An Exact Approach to Demonstrate the Detrimental Effects of Gerrymandering on a Political Balance in Democracies

Süleyman Topdemir[a,], Kemal Kılıç[a], Ş. İlker Birbil[b]

[a]*Sabanci University*
[b]*Erasmus University*

Abstract

Redrawing of the electoral district boundaries has been studied since the 60s in several democratic countries, especially in the USA. The major concern in such a process is gerrymandering which is manipulation the district's boundaries in favor of a particular party or group. In this study, two new integer programming mathematical formulations are introduced to demonstrate quantitatively how the gerrymandering malpractice may cause some detrimental results on a political balance. Single-member and multi-member electoral districts are covered in the formulations. Computational results are obtained using the past Istanbul election and referendum data to test the formulations and illustrated.

Keywords: political districting, gerrymandering, mixed integer programming, single-member district, multi-member district, d'Hondt, political balance

1. Introduction

In democracies, an elected party or a group of parties has the power to rule the country for a predetermined time. This is one of the essentialities of the modern democratic systems and their sustainability cannot be maintained easily without a proper economic and social structure. Taking into account a pre-established body of rules by all the institutions of a democracy such as government, law, an opposing party, etc. is one of the most essential conditions for the long-term democracy. Despite all, a democratic system is always under various threads. The gerrymandering is one of them and can be defined as a practice that involves manipulating the electoral district borders in favor of a particular party or group. For a fair representation and long-term democracy, this malpractice must be precluded.

In the USA, redistricting of the electoral districts is a legally required process that takes place annually. The main thread is that a governmental party can misuse this process by redrawing the electoral districts' boundaries in favor of itself. Here, the fair representation must be achieved by taking into consideration some criteria as the output of the redistricting process.

Email addresses: `stopdemir@sabanciuniv.edu` (Süleyman Topdemir), `kkilic@sabanciuniv.edu` (Kemal Kılıç), `birbil@ese.eur.nl` (Ş. İlker Birbil)

Political districting (PD) or redistricting is a specific version of the districting problem in which the aim is to aggregate the areal units into the set of districts with respect to an objective and subject to certain constraints on the zones. The design of districts for schools, social facilities, and sales/service territory are the most prevalent examples of other types of districting problems, also known as the zone design problem. Political districting consists of the partitioning of areal units, generally administrative units, into a prespecified number of districts that satisfy some criteria (Bação et al., 2005). Population equality, for example, is one of the most common criteria for the redistricting problem. Contiguity and the shapes of the districts are other important criteria for the redistricting.

This problem is particularly significant in democracies where one representative can be elected from each electoral district. This type of electoral district is called single-member district. In the literature, mostly single-member electoral district and the aforementioned fair representation are studied.

Several approaches have been developed in the political districting literature since the 60s. In the next 20 years, mostly, the multi-kernel growth approaches have been used to solve the PD problem (Vickrey, 1961; Hess et al., 1965; Bodin, 1973). Between the 1970s and 2010s, several exact approaches with different formulations have been suggested (Garfinkel and Nemhauser, 1970; Nygreen, 1988; Mehrotra et al., 1998; Nemoto and Hotta, 2003). In addition to these methods, local search metaheuristics have been extensively used in the recent literature (Bozkaya et al., 2003; Vanneschi et al., 2017).

The legitimacy of a proposed representation can be questionable in most of the democracies. Changing the current representation for the first time, switching from the multi-member district to the single-member district or vice versa or insufficiency of the most essential institutions of the democracy can be potential reasons for this investigation. In a changing period or such insufficient democracies, to show mathematically how easy the policymakers can misuse the redistricting practice to gain a political advantage can be helpful to create an awareness for the fair representation.

In this study, we aim to show how gerrymandering malpractice has a large impact that is calculated quantitatively on the election results by using two different mathematical formulations. We tried to achieve the results for not only the single-member district but also the multi-member district. Note that the fair representation, extensively studied in the literature, is beside the point in this study. The main contributions of this study are the following;

- to show numerically how the gerrymandering has a detrimental effect on the electoral system and the democracy by proposing two mathematical formulations for the single-member and the multi-member electoral districts. To the best of our knowledge, there is no such formulation that consists of a part of calculating the number of representatives and maximizing the number of representatives of a chosen party for both single-member and multi-member electoral districts. To do this, d'Hondt highest average method for allocation seats in the multi-member district is formulated as some constraints in the mathematical model for the first time in the literature.

- to enhance a heuristic approach that is used in several studies for initialization in terms of the shortest time in which achieving a feasible solution for the main approach, the exact method in this study.

- to push the limits of the current computational power by solving the relatively large problems in terms of the number of political units. In other words, we tried to show how the approach we developed in this study can be applied on real-world problems easily.

The rest of the study is organized as follows. In the next section, we present a comprehensive summary of the related literature. Details of the political districting problem and proposed mathematical formulations are given in Section 3. Section 4 is reserved for the computational results. Lastly, in Section 5, we conclude the study with the discussion of the results and future study.

2. Review of Related Literature

The desired outputs of an automated redistricting practice were explained for the first time in Vickrey (1961). According to the author, the practice must be completely mechanical, and no one can predict in any detail the outcome of the process.

Even though the problem began with Vickrey's work, the first mathematical model was introduced in Hess et al. (1965) for the problem. The proposed mathematical formulation is originated from a warehouse allocation problem formulation minimizing the assignment cost in terms of the distance between the centers and the unit territories. The authors claim was that a rapid and nonpartisan technique was needed for the PD problem. A multi-kernel growth algorithm was proposed to overcome the limited computational power at that time. In the formulation and algorithm, the contiguity is not an explicit constraint, however, the criterion is taken into account only a posteriori which means that the results will be feasible since the contiguity constraint is satisfied by eliminating noncontiguous solutions. In the mathematical model below, I is the set of units and x_{ij} is the binary decision variable, which is 1 if the j^{th} unit is assigned to the i^{th} center. $|I|$ units are tried to be divided into $|H|$ districts. Here H is the set of districts we aim to obtain. a and b represent the minimum and maximum allowable district populations, respectively, as a percent of the average district population. d_{ij} is the Euclidean distance between i^{th} and j^{th} units. s_j denotes the population of the j^{th} unit. The case of the value of variable x_{jj} is equal to 1 indicates that j^{th} unit is the center of a district. Note that $\forall i, j \in I$ for all of these parameters and variables.

$$
\begin{aligned}
& minimize \sum_{i \in I} \sum_{j \in I} d_{ij}^2 s_j x_{ij}, && (1)\\
& subject\ to \\
& \sum_{i \in I} x_{ij} = 1, \forall j \in I, && (2)\\
& \sum_{j \in I} x_{jj} = |H|, && (3)\\
& a\bar{s}x_{jj} \leq \sum_{i \in I} x_{ij} p_i \leq b\bar{s}x_{jj},\ \forall j \in I, && (4)\\
& x_{ij} \in \{0, 1\},\ \forall i, j \in I. && (5)
\end{aligned}
$$

In this formulation, the objective function is used to measure the compactness of the districts while the population equality criterion is taken into account by using constraints with number 4. These constraints also enforce that no basic unit can be assigned a unit which is the center of a district. Constraints with number 2 ensure that each unit is assigned to only one district. The prespecified number of districts is forced by the constraint with number 3. In the literature, there are some other location/allocation based studies (Hojati, 1996; George et al., 1997) pioneered by Hess et al. (1965).

