for simplicity. When we have k reports $Q = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^d$ of object i, it is well-known that the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}_i$ is the center of gravity $(1/k)\sum_{\boldsymbol{v}\in Q}\boldsymbol{v}$, since Σ is the identity matrix, and the corresponding loglikelihood is

$$-C_0 \sum_{\boldsymbol{v}\in Q} \|(1/k)\sum_{\boldsymbol{v}'\in Q}\boldsymbol{v}' - \boldsymbol{v}\|^2 + C_1 = -(C_0/k)\sum_{\{\boldsymbol{v},\boldsymbol{v}'\}\in\binom{Q}{2}} \|\boldsymbol{v} - \boldsymbol{v}'\|^2 + C_1$$

where C_0, C_1 are constants with $C_0 \geq 0$. Given a partition $\{Q_1, Q_2, \ldots, Q_n\}$ of all the reports such that each set corresponds to reports from a common potential object, the MLE of the set of n parameters are $\{(1/k)\sum_{\boldsymbol{v}\in Q_i}\boldsymbol{v} \mid i \in \{1, 2, \ldots, n\}\}$ and the corresponding conditional loglikelihood is

$$-(C_0/k)\sum_{i=1}^{n}\sum_{\{\boldsymbol{v},\boldsymbol{v}'\}\in\binom{Q_i}{2}} \|\boldsymbol{v} - \boldsymbol{v}'\|^2 + nC_1. \tag{1}$$

From the above, we can find the MLE of the set of n parameters by the following two steps; first, solve the multidimensional assignment problem and find pairwise disjoint n k-cliques $\{Q_1, Q_2, \ldots, Q_n\}$ which maximizes the loglikelihood (1); second, for each subset in $\{Q_1, Q_2, \ldots, Q_n\}$, output the center of gravity of contained reports. In this model, we need to solve multidimensional assignment problem, which minimizes sum total of weights of clique edges, under the assumptions that vertices of G are embedded in the d-dimensional space $\mathbb{R}^d$ and the weight of an edge is defined by the square of the Euclidean distance between two end vertices.

3 Formulations and Relaxations

In this section, we formulate the multidimensional assignment problem as an integer linear programming problem and (integer) quadratic programming problems. Lastly, we combine our formulations and give a second-order cone programming relaxation.

In the rest of this paper, we denote the vertex set $V_1 \cup \cdots \cup V_k$ by $\widehat{V}$. For any vertex subset $V \subseteq \widehat{V}$, $\delta(V)$ denotes the set of edges in E between V and $\widehat{V} \setminus V$. For any disjoint pair of vertex subsets $U, V \subseteq \widehat{V}$, we denote the edge subset $\delta(U) \cap \delta(V)$ by $E(U, V)$ and/or $E(V, U)$. We denote a singleton $\{v\}$ by v for simplicity, when there is no ambiguity. A sequence (e_1, e_2, e_3) of edges of G is called a *triangle* of G if the graph induced by edges $\{e_1, e_2, e_3\}$ is a 3-cycle in G. For any vector $\boldsymbol{x} \in \mathbb{R}^E$ and an edge $\{u, v\} \in E$, we denote the element $x(\{u, v\})$ by $x(u, v)$ and/or $x(v, u)$ for short.

3.1 Integer Linear Programming

We introduce 0-1 valued variable vector $\boldsymbol{x} \in \{0, 1\}^E$. For an arbitrary edge weight vector $\boldsymbol{w} \in \mathbb{R}^E$, we can formulate the multidimensional assignment problem as follows,

$$\begin{array}{lll} \text{ILP: min.} & \sum_{e \in E} w(e)x(e) & \\ \text{s. t.} & \sum_{u \in U} x(u, v) = 1 & (\forall U \in \mathcal{F},\ \forall v \in \widehat{V} \setminus U), \quad (2) \\ & x(e_1) \geq x(e_2) + x(e_3) - 1 & (\text{for each triangle } (e_1, e_2, e_3) \text{ of } G), \quad (3) \\ & x(e) \in \{0, 1\} & (\forall e \in E). \end{array}$$

We show the correctness of the above formulation. For any $\boldsymbol{x} \in \{0, 1\}^E$, we define an edge subset $E(\boldsymbol{x})$ by $\{e \in E \mid x(e) = 1\}$. Let $\boldsymbol{x}$ be a feasible solution of ILP. Then, for any pair $\{U, V\} \in \binom{\mathcal{F}}{2}$, constraints (2) imply that the edge subset $E(\boldsymbol{x}) \cap E(U, V)$ is a perfect matching of the bipartite graph $(U, V; E(U, V))$. Constraints (3) mean that if $[x(e_2) = 1 \text{ and } x(e_3) = 1]$, then $x(e_1) = 1$. Thus constraints (2) and (3) yield that each connected component of $(\widehat{V}, E(\boldsymbol{x}))$ contains k-clique. Since $E(\boldsymbol{x})$ contains $n(1/2)k(k-1)$ edges, the subgraph $(\widehat{V}, E(\boldsymbol{x}))$ consists of pairwise disjoint n k-cliques. The inverse implication is clear.

