Developments in Statistics
Andrej Mrvar and Anuška Ferligoj (Editors)
Metodološki zvezki, 17, Ljubljana: FDV, 2002

# Computing Regular Equivalence: Practical and Theoretical Issues

Martin Everett[1] and Steve Borgatti[2]

**Abstract**

Social network analysts have tried to capture the idea of social role explicitly by proposing a framework that precisely gives conditions under which grouped actors are playing equivalent roles. They term these methods positional analysis techniques. The most general definition is regular equivalence which captures the idea that equivalent actors are related in a similar way to equivalent alters. Regular equivalence gives rise to a whole class of partitions on a network. Given a network we have two different computational problems. The first is how to find a particular regular equivalence. An algorithm exists to find the largest regular partition but there are no efficient algorithms to test whether there is a regular k-partition. That is a partition into k groups that is regular. In addition, when dealing with real data, it is unlikely that any regular partitions exist. To overcome this problem relaxations of regular equivalence have been proposed along with optimisation techniques to find nearly regular partitions. In this paper we review the algorithms that have been developed to find particular regular equivalences and look at some of the recent theoretical results which give an insight into the complexity of finding regular partitions.

## 1 Introduction

Ideas of social role have been important to social theorists since the middle of the century. Social networks provide an ideal environment in which to formalise the theoretical concepts of role and position. Individuals play the same role if they relate in the same way to other individuals playing counterpart roles. Hence the

[1]School of Computing and Mathematical Science, University of Greenwich, 30 Park Row, London SE10 9LS.

[2] Department of Organisational Studies, Carroll School of Management, Boston College, Chestnut Hill, Boston, MA 02167.

role of middle managers in an organisational structure can be determined by their relationship to higher management and workers. There is no need for these individuals to be connected to each other; their role is independent of their organisation and so they can be in different components of the network. However, as we shall see, some models insist that actors playing the same role do so to exactly the same actors. Methods which try and capture notions of social role and related concepts in networks are often referred to as positional analysis techniques.

For clarity of exposition we shall assume, that all networks are graphs or digraphs without isolates. (The extension to multigraphs with isolates is relatively simple). Also we will allow self-loops for both graphs and digraphs.

A coloration of a graph or digraph is an assignment of colours to the vertices. In any coloration we shall use the notation C(v) to denote the colour of each vertex v. If S is a subset of the vertices of a graph or digraph then C(S) is defined by $C(S) = \{C(v) : v \in S\}$.

The set of all vertices of a particular colour is called the colour class; the set of all colour classes is called the colour partition.

# 2 Regular coloration

The purpose of coloration is to provide a framework in which to develop more formally the role concept. Members of the same colour class will be playing the same role provided they fulfil certain regularity conditions.

The neighbourhood N(v) of a vertex $v \in V$ is the set of all vertices which are adjacent to v. In a directed graph we define the in-neighbourhood $N_i(v)$ and the out-neighbourhood $N_o(v)$ by

$$N_i(v) = \{x: (x,v) \in E\}$$
$$N_o(v) = \{x: (v,x) \in E\}$$

It follows that in a digraph $N(v) = N_i(v) \cup N_o(v)$.

**Coloration:**

A coloration of a digraph D(V,E) is *regular* if and only if for all $u,v \in V$

$$C(u) = C(v) \text{ implies } C(N_i(u)) = C(N_i(v))$$
$$\text{and } C(N_o(u)) = C(N_o(v)).$$

Obviously for graphs we simply require C(N(u)) = C(N(v)).

Regular coloration is an alternative formulation (Everett and Borgatti, 1991) of regular equivalence (White and Reitz, 1983).