The first exact approach for the solution of PD problem was developed in Garfinkel and Nemhauser (1970). This approach has two phases. In the first phase, the authors aimed to attain all possible feasible district plans. In the second phase, a mathematical programming problem is solved to determine the best districts among the feasible solutions. The problem solved in this step is a version of the set-covering problem. The proposed exact approach influenced the future studies (Nygreen, 1988; Mehrotra et al., 1998). Besides these studies, completely new mathematical formulations have been contributed to the PD literature (Nemoto and Hotta, 2003; Li et al., 2007). Linear and quadratic programming approaches were used in these studies respectively. The contiguity criterion part of the formulation in the Nemoto and Hotta (2003) is utilized in this work.

In Bourjolly et al. (1981), an algorithm that has a typical framework of the local search approaches was used for the first time in the PD literature. The authors consider several districting criteria weighted in the fitness function. The most common metaheuristics such as genetic algorithm (Forman and Yue, 2003; Bação et al., 2005; Liu et al., 2016; Vanneschi et al., 2017), tabu search (Bozkaya et al., 2003), simulated annealing (Browdy, 1990), and artificial bee colony (Rincón-García et al., 2015) have been applied to the PD problem recently. One of the feasibility condition which is the contiguity must be checked before scoring the fitness function in some of the studies. However, the compactness and population equality criteria mostly take part in the fitness functions of the approaches.

The latest innovative progresses on the PD literature are utilizing the parallel high-performance computing (Liu et al., 2016), using a combination of more than one metaheuristics, and taking advantage of the Pareto optimality instead of tuning the parameters used in the fitness function (Vanneschi et al., 2017; Rincón-García et al., 2015).

3. Political Districting Problem and Proposed Methods

In this section, we present the details of the problem and the proposed mathematical formulations. The most important districting criteria are explained in detail before proposing the mathematical models. The enhanced initialization method is introduced in this section. In Section 3.3, the details of the mathematical model for the single-member district are provided. In this part, we also present the notation used in both formulations. In Section 3.4, a more generic version of the single-member district mathematical model is developed. The first mathematical model can work on a system in which only one representative can be elected in each electoral district. However, the mathematical model in Section 3.4 allows that more than one representative can be elected in each electoral district.

3.1. The Problem

The political districting problem has some similar features with the clustering problem in optimization (Chou and Li, 2006). Let the set of m initial territorial units be $I = x_1, x_2, ..., x_m$, where x_i is the i^{th} unit. We also need to define the set of districts which is denoted by H and let the number of districts be n. Lastly, let H_j denote the set of all the territorial units that belong to j^{th} district. Then the fundamental constraints for political districting problem are as follows:

$$H_j \neq \emptyset, j = 1, ..., n \,, \tag{6}$$

$$H_j \cap H_k = \emptyset, \; j, k = 1, ..., n, \; j \neq k \,, \tag{7}$$

$$\bigcup_{j=1}^{n} H_j = I \,. \tag{8}$$

According to the first constraint, each district must have at least one territorial unit. The second constraint satisfies that there cannot be any common unit between two districts. This constraint is called integrity in the literature. Lastly, there cannot be a unit that is not assigned to any district. In other words, each unit must be assigned to one of the districts.

In addition to the fundamental constraints, there are some geographical and demographic criteria in the political districting problem. Contiguity and compactness of the districts are the well-known geographical criteria in the literature. A district is said to be contiguous if each unit of the district is reachable from the rest of the units in the district without crossing the district boundary (Grofman, 1985). The definition of the contiguity criterion is more clear than the compactness. According to the compactness criterion, each district must be compact which is an ambiguous term. This term is defined as firmly put together, joined, or integrated; predominantly formed or filled in Dictionary (2019). The idea behind the compactness is to prevent the odd-shaped districts. In Niemi et al. (1990); Young (1988) some of the compactness measures are introduced. Each of them is developed heuristically since there is no exact score for the compactness criterion. The main idea behind the metrics for the compactness is similar. These methods aim to provide a numerical value that shows how close a geometric shape is to an ideal compact shape. Here the ideality assumption can differ among several compactness measurements. For instance, the main assumption of the Polsby-Popper measure method is that shape of a most ideal district should be a circle (Polsby and Popper, 1991).

The most important demographic constraint is population equality. In the redistricting process, equal populated districts are one of the most looked for outcomes since it's related to the *one-person-one-vote*. The main reason to utilize such a criterion is to maintain that each citizen is represented in the legislature with equal weight. Exact equality cannot be satisfied in every case; however, the values should be close enough to an average population value.

All the criteria mentioned above are required for both electoral types. However, population equality criterion must be modified for multi-member district system in which more than one representative can be elected from each district since the number of representatives of each district can be different from any other district. Therefore the population values of the districts must be proportional to their predetermined number of representatives. In Table A.5, the studies in the literature and their considered criteria are shown. Other frequently used criteria are socio-economic homogeneity, the consideration of natural boundaries such as rivers, lakes, etc., conformity of administrative boundaries and recognition of important minority groups and their share of representations (Bozkaya et al., 2003).

3.2. Initialization

There are several initialization procedures, mostly with similar fashion, in the literature. In this study, we used one of these approaches with some modifications (Vickrey, 1961). The outputs of the chosen procedure cannot be feasible considering the population equality. We tried to overcome this problem by applying some additional algorithmic approaches since we faced with this problem frequently in practice.

Solutions are constituted gradually in the Vickrey's work. In each iteration of the first step of the approach, an unassigned unit is chosen randomly and extended by adjoining neighbor units to it one at a time in such a way that total population value of the constituted district attains the average population for the first time. The adjoining procedure is also stopped if there is no adjacent unit left. As the reader might expect that the outputs of this procedure cannot be feasible. The total number of constituted districts can be less than or greater than a desired value. Another reason for infeasibility is that the total population value of a constituted district can be less than the allowable minimum value.

In the second step of this approach, the number of districts achieved is tried to be equal to a predetermined value. If the number of districts achieved in this fashion is larger than the desired value, the number of districts, the value is decreased by combining the least populated district with the least populated neighbor in an iterative manner until the desired value is attained. If the value is less than the desired one, the value is gradually increased by dividing the most populated district into two parts until the desired value is reached.