When we drop constraints (3), we can decompose the obtained problem, denoted by LP, into $(1/2)k(k-1)$ subproblems each of which is a classical assignment problem defined on a bipartite graph $(U, V; E(U, V))$ for a pair $\{U, V\} \in \binom{\mathcal{F}}{2}$. Thus we can solve LP by applying well-known Hungarian method to each subproblem. In the following, we briefly describe a randomized version of *multiple-hub heuristic* proposed by Bandelt, Crama and Spieksma in [2]. First, we solve the relaxation problem LP and obtain a 0-1 valued optimal solution $\boldsymbol{x}^{\mathrm{LP}}$. Next, we choose a vertex subset $U \in \mathcal{F}$ randomly. Lastly, we construct a graph $G' = (\widehat{V}, \delta(U) \cap E(\boldsymbol{x}^{\mathrm{LP}}))$ and output a family of vertex subsets $\{Q_1, Q_2, \ldots, Q_n\}$ of connected components in G'. Since each connected component in G' is a complete bipartite graph $\mathrm{K}_{1,k-1}$ and meets every $V \in \mathcal{F}$ in exactly one vertex, the obtained vertex subsets $Q_1, Q_2, \ldots, Q_n$ are pairwise disjoint k-cliques of G. The obtained solution corresponds to a feasible solution $\boldsymbol{X}^{\mathrm{LPR}}$ of ILP defined by

$$X^{\mathrm{LPR}}(e) = \begin{cases} x^{\mathrm{LP}}(e) & (\forall e \in \delta(V)), \\ \sum_{u \in U} x^{\mathrm{LP}}(u,v) x^{\mathrm{LP}}(u,v') & (\forall e = \{v,v'\} \in E \setminus \delta(U)). \end{cases}$$

Results of Bandelt, Crama and Spieksma [2] imply the following. Under the assumptions that (i) edge weights are non-negative and

$$\text{(ii) } \exists \tau \geq 1/2, \text{ for each triangle } (e_1, e_2, e_3) \text{ of } G,\ w(e_1) + w(e_2) \geq (1/\tau) w(e_3), \tag{4}$$

the expectation of the objective function value of X^{LPR} satisfies that

$$\mathrm{E}\left[\sum_{e \in E} w(e) X^{\mathrm{LPR}}(e)\right] \leq (2/k)((k-2)\tau + 1) z^*(\mathrm{ILP})$$

where $z^*(\mathrm{ILP})$ is the optimal value of ILP. Since we deal with the case that the weight of an edge is defined by the square of the Euclidean distance between the end points, property (4) is satisfied by setting $\tau = 2$. Thus, approximation ratio of the above algorithm is bounded by $(4 - 6/k)$ for our case.

3.2 Integer Quadratic Programming

In this subsection, we reformulate ILP as an integer programming problem with a convex quadratic objective function. We also fix a subset $U \in \mathcal{F}$ throughout this subsection. In the rest of this section, we use the assumption that vertices in $\widehat{V}$ are embedded in $\mathbb{Q}^d$ and the weight of an edge is defined by the square of the Euclidean distance between two end points. For any vertex $v \in \widehat{V}$, we denote the position (in $\mathbb{Q}^d$) of v by $\boldsymbol{v} \in \mathbb{Q}^d$. For any clique Q of G, we denote the weight of Q by $w(Q)$.

Let $\boldsymbol{x} \in \{0,1\}^E$ be a feasible solution of ILP and $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_n\}$ be the set of n k-cliques. For any vertex $u \in U$, $Q(u)$ denotes a unique clique in $\mathcal{Q}$ including u. The objective function value of ILP with respect to $\boldsymbol{x}$ is sum total of clique weights and thus equal to

$$\sum_{u \in U} w(Q(u)) = \sum_{e \in \delta(U)} w(e) x(e) + \sum_{u \in U} w(Q(u) \setminus u).$$

For any vertex $u \in U$, the clique $Q(u)$ meets every subset $V \in \mathcal{F} \setminus \{U\}$ in exactly one vertex in the singleton $Q(v) \cap V$, whose position (in $\mathbb{Q}^d$) is denoted by $\sum_{v \in V} x(v,u)\boldsymbol{v}$, because the equality $\sum_{v \in V} x(v,u) = 1$ holds. For any pair $\{V, V'\} \in \binom{\mathcal{F} \setminus \{U\}}{2}$, the clique $Q(u)$ has a unique clique edge in $E(V,V')$ connecting vertices in $Q(u) \cap V$ and $Q(u) \cap V'$. Thus the weight of the edge is equal to

$$\left\| \sum_{v \in V} x(v,u)\boldsymbol{v} - \sum_{v' \in V'} x(v',u)\boldsymbol{v}' \right\|^2.$$