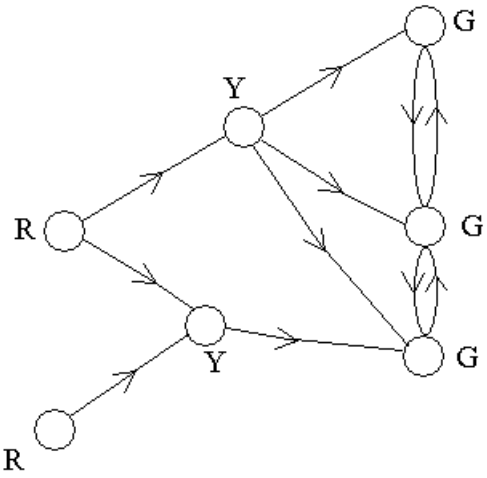



**Figure 1:** Regularly coloured digraph (R = Red, Y = Yellow, G = Green).

In the regularly coloured digraph of Figure 1 we note that the out-neighbours of the red vertices are all yellow; the out-neighbours of the yellow vertices are green and the in-neighbours are red; the out-neighbours of the green vertices are green and the in-neighbours are green and yellow. The number of vertices each is connected to does not matter, one red is connected to one yellow whilst the other is connected to two yellows. It is important that each vertex is connected to or from a vertex in its colour neighbourhood; hence every green must receive a connection from a yellow and a green.

Any colour partition induces an equivalence relation on the vertex set with the distinct equivalence classes corresponding to the colour classes. Conversely any equivalence class will induce a coloration.

**Theorem:** Let D(V,E) be a digraph and ≡ an equivalence relation on v. If ≡ is such that for all a,b,c,∈ V, a ≡ b implies both

1. aEc implies there exists d∈ V such that bEd and d≡c,
2. cEa implies there exists d∈ V such that dEb and d≡c,

then ≡ induces a regular colouring on D. Conversely if ≡ is an equivalence relation induced by a regular colouring then ≡ satisfies the conditions given above.

The above theorem gives an alternative formulation of regular colorations in terms of equivalence relations. An equivalence relation which induces a regular coloration will be called a *regular equivalence.* This is the original formulation proposed by White and Reitz.

For any graph or digraph the trivial partition w(V) which places every vertex in its own equivalence class induces a regular coloration. For any particular graph or digraph D there will be a non-empty collection of regular coloration which we shall denote by R(D). The graph in Figure 2 has 9 regular colorations out of a possible 15 colorations.

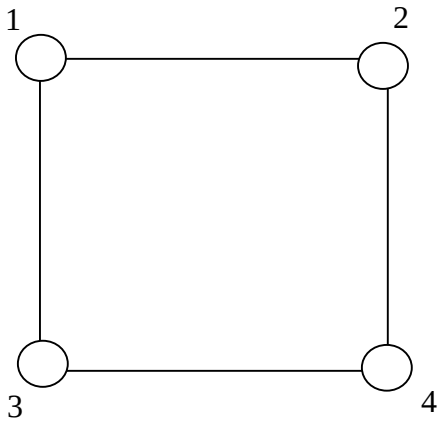



**Figure 2:** Sample graph.

The regular colorations can be partially ordered by the refinement relation <. If $P_o$ and $P_1$ are partitions of a vertex set V; then $P_o < P_1$ if for every $A \in P_o$, there exists a $B \in P_1$ such that $A \subset B$. Figure 3 is pictorial representation of this partial order for the graph in Figure 2.