The dividing process can be achieved randomly maintaining the contiguity criterion for both outputs.

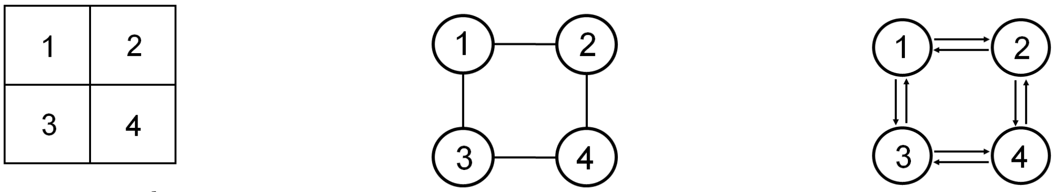
We developed two additional steps for the existing initialization procedure. In each iteration of the first additional step, one of the units in which one of the overpopulated districts is assigned to one of the underpopulated district which is adjacent to the chosen overpopulated district whenever the population of the chosen underpopulated district attains a lower bound which is less than the average population value for the first time. This step is terminated after a certain number of iterations without improvement in terms of total number of districts that satisfy the population equality criterion.

In the last step of the procedure, we aim to decrease the standard deviation among the population values of the districts by changing the outputs of the third step. In each iteration of this step, we're interested in two adjacent districts one of which is chosen randomly. Neighbor district to the seed district is also chosen randomly among all the neighbor districts. We determine the population *difference* between the two districts to be decreased in further alterations. After that, the border units in both districts and their scores which are calculated as the difference of the population values of new districts which are obtained by assigning the corresponding units to their adjacent districts are determined. The number of alternatives for the assignment-based neighborhood change is equal to the number of border units in both districts. The score values are compared with the *difference* value, which is calculated considering the seed districts. The ones which are greater than the *difference* value are eliminated. Here, we tried to eliminate some of the alternatives that will increase the standard deviation value between the populations of the districts. The alternative with the least score value is applied to the current districts and the same procedure is continued whenever no alternative assignment remains. Note that, after any change between two districts, the *difference* value, border units, and their score values are recalculated.

The last step of the initialization approach is terminated whenever all the districts satisfy the population equality criterion. Note that, it is not necessary to apply all the steps we explained since the population equality criterion can be satisfied in some previous steps of the initialization approach.

3.3. Mathematical Formulation for the Single-Member District

G=(I, A) denotes a contiguity graph with a set of nodes, I, and a set of arcs between the nodes, A. Here the political units and adjacency of the units are represented by the nodes and arcs in the graph respectively. After this representation, a new network which is originated from the graph G is introduced; $T = (\bar{I}, \bar{A})$. According to this network, each arc connecting node i and node j in G is replaced by the pair of arcs (i, j) and (j, i); n copies of this graph are formed, and node i in the h^{th} copy of the graph is denoted by w_i^h; $\forall i \in I$, $\forall h \in H$. A small illustration of the process can be seen in Figure 1 and 2. Note that H denotes the set of districts; in other words, the copies of graph. In addition, $|H|$, which is equal to n, source-nodes and $|I|$ sink-nodes a.k.a tail-nodes are introduced. Each source-node s^h, $\forall h \in H$ is connected to all the nodes of the h^{th} copy of the graph with an arc (s^h, w_i^h) $\forall i \in I$, while for each sink-node t_i, $\forall i \in I$ there exists an arc (w_i^h, t_i), $\forall h \in H$. The illustration of the whole network with original nodes, source-nodes, tail-nodes, and arcs is shown in Figure A.4. After introducing the additional nodes and arcs, $\bar{I}$ covers the source-nodes (s^h), the sink-nodes (t_i) and the nodes in the copies of the graph (w_i^h); $\forall i \in I$, $\forall h \in H$. $\bar{A}$ covers the arcs between the source-nodes and the nodes of the copies of the graph (s^h, w_i^h), the arcs between the nodes w_i^h and w_j^h , and lastly the arcs between the nodes of the copies of the graph and sink-nodes (w_i^h, t_i); $\forall i, j \in I$, $\forall h \in H$. P denotes the set of parties in the mathematical models. In order to distinguish the entering and leaving arcs, we have introduced two additional sets, $\delta^-(w_i^h)$ and $\delta^+(w_i^h)$ respectively, $\forall i \in I$, $\forall h \in H$.



(a) the units in an elementary form

(b) the contiguity graph

(c) the graph with arcs

Figure 1: Illustration of the contiguity graph process

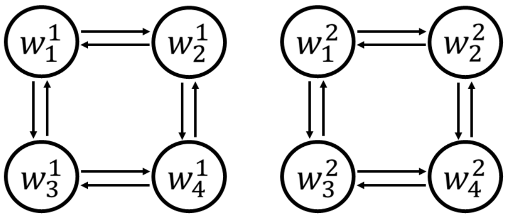


Figure 2: Illustration of the graph copies

The number of votes of a party k in a political unit i represented in the mathematical models as v_{ki}, $\forall k \in P$, $\forall i \in I$. s_i denotes the total number of voters, in other words, the total population in unit i, $\forall i \in I$. To consider the contiguity criterion in the models, we have introduced the parameter β which represents the allowable percentage deviation from the average number of voters that is calculated as follows:

$$\bar{s} = \frac{\sum_{i \in I} s_i}{|H|}, \tag{9}$$

To control the assignment of the political units to the districts, a binary decision variable with two indices is introduced, x_{ih}, $\forall i \in I$, $\forall h \in H$. Three additional binary decision variables; y_{ih}, c_{kh}, and z_{kph} are also introduced, $\forall i \in I$, $\forall h \in H$, $\forall k, p \in P, k \neq p$. y_{ih} controls whether the flow enters from the sink node s^h to the unit w_i^h, $\forall i \in I$, $\forall h \in H$. c_{kh}

controls whether the party k dominates rest of the parties in terms of of total number of votes in district h, $\forall k \in P$, $\forall h \in H$. z_{kph} is required for the comparison of the total votes of a couple of parties, $\forall k, p \in P$, $\forall h \in H$, $k \neq p$. To determine which party dominates rest of the parties in each district, we need such a decision variable.

The rest of the decision variables are non-binary. We need a decision variable that shows the number of voters in each district for the population equality criterion check. t_h is introduced for that aim, $\forall h \in H$. We also need to keep the flow amount of the arcs in the model and introduced a decision variable, $f(a)$, that shows the volume of the flow on arc a, $\forall a \in \bar{A}$. Final decision variable is an auxiliary variable, o_{kh}, that shows the total votes of party k in each district h, $\forall k \in P$, $\forall h \in H$.