From the above, sum total of clique weights, $\sum_{u \in U} w(Q(u))$, is given by

$$\sum_{e \in \delta(V)} w(e)x(e) + \sum_{u \in U} \sum_{\{V,V'\} \in \binom{\mathcal{F} \setminus \{U\}}{2}} \left\| \sum_{v \in V} x(v,u)\boldsymbol{v} - \sum_{v' \in V'} x(v',u)\boldsymbol{v}' \right\|^2.$$

Employing the above function, we obtain the following integer quadratic programming formulation of our problem;

IQP(U):

$$\begin{aligned}
\text{min.}\ & \sum_{e\in\delta(U)} w(e)x(e) + \sum_{u\in U} \sum_{\substack{\{V,V'\}\\ \in\binom{\mathcal{F}\setminus\{U\}}{2}}} \left\| \sum_{v\in V} x(v,u)\boldsymbol{v} - \sum_{v'\in V'} x(v',u)\boldsymbol{v}' \right\|^2 \\
\text{s. t.}\ & \textstyle\sum_{u\in U} x(u,v) = 1 \quad (\forall v \in \widehat{V}\setminus U), \\
& \textstyle\sum_{v\in V} x(u,v) = 1 \quad (\forall u \in U,\ \forall V \in \mathcal{F}\setminus\{U\}), \\
& x(e) \in \{0,1\} \quad (\forall e \in \delta(U)).
\end{aligned}$$

An advantage of this formulation is that the objective function is a convex quadratic function. Thus the continuous relaxation problem, obtained by substituting 0-1 constraints with non-negativity constraints, is a convex quadratic programming problem, which is solvable efficiently.

3.3 Second-Order Cone Programming Relaxation

Lastly, we combine formulations ILP and IQP(U) and construct a second-order cone programming (SOCP) relaxation, which hopefully gives a better lower bound. Here we note that we do not fix a vertex subset $U \in \mathcal{F}$ in this subsection. By introducing an artificial variable z, our relaxation problem is described as follows,

SOCPR:

$$\begin{aligned}
\text{min.}\ & z \\
\text{s. t.}\ & z \geq \textstyle\sum_{\{u,v\}\in E} \|\boldsymbol{u}-\boldsymbol{v}\|^2 x(u,v), \\
& z \geq \textstyle\sum_{\{u,v\}\in\delta(U)} \|\boldsymbol{u}-\boldsymbol{v}\|^2 x(u,v) \\
& \quad + \sum_{u\in U} \sum_{\substack{\{V,V'\}\\ \in\binom{\mathcal{F}\setminus\{U\}}{2}}} \left\| \sum_{v\in V} x(v,u)\boldsymbol{v} - \sum_{v'\in V'} x(v',u)\boldsymbol{v}' \right\|^2 \quad (\forall U \in \mathcal{F}), \\
& \textstyle\sum_{u\in U} x(u,v) = 1 \quad (\forall U \in \mathcal{F},\ \forall v \in \widehat{V}\setminus U), \\
& x(e) \geq 0 \quad (\forall e \in E).
\end{aligned}$$

As is well-known, the above problem can be transformed to a second-order cone programming problem, which can be solved within any given gap ϵ in polynomial time by using an interior point method (e.g., see a recent survey paper [1]).

4 Randomized Approximation Algorithm

In this section, we propose a randomized approximation algorithm.

Algorithm 1.
Input: Subsets $V_1, V_2, \ldots, V_k \subseteq \mathbb{Q}^d$ satisfying $|V_1| = |V_2| = \cdots = |V_k| = n$.
Output: A feasible solution $\boldsymbol{X}$ of ILP.
Step 1: Solve SOCPR and obtain an optimal solution $(z^*, \boldsymbol{x}^*)$.
Step 2: Randomly choose a vertex set $U \in \mathcal{F}$.
Step 3: For a subvector $\boldsymbol{x}^*|_{\delta(U)}$ and for each vertex subset $V \in \mathcal{F} \setminus \{U\}$, execute the followings and obtain 0-1 valued vector $\boldsymbol{X}_U$ indexed by $\delta(U)$.
Step 3-1: Represent the subvector $\boldsymbol{x}^*|_{E(U,V)}$ by a convex combination of characteristic vectors of perfect matchings in the bipartite graph $(U, V; E(U, V))$. We denote the coefficient of convex combination with respect to a perfect matching M by $\lambda(M)$.
Step 3-2: Choose a perfect matching of $(U, V; E(U, V))$ under the probability function that a perfect matching M is chosen with probability $\lambda(M)$.
Step 3-3: Set the subvector $\boldsymbol{X}|_{E(U,V)}$ be the characteristic vector of the chosen perfect matching.
Step 4: Output a 0-1 valued vector $\boldsymbol{X} \in \{0, 1\}^E$ defined by

$$X(e) = \begin{cases} X_U(e) & (\forall e \in \delta(U)), \\ \sum_{u \in U} X_U(u, v) X_U(u, v') & (\forall e = \{v, v'\} \in E \setminus \delta(U)). \end{cases}$$