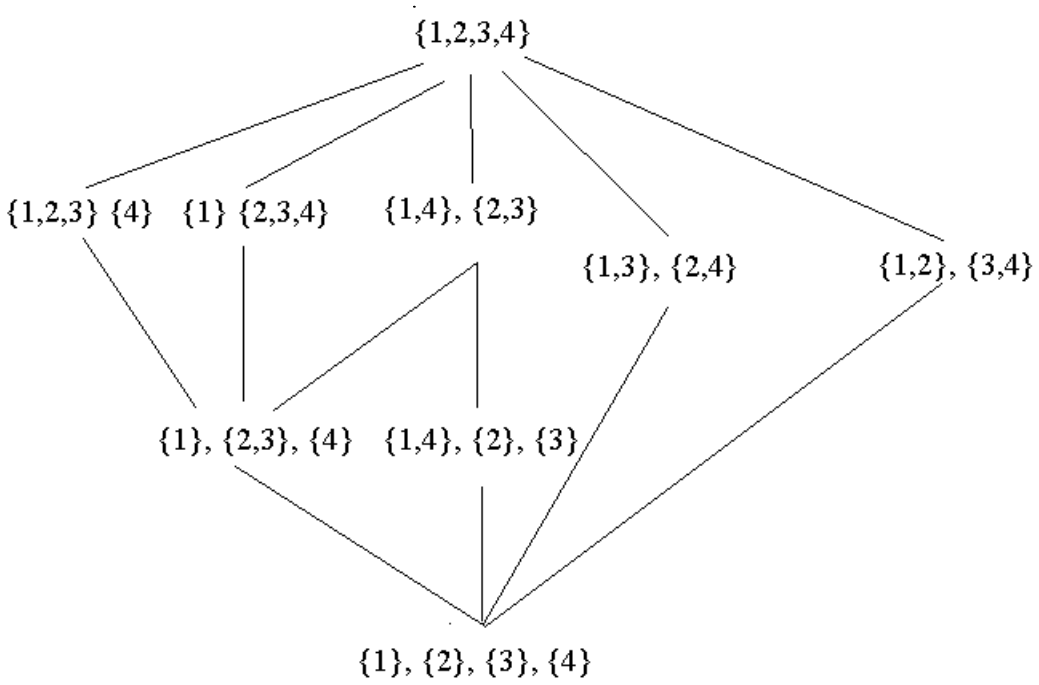



**Figure 3:** Pictorial representation of partial order.

This diagram suggests that (R(D), <) is actually a lattice. This result was proved by Borgatti and Everett (1989).

The proof describes how to join together regular colorations to make new regular colorations. Forming the meet is however more complicated; it is fairly easy to construct examples which demonstrate that unlike the join the meet is not the same as the meet of equivalence classes.

A consequence of the result above is that there will always be a maximal regular coloration. For the graph in Figure 2 this was the complete partition i(G) in which every vertex is the same colour. In general, in connected graph, the complete colour partition i(D) is regular (and obviously maximal) if and only if D contains no sources or sinks. Consequently in an undirected graph i(G) is the maximal regular coloration.

Unfortunately, for many digraphs with sources or sinks the maximal regular coloration may only be the equality partition. For graphs and digraphs without sources and sinks the complete partition is regular, but it may happen that there are no smaller partitions except the equality partition, in these cases strict regular colorations are not much use.

A graph or digraph D with three or more vertices which does not possess a regular coloration different from i(D) or w(D) is called *role-primitive.*

It is easy to see that the digraph in Figure 4 is role-primitive.

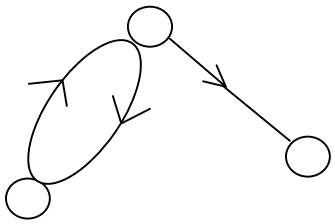


**Figure 4:** Role-primitive digraph.

The existence of graphs (as opposed to digraphs) which are role-primitive is more complicated; the graph shown in Figure 5 is an example (smaller examples do exist).