In the mathematical model, we aimed to maximize the total number of representatives of a particular party. The chosen party is shown as ρ in the mathematical models. All sets, parameters, and decision variables for the formulations are also provided in Table A.6.

In the single-member district model, there are three numerical parameters that must be predetermined; (β), (M) a sufficiently big value, and (F) the volume of the flow from each sink node. M value must be equal to or greater than the maximum total votes difference between two different parties considering all the districts. F value can be any positive value under the condition of the flow values are not integer. The mathematical model and explanation of the constraints can be seen below.

$$maximize \sum_{h \in H} c_{\rho h}, \tag{10}$$

$$subject\ to$$

$$\sum_{h \in H} x_{ih} = 1,\ \forall i \in I, \tag{11}$$

$$\sum_{i \in I} x_{ih} v_{ki} = o_{kh},\ \forall k \in P,\ \forall h \in H, \tag{12}$$

$$o_{ph} - o_{kh} + M z_{kph} \ \geq 0,\ \forall k, p \in P,\ \forall h \in H, k \neq p, \tag{13}$$

$$\sum_{p \in P, p \neq k} z_{kph} - |P| + 2 \ \leq c_{kh},\ \forall h \in H,\ \forall k \in P, \tag{14}$$

$$\sum_{k \in P} c_{kh} = 1,\ \forall h \in H, \tag{15}$$

$$\sum_{i \in I} x_{ih} s_i = t_h,\ \forall h \in H, \tag{16}$$

$$t_h \leq \bar{s}(1+\beta),\ \forall h \in H, \tag{17}$$

$$t_h \geq \bar{s}(1-\beta),\ \forall h \in H, \tag{18}$$

$$\sum_{i \in I} y_{ih} = 1,\ \forall h \in H, \tag{19}$$

$$f(s^h, w_i^h) = F y_{ih},\ \forall i \in I,\ \forall h \in H, \tag{20}$$

$$\sum_{a \in \delta^-(w_i^h)} f(a) = \sum_{a \in \delta^+(w_i^h)} f(a),\ \forall i \in I,\ \forall h \in H, \tag{21}$$

$$\sum_{a \in \delta^-(w_i^h)} f(a) \leq F x_{ih},\ \forall i \in I,\ \forall h \in H, \tag{22}$$

$$x_{ih} \leq f(w_i^h, t_i),\ \forall i \in I,\ \forall h \in H, \tag{23}$$

$$f(a) \geq 0; o_{kh} \geq 0,\ \forall a \in \bar{A}; \forall h \in H,\ \forall k \in P, \tag{24}$$

$$x_{ih}, y_{ih}, c_{kh}, z_{kph} \in \{0,1\},\ \forall h \in H,\ \forall i \in I,\ \forall k,p \in P, k \neq p. \tag{25}$$

In the objective function, the number of representatives of a particular party is maximized. Each unit must be assigned to a district in the constraints 11. The number of representatives of each party from each district is calculated by using the constraints from 12 to 15. The total number of votes of each party in each district is calculated in 12. The constraints with the number 13 are utilized to find the number of representatives of each party by considering the total number of votes of each party. To find the winner party in each district, at first we compare the total votes of each pair of the parties and keep the result information that can be 0 or 1, then we determine which party dominates the others using the information from the constraints 13. This determination part is in 14. The mentioned information is kept in the variables z_{kph} $\forall k \in P, \forall p \in P, \forall h \in H$. In 15, the constraints enforce that only one candidate or party can be elected in each district. The calculation of the population of each district is handled by the constraints with number 16. The constraints of 17 and 18 are developed for the population equality criterion.

The rest of the mathematical model is originated from the aforementioned article. The main function of these constraints is to satisfy the contiguity criterion. The constraints in 19 enforce that each source node must be assigned to one of the districts. In the constraints 20, the flow amount from each source node to the nodes we have in the original problem must be equal to a predetermined value if the corresponding arc is used. The total amount of flow on entering arcs must be equal to the total amount of flow on leaving arcs for each node (21). In the constraints 22, if node i is not assigned to copy h, then the total flow on the entering arcs to this unit must be zero. If the flow between node w_i^h and tail node

t_i is zero, then the node (unit) i cannot be assigned to the copy (district) h according to the constraints 23. The constraints with number 24 are for the nonnegativity. The last constraints are for defining the domains of the decision variables.

3.4. Mathematical Formulation for the Multi-Member District

In the single-member district system, determining how to handle the allocation of the seats is not an issue since a candidate with maximum vote percentage wins an election. However, an allocation problem arises in the multi-member district electoral system. In order to eliminate this issue, there are several methods of proportionality representation such as d'Hondt, Saint-Laguë, Hamilton, etc.. The d'Hondt method is the most prevalently used proportionality representation method in Europe (Gallagher, 1991). The required inputs to apply the method are the number of representatives that will be selected and the total votes of the each party in an election. Let o_k denote the total vote achieved and c_k the seats received so far by party k. At the initial step of the method, all c_k values are zero. The d'Hondt method compares the respective values $o_k/(c_k + 1)$ for each party, increase the corresponding c_k by one unit until no representative will be assigned. Therefore, not only o_k values, but also o_k/j, here j is called divisor and can be, at most, the total number representatives will be allocated, values are employed in the d'Hondt method.

The constraints from 12 to 15 in the first formulation are modified to the constraints that are from 29 to 34 in the second formulation for the d'Hondt method. Calculating the total number of representatives of each party in each district is handled by this group of constraints in both models. The d'Hondt method is utilized for the allocation of the number of representations of each party. In the formulation above, the constraints with number 15 enforce that only one representative can be elected from each district, however, in the multi-member case, the number of representatives from each district can be different from each other and doesn't have to be equal to one. To represent this distinction, additional sets and parameters are defined. J_h, $\forall h \in H$ indicates the set of representatives of district h. The average number of voters, $\bar{s}$, cannot be used in the multi-member district formulation due to a possible inequality between the number of representatives of the districts. A weighted average, $\bar{s}_h, \forall h \in H$, is defined for the new formulation and can be calculated as follows:

$$\bar{s}_h = \frac{|J_h| \sum_{i \in I} s_i}{\sum_{h \in H} |J_h|}, \forall h \in H. \tag{26}$$

We also defined two additional parameters, M_1 and M_2, to construct some of the constraints in the formulation. M_1 can be defined as a sufficiently big value for the difference of the total votes. In other words, the maximum total population difference among all the districts. The lowest value of the M in the previous formulation and M_1 can be calculated in the same way. M_2 is defined as a sufficiently big value for the number of representatives and parties.