For executing Step 3-1, we need to represent the subvector $\boldsymbol{x}^*|_{E(U,V)}$ by a convex combination of characteristic vectors of perfect matchings in the bipartite graph $(U, V; E(U, V))$ for each $V \in \mathcal{F} \setminus \{U\}$. We can find a set of coefficients for convex combination by applying (unweighted version of) Hungarian method $O(n^2)$ times. Thus, the time complexity of Step 3 is bounded by $O(kn^{4.5})$ [10]. The time complexity of Step 4 is bounded by $O(k^2 n^3)$.

The following theorem is our main result.

Theorem 1. *Algorithm 1 finds a feasible solution of ILP such that the expectation of the corresponding objective function value is less than or equal to* $\left(\frac{5}{2} - \frac{3}{k}\right) z^{**}$ *where* z^{**} *is the optimal value of the multidimensional assignment problem defined by subsets* $V_1, \ldots, V_k \subseteq \mathbb{Q}^d$ *satisfying* $|V_1| = \cdots = |V_k| = n$.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. *Mathematical Programming*, **95** (2003), 3–51.
2. Bandelt, H.-J., Crama, Y., Spieksma, F. C. R.: Approximation algorithms for multi-dimensional assignment problems with decomposable costs. *Discrete Applied Mathematics*, **49** (1994), 25–50.
3. Burkard, R. E., Rudolf, R., Woeginger, G. J.: Three-dimensional axial assignment problems with decomposable cost coefficients. *Discrete Applied Mathematics*, **65** (1996), 123–139.
4. Burkard, R. E., Çela, E.: Linear assignment problems and extensions. In: Du, D.-Z., Pardalos, P. M. (eds.): *Handbook of Combinatorial Optimization, Supplement Volume A*, 1999, Kluwer Academic Publishers, 75–149.
5. Crama, Y., Kolen, A. W. J., Oerlemans, A. G., Spieksma, F. C. R.: Throughput rate optimization in the automated assembly of printed circuit boards. *Annals of Operations Research*, **26** (1991), 455–480.
6. Crama, Y., Spieksma, F. C. R.: Approximation algorithms for three-dimensional assignment problems with triangle inequalities. *European Journal of Operational Research*, **60** (1992), 273–279.

7. Frieze, A. M., Yadegar, J.: An algorithm for solving 3-dimensional assignment problems with application to scheduling a teaching practice. *Journal of the Operational Research Society*, **32** (1981), 989–995.
8. Garey, M. R., Johnson, D. S.: *Computers and Intractability: A Guide to the Theory of NP Completeness.* 1979, W. H. Freeman.
9. Haley, K. B.: The multi-index problem. *Operations Research*, **11** (1963), 368–379.
10. Hopcroft, J. E., Karp, R. M.: An $n^{5/2}$ algorithm for maximum matching in bipartite graphs. *SIAM Journal on Computing*, **2** (1973), 225–231.
11. Pierskalla, W. P.: The tri-substitution method for the three-dimensional assignment problem. *Journal of the Canadian Operational Research Society*, **5** (1967), 71–81.
12. Pierskalla, W. P.: The multidimensional assignment problem. *Operations Research*, **16** (1968), 422–431.
13. Poore, A. B.: Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking. *Computational Optimization and Applications*, **3** (1994), 27–54.
14. Poore, A. B., Rijavec, N.: Partitioning multiple data sets: multidimensional assignments and Lagrangian relaxation. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, **16** (1994), 317–342.
15. Rässler, S.: *Statistical Matching.* Lecture Notes in Statistics **168**, 2002, Springer-Verlag.
16. Soong, R., de Montigny, M.: No free lunch in data fusion/integration. *Session Papers of Worldwide Audience Measurement*, June 2004, Geneva, 33–54.
17. Spieksma, F. C. R., Woeginger, G. J.: Geometric three-dimensional assignment problems. *European Journal of Operational Research*, **91** (1996), 611–618.
18. Spieksma, F. C. R.: Multi index assignment problems: complexity, approximation, applications. In: Pardalos, P. M., Pitsoulis, L. S. (eds.): *Nonlinear Assignment Problems, Algorithms and Applications*, 2000, Kluwer Academic Publishers, 1–12.