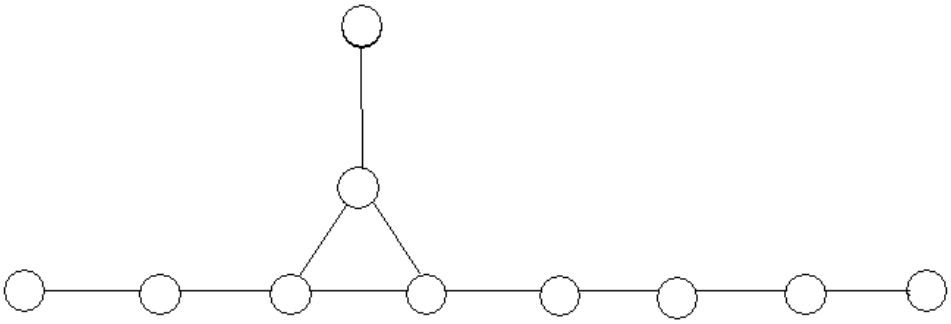


**Figure 5:** Role-primitive graph.

The lattice of all regular colorations may contain a large number of elements. For instance the complement of the graph Figure 5 has over 1,000 regular colorations. Any of these coloration is acceptable as a model for role analysis; the difficulty faced by the researcher is to identify which of the colorations is best suited to the situation in question. Generating all possible colorations is in itself a computationally difficult tasks. Consequently there is a need to develop certain

colorations which have properties which can be identified by the researcher as being important for any particular analysis. For a directed graph with sources or sinks this may be the maximal regular coloration; however obviously this would not be of use in other cases. In the following two sections we investigate the more important regular colorations.

# 3 Automorphic equivalence

In this section we begin to investigate certain types of regular colorations which group together vertices which possess identical structural properties (Winship and Mandel, 1983; Everett, 1985).

Let D(V,E) be a labelled graph. An *automorphism* of D is a permutation $\pi$ of V such that aEb if and only if $\pi(a)E\pi(b)$ $\forall a,b \in V$. Two vertices $c,d \in V$ are *automorphically equivalent* if there exists an automorphism $\pi$ such that $\pi(c)= d$.

The set of all automorphisms of a digraph D form a group under the operation of composition which we shall denote by Aut(D).

Let D(V,E) be a digraph and H a subgroup of Aut(D). Two vertices $a,b \in V$ belong to the same *orbit* of H if and only if $\pi(a) = b$ for some $\pi \in H$.

The orbits of H are therefore the maximal sets of autmorphically equivalent vertices of D for which the automorphisms are restricted to those in H. It is easy to show that belonging to the same orbit is an equivalence relation and hence the orbits partition the vertices of D.

Everett and Borgatti (1991) show that colouring each orbit of any subgroup H of Aut(D) forms a regular coloration of D.

# 4 Structural equivalence

Structural equivalence was the first mathematical formalisation of the role concept (Lorrain and White, 1971). Originally the concept was introduced using the terminology of category theory but we shall give a graph theoretic definition here. This definition makes it difficult to appreciate the use of the 'functional reductions' which were central to the concept when it was originally introduced.

Two vertices a,b of a digraph D are structurally equivalent if and only if (a,b) is an automorphism.

It is easy to see that any partition which places together structurally equivalent vertices will induce a regular coloration.

The regular coloration in which each colour class consists of all structurally equivalent vertices is called the maximal structural coloration.

The definition of structural equivalence is the strictest of all the role models introduced. Two vertices are structurally equivalent only if they connect to exactly

the same vertices in the graph or digraph. Hence two vertices in different components of a graph cannot be structurally equivalent (unless, of course, they are isolates). Consequently structurally equivalent actors not only play the same role but they play the same role to the same people. This is clearly a lot stricter than automorphic in which individuals play the same role to equivalent people. If we consider the example of parents and children then the structurally equivalent children would have the same parents; so that structural equivalence captures the brother/sister role. Automorphically equivalent children would come from families of the same size; and maximal regular equivalence would place together all children.

# 5 Computation

## 5.1 Maximal regular coloration

There is a simple polynomial time algorithm which finds the regular coloration which uses the fewest colours. This algorithm can be summarised as (Everett and Borgatti, 1997)

1. colour all vertices the same
2. for each colour pick a representative colour
3. for each representative: check the regular coloration condition against all vertices in the same colour class. Re-colour all vertices which fail the condition with a new colour. If no re-colouring occurs stop: otherwise go to 2.