In both formulations, the variable, c_{kh}, is used for the same reason. However, the values are binary in the first formulation and positive integer values are in the second one. c_{kh}

denotes the total number of representatives of party k gained in district h; $\forall k \in P, \forall h \in H$. In addition to the variable o_{kh}, we defined an additional auxiliary variable o_{kjh} to calculate the total votes of party k in district h for the j^{th} representative alternative that is required for the d'Hondt method; $\forall k \in P, \forall h \in H, \forall j \in J_h$. e_{kjh} is a new binary variable and must be equal to one if number of total votes of j^{th} representative option of k^{th} party in district h is not maximum among the all parties with same the j^{th} representative option; $\forall k \in P, \forall h \in H, \forall j \in J_h$. This variable is required for the calculation of total number of representatives in the d'Hondt method. Lastly, a binary variable with five indices, z_{pkjdh} is defined for comparison of the total votes; $\forall k, p \in P$, $\forall h \in H$, $\forall j, d \in J$, $(k \neq p, d \neq j)$. The mathematical model that covers the mentioned decision variables and parameters can be seen below.

$$maximize \sum_{h \in H} c_{ph}, \tag{27}$$

$$subject\ to$$

$$\sum_{h \in H} x_{ih} = 1,\ \forall i \in I, \tag{28}$$

$$\sum_{i \in I} x_{ih} v_{ki} = o_{ph},\ \forall p \in P,\ \forall h \in H, \tag{29}$$

$$o_{phj} = o_{ph}/j,\ \forall p \in P,\ \forall j \in J_h,\ \forall h \in H, \tag{30}$$

$$o_{pjh} - o_{kdh} + M_1 z_{pkjdh}\ \geq 0,\ \forall k, p \in P,\ \forall j, d \in J_h,\ \forall h \in H, (p \neq k, j \neq d), \tag{31}$$

$$\sum_{k \in P} \sum_{d \in J_h} z_{pkjdh} - |J_h| + 1\ \leq M_2 e_{pjh},\quad \forall p \in P,\ \forall j \in J_h,\ \forall h \in H, (p \neq k, j \neq d), \tag{32}$$

$$\sum_{p \in P} \sum_{j \in J_h} e_{pjh} = |J_h||P| - |J_h|,\ \forall h \in H, \tag{33}$$

$$c_{ph} = |J_h| - \sum_{j \in J_h} e_{pjh},\ \forall p \in P,\ \forall h \in H, \tag{34}$$

$$\sum_{i \in I} x_{ih} s_i = t_h,\ \forall h \in H, \tag{35}$$

$$t_h\ \leq \bar{s}_h (1 + \beta),\ \forall h \in H, \tag{36}$$

$$t_h\ \geq \bar{s}_h (1 - \beta),\ \forall h \in H, \tag{37}$$

$$\sum_{i \in I} y_{ih} = 1, \ \forall h \in H, \tag{38}$$

$$f(s^h, w_i^h) = F y_{ih}, \ \forall i \in I, \ \forall h \in H, \tag{39}$$

$$\sum_{a \in \delta^-(w_i^h)} f(a) = \sum_{a \in \delta^+(w_i^h)} f(a), \ \forall i \in I, \ \forall h \in H, \tag{40}$$

$$\sum_{a \in \delta^-(w_i^h)} f(a) \ \leq F x_{ih}, \ \forall i \in I, \ \forall h \in H, \tag{41}$$

$$x_{ih} \ \leq f(w_i^h, t_i), \ \forall i \in I, \ \forall h \in H, \tag{42}$$

$$f(a) \ \geq 0, \ \forall a \in \bar{A}, \tag{43}$$

$$x_{ih}, y_{ih}, z_{pkjdh}, e_{kjh} \in \{0, 1\}, \ \forall j, d \in J_h, \ \forall h \in H, \ \forall i \in I, \ \forall k, p \in P. \tag{44}$$

In the objective function, the number of representatives of a particular party is maximized. Each unit must be assigned to a district (28). The number of representatives of each party in each district is calculated by using the constraints from 29 to 34. The total number of votes of each party in each district is calculated (29). The constraints with number 30 are for the comparison. Only the information of total votes of each party in each district is not sufficient to determine the total number of representatives of each party elected in each district for multi-member district. According to the d'Hondt method, we also need to calculate the half of the total votes of each party, one third of it, and so on. At the end, we are going to compare the values as much the multiplication of the number of total representatives can be elected and total number of available parties. We are calculating these values (30). The main motivation of the constraints 31 is similar to the constraints 13 . The comparison information is obtained by using the z variables with 5 indices . Using the constraints 32 and 33 we are ordering all the values we found from 31. Note that, right hand side value of 33 is equal to a value that is the difference between the number of alternatives and the number of representatives. In other words, we are trying to determine which values are the smallest ones. Total number of representatives of each party elected in each district is calculated by summing the each party's e columns and subtract this value from the total number of representatives can be elected. The constraints from 36 to 37 are related to the population equality criterion. The rest of the mathematical model is originated from Nemoto and Hotta (2003) and explained in the previous formulation in detail.

One of the main contributions of this study is the constraints for the calculation of the number of representatives for each party according to the d'Hont method. The constraints with the number 29-34 are developed for this aim. We tried to show these steps regarding this calculation with a small example in Table A.7, A.8, and A.9. Assume that we are interested in only one district in this example.

4. Computational Results

Since finding an optimal solution is dependent on the size of a problem in terms of the number of units and districts, several feasible initial solutions are achieved to speed up the process. We aimed to extend the quality of the computational results in a limited time starting CPLEX solver with a given initial solution. The computational experiments presented in this chapter are conducted on a workstation with Intel E5-2640 v3 CPU @2.60 GHz and 16 GB RAM running on Windows 7 Professional operating system.

4.1. Data sets

In order to test the effectiveness of the models, a past election and referendum held in Turkey, in November 2015 and in April 2017 respectively, data sets are downloaded from the Turkish Statistical Institute's website (TUIK, 2019). The main reason to use two different data sets is to test the models with the different number of parties. In the raw election data, depending on the county, there can be more than ten parties and most of them had less than one percent of the total votes. To eliminate such parties from the data sets, five parties are chosen to be considered. Two of these parties that are *Justice and Development Party*(JDP) and *Republican People's Party*(RPP) had the vast majority. Quantitatively, they have 81.16 percent of the total votes in the districts we used in this section. The other parties in the preprocessed election data are *Nationalist Movement Party*, *People's Democratic Party* and other parties as one party. In the computational experiments, only JDP and RPP are used in the objective because of their total vote percentages.

On the other hand, there were only two options in the 2017 Turkish constitutional referendum; *yes* and *no*. The percentages of these choices are close to each other; % 51.41 and % 48.59 for *yes* and *no* respectively throughout Turkey. The data sets for the counties we considered in this part have similar percentages.

The data sets of four counties in Istanbul with the different number of political units, here neighborhoods, are used in this section. As we mentioned earlier several times, we are interested in maximizing the authority of a party or a group of parties to show how this process can be manipulated easily. One can claim that using the referendum data set for political representation is questionable. One of the reasons for using such a data set is that we want to show the effect of the different number of options on the execution. Another reason is that both sides in the referendum held in Turkey can be classified as political opponents of each other; a group of right parties versus a group of left parties.

4.2. Comparative Analysis

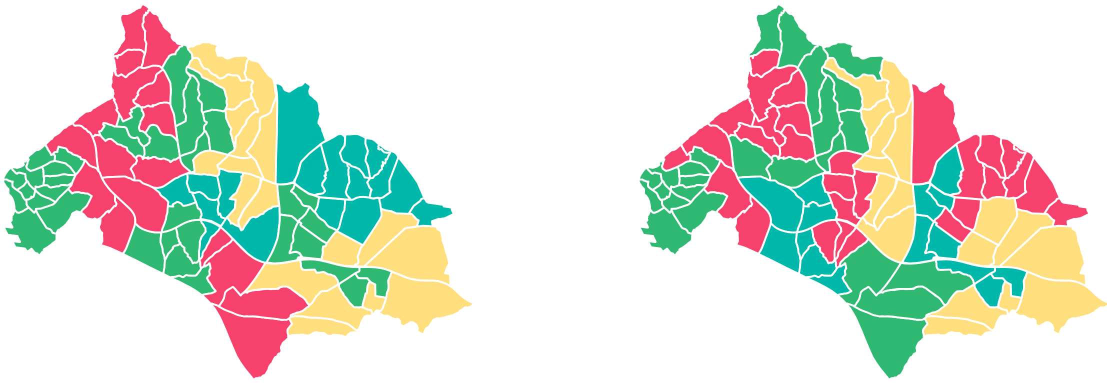
In Table 1, the results of the computational experiments can be seen. The number of districts in this table are determined considering the current number of representatives of the counties. One of the most important observations for these results is that the 2-phases approach has superior outcomes in terms of elapsed time and achieving a feasible solution. Without the initialization method, any result could not be achieved for the cases with 68 and 85 political units in a limited time. Even though we do not have any optimal results for the cases with 68 and 85 units, some satisfying objective values are achieved from the point of showing the gerrymandering can cause critical change in a political balance. For example, the county with 85 political units can be divided in such a way that seven-tenths of the districts have more *yes* percentage. On the other hand, the same county can be divided such that seven-tenths of the districts have more *no* percentage according to the last experiment. The geographical results of these two experiments are visualized in Figure 3. Note that all the population deviation values are smaller than the allowable β value which is set to 0.10 in the experiments.

Dataset			Exact Approach				Initialization + Exact			
# of units	# of districts	*y/n*	obj. value	elapsed time	opt. gap	pop. deviation	obj. value	elapsed time	opt. gap	pop. deviat.
17	2	*yes*	1	0.42	0.00 %	0.09	-	-	-	
17	2	*no*	2	0.14	0.00 %	0.05	-	-	-	-
33	4	*yes*	3	144.06	0.00 %	0.06	3	14.45	0.00 %	0.05
33	4	*no*	3	7200	33.34 %	0.09	3	7200	33.34%	0.09
68	8	*yes*	no sol.	7200	-	-	6	7200	33.34 %	0.09
68	8	*no*	no sol.	7200	-	-	4	7200	75.00 %	0.08
85	10	*yes*	no sol.	7200	-	-	7	7200	28.57%	0.07
85	10	*no*	no sol.	7200	-	-	7	7200	42.86 %	0.07

Note: In the computational experiments, β, M and F values are set to 0.10, 100000, and 200 respectively.

Table 1: Computational experiments on the past referendum data

Despite the fact that the 2-phases approach has a worse outcome in terms of the elapsed time in one experiment, see the case with 33 units and JDP objective in Table 2, superior outcomes are achieved in the rest of the experiments with this approach. In theory, a better initial solution can be achieved by using the only exact approach. However, in practice, we needed some initial feasible solutions for the cases with a relatively high number of units before the execution of the exact approach. We test our models on the data sets with different number of objective options. In the overall election data, the party JDP has a higher percentage value than the RPP. On the other hand, the options, *yes* and *no*, have close percentages in the referendum data. According to the results of both tables, the percentages of the options for the objective in the corresponding data sets have more impact on the results than the number of options in the data sets. The results for the cases with 68 and 85 units in both tables corroborate this statement.



(a) result with 7 *no* and 3 *yes* (b) result with 7 *yes* and 3 *no*

Figure 3: Results on the past referendum data for the case with 85 units

Dataset			Exact Approach				Initialization + Exact			
# of units	# of districts	party	obj. value	elapsed time	opt. gap	pop. deviat.	obj. value	elapsed time	opt. gap	pop. deviat.
17	2	JDP	2	0.27	0.00 %	0.07	-	-	-	-
17	2	RPP	1	0.16	0.00 %	0.06	-	-	-	-
33	4	JDP	4	1.20	0.00 %	0.05	4	33.13	0.00 %	-
33	4	RPP	2	10.41	0.00 %	0.09	2	1.17	0.00 %	-
68	8	JDP	7	7200	14.29%	0.08	8	5888	0.00 %	0.07
68	8	RPP	no sol.	7200	-	-	2	7200	195.46 %	0.07
85	10	JDP	no sol.	7200	-	-	10	5986	0.00 %	0.09
85	10	RPP	no sol.	7200	-	-	2	7200	400.00 %	0.04

Note: Elapsed time values are in seconds.

Table 2: Computational experiments on the past election data

In the experiments for the multi-member district model, the 2-phases approach achieves superior outcomes in terms of the elapsed time. In Table 3 and 4, all the elapsed time values are less than the ones achieved by only the exact approach. In most of the experiments in the last two tables, the results of the initialization method are optimal. This can be explained by considering the way of work of d'Hondt method and the total vote percentages of the data sets we employed. As mentioned before, we used the data sets that have close total vote percentages of parties with each other.

Dataset			Exact Approach				Initialization + Exact			
# of units	# of districts	*y/n*	obj. value	elapsed time	opt. gap	pop. deviat.	obj. value	elapsed time	opt. gap	pop. deviat.
33	2	*yes*	2	0.44	0.00 %	0.07	-	-	-	-
33	2	*no*	2	0.33	0.00 %	0.07	-	-	-	-
68	4	*yes*	4	50.58	0.00%	0.06	4	0.19	0.00 %	0.07
68	4	*no*	4	53.13	0.00 %	0.06	4	0.23	0.00%	0.07
85	5	*yes*	no sol.	7200	-	-	5	0.34	0.00 %	0.09
85	5	*no*	no sol.	7200	-	-	5	0.34	0.00 %	0.09

Note: Total number of representatives value of each experiment is same with the one used in single-member case. That is twice the number of districts of each run.