Unfortunately this algorithm is of little practical use. It requires the graph to have sources or sinks (otherwise the result is trivial) and even then it often produces a trivial result in which every vertex is coloured differently. On real data the REGE algorithm yields a measure of regular equivalence. This algorithm was developed by Lee Sailor, John Boyd, Douglas White and Karl Reitz although they never published a description. Details can be found in Wasserman and Faust (1994). This algorithm still cannot be used on un-directed data or data without sources or sinks as it will again produce trivial results.

## 5.2 Structural equivalence

It is a simple matter to find if any pair of vertices are structurally equivalent. Algorithms have been developed to measure the degree of structural equivalence between any pair of vertices. One of the most common is the profile similarity method in which the row and column of the adjacency matrix of a vertex are

connected to form a profile. The similarity of these profiles is measured by their distance apart in Euclidean space, or their correlations are taken as a measure of structural equivalence. If the data are pure adjacencies then these are sometimes converted to geodesic distances before constructing the profiles (Burt, 1976).

## 5.3 Automorphic equivalence

A 'foolproof' algorithm for finding all automorphically equivalent vertices is computationally very difficult (it is not known whether this problem is NP-complete or not) and is equivalent to solving the graph isomorphism problem. However, efficient (polynomial) algorithms have been proposed which solve the problem for all but a handful of graphs. These algorithms are based upon separating vertices into different orbits if they do not share certain properties. Eg. if vertex x has a different centrality measure than y then they must be in different orbits.

Borgatti, Everett and Freeman (1999) have implemented an algorithm based on this principle which can be seen as an extension of the profile similarity method. The profiles are constructed as in the structural equivalence method but before they are compared the elements in the profiles are sorted. If the original data consists of a simple adjacency matrix then this should be converted to a geodesic distance matrix. These distances then need to be transformed so that when a correlation or Euclideon distance is taken smaller values have greater influence. We expect immediate and close neighbours to be of more influence than those at a distance. The simplest transformation is to take the reciprocal of the non-zero elements. We demonstrate this method on the undirected graph in Figure 6.

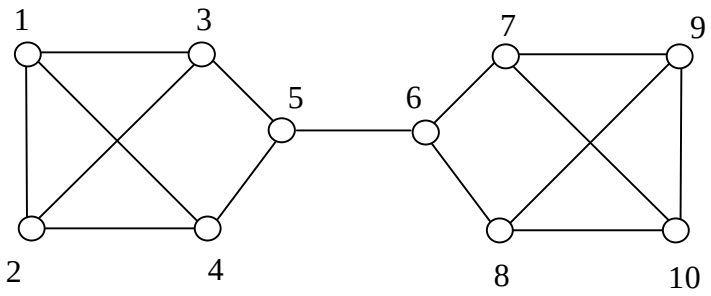



**Figure 6:** Undirected graph.

It is easy to see that 3 and 4 are structurally equivalent and that 7 (and 8) are automorphically equivalent to both 3 and 4. The standard profiles of 3 and 4 are (1,1,0,0,1,0,0,0,0,0) and since they are identical we conclude that 3 and 4 are structurally equivalent. The standard profile of 7 is (0,0,0,0,0,1,0,0,1,1) which is clearly different. The distance profiles for 3,7 and 1 are given by

profile (1) = (0,1,1,1,2,3,4,5,4,5)
profile (3) = (1,1,0,2,1,2,3,3,4,4)
profile (7) = (4,4,3,3,2,1,0,2,1,1)

If we now arrange the elements in ascending order we obtain the following

profile ′(1) = (0,1,1,1,2,3,4,4,5,5)
profile ′(3) = (0,1,1,1,2,2,3,3,4,4)
profile ′(7) = (0,1,1,1,2,2,3,3,4,4)

We immediately see that 3 and 7 are identical and that 1 is different. We can provide a measure of all pairs by taking reciprocals before comparing profiles.