Table 3: Multi-member district model computational experiments on the past referendum data

Dataset			Exact Approach				Initialization + Exact			
# of units	# of districts	party	obj. value	elapsed time	opt. gap	pop. deviat.	obj. value	elapsed time	opt. gap	pop. deviat.
33	2	JDP	2	0.38	0.00 %	0.05	-	-	-	-
33	2	RPP	2	1.13	0.00 %	0.05	-	-	-	-
68	4	JDP	4	49.16	0.00%	0.08	4	0.28	0.00%	0.07
68	4	RPP	no sol.	7200	-	-	4	0.23	0.00%	0.07
85	5	JDP	5	158.66	0.00 %	0.06	5	0.59	0.00 %	0.07
85	5	RPP	no sol.	7200	-	-	5	0.48	0.00 %	0.07

Note: In the computational experiments for the multi-member district models, β, M_1, M_2 and F values are set to 0.10, 100000, 100 and 200 respectively.

Table 4: Multi-member district model computational experiments on the past election data

5. Conclusions and Future Works

We have discussed that the proposed mathematical models can be employed to demonstrate the detrimental effects of the gerrymandering on a political balance in democracies in which no matter which type of electoral district used. As mentioned before, the gerrymandering malpractice is one of the biggest obstacles in democracies and must be precluded. The fair representation by using the current computational power can be an answer for this misuse and has been studied in the political districting literature since the 60s. However, there is no study that shows mathematically the effects of the gerrymandering on democracy in the literature. We have used the real-world data sets for solving the mathematical models and illustrated the results on a real-world map.

In the multi-member district model, J_h is the set of representatives of district $h, \forall h \in H$.

The number of representatives in each district must be assigned by the decision-maker according to the multi-member district formulation. However, an extended model can decide on these values to maximize the number of representatives of a particular party. In this model, J_h must be determined as a decision variable instead of a set. This kind of extended model can be developed in the future considering these changes and, of course, with some other alterations.

In order to achieve our goal in this study, we have changed the model that is present in the literature. The main goal of this inspired study is the fair representation. However, we have only used the contiguity part of this model and enhanced it in accordance with our aims. We have also benefited from an initialization method that is present in the literature and enhanced it. The main reason to use such a 2-phases approach is that we have tried to achieve optimal or satisfying results for the relatively large cases. The results of the study are promising to form public opinion and draw attention to decision-makers.

References

Bação, F., Lobo, V., Painho, M., 2005. Applying genetic algorithms to zone design. Soft Computing 9 (5), 341–348.

Bodin, L. D., 1973. A districting experiment with a clustering algorithm. Annals of the New York Academy of Sciences 219 (1), 209–214.

Bourjolly, J.-M., Laporte, G., Rousseau, J.-M., 1981. Decoupage electoral automatise: Application a lIle de montreal. INFOR: Information Systems and Operational Research 19 (2), 113–124.

Bozkaya, B., Erkut, E., Laporte, G., 2003. A tabu search heuristic and adaptive memory procedure for political districting. European Journal of Operational Research 144 (1), 12–26.

Browdy, M. H., 1990. Simulated annealing: an improved computer model for political redistricting. Yale Law & Policy Review 8 (1), 163–179.

Chou, C.-I., Li, S.-P., 2006. Taming the gerrymanderstatistical physics approach to political districting problem. Physica A: Statistical Mechanics and its Applications 369 (2), 799–808.

Cirincione, C., Darling, T. A., O'Rourke, T. G., 2000. Assessing south carolina's 1990s congressional districting. Political Geography 19 (2), 189–211.

Dictionary, 2019. Compactness. `https://www.dictionary.com/browse/compactness`.

Forman, S. L., Yue, Y., 2003. Congressional districting using a tsp-based genetic algorithm. Genetic and Evolutionary Computation Conference, 2072–2083.

Gallagher, M., 1991. Proportionality, disproportionality and electoral systems. Electoral studies 10 (1), 33–51.

Garfinkel, R. S., Nemhauser, G. L., 1970. Optimal political districting by implicit enumeration techniques. Management Science 16 (8), B–495.

George, J. A., Lamar, B. W., Wallace, C. A., 1997. Political district determination using large-scale network optimization. Socio-Economic Planning Sciences 31 (1), 11–28.

Grofman, B., 1985. Criteria for districting: A social science perspective. UCLA L. Rev. 33, 77.

Hess, S. W., Weaver, J., Siegfeldt, H., Whelan, J., Zitlau, P., 1965. Nonpartisan political redistricting by computer. Operations Research 13 (6), 998–1006.

Hojati, M., 1996. Optimal political districting. Computers & Operations Research 23 (12), 1147–1161.

Kalcsics, J., Nickel, S., Schröder, M., 2005. Towards a unified territorial design approach: Applications, algorithms and gis integration. Top 13 (1), 1–56.

Li, Z., Wang, R.-S., Wang, Y., 2007. A quadratic programming model for political districting problem. In: Proceedings of the firsst international symposium on optimization and system biology (OSB). Bejing, China.

Liu, Y. Y., Cho, W. K. T., Wang, S., 2016. Pear: a massively parallel evolutionary computation approach for political redistricting optimization and analysis. Swarm and Evolutionary Computation 30, 78–92.

Mehrotra, A., Johnson, E. L., Nemhauser, G. L., 1998. An optimization based heuristic for political districting. Management Science 44 (8), 1100–1114.

Nemoto, T., Hotta, K., 2003. Modelling and solution of the problem of optimal electoral districting. Communications of the OR Society of Japan 48, 300–306.

Niemi, R. G., Grofman, B., Carlucci, C., Hofeller, T., 1990. Measuring compactness and the role of a compactness standard in a test for partisan and racial gerrymandering. The Journal of Politics 52 (4), 1155–1181.

Nygreen, B., 1988. European assembly constituencies for wales-comparing of methods for solving a political districting problem. Mathematical Programming 42 (1-3), 159–169.

Polsby, D. D., Popper, R. D., 1991. The third criterion: Compactness as a procedural safeguard against partisan gerrymandering. Yale Law & Policy Review 9, 301.

Ricca, F., Simeone, B., 2008. Local search algorithms for political districting. European Journal of Operational Research 189 (3), 1409–1426.

Rincón-García, E.-A., Gutiérrez-Andrade, M.-Á., de-los Cobos-Silva, S.-G., Lara-Velázquez, P., Mora-Gutiérrez, R.-A., Ponsich, A., 2015. Abc, a viable algorithm for the political districting problem. Scientific Methods for the Treatment of Uncertainty in Social Sciences, 269–278.

TUIK, 2019. Election. http://www.turkstat.gov.tr/UstMenu.do?metod=kategorist.
Vanneschi, L., Henriques, R., Castelli, M., 2017. Multi-objective genetic algorithm with variable neighbourhood search for the electoral redistricting problem. Swarm and Evolutionary Computation 36, 37–51.
Vickrey, W., 1961. On the prevention of gerrymandering. Political Science Quarterly 76 (1), 105–110.
Yamada, T., 2009. A mini–max spanning forest approach to the political districting problem. International Journal of Systems Science 40 (5), 471–477.
Young, H. P., 1988. Measuring the compactness of legislative districts. Legislative Studies Quarterly, 105–115.

Appendix A.

Reference	Country	Pop. Equ.	Compactness	Contiguity	Other
Vickrey (1961)	USA	■	■	■	■
Hess et al. (1965)	USA	■	■	■	□
Garfinkel et al. (1970)	USA	■	■	■	□
Bodin (1973)	USA	■	□	■	□
Bourjolly et al. (1981)	USA	■	■	■	■
Nygreen (1988)	Wales	■	■	■	■
Browdy (1990)	USA	■	■	■	□
Hojati (1996)	Canada	■	■	□	□
George et al. (1997)	New Zealand	■	■	□	□
Mehrotra (1998)	USA	■	■	■	□
Cirincione et al. (2000)	USA	■	■	■	■
Bozkaya et al. (2003)	Canada	■	■	■	■
Forman and Yue (2003)	USA	■	■	■	□
Nemoto and Hotta (2003)	Japan	■	□	■	□
Bação et al. (2004)	Portugal	■	■	■	□
Kalcsics et al. (2005)	Germany	■	■	■	□
Chou et al. (2006)	Taiwan	■	■	■	□
Li et al. (2007)	USA	■	■	□	□
Ricca and Simeone (2008)	Italy	■	■	■	■
Shirabe (2009)	USA	■	■	■	□
Yamada (2009)	Japan	■	□	■	□
Rincón-García et al. (2015)	Mexico	■	■	■	□
Liu et al. (2016)	USA	■	■	■	□
Vanneschi (2017)	USA	■	■	■	□

Table A.5: Literature in political districting and main criteria

H	the set of the districts or copies of graph
I	the set of the political units
P	the set of the parties
$\bar{A}$	the set of the arcs between the pair of nodes (s^h, w_i^h), (w_i^h, w_j^h), and (w_i^h, t_i);
$\delta^-(w_i^h)$	the set of arcs entering node w_i^h; $\forall i \in I$, $\forall h \in H$
$\delta^+(w_i^h)$	the set of arcs leaving node w_i^h; $\forall i \in I$, $\forall h \in H$
J_h	the set of representatives of district h; $\forall h \in H$
v_{ki}	the number votes that party k has in unit i; $\forall k \in P$, $\forall i \in I$
s_i	number of voters in unit i; $\forall i \in I$
ρ	the party chosen
$\bar{s}$	the average number of voters
$\bar{s}_h$	the weighted average number of voters in district h, $\forall h \in H$
β	the allowable percentage deviation from the average number voters
M	a sufficiently big value
M_1	a sufficiently big value for the difference of the total votes
M_2	a sufficiently big value for the number of representatives and parties
F	the volume of the flow from each source node
x_{ih}	binary decision variable indicating if unit i is assigned to district h; $\forall i \in I$, $\forall h \in H$
$f(a)$	the volume of the flow on arc a; $\forall a \in \bar{A}$
y_{ih}	binary decision variable indicating if the h^{th} copy of G the flow enters through node i; $\forall i \in I$, $\forall h \in H$
c_{kh}	binary decision variable indicating the party k wins in district h; $\forall k \in P$, $\forall h \in H$
t_h	the number of voters in district h; $\forall h \in H$
o_{kh}	auxiliary variable for the total votes; $\forall k \in P$, $\forall h \in H$
z_{kph}	binary decision variable for the comparing of total votes of party k and p in district h; $\forall k, p \in P$, $\forall h \in H, k \neq p$
c_{kh}	number of representatives of party k gained in district h; $\forall k \in P$, $\forall h \in H$
o_{kjh}	auxiliary variable which is derived from the o_{kh}; $\forall k \in P$, $\forall h \in H, \forall j \in J_h$
e_{kjh}	binary decision variable if party k's denominator j does not have enough vote for a representative in district h; $\forall k \in P$,$\forall h \in H, \forall j \in J_h$
z_{pkjdh}	binary decision variable for the comparison of the votes $\forall k, p \in P$, $\forall h \in H, \forall j, \in J_h$, $(k \neq p, d \neq j)$

Table A.6: All sets, parameters, and decision variables

	p=1	p=2	p=3
o_{ph}	50	18	32

		p=1	p=2	p=3
o_{phj}	j=1	50	18	32
	j=2	25	9	16
	j=3	16.67	6	10.67
	j=4	12.50	4.5	8

Table A.7: Calculation of all vote percentages for the comparison

		p=1	p=2	p=3
$\sum_{k\in P}\sum_{d\in J_h} z_{pkjdh}$	j=1	0	3	1
	j=2	2	8	5
	j=3	4	10	7
	j=4	6	11	9

		p=1	p=2	p=3
e_{pjh}	j=1	0	0	0
	j=2	0	1	1
	j=3	1	1	1
	j=4	1	1	1

Table A.8: Comparison of all vote percentages

	p=1	p=2	p=3
c_{ph}	4-2=2	4-3=1	4-3=1

Table A.9: Calculation of total number of representatives gained by each party

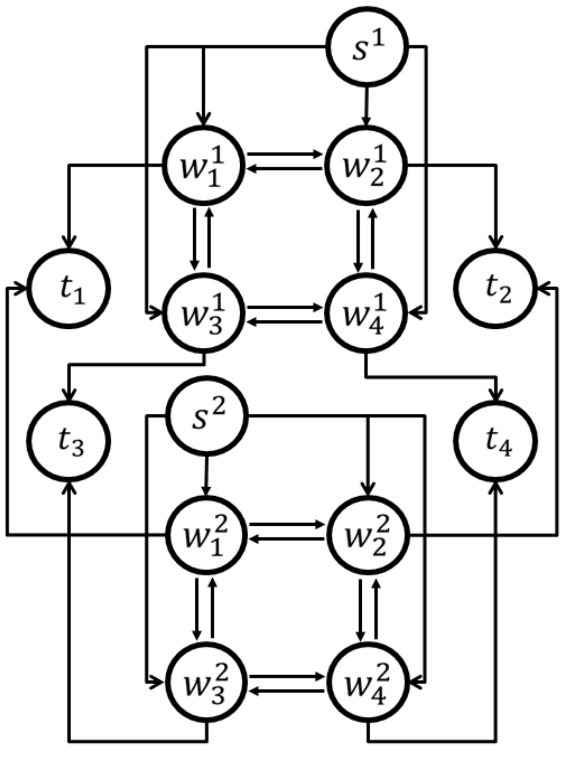


Figure A.4: Illustration of the whole network