An alternative methods has been proposed by Sparrow (1993) using numerical signatures. This is a neat method and is based upon extending the idea of degree sequences of neighbourhoods. Suppose a vertex j is adjacent to a set of nodes which have degrees 1,1,1,4 and 7. We then combine these into a unique number which is independent of any permutation of the vertices. We form $S_{j,1}$ $(1 + \pi)(1 + \pi)$ $(1 + \pi)$ $(4 + \pi)$ $(7 + \pi)$

The use of the transcendental number $\pi$ means that $S_{j,i}$ can only be formed from 1,1,1,4 and 7. This is then iterated so that $S_{j,\,k\,+\,1} = \Pi\,(\pi + S_{j,k})$

This iteration means that the degree sequence of the higher order neighbourhoods are taken into account. Finally we need to modify S so that it can take account of different cycle lengths. We use a different transcendental number (e in this case) so that our uniqueness property is preserved.

$S_{j,k+1} = (\,|\,N_{k+1}(j)\,|\,+ e)\;\Pi\;(\Pi + S_{j,k})$.

This unique (theoretically it need not be unique since the iterative nature of the construction means that it is possible to produce the same number from a different graph) can then be used to partition the graph. Unfortunately the signatures themselves do not have any meaning – but this could be a fruitful line of research.

## 5.4 Combinational optimisation

An alternative approach is to specify the number of colours in advance and find coloration which is nearly regular. The cost of a coloration is the minimum number of edge changes (deletion or addition) required to make the coloration regular. The cost of a regular coloration is therefore zero. Once a cost has been established then we can use standard combinational optimisation techniques to find the coloration which has the lowest cost. Batagelj et al. (1992) first suggested this approach and they use local search. Others here used genetic algorithms, simulated annealing or tabu search.

# 6 Algorithmic complexity

We have already mentioned the complexity of finding the automorphically equivalent vertices. We also gave a polynomial time algorithm for finding the maximal regular coloration and we noted that it is easy to find structurally equivalent vertices. The use of combinatorial optimisation routines to find colorations with a specified number of colours suggests that these are non-trivial problems.

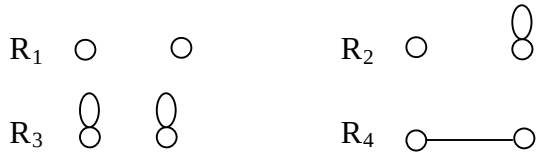
We now look at the complexity of finding if a regular coloration exists with a specified number of colours. We shall suppose that our graphs or digraphs have n vertices. The image graph of a coloration is a graph (or digraph) with colours as vertices and an edge connects colour A to colour B if there is an edge joining a vertex coloured A to a vertex coloured B in the original graph.

Suppose we wish to check for n-colours. Since this is always a regular coloration then there is nothing to do and the problem is in the class P. It is also easy to check for n-1 colours. In this case the pair of vertices coloured the same must be connected to and from the same vertices (that is they are structurally equivalent) and hence this belongs to the class P.

At the other extreme suppose we wish to use just one colour. Clearly this must be in the class P since we simply colour all the vertices and check it is regular. However as in the n-1 case we can characterise the graphs and digraph which have a 1-regular coloration. For a graph to have a 1–regular coloration it is either the null graph or is a graph with no isolates.

For a digraph it is either the null graph or it has no sources, sinks or isolates.

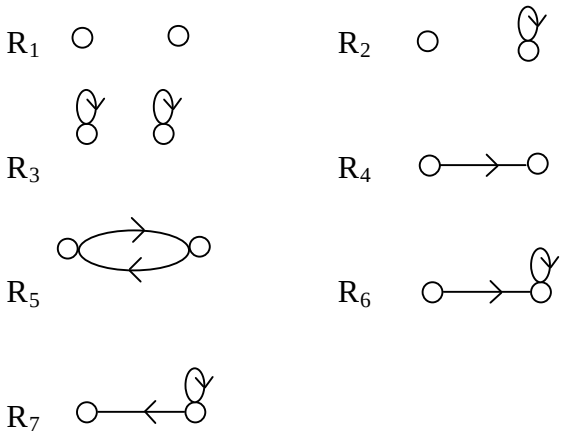
The simplest interesting case is a 2-regular coloration and this was investigated for graphs by Roberts and Sheng (1999). They show that finding a 2-regular coloration for a graph belongs to the class NP. It follows that for digraphs the problem cannot be in P and almost certainly belongs to the class NP as well. In addition they show that for certain image graphs we can characterise 2-regular colorations. These are as follows. If G was the original graph then the following image graphs are possible:




- $R_1$ is possible iff G is a null graph
- $R_2$ is possible iff G has at least one edge and at least one isolated vertex
- $R_3$ is possible iff G is disconnected with no isolates
- $R_4$ is possible iff G is bipartite with no isolates.

The problem of determining whether the two remaining images are possible both belong to the class NP.

We can similarly characterise 7 of the 10 possible digraph images. These are as follows where D was the original digraph.




- $R_1$ is possible iff D is a null graph.
- $R_2$ is possible iff D has isolates but no sources or sinks.
- $R_3$ is possible iff D is disconnected with no sources, sinks or isolates.
- $R_4$ is possible iff D has no isolates and every vertex is a source or sink.
- $R_5$ is possible iff D has no sources, sinks or isolates and the underlying graph is bipartite.
- $R_6$ is possible iff D has no isolates or sinks but contains at least one source.
- $R_7$ is possible iff D has no isolates or sources but contains at least one sink.
- We conjecture that the 3 remaining images give rise to problems which are members of the class NP.

# References

[1] Batagelj, V., Doreian, P., and Ferligoj, A. (1992): An optimisation approach to regular equivalence. *Social Networks,* **14**, 121-135.

[2] Borgatti, S. and Everett, M. (1989): The class of all regular equivalences: Algebraic structure and computation. *Social Networks*, **11**, 65-88.

[3] Borgatti, S., Everett M., and Freeman, L. (1999): *UCINET V: Software for Social Network Analysis*. Natick: Analytic Technologies.

[4] Burt, R. (1976): Positions in networks. *Social Forces*, **55**, 93-122.

[5] Everett, M. (1985): Role similarity and complexity in social networks. *Social Networks*, **7**, 353-359.

[6] Everett, M. and Borgatti, S. (1990): Role colouring a graph. *Mathematical Social Science,* **21**, 183-188.

[7] Everett, M. and Borgatti, S. (1997): The regular colouration of graphs. In C. Mitchell (Ed.): *Applications of Combinatorial Mathematics.* Oxford Clarendon Press.

[8] Lorrain, F. and White, H. (1971): Structural equivalence of individuals in networks. *Journal of Mathematical Sociology,* **1**, 49-80.

[9] Roberts, F. and Sheng, L. (1999): Role assignments. In Y. Alavi, P. Lick, and A. Schwenk (Eds): *Combinatorries, Graph Theory and Algorithms*. **II,** New Issues Press Kalamazoo, MI

[10] Sparrow, M. (1991): A linear algorithm for computing automorphic equivalence classes: The numerical signatures approach. *Social Networks,* **15**, 151-170.

[11] White, D. and White, H. (1983): Graph and semigroup homomorphisms on networks of relations. *Social Networks*, **5**, 193-235.

[12] Winship, C, and Mandel, M. (1983): Roles and positions: A critique and extension of the block modelling approach. In S. Leinhardt (Ed.): *Sociological Methodology*, 1983-1984. San Francisco: Josey-Bass.

[13] Wasserman, S. and Faust, K. (1994): *Social Network Analysis – Methods and Applications.* Cambridge, England: Cambridge University Press.